Efficiency Cost of Taxes and Empirical Welfare Analysis

Matteo Paradisi
(EIEF)

Public Finance RED - Lecture 3-5

Outline

1. Marshallian surplus
2. Excess burden with income effects
3. Harberger Approximation
4. Sufficient Statistics Approach
5. Marginal Value of Public Funds (MVPF)
6. Redistributive Concerns and Interpersonal Comparisons

Efficiency Definition

- Incidence: effect of policies on distribution of economic pie
- Efficiency or deadweight cost: effect on size of the pie
- Focus in efficiency analysis is on quantities, not prices

Taxes and Efficiency

- Government raises taxes for one of two reasons:
    1. To raise revenue to finance public goods
    2. To redistribute income
- But to generate $1 of revenue, welfare of those taxed falls by more than $1 because the tax distorts behavior
- How to implement policies that minimize these efficiency costs?
    - Start with positive analysis of how to measure efficiency cost of a given tax system

Marshallian Surplus

Marshallian Surplus: Assumptions

- Simplest analysis of efficiency costs
- Main assumptions
  1. Quasi-linear utility: no income effects, money metric
  2. Competitive production

Marshallian Surplus: Partial Equilibrium Model

- Two goods: x and y
- Quasi-linear utility: no income effects
- Consumer has wealth Z, utility $u(x) + y$, and solves

$$\max_{x,y} u(x) + y \text{ s.t. } (p + \tau)x(p + \tau, Z) + y(p + \tau, Z) = Z$$

- Firms use $c(S)$ units of the numeraire y to produce S units of x
- Marginal cost of production is increasing and convex:

$$c'(S) > 0 \text{ and } c''(S) \geq 0$$

- Firm's profit at pretax price p and level of supply S is

$$pS - c(S)$$

Equilibrium

- With perfect optimization, supply fn for x is implicitly defined by

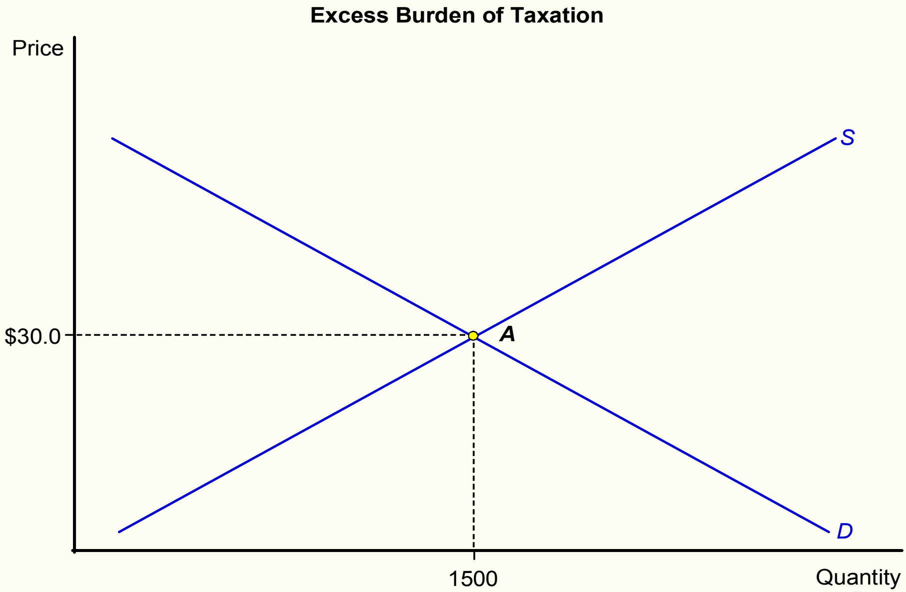
$$p = c'(S(p))$$

- Let $\eta_S = p\frac{S'}{S}$ denote the price elasticity of supply
- Let Q denote equilibrium quantity sold of good x
- Q satisfies:

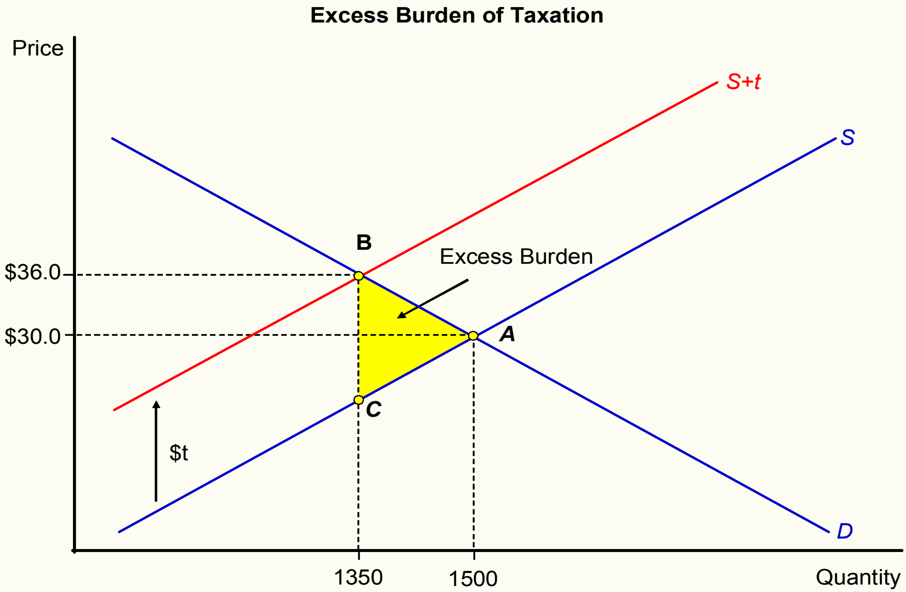
$$Q(\tau) = D(p + \tau) = S(p)$$

- Consider effect of introducing a small tax $d\tau > 0$ on Q and surplus

Equilibrium - Graphical Representation



Equilibrium - Graphical Representation

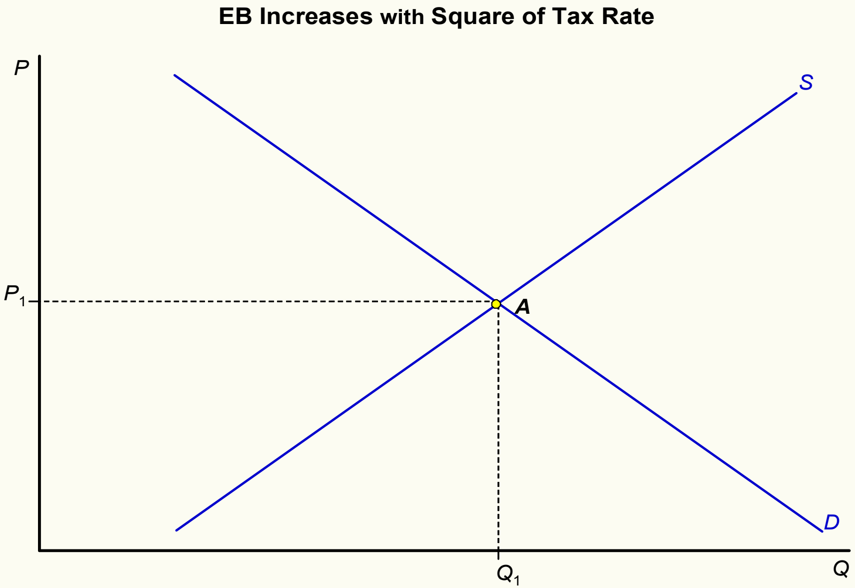


Efficiency Cost: Qualitative Properties

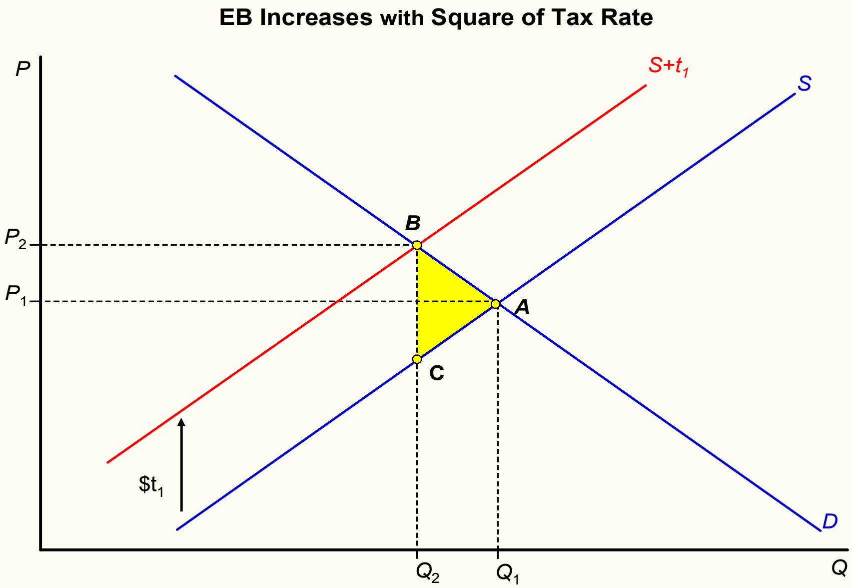
1. Excess burden increases with square of tax rate

2. Excess burden increases with elasticities

Efficiency Cost: Qualitative Properties

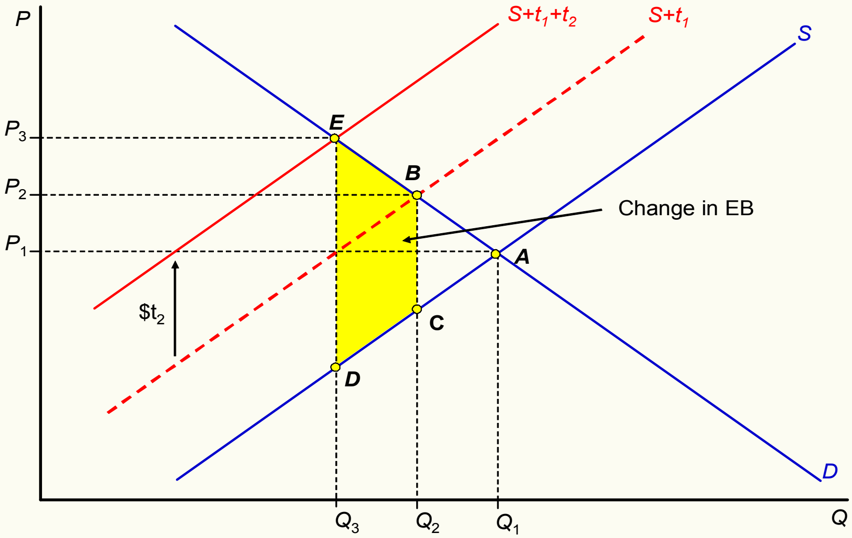


Efficiency Cost: Qualitative Properties



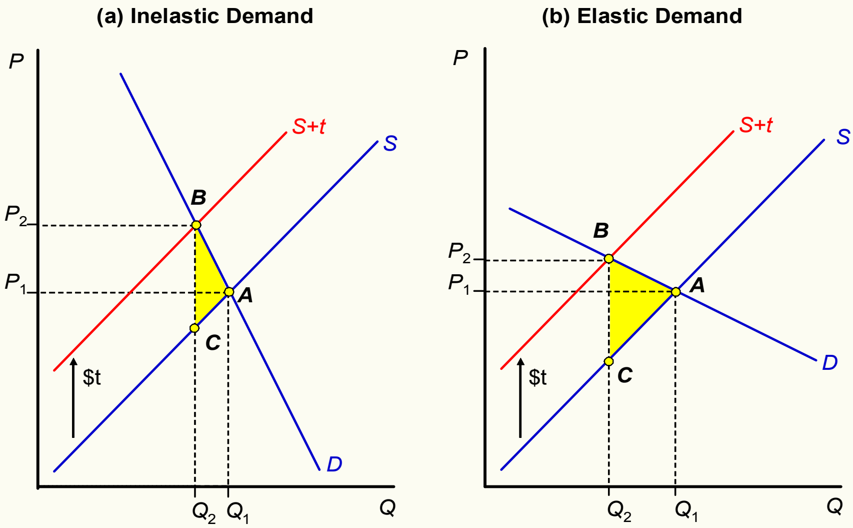
Efficiency Cost: Qualitative Properties

EB Increases with Square of Tax Rate



Efficiency Cost: Qualitative Properties

Comparative Statics



Tax Policy Implications

- With many goods, most efficient way to raise tax revenue is:
  1. Tax inelastic goods more (leads to narrow tax base)
  2. Spread taxes across all goods to keep tax rates relatively low on all goods (broad tax base)
- These are two countervailing forces; balancing them requires quantitative measurement of excess burden

Optimal Linear Taxation - Inverse Elasticity Rule

- Ramsey (1927) and long following literature analyze optimal linear taxation using this framework
  - Intuitively, seek to equalize marginal excess burden across goods
- In this simple setting (no cross-price elasticities), one generates the classic "inverse elasticity" rule:

$$\frac{\tau_i}{1+\tau_i} = \frac{\theta}{\lambda}\frac{1}{\varepsilon_i}$$

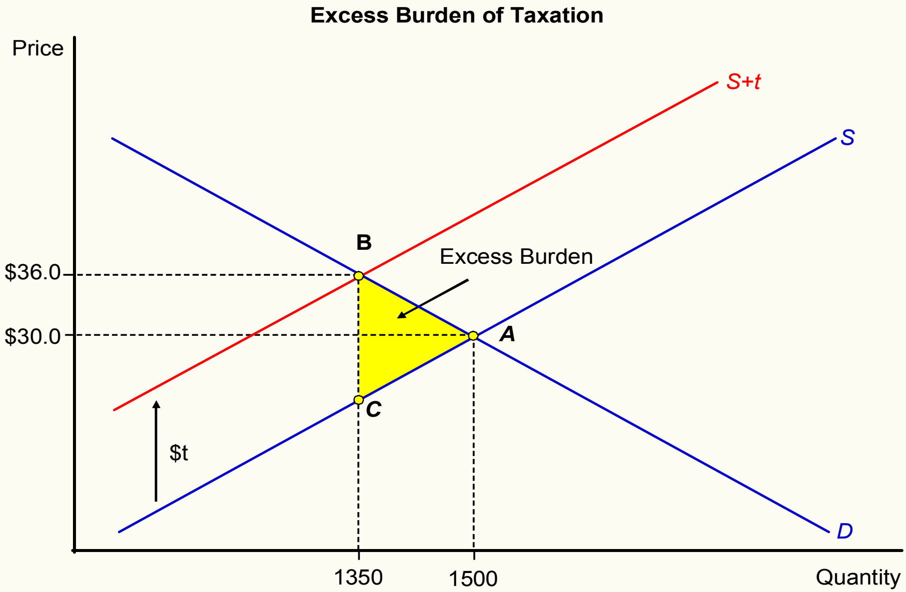
  - θ measures the net value for the government of introducing a $1 lump sum tax
  - λ is the value of government spending

Measuring Excess Burden: Marshallian Surplus

How to measure excess burden? Three empirical methods:

