Environment Field Exam Review Document

Simon Greenhill

May 2022

Contents

1 Meredith

1.1 Questions on past exams

Almost certainly will get a discrete choice-related question:

- 2021: Conditional logit, comparing mixed and conditional logit, based on Grigolon et al. (2018)
- 2020: Conditional logit, estimating household MWTP for fans in developing country context, limitations of conditional logit
- 2019: Conditional logit, Berry two-step, estimating MWTP, mixed logit, based on Ito and Shuang (2019) [not bolded on 2021 syllabus]
- 2018: Environmental justice, neighborhood sorting, based on Depro et al. (2015)
- 2017: Randomized experiment design, mixed logit, based on Ito et al., "Information frictions, inertia, and selection on elasticity: a field experiment on electricity tariff choice" [not on 2021 syllabus]

1.2 Randomized experiments and power calculations

Things to watch out for:

- Hawthorne/experimenter demand effects
  - Example: study sending people notification that they were being monitored as part of a study reduced their energy consumption

1.2.1 Treatment effects

Treatment effects in encouragement designs

$$ITT = \underbrace{\pi_C}_{\substack{\text{Compliance}\\\text{rate}}} \times \underbrace{(Y_C(1) - Y_C(0))}_{\text{LATE for compliers}}$$

1.2.2 Power calculations

In a simple RCT regression framework $Y_i = \alpha + \tau D_i + \varepsilon_i$ with $\varepsilon \sim N(0, \sigma_e^2)$, the variance of $\hat{\tau}$ is

$$\begin{aligned} V[\hat{\tau}] &= \frac{\hat{\sigma}^2_{Y|D}}{\sum(D_i - \overline{D})^2} \\ &= \frac{\hat{\sigma}^2_{Y|D}}{PN(1-P)} \end{aligned}$$

Given type I error (false rejection) rate α and type 2 error (false fail to reject) rate β which is converted into power $\kappa = 1 - \beta$, we want the t-statistic for $\hat{\tau}$ to be at least $t_\beta + t_\alpha$. Using the variance formula above, we can calculate the sample size N needed to calculate a minimum detectable effect MDE:

$$N = \frac{(t_\beta + t_\alpha)^2}{P(1-P)} \frac{\sigma_e^2}{MDE^2}$$

One issue is that we need to estimate σ_e^2. One way to do so is by taking $V[Y_i]$. For a randomized encouragement design, scale N by $\frac{1}{(c-s)^2}$, where c is the share receiving treatment in the encouraged group and s is the share receiving treatment in the control group (the always takers).

1.3 IV

1.4 Panel data and DiD

Hernandez-Cortes and Meng (2020): Do environmental markets cause environmental injustice?

Research question: How did California Cap and Trade affect the "environmental justice gap," i.e. the gap in pollution exposure between disadvantaged communities and others?
Methods: Two steps. First, estimate the impact of cap and trade of firm-level emissions using a diff-in-diff (control group is firms that were just below the cutoff for inclusion in cap and trade). This estimate yields the ATE of cap and trade on emissions. Second, use air transport model (HYSPLIT) to estimate how downwind exposure changed due to the change in pollution calculated in step 1.
Findings: EJ gap narrowed under cap and trade.
Some concerns: Differences across treatment and control. Imposing an ATE in percentage reductions mechanically leads to more reductions for larger emitters; if these are disproportionally polluting minorities, this will narrow EJ gap.

1.5 Discrete choice

1.5.1 Basic setup

Choice environment: Agents choose one item $j \in \mathcal{J} = \{1, \ldots, J\}$. The choices must be:

- Mutually exclusive (can choose only one)
- Exhaustive ($\mathcal{J}$ captures all possible realizations)
- Finite

Goal: estimate the *conditional probability* than an agent n makes choice j condition on observable characteristics X. Some factors will be unobserved, these are "nuisance parameters."

Random utility model:

$$\underbrace{U_{ni}^*}_{\substack{\text{Choice-specific}\\ \text{utility}}} = \underbrace{U(X_{ni};\beta)}_{\substack{\text{Deterministic}\\ \text{component}}} + \underbrace{\varepsilon_{ni}}_{\substack{\text{Unobserved}\\ \text{component}}}$$

- Often assume vector ε_n of choice-specific disturbances for individual n are distributed iid.

Conditional choice probabilities: Can express the conditional prob. that agent n will choose i conditional on observables X as:

$$\begin{aligned} P\left[Y_n = i | X_n = X\right] &= P\left[U_{ni}^* > U_{nj}^*\right] \forall i \neq j \\ &= P\left[U(X_{ni}) + \varepsilon_{ni} > U(X_{nj}) + \varepsilon_{nj}\right] \forall i \neq j \\ &= P\left[\varepsilon_{nj} - \varepsilon_{ni} < U(X_{ni}) - U(X_{nj})\right] \forall i \neq j \end{aligned}$$

Define an indicator for whether n chooses i over j:

$$\mathbb{I}\left[\varepsilon_{nj} - \varepsilon_{ni} < U(X_{ni}) - U(X_{nj})\right] = 1 \iff n \text{ chooses } i \text{ over } j$$

Note that $\mathbb{I}\left[\cdot\right]$ is a function of the random vector ε_n. Taking the expected value of the indicator gives the choice probability:

$$P\left[Y_n = i | X_n = X\right] = \int_{\varepsilon_n} \mathbb{I}\left[\varepsilon_{nj} - \varepsilon_{ni} < U(X_{ni}) - U(X_{nj})\right] f(\varepsilon_n) d\varepsilon_n$$

The key idea is that if we know (or assume) $f(\varepsilon_n)$, we can compute the conditional choice probabilities.

1.5.2 Conditonal logit

Setup and choice probabilities

In a *conditional logit model*, choice probabilities are a function of the attributes of choice alternatives.

$$U_{nj}^* = U(X_{nj};\beta) + \varepsilon_{nj}, \text{ where } \varepsilon_{nj} \sim EV(1)^1$$

To re-cast this utility maximization problem as a cost minimization problem, simply subtract the EV1 disturbance instead of adding it. This will yield analogous expressions to those above.

For a conditional logit, the choice probabilities are:

$$\begin{aligned} P\left[Y_n = i | X_{ni}, \varepsilon_{ni}\right] &= P\left[V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \forall i \neq j\right] \\ &= P\left[\varepsilon_{nj} < \varepsilon_{ni} + V_{ni} - V_{nj} \forall j \neq i\right], \end{aligned}$$

[1] Type I EV, aka Gumbel distribution: $f(x) = \frac{1}{\sigma} e^{\frac{x-\mu}{\sigma}} e^{-e^{\frac{x-\mu}{\sigma}}}$. μ is the location parameter; σ is the scale parameter. We often work with a simplified version where $\mu = 0$ and $\sigma = 1$ (the standard Gumbel distribution). The CDF of the standard Gumbel distribution is $F(x) = e^{e^{-x}}$. The difference between two iid EV1 RVs, $\varepsilon_{nji}^* \equiv \varepsilon_{nj} - \varepsilon_{ni}$ is distributed logistic: $F(\varepsilon_{nji}^*) = \frac{e^{\varepsilon_{nji}^*}}{1+e^{\varepsilon}_{nji}}$. (It's pretty easy to show this using characteristic functions. See here.)

where $V_{ni} = U(X_{ni}; \beta)$. Because the εs are iid, we can write this as:

$$P\left[Y_n = i | X_{ni}, \varepsilon_{ni}\right] = \prod_{j \neq i} e^{e^{-(\varepsilon_{ni} + V_{ni} - V_{nj})}}.$$

Integrating over the distribution of ε_{ni} yields the choice probability:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \tag{1}$$

Estimation

Typically done by MLE. Let y_{ni} be an indicator for whether agent n made choice i. The likelihood function is:

$$L(\beta) = \prod_n \prod_i \left(\frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \right)^{y_{ni}}$$

Taking logs and rearranging yields the log likelihood function:

$$LL(\beta) = \sum_n \sum_i y_{ni} V_{ni} - \sum_n \sum_i y_{ni} \ln \left(\sum_j e^{V_{nj}} \right) \tag{2}$$

Can then solve for $\hat{\beta}$ by maximizing (2). To evaluate "goodness of fit," can look at:

- In-sample prediction accuracy (what about out of sample/test set?)
- Compare value of log-likelihood function at $LL(\hat{\beta})$ to a "no-information" model where β is set to 0, $LL(0)$. Then construct a likelihood ratio: $LR \equiv \frac{LL(0) - LL(\hat{\beta})}{LL(0)}$

Welfare analysis

Willingness to pay. With quasi-linear utility, it is straightforward to get willingness to pay estimates for a specific attribute. An agent's WTP for an attribute x_i with coefficient β_i is the ratio of β_i to the coefficient on price. To see this, note that an agent's utility from choosing j is given by $u_j = \sum_i \beta_i x_i - \alpha p_j$. Taking the derivative with respect to the attribute of interest x_i and setting equal to 0 gives us the price change that keeps the agent's utility constant given the change in x_i: $\frac{dp_j}{dx_k} = \frac{-\beta_i}{\alpha}$.

Consumer surplus. Consumer surplus is the utility, in dollar terms, an agent receives from their choice. With quasi-linear utility, consumer surplus is $CS_n = \frac{1}{\alpha} \max_j U_{nj}$, where α_n is n's marginal utility of income.[2] Because utility is a random variable (with EV1 errors), we have to work with expectations. In this setting, expected consumer surplus is:

$$E\left[CS_n\right] = \frac{1}{\alpha_n} \ln \left(\sum_j e^{V_{nj}} \right) + C. \tag{3}$$

[2]It's important to note that this formulation assumes away income effects. That means this model isn't appropriate for setting where income effects may be important.

Note that we can calculate the first term in (3), but not the second, which captures the (unidentified) level of utility. Since we are often interested in changes in consumer surplus (e.g., due to a policy), we can difference version (3) in different states of the world, and the C term will cancel out.

