I want to thank Raj Chetty for sharing his slides on public goods, which form the basis for Section 3 of this lecture.
Suppose we have a policy that spends more on $G$ targeted towards those earning around $\$y$ of income.

Need to calculate:

- Individuals WTP out of their own income for additional $G$,
  \[ s(y) = \frac{\partial u_i}{\partial G} = \frac{u_G}{u_c} \]
  (assume homogenous WTP conditional on income)

- Total cost to the government inclusive of fiscal externalities
  \[ 1 + FE_G = \frac{d}{dG} [q], \text{ where } q \text{ is the aggregate govt budget} \]

- Construct MVPF for each individual with earnings $y$

  \[ MVPF(y) = \frac{s(y)}{1 + FE_G} \]
Recap: Aggregation

- Aggregate using either:
  - [SWF] Social marginal utilities of income, $\int \eta (y) \text{MVPF} (y)$
  - [Kaldor-Hicks/Kaplow/Mirrlees 1976] Marginal cost of redistributing to those with income $y$, $1 + FE (y)$

$$W = \int (1 + FE (y)) \text{MVPF} (y)$$

- Implicitly compare efficiency of $G$ to efficiency of redistribution through modifications to tax schedule, $T (y)$
Key Difficulty: Estimating $FE$...

- Implementing these formulae require estimating two fiscal externalities:
  - Impact of $G$ on tax revenue, $FE_G$
  - Impact of tax changes to those earning $y$ on tax revenue, $FE(y)$, for all $y$

- Why are these difficult?
  - Dynamics (impact on tax revenue in 30 years...)
  - Bases (impact of income tax changes on capital taxes, sales taxes, food stamp participation, etc...)
  - And, need rich variation in tax policies to identify $FE(y)$ for all $y$
    - Made progress in Topic 2 by assuming constant taxable income elasticity/etc.

- This lecture: potentially able to ignore all behavioral responses
  - Literature on optimal commodity taxation and optimal public goods
  - Key (weak?) assumption reduces these empirical requirements: “weak separability”
Basic Idea

- Begin with a roadmap of the basic idea
- Many economic models imply a relationship between $FE_G$ and $FE(y)$
- The social benefit of $1$ of spending on $G$ is given by:

$$W = \int (1 + FE(y)) s(y) \, dy$$

- Cost is given by $1 + FE_G$
- So, additional spending can increase welfare if and only if

$$\int (1 + FE(y)) s(y) \, dy \geq 1 + FE_G$$

or

$$\int s(y) \, dy \geq \int s(y) FE(y) \, dy - FE_G$$
Key insight: In many cases, reasonable to think that

$$\int s(y) \, FE(y) \, dy = FE_G$$

Why?
Increase $G$ to those earning near $y^*$.
Value: $dG^* (u_G / u_c)$

Consumption

Earnings ($y$)
How big are behavioral responses?

(c)

Consumption

\[ y - T(y) \]

Earnings (y)

\[ y^* \]

\[ dG \]

\[ \epsilon \]
Insight: If $G$ is "like $y$", then similar behavioral response

Distort behavior by $dG^*(u_G/u_c)$

$\text{Consumption}$

$y - T(y)$

$y^*$

$\varepsilon$
$FE_G = s(y)FE(y^*)$, where $s(y) = \frac{u_G}{u_c}$
Basic Idea

- If “G is like y”, then \( \int s(y) \text{FE}(y) \, dy = \text{FE}_G \), so that additional G can generate a potential Pareto improvement iff aggregate (unweighted!) surplus is positive:

  \[
  \int s(y) \, dy > 0
  \]

- Key question: What does it mean for G to be “like y”?
  - Will formalize as weak separability of utility
Explore these ideas in two broad context that have been focus of previous literature

- Public goods: Do we subsidize if public good disproportionately helps poor?
- Commodities / in-kind subsidies: Do we subsidize if commodity disproportionately consumed by poor?

Along the way, discuss other implications/related results

- Diamond-Mirrlees “Production efficiency” result
- Zero capital taxation result
What are Pure public goods?