1. In terms of total change in equilibrium quantity caused by tax
2. In terms of supply and demand elasticities
3. In terms of change in government revenue

Method 1: Change in equilibrium quantity



Method 1: Distortions in Equilibrium Quantity

- Define $\eta_Q = \frac{dQ}{d\tau}\frac{p_0}{Q}$
- η_Q: effect of 1% increase in price via a tax change on equilibrium quantity, taking into account the endogenous price change
- Coefficient β in a reduced-form regression:

$$\log Q = \alpha + \beta \frac{\tau}{p_0} + \varepsilon$$

- Identify β using exogenous variation in τ. Then:

$$\begin{aligned} EB &= -\frac{1}{2}\frac{dQ}{d\tau}d\tau d\tau \\ &= -\frac{1}{2}\frac{dQ}{d\tau}\left(\frac{p}{Q}\right)\left(\frac{Q}{p}\right)d\tau d\tau \\ &= -\frac{1}{2}\eta_Q pQ\left(\frac{d\tau}{p}\right)^2 \end{aligned}$$

Method 2: Supply and Demand Elasticities

- Use demand and supply elasticities to rewrite the tax-quantity elasticity $\eta_Q = \frac{dQ}{d\tau}(\frac{p}{Q})$

- Rearranging terms yields (using incidence formula for $\frac{dp}{d\tau}$):

$$\eta_Q = \frac{dQ}{d\tau}\left(\frac{p}{Q}\right) = \frac{dp}{d\tau}\frac{dQ}{dp}\left(\frac{p}{Q}\right)$$
$$\eta_Q = \frac{\varepsilon_S \varepsilon_D}{\varepsilon_S - \varepsilon_D}$$

- Yields alternative representation of EB (per \$ of revenue):

$$EB = -\frac{1}{2}\frac{\varepsilon_S \varepsilon_D}{\varepsilon_S - \varepsilon_D} pQ(\frac{d\tau}{p})^2$$
$$\frac{EB}{R} = -\frac{1}{2}\frac{\varepsilon_S \varepsilon_D}{\varepsilon_S - \varepsilon_D}\frac{d\tau}{p}$$

where $R = d\tau Q$

Marginal Excess Burden of Tax Increase

- Excess burden of a tax τ is

$$EB(\tau) = -(1/2)\frac{dQ}{d\tau}\tau^2$$

- Consider EB from raising tax by $\Delta\tau$ given pre-existing tax τ:

$$\begin{aligned} EB(\tau+\Delta\tau) - EB(\tau) &= -(1/2)\frac{dQ}{d\tau}[(\tau+\Delta\tau)^2 - \tau^2] \\ &= -(1/2)\frac{dQ}{d\tau}\cdot[2\tau\cdot\Delta\tau + (\Delta\tau)^2] \\ &= \underbrace{-\tau\frac{dQ}{d\tau}\Delta\tau}_{\text{Rectangle in trapezoid}} \quad \underbrace{-(1/2)\frac{dQ}{d\tau}(\Delta\tau)^2}_{\text{Triangles}} \end{aligned}$$

assuming
  - $dQ/d\tau$ is unchanged since locally linear D
- First term is first-order in $\Delta\tau$; second is second-order ($(\Delta\tau)^2$)
- Taxing markets with pre-existing taxes causes larger marginal EB

First vs. Second-Order Approximations

- Marginal excess burden by differentiating formula for EB:

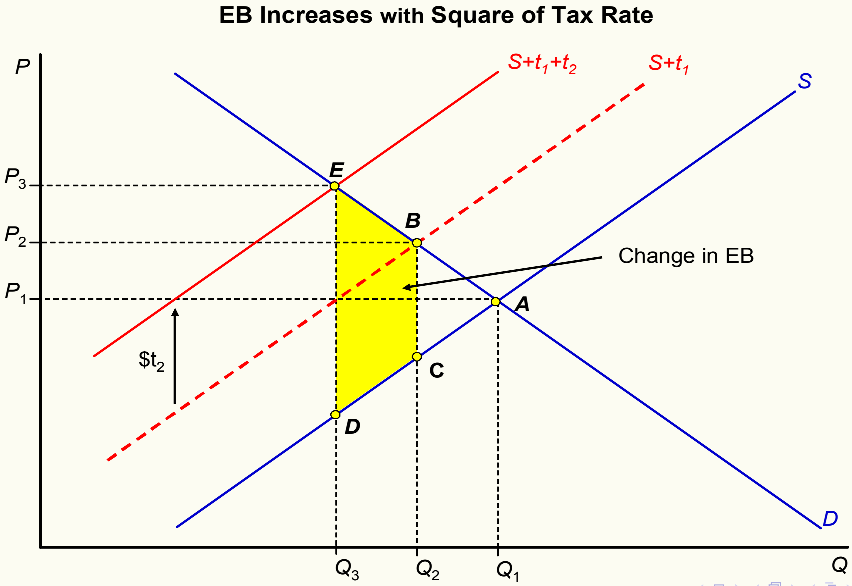
$$\frac{dEB}{d\tau} \cdot \Delta\tau = -\tau \frac{dQ}{d\tau} \cdot \Delta\tau$$

- First derivative of $EB(\tau)$ only includes first-order term in Taylor expansion:

$$EB(\tau + \Delta\tau) = EB(\tau) + \frac{dEB}{d\tau}\Delta\tau + \frac{1}{2}\frac{d^2EB}{d\tau^2}(\Delta\tau)^2$$

- First-order approximation accurate when $\tau >> \Delta\tau$
    - Ex: $\tau = 20\%$, $\Delta\tau = 5\%$ implies 1st term accounts for 90% of EB
    - But new tax ($\tau = 0$) generates EB only through 2nd-order term

Excess Burden: Increases with square of tax rate



Method 3: Leakage in government revenue

- To first order, marginal excess burden of raising τ is:

$$\frac{\partial EB}{\partial \tau} = -\tau \frac{dQ}{d\tau}$$

- Observe that tax revenue $R(\tau) = Q\tau$
  - Mechanical revenue gain: $\frac{\partial R}{\partial \tau}|_Q = Q$
  - Actual revenue gain: $\frac{\partial R}{\partial \tau} = Q + \tau \frac{dQ}{d\tau}$
- MEB is difference bw mechanical and actual revenue gain:

$$\frac{\partial R}{\partial \tau}|_Q - \frac{dR}{d\tau} = Q - [Q + \tau \frac{dQ}{d\tau}] = -\tau \frac{dQ}{d\tau} = \frac{\partial EB}{\partial \tau}$$

Excess Burden with Income Effects

General Model with Income Effects

- Drop quasilinearity assumption and consider utility

$$u(c_1, .., c_N) = u(c)$$

- Individual's problem:

$$\max_{c} u(c) \text{ s.t. } q \cdot c \leq Z$$

where $q = p + \tau$ vector of tax-inclusive prices, Z is wealth

- Labor: commodity with price w consumed in negative quantity

Demand Functions and Indirect Utility

- Let λ denote multiplier on budget constraint
  - Marginal utility of relaxing budget constraint (e.g., extra money)
- First order condition from optimization on each c_i:

$$u_{c_i} = \lambda q_i$$

- These conditions implicitly define:
  - $c_i(q, Z)$: the Marshallian ("uncompensated") demand function
  - $v(q, Z)$: the indirect utility function

Measuring Deadweight Loss with Income Effects

- How much utility is lost because of tax beyond revenue transferred to government?
- Marshallian surplus does not answer this question with income effects
  - Problem: not directly derived from utility function or a welfare measure
  - Creates various problems such as "path dependence" with taxes on multiple goods

$$\Delta CS(\tau^0 \to \tilde{\tau}) + \Delta CS(\tilde{\tau} \to \tau^1) \neq \Delta CS(\tau^0 \to \tau^1)$$

- Need units to measure "utility loss"
  - Use expenditure function to translate the utility loss into dollars
  - Known in the literature as a "money metric"

Expenditure Function

- Fix utility at U and prices at q. Find bundle that minimizes cost to reach U for q:

$$e(q, U) = \min_c q \cdot c \text{ s.t. } u(c) \geq U$$

- μ multiplier on utility constraint; FOC is:

$$q_i = \mu u_{c_i}$$

  - Generate Hicksian (or compensated) demand fns:

$$c_i = h_i(q, u)$$

- Define individual's loss from tax increase as

$$e(q^1, u) - e(q^0, u)$$

  - Money metric for welfare cost with no path dependence

Compensating and Equivalent Variation

- But where should u be measured?

- Price change from q^0 to q^1, two potential reference levels
  - utility $u^0 = v(q^0, Z)$ at old prices (q^0)
  - or $u^1 = v(q^1, Z)$ at new prices (q^1)

- Noting that $Z = e\left(q^1, u^1\right) = e\left(q^0, u^0\right)$, two alternate concepts:
  - Compensating Variation uses old utility level:

$$CV = e(q^1, u^0) - e(q^0, u^0) = e(q^1, u^0) - Z$$

  - Equivalent Variation uses new utility level:

$$EV = e(q^1, u^1) - e(q^0, u^1) = Z - e(q^0, u^1)$$

Compensating Variation

- Measures utility at initial price level u^0
- Amount agent must be compensated in order to be indifferent about tax increase

$$CV = e(q^1, u^0) - e(q^0, u^0) = e(q^1, u^0) - Z$$

- How much compensation is needed to reach original utility level at new prices?
- Ex-post cost government must cover to yield same ex-ante utility:

$$e(q^0, u^0) = e(q^1, u^0) - CV$$

Equivalent Variation

- Measures utility at new price level u^1
- Lump sum amount agent willing to pay to avoid tax (at pre-tax prices)

$$EV = e(q^1, u^1) - e(q^0, u^1) = Z - e(q^0, u^1)$$

- EV is amount extra that can be taken from agent to leave him with same ex-post utility:

$$e(q^0, u^1) + EV = e(q^1, u^1)$$

Efficiency Cost with Income Effects

- Derive empirically implementable formula analogous to Marshallian EB formula in general model with income effects
- Literature typically assumes either
  1. Fixed producer prices and income effects
  2. Endogenous producer prices and quasilinear utility
- With endogenous prices AND income effects, efficiency cost depends on how revenues are returned to consumers
  - Formulas are messy and fragile (Auerbach 1985, Section 3.2)

Efficiency Cost Formulas with Income Effects

- Derive empirically implementable formulas using Hicksian demand (EV and CV)
- Assume p is fixed $\to$ flat supply, constant returns to scale
- The envelope thm implies that $e_{q_i}(q,u) = h_i$, and so:

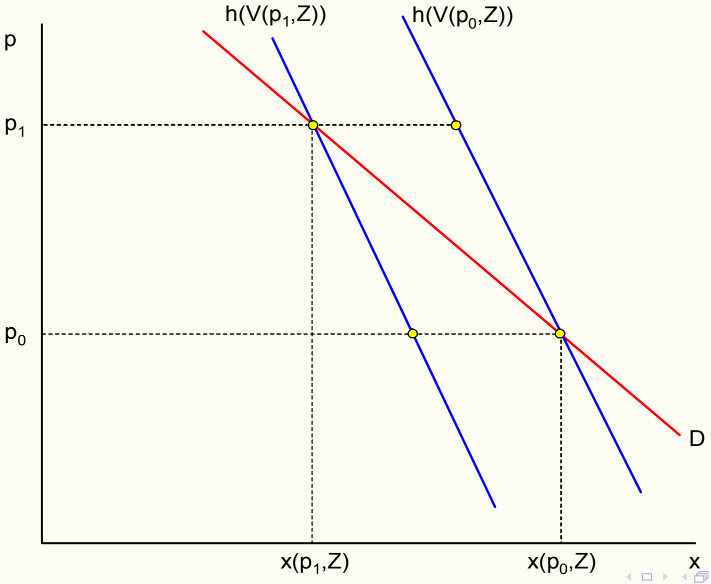
$$e(q^1,u) - e(q^0,u) = \int_{q^0}^{q^1} h(q,u)dq$$

(integral of the derivative)

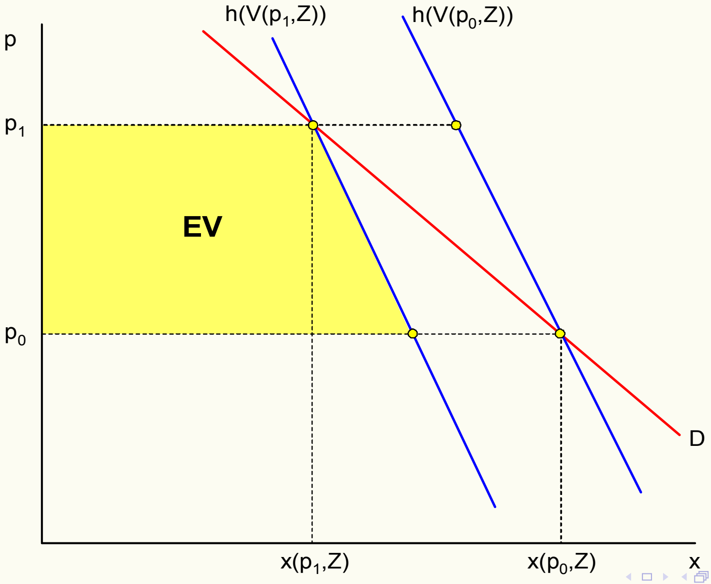
- If only one price changing, this is area under the Hicksian demand curve
- Definitions imply

$$h\left(q, v\left(q, Z\right)\right) = x\left(q, Z\right)$$

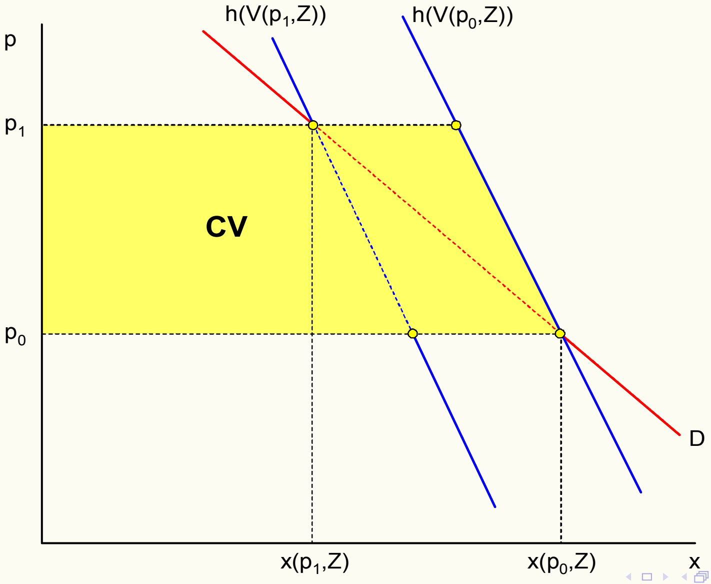
Compensating VS Equivalent Variation



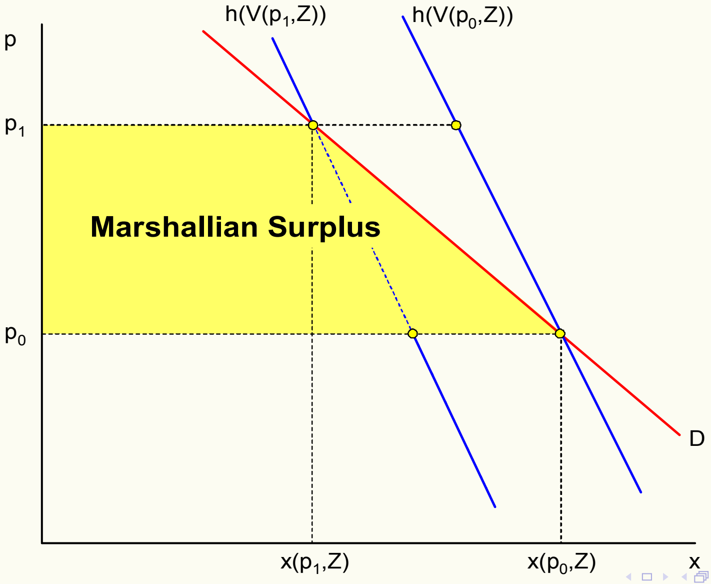
Compensating VS Equivalent Variation



Compensating VS Equivalent Variation



Compensating VS Equivalent Variation



EV, CV, and Marshallian Surplus

- With one price change:

$$\text{EV} < \text{ Marshallian Surplus} < \text{ CV}$$

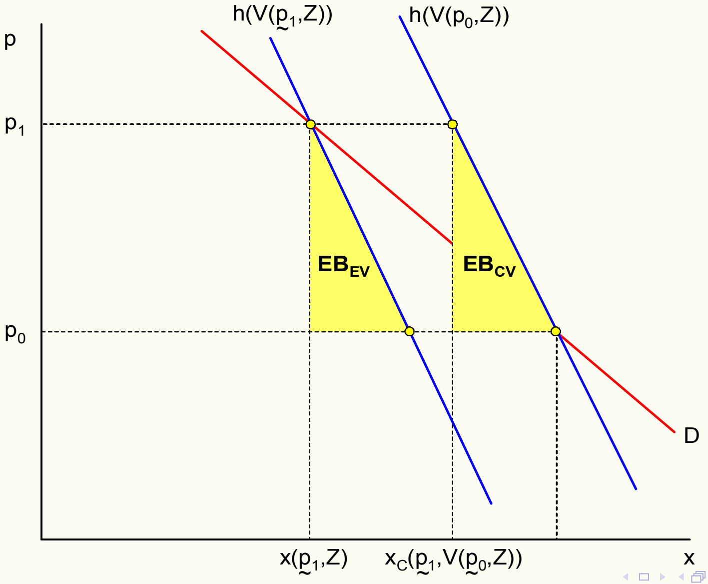
- Not true in general with multiple price changes because Marshallian Surplus is ill-defined

Excess Burden

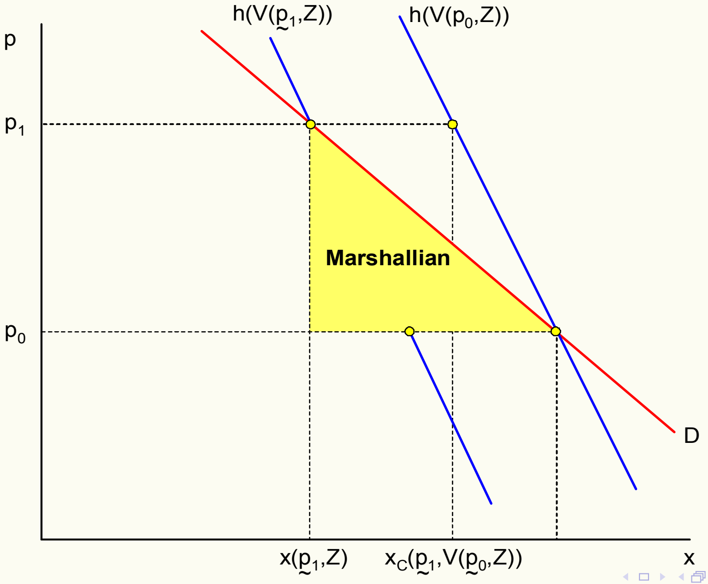
- Deadweight burden: change in consumer surplus less tax paid
- What is lost in excess of taxes paid?
- Two measures, corresponding to EV and CV:

$$EB(u^1) = EV - (q^1 - q^0)h(q^1, u^1) \text{ [Mohring 1971]}$$
$$EB(u^0) = CV - (q^1 - q^0)h(q^1, u^0) \text{ [Diamond and McFadden 1974]}$$

Compensating VS Equivalent Variation



Compensating VS Equivalent Variation



Excess Burden

- In general, CV and EV measures of EB will differ
- Marshallian measure overstates excess burden because it includes income effects
  - Income effects are not a distortion in transactions
  - Buying less of a good due to having less income is not an efficiency loss; no surplus foregone b/c of transactions that do not occur
- CV = EV = Marshallian DWL only with quasilinear utility (Chipman and Moore 1980)

Excess Burden

- Consider increase in tax τ on good 1 to $\tau + \Delta\tau$
- No other taxes in the system
- Recall the expression for EB:

$$EB(\tau) = [e(p+\tau, U) - e(p, U)] - \tau h_1(p+\tau, U)$$

- Second-order Taylor expansion:

$$\begin{aligned} MEB &= EB(\tau + \Delta\tau) - EB(\tau) \\ &\simeq \frac{dEB}{d\tau}\Delta\tau + \frac{1}{2}(\Delta\tau)^2 \frac{d^2 EB}{d\tau^2} \end{aligned}$$

Harberger Trapezoid Formula

$$\frac{dEB}{d\tau} = h_1(p+\tau,U) - h_1(p+\tau,U) - \overbrace{\tau\frac{dh_1}{d\tau}}^{\text{Leakage in gvt rev}}$$

$$= -\tau\frac{dh_1}{d\tau}$$

$$\frac{d^2EB}{d\tau^2} = \underbrace{-\frac{dh_1}{d\tau}}_{\text{Slope of Hicksian}} \overbrace{-\tau\frac{d^2h_1}{d\tau^2}}^{\text{Curvature of Hicksian}}$$

- Standard practice: assume $\frac{d^2h_1}{d\tau^2} = 0$ (linear Hicksian); not necessarily well justified b/c it does not vanish as $\Delta\tau \to 0$

$$\Rightarrow MEB = -\tau\Delta\tau\frac{dh_1}{d\tau} - \frac{1}{2}\frac{dh_1}{d\tau}(\Delta\tau)^2$$

- Formula equals area of "Harberger trapezoid" using Hicksian demands

Harberger Formula

- Without pre-existing tax, obtain "standard" Harberger formula:

$$EB = -\frac{1}{2}\frac{dh_1}{d\tau}(\Delta\tau)^2$$

- General lesson: use compensated (substitution) elasticities to compute EB, not uncompensated elasticities

- Empirically, estimate Marshallian price elasticity and income elasticity. Then apply Slutsky eqn:

$$\underbrace{\frac{\partial h_i}{\partial q_j}}_{\text{Hicksian Slope}} = \underbrace{\frac{\partial c_i}{\partial q_j}}_{\text{Marshallian Slope}} + \underbrace{c_j \frac{\partial c_i}{\partial Z}}_{\text{Income Effect}}$$

Excess Burden with Taxes on Multiple Goods

- With multiple goods and fixed prices, EB of τ_k is

$$EB = -\frac{1}{2}\tau_k^2 \frac{dh_k}{d\tau_k} - \sum_{i \neq k} \tau_i \tau_k \frac{dh_i}{d\tau_k}$$

- Second-order effect in own market, first-order effect from other markets with pre-existing taxes
- Complementarity between goods important for EB
  - Classic Ex.: With an income tax, minimize total DWL tax by taxing goods complementary to leisure (Corlett and Hague 1953)
  - Modern Ex: With a labor income tax, first-order cost of gas tax may come from reduction in work (cost of commuting) rather than distortion in gas consumption (Goulder and Williams 2003)

Goulder and Williams (2003)

- Ignoring cross effects using one-good formula can be very misleading
- Differentiate multiple-good formula wrt τ_k

$$EB = -\tau_k \frac{dh_k}{d\tau_k} - \sum_{i \neq k} \tau_i \frac{dh_i}{d\tau_k}$$

- if τ_k small (e.g. gas tax), only distortions in other markets matter
- if $\tau_k \to 0$ error in single-good formula is big

Using Sufficient Statistics

to Estimate Excess Burden
In Absence of Income Effects

Harberger vs. Hausman Approach

- Two competing approaches to calculating EB:
  - Harberger: Use elasticity-based approximate formulas
  - Hausman: exact measure, estimate of entire demand curves
- Modern literature: derive "sufficient statistics" formulas
- Distinction bw structural and sufficient statistic approaches to welfare analysis in a simple model of taxation
  - No income effects (quasilinear utility)
  - Constant returns to production (fixed producer prices)
  - But permit multiple goods (GE)

Sufficient Statistics vs. Structural Methods

- N goods: $x = (x_1, ..., x_N)$; prices $(p_1, ...p_N)$; wealth Z
- Normalize $p_N = 1$ (x_N is numeraire)
- Government levies a tax t on good 1
- Individual takes t as given and solves

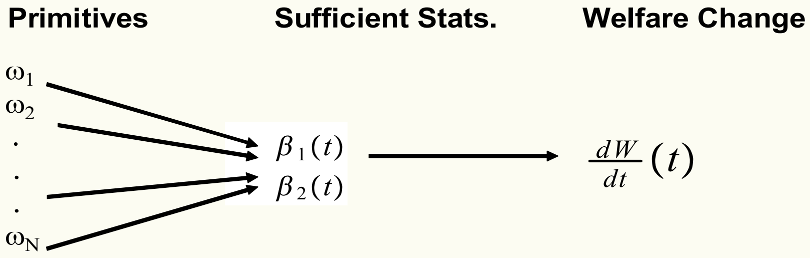
$$\max u(x_1, ..., x_{N-1}) + x_N \text{ s.t. } (p_1 + t)x_1 + \sum_{i=2}^{N} p_i x_i = Z$$

- Social welfare: sum of individual's utility and tax revenue:

$$W(t) = \{\max_x u(x_1, ..., x_{N-1}) + Z - (p_1 + t)x_1 - \sum_{i=2}^{N-1} p_i x_i\} + tx_1$$

- Goal: $\frac{dW}{dt}$ = loss in social surplus caused by tax change

Sufficient Statistics Graphical Illustration



ω=preferences, constraints

ω not uniquely identified

$\beta = f(\omega, t)$
$y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

β identified using program evaluation

dW/dt used for policy analysis

Source: Chetty (2009)

Sufficient Statistics vs. Structural Methods

- Structural: estimate N good demand system, recover u
  - Ex: use functional assumption on utility to recover preference parameters; then calculate "exact consumer surplus"

- Alternative: Harberger's deadweight loss triangle formula
  - Private sector choices maximize term in red (private surplus)

$$W(t) = \{\max_{x} u(x_1, ..., x_{N-1}) + Z - (p_1 + t)x_1 - \sum_{i=2}^{N-1} p_i x_i\} + tx_1$$

  - Envelope conditions for $(x_1, ..., x_N)$: ignore behavioral responses ($\frac{dx_i}{dt}$) in term in red

$$\frac{dW}{dt} = -x_1 + x_1 + t\frac{dx_1}{dt} = t\frac{dx_1}{dt}$$

$\rightarrow \frac{dx_1}{dt}$ is a "sufficient statistic" for $\frac{dW}{dt}$

Heterogeneous Agents

- Benefit of suff stat evident with heterogeneity
- K agents, each with utility $u_k(x_1, ..., x_{N-1}) + x_N$
- Social welfare function under utilitarian criterion:

$$W(t) = \{\max_x \sum_{k=1}^{K} [u_k(x_1^k, ..., x_{N-1}^k) + Z$$

$$- (p_1 + t)x_1^k - \sum_{i=2}^{N-1} p_i x_i^k]\} + \sum_{k=1}^{K} t x_1^k$$

- Structural method: estimate demand systems for all agents
- Sufficient statistic formula unchanged: need only slope of aggregate demand $\frac{dx_1}{dt}$

$$\frac{dW}{dt} = -\sum_{k=1}^{K} x_1^k + \sum_{k=1}^{K} x_1^k + t\frac{d\sum_{k=1}^{K} x_1^k}{dt} = t\frac{dx_1}{dt}$$

Discrete Choice Model

- Harberger sufficient statistic also works with discrete choice
- Agents have value V_k for good 1; can either buy or not buy
- Let $F(V)$ denote distribution of valuations
- With 2 goods, utility of agent k is

$$V_k x_1 + Z - (p + t)x_1$$

- Social welfare:

$$W(t) = \{\int_{V_k} \max_{x_1^k} [V_k x_1^k + Z - (p_1 + t)x_1^k] dF(V_k)\}$$
$$+ \int_{V_k} t x_1^k dF(V_k)$$

- This problem is not smooth at individual level, so cannot directly apply envelope thm. as stated

Discrete Choice Model

- Recast: planner chooses threshold above which agents are allocated good 1:

$$W(t) = \left\{ \max_{\bar{V}} \int_{\bar{V}}^{\infty} \left[V_k - (p_1 + t)\right] dF\left(V_k\right) + Z \right\} + t \int_{\bar{V}}^{\infty} dF\left(V_k\right)$$

- Harberger formula: fn of slope of aggregate demand curve $\frac{dx_1}{dt}$

$$\frac{dW}{dt} = -\left(1 - F\left(\bar{V}\right)\right) + \left(1 - F\left(\bar{V}\right)\right) + t\frac{d\int_{\bar{V}}^{\infty} dF\left(V_k\right)}{dt}$$

$$\Rightarrow \frac{dW}{dt} = t\frac{dx_1}{dt}$$

Economic Intuition for Robustness of Harberger Result

DWL fully determined by difference between marginal willingness to pay for good x_1 and its cost (p_1)

- Recovering marginal willingness to pay requires an estimate of the slope of the demand curve because it coincides with marginal utility:

$$p = u'(x(p))$$

- Slope of demand is therefore sufficient to infer efficiency cost of a tax, without identifying rest of the model

Efficiency Cost Applications to Income Taxation

1. Feldstein (1995, 1999): taxable income
2. Chetty (2009): taxable and earned income
3. Gorodnichenko et al. (2009): implement Chetty's formula

Efficiency Cost Applications: Feldstein 1995, 1999

- Following Harberger, literature in labor estimated effect of taxes on hours worked to assess efficiency costs of taxation
- Feldstein: labor supply involves multiple dimensions, not just choice of hours: training, effort, occupation
  - Taxes also induce inefficient avoidance/evasion behavior
- Feldstein: elasticity of taxable income with respect to taxes is a sufficient statistic for calculating deadweight loss
- Powerful results: taxable income is measurable in data