Limitations of the conditional logit

Conditional logit breaks down in conditions where there might be:

- Preference heterogeneity (random only)
  - Systematic taste variation can be accommodated by including variable(s) that drive the taste variation (e.g. firm size and energy consumption choice)
  - This breaks down if taste variation has *any* idiosyncratic/random component *and* that variation is correlated with individuals' characteristics
- Repeated choices
- IIA doesn't hold
  - One implication of this is that the probability of choosing one thing over another depends only on the attributes of the two goods being compared.
  - This can imply unrealistic substitution patterns:[3] when a new good is introduced, it will draw evenly from the market shares of all the other goods, regardless of how similar or different it might be from the other goods already on the market.[4]

These limitations motivate the mixed logit.

Applications of the conditional logit model

Application 1: Davis (2021)

- Research question: What are the determinants of home heating electrification in the US over the last 70 years? How much are households willing to pay for the status quo (where they have an option to choose) over an electrification mandate?
- Starts with linear probability model, regresses household electrification on covariates (fuel prices, climate, home size, HH income)
  - Fuel prices are the most important predictor of electrification
- Then moves on to a random utility model, to be estimated with a conditional logit:

$$u_{ij} = \alpha_{0j} + \alpha_1 \underbrace{x_{ij}}_{\text{Price}} + \alpha_{2j} \underbrace{z_i}_{\text{HH characteristics}} + \varepsilon_{ij}$$

Normalization: define gas as the baseline choice, so coefficients are interpreted relative

[3]Conditional logit imposes substitution patterns as a structural assumption, rather than learning them from the data.

[4] An example. Red bus, blue bus (Train, p. 46): A traveler picks between a blue bus and a car. The two have equal market share. Introducing a red bus which the traveler sees as identical to the blue bus will lead the market shares of the red bus, blue bus, and car to be split evenly. This is clearly unrealistic: we would think that the car's market share would be unchanged while the red bus and the blue bus would split the blue bus' previous market share evenly.

to gas. Furthermore, normalize the gas-specific coefficients to 0: $\alpha_{0g} = \alpha_{2g} = 0$. Then the indirect utility functions for each of the two choices (electricity, gas) are:

$$u_{ie} = \alpha_{0e} + \alpha_1 x_{ie} + \alpha_{2e} z_i + \varepsilon_{ie}$$
$$u_{ig} = \alpha_1 x_{ig}$$

- Welfare analysis: want to estimate households' willingness to pay to avoid an electrification mandate.[a] Applying the formula in (3), WTP is:

$$WTP_i = \frac{1}{|\alpha_1|}\left[\ln(e^{\alpha_{0e}+\alpha_1 x_{ie}+\alpha_{2e}z_i} + e^{\alpha_1 x_{ig}}) - \ln(e^{\alpha_{0e}+\alpha_1 x_{ie}+\alpha_{2e}z_i})\right]$$

- Results:
  - Electric heating has increased dramatically over last 70 years
  - Energy prices explain > 70% of this increase
  - HH in warmer places are close to indifferent between electricity and gas, but HHs in colder places vastly prefer natural gas, with WTP up to \$2000 per year.

[a] Note that the language used in the paper is WTP, but this matches the consumer surplus math shown above. We are thinking about a ban here as opposed to a marginal change in an attribute, so the relevant welfare metric is consumer surplus across states of the world.

Application 2: Burgess et al. (2020)

Research question: What does demand for electrification look like in places at the "global electrification frontier"? What does demand for specific sources (grid, off-grid) look like?
Methods: Nested logit model combined with experimental variation

- Experimental variation comes from a field experiment in which households in Bihar, India were randomly offered different tariffs (subsidized price, normal price, no offer) for solar microgrids. This allows them to calculate well-identified demand elasticities for solar microgrids. However, this is not enough for other policy-relevant questions. For example, this does not tell us WTP for electricity, because if the price of solar microgrids changes, agents could switch to a different offering.
- Adding a discrete choice model (conditional logit) to the experimental estimates allows for welfare analysis, specifically WTP for electricity access
  - CL may be too restrictive, however. Model is estimated with panel data, but CL does not allow for state dependence.
  - The experimental variation provides an instrument for price. However, it relies on the assumption that microgrid price is perceived in the same way that any other electricity source's price is perceived, which may not be a good assumption.
  - They can also back out measures of unobserved quality by subtracting the calibrated mean utilities of consuming option j in village v in survey wave t, δ_{jtv}, from the fitted mean utility from observed characteristics $x'_{jtv}\hat{\bar{\beta}}$. In other words, this is the residual.
- With the calibrated demand model, they can calculate changes in surplus from changing the price of a source

Results

- Households value electricity; demand for any one source is highly elastic
- Richer households greatly prefer grid electricity
- Future growth in electrification will mainly be due to new grid connections

1.5.3 Mixed logit

The *mixed logit* generalizes the conditional logit, allowing for random taste variation across people and correlation in unobserved attributes over choices and time.

Setup and choice probabilities

The setup is similar to that of the conditional logit, with a crucial distinction:

$$U_{ni} = \beta_n X_{ni} + \varepsilon_{ni} \text{ where } \varepsilon_{ni} \sim EV(1) \text{ and } \beta_n \sim f(\beta|\theta) \tag{4}$$

The difference is that each individual has their own set of taste parameters β_n. We assume β_n takes distribution $f(\beta|\theta)$.[5]

The choice probabilities are obtained with similar steps as in the conditional logit, but now we also need to integrate over the distribution of β_n:

$$P_{ni} = \int \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} dF(\beta|\theta)$$

Estimation

Done by simulated MLE. For a particular value of θ, simulate draws from $f(\cdot)$ many times to calculate an average of simulated choice probabilities $\check{P}$. Then, plug these into the log-likelihood function from the conditional logit (equation 2), which we now call the simulated log-likelihood (SLL). Maximize over θ to get the simulated MLE estimate $\hat{\theta}$.

Mixed logit with panel data

It's pretty straightforward to extend the mixed logit model in (4) for a panel setting:

$$U_{nti} = \beta_n X_{nti} + \varepsilon_{nti},$$

with the same distributional assumptions as in (4). Notice that now the EV(1) errors are IID across individuals, choices, and time, but the idiosyncratic taste parameters β_n are fixed across time for each individual. In this setup, the probability of observing a sequence of decisions by a single individual given a value of β is:

$$L_{ni} = \prod_t \frac{e^{\beta_n' X_{nti}}}{\sum_j e^{\beta_n' X_{ntj}}}$$

[5] The particular choice of distribution depends on the setting and is chosen by the researcher.

To get the unconditional choice probabilities, integrate over the distribution of β:

$$P_{ni} = \int L_{ni} dF(\beta)$$

Estimating individual-specific parameters

[Leaving this out for now. See slide 70- of Discrete Choice Lecture 2.]

1.5.4 Dealing with endogeneity

In cases where we are interested in estimating in a parameter such as the marginal utility of income (recall this is the coefficient on a choice's price in a CL framework with quasi-linear utility), we might be concerned that price is correlated with unobserved attributes of the choice. To deal with this, we need methods that will allow us to isolate exogenous price variation.

The Berry transformation

In a CL framework, the market shares of each option are just the choice probabilities for that option. Adding an "outside option" with utility normalized to 0, we can write the market shares as follows:

$$\begin{aligned} s_0 &= \frac{1}{1 + \sum_{j \neq 1} e^{V_j}} \\ s_i &= \frac{e^{V_i}}{1 + \sum_{j \neq 1} e^{V_j}} \text{ for } i \neq 1 \end{aligned}$$

Taking logs yields:

$$\begin{aligned} \log s_0 &= -\log \left(1 + \sum_{j \neq 1} e^{V_j} \right) \\ \log s_i &= V_i - \log \left(1 + \sum_{j \neq 1} e^{V_j} \right) \text{ for } i \neq 1 \end{aligned}$$

Rearranging yields:

$$\delta_i \equiv \log s_i - \log s_0 = V_i = X_i' \beta + \alpha p_i + \xi_i \tag{5}$$

The key idea is that we have reformulated the market shares into an equation that is linear in the parameters and thus can be estimated via OLS. Most importantly, we can then apply standard IV tools to instrument for endogenous regressors in (5).

1.5.5 Discrete choice with aggregated data: BLP

Basic setup

Starting-point assumptions:

- Firms sell directly to consumers
- No price discrimination
- Perfect information (about prices and attributes)
- No dynamics
- All non-price attributes are exogenous (only price is endogenous)[6]
- Each consumer purchases one item
- Markets are defined such that all consumers within a market face the same products and attributes

Random utility model:

$$U_{njm} = V(p_{jm}, x_{jm}, \underbrace{d_n}_{\substack{\text{HH}\\\text{covars}}}, \beta_n) + \underbrace{\xi_{jm}}_{\substack{\text{Avg.}\\\text{utility}\\\text{of prod. } j\\\text{in mkt } m}} + \underbrace{\varepsilon_{njm}}_{EV(1)}$$

BLP assume utility is Cobb-Douglas and work with the utility function in logs:

$$U_{njm} = \alpha \ln(y_n - p_{jm}) + \sum_k x_{jmk}\beta_{nk} + \xi_{jt} + \varepsilon_{njm}$$

To add preference heterogeneity,[7] the setup is similar to a mixed logit:

$$\begin{aligned}
\alpha_n &= \alpha + \eta_{n\alpha} \\
\beta_{nk} &= \beta_k + \sum_r \mu_{kr} d_{nr} + \eta_{nk} \\
&\text{where } \eta_n | d_n \sim N(0, \Sigma) \text{ and } d_n \sim f(d_n)
\end{aligned}$$

Putting this together, we have:

$$\begin{aligned}
U_{njm} &= \alpha p_{jm} + X_{jm}\beta + \xi_{jm} + \left(\sum_r \mu_{kr} d_{nr} + \eta_{nk}\right)' X_{jt} \ldots && (6) \\
&= \underbrace{\delta_{jm}(X_{jm}, p_{jm}, \xi_{jm}; \theta_1)}_{\substack{\text{mean utility for all consumers}\\\text{in market } m}} + \underbrace{\nu_{njm}(X_{jm}, p_{jm}, \eta_n, d_{nr}; \theta_2)}_{\text{Individual-specific deviatons from } \delta_{jm}} + \underbrace{\varepsilon_{njm}}_{EV1} && (7)
\end{aligned}$$

Confused about the above. The parameters are broken down into a set of linear parameters $\theta_1 \equiv \{\alpha, \beta\}$ and a set of nonlinear parameters $\theta_2 \equiv \{\mu, \Sigma\}$.