- Non-rival: My consumption doesn’t prevent your consumption
- Non-excludable: Provider can’t prevent consumption by those who don’t pay

Public Goods benefit several individuals simultaneously

- Lowers effective cost of additional $G$

Why might the free market under-provide public goods?

- Free-riding
- Public goods create positive externalities, individuals under-provide
First Welfare Theorem: Any market equilibrium is Pareto Optimal

- With public goods, this fails
- Samuelson (1954) derives condition for a Pareto Optimum

Consider First Welfare Theorem setup:

- Individuals indexed by $i$, two goods, $X$ and $G$
- Utility functions $U^i(x_i, G_i)$, standard budget constraint
- $c$ is the dollar cost of producing $G$. (Normalize price of $x$ to 1 so $\frac{p_G}{p_x} = c$)

Condition for private optimality

$$ s_i = \frac{U_G(x_i, G_i)}{U_X(x_i, G_i)} = c \iff s_i = c \forall i $$
Now, suppose G is a public good
- So each person purchases $G_i$, but values $G = \sum_i G_i$
- Utility is $U(x_i, G) = U(x_i, G_i + \sum_{j \neq i} G_j)$

Condition for private optimality
- Still $\frac{U_G(x_i, G)}{U_x(x_i, G)} = c \iff s_i = c \ \forall i$
  - FOC will determine private contribution to public good

But, unweighted social surplus is maximized when

$$\sum_i s_i = c$$
Can the government help?
- Direct provision can avoid the free-rider problem

What is the optimal level of public provision of $G$?
- Samuelson (1954): Pareto efficiency requires maximizing surplus:
  $$\sum_i s_i = c$$

How can we decentralize this?
- If $\sum MRS_i = c$, then government can find transfers, $t_i$, and a change in $g$ to make everyone better off
  - Set $t_i = MRS_i$
- But, if we have individual specific lump-sum transfers, what does this say about the social marginal utility of income for rich and poor?
  - Should be equalized!
But, we transfer based on observed income
- Implies transfers are distortionary!

What does this mean for optimal public goods? Can still consider taxing back the benefits to each individual $i$:

$$\int s_i \left( 1 + FE(y_i) \right) \, di \geq 1 + FE_G$$

But, now we need to estimate $FE(y)$ and $FE_G$!
- Can we do something simpler?
Utility is a function of:
- A (private) consumption good, $c$
- The level of government expenditure on a publicly provided good, $g$
  (same as “$G$” in previous lectures)
- Labor supply $l$

Utility satisfies weak separability: there exists a function $v$ (common to all individuals) such that utility is given by

$$u(v(c,g), l)$$

Individuals differ in their wage, $w$

Consumption given by budget constraint

$$c = wl - T(wl, g)$$

where $T(wl, g)$ is the tax/transfers to individuals with earnings $wl$

- Cannot transfer based on (unobserved) wage, $w$
Social welfare given by

\[ SW = \int W(U(v(c, g), l)) f(w) \, dw \]

Government revenue given by

\[ R = \int T(wl(w), g) f(w) \, dw \]

where \( l(w) \) is the labor supply choice of type \( w \)

Social objective: Choose \( g \) to maximize \( SW \) subject to \( R = g \)
Kaplow (2006): Benefit Absorbing Tax

- What is the optimal level of $g$?
- Consider a policy that increases $g$ by a small amount
- Define a “benefit-absorbing tax” (analogous to last lecture...)
  - Change $T$ such that utility does not change when both $g$ and $T$ are simultaneously changed
  - Assume for now that $l$ will not change (will verify later)
  - Will solve implicitly for what the change in the tax schedule must be

The total derivative from the policy is given by:

$$\frac{\partial U}{\partial g} = \frac{\partial U}{\partial v} \left[v_c c_g + v_g\right]$$

- $v_c = \frac{\partial v}{\partial c}$ and $v_g = \frac{\partial v}{\partial g}$
- $c_g = -\frac{\partial T(wl,g)}{\partial g}$ is the partial derivative of how much consumption changes in response to the policy that simultaneously increases $g$ and changes taxes so that utility is unchanged

We assume that the change in $g$ and increase in $T$ is defined such that $\frac{\partial U}{\partial g} = 0$
What must the tax adjustment look like to set $\frac{\partial U}{\partial g} = 0$?

i.e. how do we change $T$ in response to the increase in $g$ to hold utility constant for everyone?