Feldstein Model: Setup

- Government levies linear tax t on reported taxable income
- Agent makes N labor supply choices: $l_1, \ldots l_N$
- Each choice l_i has disutility $\psi_i(l_i)$ and wage w_i
- Agents can shelter $\$e$ of income from taxation by paying cost $g(e)$
- Taxable Income (TI) is

$$TI = \sum_{i=1}^{N} w_i l_i - e$$

- Consumption is given by taxed income plus untaxed income:

$$c = (1 - t)TI + e$$

Feldstein Taxable Income Formula

- Agent's utility is quasi-linear in consumption:

$$u(c, e, l) = c - g(e) - \sum_{i=1}^{N} \psi_i(l_i)$$

- Social welfare:

$$W(t) = \{(1-t)TI + e - g(e) - \sum_{i=1}^{N} \psi_i(l_i)\} + tTI$$

- Envelope conditions for l_i ($(1-t)w_i = \psi_i'(l_i)$) and e ($g'(e) = t$) imply

$$\frac{dW}{dt} = -TI + TI - t\frac{dTI}{dt} = -t\frac{dTI}{dt}$$

- Intuition: marginal social cost of reducing earnings through each margin is equated at optimum $\rightarrow$ irrelevant what causes change in TI

Taxable Income Formula

- Simplicity has led to a large literature estimating elasticity of taxable income
- But since primitives not estimated, assumptions of model used to derive formula are never tested
- Chetty (2009) questions validity of assumption that $g'(e) = t$
  - Costs of some avoidance/evasion behaviors are transfers to other agents in the economy, not real resource costs
    - Ex: cost of evasion is fine imposed by government

Chetty Transfer Cost Model: Setup

- Individual chooses e (evasion/shifting) and l (labor supply) to

$$\max_{e,l} u(c,l,e) = c - \psi(l)$$
$$\text{s.t. } c = y + (1-t)(wl - e) + e - z(e)$$

- Social welfare is now:

$$\begin{aligned} W(t) = \{& y + (1-t)(wl - e) + e \\ & - z(e) - \psi(l)\} \\ & + z(e) + t(wl - e) \end{aligned}$$

- Difference: $z(e)$ now appears twice in SWF, with opposite signs

Excess Burden with Transfer Costs

- $LI = wl$: total (pretax) earned income
- $TI = wl - e$: taxable income
- Exploit the envelope condition:

$$\frac{dW}{dt} = -(wl - e) + (wl - e) + \frac{dz}{de}\frac{de}{dt} + t\frac{d[wl - e]}{dt}$$
$$= t\frac{dTI}{dt} + \frac{dz}{de}\frac{de}{dt}$$
$$= t\frac{dLI}{dt} - t\frac{de}{dt} + \frac{dz}{de}\frac{de}{dt}$$

- First-order condition for individual's choice of e:

$$t = \frac{dz}{de}$$
$$\Rightarrow \frac{dW}{dt} = t\frac{dLI}{dt} \tag{1}$$

- Intuition: MPB of raising e by \$1 (saving \$t) equals MPC

Chetty (2009) Formula

- With both transfer cost $z(e)$ and resource cost $g(e)$ of evasion:

$$
\begin{aligned}
\frac{dW}{dt} &= t\frac{dLI}{dt} - g'(e)\frac{de}{dt} \\
&= t\{\mu\frac{dTI}{dt} + (1-\mu)\frac{dLI}{dt}\} \\
&= -\frac{t}{1-t}\{\mu TI\varepsilon_{TI} + (1-\mu)wl\varepsilon_{LI}\}
\end{aligned}
$$

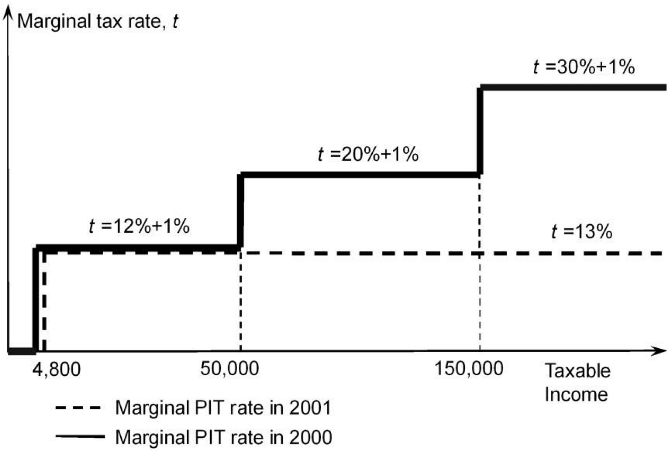
- EB depends on weighted average of taxable income (ε_{TI}) and total earned income elasticities (ε_{LI})
  - Practical importance: even though reported taxable income is highly sensitive to tax rates for rich, efficiency cost may not be large!
- Most difficult parameter to identify: weight μ, which depends on marginal resource cost of sheltering, $g'(e)$

Gorodnichenko, Martinez-Vazquez, and Peter 2009

- Estimate $\varepsilon_{\mathrm{LI}}$ and $\varepsilon_{\mathrm{TI}}$ to implement formula that permits transfer costs
- Insight: consumption data can be used to infer $\varepsilon_{\mathrm{LI}}$
- Estimate effect of 2001 flat tax reform in Russia on gap between taxable income and consumption, which they interpret as evasion

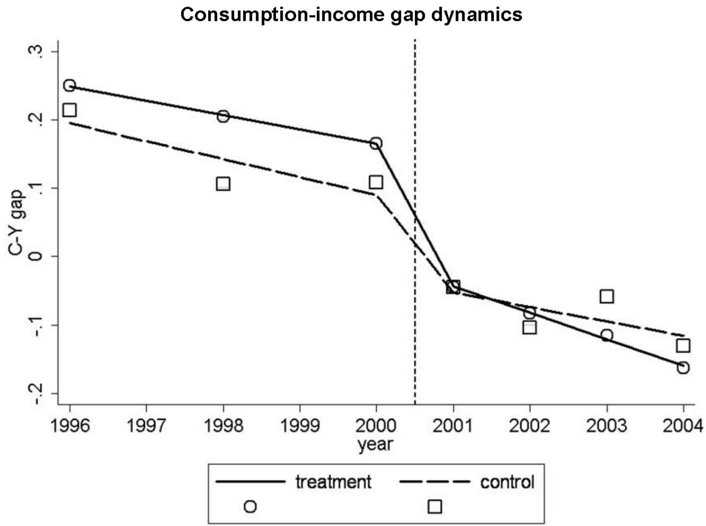
Gorodnichenko, Martinez-Vazquez, and Peter 2009

Marginal personal income tax rate before and after the reform



Source: Gorodnichenko, Martinez-Vazquez, and Peter 2009

Gorodnichenko, Martinez-Vazquez, and Peter 2009



Source: Gorodnichenko, Martinez-Vazquez, and Peter 2009

Gorodnichenko et al: Results

- Taxable income elasticity $\frac{dTI}{dt}$ is large, whereas labor income elasticity $\frac{dLI}{dt}$ is not

  $\rightarrow$ Feldstein overestimates the efficiency costs relative to more general formula
  - (under plausible value of $g'(e)$)

- Can we estimate $g'(e)$ from consumption data itself?

Empirical Welfare Analysis

Empirical Welfare Analysis: Motivation

- Nest causal effects into normative welfare framework
- Key questions:
  - What types of causal effects do we need?
  - What else do we need to know?
  - What are the key assumptions needed?
- Build Marginal Value of Public Funds (MVPF) of a policy

$$\text{MVPF} = \frac{\text{Aggregate Benefits}}{\text{Government Costs}}$$

- Literature: Mayshar (1990), Slemrod and Yitzhaki (1996, 2001), Kleven and Kreiner (2006), Hendren (2016), Hendren and Sprung-Keyser (2019)

General Welfare Framework

- Goal: Measure welfare impact of government policy changes
- Environment:
  - Government chooses vector of policies P that affect
    - constraints
    - preferences
  - Individuals maximize their utility subject to a budget constraint, indirect utility $U_i(P)$

$$U_i(P) = \max_{x \in B_i(P)} u_i(x; P)$$

- Can incorporate different elements
  - public good spending: enters utility
  - taxes: change prices and budget
  - uncertainty and dynamics: $u_i = E\left[\sum_{t \geq 0} \beta^t v_{it}\right]$

Social Welfare Function

Social welfare function (SWF)

$$W(P) = \sum_i \psi_i U_i(P)$$

- ψ_i is i's Pareto weight
  - Nests the case of a classic welfare function

    $$W(P) = \sum_i \omega(U_i)$$

    by setting $\psi_i = \omega'(U_i)$
  - This is suitable for small policy changes.

General Welfare Framework

- Define social marginal utilities of income

$$\eta_i = \psi_i \lambda_i$$

where λ_i is marginal utility of income

$$\lambda_i = \frac{dU_i}{dy_i} = \frac{\partial u_i}{\partial c}$$

- Ratios η_i / η_j correspond to "Okun's leaky Bucket" (Okun, 1976)
  - $\eta_i / \eta_j = 2$ means one is willing to take \$1 from j to give \$0.50 to i
  - society willing to lose money to give \$0.50 to i
- Social marginal utilities of income provide general representation of social preferences
- One restriction:
  - no behavioral biases in utility function: willingness to pay congruent with choices that maximize well-being

Impact of Policy Change on Social Welfare

- $W(P)$: measure used to evaluate government policy changes
- Consider policy change dp (e.g. change in tax rate)
- First-order welfare impact:

$$\frac{dW}{dp} = \sum_i \psi_i \frac{dU_i}{dp} = \bar{\eta}_p \sum_i WTP_i$$

- $WTP_i = \frac{dU_i}{dp} / \lambda_i$ is willingness to pay of i
- $\bar{\eta}_p = \sum_i \eta_i \frac{WTP_i}{\sum_j WTP_j}$: incidence-weighted average social marginal utility of income

Non-Budget Neutral Policies

- Most policies (i.e. reduced-form variations, dp) are not budget neutral
- Identification condition in the studies identifying the causal effects does not specify how the budget constraint is closed
- Traditional approaches: transform policy into a new policy that uses lump-sum taxation
  - need compensated responses rather than estimated ones

Non-Budget Neutral Policies

How we close budget is core issue when comparing methods of welfare analysis:

- Marginal Excess Burden (Auerbach and Hines, 2002):
  - deficit is rebated through hypothetical individual-specific transfers
- CBA with MCPF Adjustment (e.g. Heckman et al. 2010)
  - budget closed through increase in linear tax rate
  - E.g. spend on Perry Preschool and raise taxes through distortionary taxation
- MVPF
  - budget closed through another policy of your choosing

The MVPF: Compare Policies by Normalizing by Cost

- $R(P)$: government budget:

$$R(P) = g(P) - \sum_i T(y_i; P)$$

  - $g(P)$: government spending
  - $T(\cdot, P)$: tax schedule
  - y_i: taxable earnings

- Let $G = \frac{dR}{dp}$: net impact of the policy on government budget
- For instance,

$$G = -\sum_i \left[T'(y_i; P) \frac{dy_i}{dp} + \frac{\partial T(y_i; P)}{\partial p} \right] + \frac{dg(P)}{dp}$$

G includes fiscal externalities from behavioral responses to the policy

The MVPF: Compare Policies by Normalizing by Cost

The Marginal Value of Public Funds (MVPF) of policy p is given by:

$$\text{MVPF}_p = \frac{\sum_i \text{WTP}_i}{G} = \frac{\text{Total Willingness to Pay}}{\text{Net Costs}}$$

- History:
  - Originally defined in Mayshar (1990) and referred to as MEB;
  - later defined in Slemrod and Yitzhaki (1996; 2001) and Kleven and Kreiner (2006) where it was referred to the MCPF
  - Hendren (2016) formalizes the idea and introduces name
- \$1 of spending on the policy $\rightarrow$ \$MVPF benefits to beneficiaries
- \$1 of govt spending $\rightarrow$ $\bar{\eta}_p \text{MVPF}_p$ in social welfare
  - Recall $\left(dW/dp\right)/G = \bar{\eta}_p \text{MVPF}_p$ where $\bar{\eta}_p = \sum_i \eta_i \left(\text{WTP}_i / \sum_j \text{WTP}_j\right)$

MVPF Helps Construction of Policies that Increase Welfare

- Take two (non-budget neutral) policies: policy 1 and policy 2
- Consider budget neutral policy, dp:
  - increase spending on policy 1
  - financed from less spending (greater revenue) from policy 2
- To first order, combined policy increases social welfare ($dW/dp > 0$) iff

$$\bar{\eta}_1 \text{MVPF}_1 > \bar{\eta}_2 \text{MVPF}_2$$

- MVPF normalizes net government spending on each policy
  - construct hypothetical budget neutral policies
- Motivates comparing policies with similar distributional incidence ($\bar{\eta}_1 \approx \bar{\eta}_2$)
  - compare the effectiveness of delivering welfare to the same beneficiaries

MVPF and Welfare-Improving Policy Combination

- Take 2 non-budget neutral policies p_k and p_j
- Change them such that $p_i \to p_i + \Delta p_i, \; i = k, j$
- Choose Δp_j and Δp_k so that sum of two changes budget neutral
  - $\Delta p_k = (\partial R / \partial p_k)^{-1}$ and $\Delta p_j = -\left(\partial R / \partial p_j\right)^{-1}$
- Change in welfare is

$$\Delta W = \bar{\eta}_k MVPF_{p_k} - \bar{\eta}_j MVPF_{p_j}$$

- Policy combination welfare improving if $\Delta W > 0$

Example: MVPF for Tax Rate Change

- MVPF of reduction in marginal income tax rate, τ, by $d\tau$
  - τ applies to earnings y_i
  - average earnings in population $E[y_i]$
- Government revenue is

$$R(\tau) = \tau E[y_i]$$

- Change $d\tau$ leads to

$$-\frac{dR}{d\tau} = E[y_i] + \tau \frac{dE[y_i]}{d\tau}$$

  depends on causal effect of $d\tau$ on $E[y_i]$

Example: MVPF for Tax Rate Change

- Consider the WTP

$$U_i(\tau) = \max_{x_i, y_i \in B_i(\tau)} u_i(x_i, y_i) \text{ s.t. } px_i \leq (1-\tau) y_i$$
$$= u(x_i^*, y_i^*) + \lambda_i ((1-\tau) y_i^\star - px_i^\star)$$

- By envelope theorem (recall: decrease in tax)

$$\frac{dU_i(\tau)}{d\tau} = \lambda_i y_i$$

  - To first order, individuals do not value their change in incomes
  - If earn \$100 and taxes go from 10% to 9%, WTP \$1 for the decrease regardless of how you change earnings

