# Econ 230A: Public Economics Lecture: Deadweight Loss & Optimal Commodity Taxation <sup>1</sup>

Hilary Hoynes

UC Davis, Winter 2012

<sup>1</sup>These lecture notes are partially based on lectures developed by Raj Chetty and Day Manoli. Many thanks to them for their generosity.

#### Outline

- Deadweight Loss
- What is deadweight loss?
- Ø Marshallian Surplus & the Harberger Formula
- General Model with income effects
- Empirical Applications
  - sufficient statistics: structural vs. reduced form approaches
  - Marion and Muehlegger
  - Optimal Commodity Taxation
- What is the problem?
- ② Ramsey Tax Problem (Representative Agent)
- Production Efficiency

( )

### 1. What is deadweight loss?

- Thus far, we have focused on the incidence of government policies: how price interventions affect equilibrium prices and factors returns.
  - that is, we determined how policies affect the distribution of the pie.
- A second general set of questions is how taxes affect the size of the pie.
- Example: income taxation
  - Government raises taxes:
    - \* to raise revenue to finance public goods (roads, defense ...)
    - ★ to redistribute income from rich to poor.
  - But raising tax revenue generally has an efficiency cost: to generate \$1 of revenue, need to reduce welfare of the taxed individuals by more than \$1
  - Efficiency costs come from distortion of behavior.

- 3

イロト 人間ト イヨト イヨト

#### 1. What is deadweight loss? (cont)

- Large set of studies on how to implement policies that minimize efficiency costs (optimal taxation). This is the core theory of public finance, which is then adapted to the study of transfer programs, social insurance, etc.
- We begin with positive analysis of how to measure efficiency cost ("excess burden" or "deadweight cost") of a given tax system.
  - Computing EB gives you the cost of taxation (often referred to as the marginal cost of public funds).
  - We will see that this number is not uniquely defined
- Note: EB does not tell you anything about the <u>benefit</u> of taxation (redistribution, raise money for public goods,...).
  - Ultimately we will weigh DWL and the benefits of what is done with taxes raised.

# 2. Marshallian Surplus & the Harberger Formula (triangle)

- Start with simplest case
- Two good model with representative consumer & firm
  - x = taxed good, y =(untaxed numeraire), p =producer (before tax) price of x, t =tax on x, Z =income
- Key assumptions: quasilinear utility (<u>no income effects</u>), competitive production.
- No income effect means marshallian can give us the welfare effects; can examine in simple S/D setting
- Consumer solves

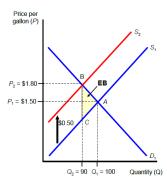
$$\max_{x,y} u(x) + y \text{ s.t. } (p+\tau)x(p+\tau,Z) + y(p+\tau,Z) = Z$$

イロト 不得下 イヨト イヨト 二日

- Price-taking firms use c(S) units of the numeraire y to produce S units of x
  - c'(S) > 0 and  $c''(S) \ge 0$
  - firm maximizes profit pS c(S)
  - ► supply function for good x is implicitly defined by the marginal condition (MR=MC) p = c'(S(p)).
- Equilibrium:  $Q(p) = D(p + \tau)$

イロト 不得下 イヨト イヨト 二日

- Consider Introduction of a small tax:  $d\tau > 0$ . See figure (Gruber)
  - On the graph we can see cons surplus (area under demand above price), producer surplus (revenue - area under supply), tax revenue, and DWL.
  - DWL ("deadweight loss" or "excess burden") is what is lost on top of what is collected in taxes. This is the small triangle in the picture.



- There are 3 ways of measuring the area of the triangle:
  - In terms of supply and demand elasticities:
  - In terms of total change in equilibrium quantity caused by tax.
  - In terms of change in government revenue (this will be a first-order approximation)

(日) (周) (三) (三)

- 2. Marshallian Surplus & the Harberger Formula
  - <u>Method 1</u>: Measuring EB in terms of supply and demand elasticities:

$$EB = (\frac{1}{2})dQd\tau$$

$$EB = (\frac{1}{2})S'(p)dpd\tau = (\frac{1}{2})(\frac{pS'}{S})(\frac{S}{p})(\frac{\eta_D}{\eta_S - \eta_D})d\tau^2$$

$$EB = (\frac{1}{2})(\frac{\eta_S\eta_D}{\eta_S - \eta_D})(pQ)(\frac{d\tau}{p})^2$$

- 2nd line uses incidence formula  $dp = (\frac{\eta_D}{\eta_S \eta_D}) d\tau$
- 3rd line uses definition of  $\eta_S$  .
- Third line shows common intuition that EB increases with the square of the tax and with elasticities of S and D.

- <u>Method 1</u>: Measuring EB in terms of supply and demand elasticities (cont)
- Tax revenue R = Qdτ, so useful expression is deadweight burden per dollar of tax revenue:

$$\frac{EB}{R} = \frac{1}{2} \frac{\eta_S \eta_D}{\eta_S - \eta_D} \frac{d\tau}{p}$$

(日) (周) (三) (三)

- <u>Method 2</u>: Measuring EB in terms of total change in equilibrium quantity caused by tax:
- Define  $\eta_Q = -\frac{dQ}{d\tau}\frac{p}{Q}$  as the effect of a 1% increase in the initial price via a tax change on equilibrium quantity (elas version of incidence formula)
- Then defining EB using change in quantity and change in price:

$$EB = -(\frac{1}{2})dQd\tau$$
$$= -(\frac{1}{2})\frac{dQ}{d\tau}(\frac{p}{Q})(\frac{Q}{p})d\tau d\tau$$
$$= (\frac{1}{2})\eta_Q(pQ)(\frac{d\tau}{p})^2$$

 Again, the EB is a function of the square of the tax and the sensitivity to price changes  $\eta_Q$ 

イロト イポト イヨト イヨト

- <u>Method 3</u>: Measuring EB in terms of change in government revenue
- This is a first-order approximation -> use to calculate marginal DWL given pre-existing taxes.
- Start with a tax of  $\tau$  per unit. Then from method 1 (replacing  $d\tau$  by  $\tau$ ) we have:

$$DWL(\tau/p) = (\frac{1}{2})\frac{\eta_{S}\eta_{D}}{\eta_{S} - \eta_{D}}(pQ)(\frac{\tau}{p})^{2}$$

• Marginal DWB (first-order approximation) is:

$$\frac{\partial DWL}{\partial (\tau/p)} = \frac{\eta_S \eta_D}{\eta_S - \eta_D} (pQ) \frac{\tau}{p} = \eta_Q Q \tau$$

• Uses incidence formula for impact of tax on equil Q.