For each level of labor earnings, $wl$, define the marginal change in the tax schedule by

$$\frac{\partial T (wl, g)}{\partial g} = \frac{v_g}{v_c}$$

Note that this is the individual’s WTP for $g$ in units of $g$.

We “tax back the benefits”

Notice that if we substitute $\frac{\partial T(wl,g)}{\partial g} = \frac{v_g}{v_c}$ into $\frac{\partial U}{\partial g} = \frac{\partial U}{\partial v} [v_c C_g + v_g]$ we obtain $\frac{\partial U}{\partial g} = 0$ for each type $w$!
Kaplow (2006): Benefit Absorbing Tax

- But, does the benefit-absorbing tax affect labor supply choices, \( l \)?
  - We assumed these were constant...need to verify.
  - This is where weak separability helps
- Define \( v(l) = v(wl - T(wl, g), g) \) to be the level of \( v(c, g) \) experienced by type \( w \) if she chooses \( l \)
- Labor supply \( l \) maximizes
  \[
  l(w) = \arg\max U(v(l), l)
  \]
- Kaplow: Notice that when the policy changes, \( v(l) \) is unaffected by the policy change!
  \[
  \frac{dv}{dg}(l) = \nu_c \frac{\partial T}{\partial g} + \nu_g = 0 \quad \forall l
  \]
- Therefore solution to \( \arg\max U(v(l), l) \) is not affected by the policy change
  - Graphically: Blue arrows for tax adjustment perfectly offset blue arrows from change in \( g \)
  - Exercise: Verify this by solving for \( l(w, g) \) and showing that \( \frac{\partial l}{\partial g} = 0 \) for all \( w \) in this policy change.
What is the optimal level of public expenditure on $g$?

Dual: Maximize government revenue subject to utility held constant

$$\frac{dR}{dg} = \int \frac{dT(wl, g)}{dg} f(w) dw$$

Revenue from Benefit-Tax

But, note that

$$\frac{dT(wl, g)}{dg} = \frac{v_g}{v_c} = \frac{U_v}{U_c} \frac{v_g}{v_c} = \frac{dU}{dg} = \frac{dU}{dc} = s(y)$$ is each type’s willingness to pay ($y = wl$)

Re-writing in notation from last class, optimal to increase $g$ whenever aggregate (unweighted!) surplus is positive

$$\int s(y) \, dy \geq 1$$
What is the role of weak separability? $U(c, g, l) = U(v(c, g), l)$?

Ensures behavioral response to $g$ is similar to behavioral response for tax cut:

$$FE_G = \int s(y) FE(y) \, dy$$

Why might weak separability be violated?

Suppose $g$ is:

- Job training
- Medical care
- Education
- Food stamps
Commodity Taxation

- What about commodity taxes? Or taxes on other goods?
  - Subsidize food vs. expensive cars?
  - Follow Kaplow (2006, JPubEc) for a nice proof
Setup: individuals indexed by $h$

Individuals choose commodities $\{c_1, c_2, \ldots\}$ and labor effort, $l$

Maximize utility function

$$u_h(c_1, c_2, \ldots, l) = \tilde{u}_h(v(c_1, \ldots), l)$$

**Key assumption:** $g$ is the same across people (but $\tilde{u}_h$ can be heterogeneous)

Subject to budget constraint

$$\sum (p_i + \tau_i) c_i \leq wl - T(wl)$$

where $w$ is an individual’s wage (heterogeneous in population)

$wl$ is earnings and $T(wl)$ is the (nonlinear) tax on earnings
Suppose there is a commodity tax
\[ \frac{p_i + \tau_i}{p_j + \tau_j} \neq \frac{p_i}{p_j} \]

for some \( i \) and \( j \)

Can welfare be improved by re-setting \( \tau_i = \tau_j = 0 \) and suitably augmenting the tax schedule \( T \)?
- Atkinson-Stiglitz/Kaplow: YES.

