Reviving the First Best

Or, When Can We Overcome Imperfect Targeting?

Lecture 9

ARE 264

February 15, 2022

Preparing for lecture 10

- Next lecture is our "modeling workshop"
- There is a prompt on bCourses. Read it, think about the problem, and upload some notes before class

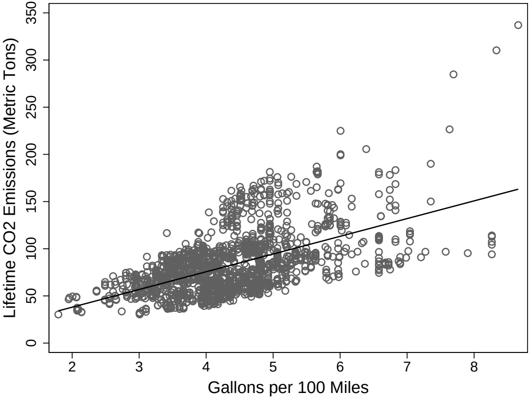
Lecture Recap

1. **What is the second-best tax rate when damages differ across sources of a pollutant but the tax rate must be uniform?**
   - Diamond (1973) shows that the second-best tax rate is a weighted-average of marginal damages, where weights are demand derivatives
2. **Discuss Knittel and Sandler (2018)**
3. **When can we estimate the welfare loss of using a restricted second-best policy compared to the first best?**
   - Jacobsen, Knittel, Sallee and van Benthem (2018) shows that, in some cases, can estimate this via simple regression statistics

Some limits of the Pigouvian prescription

- There seem to many situations where the marginal externality from the consumption or production of a particular good varies significantly across users or producers
- Examples?
- Does the Diamond model present a useful framework for any of these?
- Does Knittel and Sandler provide a blueprint for analysis?

Jacobsen, Knittel, Sallee and van Benthem



- When there is heterogeneity, under intuitive conditions, the R^2 of this figure is a sufficient statistic for the inefficiency of a fuel-economy policy that imposes linear function of gpm but ignores durability heterogeneity

$$\text{DWL} = -\frac{1}{2}\left[\underbrace{\sum_{j=1}^{J} e_j^2 \frac{\partial x_j}{\partial t_j}}_{\text{"own effects"}} + \underbrace{\sum_{j=1}^{J}\sum_{k\neq j} e_j e_k \frac{\partial x_j}{\partial t_k}}_{\text{"cross effects"}}\right] \quad \text{where} \quad e_j \equiv (\tau_j - \phi_j)$$

- Decompose formula into own and cross effects
- This is re-derivation of Harberger (1964), with externalities
- In some plausible cases, cross-effects will be zero (or small)

R^2 and SSR as Sufficient Statistics

When cross-effects cancel out, $\text{DWL} = -\frac{1}{2}\sum_{j=1}^{J} e_j^2 \frac{\partial x_j}{\partial t_j}$

- DWL is Sum of Squared "Tax Errors", weighted by own-price demand derivatives
- When demand derivative uncorrelated with errors, OLS is the second-best policy, and e_j will be a residual

- Example: CAFE imposes (implicit) tax vector that is linear in fuel consumption: $\tau_j = \alpha + \beta \cdot \text{gpm}_j$
- SB policy chooses α and β to minimize DWL $\Rightarrow$ solution is OLS
- DWL is demand-elasticity-weighted SSR

R^2 and SSR as Sufficient Statistics

- When derivative uncorrelated with errors and cross-effects cancel, the R^2 is the fraction of possible welfare gain achieved by the linear policy, where baseline is flat tax on all products:

$$R^2 = \frac{\text{DWL(OLS)} - \text{DWL(Constant)}}{\text{DWL(Pigouvian)} - \text{DWL(Constant)}} = \frac{\text{ESS}}{\text{TSS}}$$

- R^2 and the sum of squared residuals are sufficient statistics with a welfare interpretation

When are cross-effects zero?

$$\text{Cross effects} = \sum_{j=1}^{J} \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}$$