- A TE N - A TE N

- Measuring EB in terms of change in government revenue (continued)
- Alternative representation of <u>∂DWB</u>/∂(τ/ρ): use data on government budget: DWL equals the difference between the "mechanical" revenue gain (no change in price) and the actual revenue gain.
- Note: This is theoretically interesting, but in practice the difference between mechanical and actual could be due to lots of factors changing in the economy. Not empirically feasible.

イロト イポト イヨト イヨト 二日

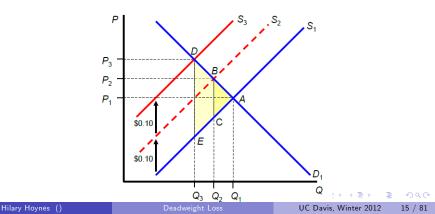
- Measuring EB in terms of change in government revenue (continued)
- Note that  $\frac{\partial DWL}{\partial \tau} = \frac{\tau}{p} \eta_Q Q$  is a first-order approximation to MDWL.
- It includes loss in govt revenue due to behavioral response (the rectangle in the Harberger trapezoid, proportional to τ), but not the second-order term (proportional to τ<sup>2</sup>).
- Second-order approximation includes triangles at the end of the Harberger trapezoid :

$$EB = x(\tau)\eta_Q\tau(\Delta\tau) + \frac{1}{2}x(\tau)\eta_Q(\Delta\tau)^2.$$

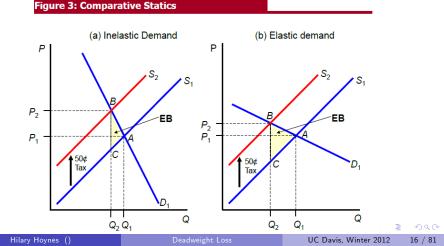
イロト 不得下 イヨト イヨト 二日

• Key Result 1: Deadweight burden is increasing at the rate of the square of the tax rate and deadweight burden over tax revenue increases linearly with the tax rate. See figure (Gruber).

Figure 2: EB Increases with Square of Tax Rate



• Key Result 2: Deadweight burden Deadweight burden increases with the absolute value of the elasticities (note that if either elasticity is zero, there is no DWB). See figure (Gruber).



- Important consequence: With many goods the most efficient (efficient = keeping DWB as low as possible) way to raise tax revenue is
  - tax relatively more the inelastic goods. E.g. medical drugs, food. But what's the tradeoff?
  - spread the taxes across all goods so as to keep tax rates relatively low on all goods (because DWB increases with the square of the tax rate)

イロト 不得下 イヨト イヨト

## 3. General Model

- Drop quasilinearity assumption and consider an individual with utility  $u(c_1, ..., c_N) = u(c)$
- Individual program:  $\max_{c} u(c)$  s.t.  $q \cdot c \leq Z$ 
  - ▶ where q = p + t denotes vector of tax-inclusive prices and Z is wealth (can be zero).
- Multiplier of the budget constraint is  $\lambda$
- FOC in  $c_i$ :  $u_{c_i} = \lambda q_i$
- FOCs + budget constraint determine Marshallian (or uncompensated) demand functions  $c_i(q, Z)$  and an indirect utility function v(q, Z).
  - useful property is Roy's identity:  $v_{q_i} = -\lambda c_i$ : welfare effect of a price change  $dq_i$  is the same as taking  $dZ = c_i dq_i$  from the consumer
  - adjustment of c<sub>j</sub> do not produce a first order welfare effect because of the envelope theorem

3. General Model: Income Effects & Path Dependence Problem (Auerbach 1985)

- Start from a price vector  $q^0$  and move to a price vector  $q^1$  (by levying taxes on goods).
- Marshallian surplus is defined as:

$$CS = \int_{q^0}^{q^1} c(q, Z) dq$$

- Problem with this definition?
  - the consumer surplus is path dependent when more than one price changes: q<sup>0</sup> to q' and q' to q<sup>1</sup>.

$$CS(q^0 \rightarrow q') + CS(q' \rightarrow q^1) \neq CS(q^0 \rightarrow q^1)$$

 In other words, it matters the order that you vary the taxes (unappealling property)

- A TE N - A TE N

# 3. General Model: Income Effects & Path Dependence Problem

• Example with taxes on two goods: CS defined in two ways

$$CS = \int_{q_1^0}^{q_1^1} c_1(q_1, q_2^0, Z) dq_1 + \int_{q_2^0}^{q_2^1} c_2(q_1^1, q_2, Z) dq_2.$$
  
or  $CS = \int_{q_2^0}^{q_2^1} c_2(q_1^0, q_2, Z) dq_2 + \int_{q_1^0}^{q_1^1} c_1(q_1, q_2^1, Z) dq_1.$ 

- Mathematical problem: for these to be equivalent (path-independent), need cross-partials to be equal, i.e.  $\frac{dc_2}{da_1} = \frac{dc_1}{da_2}$ .
- This will not be satisfied for Marshallian demand functions unless there are no income effects, b/c income effects and initial consumption levels differ across goods
- But they are equal for Hicksian (compensated) demand [Slutsky is symmetric]

# 3. General Model: Income Effects & Path Dependence Problem

- Bottom line:
  - Marshallian EB is appealing, since it is easy. But unappealing because of path dependence.
  - Hicksian EB is appealing because there is no path dependence. But unappealing because it is not observable and depends on utility measure utility h(q, u).
  - What utility to measure Hicksian EB at? Two natural candidates (pre tax utility, post tax utility). This gets us to compensating variation and equivalent variation measures.