Define \( V(\tau, T, w) \) to be
\[ V(\tau, T, w) = \max_v (c_1, c_2, ...) \]
\[ \text{s.t. } \sum (p_i + \tau_i) c_i \leq w - T(w) \]

\( V \) is the value of the consumption argument of the utility function – holds independent of labor effort \( l \! \! \! \! \).  
- Consumption allocations don’t reveal any information about labor supply type \( w \) conditional on \( w \! \! \! \! \).
Proof

- Define intermediate environment:
  - Start with commodity taxes $\tau$
  - Define new taxes at zero $\tau_i^* = 0$
  - Augment the tax schedule
    - Define $T^*$ to offset the impact on utility so that utility is held constant in this intermediate world
  - Specifically, $T^*$ satisfies
    \[
    V(\tau, T, w) = V(\tau^*, T^*, w)
    \]
    for all $w$.
Lemma 1: Every type \( w \) chooses the same level of labor effort under \( \tau^*, T^* \) as under \( \tau, T \).

Proof:

Note that

\[
U(\tau, T, w, l) = u(V(\tau, T, wl), l) = u(V(\tau^*, T^*, wl), l) = U(\tau^*, T^*, w, l)
\]

The utility function (as a function of \( l \)) is the same in both environments.

Therefore, the \( l \) that maximizes utility in the original world maximizes utility in the intermediate world.
Lemma 2: The augmented world raises more revenue than the original world

Proof:
- Will show that no individual in the intermediate regime can afford the original consumption vector
  - Implies they pay more taxes in intermediate regime
- Suppose type \( w \) can afford original vector when there is no commodity tax, \( \tau_i^* = 0 \).
  - Then she strictly prefers a different vector because of change in relative price
    - Utility level hasn’t changed, but relative prices have
  - But this would imply intermediate environment is strictly better off
    - Choosing a better bundle than the old bundle would strictly increase utility
  - Contradicts definition of intermediate environment holding utilities constant
    - Therefore, type \( w \) cannot afford the original bundle
Proof Cont’d

- Next: If type $w$ cannot afford original bundle, then aggregate tax revenue must be higher in the intermediate environment.
- Because the original bundle is unaffordable, we have:
  \[ \sum (p_i) c_i > wl - T^*(wl) \]
  for all $wl$ (note $\tau^* = 0$)
- Budget constraint in initial regime implies
  \[ \sum_i (p_i + \tau_i) c_i = wl - T(wl) \]
  so that
  \[ \sum p_i c_i = - \sum_i \tau_i c_i + wl - T(wl) \]
- So that
  \[ - \sum_i \tau_i c_i + wl - T(wl) > wl - T^*(wl) \]
- or
  \[ T^*(wl) > \sum \tau_i c_i + T(wl) \]
So, the intermediate world generates more tax revenue and holds utility constant.

Why does this mean one can have a Pareto improvement from no commodity tax?

Generate a third world that gives $\epsilon$ benefits to everyone through lowering the tax schedule.

- Implies everyone better off.
Implications of Atkinson Stiglitz

- Result generally known as the “Atkinson-Stiglitz” theorem
  - Arguably first shown by Hylland and Zeckhauser (1979)
- Incredibly powerful theorem
- Nesst many other results:
  - Zero capital taxes in the standard model
  - “Production efficiency” theorem of Diamond and Mirrlees (1971)
Should we have a tax on capital?
- Capital owners are rich, doesn’t this mean we should tax them if we have redistributive preferences?