- Assumption: conditional on the policy variable, differences in externalities are uncorrelated with substitutability
- If errors are "white noise", then this condition will be met
- If closer substitutes have more similar errors, then cross-effects will not be exactly zero
- Explore this quantitatively; intuition differs across applications
    - Reasonable approximation for noisy MPG & electricity
    - Plausible for CAFE; test robustness
    - Fridges: important violation, derive alternative statistic, which is within-R^2 from panel regression

Paper's history

- Started by extending Diamond model to multiple goods
- Struggled to match empirical material to the theory
- Tried to just follow steps in Chetty (2009) and couldn't because it was too complicated, so first removed all the consumer heterogeneity but left the heterogeneity across goods
- Stumbled upon the analogy to regression statistics

- Paper had just the vehicle longevity application. Seminar feedback suggested maybe we needed to make a more general point. So we added two more applications (one was the original issue that I was thinking about in noisy ratings)
- Titled paper: "Sufficient Statistics for Imperfect Externality-Correcting Policies"
- Referee told us that we were not doing sufficient statistics (title change)
- Referees told us that our main application did not match our theory because we were ignoring other externalities/second-best considerations. We added fourth application, electricity pricing

Some limits of the Pigouvian prescription

- The Pigouvian prescription is a useful default (reference point), but there are reasons why it needs modification:

- What if I can't tax the externality directly?
- What about general equilibrium?
- What if there is another market failure?
- What if the market already fixed the problem? (Coase)
- What about equity?

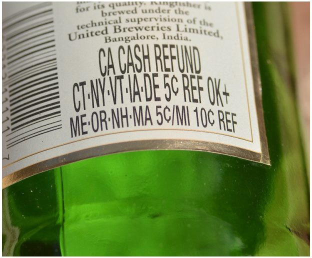
- Last lecture we discussed two cases where imperfect targeting implied second-best outcomes were optimal; **today we emphasize special situations where imperfect targeting can be overcome to return to the first best**

Outline

1. What is a **multi-part tariff** and what is it good for?
   - Multi-part tariffs can sometimes mimic an unavailable Pigouvian tax
2. How can we base taxes on **ambient pollution** levels?
   - Segerson (1988) points out that taxing all polluters with respect to total pollution can induce desired behavior.
3. What is the **additivity property** and when is it relevant?
   - Second-best tax on a dirty good often has additively separable Pigouvian tax with second-best ("Ramsey") component
4. What is the **marginal cost of public funds**?
   - MCPF represents welfare loss required to raise $1 revenue

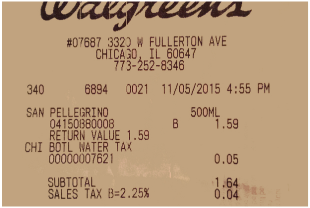
- Often we'd like to tax waste, but we won't be able to observe waste
- Basic idea: there are two ways to dispose of trash. One is socially desirable, but more costly. How can we get people to choose expensive option?

Two-part instruments



- Consider recycling bottles: you can throw it away, or recycle it
- Recycling takes more effort
- You could put Pigouvian tax (say \$0.05) on **trash**
- But, there is no credible way to observe bottles in trash—you can't measure the waste
- What could you do instead?

Two-part instruments



- You could subsidize recycling—pay \$0.05 per bottle recycled
- But, this is not optimal because it makes bottles cheaper—will lead to too many bottles in economy
- The solution is a **two-part instrument**, in this case a **deposit-refund**
- You tax \$0.05 per bottle when sold, and then you refund the \$0.05 when recycled
- This is identical to a Pigouvian \$0.05 tax on waste!

Multi-part instruments

- Sometimes you can **mimic** an unavailable tax on externality with several different taxes
  - Can be either a first-best or second-best version

- E.g., you want to tax people who throw away recyclable cans
  - Don't observe trash
  - Use **deposit-refund** system
  - Employ **presumptive tax** on all cans: charge as if they will all be wasted
  - Give a **refund** for recycling

- A subsidy for recycling alone gives the wrong incentives: it makes cans cheaper, when they should be more expensive
- Aside: what happens if one of the instruments is more **salient** than the other?

