Public Economics Lectures
Efficiency Cost of Taxation and Optimal Taxation

John Karl Scholz (borrowing from Raj Chetty and Gregory A. Bruich)

University of Wisconsin – Madison
Fall 2011
Outline

1. Marshallian surplus
2. Definitions of EV, CV, and excess burden
3. Welfare Analysis in Behavioral Models
4. Optimal Taxation (once over lightly)
Definition

- Incidence analysis: effect of policies on **distribution** of economic pie
- Efficiency or deadweight cost: effect of policies on **size** of the pie
- Focus in efficiency analysis is on quantities, not prices
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 is reduced by more than $1 because the tax distorts incentives and behavior.

Core theory of public finance: how to implement policies that minimize these efficiency costs

- This basic framework for optimal taxation is adapted to study transfer programs, social insurance, etc.
- Start with positive analysis of how to measure efficiency cost of a given tax system
Marshallian Surplus: Assumptions

- Most basic analysis of efficiency costs is based on Marshallian surplus
- Two critical assumptions:
  1. Quasilinear utility (no income effects)
  2. Competitive production
Partial Equilibrium Model: Setup

- Two goods: $x$ and $y$

- Consumer has wealth $Z$, utility $u(x) + y$, and solves

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

- 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

$$pS - c(S)$$
Model: Equilibrium

- With perfect optimization, supply fn for $x$ is implicitly defined by the marginal condition
  \[ p = c'(S(p)) \]
- Let $\eta_S = p \frac{S'}{S}$ denote the price elasticity of supply
- Let $Q$ denote equilibrium quantity sold of good $x$
- $Q$ satisfies:
  \[ Q(t) = D(p + t) = S(p) \]
- Consider effect of introducing a small tax $d\tau > 0$ on $Q$ supplied
Excess Burden of Taxation

Price

$30.0

Quantity

1500

Excess Burden of Taxation

Public Economics Lectures

Efficiency & Optimal Tax

8 / 81
Excess Burden of Taxation

Price

$36.0

$30.0

Excess Burden

1500 1350

Excess Burden of Taxation

Price

Quantity

A

D

S

S+t

B

C

$36.0

$t

$30.0

Public Economics Lectures

Efficiency & Optimal Tax

9 / 81
1. In terms of supply and demand elasticities:

\[
EB = \frac{1}{2} dQ d\tau \\
EB = \frac{1}{2} S'(p) dp d\tau = \frac{1}{2} \left( \frac{pS'}{S} \right) \frac{\eta_D}{\eta_S - \eta_D} d\tau^2 \\
EB = \frac{1}{2} \frac{\eta_S \eta_D}{\eta_S - \eta_D} pQ \left( \frac{d\tau}{p} \right)^2
\]

- Note: second line uses incidence formula \( dp = \left( \frac{\eta_D}{\eta_S - \eta_D} \right) d\tau \)
- Tax revenue \( R = Q d\tau \)
- Useful expression is deadweight burden per dollar of tax revenue:

\[
\frac{EB}{R} = \frac{1}{2} \frac{\eta_S \eta_D}{\eta_S - \eta_D} \frac{d\tau}{p}
\]
Efficiency Cost: Qualitative Properties

\[ EB = \frac{1}{2} \frac{\eta_S \eta_D}{\eta_S - \eta_D} pQ \left( \frac{d\tau}{p} \right)^2 \]

1. Excess burden increases with square of tax rate
2. Excess burden increases with elasticities
3. Excess burden increases with the budget shares of the taxed commodity
EB Increases with Square of Tax Rate

\[ P \]

\[ Q \]

\[ D \]

\[ S \]

\[ A \]

\[ Q_1 \]

\[ P_1 \]
EB Increases with Square of Tax Rate

Public Economics Lectures
Efficiency & Optimal Tax
EB Increases with Square of Tax Rate

Change in EB

$\Delta t_2$

EB Increases with Square of Tax Rate
Comparative Statics

(a) Inelastic Demand

(b) Elastic Demand

Public Economics Lectures

Efficiency & Optimal Tax

15 / 81
With many goods, formula suggests that the most efficient way to raise tax revenue is:

1. Tax relatively more the inelastic goods (e.g. medical drugs, food)

2. Spread taxes across all goods so as to keep tax rates relatively low on all goods (broad tax base)
General Model with Income Effects

- Marshallian surplus is an ill-defined measure with income effects

- Question of interest: how much utility is lost because of tax beyond revenue transferred to government?