Suppose

\[ U(c_1, c_2, \ldots, l) = u(c_1) - v(l_1) + \beta [u(c_2) - v(l_2)] + \ldots \]

With budget constraint

\[ \sum_i (p_i + \tau_i) c_i \leq \sum_i w_i l_i \]

So

\[ g(c_1, c_2, \ldots) = u(c_1) + \beta u(c_2) + \ldots \]

Implies no distortion in relative price of \( c_1 \) and \( c_2 \)
- You should prove extension to case with \( l_i \) instead of just \( l \).
- What if more productive types have higher preferences for bequests?
Production Efficiency

- Should we let firms deduct the price of inputs
  - E.g. firms don’t pay sales tax on their inputs?
- Diamond and Mirrlees (1971) show a surprising result:
- Suppose $C$ is produced with a bunch of intermediate inputs, $x_i$

$$C = f(x_1, ..., x_n)$$

- Question: would you ever want to tax these inputs?
- Answer: No if $C$ is all people care about

$$u(x, l) = U(C(x), l)$$

- The production function for $C$ is the same for all people
  - Weak separability holds
  - Implies no taxes on intermediate inputs
When does weak separability fail?

- When does this fail?
  - Is labor supply an “intermediate input”
    - No taxes on earnings!?
  - What if we can’t tax profits of an intermediate producer?
Relation to Mirrlees

- Another way of seeing this: Mirrlees information logic:
  - When commodity choices have desirable information about type conditional on earnings?
    - See Mirrlees (1976, JPubEc)

- What constitutes “desirable information”? (Saez 2002 JPubEc)
  - Information about social welfare weights: Society likes people that consume $x_1$ more than $x_2$ conditional on earnings
    - Implement subsidy on good $x_1$ financed by tax on $x_2$
    - First order welfare gain (b/c of difference in social welfare weights)
    - Second order distortionary cost starting at $\tau = 0$

- Information about latent productivity: More productive types like $x_1$ more than $x_2$ conditional on earnings
  - e.g. $x_1$ is books; $x_2$ is surf boards
  - Then, tax the goods rich people like but reduce the marginal tax rate
  - Leads to increase in earnings!
  - Depends on covariance
Key Lessons

- In general, need to estimate fiscal externalities associated with policy changes.
- But, if willing to assume weak separability of utility, can just assume that the FE is the same as an income tax.
- Motivates only needing to calculate whether the aggregate surplus is positive.
  - Are people WTP for the policy change out of their own income?
Two empirical literatures on public goods:
- Measuring willingness to pay
- Measuring private crowd-out of government provision
Measuring WTP

Two methods:
- Infer based on behavior / prices
- Ask people (Contingent valuation)
More recently, Keiser and Shapiro (2017): “Consequences of the Clean Water Act and the Demand for Water Quality”

- Cost-benefit analysis of the Clean Water Act

Three analyses

- Estimate water pollution from 1962-2001
- Estimate impact of clean water act grants to wastewater treatment plants on pollution
- Estimate WTP for clean water grants from house prices within 25 mi of plants
Figure 2. Water Pollution Trends, 1962-2001
Panel A. Dissolved Oxygen Deficit

Notes: Graphs show year fixed effects plus a constant from regressions which also control for monitoring site fixed effects, a day-of-year cubic polynomial, and an hour-of-day cubic polynomial, corresponding to equation (2) from the text. Connected dots show yearly values, dashed lines show 95% confidence interval, and 1962 is reference category. Standard errors are clustered by watershed.
Keiser and Shapiro (2017)

\[ Q_{pdy} = \sum_{\tau=-25}^{\tau=25} \gamma_{\tau} \mathbb{1}\left[G_{p,y+\tau} = 1\right]d_d + X'_{pdy} \beta + \eta_{pd} + \eta_{py} + \eta_{dwy} + \epsilon_{pdy} \]

