Section 7: Optimal Tax Insights

Jon Cohen

November 13, 2021
Recitation Takeaways

- Traditional welfare analysis is about compensated elasticities
Recitation Takeaways

- Traditional welfare analysis is about compensated elasticities
- Optimal income taxation is about “type” incentive compatibility constraints preventing full redistribution
Recitation Takeaways

- Traditional welfare analysis is about compensated elasticities.
- Optimal income taxation is about “type” incentive compatibility constraints preventing full redistribution.
- Optimal (non-Pigouvian) commodity “taxation” is about whether consumption choices have residual info about “type.”
Outline

Harberger-Style DWL Analysis

Optimal Nonlinear Income Taxation

Optimal Commodity Taxation
Old School Analysis

- Contrast with the MVPF framework at each step
- Requires hard to estimate objects (e.g. compensated demands)...
- ...but highlights useful insights using standard micro theory
- Builds straw men for optimal commodity/nonlinear income taxation to dunk on
(Indirect) utility: \( v(p_1, w) - v(p_0, w) \)

Challenge: Utility is ordinal, not cardinal
Back to Basics: Welfare Effect of Price Change

- **(Indirect) utility**: $v(p_1, w) - v(p_0, w)$

- **Challenge**: Utility is ordinal, not cardinal

- **Solution**: Use money-metric utility
  - i.e. expenditure function $e(\bar{p}, v(p, w))$
  - Valid representation of $v(p, w)$ because $e(p, u)$ strictly incr. in $u$
• (Indirect) utility: \( v(p_1, w) - v(p_0, w) \)

• Challenge: Utility is ordinal, not cardinal

• Solution: Use money-metric utility
  - i.e. expenditure function \( e(\tilde{p}, v(p, w)) \)
  - Valid representation of \( v(p, w) \) because \( e(p, u) \) strictly incr. in \( u \)

• New Challenge: \( \tilde{p} = p_1 \) or \( \tilde{p} = p_0 \)?
Compensating Variation

- *Compensate* at new prices
- Define $u_t = v(p_t, w)$ for $t \in \{0, 1\}$

$$CV = e(p_1, u_1) - e(p_1, u_0)$$  \hspace{1cm} (1)

(2)

(3)

$$CV = w - e(p_1, u_0)$$  \hspace{1cm} (4)

where the last line follows Shephard's Lemma

- Analogous EV is the transfer to get equivalent utility at old prices
Compensating Variation

- *Compensate* at new prices
- Define $u_t = v(p_t, w)$ for $t \in \{0, 1\}$

\[ CV = e(p_1, u_1) - e(p_1, u_0) \]  
\[ = w - e(p_1, u_0) \]  
\[ = e(p_0, u_0) - e(p_1, u_0) \]  
\[ = \int_{p_0}^{p_1} h(p, u_0) \, dp \]  

where the last line follows Shephard's Lemma

- Analogous EV is the transfer to get equivalent utility at old prices
Compensating Variation

- *Compensate* at new prices
- Define $u_t = v(p_t, w)$ for $t \in \{0, 1\}$

\[ CV = e(p_1, u_1) - e(p_1, u_0) \]  (1)
\[ = w - e(p_1, u_0) \]  (2)
\[ = e(p_0, u_0) - e(p_1, u_0) \]  (3)
\[ = \int_{p_0}^{p_1} h(p, u_0) \, dp \]  (4)

where the last line follows Shephard's Lemma

- Analogous EV is the transfer to get equivalent utility at old prices
Compensating Variation