- Need units to measure “utility loss”

- Introduce expenditure function to translate the utility loss into dollars (money metric)
Expenditure Function

- Fix utility at $U$ and prices at $q$ where $q = p + t$ denotes vector of tax-inclusive prices
- Find bundle that minimizes cost to reach $U$ for $q$:
  \[ e(q, U) = \min_C q \cdot c \quad \text{s.t.} \quad u(c) \geq U \]
- Let $\mu$ denote multiplier on utility constraint
- First order conditions given by:
  \[ q_i = \mu u_{c_i} \]
- These generate Hicksian (or compensated) demand fns:
  \[ c_i = h_i(q, u) \]
- Define individual’s loss from tax increase as
  \[ e(q^1, u) - e(q^0, u) \]
- Single-valued function $\rightarrow$ coherent measure of welfare cost, no path dependence
But where should $u$ be measured?

Consider a price change from $q^0$ to $q^1$

Initial utility:

$$u^0 = v(q^0, Z)$$

Utility at new price $q^1$:

$$u^1 = v(q^1, Z)$$

Two concepts: compensating ($CV$) and equivalent variation ($EV$) use $u^0$ and $u^1$ as reference utility levels.
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?

- \(CV\) is amount of ex-post cost that must be covered by government 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
- 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:

\[ e(q^0, u^1) + EV = e(q^1, u^1) \]
Goal: derive empirically implementable formula analogous to Marshallian EB formula in general model with income effects

Existing literature assumes either

1. Fixed producer prices and income effects
2. Endogenous producer prices and quasilinear utility

With both endogenous prices and income effects, efficiency cost depends on how profits are returned to consumers

Formulas are very messy and fragile (Auerbach section 3.2)
Efficiency Cost Formulas with Income Effects

- Derive empirically implementable formulas using Hicksian demand ($EV$ and $CV$)

- Assume $p$ is fixed → flat supply, constant returns to scale

- The envelope thm implies that $e_{q_i}(q, u) = h_i$, and so:

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

- If only one price is changing, this is the area under the Hicksian demand curve for that good

- Note that optimization implies that

$$h(q, v(q, Z)) = c(q, Z)$$
Compensating vs. Equivalent Variation

\[ h(V(p_1, Z)) \quad h(V(p_0, Z)) \]

\[ D \]

\[ p \quad p_0 \quad p_1 \quad h(V(p_1, Z)) \quad h(V(p_0, Z)) \]

\[ x(p_1, Z) \quad x(p_0, Z) \quad x \]

Public Economics Lectures

Efficiency & Optimal Tax

24 / 81
Compensating vs. Equivalent Variation

\[ h(V(p_1, Z)) h(V(p_0, Z)) \]

\[ x(p_0, Z) x(p_1, Z) \]
Compensating vs. Equivalent Variation

\[ h(V(p_1, Z)) \quad h(V(p_0, Z)) \]

\[ x(p_0, Z) \quad x(p_1, Z) \]

\[ CV \]

\[ D \]

Public Economics Lectures
Efficiency & Optimal Tax
26 / 81
Marshallian Surplus

\[ h(V(p_1, Z)) \quad \text{and} \quad h(V(p_0, Z)) \]

\[ x(p_0, Z) \quad \text{and} \quad x(p_1, Z) \]
Excess Burden

- Deadweight burden: change in consumer surplus less tax paid

- Equals what is lost in excess of taxes paid

- Two measures, corresponding to $EV$ and $CV$:

\[
EB(u^1) = EV - (q^1 - q^0)h(q^1, u^1) \quad [Mohring 1971]
\]

\[
EB(u^0) = CV - (q^1 - q^0)h(q^1, u^0) \quad [Diamond and McFadden 1974]
\]
Public Economics Lectures ()