Given distributional assumptions about η_i, can compute the unconditional choice probabilities:

$$P_{nim} = \int \frac{e^{\delta_{im} + \nu_{nim}}}{1 + \sum_j e^{\delta_{jm} + \nu_{njm}}} dF(\eta_n | \theta_2)$$

[6] I think this means exogenous to unobserved quality. Is this plausible?

[7] Preference heterogeneity is important in differentiated product markets. If there was no preference heterogeneity, why would there be differentiated products?

Estimation using market-level data

Using similar steps/logic to the Berry transformation, can write market shares as a function of mean utilities and non-linear parameter estimates θ_2, prices, non-price attributes and consumer demographics.

Dealing with endogeneity

We are back to the issue that unobserved characteristics may be correlated with price and observed characteristics. Recall that $\delta_{jt} = X_{jm}\beta + \alpha p_{jm} + \xi_{jt}$. Thus, when including product-specific constants in (7), ε_{njm} no longer is contaminated by unobserved product characteristics. To get unbiased estimates of α, we can use IV.[8]

A problem here is that estimating all the δ_{jt}s is complicated to solve numerically. BLP come up with a *contraction mapping* allowing for faster convergence in estimating the δ_{jt}s.

Pairing the above analysis with a model of the supply side motivates additional instruments ("BLP instruments"):

- Own-product non-price attributes
- Sum of non-price attributes for other products offered by the same firm
- Sum of non-price attributes for products offered by other firms

Applications of BLP

Application 1: Grigolon et al. (2018)

Research question: To what extent do car buyers undervalue future fuel costs, and what does this imply for policy (specifically, fuel taxes vs. product taxes)?
Fuel taxes vs. product taxes: Ex ante, fuel taxes will impact intensive margin directly and may also impact vehicle choice. Fuel taxes will be an efficient way of reducing GHG emissions if driving behavior is elastic and if consumers are forward-looking, i.e. they account for future fuel costs when they buy a new car. Product taxes could be more efficient if investment inefficiency is a result of consumer myopia.
Methods: BLP. Estimate a model of vehicle choice across models j and engine variants k (diesel or gasoline). Fit to data from 7 EU countries 1998-2011.
Model: T markets (country/year). Each consumer chooses a car, or the outside option. Slight twist on BLP: add discounted expected fuel costs as a non-price attribute, with an attention parameter γ. This is important in this setting because it allows for consumer-specific heterogeneity in mileage. The model:

$$u_{ijk} = x_{jk}\beta_i^x - \alpha_i(p_{jk} + \underbrace{\gamma\rho}_{\substack{\text{Attn}\\ \text{param}\\ \text{+discounting}}} \underbrace{\beta_i^m e_{jk} g_k}_{\substack{\text{Expected}\\ \text{annual}\\ \text{driving}\\ \text{cost}}}) + \xi_{jk} + \varepsilon_{ijk}$$

Note that $\alpha_i\gamma\rho$ are only jointly identified. Given an assumption on the interest rate and the time horizon/life of the car, can back out ρ.

[8] The actual instruments used are often sketchy: things like prices of other goods or in other markets.

Estimation: Following BLP, the predicted market shares are:

$$s_{jk} = \int_{\beta} \frac{\exp\left(x_{jk}\beta - \alpha(p_{jk} + \gamma\rho\beta^m e_{jk} g_k) + \xi_{jk}\right)}{1 + \sum_{j'}\sum_{k'} \exp\left(x_{j'k'}\beta^x - \alpha(p_{j'k'} + \gamma\rho\beta^m e_{j'k'} g_{k'}) + \xi_{j'k'}\right)} dF_{\beta}(\beta;\theta)$$

Authors simulate parameters that minimize distance between observed and predicted shares. Assumed distributions of parameters are $\beta_i = (\beta_i^x, \alpha_i, \beta_i^m)$ such that:

- β_i^m follows the observed empirical distribution of miles driven
- β_i^x is normally distributed
- α_i is constant across individuals and inversely proportional to income y_t in market t.

Use BLP instruments to deal with price endogeneity
Results: In preferred model, $\hat{\gamma} = 0.91$, a small amount of undervaluation. Under policy counterfactuals, find that fuel tax is more effective at reducing total fuel usage *when mileage heterogeneity is accounted for.* As a result, the fuel tax dominates a product tax (eg a tax on new cars): a lower tax can achieve the same level of emissions.

Application 2: Ito and Zhang (n.d.)

Research question: What is Chinese households' WTP for clean air?
Contribution: First WTP estimates coming from air purifier purchases
Methods: Combine BLP-style demand estimation with RD along Huai River boundary generating exogenous variation in pollution. Additionally instrument for price with distance to manufacturing plant.
Data: Scanner-level data on air purifier purchases across cities (aggregate so that markets are cities)
Model: Standard BLP, where the main attribute of interest is the reduction in air pollution from buying a filter j in city c: $x_{jc} = x_c e_j$, where x_c is the city's pollution level and e_j is the filter's effectiveness.

$$u_{ijc} = \beta_i x_{jc} + \alpha_i p_{jc} + \underbrace{\eta_j + \lambda_c}_{\text{FEs}} + \xi_{jc} + \varepsilon_{ijc}$$

Note that while they write this like a mixed logit, they will also implement a conditional logit (i.e. $\beta_i = \beta$ and $\alpha_i = \alpha\ \forall i$).
Estimation: Because they are using a conditional logit, the Berry transformation allows them to get linear equations for the market shares:

$$\ln s_{jc} = \beta x_{jc} + \alpha p_{jc} + \eta_j + \lambda_c + \xi_{jc}$$

Note that this differs slightly from equation (5) because the market share of the outside option $\ln s_{0c}$ is absorbed by the city fixed effect λ_c. Here, β is the marginal utility from an improvement in air quality and α is the marginal utility of income. Thus, $-\frac{\beta}{\alpha}$ is MWTP for a one-unit reduction in air pollution. Authors also implement an analogous mixed logit specification.
Results: Mixed logit results imply that median MWTP $-\frac{\hat{\beta}}{\hat{\alpha}}$ is \$1.19 per $\mu g m^{-3}$. Note that results are a lower bound: other avoidance behaviors (other than buying an air purified) are

possible, people may have limited understanding of pollution levels and their health effects, and indoor air pollution may already be lower than outdoor air pollution

Application 3: Depro et al. (2015)

Identification problem in traditional models of residential mobility: Starting point of this paper is the "ecological fallacy" (you can't draw conclusions about individual behavior from group behavior). The authors argue that EJ studies that regress changes in populations on changes in amenities/pollution exposure to try to say whether the dominant force is "white flight" or "coming to the nuisance" are not identified: the differences in populations do not identify the transition matrix describing moving probabilities.
Methods: Spatial version of BLP model where households make location choices based on neighborhood characteristics and individual preferences. Calibrate the model to observed data.
Model:

$$U_{ik} = \underbrace{f(X_k, \xi_k; \beta)}_{\equiv \delta_k} + \underbrace{\eta_{ik}}_{EV1}$$

Interpret δ_k as the mean utility of living in location k. Utility for individual i of moving from j to k:

$$U_{ij} - U_{ik} = (\delta_j - \delta_k) - \mu \underbrace{MC_{jk}}_{\substack{\text{Moving} \\ \text{costs}}} + (\eta_{ij} - \eta_{ik})$$

This plugs right into the familiar logit probabilities:

$$s_{jk} = \frac{e^{\delta_j - \delta_k - \mu MC_{jk}}}{\sum_l e^{\delta_l - \delta_k - \mu MC_{lk}}}$$

Where s_{jk} is the share of movers from j to k.
Estimation: Estimation is done separately for each race. Given a guess of μ, use a contraction mapping to solve for the vector of δs that best matches observed neighborhood shares. Then search over μs to match the percent of stayers. Given race-specific δs, divide by the marginal utility of income ($\hat{\mu}$) for each group and compare across groups. Finally, regress fitted values $\hat{\delta_k} \equiv \frac{\delta_k}{\mu_k}$ on fixed effects and census tract characteristics. Interpret the coefficients on the characteristics as the difference in WTP across the groups.
Results: Find that Hispanics are less WTP to avoid cancer risk. This is more consistent with a residential mobility explanation of the exposure/race correlation.