- *Compensate* at new prices
- Define $u_t = v(p_t, w)$ for $t \in \{0, 1\}$

\[
CV = e(p_1, u_1) - e(p_1, u_0) \quad (1)
\]
\[
= w - e(p_1, u_0) \quad (2)
\]
\[
= e(p_0, u_0) - e(p_1, u_0) \quad (3)
\]
\[
= \int_{p_1}^{p_0} h(p, u_0) dp \quad (4)
\]

where the last line follows Shephard's Lemma
Compensating Variation

- Compensate at new prices
- Define $u_t = v(p_t, w)$ for $t \in \{0, 1\}$

\[
CV = e(p_1, u_1) - e(p_1, u_0) \tag{1}
\]
\[
= w - e(p_1, u_0) \tag{2}
\]
\[
= e(p_0, u_0) - e(p_1, u_0) \tag{3}
\]
\[
= \int_{p_1}^{p_0} h(p, u_0) dp \tag{4}
\]

where the last line follows Shephard’s Lemma

- Analogous EV is the transfer to get *equivalent* utility at old prices
Visualizing Compensated Demands

Aside: When is uncompensated demand flatter than compensated, as depicted above?
Aside: When is uncompensated demand flatter than compensated, as depicted above?
Visualizing Compensated vs. Uncompensated DWL

- Substitution effects (not income effects) matter for DWL
- Efficiency lost from forgone transactions due to relative price $\Delta p$, not income $\Delta income$
- Most applied papers assume away income effects. When is this more reasonable?
• Substitution effects (not income effects) matter for DWL
• Efficiency lost from forgone transactions due to relative price $\Delta$, not income $\Delta$
• Most applied papers assume away income effects. When is this more reasonable?
Partial Equilibrium DWL: New tax $\tau$ w/ pre-tax prices $p_0$ fixed

\[
\text{Compensated rev.} = \sum_j \tau_j h_j(p_0 + \tau, u_0)
\]

Harberger approximation shown above:

\[
\text{DWL} = \frac{1}{2} \Delta Q \Delta P
\]

\[
\Delta P = \tau, \text{ new tax}
\]

\[
\eta_D \equiv \Delta Q_D \Delta P
\]

\[
\Delta Q_D = \tau \eta_D Q_D
\]

\[
\text{DWL} = \frac{1}{2} \tau^2 \eta_D Q_D
\]

How could we think about the MVPF of this new tax?
Partial Equilibrium DWL: New tax $\tau$ w/ pre-tax prices $p_0$ fixed

- $DWL = CV(p_0 + \tau, p_0, u_0) - R(p_0 + \tau, p_0, u_0)$
- Compensated rev. $= \sum_j \tau_j h_j(p_0 + \tau, u_0)$
Partial Equilibrium DWL: New tax $\tau$ w/ pre-tax prices $p_0$ fixed

- $DWL = CV(p_0 + \tau, p_0, u_0) - R(p_0 + \tau, p_0, u_0)$
- Compensated rev. $= \sum_j \tau_j h_j(p_0 + \tau, u_0)$
- Harberger approximation shown above: $DWL = \frac{1}{2} \Delta Q \Delta P$
Partial Equilibrium DWL: New tax $\tau$ w/ pre-tax prices $p_0$ fixed

- $DWL = CV (p_0 + \tau, p_0, u_0) - R (p_0 + \tau, p_0, u_0)$
  - Compensated rev. $= \sum_j \tau_j h_j (p_0 + \tau, u_0)$
  - Harberger approximation shown above: $DWL = \frac{1}{2} \Delta Q \Delta P$
  - Fixed pre-tax price $\Delta P = \tau$, new tax $\tau = \Delta \tau$, $\eta^D \equiv \frac{\Delta Q^D}{\Delta P} \frac{P}{Q^D} \Rightarrow \Delta Q^D = \tau \eta^D \frac{Q^D}{P}$
Partial Equilibrium DWL: New tax $\tau$ w/ pre-tax prices $p_0$ fixed

\[ DWL = CV(p_0 + \tau, p_0, u_0) - R(p_0 + \tau, p_0, u_0) \]

Compensated rev. = $\sum_j \tau_j h_j(p_0 + \tau, u_0)$

Harberger approximation shown above: \( DWL = \frac{1}{2} \Delta Q \Delta P \)