イロト イポト イヨト イヨト 二日

### 3. EV and CV Meaures - Definitions

- To translate the utility loss into dollars, introduce the expenditure function.
- Fix utility and prices, and look for the bundle that minimizes cost to reach that utility for these prices:

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

• Let  $\mu$  denote multiplier on utility constraint, then the FOCs given by

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

 FOCs & constraint generate Hicksian (or compensated) demand functions h which map prices and utility into demand

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

• Now define the loss to the consumer from increasing tax rates as

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

### 3. EV and CV Meaures - Definitions

- e(q<sup>1</sup>, u) e(q<sup>0</sup>, u) is a single-valued function and hence is a coherent measure of the welfare cost of a tax change to consumers. So no path dependence problem.
- But now, which *u* should we use? Consider change of prices  $q^0$  to  $q^1$  and assume that individual has income *Z*.

• 
$$u^0 = v(q^0, Z)$$
 (initial utility)

- $u^1 = v(q^1, Z)$  (utility at new price  $q^1$ ).
- Using these, we define EV and CV (next page)

イロト 不得 トイヨト イヨト 二日

### 3. EV and CV Meaures - Definitions

• Compensating variation:

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

- How much you need to compensate the consumer for him to be indifferent between having the tax and not having the tax (to reach original utility level at new prices).
- ► Logic: e(q<sup>0</sup>, u<sup>0</sup>) = e(q<sup>1</sup>, u<sup>0</sup>) CV where CV is amount of your ex-post expenses I have to cover to leave you with same ex-ante utility.
- Equivalent variation:  $EV = e(q^1, u^1) e(q^0, u^1) = Z e(q^0, u^1)$ :
  - How much money would the consumer be willing to pay as a lump sum to avoid having the tax (and reach new post-tax utility level at original price).
  - ▶ Logic: e(q<sup>0</sup>, u<sup>1</sup>) + EV = e(q<sup>1</sup>, u<sup>1</sup>) where EV is amount extra I can take from you and leave you with same ex-post utility.
- We use these to define the Excess Burden
   → EB is the excess of EV (CV) over revenue collected.

# 3. General Model: Harberger Formula with Income Effects (following Auerbach 1985)

- $\bullet$  How to see EB using CV/EV graphically
- Start with Hicksian (compensated) demand functions *h*. First note that envelope theorem implies

$$e_{q_i}(q,u)=h_i$$

• Hence can define CV or EV as:

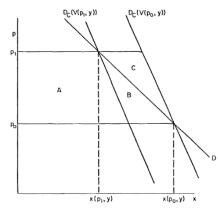
$$e(q^1, u) - e(q^0, u) = \int_{q^0}^{q^1} h(q, u) dq$$

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

(日) (周) (三) (三)

# 3. Comparing Surplus Measures Graphically (Auerbach 1985)

 $\bullet\,$  From Auerbach 1985 we have that Marshallian CS is A+B, CV is A+B+C and EV is A



- Note that h(q, v(q, Z)) = c(q, Z) because of duality (solution to utility max problem must coincide with solution to expenditure min problem at the same indirect utility level).
- Hence Hicksians corresponding to CV and EV must intersect Marshallians at the two prices (CV: q<sub>0</sub> and EV: q<sub>1</sub>).
- Intuition for why h(q, u) has a steeper slope than c(q, Z): only price effect, not price + income effects.
- ullet Note that with one price change EV < Marshallian Surplus < CV
  - but not true with multiple price changes b/c Consumer (Marshallian) Surplus not well-defined.

- Key: No path depedence with Hicksian measures:
- Why? Slutsky equation:

$$rac{\partial h_i}{\partial q_j} = rac{\partial c_i}{\partial q_j} + c_j rac{\partial c_i}{\partial Z}$$

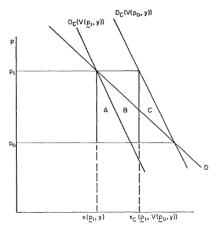
• Symmetry of the matrix  $(\frac{\partial h_i}{\partial q_j})_{ij} \rightarrow$  no path-dependence problem in this integral.

- Defining the EB measures
- Deadweight burden: change in consumer surplus less tax paid; what is lost in excess of taxes paid.
- In addition to Marshallian measure, two measures of EB, corresponding to *EV* and *CV*:

• 
$$EB(u_1^1) = EV - (q_1^1 - q_2^0)h(q_1^1, u_1^1)$$
 [Mohring 1971]

•  $EB(u^0) = CV - (q^1 - q^0)h(q^1, u^0)$  [Diamond & McFadden 1974]

• Figure from Auerbach 1985 shows different EB measures (A=EV, A+B=marshallian, C=CV)





3

(日) (周) (三) (三)

- Observations from this figure:
  - ▶ In general the three measures of *EB* will differ.
  - EV and CV no longer bracket the Marshallian one.
  - Key point is that Marshallian measure overstates EB.
- In the special case with no income effects (quasilinear utility) then CV = EV and there is a unique definition of consumer surplus and DWB

A B F A B F

# 3. Deriving Empirically Implementable Formula for EB based on EV and CV

- Suppose tax on good 1 is increased by Δτ units from a pre-existing tax of τ. No other taxes in the system.
- Recall

$$\textit{EB} = [\textit{e}(\textit{p} + \textit{\tau}, \textit{U}) - \textit{e}(\textit{p}, \textit{U})] - \textit{\tau}\textit{h}_1(\textit{p} + \textit{\tau}, \textit{U})$$