1.5.6 Things to remember: how logit differs from OLS

Logit estimates are relative

- Only relative differences in utility matter (implication: cannot estimate coefficients on characteristics that do not vary across choices, such as individual-specific predetermined characteristics like gender)

- If you are interested in differences across choices, you need to normalize. For example, in the following:

$$U^*_{1n} = \delta_1 + \beta'(X_{1n}) + \varepsilon_{1n}$$
$$U^*_{2n} = \delta_2 + \beta'(X_{2n}) + \varepsilon_{2n},$$

  the δs are not identified. To estimate δ_2, normalize:

$$U^*_{1n} = \beta'(X_{1n}) + \varepsilon_{1n}$$
$$U^*_{2n} = \delta_2 + \beta'(X_{2n}) + \varepsilon_{2n},$$

  Here, δ_2 captures the *average effect of unobserved attributes of choice 2 relative to the unobserved attributes of the omitted choice.*

- Discrete choice models are identified only up to scale
  - To interpret parameters, you must *normalize* the model. A simple example where errors are iid (from Train, p.24):

    Start with the model

$$U^0_{nj} = x'_{nj}\beta + \varepsilon^0_{nj} \text{ where } V\left[\varepsilon^0_{nj}\right] = \sigma^2$$

    Divide through by σ and define ε^1_{nj} where $V\left[\varepsilon^1_{nj}\right] = 1$. The new model is:

$$U'_{nj} = x'_{nj}\frac{\beta}{\sigma} + \varepsilon^1_{nj}$$

    Upon estimation, the coefficients $\frac{\hat{\beta}}{\sigma}$ are to be interpreted as the effect of the observed variable *relative* to the standard deviation of the unobserved variables σ.

    * In a standard logit model, error variances are typically normalized to $\frac{\pi^2}{6}$, which is the variance of the standard Gumbel distribution.
  - The choice of normalization doesn't matter, but it does matter for interpreting the *magnitudes* of the coefficients, especially across models (e.g. the same model estimated on different data or different models estimated on the same data)
  - A standard interpretation of logit coefficients: *a large coefficient on an attribute x implies that x is important in explaining variation in choices probabilities relative to unobserved factors.*

- Because discrete choice model coefficients are hard to interpret on their own, two other ways of interpreting: **marginal effects** and **elasticities**:
  - **Marginal effects** tell us the change in the choice probabilities given a marginal change in some attribute z_{ni}. Taking the derivative of (1) with respect to z_{ni} yields:[9]

$$\frac{\partial P_{ni}}{\partial z_{ni}} = \frac{\partial V_{ni}}{\partial z_{ni}} P_{ni}(1 - P_{ni}) \tag{8}$$

[9] See Train p. 58 for the derivation. It's a straightforward application of the chain rule and product rules.

If utility is linear in z_{ni} with coefficient β, (8) becomes something we can estimate from the data:

$$\frac{\partial P_{ni}}{\partial z_{ni}} = \beta P_{ni}(1 - P_{ni})$$

Note that this is maximized at $P_{ni} = 0.5$. That is, the marginal effect of changing an attribute is largest when the choice is most uncertain.

Using analogous machinery, we can estimate the marginal effect of a change in another choice's attribute z_{nj}:[10]

$$\frac{\partial P_{ni}}{\partial z_{nj}} = -\beta P_{ni} P_{nj}$$

– **Elasticities** are just the marginal effects normalized so we can interpret them in terms of percent changes:[11]

$$\begin{aligned} E_{iz_{ni}} &= \frac{\partial P_{ni}}{\partial z_{ni}} \frac{z_{ni}}{P_{ni}} \\ &= -\beta(1 - P_{ni}) z_{ni} \\ E_{iz_{nj}} &= -\beta P_{nj} z_{nj} \end{aligned}$$

Interaction terms are not easy to interpret

Interpreting the coefficients on interaction terms is not possible in the way it is in OLS: because they vary across values of all covariates, interactions' importance might vary significantly across units in the data. One way to address this is to look at a scatterplot of the interaction effect across the distribution of probabilities predicted by the logit. This gives a sense of the underlying heterogeneity not captured by looking at the interaction effect alone.

Comparisons across sub-groups are fraught

Coefficients on covariates in a discrete choice model are confounded by unobserved heterogeneity in the error term. This is not a big deal if errors are truly homoskedastic. If they are heteroskedastic, we cannot interpret differences in coefficient values across sub-groups cleanly: these may reflect differences in residual variation across groups.

One possible fix: look at group-specific ratios of coefficients (for these, the group-specific scale parameters cancel out). This has the possible pitfall that if the support of the distributions of the parameters being compared is 0, the variance of the ratio is not well-defined.

[10] See the bottom of Train p. 58 for this one.

[11] That is, the elasticity of y with respect to x is the percent change in y caused by a one percent change in x.

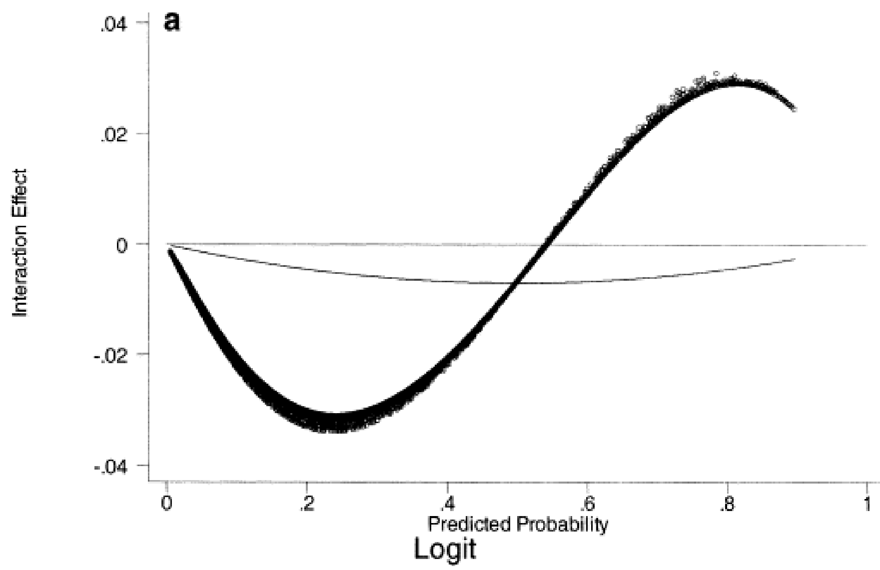


Figure 1: From Ai and Norton (2003). Shows coefficient on interaction between age and number of mobility limitations on probability of choosing HMO across the sample.

2 Joe

2.1 Joe's advice

I think you can look at exams from previous years to see what they are like, which may be a helpful guide. While doing all this, I would encourage you to write down ideas for your own research, so that after the field exam you have a list of ideas to start from. I think a field exam is useful if it encourages you to think broadly about the field, see connections between topics, get a better grasp of key conceptual ideas, but not to encourage rote memorization. Beyond that unfortunately I don't have much to advise. Good luck!

2.2 Questions on past exams

Likely to be hedonics and/or other revealed preference valuation methods

- 2021: Hedonics and revealed preference valuation (via health effects) for value of drinking water quality
- 2020: Hedonic model for value of timber
- 2019: Production functions and regulations, plus "describe a paper on this topic." [More similar to Reed's past questions]

2.3 Demand for environmental goods

2.3.1 Pollution

2.3.2 Climate

2.3.3 Water

Keiser and Shapiro (2019)

Setting: This is a very general paper assessing the impact of the 1972 Clean Water Act and the demand for clean water.
Methods: Use a triple difference to estimate the effect of CWA grants on pollution. Difference between pre- and post, treated and untreated, upstream and downstream. Use a DiD to look at the effect of CWA on municipal spending and housing prices. Specification for housing prices:

$$V_{pt} = \gamma G_{pt} + X_{pt}\beta + \eta_p + \eta_{wt} + \varepsilon_{pt}$$

where V_{pt} is log mean value of homes within a certain radius (0.25, 1, 25 mi) of the portion of the river that is 25 miles downstream from plant p.
Three main findings:

1. Water pollution has fallen substantially
2. CWA's grants to municipal wastewater treatment plants caused some of the declines (CWA funding did not crowd out municipal spending on clean water)
3. Grants' effects on housing values are smaller than their costs (overall ratio: 0.25)

Interpreting hedonic estimates: Hedonic estimates may be lower bound for WTP, because of:

- Incomplete information (people don't know the benefits of clean water)
- May exclude some types of demand, especially non-use value
- General equilibrium channels are ignored; in particular, if hedonic price schedule shifted, this could be a lower bound (Banzhaf, 2021).

2.3.4 Natural resources

2.3.5 Hedonics (Rosen, 1974)

Basic setup

n attributes, attribute vector $z \equiv (z_1, \ldots, z_n)$, and price function $p(z)$ that maps attributes to eqm prices. z assumed to be exhaustive, continuous, and accurately perceived by consumers. Consumers purchase a single good. Prices are "parametric"–i.e. there is no market power.

Demand side

Consumers' FOCs imply:

$$\frac{\partial p}{\partial z_i} = p_i = \frac{\frac{\partial U}{\partial z_i}}{\frac{\partial U}{\partial x}} \forall i$$

Given income y, define the bid function $\theta(z_1, \ldots, z_n; y, u)$ such that

$$U(y - \theta, z_1, \ldots, z_n) = u$$

The bid function is an indifference curve: holding utility constant, it reveals the maximum amount an individual would pay for an attribute z_j holding utility and the other attributes z_{-j} constant. Utility is maximized when the bid function is tangent to the hedonic price function: $\frac{\partial \theta(z^*, u^*, y)}{\partial z_i} = \frac{\partial p(z^*)}{\partial z_i}$.

Supply side

On the supply side, firms with heterogeneous costs sell housing services. Firms maximize profits subject under perfect competition. Inverting their profit function yields their offer curve $\phi(z^*; \pi)$. At the optimum, the offer function is tangent to the hedonic price function: $\frac{\partial p(z^*)}{\partial z_i} = \frac{\partial \phi(z^*; \pi)}{\partial z_i}$

The Hedonic Price Function/Schedule

The Hedonic Price Schedule (HPS) is the locus of tangencies between the bid curves and the offer curves. The slope of the HPS at any point is the marginal willingness to pay (MWTP) for that attribute for consumer(s) at that point. As a whole, the HPS reveals average MWTP in the population. However, the HPS is not useful for analyzing nonmarginal changes (see fig. 2) for an example: the HPS would suggest that consumer 1's WTP for a change in air quality from c_j^1 to c_j^2 is $p_2 - p_1$, when in reality it is $p_1^2 - p_1$. To understand WTP for non-marginal changes, need knowledge of both the HPS and the bid/offer curves.