- It follows that

$$E[WTP_i] = E\left[\frac{\frac{dU_i(\tau)}{d\tau}}{\lambda_i}\right] = E[y_i]$$

Example: MVPF for Tax Rate Change

- MVPF is therefore

$$\mathsf{MVPF}_{d\tau} = \frac{E[y_i]}{E[y_i] + \tau \frac{dE[y_i]}{d\tau}} = \frac{1}{1 + FE(d\tau)}$$

- $FE(d\tau)$: fiscal externality of the policy

$$FE(d\tau) = \frac{\tau}{E[y_i]} \frac{dE[y_i]}{d\tau} = \varepsilon$$

  - ratio of behavioral to mechanical effect
  - equal to elasticity of income here

- Key statistics: causal effect of tax rate on income

Infinite MVPFs

- Infinite MVPF: policy pays for itself
  - government saves and generates WTP
  - hence, WTP $>$ 0 and G $<$ 0
  - know as "Laffer effect" (we'll see it again)
- Preserves ordering: MVPF $= \infty$ better than finite MVPF
- MVPF $= -\infty$ when WTP $<$ 0 and G $>$ 0
- MVPF generalizes Laffer effects to policies $\neq$ income tax
- In previous example: MVPF $= \infty$ if $\varepsilon < -1$
  - for 1% increase in tax lose more than 1% income

An Aside: Optimized Social Welfare

- Full problem:

$$\max_{p \in P} W(P) \text{ s.t. } R(P) \geq 0$$

- FOCs:

$$\frac{\partial W}{\partial p} + \lambda \frac{\partial R(p)}{\partial p} = 0$$

where λ is multiplier on govt budget. For any p

$$MVPF_p = \frac{1}{\lambda} \frac{\partial W / \partial p}{-\partial R / \partial p} = 1$$

- Social optimum: equate benefit-cost ratios across policies
- Full optimum: $MVPF_p = 1$ for each p (in govt dollars)
- Important: MVPF useful even far from full optimum
  - tells us which policies should be implemented first

Empirical Estimates of MVPF

- Hendren and Sprung-Keyser (2019): 133 MVPFs for policies in social insurance, education and job training, taxes and cash transfers, and in-kind transfers
- Construct sample from survey and review articles
- Assess robustness to range of assumptions
  - Program Parameters (discount rate, tax rate, etc.)
  - Forecasting/Extrapolation of Observed Effects Validity of Empirical Designs (RCTs/RDs vs. Diff-in-Diff; Peer Reviewed vs. not; etc.)
  - Publication Bias (Andrews and Kasy, 2018)
  - Missing Causal Estimates (e.g. restrict to subsets of policies with different sets of observed effects)

Example: Admission to Florida International University

- Florida International University (FIU): minimum GPA threshold for admission
  - created a fuzzy discontinuity
- Zimmerman (2014): use discontinuity to examine impact of FIU admission on earnings for 14 years after admission.

Zimmerman (2014)

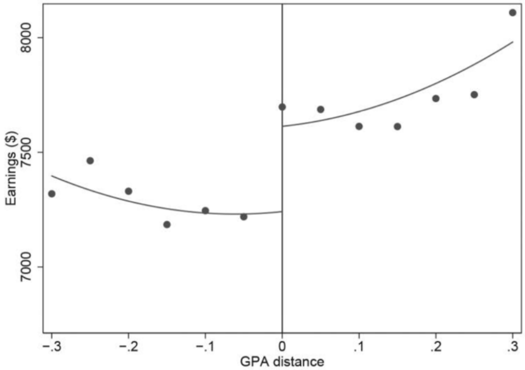
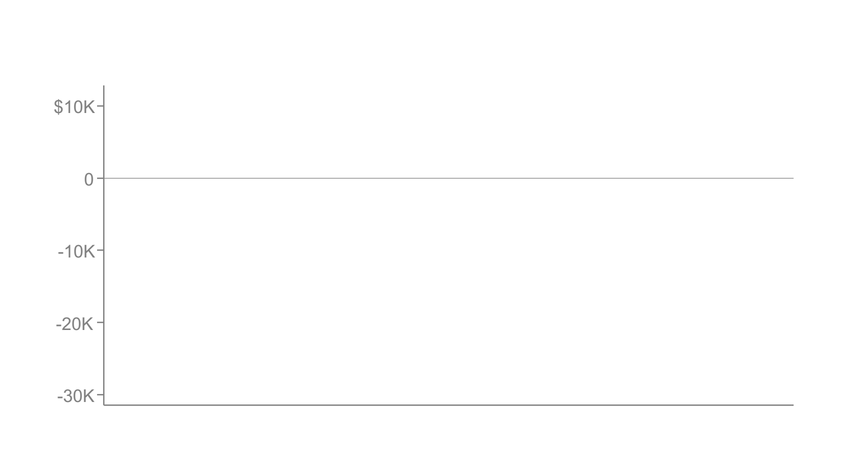
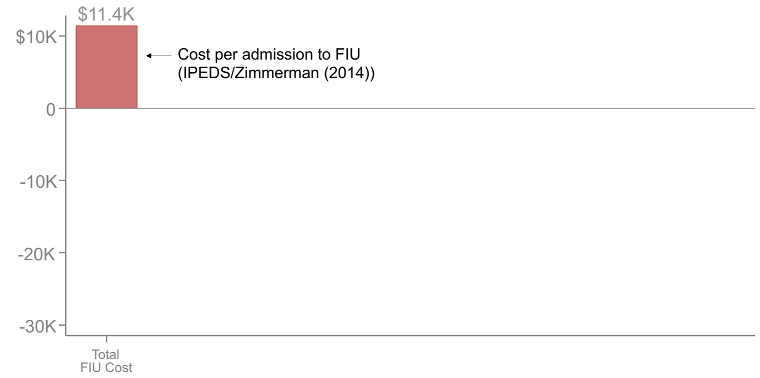


FIG. 8.—Quarterly earnings by distance from GPA cutoff. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

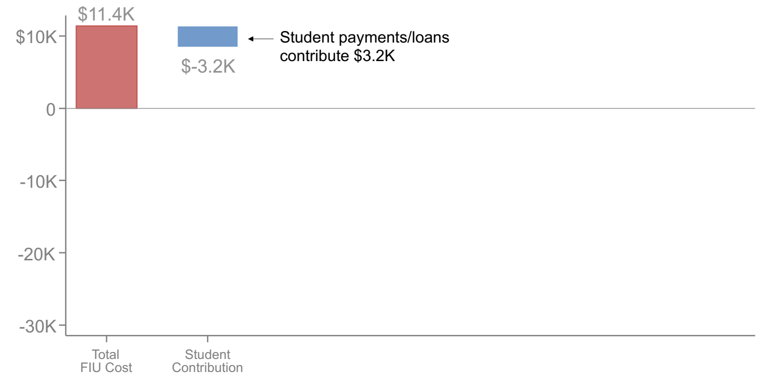
Zimmerman (2014)



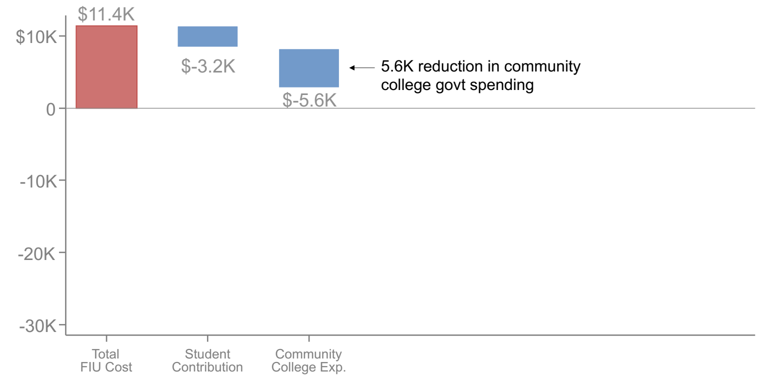
Zimmerman (2014)



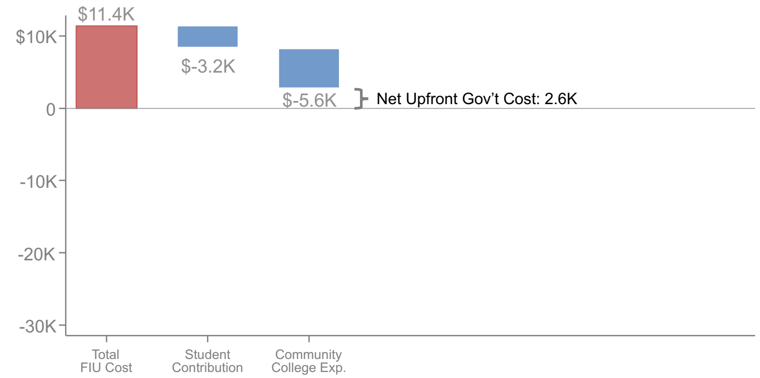
Zimmerman (2014)



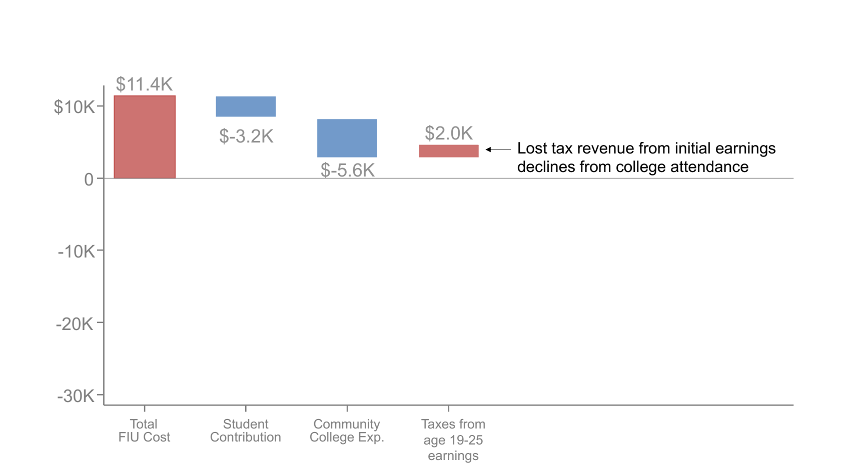
Zimmerman (2014)



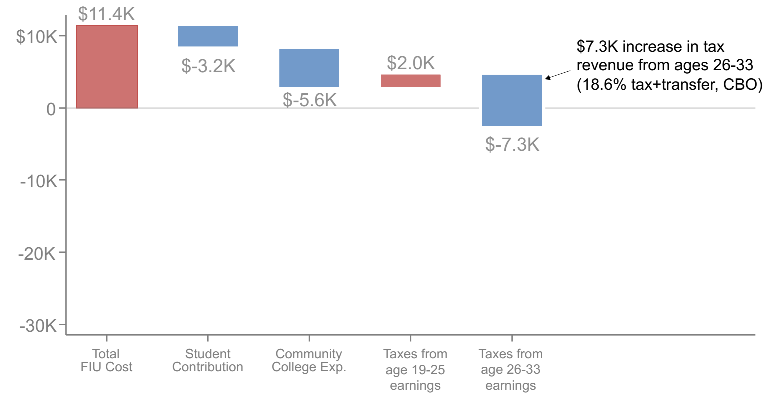
Zimmerman (2014)



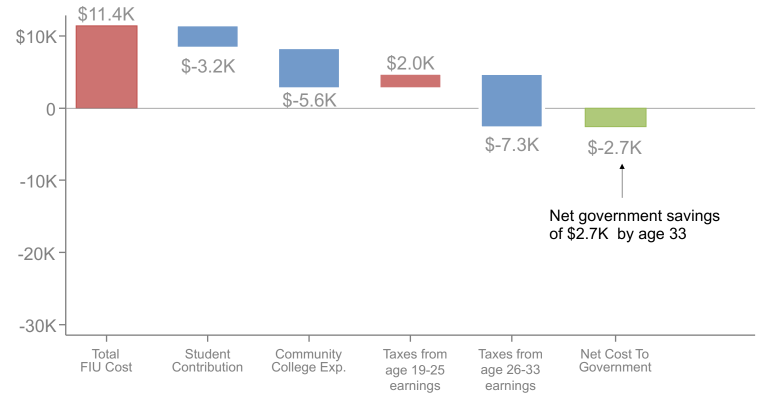
Zimmerman (2014)



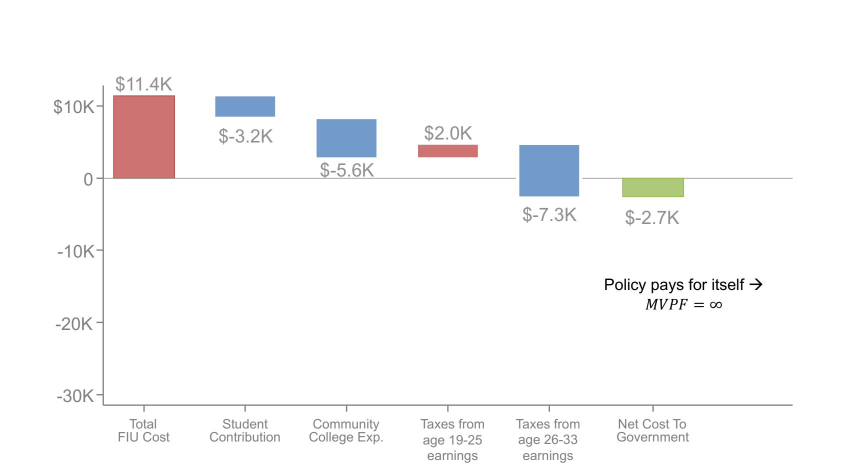
Zimmerman (2014)



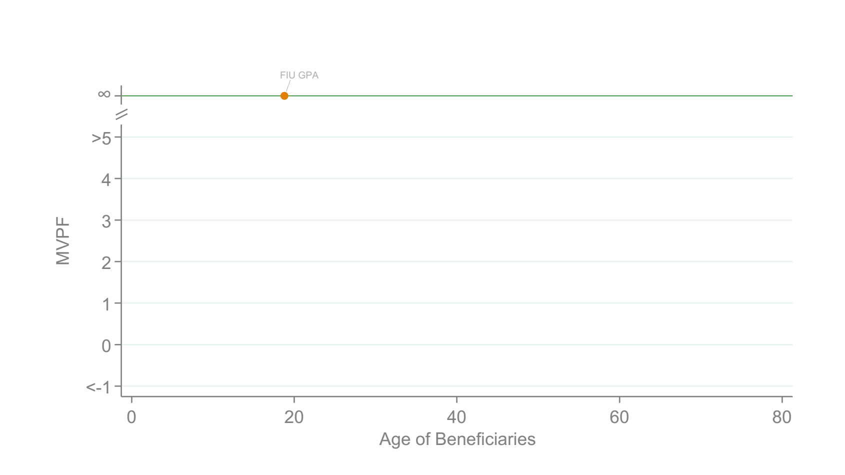
Zimmerman (2014)