Fixed pre-tax price $\Delta P = \tau$, new tax $\tau = \Delta \tau$, $\eta^D \equiv \frac{\Delta Q^D}{\Delta P} \frac{P}{Q^D}$ \( \Rightarrow \Delta Q^D = \tau \eta^D \frac{Q^D}{P} \)

\( \Rightarrow DWL = \frac{1}{2} \tau^2 \eta^D \frac{Q^D}{P} \)
Partial Equilibrium DWL: New tax $\tau$ w/ pre-tax prices $p_0$ fixed

$\text{DWL} = CV(p_0 + \tau, p_0, u_0) - \underbrace{R(p_0 + \tau, p_0, u_0)}_{\text{Compensated rev.} = \sum_j \tau_j h_j(p_0+\tau,u_0)}$

Harberger approximation shown above: $\text{DWL} = \frac{1}{2} \Delta Q \Delta P$

Fixed pre-tax price $\Delta P = \tau$, new tax $\tau = \Delta \tau$, $\eta^D \equiv \frac{\Delta Q^D \Delta P}{Q^D P} \Rightarrow \Delta Q^D = \tau \eta^D \frac{Q^D}{P}$

$\Rightarrow \text{DWL} = \frac{1}{2} \tau^2 \eta^D \frac{Q^D}{P}$

How could we think about the MVPF of this new tax?
What if pre-tax prices can adjust?

\[ (p + \tau) = S(p) \]

Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change.

Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities:

\[ \frac{\partial p}{\partial \tau} = \eta_D p + \tau p \eta_S - \eta_D S \]

where \( \eta_D < 0 < \eta_S \).

Need to translate producer incidence into equil. Q response w/ S curve:

Chain rule:

\[ \frac{\partial Q}{\partial \tau} = \frac{\partial Q}{\partial p} \frac{\partial p}{\partial \tau} \]

Substitute into DWL calculation from before:

Recall \( DWL = \frac{1}{2} \Delta P \Delta Q \)

For simplicity suppose there was no pre-existing tax:

\[ DWL = \frac{1}{2} \tau^2 \eta_D \eta_S - \eta_D Q P \]

How does your previous answer change about thinking of the tax change’s MVPF?
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)
- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change

\[
\frac{\partial p}{\partial \tau} = \eta_D p + \tau \eta_S - \eta_D \eta_S Q P
\]
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)
- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[
    \Rightarrow \frac{\partial p}{\partial \tau} = \frac{\eta_D}{p+\tau} \frac{\eta_D}{\eta_S - \eta_D} \quad \text{where} \quad \eta_D < 0 < \eta_S
    \]
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)

- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[ \Rightarrow \frac{\partial p}{\partial \tau} = \frac{\eta_D}{p + \tau} \frac{\eta_D}{\eta_S - \eta_D} \]
    where \( \eta_D < 0 < \eta_S \)

- Need to translate producer incidence into equil. Q response w/ S curve
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)
- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[ \frac{\partial p}{\partial \tau} = \frac{\eta_D}{\frac{p+\tau}{p} \eta_S - \eta_D} \]
    \[ \text{where } \eta_D < 0 < \eta_S \]
- Need to translate producer incidence into equil. Q response w/ S curve
  - Chain rule: \( \frac{\partial Q^S}{\partial \tau} = \frac{\partial Q^S}{\partial p} \frac{\partial p}{\partial \tau} \)
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)

- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[ \Rightarrow \frac{\partial p}{\partial \tau} = \frac{\eta_D}{p+\tau} \frac{p}{\eta_S-\eta_D} \] where \( \eta_D < 0 < \eta_S \)

- Need to translate producer incidence into equil. Q response w/ S curve
  - Chain rule: \( \frac{\partial Q^S}{\partial \tau} = \frac{\partial Q^S}{\partial p} \frac{\partial p}{\partial \tau} \)

