Incidence and Equity (Continued)

Lecture 7

ARE 264

February 8, 2022

Preparing for lecture 8

- Complete the bCourses prompt on heterogeneituy
- Read Knittel & Sandler (2018), Jacobsen, Knittel, Sallee and van Benthem (2020)

Lecture 7 Recap

1. **How do we incorporate equity into our analysis of environmental policy?**
   - The "optimistic separability" view suggests that efficiency and equity can be divorced and considered separately. This isn't entirely true, but it is an important perspective
2. **If we do care about equity, how might pollution and policy affect the distribution of welfare?**
   - Fullerton suggests several channels. This flags the need to understand incidence, including in general equilibrium
3. **Quick notes on incidence**
4. **How can we model the incidence of environmental policy?**
   - Fullerton suggests several channels. This flags the need to understand incidence in a general equilibrium setting

Outline

1. **Remaining material on incidence**
   - Fullerton and Heutel demonstrate how to translate Harberger model to apply to environmental taxes, delineate channels of impact in general equilibrium
2. **Are environmental taxes regressive?**
   - Many have a direct price effect that is regressive, but this depends on revenue recycling, as well as the welfare base (income versus consumption)
3. **What are the roots of optimistic separability?**
   - Optimistic separability owes in large part to Atkinson and Stiglitz, but recent work demonstrates various ways that this breaks down due to preference heterogeneity
4. **Can we really compensate losers?**
   - Heterogeneity implies that it is hard to target losers, implying that we can't really get Pareto improvements from Pigouvian taxes

Tax theory and environmental policies

- Two closely related papers on the syllabus by Fullerton and Heutel modify the Harberger model to assess the incidence of environmental policy
- The first addresses and environmental tax; the second addresses non-tax policies

What is the key point in Fullerton and Heutel models?

- Some of these models will get a little confusing (at least for me!)
- The main take away I want you to get:
  - A sense of the different **channels**—output effect, factor substitution effects, etc.—that are relevant for general equilibrium
  - A sense of how to adapt alternative policies into an incidence model
  - An example of "cross field arbitrage"
  - Some humility—"in general, anything can happen" (see Goulder, Hafstead and Williams in future lecture)
- But, also, I wonder if the GE models from trade have largely superseded this approach

Fullerton and Heutel 2007

- How might we introduce pollution into the Harberger model?

- Fullerton and Heutel add pollution Z as an input

$$X = X(K_X, L_X)$$
$$Y = Y(K_Y, L_Y, Z)$$

- Price of pollution is just tax τ_Z
- No limit on pollution

- Fixed K and L (just shift between sectors); perfect competition; constant returns to scale
- Use log-linearization: hats denote proportional changes; e.g., $\hat{K_X} \equiv dK_X/K_X$

Key objects that allow us to interpret results

- Allen elasticity (constant output) of substitution for inputs $\mathbf{e_{ij}}$
  - I.e., how much does demand for i rise in response to rise in price of j, holding output fixed
  - $e_{ij} > 0$ for substitutes; $e_{ij} < 0$ for complements
- Factor shares of revenue by sector $\theta_{sector,factor}$
  - E.g., $\theta_{YZ} = \tau_Z Z/(p_Y Y)$
  - Analogous factor shares for other factors, sectors
  - Factor shares within sector add to 1 ($\theta_{YK} + \theta_{YL} + \theta_{YZ} = 1$)
- Factor intensities, written as ratios of dirty/clean $\gamma_{\mathbf{factor}}$
  - $\gamma_K \equiv K_Y/K_X$; higher ratio implies more of K goes to Y
  - If $\gamma_{\mathbf{K}} - \gamma_{\mathbf{L}} > 0$ then dirty good is <u>K intensive</u>
- Elasticity of demand between goods Y and Z: $\sigma_{\mathbf{u}}$

$$\hat{p}_Y = \frac{(\theta_{YL}\theta_{XK} - \theta_{YK}\theta_{XL})\theta_{YZ}}{D}[A(e_{ZZ} - e_{KZ}) - B(e_{ZZ} - e_{LZ}) + (\gamma_K - \gamma_L)\sigma_u]\hat{\tau}_Z + \theta_{YZ}\hat{\tau}_Z \tag{11a}$$

