Second-Best Policy I

Or, What if You Cannot Directly Target an Externality?

Lecture 8

ARE 264

February 10, 2022

Preparing for lecture 9

- bCourses prompt on Ambient pollution due before class
- Additional reading is Segerson 1988

Lecture 8 Recap

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

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?

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?

- In various situations, a policy will not be able to perfectly target emissions, so the Pigouvian prescription cannot be applied. There is no general case that describes all such situations, but we will consider two possibilities today

Outline

1. What is the second-best tax rate when damages differ across sources of a pollutant but the tax rate must be uniform?

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?

Outline

1. What is the second-best tax rate when damages differ across sources of a pollutant but the tax rate must be uniform?
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?

Example of Pigouvian tax that is difficult to implement: tailpipe emissions

- What pollution externalities come out of this truck?

Example of Pigouvian tax that is difficult to implement: tailpipe emissions

- What pollution externalities come out of this truck?
- **Greenhouse gases**: carbon dioxide (CO_2)
- **Local air pollution**: carbon monoxide (CO), hydrocarbons (HC), particulate matter (PM), nitrous oxide (NO_X)

Example of Pigouvian tax that is difficult to implement: tailpipe emissions

- What pollution externalities come out of this truck?
- **Greenhouse gases**: carbon dioxide (CO_2)
- **Local air pollution**: carbon monoxide (CO), hydrocarbons (HC), particulate matter (PM), nitrous oxide (NO_X)
- You can target GHG with a gas tax; but very hard to target local air pollution

Figure 3.21: Vehicle A instrumentation setup

- The Pigouvian prescription tells us to tax each driver according to their **emissions** (multiplied by damages per ton of emissions)
- But, we can't directly monitor tailpipe emissions—it is prohibitively expensive (administrative costs!); and damages depend on location as well as emissions quantities
- If we are limited to a gasoline tax, what is the optimal gasoline tax to target local air pollution?

What research questions might one ask about second-best settings like the one described here?

What research questions might one ask about second-best settings like the one described here?

- If a first-best Pigouvian tax is unobtainable, what is the second-best tax scheme?

- If we impose a second-best policy, how much welfare is lost compared to the first-best?
  - I.e., how bad is the second-best?

- If we are stuck with the second-best, is a different instrument better than a tax (tradable permits)?

Heterogeneous consumers (Diamond 1973)

- Diamond considers a model with one externality causing good and n consumers, each of whom might create a heterogeneous marginal damage
- Good x causes externality
- n consumers
- Quasilinear utility; with separable damage function
- Each consumer has different externality
- Utility for consumer i: $U^i(x_1, ..., x_n) + \mu_i$
- Where μ_i is income remaining for consumption on numeraire (after x_i is purchased); Y_i is total income
- Assume $\partial^2 U^i / \partial x_i \partial x_j = 0$ for $i \neq j$
- Assume all externalities negative: $\partial U^i / \partial x_j$ for $i \neq j$
- x produced with constant marginal cost; price p and tax t

Consumer's problem:

$$\max_{x_i} \; U^i(x_1, ..., x_n) + \mu_i$$
$$\text{s.t. } (p + t)x_i + \mu_i = Y_i$$

- Given all the simplifying assumptions, demand for x_i depends only on $p_i + t_i$
- Write as $x_i(p_i + t_i)$
- Own-price derivative denoted x_i'

Planner's problem:

- Suppose planner can set single tax rate on x common to all n
- Utilitarian welfare function
- Maximize with respect to tax rate, subject to total resource constraint in economy
- Math below implies lump-sum revenue recycling

$$\max_t SWF = \sum_{i=1}^{n} U^i(x_1(p+t), ..., x_n(p+t))$$

$$\text{s.t. } p\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \mu_i = \sum_{i=1}^{n} Y_i$$

Differentiate planner's problem wrt t:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial U^i}{\partial x_j}x_j' - p\sum_{i=1}^{n}x_i' = 0$$

Substitute consumer's optimality condition: $U_i^i = p + t$

$$\sum_{i=1}^{n}\sum_{j\neq i}\frac{\partial U^i}{\partial x_j}x_j' + t\sum_{i=1}^{n}x_i' = 0$$

Solve for t:

$$t^* = \frac{-\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{\partial U^i}{\partial x_j}x_j'}{\sum_{i=1}^{n}x_i'} = \sum_{j=1}^{n}\left(\frac{x_j'}{\sum_{k=1}^{n}x_k'}\sum_{i\neq j}\frac{\partial U^i}{\partial x_j}\right)$$

Diamond (1973) result

$$t^* = \sum_{j=1}^{n} \left(\frac{x_j'}{\sum_{k=1}^{n} x_k'} \sum_{i \neq j} \frac{\partial U^i}{\partial x_j} \right)$$

- For j, externality is $\sum_{i \neq j} \frac{\partial U^i}{\partial x_j}$
- Optimal tax is a weighted average of the Pigouvian taxes that would apply to each type
- Weights are demand derivatives
- Second-best tax is the weighted average of Pigouvian taxes; individuals who are more responsive to prices are weighted more heavily
- Now, SB tax depends on demand derivatives, but only as weights

Note that this example contains a basic algorithm for determining the optimal second-best tax rate.