- Substitute into DWL calculation from before
  - Recall \( DWL = \frac{1}{2} \Delta P \Delta Q \)
  - For simplicity suppose there was no pre-existing tax
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)
- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[ \Rightarrow \frac{\partial p}{\partial \tau} = \frac{\eta_D}{p+\tau} \frac{p}{\eta_S-\eta_D} \]
    where \( \eta_D < 0 < \eta_S \)
- Need to translate producer incidence into equil. Q response w/ S curve
  - Chain rule: \( \frac{\partial Q^S}{\partial \tau} = \frac{\partial Q^S}{\partial p} \frac{\partial p}{\partial \tau} \)
- Substitute into \( DWL \) calculation from before
  - Recall \( DWL = \frac{1}{2} \Delta P \Delta Q \)
  - For simplicity suppose there was no pre-existing tax
    \[ \Rightarrow DWL = \frac{1}{2} \tau^2 \frac{\eta_D \eta_S}{\eta_S-\eta_D} \frac{Q}{P} \]
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)
- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[
    \Rightarrow \frac{\partial p}{\partial \tau} = \frac{\eta_D}{p+\tau} \frac{p}{\eta_S-\eta_D} \quad \text{where } \eta_D < 0 < \eta_S
    \]
- Need to translate producer incidence into equil. Q response w/ S curve
  - Chain rule: \( \frac{\partial Q^S}{\partial \tau} = \frac{\partial Q^S}{\partial p} \frac{\partial p}{\partial \tau} \)
- Substitute into \( DWL \) calculation from before
  - Recall \( DWL = \frac{1}{2} \Delta P \Delta Q \)
  - For simplicity suppose there was no pre-existing tax
    \[
    \Rightarrow DWL = \frac{1}{2} \tau^2 \frac{\eta_D \eta_S}{\eta_S-\eta_D} \frac{Q}{P}
    \]
What if pre-tax prices can adjust?

- \( D(p + \tau) = S(p) \)
- Need to calculate producer incidence \( \frac{\partial p}{\partial \tau} \) for price change
  - Differentiate equality to get \( \frac{\partial p}{\partial \tau} \) and rewrite in terms of elasticities
    \[
    \Rightarrow \frac{\partial p}{\partial \tau} = \frac{\eta_D}{p+\tau} \eta_S - \eta_D
    \]
    where \( \eta_D < 0 < \eta_S \)
- Need to translate producer incidence into equil. Q response w/ S curve
  - Chain rule: \( \frac{\partial Q^S}{\partial \tau} = \frac{\partial Q^S}{\partial p} \frac{\partial p}{\partial \tau} \)
- Substitute into \( DWL \) calculation from before
  - Recall \( DWL = \frac{1}{2} \Delta P \Delta Q \)
  - For simplicity suppose there was no pre-existing tax
    \[
    \Rightarrow DWL = \frac{1}{2} \tau^2 \frac{\eta_D \eta_S}{\eta_S - \eta_D} \frac{Q}{P}
    \]
How does your previous answer change about thinking of the tax change’s MVPF?
What if there’s already an existing tax?
(Even More) General DWL

What if there’s already an existing tax?

- Recall $\text{DW}L(\tau) = [e(p + \tau, u) - e(p, u)] - \tau h(p + \tau, u)$

  - CV
  - Tax Revenue

Note $\frac{\partial \text{DW}L}{\partial \tau} = 0$ by env. thm.

$\frac{\partial}{\partial \tau} - \tau \frac{\partial h}{\partial \tau} = -\tau \frac{\partial h}{\partial \tau}$ and $\frac{\partial^2 \text{DW}L}{\partial \tau^2} = -\frac{\partial h}{\partial \tau} - \tau \frac{\partial^2 h}{\partial \tau^2}$

$\Rightarrow \text{MDWL}(\tau) = -\tau \Delta \tau \frac{\partial h}{\partial \tau}$

New 1st order distortion $-\frac{1}{2} \tau (\Delta \tau)^2 \frac{\partial h}{\partial \tau}$

Standard Harberger triangle
(Even More) General DWL

What if there's already an existing tax?