Efficiency & Optimal Tax

29 / 81
Marshallian

$h(V(p_1, Z))$

$h(V(p_0, Z))$

$x(p_1, Z)$

$x_C(p_1, V(p_0, Z))$

$p$

$p_0$

$p_1$

$x$

$D$
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

- This is a big deal – DWL arises solely from substitution effects!
- Chipman and Moore (1980): $CV = EV = $ Marshallian DWL only with quasilinear utility
Consider increase in tax $\tau$ 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:

$$MEB = EB(\tau + \Delta \tau) - EB(\tau)$$

$$\approx \frac{dEB}{d\tau}(\Delta \tau) + \frac{1}{2}(\Delta \tau)^2 \frac{d^2EB}{d\tau^2}$$
Harberger Trapezoid Formula

\[
\frac{dEB}{d\tau} = h_1(p + \tau, U) - \tau \frac{dh_1}{d\tau} - h_1(p + \tau, U)
\]

\[
= -\tau \frac{dh_1}{d\tau}
\]

\[
\frac{d^2 EB}{d\tau^2} = -\frac{dh_1}{d\tau} - \tau \frac{d^2 h_1}{d\tau^2}
\]

- Standard practice in literature: assume \( \frac{d^2 h_1}{d\tau^2} \) (linear Hicksian); not necessarily well justified because 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
Without pre-existing tax, obtain “standard” Harberger formula:

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

Observe that first-order term vanishes when \( \tau = 0 \)

A new tax has second-order deadweight burden (proportional to \( \Delta \tau^2 \) not \( \Delta \tau \))

Bottom line: need compensated (substitution) elasticities to compute \( EB \), not uncompensated elasticities

Empirically, need estimates of income and price elasticities (using the slutsky relationship to uncover compensated elasticities)
Excess Burden with Taxes on Multiple Goods

- Previous formulas apply to case with tax on one good

- With multiple goods and fixed prices, excess burden of introducing a tax $\tau_k$

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

- Second-order effect in own market, first-order effect from other markets with pre-existing taxes

- Hard to implement because we need all cross-price elasticities

- Complementarity between goods important for excess burden calculations

- Ex: with an income tax, minimize total DWL tax by taxing goods complementary to leisure (Corlett and Hague 1953)
Chetty, Looney, and Kroft (2009) section 5

Derive partial-equilibrium formulas for incidence and efficiency costs

Focus here on efficiency cost analysis
Welfare Analysis with Salience Effects: Setup

- Two goods, \( x \) and \( y \); price of \( y \) is 1, pretax price of \( x \) is \( p \).

- Taxes: \( y \) untaxed. Unit sales tax on \( x \) at rate \( t^S \), which is not included in the posted price.

- Tax-inclusive price of \( x \): \( q = p + t^S \)
Representative consumer has wealth $Z$ and utility $u(x) + v(y)$

Let $\{x^*(p, t^S, Z), y^*(p, t^S, Z)\}$ denote bundle chosen by a fully-optimizing agent

Let $\{x(p, t^S, Z), y(p, t^S, Z)\}$ denote empirically observed demands

Place no structure on these demand functions except for feasibility:

$$(p + t^S)x(p, t^S, Z) + y(p, t^S, Z) = Z$$
Price-taking firms use $y$ to produce $x$ with cost fn. $c$

Firms optimize perfectly. Supply function $S(p)$ defined by:

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

Let $\varepsilon_S = \frac{\partial S}{\partial p} \times \frac{p}{S(p)}$ denote the price elasticity of supply
Define excess burden using EV concept

Excess burden (EB) of introducing a revenue-generating sales tax $t$ is:

$$EB(t^S) = Z - e(p, 0, V(p, t^S, Z)) - R(p, t^S, Z)$$
Preference Recovery Assumptions