• Use a second-order Taylor expansion of formula for marginal excess burden *MEB* 

$$MEB = rac{dEB}{d au}(\Delta au) + rac{1}{2}(\Delta au)^2 rac{d^2EB}{d au}$$

Note that

$$\frac{dEB}{d\tau} = h_1(p+\tau, U) - \tau \frac{dh_1}{d\tau} - h_1(p+\tau, U) = -\tau \frac{dh_1}{d\tau}$$
$$\frac{d^2 EB}{d\tau^2} = -\frac{dh_1}{d\tau} - \tau \frac{d^2 h_1}{d\tau^2}$$

Derivations in first line come from envelope theorem

Hilary Hoynes ()

- 3. General Model: Harberger Formula with Income Effects
  - Ignoring  $\frac{d^2h}{d\tau^2}$  term (common practice but not well justified does not go to zero as  $\Delta \tau$  approaches zero), we get

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

- Same formula as the Harberger trapezoid derived above, but using Hicksian demands.
- Note that first-order term vanishes when τ = 0; this is the precise sense in which introduction of a new tax has "second-order" deadweight burden (proportional to Δτ<sup>2</sup> not Δτ).
- Without pre-existing tax, obtain "standard" Harberger formula:

$$EB=-rac{1}{2}rac{dh_1}{d au}(\Delta au)^2$$

イロト 不得下 イヨト イヨト 二日

## 3. General Model: Harberger Formula with Income Effects

- Bottom line: need to estimate compensated (substitution) elasticities to compute EB, not uncompensated elasticities.
- How to do this empirically?
  - Need estimates of income and price elasticities (and subtract off the income effect).
- Why do income effects not matter?
  - Not a distortion in transactions: if you buy less of a good because you are poorer, this is not an efficiency loss (no surplus left on table b/c of incomplete transactions).

(日) (周) (三) (三)

# 3. General Model: Harberger Formula with Taxes on Multiple Goods

- Previous case had only tax on good 1.
- With multiple taxed goods and fixed producer prices, can extend formula above to

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

- Problem: very hard to implement b/c you need to know all cross-price elasticities, so people usually just implement one-good Harberger formula.
- Goulder and Williams (JPE 2003) argue that one-good formula can be very misleading.

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

# 4. Empirical Applications: Structural vs. Reduced-Form

- Harberger formulas are empirically implementable, but approximations.
- Why use approximate formulas as above at all? Alternative approach: full (structural) estimation of demand model
- Comparison between two methods highlights some advantages & disadvantages of "sufficient statistic" approach

#### 4. Empirical Applications: Structural vs. Reduced-Form

- Consider loss in social surplus from tax change (previously focused only on individual)
- Simplifying assumptions made here:
  - No income effects (quasilinear utility)
  - Constant returns to production (fixed producer prices)
- N goods:  $x = (x_1, ..., x_N)$ . (Pre-tax) Prices:  $(p_1, ..., p_N)$ . Z = wealth
- Normalize  $p_N = 1$  ( $x_N$  is numeraire)
- Government levies a tax t on good 1

イロト イポト イヨト イヨト 二日

#### 4. Empirical Applications: Structural vs. Reduced-Form

• Individual takes t as given and solves

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

• To measure excess burden of tax, define social welfare as sum of individual's utility and tax revenue:

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

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

(日) (同) (三) (三)

# 4. Structural, Reduced Form and Sufficient Statistics (Chetty 2009)

- Basic structure: primitives -> Sufficient statistics -> Welfare change
- Primitives =  $\omega_1$ ,  $\omega_2$ , ...  $\omega_n$ 
  - preferences, constraints
  - not uniquely identified]
- Sufficient statistics =  $\beta_1(\omega, t), \beta_2(\omega, t)$ ,etc
  - functions of primitives, taxes
  - Possibly ientified using quasi-experimental variation  $y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$
- Welfare Change  $\frac{dW}{dt}(t)$ 
  - Used for policy analysis

#### 4. Empirical Applications: Structural vs. Reduced-Form

- Structural method: Estimate N-1 good demand system and recover u (Hausman 1981)
- Alternative: deadweight loss "triangle" popularized by Harberger (1964)
  - Envelope conditions for  $(x_1, ..., x_N)$  yield simple formula

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

- $\frac{dx_1}{dt}$  is a sufficient statistic for calculating change in welfare
- > Do not need to identify full demand system, simplifying identification

#### 4. Empirical Applications: Structural vs. Reduced-Form

- Benefit of sufficient statistic approach is particularly evident in a model that permits heterogeneity across individuals
  - Structural method requires estimation of demand systems for all agents
  - Sufficient statistic formula is unchanged still need only slope of aggregate demand <sup>dx1</sup>/<sub>dt</sub>
- Economic intuition for robustness of sufficient statistic approach:
  - ▶ Key determinant of deadweight loss is difference between marginal willingness to pay for good x<sub>1</sub> and its cost (p<sub>1</sub>).
  - Recovering marginal willingness to pay requires an estimate of the slope of the demand curve because MWTP coincides with marginal utility
- Many more applications of this type of reasoning throughout the course
- Modern public finance theory literature basically aims to connect theory with evidence using "sufficient statistics."