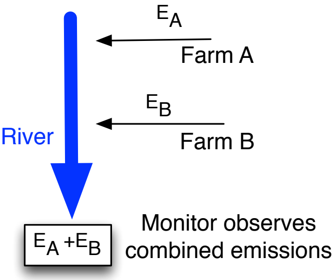
What's in the bag? Pareto efficiency.
Source: Stock Unlimited

- I like to call such a solution a **sneaky first-best** policy
- That is meant to denote something where the first-best incentives can be created, even though we can't directly tax emissions

- Fullerton and Wolverton (1997) emphasize that, in some cases, a two-part instrument can **exactly** recreate the unavailable Pigouvian tax
- Intuition in Fullerton and Wolverton:
  - Consider waste (pollution, etc.) as an input to production, along with other inputs like labor and capital
  - Pigouvian tax raises cost of final good by raising input prices: reduction via **output effect**
  - Also substitute to cleaner inputs via **substitution effect**
  - Intuition of a two-part instrument is to achieve these two effects separately in two instruments
- Multi-part instruments useful in second-best cases as well (see Fullerton and Wolverton 2005 JPubE)
- Eskeland and Devarajan (1996) argue for combining a gasoline tax with emissions regulations to mimic Pigouvian tax on auto emissions

Outline

1. What is a **multi-part tariff** and what is it good for?
   - Multi-part tariffs can sometimes mimic an unavailable Pigouvian tax
2. How can we base taxes on **ambient pollution** levels?
   - Segerson (1988) points out that taxing all polluters with respect to total pollution can induce desired behavior.
3. What is the **additivity property** and when is it relevant?
   - Second-best tax on a dirty good often has additively separable Pigouvian tax with second-best ("Ramsey") component
4. What is the **marginal cost of public funds**?
   - MCPF represents welfare loss required to raise $1 revenue



- In lots of situations, multiple polluters contribute to create an ambient pollution level
- We can measure the ambient level of pollution, but not emissions from each source
- Can we apply the Pigouvian prescription?

Segerson (1988) JEEM

- Consider a tax scheme imposed on (possible) polluter i, based on ambient pollution x:

$$T_i(x) = \begin{cases} t_i(x - \bar{x}) + k_i, & \text{if } x > \bar{x} \\ t_i(x - \bar{x}) & \text{if } x \leq \bar{x} \end{cases}$$

- where $\bar{x}$ is a constant (not a mean) that you can think of as the target pollution level
- Can write polluter's problem as:

$$\max_{a_i, q_i} \pi_i = pq_i - E[T_i(x(a, e))] - C_i(q_i, a_i)$$

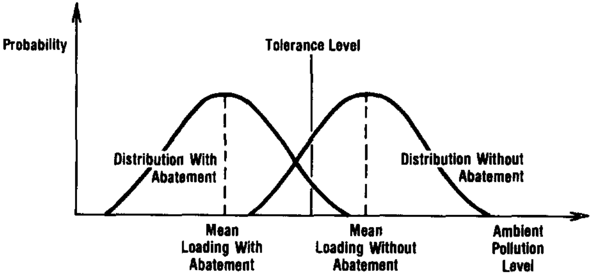
- Where q is output, a is abatement and C is total cost, with a and e without subscripts representing the vector of all polluters

Segerson (1988) JEEM

- Under a Cournot-Nash assumption about beliefs, efficient output can be achieved by choices of k_i alone, or t_i alone, or some combination

- The idea is that you can tax everyone according to aggregate emissions
- This gives everyone the right marginal incentives (so it creates the right behavior)
- But, this potentially taxes people who did not pollute, which runs into legal problems and is certainly not politically popular

Segerson (1988) JEEM



- Segerson considers multiple polluters as one example, but also points out the case of uncertainty about ambient measurements from a given level of emissions
- There are numerous papers that think about alternative cases, competition, etc.