**A1** Taxes affect utility only through their effects on the chosen consumption bundle. Agent’s indirect utility given taxes of \((t^E, t^S)\) is

\[
V(p, t^S, Z) = u(x(p, t^S, Z)) + v(y(p, t^S, Z))
\]

**A2** When tax inclusive prices are fully salient, the agent chooses the same allocation as a fully-optimizing agent:

\[
x(p, 0, Z) = x^*(p, 0, Z) = \arg \max_x u(x) + v(Z - px)
\]

- A1 analogous to specification of ancillary condition; A2 analogous to refinement
Two steps in efficiency calculation:

1. Use price-demand $x(p, 0, Z)$ to recover utility as in standard model.

2. Use tax-demand $x(p, t^S, Z)$ to calculate $V(p, t^S, Z)$ and EB.
Excess Burden with No Income Effect for Good $x \times (\frac{\partial x}{\partial Z} = 0)$

Source: Chetty, Looney, and Kroft (2009)
In the case without income effects \((\frac{\partial x}{\partial Z} = 0)\), which implies utility is quasilinear, excess burden of introducing a small tax \(t^S\) is

\[
EB(t^S) \approx -\frac{1}{2} (t^S)^2 \frac{\partial x/\partial t^S}{\partial x/\partial p} \frac{\partial x}{\partial t^S}
\]

\[
= \frac{1}{2} (\theta t^S)^2 \frac{\varepsilon_D}{p + t^S}
\]

Inattention reduces excess burden when \(dx/dZ = 0\).

Intuition: tax \(t^S\) induces behavioral response equivalent to a fully perceived tax of \(\theta t^S\).

If \(\theta = 0\), tax is equivalent to a lump sum tax and \(EB = 0\) because agent continues to choose first-best allocation.
Efficiency Cost with Income Effects

- Same formula, but all elasticities are now compensated:

\[
EB(t^S) \approx -\frac{1}{2}(t^S)^2 \frac{\partial x^c}{\partial x^c / \partial p} \frac{\partial x^c / \partial t^S}{\partial x^c / \partial p} \\
= \frac{1}{2}(\theta^c t^S)^2 \frac{\varepsilon_D^c}{p + t^S}
\]

- Compensated price demand: \(dx^c / dp = dx / dp + xdx / dZ\)

- Compensated tax demand: \(dx^c / dt^S = dx / dt^S + xdx / dZ\)

- Compensated tax demand does not necessarily satisfy Slutsky condition \(dx^c / dt^S < 0\) b/c it is not generated by utility maximization
Efficiency Cost with Income Effects

\[ EB(t^S) \approx -\frac{1}{2} (t^S)^2 \frac{\partial x^c / \partial t^S}{\partial x^c / \partial p} \frac{\partial x^c / \partial t^S}{\partial p} \]

\[ = \frac{1}{2} (\theta^c t^S)^2 \frac{\varepsilon_c^c}{p + t^S} \]

- With income effects \((dx / dZ > 0)\), making a tax less salient can **raise** deadweight loss.

- Tax can generate \(EB > 0\) even if \(dx / dt^S = 0\)

- Example: consumption of food and cars; agent who ignores tax on cars underconsumes food and has lower welfare.

- Intuition: agent does not adjust consumption of \(x\) despite change in net-of-tax income, leading to a positive compensated elasticity.
Brief Look at Optimal Commodity (and Income) Taxation: Introduction

- Now combine lessons on incidence and efficiency costs to analyze optimal design of commodity taxes
- What is the best way to design taxes given equity and efficiency concerns?
- Optimal commodity tax literature focuses on linear ($t \cdot x$) tax system
- Non-linear ($t(x)$) tax systems considered in income tax literature
Second Welfare Theorem

- Starting point: second-welfare theorem
- Can achieve any Pareto-efficient allocation as a competitive equilibrium with appropriate lump-sum transfers
- Requires same assumptions as first welfare theorem plus one more:
  1. Complete markets (no externalities)
  2. Perfect information
  3. Perfect competition
  4. Lump-sum taxes/transfers across individuals feasible
- If 1-4 hold, equity-efficiency trade-off disappears and optimal tax problem is trivial
  - Simply implement lump sum taxes that meet distributional goals given revenue requirement
- Problem: information
To set the optimal lump-sum taxes, need to know the characteristics (ability) of each individual