- Recall $DWL(\tau) = \left[ e(p + \tau, u) - e(p, u) \right] - \tau h(p + \tau, u)$
  - CV
  - Tax Revenue

- Take 2\textsuperscript{nd} order Taylor expansion around $DWL(\tau)$ to get marginal $DWL$:

$$MDWL(\tau) = \frac{\partial DWL(\tau)}{\partial \tau} \Delta \tau + \frac{1}{2} \frac{\partial^2 DWL(\tau)}{\partial \tau^2} \tau^2$$

Note $\frac{\partial DWL(\tau)}{\partial \tau} = h - h \overset{\text{env. thm.}}{=} 0$ by env. thm.

$-\tau \frac{\partial h}{\partial \tau} = -\tau \frac{\partial h}{\partial \tau}$ and $\frac{\partial^2 DWL(\tau)}{\partial \tau^2} = -\frac{\partial h}{\partial \tau} - \tau \frac{\partial^2 h}{\partial \tau^2}$

Assume $= 0$ for simplicity

⇒ $MDWL(\tau) = -\tau \Delta \tau \frac{\partial h}{\partial \tau} \overset{\text{New 1\textsuperscript{st} order distortion}}{=} -\frac{1}{2} \tau (\Delta \tau)^2 \frac{\partial h}{\partial \tau}$

Standard Harberger triangle
What if there’s already an existing tax?

- Recall $DWL(\tau) = \left[ e(p + \tau, u) - e(p, u) \right] - \tau h(p + \tau, u)$
  - CV
  - Tax Revenue

- Take $2^{nd}$ order Taylor expansion around $DWL(\tau)$ to get marginal $DWL$:

\[
MDWL(\tau) = \frac{\partial DWL(\tau)}{\partial \tau} \Delta \tau + \frac{1}{2} \frac{\partial^2 DWL(\tau)}{\partial \tau^2} \tau^2
\]

Note $\frac{\partial DWL(\tau)}{\partial \tau} = h - h - \tau \frac{\partial h}{\partial \tau} = -\tau \frac{\partial h}{\partial \tau}$ and $\frac{\partial^2 DWL(\tau)}{\partial \tau^2} = -\frac{\partial h}{\partial \tau} - \tau \frac{\partial^2 h}{\partial \tau^2}$

=0 by env. thm.

assume =0 for simplicity
(Even More) General DWL

What if there’s already an existing tax?

- Recall $DWL(\tau) = [e(p + \tau, u) - e(p, u)] - \tau h(p + \tau, u)$

- Take $2^{nd}$ order Taylor expansion around $DWL(\tau)$ to get marginal $DWL$:

$$MDWL(\tau) = \frac{\partial DWL(\tau)}{\partial \tau} \Delta \tau + \frac{1}{2} \frac{\partial^2 DWL(\tau)}{\partial \tau^2} \tau^2$$

Note $\frac{\partial DWL(\tau)}{\partial \tau} = h - h - \tau \frac{\partial h}{\partial \tau} = -\tau \frac{\partial h}{\partial \tau}$ and $\frac{\partial^2 DWL(\tau)}{\partial \tau^2} = -\frac{\partial h}{\partial \tau} - \tau \frac{\partial^2 h}{\partial \tau^2}$

$= 0$ by env. thm.

$\Rightarrow MDWL(\tau) = -\tau \Delta \tau \frac{\partial h}{\partial \tau} - \frac{1}{2} \tau (\Delta \tau)^2 \frac{\partial h}{\partial \tau}$

New $1^{st}$ order distortion

Standard Harberger triangle
Visualizing the Harberger Triangle (Courtesy of Chetty)

Excess Burden of a Tax Increase: Harberger Trapezoid

- Lost cons. surplus (2nd order)
- Lost govt. revenue (1st order)
- Lost producer surplus (2nd order)
How does this interact with other markets?
How does this interact with other markets?

- Analogous to the analysis when there are existing taxes
- Behavioral response has 1\textsuperscript{st} order welfare effect if there’s a pre-existing distortion
How does this interact with other markets?