1. Derive the consumer's optimality conditions, conditional on the tax rate
2. Write the planner's problem, assuming they have a choice over the tax rate and differentiate the planner's problem, allowing each endogenous choice to be a function of the tax rate
3. Substitute the consumer's optimality condition into the planner's FOC and (hopefully) solve

What are some other settings that have a Diamond "flavor"?

- SO_2 or NO_X damages (spatial differentiation)
  - Literature on tradable permits considers **trading ratios**

Outline

1. What is the second-best tax rate when damages differ across sources of a pollutant but the tax rate must be uniform?
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?

Outline

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?

Outline

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?

Knittel and Sandler (2018)

- What is the research question?
- What is the main contribution?
- What are the strengths and weaknesses of the paper?

Knittel and Sandler (2018)

- Asks how effective is a gasoline tax at reducing local air pollution
- Why is this a second-best problem (rather than a first-best Pigouvian tax)?

Knittel and Sandler (2018)

- Asks how effective is a gasoline tax at reducing local air pollution
- Why is this a second-best problem (rather than a first-best Pigouvian tax)?
- Great data on emissions for each vehicle from CA Smog Check
- Use panel nature of data to estimate mileage response to gasoline price fluctuations

Knittel and Sandler (2018)

- Draw explicit analogy to Diamond model
- Show that gasoline price elasticity is positively correlated with emissions (older, dirtier cars have bigger mileage response to gasoline price)
- What does this imply about the second-best uniform tax rate, as compared to the Pigouvian benchmark?

Knittel and Sandler (2018)

- Draw explicit analogy to Diamond model
- Show that gasoline price elasticity is positively correlated with emissions (older, dirtier cars have bigger mileage response to gasoline price)
- What does this imply about the second-best uniform tax rate, as compared to the Pigouvian benchmark?
- Second-best gasoline tax is higher than “average marginal damages per gallon”

Knittel and Sandler (2018)

- Draw explicit analogy to Diamond model
- Show that gasoline price elasticity is positively correlated with emissions (older, dirtier cars have bigger mileage response to gasoline price)
- What does this imply about the second-best uniform tax rate, as compared to the Pigouvian benchmark?
- Second-best gasoline tax is higher than "average marginal damages per gallon"
- The Diamond formula, however, does not indicate welfare gains (just optimal stringency)

Knittel and Sandler (2018)

- To get welfare gains, need some additional assumptions
- Suppose that mileage distribution follows $m = \beta_0 + \beta_1 dpm(p_g + \tau)$
- Suppose the externality per mile E is log normal with pdf $\psi(E_i)$
- Then, DWL absent policy is $D = (2\beta_1)^{-1} e^{2\mu_E + 2\sigma_E^2}$
- Then, ratio of remaining DWL after imposing uniform tax is:

$$R = 1 - e^{-\sigma_E^2}$$

Knittel and Sandler (2018)

- To get welfare gains, need some additional assumptions
- Suppose that mileage distribution follows $m = \beta_0 + \beta_1 dpm(p_g + \tau)$
- Suppose the externality per mile E is log normal with pdf $\psi(E_i)$
- Then, DWL absent policy is $D = (2\beta_1)^{-1} e^{2\mu_E + 2\sigma_E^2}$
- Then, ratio of remaining DWL after imposing uniform tax is:

$$R = 1 - e^{-\sigma_E^2}$$

- The point of this is NOT to argue that this is a correct or special way of approaching the question of quantifying welfare gains
- Rather, Knittel and Sandler is a nice example of a paper that combines great reduced form empirical analysis with a simple welfare model in order to translate reduced form results into a welfare conclusion

Outline

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?