$$\hat{w} = \frac{\theta_{XK}\theta_{YZ}}{D}[A(e_{ZZ} - e_{KZ}) - B(e_{ZZ} - e_{LZ}) + (\gamma_K - \gamma_L)\sigma_u]\hat{\tau}_Z, \tag{11b}$$

$$\hat{r} = \frac{\theta_{XL}\theta_{YZ}}{D}[A(e_{KZ} - e_{ZZ}) - B(e_{LZ} - e_{ZZ}) - (\gamma_K - \gamma_L)\sigma_u]\hat{\tau}_Z, \tag{11c}$$

$$\begin{aligned}\hat{Z} = -\frac{1}{C}[&\theta_{YK}(\beta_K(e_{KK} - e_{ZK}) + \beta_L(e_{LK} - e_{ZK}) + \sigma_u)\hat{r} + \theta_{YL}(\beta_K(e_{KL} - e_{ZL}) \\ &+ \beta_L(e_{LL} - e_{ZL}) + \sigma_u)\hat{w} + \theta_{YZ}(\beta_K(e_{KZ} - e_{ZZ}) \\ &+ \beta_L(e_{LZ} - e_{ZZ}) + \sigma_u)\hat{\tau}_Z]\end{aligned} \tag{11d}$$

- A, B, C and D are additional terms, substituted for readability
- Results describe incidence of a change in tax rate on pollution τ_Z
- To understand incidence, just want to know how prices change

- Fullerton and Heutel (2007) considers a tax
- But, lots of policies are not a tax
- **Lots of policies that are not actually a tax can still be understood as a tax**

Fullerton and Heutel (2010)

- Fullerton and Heutel (2010) study general equilibrium incidence of environmental mandates
- Same setup as Fullerton and Heutel (2007), but different policies
- Two sectors: one clean and one dirty

$$X = X(L_X, K_X) \quad Y = Y(L_Y, K_Y, Z)$$

where Z is pollution

- Model **performance standard** as limit on emissions per unit of output: $Z/Y \leq \delta$
- Model **technology mandate** as limit on emissions per unit of capital: $Z/K_Y \leq \zeta$

Performance standard

- Dirty firm solves:

$$\max_{K_Y, L_Y, Z} p_Y Y(K_Y, L_Y, Z) - rK_Y - wL_Y$$
$$\text{s.t. } Z/Y \leq \delta$$

- Pollution tax model (and original Harberger model) has **output effect**; a pollution tax will shift demand from Y toward X; this lowers relative factor price for factor used intensively in dirty production
- Performance standard creates new **output-subsidy effect**: constraint is relaxed by increasing Y; which benefits factor used intensively in Y
- Performance standard taxes pollution **but subsidizes output** (same insight as Holland, Hughes and Knittel, to be discussed later)

Technology mandate

- Dirty firm solves:

$$\max_{K_Y, L_Y, Z} p_Y Y(K_Y, L_Y, Z) - rK_Y - wL_Y$$
$$\text{s.t. } Z/K_Y \leq \zeta$$

- This subsidizes K: K directly increases production, but it also allows more pollution (which raises production)
- A **capital-subsidy effect** creates an effect that favors K over L in the dirty sector
- By subsidizing K used in dirty sector, this creates the same type of output-subsidy effect as under the performance mandate

Outline

1. **Remaining material on incidence**
   - Fullerton and Heutel demonstrate how to translate Harberger model to apply to environmental taxes, delineate channels of impact in general equilibrium
2. **Are environmental taxes regressive?**
   - Many have a direct price effect that is regressive, but this depends on revenue recycling, as well as the welfare base (income versus consumption)
3. **What are the roots of optimistic separability?**
   - Optimistic separability owes in large part to Atkinson and Stiglitz, but recent work demonstrates various ways that this breaks down due to preference heterogeneity
4. **Can we really compensate losers?**
   - Heterogeneity implies that it is hard to target losers, implying that we can't really get Pareto improvements from Pigouvian taxes

Our two views of regressivity

- Conventional view
- Optimistic separability

- In either view, we need to know the distribution of burdens created by a policy

Are environmental taxes regressive?

- There is (obviously) no single answer
- Many papers study this question focusing on only the price channel
- Even then, regressivity will change completely depending on assumptions about revenue recycling
- There are also some subtleties that are worth noting here

- Is a carbon tax (or an energy tax) regressive (before considering distribution of environmental benefits)?