Estimation

In Rosen (1974), estimation takes place in two steps:

1. Estimate the HPS by regressing prices on characteristics:

$$p = \alpha + \underbrace{f(z_1, \ldots, z_n)}_{\text{Nonlinear fn of attributes}} + \varepsilon$$

   Estimating the HPS is subject to the usual challenges of causal inference (particularly OVB). The last 20 years have seen a big advance in doing this successfully. A leading early example is **?**.

2. Given the estimated price function, regress the derivative of the price function with respect to the attribute of interest z_j on values of that attribute to trace out the bid/offer functions:[12]

$$\frac{\partial p}{\partial z_j} = \alpha' + g(z_j) + \varepsilon'$$

   Credible estimation of the second step is even harder than the first. Despite some attempts, there hasn't been much success on this front.

[12]To do this, must be able to observe the same consumer at different values of z_j–a tall order.

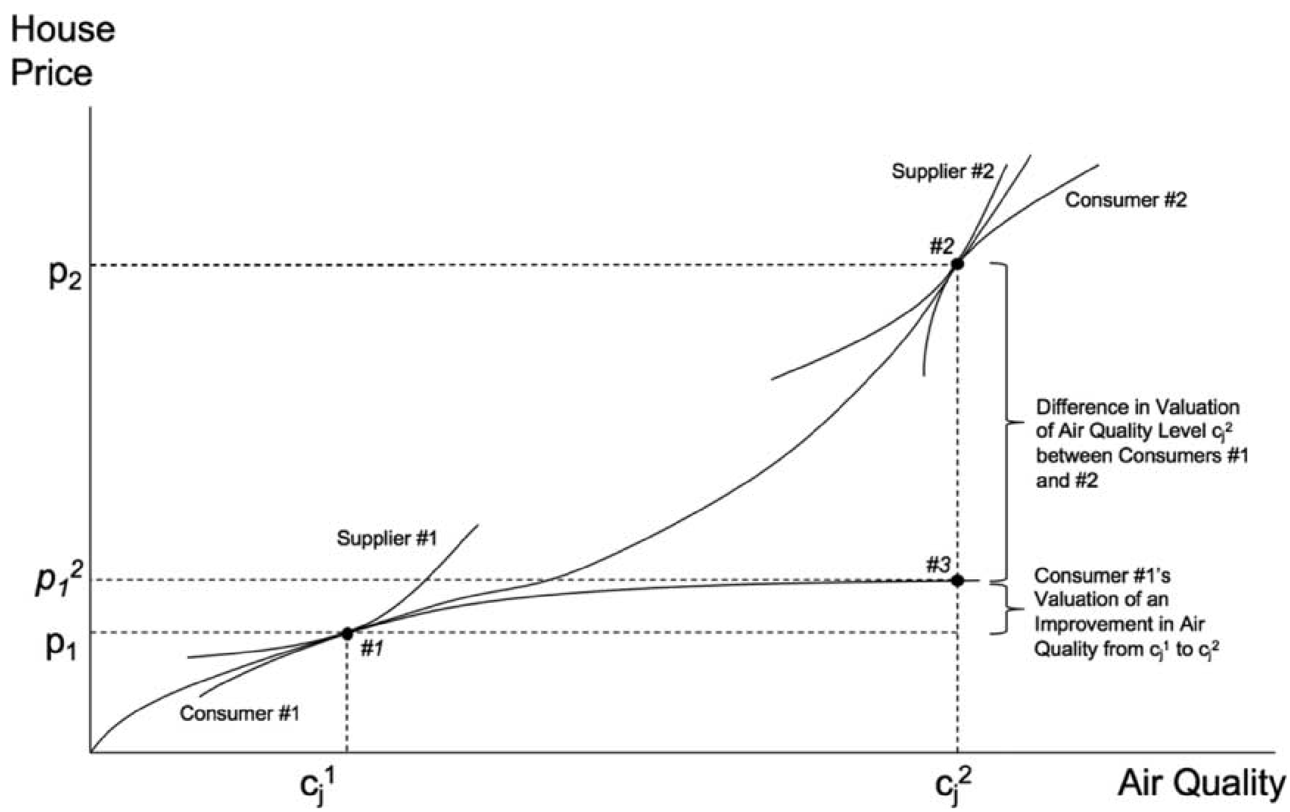


FIG. 1.—Bid curves, offer curves, and the equilibrium HPS in a hedonic market for air quality.

Figure 2: From Greenstone (n.d.)

Step (1) allows for the estimation of WTP for a marginal change in an attribute. With the second step, we are able to get a full description of the HPS *and* the bid and offer functions, allowing us to calculate WTP for non-marginal changes.

Four cases to keep in mind in the hedonic model:

1. If firms have identical costs but consumers differ, the HPS coincides with the offer function
2. If consumers have identical costs but firms differ, the HPS coincides with the bid function
3. If consumers are identical and firms are identical, the HPS collapses to a single point, and the market offers only one quality (no product differentiation)
4. If there is a distribution of buyers and sellers, the HPS is just the locus of tangencies and does not (fully) identify either the offer curves or the bid curves

Applications of hedonic models

Black (1999)

Methods: Uses a spatial regression discontinuity design across school "walk zones" to estimate WTP for school quality. This ID strategy solves the OVB issue that plagued this literature previously–school quality is correlated with all kinds of other local public goods (and other variables) that are hard to account for.
Results: homeowners are WTP for school quality, but the value is about half of what you would recover from a hedonic regression. This paper is primarily a well-identified measurement paper; little to say about what the optimal level of school spending is, does not attempt to estimate the second step in the hedonic method.
Extension: An extension of this work by Bayer et al. (2007) embeds the same identification strategy in a BLP-style discrete choice framework. Find that much of observed WTP is due to homophily.

Davis (2004)

Methods: DiD on housing prices across two counties in Nevada, one of which experienced a cluster of pediatric leukemia. Calculates WTP to avoid leukemia risk and value of a statistical case of pediatric leukemia.

2.3.6 Roback models

Basic setup

- Representative agent[13] consumes location-specific amenities $s \in \{S_1, S_2\}$ and a numeraire X.
- Capital and labor are fully mobile across cities
- No commuting costs within cities; infinite commuting cross across cities (you have to live and work in the same place)
- Perfectly competitive production with CRS technology

Consumer's problem is

$$\max U(x, l^C; s) \text{ s. t. } w + I = X + l^C r$$

where l^C is consumption of residential land, r is the rental rate, and I is non-labor income.

Key idea for consumers' problem: In equilibrium, wages and rents equalize utility across all occupied locations:

$$\underbrace{V(w, r; s)}_{\substack{\text{Indirect} \\ \text{utility}}} = \underbrace{k}_{\text{constant}} \tag{9}$$

[13] Extensions of this model might think about heterogeneous consumers, but the indifference condition in (9) still holds for a marginal consumer.

Production equilibrium implies

$$C(w, r; s) = 1. \tag{10}$$

Where does this come from? First, recall that firms produce only the composite X, which is numeraire so has price 1. Under perfect competition, price = marginal cost. Finally, under CRS, marginal cost = average cost.

Key idea for firms' problem: Firms earn the same profit in all locations.

Equilibrium

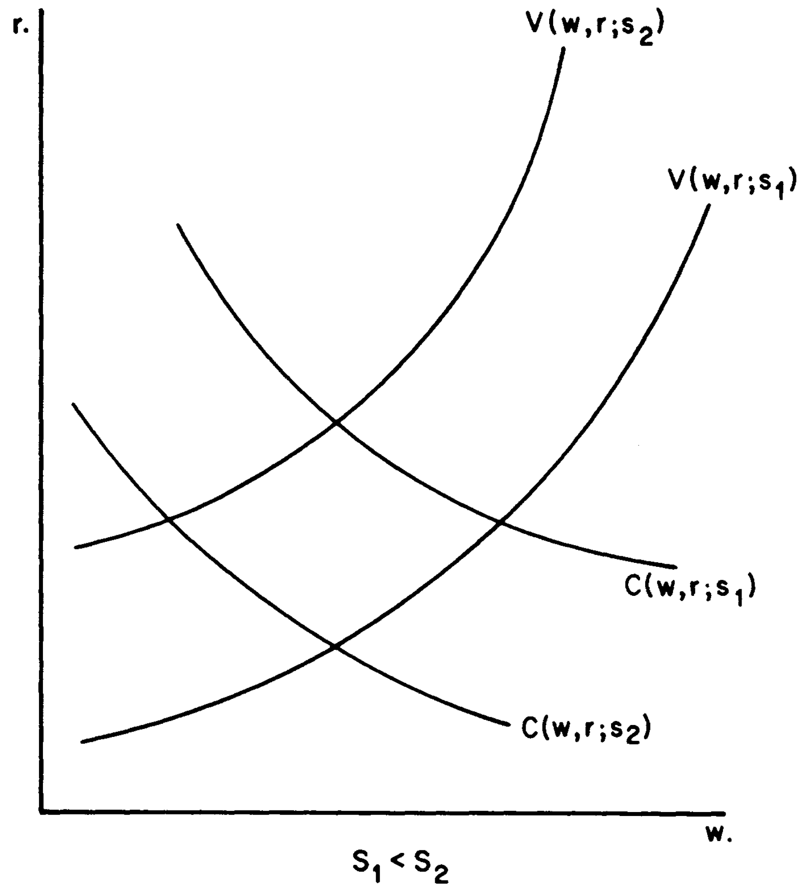


Figure 3: Figure 1 from Roback (1982).