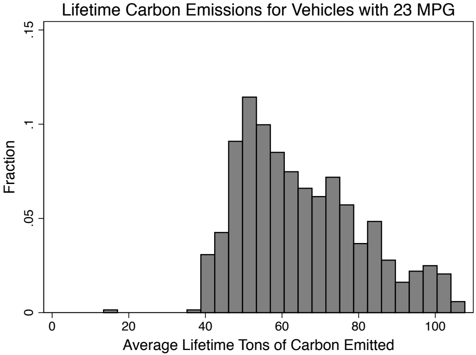
Outline

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

Jacobsen, Knittel, Sallee and van Benthem

- Many policies take form that there is an externality that is a function of many factors $e(A, B)$
- Policy can be contingent on only a subset of them $t(A)$
- What is welfare cost of basing policy on A instead of A and B?

Automobiles and durability



Source: CA Smog Check data

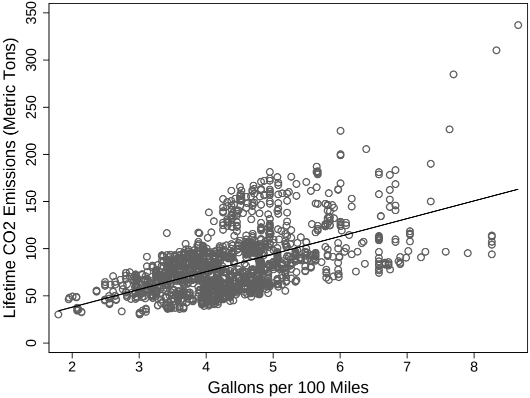
- Policy favors fuel economic cars because they consume less gasoline
- But, emissions depend on both fuel economy and lifetime miles traveled (durability)

Sufficient statistics

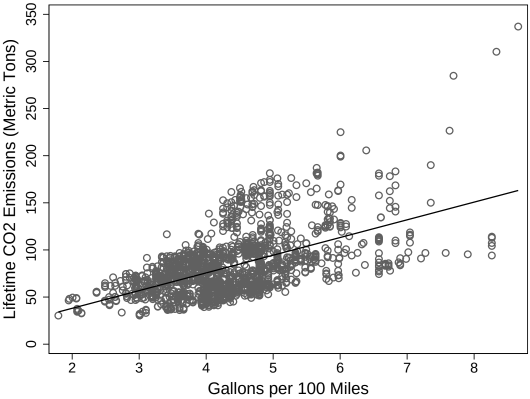
- **Goal:** build framework to characterize welfare costs of policies that regulate/tax based on limited number of dimensions that determine total externality, which requires limited market information (sufficient statistics, in spirit of Chetty (2009))

- Consider linear tax on fuel consumption: $t_j = \alpha + \beta gpm_j$
- This approximates actual policy (CAFE)
- With no intensive margin, if all cars have common lifetime mileage this is Pigouvian tax on emissions

Sufficient statistics



Sufficient statistics



- 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

Recall our "questions one might ask about second-best settings"

Recall our "questions one might ask about second-best settings"

- If a first-best Pigouvian tax is unobtainable, what is the second-best tax scheme? Diamond, Knittel and Sandler

- If we impose a second-best policy, how much welfare is lost compared to the first-best? Knittel and Sandler, JKSvB
  - I.e., how bad is the second-best?

- If we are stuck with the second-best, is a different instrument better than a tax (tradable permits)?

Theory: setup

Theory: setup

- Representative consumer buys portfolio of $j = 1, \ldots, J$ products (e.g., cars); quantity is x_j
- Utility over bundle of goods $U(x_1, \ldots, x_J)$
- Each product generates an externality ϕ_{j} (linear in x_j); represents marginal damage from product j
- Production cost function $C(x_1, \ldots, x_J)$
- Perfectly competitive supply
- Exogenous income M; consumer buys products and quasi-linear numeraire n
- Planner can place a tax on each product t_j
- Consumer prices are $p_j + t_j$

Theory: notes and caveats

Theory: notes and caveats

- Fixed products (short run, no technology change)
- No intensive margin $\Rightarrow$ no utilization effect (for car case: expected lifetime VMT when new is exogenous)
- Model can be recast as including random failure rates for durable goods, so long as all consumers share ex ante distribution

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

- Decompose formula into own and cross effects
- This is re-derivation of Harberger (1964), with externalities

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

- 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

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

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 below; intuition differs across applications
  - Reasonable approximation for noisy MPG
  - Plausible for CAFE; test robustness
  - Fridges: important violation, derive alternative statistic, which is within-R^2 from panel regression

Application 2: spatial heterogeneity

Spatial heterogeneity in emissions: many policies give same subsidy/tax to externality-generating products across space; but marginal damages are very different in different places (and times)