- Answer depends on:
  - Revenue recycling will largely determine final impact
  - Consumption base looks less regressive than income base
  - Existing policy already creates some automatic feedback (tax code)

Summary of direct price effect literature on energy taxes

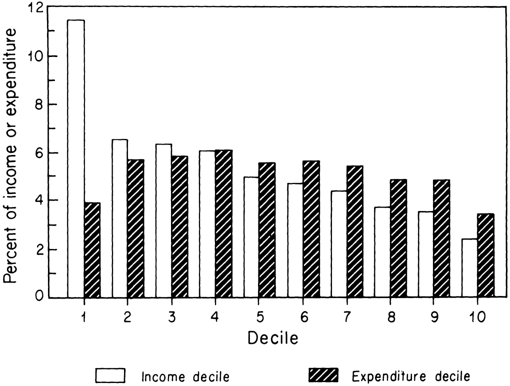

FIGURE 1. ***Gasoline share of income or expenditure.***

- Poterba (1991 Tax Policy and Economy): gasoline tax regressive based on income
- Relationship much weaker when using **consumption** as base

- Initial literature focuses on **direct** effects (i.e., fossil fuels)
- **Indirect** effects consider how prices of all other goods change due to energy price changes

- Standard approach is to use input-output table
- Assume full pass through of all prices
- Do not model affect on workers
- Hassett, Mathur and Metcalf one example

Table 1. Distribution of Burden: Annual Income

Decile	1987			1997			2003		
	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total
Bottom	2.77	1.23	4.04	3.07	1.42	4.53	2.12	1.60	3.74
Second	2.18	1.10	3.29	2.25	1.19	3.47	1.74	1.31	3.06
Third	1.68	0.96	2.64	1.90	1.08	2.99	1.36	0.99	2.36
Fourth	1.59	0.84	2.45	1.52	0.89	2.42	1.19	0.88	2.06
Fifth	1.19	0.74	1.92	1.21	0.75	1.98	0.97	0.78	1.76
Sixth	1.02	0.66	1.69	1.00	0.69	1.69	0.85	0.68	1.53
Seventh	0.94	0.58	1.51	0.89	0.64	1.53	0.69	0.61	1.30
Eighth	0.83	0.57	1.40	0.78	0.56	1.34	0.61	0.63	1.23
Ninth	0.68	0.52	1.19	0.63	0.51	1.14	0.53	0.49	1.01
Top	0.53	0.49	1.01	0.43	0.43	0.85	0.36	0.45	0.81

Source: Authors' calculations. The table reports the within decile average ratio of carbon tax burdens to income

Hasset, Mathur and Metcalf: measuring with income—table shows carbon tax burden as percent of income

Table 2. Distribution of Burden: Current Consumption

Decile	1987 Direct	1987 Indirect	1987 Total	1997 Direct	1997 Indirect	1997 Total	2003 Direct	2003 Indirect	2003 Total
Bottom	1.41	0.51	1.92	1.37	0.50	1.90	0.98	0.50	1.49
Second	1.37	0.50	1.89	1.18	0.50	1.68	0.92	0.49	1.41
Third	1.12	0.51	1.64	1.09	0.49	1.58	0.84	0.50	1.34
Fourth	1.09	0.52	1.61	0.95	0.51	1.46	0.79	0.50	1.29
Fifth	1.02	0.53	1.55	0.89	0.51	1.40	0.73	0.51	1.24
Sixth	0.91	0.55	1.46	0.82	0.51	1.34	0.65	0.52	1.16
Seventh	0.86	0.55	1.40	0.75	0.51	1.26	0.65	0.51	1.17
Eighth	0.75	0.56	1.31	0.66	0.51	1.17	0.54	0.53	1.07
Ninth	0.71	0.56	1.28	0.58	0.52	1.09	0.48	0.52	1.00
Top	0.57	0.57	1.14	0.46	0.51	0.97	0.37	0.52	0.89

Hasset, Mathur and Metcalf: measuring with current consumption
Results stable over time
Regressivity from direct effects (indirect flat)