- Analogous to the analysis when there are existing taxes
- Behavioral response has 1\textsuperscript{st} order welfare effect if there's a pre-existing distortion
- Tax $\tau_i$ in market $i$

\[
\Rightarrow DWL = \frac{1}{2} \tau_i dh_i + \sum_{j \neq i} \tau_i \tau_j \frac{dh_j}{d\tau_i}
\]
How does this interact with other markets?

- Analogous to the analysis when there are existing taxes
- Behavioral response has 1st order welfare effect if there’s a pre-existing distortion
- Tax $\tau_i$ in market $i$

\[
\Rightarrow DWL = \frac{1}{2} \tau_i dh_i + \sum_{j \neq i} \tau_i \tau_j \frac{dh_j}{d\tau_i}
\]

- Atkinson-Stiglitz Theorem says uniform commodity taxation is optimal under certain conditions
- The above formula foreshadows a “Law of the 2nd Best” application that you might want to subsidize/tax goods that have spillover effects on already distorted markets
Outline

Harberger-Style DWL Analysis
Optimal Nonlinear Income Taxation
Optimal Commodity Taxation
The 16th Amendment was Ratified in 1913...Get With the Times!

- (Income taxation)
- Potentially more natural to redistribute using this
- Focus on two type case (Stiglitz 1982) to highlight intuition
  1. Incentive compatibility
  2. Impossibility of Laffer effects at optimum

- **Worker preferences:** $U^i(c, Y) = U(c, Y; \theta^i)$
  - consumption $c$, heterogeneous productivity $\theta^i$, labor $l = \frac{Y}{\theta^i}$
  - $\theta^H > \theta^L$

- **Aggregate resource constraint:** $\sum_i c(\theta^i) \leq \sum_i Y(\theta^i)$

- **Work budget constraint:** $B = \{(c, Y) | c \leq Y - T(Y)\}$
  - Potentially nonlinear income tax schedule $T(Y)$

- **Objective:** Redistribution across types $\theta^i$

- **Challenge:** $\theta^i$ unobserved
What If Worker Type IS Observed?

- With perfect observability, can use type-specific policy (i.e. lump-sum transfers)
- Only constraint is the aggregate resource constraint

But alas...
• **Key idea:** Unobservability of type $\Rightarrow$ any type-specific policy will have to get each type to *willingly* reveal themselves
  - Lets you consider allocations as function of $\theta$
  - Related to *revelation principle* from mechanism design

• **Incentive compatibility (IC) constraint:** $u(c(\theta), Y(\theta); \theta) \geq u(c(\tilde{\theta}), Y(\tilde{\theta}); \theta) \ \forall \theta, \tilde{\theta}$

• **Worker FOC:**
  $$MRS(c, Y; \theta) \equiv -\frac{U_Y(c, Y; \theta)}{U_c(c, Y; \theta)} = 1 - T'(Y)$$
Visualizing $MRS$

- Draw indifference curves for each type
  - $Y$ implicitly defines labor
  - $c, Y$ together implicitly define $T(Y)$ and $T'(Y)$

Model can be generalized, but key is $MRS(c, Y; \theta)$ decreasing in $\theta$ (single crossing)
Non-Type Specific Policy Cannot Be Along Pareto Frontier

- Pooling equilibrium represented by both allocation at same point along $c = Y$
- **Key idea**: Check feasible Pareto improvements along $c = Y$

Case #1: No Pooling w/ $MRS_H < MRS_L < 1$

Case #2: No Pooling w/ $MRS_H < 1 < MRS_L$

Case #3: No Pooling w/ $1 < MRS_H < MRS_L$

Above are graphs showing efficiency $\Rightarrow MRS(c, Y; \theta) = 1 - T'(Y)$ but $\theta^H > \theta^L$

*Counterintuitive corollary: $T'(Y(\theta^H)) > 0$ is not optimal because of Laffer effects*
So What Can Be Pareto Optimal?