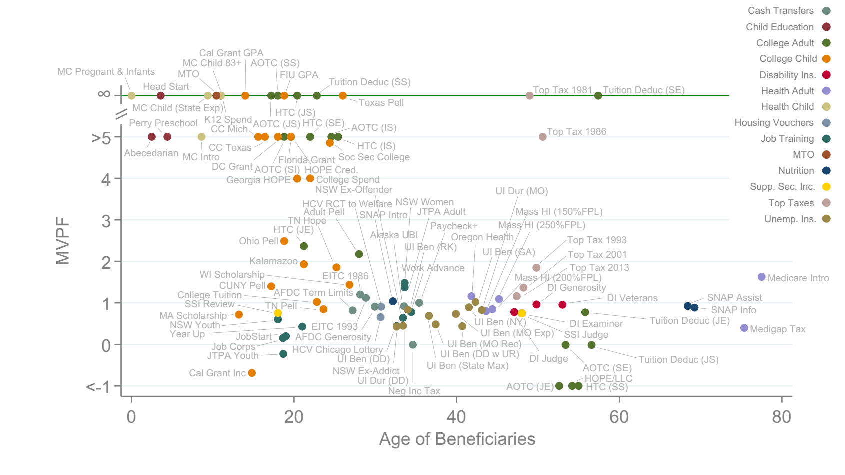
Zimmerman (2014)



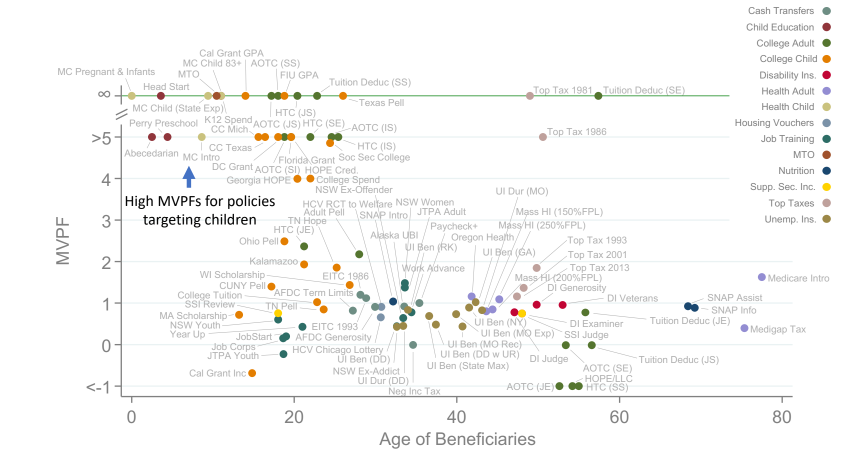
Hendren and Sprung-Keyser (2019)



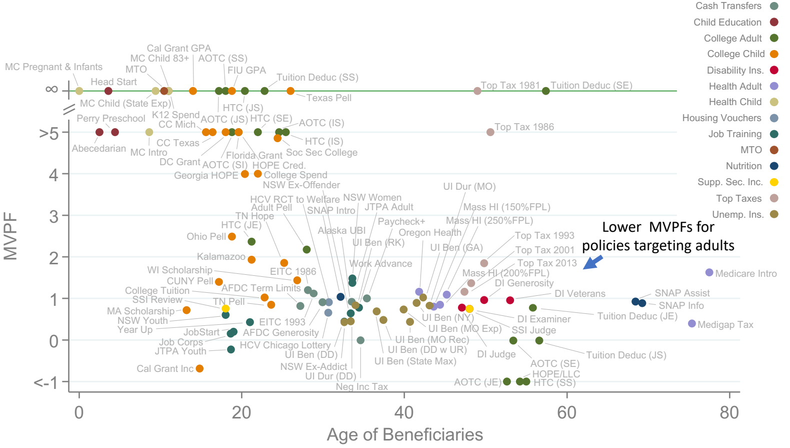
Hendren and Sprung-Keyser (2019)



Hendren and Sprung-Keyser (2019)

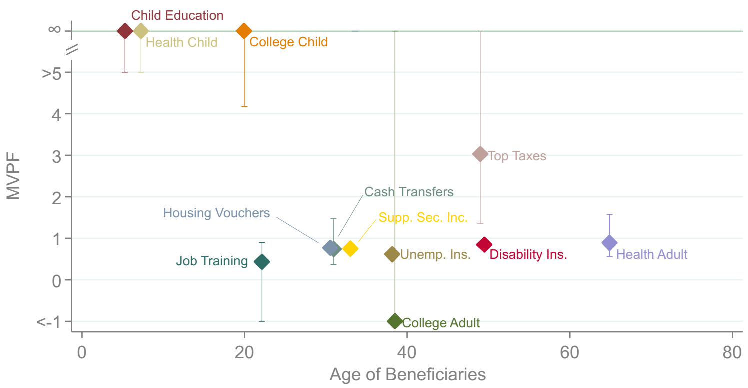


Hendren and Sprung-Keyser (2019)



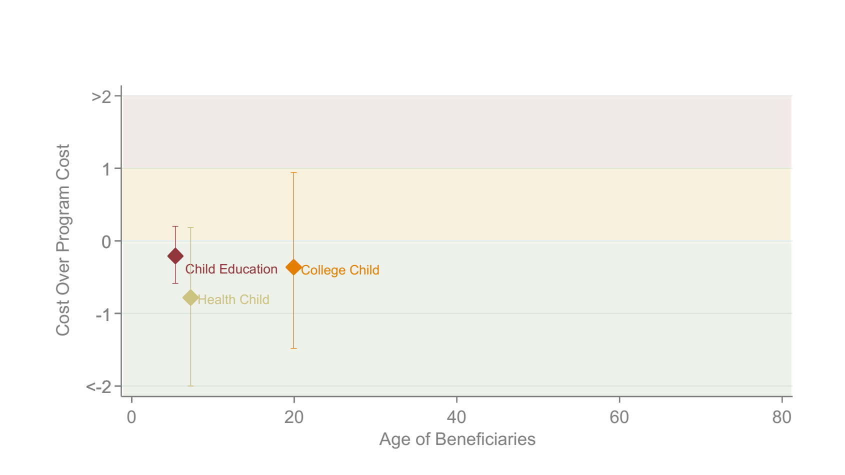
Hendren and Sprung-Keyser (2019)

Policy Categories



Hendren and Sprung-Keyser (2019)

Policy Categories - Net Costs



Hendren and Sprung-Keyser (2019)

Policy Categories - Net Costs

Comparison with MEB

- MEB corresponds to another conceptual policy experiment
- Historically, two ways of defining MEB (Auerbach and Hines (2002))
    1. How much additional revenue could the government get if the policy change is implemented but utility is held constant using individual specific lump-sum transfers?
        - Policies better if the government can increase revenue by doing the policy
    2. How much are individuals willing to pay for the policy if the government closes the budget constraint through taxes using individual-specific (non-distortionary) transfers

Comparison with MEB

- Use EV definition of MEB
- Suppose status quo is policy vector P
- Indirect utility is $v_i^{init} = U_i(P)$
- Compensated policy path $\hat{P}$ such that $U_i(\hat{P}) = v_i^{init}$
  - $\hat{P}$ is P with some policy changes dp
  - AND with individual-specific transfers
- $\hat{t}_i(\hat{P})$: level of resources allocated by govt to i under $\hat{P}$
- MEB of policy dp for individual i is

$$MEB_{dp,i} = \frac{d\hat{t}_i(\hat{P})}{dp}$$

MEB and WTP Link

- How does MEB relate to WTP?
- Define income effect as $INC_i = \tau_i^x \left(\frac{dx_i}{dp} - \frac{dx_i}{dp}^c \right)$, where x_i^c are choices under compensated path
- Then MEB can be defined as

$$MEB_{dp,i} = \underbrace{\frac{\frac{dU_i}{dp}}{\lambda_i} - \frac{dt_i(P)}{dp}}_{\text{WTP Above Resource Cost}} - INC_i$$

where $t_i(P)$ is resources to i without compensation

- MEB and WTP related through income effects induced by the lump-sum taxation that hold individuals' utility constant in the MEB experiment
- If $\frac{\frac{dU_i}{dp}}{\lambda_i} = 0$, no compensation and $\frac{dx_i}{dp}^c = \frac{dx_i}{dp} = 0 \rightarrow MEB_{dp,i} = 0$

MEB in Tax Hike Example

- Individual budget constraint

$$c_i \leq (1-\tau) y_i + t_i$$

- Revenue impact of compensated tax change

$$\frac{dR}{d\tau^c} = \underbrace{E[y] + \tau \frac{dE[y]}{d\tau}}_{\text{Tax Hike}} \underbrace{- E[y] - \tau E[y] \frac{dE[y]}{dt}}_{\text{Lump-sum}}$$

WTP to compensate is $E[y]$, $dt = E[y]$

- Hence

$$\frac{dR}{d\tau^c} = \tau \underbrace{\left(\frac{dE[y]}{d\tau} - E[y] \frac{dE[y]}{dt} \right)}_{\text{Compensated Response: Slutsky}}$$

Issues with the MEB Approach

Two fundamental problems with MEB

- Requires compensated, not causal effects
  - Income effects hard to measure (especially if they are not invariant across environments)
- Individual specific transfers are not feasible (this is the core idea behind Mirrlees' optimal income tax work).
  - E.g. distortionary taxes will always look "bad"
- It is still possible to compare MEBs across policies
  - Appropriately defined, this will characterize changes in social welfare
  - But, requires compensated effects bc both policy changes need to add in, then subtract the income effects

Cost-Benefit Analysis

- Benefit Cost Ratios are another method of policy comparison
  - (Boardman et al. (2018), Garcia, Heckman, et al (2017), Heckman et al. (2010))
- Compare the total benefits to the upfront programmatic cost:

$$\text{BCR} = \frac{\text{Social Benefits - Social Costs}}{\text{Programmatic Costs}\,(1+\phi^{\text{DWL}})}$$

- Multiply costs by an adjustment for the EB of taxation
- Benefits accruing to the government are included as social costs

Key Problem with Cost-Benefit Analysis

- BCR suffers from three related conceptual problems
  1. Revenue impacts included in numerator but they reduce the need to raise revenue and thus the excess burden of taxation!
     - But the excess burden only multiplies the upfront cost
  2. Forces a particular method of closing the budget constraint (linear taxation)
  3. Does not account for differential distributional incidence of the policy relative to the method used to raise revenue
     - one can incorporate distributional weights
- On the contrary, MVPF
  - puts the net government cost in the denominator,
  - allows the researcher to compare the MVPF to other policies,
  - uses Okun's bucket

Redistributive Concerns and Interpersonal Comparisons

Distributional Incidence

- Benefit-cost ratio for welfare analysis

$$\text{MVPF} = \frac{\text{WTP}}{\text{Costs}}$$

- Welfare comparisons, what we said so far
  - Pareto comparisons when policies have same distributional incidence
  - Okun's bucket or social welfare weights when different incidence
- When policies have different distributional inference can we do more?
  - Kaldor-Hicks efficiency tests

Interpersonal Comparisons

Two methods for resolving interpersonal comparisons

1. Social welfare function (Bergson (1938), Samuelson (1947), Diamond and Mirrlees (1971), Saez and Stantcheva (2015))
   - Allows preference for equity
   - Do the policy only if \$1.50 to rich valued more than \$0.5 to the poor:

$$\frac{\eta^{\text{rich}}}{\eta^{\text{poor}}} > \frac{1}{3}$$

   - Subjective choice of researcher or policy-maker

2. Kaldor Hicks Compensation Principle (Kaldor (1939), Hicks (1939, 1940))
   - Aggregate surplus, or “efficiency”, as normative criteria

$$\$1.5 - \$0.5 = \$1 > 0 \iff \text{do the policy}$$

   - Ignores issues of “equity”

Kaldor-Hicks: Motivating Aggregate Surplus

- Individuals are willing to pay s_i for alternative environment
- Pareto improvement only if $s_i > 0$ for all i
  - In general, $s_i > 0$ and $s_j < 0$ for some i and j
- Kaldor Hicks: consider alternative policy that taxes/transfers t_i to individuals
- How much can we tax each i and break even?

$$t_i^{max} = s_i$$

- Potential Pareto improvement if

$$\sum_i t_i^{max} > 0 \iff \sum_i s_i > 0$$

- If total surplus is positive, govt can use policy + tax/transfers to make everyone better off
- Can winners compensate losers through transfers?
  - Interpreted as lump-sum transfers (individual specific)

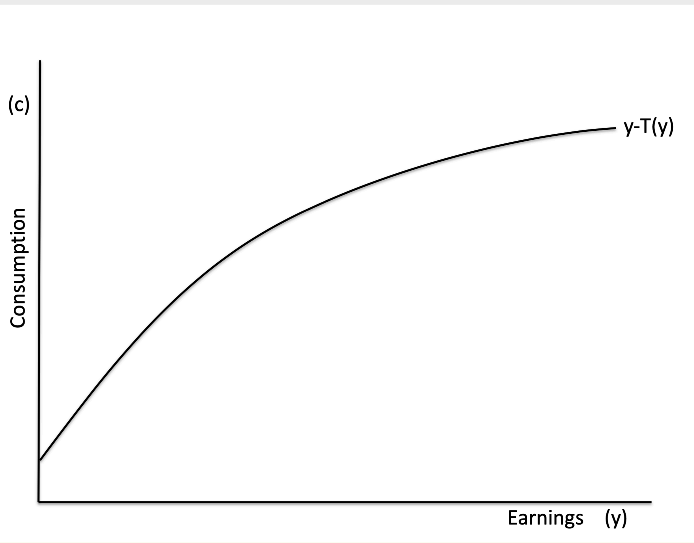
Kaldor-Hicks Limitations

- Kaldor-Hicks: use individual-specific lump-sum transfers to neutralize interpersonal comparisons
- BUT, key insight of Mirrlees and optimal tax literature
  - Can't do individual-specific lump-sum taxes
  - We use distortionary taxes since individual transfers not feasible
- Modify Kaldor-Hicks to make transfers incentive compatible
- Kaldor-Hicks motivates comparing MVPF of policy to MVPF of distributionally-equivalent tax cut

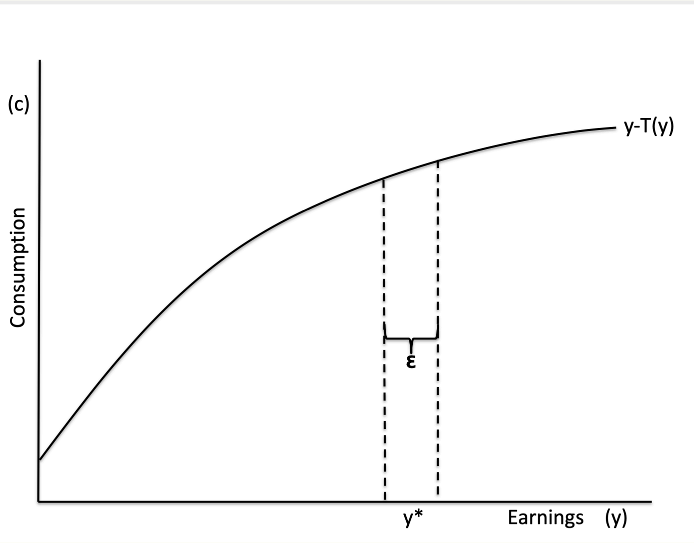
Hendren (2020): Efficient Inverse-Optimum Weights