Kotchen and Segerson (2018)

- Group rewards/penalties create a "local public good" for the affected firms
- Basic model extends to when firms collude, but only if planner knows firm collusion
  - This would seem to require too much information
  - Generally, ideas point to important interaction between policy design and Ostrom style questions on institutions and resource management
- Threshold policies do not have full efficiency and can be problematic, but also can be modeled concisely
  - Clean Air Act (where is the good theoretical paper describing the structure of the CAA?)
  - Threat of regulation as a threshold problem
- Paper has a collection of citations pointing to empirical work, mostly in resource management

Outline

1. What is a **multi-part tariff** and what is it good for?
   - Multi-part tariffs can sometimes mimic an unavailable Pigouvian tax
2. How can we base taxes on **ambient pollution** levels?
   - Segerson (1988) points out that taxing all polluters with respect to total pollution can induce desired behavior.
3. What is the **additivity property** and when is it relevant?
   - Second-best tax on a dirty good often has additively separable Pigouvian tax with second-best ("Ramsey") component
4. What is the **marginal cost of public funds**?
   - MCPF represents welfare loss required to raise $1 revenue

- Our discussion of **general equilibrium** issues and the Pigouvian prescription begins with Sandmo (1975), which sought to understand whether the existence of other markets changed the Pigouvian prescription
- Sandmo sets out a general equilibrium framework that makes some simplifying assumptions (linear production, perfect competition, etc.)

- First shows that in general equilibrium, but with **no distortions other than the externality**, Pigouvian prescription holds. The fact that there are GE effects does not alter the Pigouvian prescription
  - E.g., if taxing alcohol reduces demand for food eaten at restaurants, this is not a reason for tax to differ from marginal damages...unless restaurant food is not priced at marginal cost
- Second shows how second-best tax changes when there are distortions (i.e., other goods have P$\neq$MC)

Simple General Equilibrium version (Sandmo 1975)

Consumer/workers: n identical

Goods: $j = 0, ..., J$ with quantities x_j and aggregate q X_j

Good zero is labor: $x_0 =$ hours worked, $1 - x_0 =$ leisure

Good J is dirty: damages depend on aggregate X_J

Utility: $u(1 - x_0, x_1, ..., x_J, X_J)$

$\partial u / \partial x_j \equiv u_j > 0 \ (j = 0, ..., J)$

Marginal damage: $u_{J+1} < 0$

Production: $$-X_0 + \sum_{j=1}^{J} a_j X_j \leq Y \quad (a_i > 0 \forall i)$$

normalizes productivity of labor to 1

Utilitarian: $SWF = nu(\cdot)$

First-best allocation

Characterize first-best by choosing quantities to maximize social welfare function.

$$\mathcal{L}_P = nu(1 - x_0, x_1, ..., x_J, X_J) - \alpha \left(-X_0 + \sum_{j=1}^{J} a_i X_i \right)$$

FOCs imply:

$$\frac{u_j}{u_0} = a_j \quad (j = 1, ..., J-1)$$
$$\frac{u_J}{u_0} + n\frac{u_{J+1}}{u_0} = a_J$$

- Regular goods equate marginal utility ratio to marginal rate of technical substitution
- Dirty good has two terms: MU ratio and externality

Sandmo (1975)

1. Characterize first-best allocation via SWF assuming planner chooses quantities
2. Show that first-best allocation can be decentralized via Pigouvian tax
   - Optimal tax vector is $t_j = 0$ $(j = 1, ..., J-1)$ and $t_J = -n\frac{u_{J+1}}{u_0}$
   - No need to tax inputs or complements; **general equilibrium does not seem to matter** for tax
   - This works out because there are no other distortions
3. Impose revenue requirement
   - $T \geq \sum_{j=1}^{J}(P_j - p_j)x_j = \sum_{j=1}^{J} t_j x_j$
   - Now this is a second-best problem. The other goods will be distorted in order to raise revenue, so there are **pre-existing distortions** that might be exacerbated by a tax on the dirty good

What should we expect in the second best?

- Until now, there are no other market failures or distortions
- Next we suppose the government needs to raise revenue: $T \geq \sum_{j=1}^{J} t_j x_j$
- Planner chooses taxes on each good; <u>no lump-sum tax</u>
- Will need to tax commodities in order to raise revenue
- How will this change the tax on the dirty good?

- Intuition 1: **Ramsey taxation**—put higher taxes on less elastic good
- Intuition 2: **Corlett-Hague Rule (1953)**—model with no tax on labor; put higher taxes on goods that are more complementary to leisure
- Do these tell us anything about taxing the dirty good?