Desired allocation #1

\[ \bar{U}_1^L \]

\[ \bar{U}_1^H \]
...Good to Go if IC’s are Satisfied

\[ T(Y(\theta^L)) < 0 \]

\[ T(Y(\theta^H)) > 0 \]

\[ \text{slope} = 1 - T'(Y(\theta^L)) < 1 \]

\[ \text{slope} = 1 - T'(Y(\theta^H)) = 1 \]
Can Anything Be Pareto Optimal?

- Desired allocation #2

Diagram showing a curve with axes $U^L$ and $U^H$, and points $\bar{U}_2^L$ and $\bar{U}_2^H$. The text describes the diagram in the context of Pareto optimality.
...Not If IC's Can’t be Satisfied

\[ T'(Y(\theta^H)) = 0 \]

candidate cannot be IC
Adjustment Necessary for Feasibility (with Unobservability)

- Previous slide demonstrates ICH (incentive compatibility constraint for $\theta^H$ binds)
- **2nd best allocation**: Maximal $\bar{U}_2^L$ along $\bar{U}_2^H$ s.t. resource constraint holds.
IC Constraint Prevents Reaching Pareto Frontier

- Desired allocation #2
- Feasible allocation #2
More General 2\textsuperscript{nd} Best Frontier (Due to Unobservability)
Super Aside: Alternative Variational Approach to Optimal Income Taxation

- Consider marginal increase in $T'(Y_0)$ at given $Y_0$

  1. **Mechanical Effect**: Increase $T(Y) \forall Y \geq Y_0$
     - Avg welfare weight $\bar{\lambda} \equiv E[\lambda(Y)|Y \geq \bar{Y}]$
     - Raising $1$ revenue good but comes at welfare cost of $\bar{\lambda}$

  2. **Behavioral Effect**: People previously with $Y \geq Y_0$ adjust earnings down
     - Fiscal externality with no 1\textsuperscript{st} order private welfare effect (*envelope theorem!*)
     - Sum of income and substitution effects (*Inverse elasticity rule!*)

**Optimum equates effects** $\Rightarrow (1 - \lambda(Y_0))M + B = 0$
Super Aside: Alternative Variational Approach to Optimal Income Taxation

- Consider marginal increase in $T'(Y_0)$ at given $Y_0$

1. **Mechanical Effect**: Increase $T(Y)$ $\forall Y \geq Y_0$
   - Avg welfare weight $\bar{\lambda} \equiv E[\lambda(Y) | Y \geq \bar{Y}]$
   - Raising $1$ revenue good but comes at welfare cost of $\bar{\lambda}$

2. **Behavioral Effect**: People previously with $Y \geq Y_0$ adjust earnings down
   - Fiscal externality with no 1st order private welfare effect (*envelope theorem!*)
   - Sum of income and substitution effects (*Inverse elasticity rule!*)

**Optimum equates effects** $\Rightarrow (1 - \lambda(Y_0))M + B = 0$

How can we see the MVPF in the above formula?
Takeaways from Mechanism Design Approach

• Desire to redistribute to one type is constrained by an incentive compatibility condition on the other
• It would be really great if we could relax that...
Outline

Harberger-Style DWL Analysis

Optimal Nonlinear Income Taxation

Optimal Commodity Taxation
What should commodity taxes be?

• You have nonlinear income taxes. Should you *differentially* tax commodities, too?
  ■ If it relaxes information constraints, then yes!
  ■ If it doesn’t, then you can generate a Pareto improvement by removing the commodity tax distortion and compensating through the income tax.
What should commodity taxes be?

- You have nonlinear income taxes. Should you *differentially* tax commodities, too?
  - If it relaxes information constraints, then yes!
  - If it doesn’t, then you can generate a Pareto improvement by removing the commodity tax distortion and compensating through the income tax

- What is the formal condition determining whether consumption choices reveal information?
  - **Preference restriction**: Weak separability of all consumption choices with respect to labor supply
  - **Key PF result**: Atkinson-Stiglitz
What should commodity taxes be?