- **Event-study design:**
  - Two observations for each treatment plant: one upstream and one downstream
    - \( G_{p,y+\tau} \) indicator for grant received in year \( y + \tau \), where \( \tau \) indexes years since grant received
    - \( d_d \) is an indicator for being downstream from the treatment facility
    - \( X_{pdy} \) are controls for temperature and precipitation
    - plant-downstream fixed effects, \( \eta_{pd} \) allow for different mean levels up and down-stream
    - plant-year fixed effects, \( \eta_{py} \), control for forces like growth of local industry/etc that affect water quality
    - downstream-by-basin-by-year, \( \eta_{dwy} \), allow upstream and downstream water quality to differ by year in ways common to all plants in a river basin
Figure 3. Effects of Clean Water Act Grants on Water Pollution: Event Study Graphs

Panel A. Dissolved Oxygen Deficit

Panel B. Share Not Fishable

Notes: Graphs show coefficients on downstream times year-since-grant indicators from regressions which correspond to the specification of Table 3. These regressions are described in equation (5) from the main text. Data cover years 1962-2001. Connected dots show yearly values, dashed lines show 95% confidence interval. Standard errors are clustered by watershed.
Figure 4. Effects of Clean Water Act Grants on Log Mean Home Values: Event Study Graphs

Panel A. Homes Within 0.25 Miles of River

Panel B. Homes Within 25 Miles of River

Notes: Graphs show coefficients on year-since-grant indicators from regressions corresponding to the specification of Table 6, column (3). Connected dots show yearly values, dashed lines show 95% confidence interval. Standard errors are clustered by watershed. Panels A and B show different ranges of values on their y-axes. Data cover decennial census years 1970-2000.
Conclusion: Impact on house prices in 25 mile radius is $< 1/3$ of the costs

Concerns?

Distributional incidence?
Value of Clean Air

- Clean Air Act enacted in 1963; 1970 amendment established national ambient air quality standards (NAAQS)
  - Specifies minimum level of air quality for six pollutants
  - Some counties get affected, others are OK
  - Leads to difference in difference design
- **Chay and Greenstone (2003)** look at impacts on infant mortality
- **Isen, Rossin-Slater, and Walker (JPE 2017)** use the impact of the Clean Air Act to generate variation in childhood exposure to pollution and study its impact on adult outcomes
Figure 2: Test of Parallel Trends Assumption, TSP Exposure, Younger Cohort

Isen, Rossin-Slater, and Walker (JPE 2017)
Figure 4: Graphical Summary of IV Results: College Attendance

Effect of Pollution Exposure at Birth on College Attendance

Exposure in Adolescence
- PM2.5, HS

Early Life Exposure
- TSP, In Utero
- TSP, Infancy

Value of Clean Air

- How would you measure the WTP for clean air?
- Brookshire et al. (1982)
  - Infer willingness to pay for clean air using effect of pollution on property prices (capitalization)
- Let $P_i$ denote house price of house $i$, regress
  \[ P_i = \alpha + \beta Pollution_i + \gamma X_i + \epsilon_i \]
  for range of controls, $X_i$.
- Concerns?
- Chay and Greenstone (2005) look at county-level housing prices using non-attainment status as IV
  - “Improvements in air quality induced by the mid-1970s TSPs nonattainment designation are associated with a $45$ billion aggregate increase in housing values in nonattainment counties between 1970 and 1980.”
Optimal Taxation in Ramsey (1927)

- Ramsey (1927): How should commodities be taxed to raise revenue, $R > 0$.
  - Modeled by Diamond and Mirrlees (1971)
- Key result: Tax-weighted Hicksian price derivatives are equated across goods
  - “Inverse elasticity rule”: tax goods with smaller compensated behavioral responses
Setup

• Representative Agent (drop $i$ subscripts).
• Commodities, $x_k$, indexed by $k$
• Government imposes taxes on commodities, $\tau_k$.
• Necessary condition for optimality

$$\frac{d\hat{V}_P}{d\theta}|_{\theta=0} = 0$$

for all feasible policy paths $P$.
• Optimal tax would be lump-sum of size $R$
  • Assumed to not exist
Consider policy $P(\theta)$ that changes commodity taxes (e.g. lowers tax on good 1 and raises tax on good 2)