Table 3. Distribution of Burden: Lifetime Consumption

Decile	1987			1997			2003		
	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total
Bottom	1.11	0.52	1.63	1.14	0.52	1.61	0.67	0.50	1.16
Second	0.94	0.53	1.48	0.90	0.51	1.42	0.66	0.51	1.16
Third	0.96	0.53	1.49	0.85	0.50	1.36	0.76	0.51	1.24
Fourth	1.09	0.53	1.62	0.90	0.50	1.40	0.72	0.51	1.23
Fifth	1.05	0.53	1.59	0.91	0.51	1.42	0.81	0.50	1.30
Sixth	0.97	0.53	1.51	0.90	0.51	1.41	0.71	0.50	1.22
Seventh	0.95	0.53	1.49	0.87	0.51	1.37	0.66	0.51	1.18
Eighth	0.94	0.55	1.48	0.83	0.51	1.34	0.56	0.51	1.08
Ninth	0.73	0.55	1.28	0.71	0.51	1.22	0.49	0.53	1.02
Top	0.65	0.56	1.21	0.60	0.52	1.11	0.41	0.52	0.93

Hasset, Mathur and Metcalf: measuring with "lifetime consumption"
Lifetime consumption even less regressive than current consumption

Cronin, Fullerton and Sexton JAERE 2019

Notwithstanding these limitations, this analysis yields three key findings. First, despite the fact that electricity constitutes a high fraction of spending for poor families, we find that a US carbon tax is progressive, not regressive as commonly assumed. In fact, its progressivity is a necessary consequence of the following four basic points: (1) once consumption is adopted as the measure of well-being, then a uniform consumption tax is not regressive but proportional; (2) as shown below, our calculated family total carbon consumption is not clearly concentrated in high or low consumption deciles, which, with the first point, makes a carbon tax nearly proportional;[8] (3) transfers in the United States are indexed to correct for increases in consumer prices that accompany a carbon tax; and (4) transfers are a larger fraction of income for lower deciles.[9]

- Paper highlights variance in impacts within income groups (horizontal inequity), show that rebating <u>exacerbates</u> variation
- Indexing of many policies to prices creates automatic feedback loop in existing welfare state that has significant impact on regressivity

A few other touch points

- Williams, Gordon, Burtraw, Carbone and Morgenstern (NTJ 2015) examine incidence with similar tools
- Goulder, Hafstead, Kim and Long, "Impacts of a Carbon Tax Across US Household Income Groups: What are the Equity-Efficiency Trade-Offs?" JPubE 2019 consider factor price channel and price impacts together
- Borenstein and Davis (TPE 2015) show regressivity of green subsidies
- On distribution of benefits: key earlier paper Fowlie, Holland and Mansur AER 2012, recent example Hernandez-Cortes and Meng WP, see Fowlie, Walker and Wooley (Brookings 2020)
- JAERE 2019 special issue on equity

Outline

1. **Remaining material on incidence**
   - Fullerton and Heutel demonstrate how to translate Harberger model to apply to environmental taxes, delineate channels of impact in general equilibrium
2. **Are environmental taxes regressive?**
   - Many have a direct price effect that is regressive, but this depends on revenue recycling, as well as the welfare base (income versus consumption)
3. **What are the roots of optimistic separability?**
   - Optimistic separability owes in large part to Atkinson and Stiglitz, but recent work demonstrates various ways that this breaks down due to preference heterogeneity
4. **Can we really compensate losers?**
   - Heterogeneity implies that it is hard to target losers, implying that we can't really get Pareto improvements from Pigouvian taxes

Back to our two views of regressivity

- Conventional view
- Optimistic separability

- A critical component of the optimistic separability view is a form of preference homogeneity
  - With heterogeneous preferences, the results (taken literally) break down in some ways

Roots of optimistic separability

- Atkinson and Stiglitz (1976) shows that, given an optimized nonlinear income tax, commodity taxes should not be used to achieve redistribution

- Implies we should not try to cut prices on "essential goods", but instead rely on the income tax, for example

- Based on a model with a form of separability between labor supply and other consumption and on preference heterogeneity (the only reason consumption bundles differ is income)

Roots of optimistic separability

- Atkinson and Stiglitz (1976) has some restrictive assumptions, but it has been enormously influential

- Specifically, another pillar of tax theory is Ramsey (1927), which says the optimal commodity tax should be inversely related to its elasticity (to minimize DWL)
- Diamond (1975) generalized that to many types of consumers, showing an intuitive correlation between higher commodity taxes on goods consumed disproportionately by the rich