- Turn unequal surplus into equal surplus using modifications to the tax schedule
  - Not individual-specific lump-sum transfers
- Cost of moving $1 of surplus differs from $1 because of how behavioral response affects government budget
- Back out weights to be used for "reasonable" comparisons to reach "universal agreement"
- Suppose we want to provide transfers to those earning near $y^\star$

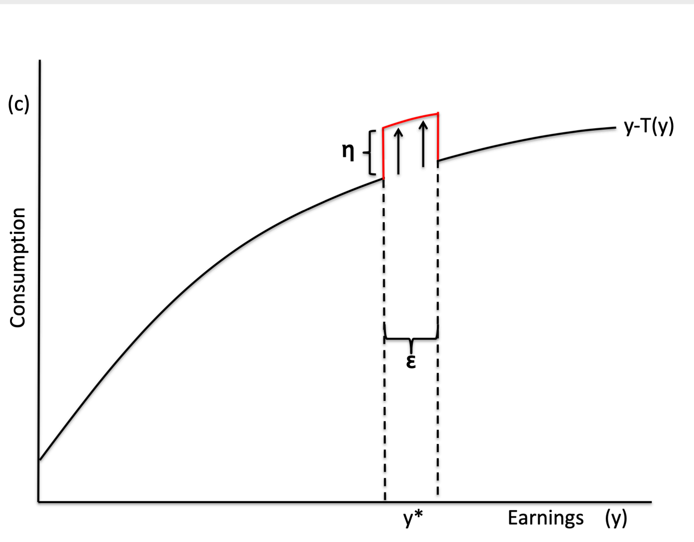
Transferring Money to $y^\star$



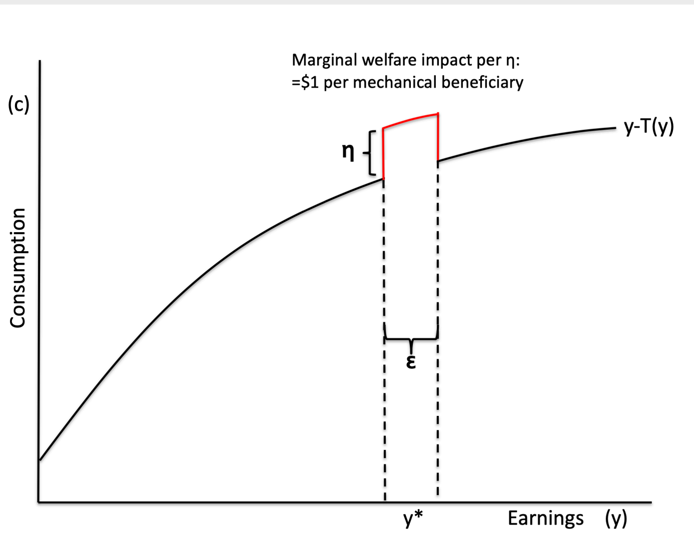
Transferring Money to y^*



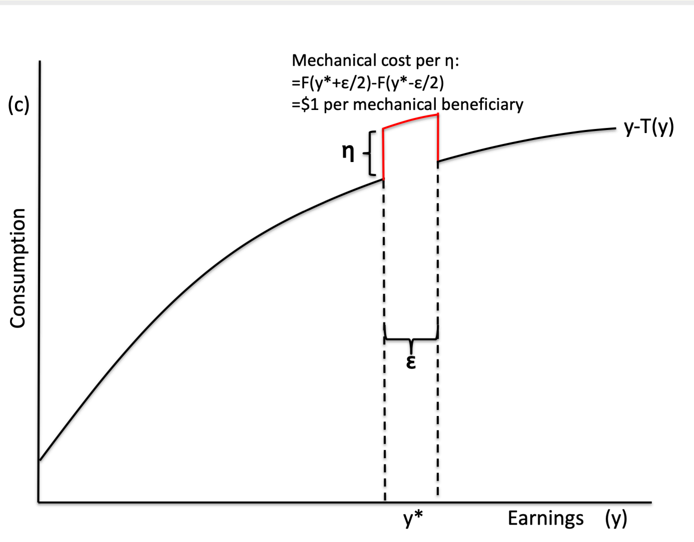
Transferring Money to y^*



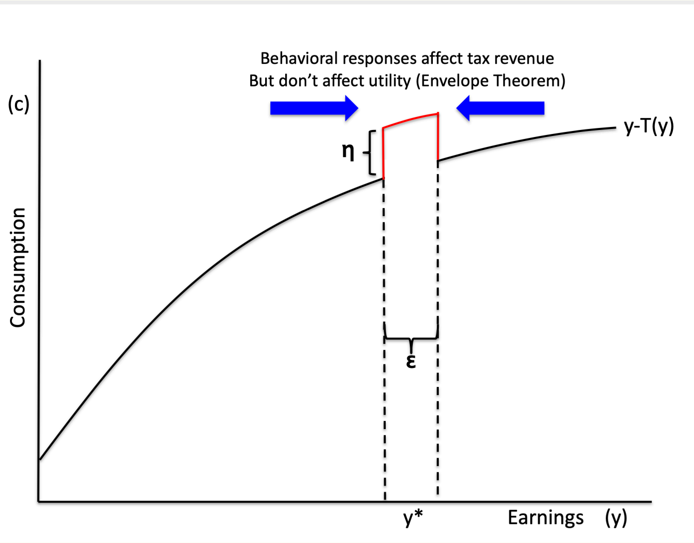
Transferring Money to y^*



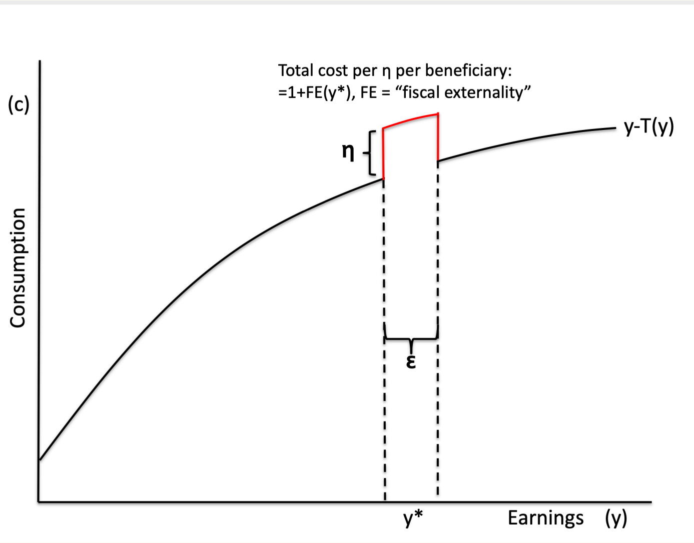
Transferring Money to y*



Transferring Money to $y^\star$



Transferring Money to y^*



Weights

- Consider the function:

$$g(y) = 1 + FE(y)$$

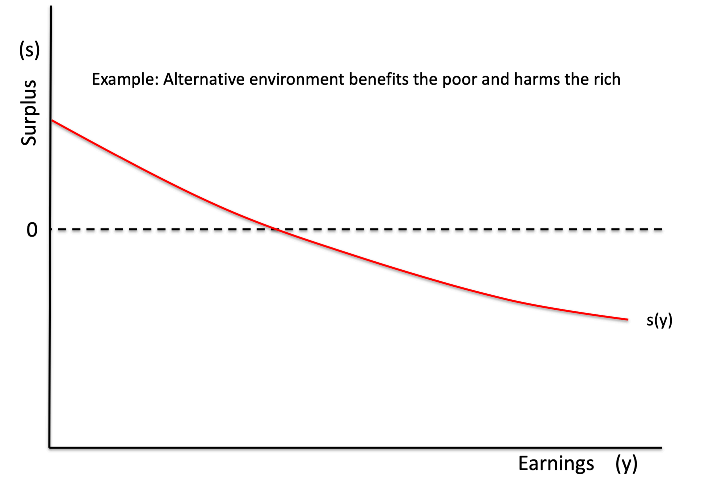
and the normalized function

$$\tilde{g}(y) = \frac{1 + FE(y)}{E[1 + FE(y)]}$$

- Interpretation: \$1 surplus to those earning y can be turned into \$$\tilde{g}(y)/n$ surplus to everyone through modifications to tax schedule
- Fiscal externality logic does not rely on functional form assumptions
  - Allows for each person to have her own utility function and arbitrary behavioral responses
  - Extends to multiple policy dimensions
- Let's use fiscal-externality weighted surplus

$$S = E[g(y)\,s(y)]$$

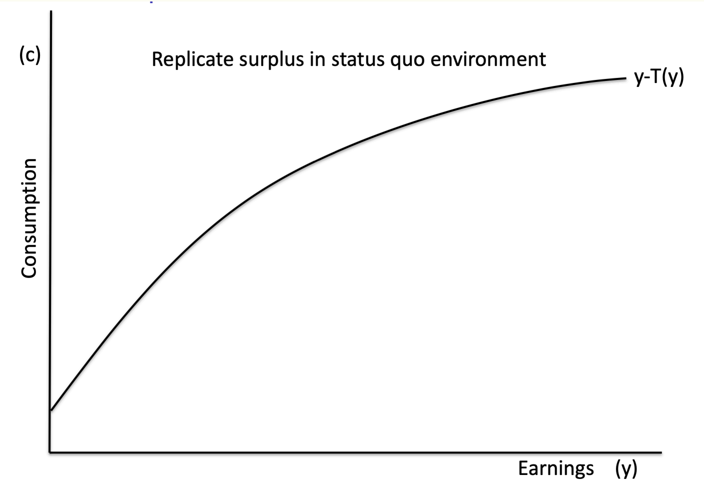
Example of alternative environment



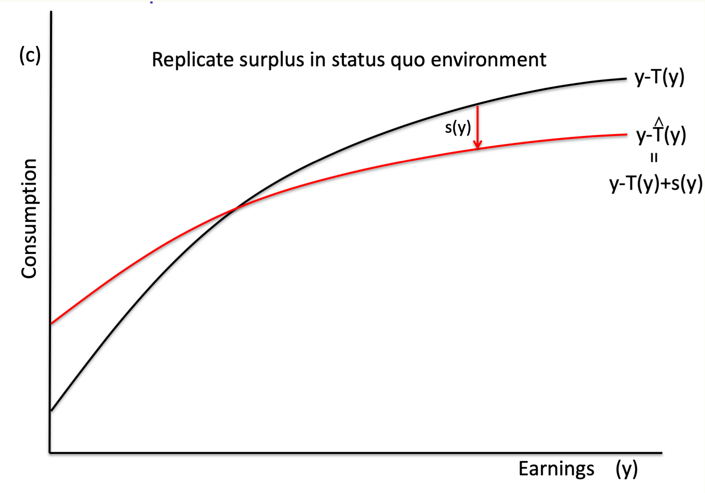
EV and CV Measures: Kaldor and Hicks Experiments

- Given $s(y)$ let's consider a modified policy that neutralizes distributional comparisons
- $s(y)$ is defined as equivalent variation
  - how much money you need in status-quo to be indifferent wrt alternative
- Ways of neturalizing distributional comparisons: EV and CV
- "EV": modify status quo tax schedule
  - by how much can everyone be made better off in modified status quo relative to alternative environment
- "CV": modify alternative environment tax schedule
  - by how much can everyone be made better off in modified alternative environment relative to status quo?