But no way to make people reveal their ability at no cost

- Incentive to misrepresent skill level

Tax instruments are therefore a fn. of economic outcomes

- E.g. income, property, consumption of goods

→ Distorts prices, affecting behavior and generating DWL

Information constraints force us to move from the 1st best world of the second welfare theorem to the 2nd best world with inefficient taxation

- Cannot redistribute or raise revenue for public goods without generating efficiency costs
Four Central Results in Optimal Tax Theory

1. Ramsey (1927): inverse elasticity rule


3. Atkinson and Stiglitz (1976): no consumption taxation with optimal non-linear (including lump sum) income taxation


I will briefly mention features of these results but we will not spend the time to derive them.
Ramsey (1927) Tax Problem

- Government sets taxes on uses of income in order to accomplish two objectives:
  
  1. Raise total revenue of amount $E$
  
  2. Minimize utility loss for agents in economy

- Originally a problem set that Pigou assigned Ramsey
Ramsey Model: Key Assumptions

1. Lump sum taxation prohibited
2. Cannot tax all commodities (leisure untaxed)
3. Production prices fixed (and normalized to one):

   \[ p_i = 1 \]

   \[ \Rightarrow q_i = 1 + \tau_i \]
Ramsey Model: Setup

- One individual (no redistributive concerns) with utility
  \[ u(x_1, \ldots, x_N, l) \]

subject to budget constraint

\[ q_1 x_1 + \ldots + q_N x_N \leq wl + Z \]

- \( Z = \) non wage income, \( w = \) wage rate

- Consumption prices are \( q_i \)
Ramsey Model: Consumer Behavior

- Lagrangian for individual’s maximization problem:

\[ \mathcal{L} = u(x_1, \ldots, x_N, l) + \alpha(wl + Z - (q_1x_1 + \ldots + q_Nx_N)) \]

- First order condition:

\[ u_{x_i} = \alpha q_i \]

Where \( \alpha = \partial V / \partial Z \) is marginal value of money for the individual.

- Yields demand functions \( x_i(q, Z) \) and indirect utility function \( V(q, Z) \) where \( q = (w, q_1, \ldots, q_N) \).
Ramsey Model: Government’s Problem

- Government solves either the maximization problem

\[ \max V(q, Z) \]

subject to the revenue requirement

\[ \tau \cdot x = \sum_{i=1}^{N} \tau_i x_i(q, Z) \geq E \]

- Or, equivalently, minimize excess burden of the tax system

\[ \min EB(q) = e(q, V(q, Z)) - e(p, V(q, Z)) - E \]

subject to the same revenue requirement
Ramsey Model: Government’s Problem

For maximization problem, Lagrangian for government is:

\[ \mathcal{L}_G = V(q, Z) + \lambda \left[ \sum_i \tau_i x_i(q, Z) - E \right] \]

\[ \implies \frac{\partial \mathcal{L}_G}{\partial q_i} = \frac{\partial V}{\partial q_i} + \lambda \left[ x_i + \sum_j \tau_j \frac{\partial x_j}{\partial q_i} \right] = 0 \]

Using Roy’s identity (\( \frac{\partial V}{\partial q_i} = -\alpha x_i \)):

\[ (\lambda - \alpha) x_i + \lambda \sum_j \tau_j \frac{\partial x_j}{\partial q_i} = 0 \]
Optimal tax rates satisfy system of $N$ equations and $N$ unknowns:

$$
\sum_j \tau_j \frac{\partial x_j}{\partial q_i} = -\frac{x_i}{\lambda} (\lambda - \alpha)
$$

Same formula can be derived using a perturbation argument, which is more intuitive.
Suppose government increases $\tau_i$ by $d\tau_i$

Effect of tax increase on social welfare is sum of effect on government revenue and private surplus