- Atkinson and Stiglitz says to ignore all that and just use the income tax
- There are various ways of breaking down the result, but for a long time they focused on the labor supply issue, less attention on the preference heterogeneity

Allcott, Lockwood and Taubinsky “Ramsey”

- The “Ramsey Strikes Back” paper is a nice intro to these issues
- It basically restores the reasoning of Diamond (1975) in the presence of nonlinear taxation, assuming that there is imperfect tax salience
- My point in assigning this is not really to focus on that result, but to introduce you to the material

Allcott, Lockwood and Taubinsky "Soda Taxes"

- The related "Soda Tax" paper, QJE 2019 does a much more thorough job of providing the full theory model for a related problem

- It is in the QJE because the theory is good, but also because it combines an empirical application that is fairly impressive

- What is really impressive is the fact that the theory is pretty general and comprehensive, but they are able to translate it into "sufficient statistics" that provide a roadmap for empirical estimation, and then to do a "good enough" job of quantifying those empirical elements

Outline

1. **Remaining material on incidence**
   - Fullerton and Heutel demonstrate how to translate Harberger model to apply to environmental taxes, delineate channels of impact in general equilibrium
2. **Are environmental taxes regressive?**
   - Many have a direct price effect that is regressive, but this depends on revenue recycling, as well as the welfare base (income versus consumption)
3. **What are the roots of optimistic separability?**
   - Optimistic separability owes in large part to Atkinson and Stiglitz, but recent work demonstrates various ways that this breaks down due to preference heterogeneity
4. **Can we really compensate losers?**
   - Heterogeneity implies that it is hard to target losers, implying that we can't really get Pareto improvements from Pigouvian taxes

- Kaplow's reasoning assumes the ability to design a well-targeted transfer system
- Such schemes may not be **"technologically feasible"** because:

- There is substantial heterogeneity in consumption of externality-creating goods
- Heterogeneity poorly predicted by variables that determine the transfer function
- Transfer functions therefore necessarily imprecise, which makes Pareto improvements infeasible
- If you base transfers on consumption, you distort incentives

- This is a "technical" (information) problem, not a lack of political will or an obstacle created by an interest group
  - Same problem relevant for other efficient policies (e.g., free trade, eliminating the mortgage interest tax deduction, etc.)

Sallee (2020) Can you actually make everyone better off?

- **Research question:** When is it truly possible to achieve a Pareto improvement from a Pigouvian tax?

- **Main idea:** background transfers can compensate losers, but they must be **well targeted**, and targeting may be impossible when there is a lot of heterogeneity
- **Theory:** what factors make a Pareto improvement possible/impossible?
- **Data:** test our ability to design a Pareto improving system for a gas tax in US

- If we cannot make everyone better off, must acknowledge the creation of losers in discussing policies

Visual depiction of theory

σ

$>$ +

Δ

Heterogeneity
>> DWL triangle

Pigou

Pareto

A straw man?

Does anyone really think you can make everyone better off?

- It is an important straw man:
  - Influential tradition separates equity and efficiency issues (Musgrave 1959, Atkinson-Stiglitz 1976, Kaplow 2012)
  - Casual arguments that "everyone can win from efficiency" abound
  - Some careful authors do note that real policies will create some losers
- This paper provides theoretical framework; derives testable condition; has original empirical estimates

A red herring?

Who cares about creating a Pareto improvement?

- Aiming for a Pareto improvement is not at all the same as maximizing social welfare
- You can preserve "average" progressivity and have efficient policy—so, who cares if someone loses?

- Political economy motivation:
  - Political systems show a bias towards the status quo
  - Small groups may often be able to veto change (Olson 1965, 1982)

- Questions/comments about the referee reports and letter from the editor?

Outline

1. **Remaining material on incidence**
   - Fullerton and Heutel demonstrate how to translate Harberger model to apply to environmental taxes, delineate channels of impact in general equilibrium
2. **Are environmental taxes regressive?**
   - Many have a direct price effect that is regressive, but this depends on revenue recycling, as well as the welfare base (income versus consumption)
3. **What are the roots of optimistic separability?**
   - Optimistic separability owes in large part to Atkinson and Stiglitz, but recent work demonstrates various ways that this breaks down due to preference heterogeneity
4. **Can we really compensate losers?**
   - Heterogeneity implies that it is hard to target losers, implying that we can't really get Pareto improvements from Pigouvian taxes