Consider two places 1 and 2, where 2 has a higher level of the unproductive amenity than 1: $s_1 < s_2$. In equilibrium, 2 will have lower wages. Effect on rents is ambiguous (intuitively, both firms and consumers dislike high rents, so it's not clear which way it will go). To see, this, differentiate the indifference conditions (9) and (10) wrt s and solve for $\frac{dw}{ds}$ and $\frac{dr}{ds}$:

$$\frac{dw}{ds} = \frac{1}{\Delta}(-V_s C_r + C_s V_r) < 0$$
$$\frac{dr}{ds} = \frac{1}{\Delta}(-V_w C_s + V_s C_w) \lesseqgtr 0$$

where $\Delta = V_w C_r - V_r C_w > 0$.[14]

Applications of Roback models

Kline and Moretti (2014)

Research question: What is the long-run effect of the Tennessee Valley Authority (TVA) on economic development (specifically agricultural and manufacturing employment)?
Methods: Combine a DiD (controls are areas that were not selected as authorities) with a separate structural model (based on a Roback model) trying to get at spillovers, externalities
Model: Counties are SOEs, homogeneous workers with perfect mobility. Counties are heterogeneous in exogenous amenity and productivity and endogenous agglomeration. In the model, the TVA has two *local* impacts: it increases productivity directly (through infrastructure investments, etc.) and through increased agglomeration due to labor reallocation. The aggregate or *national* impact depends on the shape of the agglomeration forces. In particular, impact will be positive if it reallocates people from places with lower agglomeration forces to places with higher agglomeration forces.

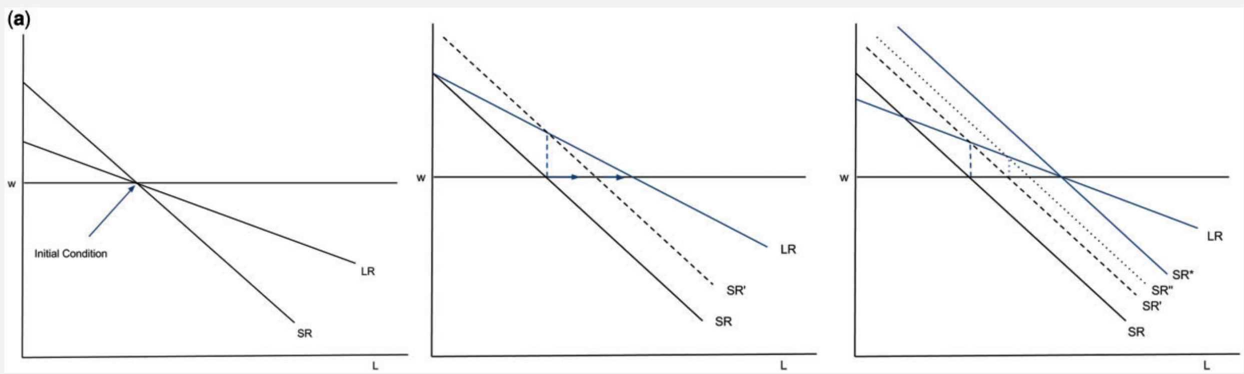


Figure IV

Panel A. Dynamics under Linear Agglomeration

In each panel, the horizontal axis is log manufacturing density and the vertical axis is the log manufacturing wage. SR and LR refer to short-run and long-run inverse demand curves, respectively (see Section IV.C of text). Panel A depicts convergence from initial condition to the new unique steady state under linear agglomeration after a permanent productivity shift. Panel B depicts effects of transitory productivity shift on steady state in the presence of nonlinear agglomeration effects.

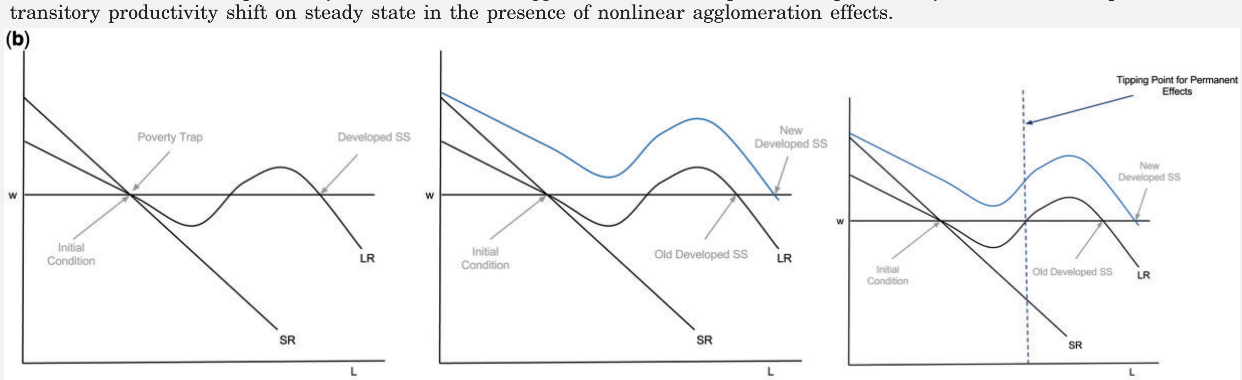


Figure IV

Panel B. Dynamics under Nonlinear Agglomeration

The authors approximate the (non-linear) agglomeration function $g(\cdot)$ using a spline in

[14] The relevant partial derivatives are: $V_w > 0, V_r < 0, C_r > 0, C_w > 0, V_s > 0, C_s > 0$.

manufacturing density. Instrument for contemporaneous manufacturing density with density 20 years before.
Results: DD results: Long-run gains in manufacturing employment; short-run gains in agricultural employment that are eventually reversed after the end of the subsidies.
Structural estimation results: agglomeration function is concave (increases in manufacturing density have a larger effect in low-density areas relative to high-density areas). Agglomeration elasticities are approximately constant.

2.4 Supply of environmental goods

2.4.1 Trade, leakage, GE

2.4.2 Spatial models

3 Jim

3.1 Jim's advice

Best practice is to ponder old field exams, then to recap the outlines/bullets from lectures. Reading papers is not a great use of time. My questions emphasize intuition and big theme ideas on policy design, and typically ask you to setup a math problem, but dont emphasize solving all the way through.

3.2 Questions on past exams

Likely to be a combination of extensions of homework problems and review of general results/big ideas from class on efficiency of different policy instruments, etc.

- 2021:
  - Optimistic separability
  - Externalities, second-best policies and related information requirements, comparison of policy instruments
  - Extension of homework problem with monopolistic competition, planner problem, optimal second-best tax
- 2019:
  - T/F questions about general class concepts (social cost of carbon, optimal level of tax)
  - Planner problem, comparing policy instruments, efficiency of different instruments
- 2018:
  - Comparison of policy instruments
  - Planner problem + decentralization of planner's solution
  - Extension of homework-style problem
- 2017:

3.3 Pigouvian taxes

The basics

- An *externality* is something that affects costs or utility directly, not through prices
- Pigouvian taxes *internalize* externalities by taxing them at their marginal damages at the optimum quantity
  - Jim says: "a lot of bodies are buried in the phrase 'at the optimum"'

  - This implies all other choices are made efficiently, e.g. population is distributed optimally, defensive investments are optimal

- Outcomes under Pigouvian taxes are pareto efficient if the tax is set at the right level.
  - Note that this does not depend on any elasticities.
  - The welfare gain from a Pigouvian tax does depend on elasticities.
  - The transition from the pre-tax to the post-tax allocations is not necessarily a Pareto improvement (i.e. some people are worse off), even though the resulting equilibrium is Pareto efficient

- The key appeal of Pigouvian taxes, despite their limitations (requirement for being set at optimal level and being the only market failure) is that they are *cost effective*[15] even if they are set at the wrong level.
  - Why? Under a price instrument, everyone has the same marginal incentive to reduce emissions across all relevant margins.
  - In math, this is the *equimarginal principal*: the marginal cost of abatement is equalized across all agents. Can check this from FOCs.
  - Alternative instruments can also be cost effective, but they require much more information (e.g. about cost and utility functions)

Solving social planner problems

Basic algorithm: set up the social planner's problem (maximize benefits net of costs) and get FOCs. Then set up the consumer/producer's problem, where prices include the tax (or whatever other policy instrument is under consideration). If you can get the consumer/produce FOCs to match the planner's, you've decentralized the optimal allocation.

Quantity instruments (cap and trade): an efficient alternative to taxes

- Initial permit allocation does not affect efficiency, only incidence

3.4 Instrument choice

Criteria for instrument choice (Goulder and Parry, 2008):

- efficiency
- cost effectiveness
- equity/distribution
- robustness to uncertainty
- political feasibility, flexibility, and robustness

[15] Formal definition of cost effectiveness: for a given level of abatement, that same level of abatement cannot be achieved at a lower cost.

- administrative costs
- enforcement

Weitzman P vs Q model

Basic idea: relative slopes of supply and demand determine whether a price or quantity instrument is more efficient when there is uncertainty about costs. See figure 4: quantity regulation is more efficient when MC (supply) curve is flatter, price instrument is more efficient with MC is steeper. Mathematical results: expected benefit Δ^O of choosing prices over quantities is:

$$\Delta^O = \frac{\sigma_\theta^2}{2c_2^2}(c_2 - b_2),$$

where c_2 is (abs value of) slope of MC, b_2 is (abs value of) slope of MB, σ_θ^2 is the variance of the shift parameter on MC (uncertainty about MC).

> **Pizer and Prest (2020)**
>
> Update to Weitzman in dynamic setting with policy updating. When firms can bank permits and the policy updates, quantity instruments can achieve first best. This arises from an intertemporal arbitrage condition that exists under bankable quantity instruments but not under taxes. The key driver is firms' expectations of future policy. The study was motivated by permit prices in the SO2 trading program reflecting expected future prices up to 6 years out after a change in the policy was announced.

The Coase Theorem

If property rights are clear and there are no transaction costs, bargaining will lead to an efficient allocation. This does not depend on the initial allocation of property rights.

3.5 Performance standards

The rebound effect (aka the Jevons paradox)

Raising efficiency lowers cost, leading to an increase in quantity demanded. This reduces the amount of mitigation.