Marginal effect on government revenue:

$$dR = x_i d\tau_i + \sum_j^\tau_j dx_j$$

Marginal effect on private surplus:

$$dU = \frac{\partial V}{\partial q_i} d\tau_i = -\alpha x_i d\tau_i$$

Optimum characterized by balancing the two marginal effects:

$$dU + \lambda dR = 0$$
Rewrite in terms of Hicksian elasticities to obtain further intuition using Slutsky equation:

\[ \frac{\partial x_j}{\partial q_i} = \frac{\partial h_j}{\partial q_i} - x_i \frac{\partial x_j}{\partial Z} \]

Substitution into formula above yields:

\[
(\lambda - \alpha)x_i + \lambda \sum_j \tau_j \left[ \frac{\partial h_j}{\partial q_i} - x_i \frac{\partial x_j}{\partial Z} \right] = 0
\]

\[
\Rightarrow \frac{1}{x_i} \sum_j \tau_j \frac{\partial h_i}{\partial q_j} = -\frac{\theta}{\lambda}
\]

where \( \theta = \lambda - \alpha - \lambda \frac{\partial}{\partial Z} (\sum_j \tau_j x_j) \)
\( \theta \) is independent of \( i \) and measures the value for the government of introducing a $1 lump sum tax

\[
\theta = \lambda - \alpha - \lambda \frac{\partial (\sum_j \tau_j x_j)}{\partial Z} 
\]

Three effects of introducing a $1 lumpsum tax:

1. Direct value for the government is \( \lambda \)
2. Loss in welfare for the individual is \( \alpha \)
3. Behavioral effect \( \Rightarrow \) loss in tax revenue of \( \frac{\partial (\sum_j \tau_j x_j)}{\partial Z} \)

Can demonstrate that \( \theta > 0 \Rightarrow \lambda > \alpha \) at the optimum using Slutsky matrix
Intuition for Ramsey Formula: Index of Discouragement

\[ \frac{1}{x_i} \sum_j \tau_j \frac{\partial h_i}{\partial q_j} = -\frac{\theta}{\lambda} \]

- Suppose revenue requirement $E$ is small so that all taxes are also small.
- Then tax $\tau_j$ on good $j$ reduces consumption of good $i$ (holding utility constant) by approximately

\[ dh_i = \tau_j \frac{\partial h_i}{\partial q_j} \]

- Numerator of LHS: total reduction in consumption of good $i$.
- Dividing by $x_i$ yields \% reduction in consumption of each good $i = \text{“index of discouragement”}$ of the tax system on good $i$.
- Ramsey tax formula says that the indexes of discouragements must be equal across goods at the optimum.
Special Case 1: Inverse Elasticity Rule

- Introducing elasticities, we can write formula as:

\[
\sum_{j=1}^{N} \frac{\tau_j}{1 + \tau_j} \epsilon_{ij}^c = \frac{\theta}{\lambda}
\]

- Consider special case where \( \epsilon_{ij} = 0 \) if \( i \neq j \)

- Slutsky matrix is diagonal

- Obtain classic inverse elasticity rule:

\[
\frac{\tau_i}{1 + \tau_i} = \frac{\theta}{\lambda} \frac{1}{\epsilon_{ii}}
\]
Special Case 2: Uniform Taxation

- Suppose $\varepsilon_{ij} = 0$ if $i \neq j$ and $\varepsilon_{x_i, w} = \frac{\partial h_i}{w} \text{ constant}$

- Using following identity, $\sum_j \frac{\partial h_i}{\partial q_j} q_j + \frac{\partial h_i}{\partial w} w = 0$, we obtain

  $$\frac{\partial h_i}{\partial q_i} q_i = - \frac{\partial h_i}{\partial w} w$$

- Proof of identity ($J$ good economy, no labor):

  $$\sum_j \frac{\partial h_i}{\partial q_j} q_j = \sum_{j \neq i} \frac{\partial h_j}{\partial q_i} q_j + \frac{\partial h_i}{\partial q_i} q_i$$