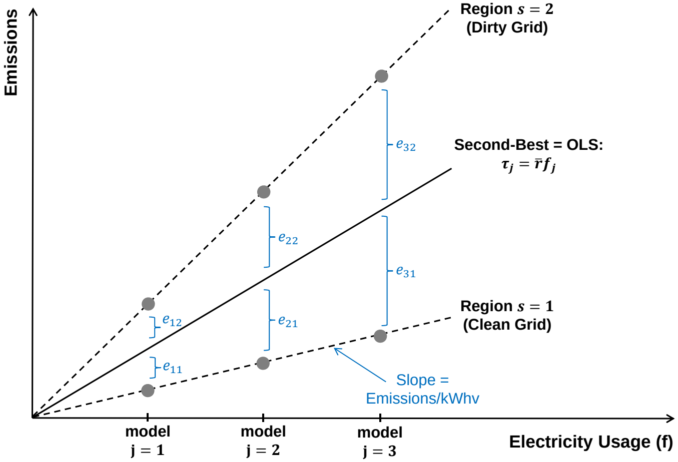
Application 2: spatial heterogeneity

Spatial heterogeneity in emissions: many policies give same subsidy/tax to externality-generating products across space; but marginal damages are very different in different places (and times)

- Emissions standards for new vehicles: same for large parts of the U.S., but damages from NO_x and VOCs differ spatially
- Appliances: subsidized based on electricity conservation, but emissions from electricity vary widely
- Use refrigerators as example

- Do same sufficient statistics apply?

Average and relative mis-pricing across regions



Modified R^2 for spatial heterogeneity

- Assume an inelastic outside good
  - Isolates relative price distortion
  - May be good assumption for refrigerators

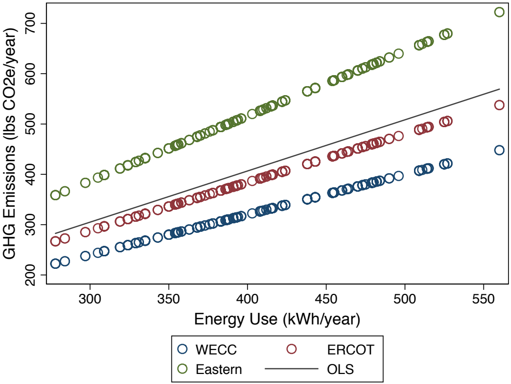
Modified R^2 for spatial heterogeneity

- Assume an inelastic outside good
    - Isolates relative price distortion
    - May be good assumption for refrigerators
- Fraction of first-best welfare gain achieved by the second-best policy over a policy of a constant unbiased tax on all products:

$$\frac{DWL(\tau = \bar{r} f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)} = 1 - \frac{var(r_s)}{E[r_s^2]}.$$

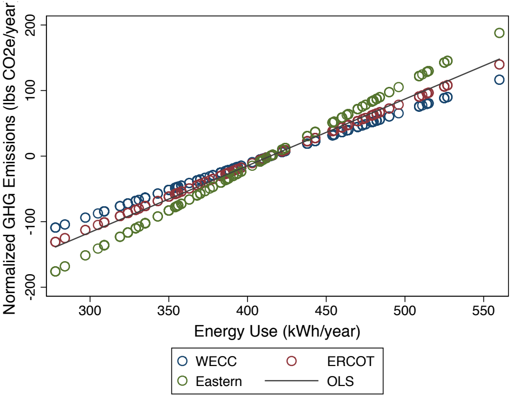
- This is equal to the R^2 from an OLS regression of regionally demeaned damages $\phi_{js} - r_s \bar{f}$ on the product attribute f_j
- This result does not rely on any further assumptions about the demand system (e.g., on cross-effects)

Refrigerator example



- f_j from cross-section of 564 refrigerators in same class
- r_s from electricity regions (Graff Zivin, Kotchen and Mansur 2014)
- Raw R^2 low (0.33); lots of error from regional bias

Refrigerator example



- Modified $\tilde{\phi}_{js}$ is residual from ϕ_{js} regressed on s fixed effects
- Modified R^2 high (0.90-0.96, depending on number of regions)
- **Conclusion:** relative price distortion surprisingly innocuous; national policy very efficient unless extensive margin large

Intellectual history

- I had been wondering about the implications of 2008 MPG fuel reform, which created a noisy re-ordering of ratings
- I discussed Knittel and Sandler in 2012: "heterogeneity is great, but where is the scrappage?"
- I then discussed Jacobsen and van Benthem (2015), which is about scrappage: "scrappage is great, but isn't heterogeneity really important?"

- Started by extending Diamond model to multiple goods
- Struggled to match empirical material to the theory
- I tried to just follow steps in Chetty (2009) and couldn't because it was too complicated, so I 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?
- What about equity?

Outline

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