What's wrong with performance standards?

Performance standards:

- Cause rebound
- Fail to correct market size (i.e., total consumption will still be too high)
  - Work on extensive margin, but not on intensive margin
- Interaction with the used durables market lead to leakage (Gruenspecht Effect)
  - Fails to create incentives to scrap existing products optimally

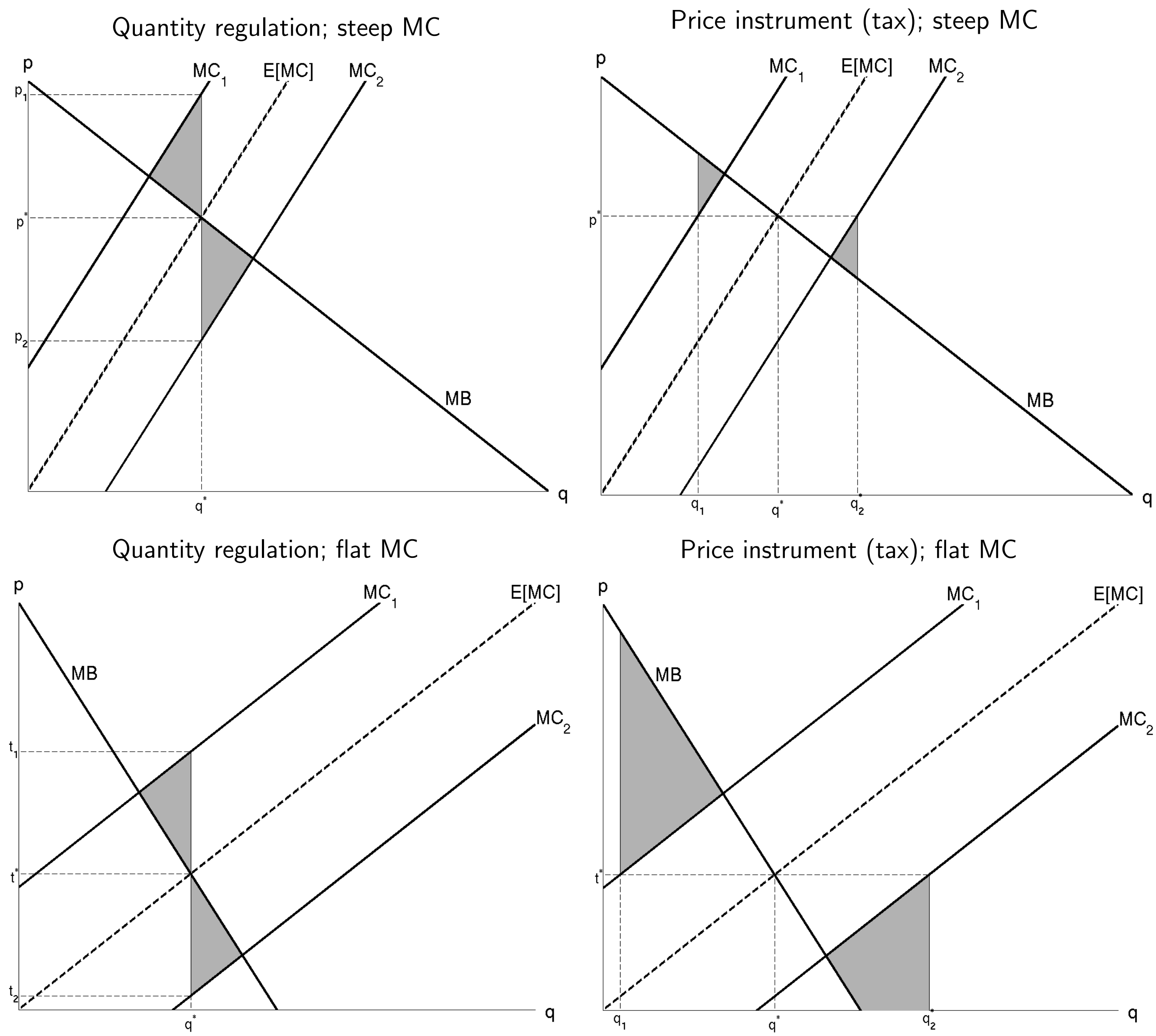


Figure 4: DWL from setting taxes and quantities away from optimum for different MB/MC curve slopes

- Example: CAFE raises prices for new trucks more than it does for new cars. This will make used trucks more expensive, assuming that used trucks are the best substitute for new trucks. This reduces scrap rates.

The math: Assume there are just two products, high emissions and low emissions, each with marginal emissions β_i. The planner's problem is:

$$\max_{q_H, q_L} U(q_H, q_L) - C_H(q_H) - C_L(q_L) - \tau(\beta_H q_H + \beta_L q_L)$$

The performance standard requires average emissions to be below a threshold σ. Then the standard is:

$$\frac{\beta_H q_H + \beta_L q_L}{q_H + q_L} \leq \sigma$$

The consumer's problem is:

$$\max_{q_L, q_H} U(q_H, q_L) - C_H(q_H) - C_L(q_L) + \lambda[\sigma(q_H + q_L) - \beta_H q_H - \beta_L q_L]$$

Notice that this matches the planner's problem, except for an additional term $\lambda\sigma(q_H + q_L)$. This acts as a *subsidy* on total output. Thus, the relative price of the high and low emissions goods is right, but the total size of the market is wrong.

Attribute-based standards

Attribute-based standards are performance standards targeting one characteristic and creating a sliding scale based on a secondary attribute

3.6 Incidence and equity

Recall that Pareto optimality and social optimality are not the same. A Pigouvian tax could lead to unfair outcomes.

Two views on equity and policy:

- *Optimistic separability*: equity and efficiency issues can be separated completely (tax externality, then redistribute).
    - One popular approach (proposed in Kaplow (2004)) is to implement Pigouvian taxes to deal with externalities, then adjust the income tax to achieve desired distribution
        * Note that this is much easier to implement with homogeneous preferences, where heterogeneity in benefits/costs of the tax are due only to differences in income, and not differences in preferences.
    - Note that if you adjust income tax schedule so that there is no change in individuals' labor supply choices, there is no distortion; the Pigouvian tax has raised revenue without changing anything else.[16] See figure 5 for an illustration.

[16] Basically, a free lunch.

Gasoline Tax Increase and Offsetting Income Tax Adjustment

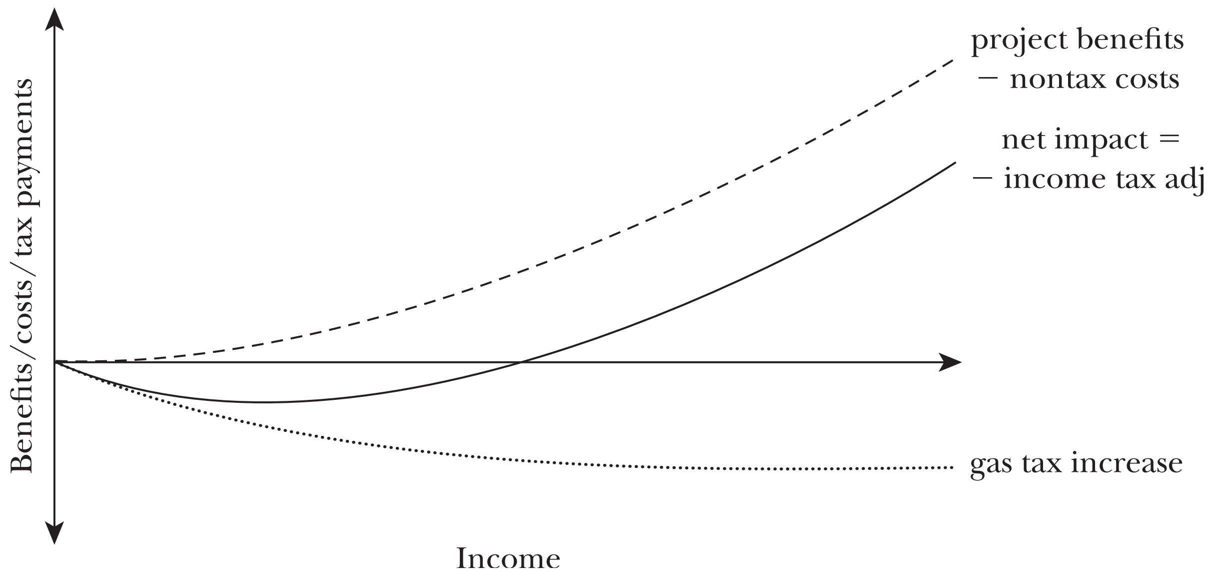


Figure 5: From Kaplow (2004).

- Optimistic separability starts to break down in realistic environments.
  - Sallee (2019) shows that you need pretty good information about the incidence of the tax to correctly implement the compensating transfers; in realistic settings this seems infeasible
  - Allcott et al. (2018) make the point that Atkinson-Stiglitz[17] only holds under full salience of all commodity taxes, which may not hold in real life.

- Conventional view: modify the Pigouvian prescription to account for equity concerns

Efficiency and social welfare functions

Kaldor-Hicks efficiency: a change in allocation is preferred if there exists a hypothetical set of transfers such that no one is worse off under the new allocation and someone is better off.

Four principles of incidence (come back to these and re-watch the supplemental video, time permitting)

1. Who cares who pays?
2. People pay taxes
3. Inelasticity is expensive
4. In general, anything can happen

[17] The Atkinson-Stiglitz theorem says that all redistribution should happen through the labor income tax. This stands in contrast to many policies (taxes on foods, healthcare and education subsidies) which implement redistribution through taxes on specific commodities.

Another: estimating incidence is challenging (e.g., expenditure as a function of income or as a function of total consumption, PIH).

3.7 Second best policy

Sometimes it's not possible to tax the externality directly. This puts us in "second-best" world.