イロト 不得 トイヨト イヨト 二日

## 4. Empirical Applications: Feldstein JPE 1995 & REStat 1999

- Following Harberger, large literature in labor estimated effect of taxes on hours worked to assess efficiency costs of taxation
- Feldstein observed that labor supply involves multiple dimensions, not just choice of hours: training, effort, occupation
- Taxes also induce inefficient avoidance/evasion behavior
- As such, if you want to examine the full DWL you somehow have to deal with all these dimensions. Two approaches:
- Structural (or explicit) approach: account for each of the potential responses to taxation separately (separate elasticities) and then aggregate
- Reduced form (sufficient statistic): Feldstein shows that the elasticity of taxable income with respect to taxes is a sufficient statistic for calculating deadweight loss

Hilary Hoynes ()

#### 4. Empirical Applications: Deriving Feldstein 1999 Result

Model Setup

- Government levies linear tax t on (reported taxable) income
- Agent makes N labor supply choices: l<sub>1</sub>,...l<sub>N</sub> (hours, training, occupation, etc.)
- Each choice  $I_i$  has disutility  $\psi_i(I_i)$  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 post-tax taxable income plus untaxed income:  $x_N = (1 t)TI + e$

▲ロト ▲圖ト ▲画ト ▲画ト 三直 - のへで

#### 4. Empirical Applications: Deriving Feldstein 1999 Result

- With this setup, Feldstein shows that the DWL of the income tax is equivalent to the DWL of an excise tax on ordinary consumption. Intuition is that since taxes do not change the relative price of the different margins of labor supply, then it is not necessary to know the elasticities of each margin.
- In terms of the model, he shows that:

$$\frac{dW}{dt} = t\frac{dTI}{dt}$$

Key intuition: marginal social cost of reducing earnings through each margin is equated at optimum → irrelevant what causes change in TI.

イロト イポト イヨト イヨト

#### 4. Empirical Applications: Deriving Feldstein 1999 Result

• He then shows that

$$DWL = -0.5 \frac{t^2}{1-t} \epsilon_C C = -0.5 \frac{t^2}{1-t} \epsilon_{TI} TI$$

- ► Therefore: to eval the full DWL of taxation we can use the estimated elasticity of taxable income →sufficient statistic
- - See for example Gruber & Saez JPubE 2002. We will talk about this literature later in the course.
- A disadvantage of this sufficient statistic approach: primitives (eg  $g(e), \psi(l)$ ) are not estimated, assumptions never tested

- 本部 ト 本 ヨ ト - オ ヨ ト - - ヨ

- Marion and Muehlegger study DWL of diesel tax
- Two uses of diesel fuel: business/transport and residential (heating homes).
  - Residential use untaxed
  - Business use taxed (Federal: 24.4 cents/gallon, State: 8-32]
- Low to no cost to move between two uses (can buy for home use and resell for truck use and thus evade the tax)
- Substantial scope for evasion
- Oct 1, 1993: Government added red dye to residential diesel fuel. Easy to check if a truck is using illegal fuel by just opening the gas tank.
  - Evasion effectively much more costly.
  - Sharp time setting?

- This paper is interesting because: High MTR can lead to DWL through (at least) two channels:
  - Changes in quantity demanded (or supplied)
  - ② Evasion (no change in quantity demanded, but behavior changes)
- It is hard to differentiate between these two sources. Suppose you observe taxes increasing and taxable income declining. You do not know if true economic activity has changed or if money has just been moved between taxable and untaxable sources. Surely both matter for DWL (that is what Feldstein's method is a useful one) but it is interesting to know which source is the one that matters.
- Their setting allows for a direct test of evasion, which is unusual in the literature
  - Most common is using audit study data

イロト 不得 トイヨト イヨト 二日

#### Two strategies:

- directly document evidence of change in evasive behavior: examine discontinuity in sales following regulatory change; look for differences in response by state using differences in state tax and state initial monitoring cost.
- estimate price and tax elasticities before and after reform (using cross-state variation in tax rates and world price series).

#### Data:

 state level data from EIA and Fed Hwy Admin by type of fuel use; both price and quantity, 1983-2003

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

• Time series evidence (national event study)

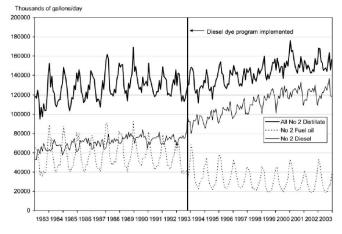
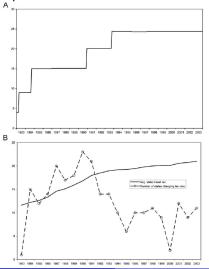


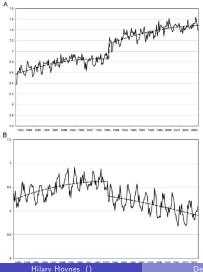
FIG. 2.-U.S. sales of No. 2 distillate

• Fig 3A, 3B, 4: show fed tax over time as well as ave state tax. The paper is not about variation in taxes over time. they are pretty stable



- Model: In  $q_{it} = \beta_0 + \delta_1 postdye_t + \Pi X_{it} + f(t) + \rho_i + \epsilon_{it}$
- f(t) quadratic in time, spline in pre and post period
- also includes calendar month fixed effects
- why not full set of month-year dummies? Look for discontinuity? Why not as nonparametric event time?
- why not include state specific seasonality controls?
- other data collected: weather, fraction of households using fuel oil
- Table 2: shows results of this model, 26% decline in diesel, 39% increase in fuel oil

• Regression adjusted smoothed national event study (Fig 5). What is the identifying assumption?



#### Other results

- (Tab 3) Estimate in levels. Full offset of decrease in fuel oil and increase in diesel oil. (Table 2, estimate in logs does not allow for testing for one-for-one offset)
- (Tab 4) Larger effects in states with high usage of home heating oil, larger effects in states with higher tax on diesel fuel.
- (Tab 5) More seasonality in demand for fuel oil in postdye period

- Estimation of elasticies: comparing elasticity of price and tax
- $\ln q_{it} = \beta_0 + \beta_1 \ln(p_{it}) + \beta_2 \ln(1 + \frac{\tau_{it}}{p_{it}}) + \Pi X_{it} + f(t) + \alpha_i + \epsilon_{it}$
- Estimate diesel fuel sales as a function of price of fuel and tax of fuel
- Instrument for price with world market shifters (Iraq war, Venezuela oil strike).
  - I am not sure why they express tax as fraction of the price. This requires instrumenting for both components.
  - Given the variation in state tax rates over time, you would think you could just use that variation (maybe some are 0?)
- Tests:
  - If no evasion then  $\beta_1 = \beta_2$ ; they allow for the elasticities to vary pre and post regulatory change.
  - You expect the tax to have no impact on demand for fuel sales.