- You have nonlinear income taxes. Should you *differentially* tax commodities, too?
  - If it relaxes information constraints, then yes!
  - If it doesn’t, then you can generate a Pareto improvement by removing the commodity tax distortion and compensating through the income tax

- What is the formal condition determining whether consumption choices reveal information?
  - **Preference restriction**: Weak separability of all consumption choices with respect to labor supply
  - **Key PF result**: Atkinson-Stiglitz

- If there’s a revenue requirement, is the *DWL* analysis from before relevant?
  - Uniform taxation on *all goods* is like a lump-sum tax!
Weak Separability

- **(Semi-)formal definition:** \( u(x_1, x_2, ..., x_n) = u(x_1, G(x_2, ..., x_n)) \Rightarrow u(\cdot) \) weakly separable between \( x_1 \) and \((x_2, ..., x_n)\)

- **Intuition:** Two-stage budgeting (weak separability is necessary and sufficient for this)
  - First decide upper-level (i.e. $ towards \( x_1 \) vs. $ towards \((x_1, ..., x_n)\))
  - Next decide lower-level (i.e. $ towards \( x_2 \) vs. \( x_3 \) vs. \( x_4 \) and so on)
  - \( MRS \) between goods in subutility \( G(\cdot) \) unaffected by level of \( x_1 \)
Goldman and Uzawa (1964) derived that weak separability ⇒ Slutky substitution terms ∝ income effects

Afriat (1970) and Varian (1983) develop non-parametric tests

- Very limited intuition: Similar to GARP tests about whether choice data can be rationalized with certain preferences
Weak Separability as Required by Atkinson-Stiglitz

- Consider $u^i(x_1, ..., x_n, l)$ over goods $(x_1, ..., x_n)$ and labor $l$
- Require $u^i(x_1, ..., x_n, l) = u^i(G(x_1, ..., x_n), l)$
  - Allow heterogeneity in consumption vs. leisure decisions, but not in $MRS$'s between common $G(\cdot)$ with respect to labor
Proof Sketch of Uniform Commodity Taxation Under WS

By contradiction (see Kaplow 2006 for details)

1. Remove differential commodity taxes and adjust (arbitrary) nonlinear income tax so that indirect utility is constant for everyone
2. By weak separability, labor is also constant for everyone
3. Show old consumption bundle now isn’t affordable
4. Therefore government gain revenue leaving everyone indifferent
5. PROFIT!
• Weak separability implies, conditional on income, relative consumption decisions reveal no information about type
• Therefore *differential* taxation doesn’t relax IC constraints but does introduce distortions
How Atkinson-Stiglitz Fails

• Just because a benchmark is useful doesn’t mean it’s always true

• Intuitive violations of weak separability (Saez 2002):
  1. Conditional on income, owning a yacht reveals info about hidden assets (i.e. income $Y$) or that you’re an insufferable person (i.e. welfare weight $\lambda$)
     $\Rightarrow$ tax it!
  2. Child care is a complement to labor and thus, conditional on income, reveals info on unobserved productivity $\theta$
     $\Rightarrow$ subsidize it!
     - Intuition related to multi-market DWL from slide 12 (Corlett and Hague 1953)

• More generally, any induced relaxation of information constraints is efficient

• How is this related to the MVPF?
Fitting in-kind provision into each MVPF term

\[ MVPF = \bar{\eta} \frac{WTP}{1 + FE} \]
Relation to Mysterious Nichols and Zeckhauser Figure

- $A$ is “advantaged” and $B$ is “broke”
- Plot indifference curve of both $A$ and $B$ with respect to consumption of indicator good
  - Use residual income to consume “everything else”
- Different slopes at same level of indicator good (i.e. $MRS$ heterogeneity) is a violation of weak separability
  - Consumption choice reveals info on type
  - Distorting consumption relaxes info constraint
  - Redistribution with the indicator good can improve on redistribution through cash tax and transfers alone
Nichols and Zeckhauser Figure!