Budget neutral: $\frac{d\hat{t}}{d\theta} = 0$

No change in public goods

So, optimality condition only involves behavioral response:

$$\sum_k \hat{\tau}_k \frac{d\hat{x}_k}{d\theta} \bigg|_{\theta=0} = 0$$
Hicksian Elasticity

- Diamond and Mirrlees (1971): At the optimum, expand the behavioral response using the Hicksian demands, $x_h^k$,

$$
\frac{dx_k}{d\theta} = \frac{\partial x_h^k}{\partial \tau_1} d\tau_1 + \frac{\partial x_h^k}{\partial \tau_2} d\tau_2
$$

- Additional term, $\frac{\partial x_h^k}{\partial u} dV_p$, but this vanishes at the optimum.

- Optimality condition is given by

$$
\sum_k \tau_k \frac{\partial x_h^k}{\partial \tau_1} d\tau_1 = \sum_k \tau_k \frac{\partial x_h^k}{\partial \tau_2} \left( - \frac{d\tau_2}{d\theta} \right)
$$

- Tax-weighted Hicksian responses are equated across the tax rates
  - Inverse elasticity rule

- What are the needed elasticities?
Assume cross elasticities are zero:

\[ BC = x_1 \frac{d\tau_1}{d\theta} + \tau_1 \frac{dx_1}{d\theta} + x_2 \frac{d\tau_2}{d\theta} + \tau_2 \frac{dx_2}{d\theta} = 0 \]

so

\[ x_1 \left(1 + \frac{\tau_1}{x_1} \frac{\partial x_1^h}{\partial \tau_1}\right) \frac{d\tau_1}{d\theta} = x_2 \left(1 + \frac{\tau_2}{x_2} \frac{\partial x_2^h}{\partial \tau_2}\right) \left(-\frac{d\tau_2}{d\theta}\right) \]

And optimality implies

\[ x_1 \left(\frac{\tau_1}{x_1} \frac{\partial x_1^h}{\partial \tau_1}\right) \frac{d\tau_1}{d\theta} = x_2 \left(\frac{\tau_2}{x_2} \frac{\partial x_2^h}{\partial \tau_2}\right) \left(-\frac{d\tau_2}{d\theta}\right) \]
Inverse Elasticity Rule

- So

\[
\left( \frac{\tau_1 \partial x_1^h}{x_1 \partial \tau_1} \right) = \left( \frac{\tau_2 \partial x_2^h}{x_2 \partial \tau_2} \right) = \kappa
\]

- Translating to price \((1+\tau)\) instead of tax \((\tau)\) elasticities:

\[
\frac{\tau_j}{1 + \tau_j} \epsilon^h_{j,(1+\tau_j)} = \kappa
\]

Or

\[
\frac{\tau_j}{1 + \tau_j} = \frac{\kappa}{\epsilon^h_{j,(1+\tau_j)}}
\]

which is the “inverse elasticity rule”.

Main result of Ramsey model: Inverse elasticity rule

Key Assumptions:
- Representative agent
- No lump sum taxation
Diamond and Mirrlees (1971) also consider the issue of production efficiency.

Commodities, \( x_k \), indexed by \( k \), transformed into one another (produced) by firms and government.

Producer prices \( p_k \), Consumer prices \( q_k \)

- Tax is wedge \( \tau_k = q_k - p_k \)

Consumer \( i \) solves \( \max u_i(\mathbf{x}) \) s.t. \( \sum q_k x_k \leq 0 \)

- Defines consumer (final) demand for each commodity \( x_k^i(\mathbf{q}) \)
- and indirect utility \( V_i(\mathbf{q}) = u(x_i^i(\mathbf{q})) \)

Note: Consumers are the ones endowed with the initial commodity supply.