  $$= \sum_{j \neq i} \frac{\partial h_j q_j}{\partial q_i} + \frac{\partial h_i q_i}{\partial q_i} - h_i$$

  $$= \frac{\partial e}{\partial q_i} - h_i = 0$$
Special Case 2: Uniform Taxation

- Then immediately obtain

\[ \frac{1}{x_i} \tau_i = -\frac{\theta}{\lambda} \frac{1}{\frac{\partial h_i}{\partial q_i}} = \frac{\theta}{\lambda} \frac{1}{\frac{\partial h_i}{\partial w} w} \]

\[ \frac{\tau_i}{q_i} = \frac{\theta}{\lambda} \frac{1}{\frac{\partial h_i}{\partial w} w x_i} = \frac{\theta}{\lambda} \varepsilon_{x_i,w} \]

- With constant \( \varepsilon_{x_i,w} \), \( \frac{\tau_i}{q_i} \) is constant \( \rightarrow \) uniform taxation

- Corlett and Hague (1953): 3 good model, uniform tax optimal if all goods are equally complementary with labor (and labor is untaxed)

- More generally, lower taxes for goods complementary to labor
Ramsey Formula: Limitations

- Ramsey solution: tax inelastic goods to minimize efficiency costs
- But does not take into account redistributive motives
- Presumably necessities are more inelastic than luxuries
- Therefore, optimal Ramsey tax system is likely to be regressive
- Diamond (1975) extends Ramsey model to take redistributive motives into account
Diamond and Mirrlees (1971)

- Previous analysis assumed fixed producer prices
- Diamond and Mirrlees (1971) relax this assumption by modelling production

Two major results

1. Production efficiency: even in an economy where first-best is unattainable, optimal policy maintains production efficiency
2. Characterize optimal tax rates with endogenous prices and show that Ramsey rule can be applied
Lipsey and Lancaster (1956): Theory of the Second Best

- Standard optimal policy results only hold with single deviation from first best
  - Ex: Ramsey formulas invalid if there are pre-existing distortions, imperfect competition, etc.

- In second-best, anything is possible
  - Policy changes that would increase welfare in a model with a single deviation from first best need not do so in second-best
    - Ex: tariffs can improve welfare by reducing distortions in other part of economy

- Destructive result for welfare economics
Diamond and Mirrlees result was an advance because it showed a general policy lesson even in second-best environment.

Example: Suppose government can tax consumption goods and also produces some goods on its own (e.g. postal services).

May have intuition that government should try to generate profits in postal services by increasing the price of stamps.

This intuition is wrong: optimal to have no distortions in production of goods.

Bottom line: only tax goods that appear directly in agent’s utility functions.

Should not distort production decisions via taxes on intermediate goods, tariffs, etc.
Policy Consequences: Public Sector Production

- Public sector production should be efficient

- If there is a public sector producing some goods, it should:
  - Face the same prices as the private sector
  - Choose production with the unique goal of maximizing profits, not generating government revenue

- Ex. postal services, electricity, health care, ...
Policy Consequences: No Taxation of Intermediate Goods

- **Intermediate goods**: goods that are neither direct inputs or outputs to indiv. consumption
- Taxes on transactions between firms would distort production
Computers:
- Sales to firms should be untaxed
- But sales to consumers should be taxed

In practice, tax policy often follows precisely the opposite rule

Ex. Diesel fuel tax studied by Marion and Muehlegger (2008)
Result hinges on key assumptions about govt’s ability to:

1. Set a full set of differentiated tax rates on each input and output
2. Tax away fully pure profits (or production is constant-returns-to-scale)

A2 rules out improving welfare by taxing profitable industries to improve distribution at expense of production efficiency.