- Results: elas of tax is much higher than elas of price before the regulatory change; after dye the elas of tax falls considerably. Also, impact of tax on fuel sales varies with pre and post period.
- Note: nothing about first stage of IV; no testing for difference in elasticities in post period

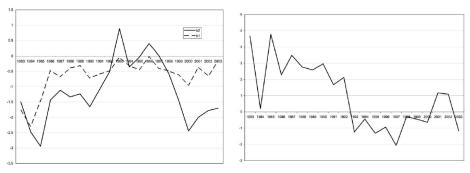
	Log Diesel Sales		LOG FUEL OIL SALES	
	OLS (1)	IV (2)	OLS (3)	IV (4)
Log(1 + tax rate)	-1.409 (.506)***	-1.333 (.511)**	2.184 (1.246)*	2.162 (1.311)*
Postdye × $log(1 + tax rate)$	.579	.597 (.965)	-1.995 (2.246)	-2.144 (2.343)
Log(price)	521 (.148)***	486 (.144)***	.442 (.407)	.862
Postdye × log(price)	.241 (.335)	.258 (.359)	.077 (.900)	657 (.829)
Observations $R^2$	10,236 .96	10,236 .96	9,866 .83	9,866 .83

#### TABLE 6 Price and Tax Elasticity Estimates

UC Davis, Winter 2012 55 / 81

イロト 不得下 イヨト イヨト 二日

• Expand the equation to estimate elas year by year:







 Diffs in elas close after change; expand again (new kind of evasion?), untaxed good elas falls to zero (so new evasion is not about untaxed good)

(4) (2) (4) (4) (4)

- Conclusions:
- Elasticities imply that 1% increase in tax rate raised revenue by 0.60% before reform vs. 0.71% after reform.
- Using revenue formulation of DWL, implies that DWL reduced from 40 cents to 30 cents b/c demand became more inelastic (25% reduction in excess burden!)

#### 5. Optimal Commodity Taxation: What is the problem?

- Goal is to maximize social welfare (minimize DWL) subject to revenue constraint
- First best:
  - Suppose we have perfect information, complete markets, perfect competition, lump sum taxes feasible at no cost.
  - Result: Second welfare theorem implies that any Pareto-efficient allocation can be achieved as a competitive equilibrium with appropriate lump-sum transfers (or taxes).
  - Economic policy problem reduces to the computation of the lump-sum taxes necessary to reach the desired equilibrium. Equity-efficiency trade-off disappears.
- Problems with first best:
  - No way to make people reveal their characteristics at no cost: to avoid paying a high lump-sum, a skilled person would pretend to be unskilled.
  - So govt has to set taxes as a function of economic outcomes: income, property, consumption of goods →distortion and DWL

#### 5. Optimal Commodity Taxation: What is the problem?

- So we end up with 2nd best world with inefficient taxation
  - cannot redistribute or raise revenue for public goods without generating efficiency costs.
- Here we discuss <u>optimal commodity tax</u> [optimal income taxation later]
- Four main qualitative results in optimal tax theory:
  - Ramsey inverse elasticity rule
  - Diamond and Mirrlees: production efficiency [not covered here]
  - Atkinson and Stiglitz: no consumption taxation with optimal non-linear (including lump sum) income taxation [not covered here]
  - Chamley/Judd: no capital taxation in infinite horizon models [not covered here]

イロト 不得下 イヨト イヨト 二日

Model

- Ramsey (1927) Tax problem: Government sets taxes on uses of income so as to raise a given amount of revenue *E* and minimize utility loss.
- Structure of problem: Govt is maximizing subject to S, D (and each S and D are solving contstraint opt problem)
- One individual or homogeneous individuals (no redistributive concerns):
- Agent's problem:
  - utility function:  $u(x_1, ..., x_N, I)$
  - Z = non wage income; I = leisure
  - w = wage rate,  $q_i =$  consumption (gross of tax) prices.
  - Utility maximized subject to:  $q_1x_1 + ... + q_Nx_N \le wl + Z$

Individual maximization

$$L = u(x_1, ..., x_N, I) + \alpha(wI + Z - (q_1x_1 + ... + q_Nx_N))$$

- FOCs  $u_{x_i} = \alpha q_i$
- Get demand functions  $x_i(q, Z)$  and indirect utility function V(q, Z) where  $q = (w, q_1, ..., q_N)$ .
  - $\alpha = \partial V / \partial Z$  is marginal utility of income for the individual.
  - Roy's identity:  $\partial V / \partial q_i = -x_i \partial V / \partial Z$

More Assumptions

- Z = 0, no exogenous income.
- Assume that producer prices are fixed (constant returns to scale and input prices fixed)
- Therefore WLG production prices normalized to one:  $p_i = 1$  and so  $q_i = 1 + \tau_i$ .
- We assume that labor is untaxed.
  - ▶ No loss of generality. Any tax system can be identically described by a tax on N 1 goods.