Endowments allow them to exchange, consumers are on budget constraint.
Price-taking firms $j$ transform commodities

Production possibilities represented by input output function $f^j(y) = 0$

- for example, $y_1 = y_2^3 \cdot y_3^7 \iff y_1 - (-y_2^3) \cdot (-y_3^7) = 0$
- Can turn $y_2$ and $y_3$ into $y_1$ (or vice versa, depending on domain)
- Negative arguments are inputs, positives are outputs
Firm side: CRS Production

- Assumption: constant returns to scale
- Then each firm can produce “as much” or “as little” as desired in fixed proportions
  - Together, many CRS firms define an aggregate production function $f(y) = 0$
  - No profits for any firm (otherwise infinite production) in equilibrium
  - $p \cdot y^j = 0$ must hold in equilibrium, and thus $p \cdot y = p \cdot (\sum y_j) = 0$
- Under CRS, behavior of many optimizing firms same as one aggregate firm
Objective: Choose point on frontier to maximize output prices - input prices

\[ \max p \cdot y \text{ s.t. } f(y) = 0 \]

Optimality condition: \( \frac{\partial f}{\partial y_k} = p_k \iff MRT = \frac{\frac{\partial f}{\partial y_k}}{\frac{\partial f}{\partial y_{k'}}} = \frac{p_k}{p'_{k}} \)

Why can we ignore lagrange multiplier on \( f(y) = 0 \) condition? Because we can normalize the units of \( f \) to be in terms of one of the commodities...see Diamond-Mirrlees (1971).
D&M think of Gov’t as a planner with a distributive objective but:

- Can’t just pick point on PPF
- Must deal with consumers through market place using uniform prices
- Uses:
  - a.) linear commodity taxes to set prices and
  - b.) public production to adjust quantities above and beyond what private sector does given prices

Public production follows PPF given by $g(z) \leq 0$
Objective

What is the objective here?
- redistribution—different than Ramsey, since no revenue requirement

Why would commodity taxes help with no lump sum transfers?
- differential wealth levels are due to endowment differences
- Commodity taxes target:
  - Different tastes
  - Value of endowment

But commodity taxes cause DWL
Objective

- Solve

\[ \max_{q,p,z} \sum_i W(V_i(q)) \text{ s.t. } \sum_i x^i_k(q) = y_k(p) + z_k, \quad f(y) = 0, \text{ and } g(z) = 0 \]

- Lagrangian

\[ \max_{q,p,z} \sum_i W(V_i(q)) + \sum_k \lambda_k (y_k(p) + z_k - \sum_i x^i_k(q)) + \gamma^f f(y(p)) + \gamma^g g(z) \]
Objective

- Production-side and consumer-side variables are additively separable

\[
\max_{q,p,z} \sum_{i} W\left(V_i(q)\right) \quad - \quad \sum_{k} \lambda_k \sum_{i} x_i^k(q) + \sum_{k} \lambda_k (y_k(p) + z_k) + \gamma^f f(y(p)) + \gamma^g g(z)
\]

consumption

production

- Note that FOC for producer prices and government production depend on \( W \) only through the shadow value of an endowment unit of \( k \).

- Also, choice of \( p \) directly implements \( y \), so we can choose \( y \) directly

\[
\max_{q,y,z} \sum_{i} W\left(V_i(q)\right) \quad - \quad \sum_{k} \lambda_k \sum_{i} x_i^k(q) + \sum_{k} \lambda_k (y_k + z_k) + \gamma^f f(y) + \gamma^g g(z)
\]

consumption

production
Result

- $[\text{FOC } y_k] \lambda_k = \gamma^f \frac{\partial f}{\partial y_k}$
- $[\text{FOC } g_k] \lambda_k = \gamma^g \frac{\partial g}{\partial z_k}$

Taking ratio, for any social welfare objective, it must be the case that:

$$\frac{\partial g}{\partial z_k} = \frac{\partial f}{\partial y_k} \frac{\partial g}{\partial z'_k} = \frac{\partial f}{\partial y'_k} = p_k$$

The government’s decision to intervene in the economy is independent of the objective. MRTs are always equalized, and the only wedge is between consumer and producer prices. Production-side and consumer-side variables are additively separable.