These assumptions effectively separate the production and consumption problems.
Diamond and Mirrlees Result: Limitations

- **Practical relevance** of the result is a bit less clear

- Ex. Assumption 1 is not realistic (Naito 1999)

- Skilled and unskilled labor inputs ought to be differentiated

- Not the case in current income tax system

- In such cases, may be optimal to:

  1. Subsidize low skilled intensive industries

  2. Set tariffs on low skilled intensive imported goods (to protect domestic industry)
Muirreles 1971: Optimal Income Tax Problem Incorporating Behavioral Responses

- Standard labor supply model: Individual maximizes
  \[ u(c, l) \text{ s.t. } c = wl - T(wl) \]
  where \( c \) is consumption, \( l \) labor supply, \( w \) wage rate, \( T(.) \) income tax

- Individuals differ in ability \( w \) distributed with density \( f(w) \)

- Govt social welfare maximization: Govt maximizes
  \[ SWF = \int G(u(c, l))f(w)dw \]
  s.t. resource constraint
  \[ \int T(wl)f(w)dw \geq E \]
  and individual FOC
  \[ w(1 - T')u_c + ul = 0 \]
  where \( G(.) \) is increasing and concave
**Mirrlees 1971: Results**

- Optimal income tax trades-off redistribution and efficiency
  - \( T(.) < 0 \) at bottom (transfer)
  - \( T(.) > 0 \) further up (tax) [full integration of taxes/transfers]

- Mirrlees formulas are a complex fn. of primitives, with only a few general results
  1. \( 0 \leq T'(.) \leq 1, \quad T'(.) \geq 0 \) is non-trivial and rules out EITC [Seade 1976]
  2. Marginal tax rate \( T'(.) \) should be zero at the top if skill distribution bounded [Sadka-Seade]
Mirrlees: Subsequent Work

- Mirrlees model had a profound impact on information economics
  - Ex. models with asymmetric information in contract theory
- But until late 1990s, Mirrlees had little impact on practical tax policy
- Recently, Mirrlees model connected to empirical literature
  - Diamond (1998), Piketty (1997), and Saez (2001)
  - See Bernheim paper on Saez’s work for one nice description.
Commodity vs. Income Taxation

Now combine commodity tax and income tax results to analyze optimal combination of policies.

In practice, government levies differential commodity taxes along with non-linear income tax.

1. Exempts some goods (food, education, health) from sales tax
2. Imposes additional excise taxes on some goods (cars, gasoline, luxury goods)
3. Imposes capital income taxes

What is the best combination of taxes?
• With separability and homogeneity, conditional on earnings $z$, consumption choices $c = (c_1, \ldots, c_K)$ do not provide any information on ability.

• Differentiated commodity taxes $t_1, \ldots, t_K$ create a tax distortion with no benefit.
  
  • Better to do all the redistribution with the individual income tax.

• With only linear income taxation (Diamond-Mirrlees 1971, Diamond 1975), diff. commodity taxation can be useful to “non-linearize” the tax system.
  
  • But not if Engel curves for each $c_k$ are linear in $y$ (Deaton 1981).
Failures of A-S Assumptions

- If higher ability consume more of good $k$ than lower ability people, then taxing good $k$ is desirable. Examples:

  1. High ability people have a relatively higher taste for good $k$ (at a given income)
     - Luxury chocolates or museums; violates homogeneous $v(c)$ assumption

  2. Good $k$ is positively related to leisure (consumption of $k$ increases when leisure increases at a given income)
     - Tax on travel, subsidy on computers and work related expenses

- In general Atkinson-Stiglitz assumptions are viewed as a good starting place for most goods
Judd (1985) and Chamley (1986) examine capital taxation

Consider a Ramsey model where govt. is limited to linear distortionary taxes

Result: optimal capital tax converges to zero in long run

Intuition: DWL rises with square of tax rate

- With non-zero capital tax, have an infinite price distortion between $c_0$ and $c_t$ as $t \to \infty$

- Undesirable to have such large distortions on some margins
Maybe capital taxes shouldn’t be zero for everyone.

When preferences are separable between consumption and leisure, the wealthy in \((t+1)\) are harder to motivate.

At the margin, therefore, good tax systems will deter wealth accumulation from period \(t\) to period \((t+1)\) to provide better work incentives in the later period.

More generally, when people are working hard, you want low asset tax rates. When people have little labor income, you want high asset tax rates.

In general, results on optimal capital income taxation are pretty fragile.