Second-best result

$$t_j = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) \;\; (j = 1, ..., J-1)$$

$$t_J = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) + \frac{-\mu}{\lambda}\left(-n\frac{u_{J+1}}{u_0}\right)$$

- Define D^* as the Jacobian of the demand matrix (e.g., the matrix of demand derivatives)
- D_{jk} is the cofactor; and D is the determinant of D^*

Second-best result

$$t_j = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) \ (j = 1, ..., J-1)$$

$$t_J = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) + \frac{-\mu}{\lambda}\left(-n\frac{u_{J+1}}{u_0}\right)$$

- λ is the shadow price on the planner's revenue constraint
- μ is marginal utility of income to agent
- μ/λ is the inverse of **marginal cost of public funds**
- MCPF = ratio of value of \$1 to government over \$1 to consumer
- $\mu/\lambda < 1$ implies we scale down Pigouvian element

Marginal cost of public funds

- What is the cost of raising \$1 of public funds?
- Answer determines the optimal provision of public goods, or optimal level of redistribution
- Cost is called the **marginal cost of public funds**
- MCPF can be defined as ratio of Lagrange multipliers
  - Multiplier on government's revenue requirement λ; is dW/dR
  - Multiplier on consumer's budget constraint μ; is marginal utility of income
  - MCPF $= \lambda/\mu$ is ratio of value of \$1 to government over \$1 to consumer
- E.g., in simple model with one factor L, constant productivity
  - With lump-sum taxation, government raises revenue at marginal utility of income: MCPF=1
  - With distortionary taxation on L: MCPF $= 1/(1-\varepsilon_L) > 1$

Marginal cost of public funds

- Generally, assume MCPF > 1
- Common estimates are around ≈ 1.3
- There are a lot of details debated in the literature, and competing definitions (e.g., does revenue use effect factor supply?)
- Start with Hakonsen (International Tax and Public Finance 1998)
- Will discuss a related concept: marginal value of public funds next week

Second-best result

$$t_j = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) \quad (j = 1, ..., J-1)$$

$$t_J = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) + \frac{-\mu}{\lambda}\left(-n\frac{u_{J+1}}{u_0}\right)$$

- Clean goods have a Ramsey tax
- **Additivity property:** dirty good has a Ramsey tax plus the Pigouvian tax

$$\frac{t_j}{P_j} = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{-1}{\varepsilon_j}\right) \;\; (j = 1, ..., J-1)$$
$$\frac{t_J}{P_J} = \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{-1}{\varepsilon_J}\right) + \frac{-\mu}{\lambda P_J}\left(-n\frac{u_{J+1}}{u_0}\right)$$

- Suppose all cross-price derivatives are zero, with ε the own-price elasticity
- This delivers Ramsey intuition: proportional tax higher for more elastic goods

Additivity property

- This is a second best setting, but you get back some "first best properties"
- The tax on the dirty good moves with externality; scaled only by MCPF
- The tax on clean goods is independent of their relation to the dirty good; i.e., you do not tax complements more
- Dixit (1985) calls this a **Principle of Targeting**—you want to correct an externality at its source; target directly and do not worry about correlated margins of choice

- Sandmo (1975) employs specific assumptions—representative consumer, fixed coefficients production, no nonlinear income taxation, no model of expenditures, etc.
- How general is additivity?

Generalizing additivity property (Kopczuk 2003)

- Kopczuk (2003) shows additivity (Principle of Targeting) is very general
  - Key requirement is that you be able to **tax the externality directly** (in contrast to material last class)
- You want to add a Pigouvian tax, and then optimize as if you were in a problem with no externalities; "correct the externality and then ignore it"

Outline

1. What is a **multi-part tariff** and what is it good for?
   - Multi-part tariffs can sometimes mimic an unavailable Pigouvian tax
2. How can we base taxes on **ambient pollution** levels?
   - Segerson (1988) points out that taxing all polluters with respect to total pollution can induce desired behavior.
3. What is the **additivity property** and when is it relevant?
   - Second-best tax on a dirty good often has additively separable Pigouvian tax with second-best ("Ramsey") component
4. What is the **marginal cost of public funds**?
   - MCPF represents welfare loss required to raise $1 revenue