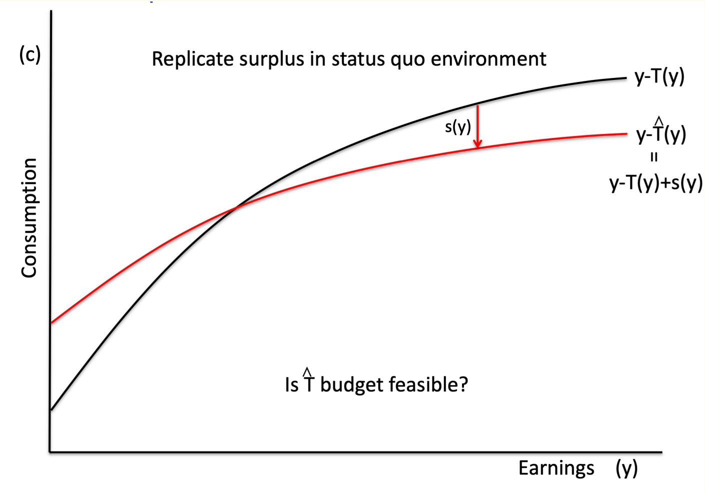
EV Measure - Hicks Experiment



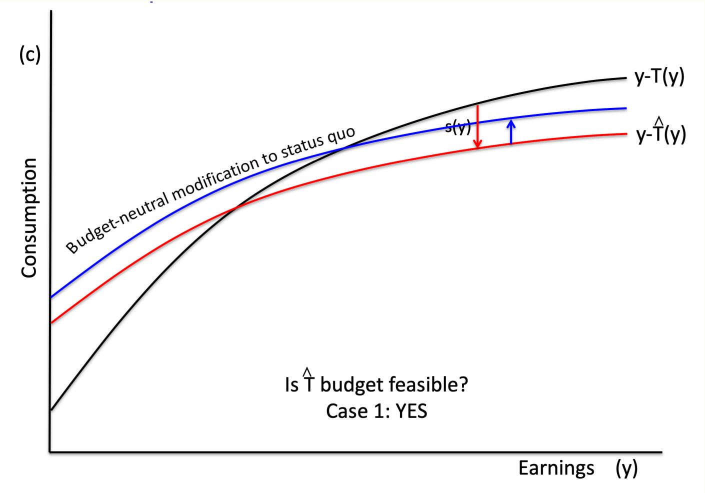
EV Measure - Hicks Experiment



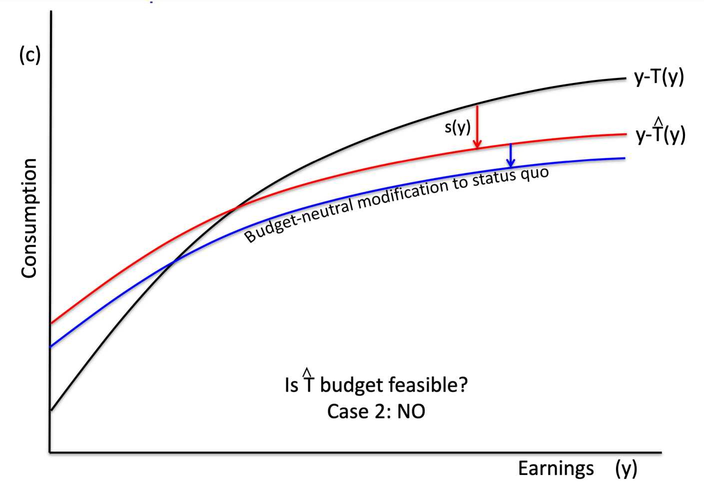
EV Measure - Hicks Experiment



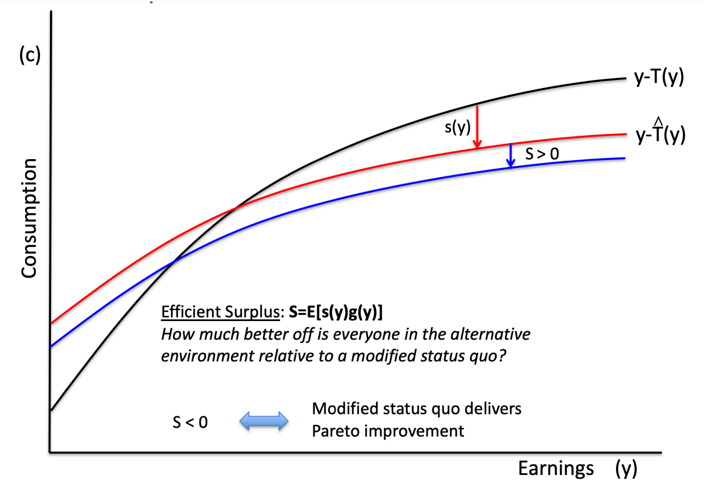
EV Measure - Hicks Experiment



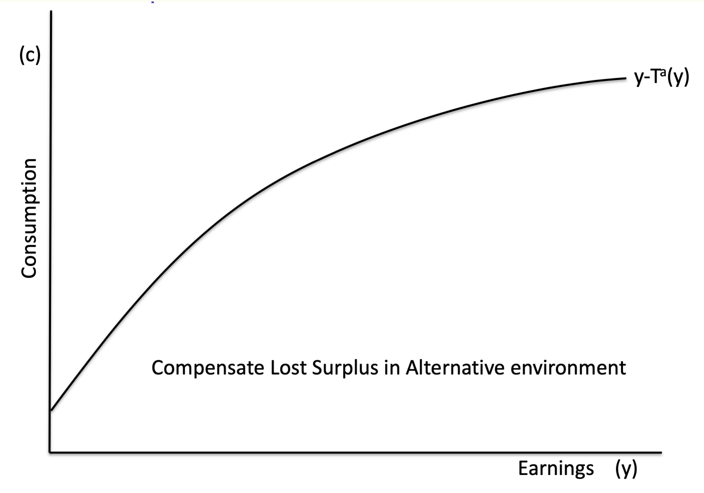
EV Measure - Hicks Experiment



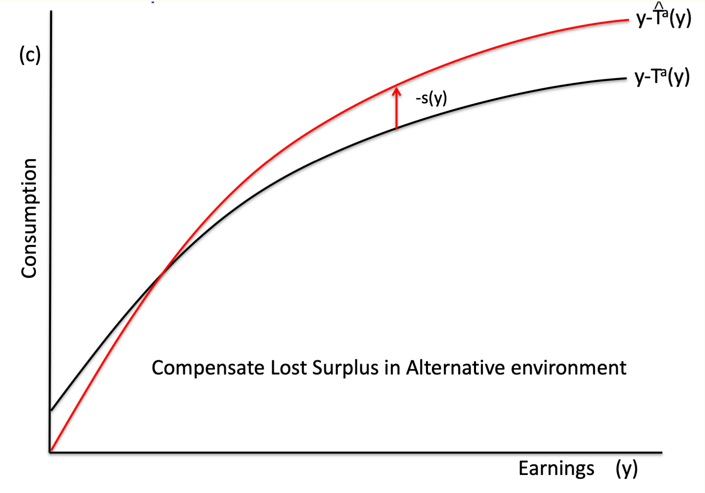
EV Measure - Hicks Experiment



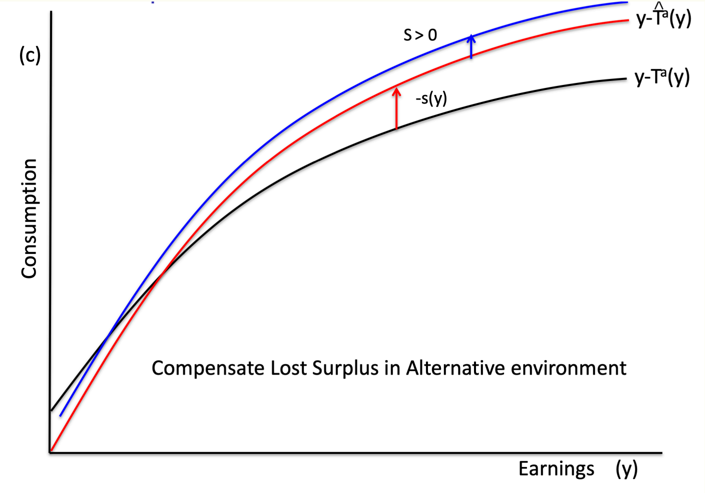
CV Measure - Kaldor Experiment



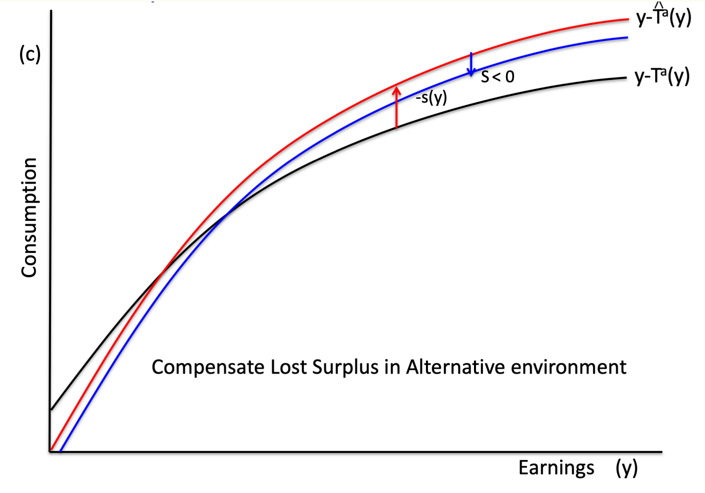
CV Measure - Kaldor Experiment



CV Measure - Kaldor Experiment



CV Measure - Kaldor Experiment



Kaldor and Hicks Experiments: Summary

Table 1

	$S > 0$	$S < 0$
Hicks Experiment: Possible to replicate s(y) using tax cut in status quo?	No	Yes
Kaldor Experiment: Possible to modify alternative environment tax schedule to make everyone better off relative to status quo?	Yes	No

- $S < 0$ alternative environment inefficient, Hicks logic:
  - feasible modification to tax schedule in the status quo leads to a Pareto superior allocation to the alternative environment
- $S > 0$ alternative environment is preferred, Kaldor logic:
  - modified alternative environment with compensations offers Pareto superior allocation relative to status quo

Remarks on Weights $g(y)$

- $g(y)$: cost of providing \$1 of welfare to those earning y through change in tax schedule
  - higher weight to those more costly to reach!

- Two reasons to use these weights:

  1. Positive (Hendren 2020):
     - Can augment policy with benefit tax to make Pareto improvement
     - Prefer the policy by the (potential) Pareto principle
  2. Normative (Bourguignon and Spadaro):
     - $g(y)$ welfare weights that rationalize status quo as optimum

Inverse-Optimum Weights

- $\chi(\mathsf{y})$: social marginal utility of income to y
  - social value of giving \$1 to y
- $\chi^*(\mathsf{y})$ is set of weights that rationalize tax schedule as optimal
- Now, consider giving \$1 to y′
  - benefit: $\chi^*(\mathsf{y}')$
  - cost: $\mathsf{g}(\mathsf{y}')$
- If schedule optimal

$$\frac{\chi^*(\mathsf{y}')}{\mathsf{g}(\mathsf{y}')} = \kappa, \ \forall \mathsf{y}'$$

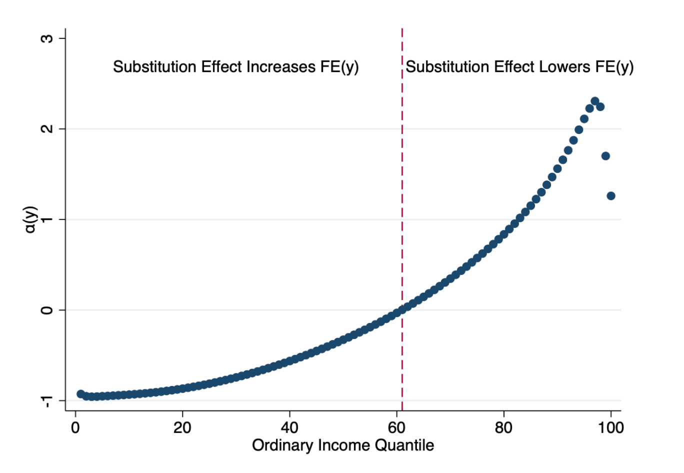
- $\chi^*(\mathsf{y})$ are defined up to a constant
- Hence, $\mathsf{g}(\mathsf{y})$ is the only set of weights that rationalizes tax schedule

Empirical Implementation

- Derive a formula for fiscal externality based on four parameters
  - compensated elasticity
  - income effect (elasticity of labor supply to income)
  - participation elasticity
  - shape of income distribution ($\alpha = -\left(1 + \frac{yf'(y)}{f(y)}\right)$)

- Calibrate the formula using estimates in literature

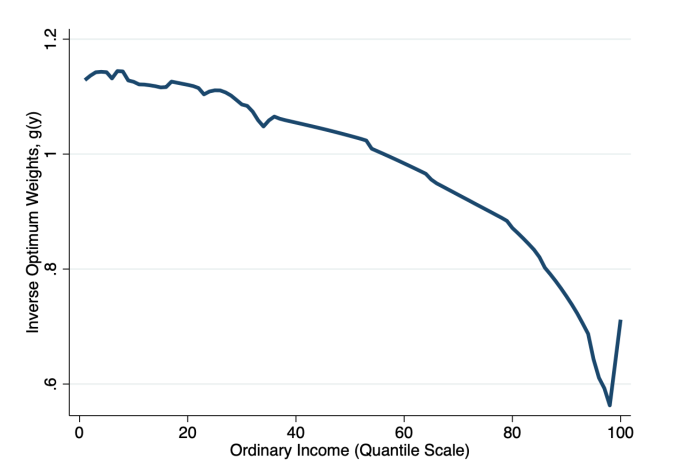
Shape of income distribution: $\alpha\,(y)$



After local tax cut η at y:

- $\alpha < 0$: more people decrease income
- $\alpha > 0$: more people increase income

Shape of $g(y) = 1 + FE(y)$



- $g(y) = 1.15$ at the bottom, 0.65 at the top
- Transferring \$1 from top generate 0.65/1.15 = \$0.57 of welfare to someone at the bottom

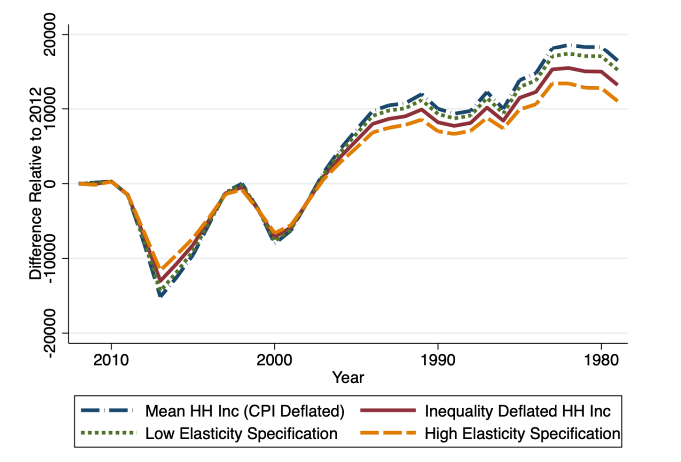
Deflating Economic Growth in the US

- Income inequality has raised over time, especially at the top
- Experiment (Kaldor original exercise, compare income distributions):
  - suppose we remove changes in income inequality
  - how richer incomes would be relative to reference year?
- Use the following surplus

$$S_{t,t-k} = \int_{\alpha} \left(Q_t(\alpha) - Q_{t-k}(\alpha) \right) g\left(Q_t(\alpha) \right) d\alpha$$

- $Q_t(\alpha)$: income in year t in α-quantile

Deflating Economic Growth in the US



- Increase of $18.3K since 1979, but only $15K when deflated
- 15-20% lower growth overall when weighted

Targeted Policies and MVPF

- Previous application compared income distribution
- What about non-budget neutral policy p on income y?
- WTP might be positive for everyone, but there is a cost!
- MVPF of p is

$$\mathrm{MVPF}_p = \frac{s^*(y)}{1 + FE^p}$$

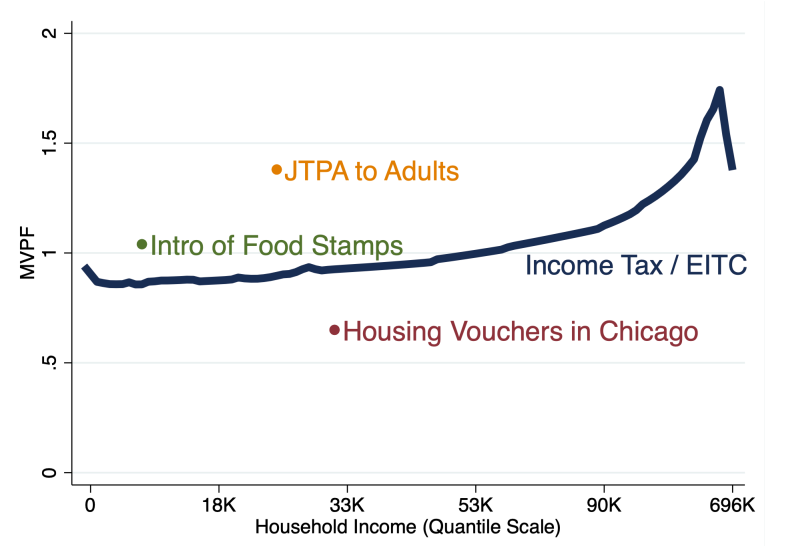
  depends on WTP and causal effects in FE^p

- $s^*(y)$: WTP of non-budget neutral policy
- Additional spending on p is good if

$$\underbrace{\mathrm{MVPF}_p}_{\text{Value of p}} \geq \underbrace{\frac{1}{1 + FE(y)}}_{\text{Value of } T(y)}$$

- Compare policy to value of tax cut to same y

Targeted Policies and MVPF



Limits of this approach

- Non-marginal transfers
  - marginal cost of the first dollar may not equal marginal cost of the last dollar of the transfers
  - $E[g(y)s(y)]$ might not be accurate
- No general equilibrium effects
  - if wage changes then WTP for \$1 tax cut $\neq 1$
- Weights are not structural
  - they depend on the current tax environment
  - e.g. Bourdignon and Spataro on France $g(y^{top}) \approx 0$