#### 6. Ramsey Tax Problem: Government's problem

- Two equivalent ways to set-up the Ramsey (government's) problem:
- max V(q, Z = 0) subject to  $\tau \cdot x = \sum_i \tau_i x_i (q, Z = 0) \ge E$ 
  - Max utility of rep agent subject to revenue constraint
  - [We will use this approach]
- min EB(q) = e(q, V(q, Z = 0)) e(p, V(q, Z = 0)) E subject to  $\tau \cdot x = \sum_i \tau_i x_i(q, Z = 0) \ge E$ 
  - Min EB (eval at post-tax utility) subject to revenue constraint
- Note equivalence with EV, not with CV (need to use actual post-tax price measure to identify optimum)

#### 6. Ramsey Tax Problem: Government's problem

max V(q, Z=0) subject to  $au \cdot x = \sum_i au_i x_i (q, Z=0) \geq E$ 

- Solve by *perturbation argument* (important to know method, allows for more intuitive grasp)
- General idea: suppose government increases  $\tau_i$  by  $d\tau_i$ .
  - changes in gov's objective since tax revenue changes (+)
  - changes in gov's objective since private welfare changes (-)
  - the optimum is characterized by balancing effects from tax revenue changes with effects from private welfare changes.

#### 6. Ramsey Tax Problem: Government's problem

• Effects on revenue:





• Effect on Private Welfare (Utility):

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

- Intuition: private welfare cost equivalent to taking lump sum of x<sub>i</sub>dτ<sub>i</sub> (envelope condition - Roy's identity).
- λ = marginal social welfare from additional revenue requirement (multiplier on gov budget)

• Optimum:

$$dU + \lambda dE = 0$$

 Of course can arrive at the same thing using FOC from the Lagrangian)

Hilary Hoynes ()

- 6. Ramsey Tax Problem:  $dU + \lambda dE = 0$ 
  - Substituting previous expressions for dU and dE and simplifying (divide by  $d\tau_i = dq_i$ ) gives you:

$$(\lambda - \alpha)x_i + \lambda \sum_j \tau_j \frac{\partial x_j}{\partial q_i} = 0$$

Optimal tax rates satisfy the Ramsey Formula

$$\sum_{j} \tau_{j} \frac{\partial x_{j}}{\partial q_{i}} = -\frac{x_{i}}{\lambda} (\lambda - \alpha)$$

for i = 1, ..., N defines a system of N equations and N unknowns.

- Connection with excess burden:
  - minimizing excess burden across goods for each tax
  - consider  $\lambda = 1$  and  $\alpha = 1$ :  $\sum_j \tau_j \frac{\partial x_j}{\partial q_i} = 0$

- Ramsey rule is often written in terms of Hicksian (compensated) elasticities to obtain further intuition.
- To do this, start by defining

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

- Note that  $\theta$  is independent of *i* (constant across goods).
- Interpretation of  $\theta$ :
  - $\blacktriangleright$   $\theta$  measures the value for the government of introducing a \$1 lumpsum tax:
  - Say the government introduces a \$1 lumpsum tax:
    - ① Direct value for the government is  $\lambda$
    - 2 Loss in welfare for the individual is  $\alpha$
    - Sehavioral loss in tax revenue because of the response dx<sub>j</sub> due to the income effect for the individual. This affects tax revenue by ∂(Σ<sub>j</sub> τ<sub>j</sub>x<sub>j</sub>)/∂Z

• Can demonstrate  $\theta > 0$  at the optimum using Slutsky matrix.

Hilary Hoynes ()

67 / 81

• Use  $\theta$  and Slutsky equation:

$$\partial x_j / \partial q_i = \partial h_j / \partial q_i - x_i \partial x_j / \partial Z$$

• After substituting & rearranging (and using symmetry of Slutsky,  $S_{ij} = S_{ji}$ ), get compensated representation of Ramsey tax formula:

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

 "Sum of price elasticities weighted by tax rates are constant across goods."

イロト 人間ト イヨト イヨト

$$rac{1}{ extsf{x}_i}\sum_j au_jrac{\partial h_i}{\partial q_j}=-rac{ heta}{\lambda}$$

- Intuition: 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  $dh_i = \tau_j \partial h_i / \partial q_j$ .
  - Therefore the total reduction in consumption of good *i* due to the tax system (all taxes together) is Σ<sub>j</sub> τ<sub>j</sub> ∂h<sub>i</sub>/∂q<sub>i</sub>.
  - Divide by x<sub>i</sub>, and you get the percentage reduction in consumption of each good *i* (normalization for revenue) due to the tax system:
  - $\frac{1}{x_i} \sum_j \tau_j \frac{\partial h_i}{\partial q_j}$  is called the 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.

Hilary Hoynes ()

イロト 不得下 イヨト イヨト 二日

• Alternate representation: compensated elasticities representation:

$$\sum_j rac{ au_j}{1+ au_j} arepsilon_{ij}^{m{c}} = rac{ heta}{\lambda}$$

 Important case 1:. e<sub>ij</sub> = 0 for all i ≠ j (cross price elas =0). Then obtain classic inverse elasticity rule:

$$\frac{\tau_i}{(1+\tau_i)} = \frac{\theta}{\lambda \epsilon_{ii}^c}$$

- higher the elasticity then the lower the optimal tax
- Intuition: link with DWB in partial eq. model.

イロト 不得下 イヨト イヨト 二日

- Important case 2: Suppose all cross elasticities are zero:  $\frac{\partial h_j}{\partial q_i} = 0$  for all  $i \neq j$  and all goods have same complementarity with leisure ( $\varepsilon_{x_i,w}$  constant)
- Then it turns out that we obtain

$$rac{ au_i}{ extsf{q}_i} = rac{ heta}{\lambda} rac{1}{rac{\partial h_i}{\partial w} rac{ extsf{w}}{ extsf{x}_i}} = rac{ heta}{\lambda} arepsilon_{ extsf{x}_i, w}$$

- Main point of Important Case 2: If all the goods have the same degree of complementarity with labor then <sup>τ<sub>i</sub></sup>/<sub>q<sub>i</sub></sub> is constant – uniform taxation.
- More generally, under assumptions of Case 2, want to tax the goods that are more complementary with labor less.
- Result depends critically on assumption of no cross-elasticities across other goods.

- COVER ONLY IF HAVE TIME
- Previous analysis essentially ignored production side of economy by assuming that producer prices are fixed.
- Diamond-Mirrlees AER 1971 tackle the optimal tax problem with endogenous production.
- D-M Result: even in an economy where first-best is unattainable (i.e. 2nd Welfare Thm breaks down), it is optimal to have production efficiency that is, no distortions in production of goods..
- The result can also be stated as follows. Suppose there are two industries, x and y and two inputs, K and L. Then with the optimal tax schedule, production is efficient:

$$MRTS_{KL}^{x} = MRTS_{KL}^{y}$$

even though allocation is inefficient:

$$MRT_{xy} \neq MRS_{xy}$$

イロト 不得下 イヨト イヨト 二日

- Example: Suppose gov can tax consumption goods and also produces some goods on its own (e.g. postal services).
  - May have intuition that gov should try to generate profits in postal services by increasing the price of stamps.
  - This intuition is wrong: optimal to have production efficiency!
- Before D-M, was suggested that optimal policy is highly dependent on particular market failures (e.g. monopolies, information failures, externalities, etc.).
- Their result: independent of market failures, optimal policy involves no distortion in production
- Bottom line: gov should only tax things that appear in agent's utility functions and should not distort production decisions via taxes on intermediate goods, tariffs, etc.

- Many consumers (index h), many goods (i) and inputs.
- Producer prices are not constant: production set that represents the production possibilities of the economy.
- Important assumption: profits do not enter into social welfare.
  - either constant returns to scale in production (no profits) or pure profits can be fully taxed.
- Government chooses different tax rates on all the different goods  $(\tau_1, ..., \tau_N)$  (that is, chooses the vector  $q = p + \tau$ ):  $\max_{q} W(V^1(q), ..., V^H(q)) s.t. \sum_{i} \tau_i \cdot X_i(q) \ge E.$

where  $X_i(q) = \sum_h x_i^h(q)$  sum of demands • Constraint can be replaced by

$$X(q) = \sum_{h} x^{h}(q) \in Y$$

where Y = production set (accounts for gov's requirement E) = 2000Hilary Hoynes () Deadweight Loss UC Davis, Winter 2012 74 / 81

- Production efficiency result: at the optimum level of taxes q\* that solves the problem, the allocation X(q\*) is on the boundary of Y.
- Proof by contradiction: Suppose  $X(q^*)$  is in the interior of Y.
- Then take a commodity that is desired by everybody (say good *i*), and decrease the tax on good *i* a little bit.
- Then X(q<sup>\*</sup> − dτ<sub>i</sub>) ∈ Y for small dτ<sub>i</sub> by continuity of demand functions. So it is a feasible point.
- Everybody is better because of that change:

$$dV^{h} = -V^{h}_{q_{i}}d\tau_{i} = V^{h}_{R}x^{h}_{i}d\tau_{i}.$$
  

$$d\tau_{i} < 0 \Rightarrow dV^{h} > 0 \forall h$$
  

$$\Rightarrow q^{*} \text{ is not the optimum.} Q.E.D$$

イロト 不得下 イヨト イヨト 二日

Important policy consequences of this result

- Public Sector production should be efficient.
  - If there is a public sector producing some goods (postal services, electricity,...): it should face the same prices as the private sector and choose production with the unique goal of maximizing profits, not generating government revenue.

- 4 週 ト - 4 三 ト - 4 三 ト

Important policy consequences of this result (continued)

- No taxation of intermediate goods (goods that are neither direct inputs or direct outputs consumed by individuals).
- Goods transactions between firms should go untaxed because taxing these transactions would distort (aggregate) production and destroy production efficiency.
- Example: Computer produced by IBM but sold to other firms should be untaxed
  - but the same computer sold to direct consumers should be taxed.
- Government sales of publicly provided good (such as postal services) to firms should be untaxed
  - but government sales to individual consumers should be taxed.
- Note: Marion-Muehlegger diesel fuel example is precisely the opposite of this!

イロト 不得 トイヨト イヨト 二日

Important policy consequences of this result (continued)

- Trade and Tariffs:
- In open economy, the production set is extended because it is possible to trade at linear prices (for a small country) with other countries.
- Diamond-Mirrlees result states that the small open economy should be on the frontier of the extended production set.
- Implies that no tariffs should be imposed on goods and inputs imported or exported by the production sector.
- Examples:
  - If IBM sells computers to other countries, that transaction should be untaxed.
  - If the oil companies buy oil from other countries, that should be untaxed.
  - If US imports cars from Japan, there should be no special tariff but should bear same commodity tax as cars made in US.

D-M Result hinges on two key assumptions:

- government needs to be able to set a full set of differentiated tax rates on each input and output
- government needs to be able to tax away fully pure profits (or production is constant-returns-to-scale);
  - otherwise can improve welfare by taxing industries that generate a lot of profits to improve distribution at the expense of production efficiency.
  - These two assumptions effectively separate the production and consumption problems.
    - Govt can vary prices of consumption goods without changing prices of production.
    - Even though govt is constrained to second-best situation in consumption problem, no reason to adopt second-best solution in production problem.
  - This separation of the consumption and production problems is why the results make sense in light of theory of the 2nd best.

Hilary Hoynes ()

Deadweight Loss

79 / 81

Practical relevance of the result is a bit less clear.

- Assumption 1 (differentiated tax rates) is not realistic.
- Example: skilled and unskilled labor inputs ought to be differentiated.
  - When they cannot (as in the current income tax system) then it might be optimal to subsidize low skilled intensive industries or set tariffs on low skilled intensive imported goods (to protect domestic industry). Naito JPubE 1999 develops this point in detail.

- A TE N - A TE N

Second Result of Diamond-Mirrlees:

- Optimal tax formulas even where producer prices are not constant take the same form as the Ramsey many persons problem.
  - Same formulas as in Ramsey just by replacing the p's by the actual p's that arise in equilibrium.
  - Incidence in the production sector can be completely ignored.

- A TE N - A TE N