3.7.1 Diamond model: Heterogeneous externalities, homogeneous tax

Diamond (1973) solves a model where the externality is heterogeneous across consumers but the policymaker can only set a single tax rate. The key result from the model is:

$$t^* = \sum_{j=1}^{n} \underbrace{\frac{x_j'}{\sum_{k=1}^{n} x_k'}}_{\substack{\text{Consumer } j\text{'s} \\ \text{demand derivative} \\ \text{relative to others}}} \underbrace{\sum_{i \neq j} \frac{\partial U_i}{\partial x_j}}_{\substack{\text{Optimal tax} \\ \text{on } j}}$$

The intuition here is that you tailor the tax to the consumers who will be most responsive to it–that is, you put higher weights on consumers whose demand is more price responsive.

Diamond also offers an algorithm for finding optimal second-best taxes (it's similar to the first-best version):

- Derive consumer optimality conditions conditional on the tax rate
- Write down the planner's problem, differentiate w.r.t the tax rate, allowing each endogenous choice (i.e. consumption of each good) to be a function of the tax rate
- Sub in consumer optimality conditions into planner FOC and solve

Jacobsen et al. (2020): Estimating inefficiency of second-best taxes

Key idea is that the R^2 of a regression of the externality on the attribute(s) being targeted is a sufficient statistic for the inefficiency of an optimal second-best policy based on that attribute.

3.7.2 Multi-part tariffs

When a Pigouvian tax is unavailable, you can sometimes achieve the first-best using a two-part tariff (recycling tax refunds are a good example):

1. Tax at sale at level of externality
2. Refund tax payment if object is disposed of in a way that does not cause an externality

3.7.3 Taxing all polluters with respect to the total (Segerson, 1988)

Proposed solution to the problem where the ambient level of an externality is observable, but each polluter's contribution to it is not. Segerson (1988) proposes the tax scheme:

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

where x is the ambient pollution level and $\bar{x}$ a constant (the target pollution level). Segerson proves that under a Cournot-Nash beliefs assumption[18] you can get the first best by choosing only k_i, only t_i, or both.

3.7.4 The additivity property

Based on Sandmo (1975). Adding general equilibrium and other distortions and seeing if the Pigouvian prescription holds. First, without any other distortions, the Pigouvian prescription goes through. Then, add a revenue requirement. With good $j \in \{0, \ldots, J\}$, x_0 is labor, so marginal utility of income is u_0. Good J is the polluting good (all others non-polluting), and damages depend on the aggregate of consumption of the dirty good X_J. Marginal damages are u_{j+1} The second-best optimal policy is:

$$\begin{aligned} t_j &= \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) \\ t_J &= \left(1 - \frac{-\mu}{\lambda}\right)\left(\frac{\sum_{k=1}^{J} x_k D_{jk}}{D}\right) + \underbrace{\frac{-\mu}{\lambda}\left(-n\frac{u_{j+1}}{u_0}\right)}_{\text{“adjustment term”}} \end{aligned}$$

where:

- D^* is the matrix of demand derivatives, D_{jk} is an entry, and D is its determinants
- λ is the shadow price on the planner's revenue constraint, μ is the marginal utility of income for the agent, and $\frac{\mu}{\lambda}$ is the inverse of the marginal cost of public funds (ratio of value of \$1 for the government vs. \$1 for consumers)
- $-n\frac{u_{j+1}}{u_0}$ is the marginal damage from the externality
- First term is the Ramsey taxation term: put higher taxes on less elastic goods

The basic idea (which generalizes beyond Sandmo's model) is you want to tax the externality separately from how you tax all other goods–the solution is additively separable between the tax on the externality and the tax on everything else

[18] Don't ask me what this assumption is...

3.7.5 Fiscal interactions

Revenue recycling and the double dividend

This is the idea that if you use the revenue from a Pigouvian tax to reduce distortionary taxes on other things in the economy, you can achieve a "double dividend" of both making the tax system more efficient and correcting the externality. .

Tax-interaction effect

This is a balancing point to the excitement of the double dividend. Any tax is an implicit tax on labor, because it lowers the real wage (because the cost of all things that use the externality-generating product of the input will go up). This additional tax compounds the existing distortion int he labor market caused by income taxes, adding a new efficiency cost. Note that this effect is larger for larger pre-existing distortions.

4 Reed

4.1 Questions on past exams

Likely to be a combination of specific questions about papers we (should have) read and something about production functions:

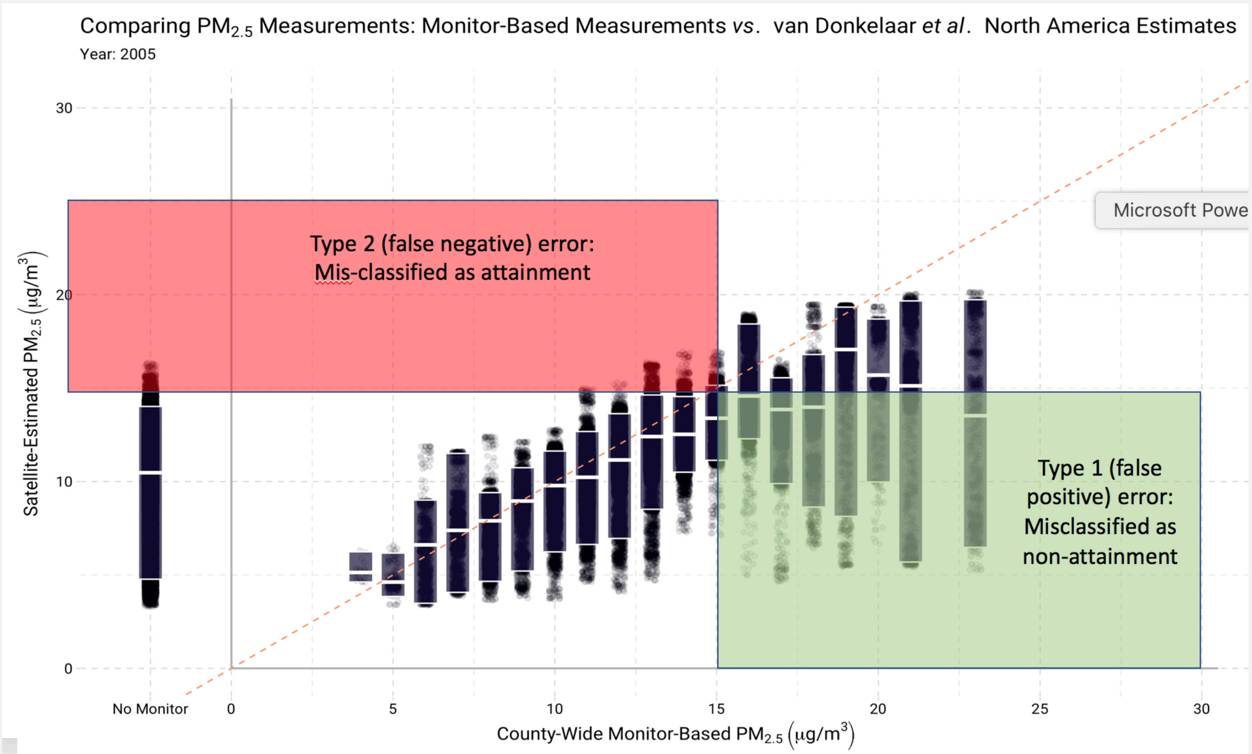
- 2018:
  - Comparison of Dell et al. (2012) and Burke et al. (2015)
  - Question about research design and welfare analysis in Greenstone and Gallagher (2008)
  - Simple cost function modeling problem
  - Question about duopoly model in Fowlie (2009) [not on 2022 syllabus]
  - Interpretation of key result from Ganapati et al. (2020)
- 2017:
  - Production function estimation question (detailed, including pros and cons of different empirical strategies)
  - "Describe a paper" question on welfare effects of regulation that uses a micro-founded structural model
  - Comparing structural and reduced form approaches;

4.2 Air pollution

4.2.1 Measurement

Fowlie et al. (2019): Bringing satellite-based air quality estimates down to Earth

Research question: To what extent does the limited network of EPA monitors lead to over/under detection of CAA non-attainment? What are health and distributional implications of this? What are implications of imprecision in satellite-based estimates for policy calibration?



Results: Type 1 errors (false positive) appear more common than type 2 errors (false negative)

Currie et al. (2020): What caused racial disparities in particulate exposure to fall?

Research question: Do people of different races have different exposure to particulate pollution? Existing evidence is "piecemeal" (because monitor networks are sparse) and "indirect" because researchers are often forced to use proxies for air pollution instead of direct measurements.

Methods/data: this paper solves both the "piecemeal and indirect" issue by combining satellite air pollution data with restricted-access individual-level census data. They estimate the explained and unexplained components of the racial exposure gap, then use an event studies and quantile regressions to estimate the causal effect of the CAA on the vintiles of the pollution distribution. Combining these with data on exposure data by race allows them to calculate the counterfactual distribution of particulate pollution without the CAA and estimate the CAA's contribution to the change in the gap.

Findings:

- The Black/white gap in pollution exposure has declined significantly in the past 20 years.
- Very little of the exposure gap can be accounted for by neighborhood characteristics
- Migration also explain only a small portion of the change in the gap (13%)
- About 60% of the reduction in the gap is attributable to differential impacts of the CAA.

4.2.2 5 main challenges to measuring the social cost of air pollution externalities

1. Causal inference is hard (recent papers making progress here include Chay and Greenstone (2003); Deryugina et al. (2019))
2. Multiple pollutants are correlated, making it difficult to attribute welfare effects to any single pollutant
3. Short run vs. long run exposure (Chen et al. (2013) makes progress here)
4. Dose response function may be both heterogeneous and non-linear
   - This makes things extra difficult because optimal policy needs information on marginal damages at optimal level of output. If damage function is non-linear and/or heterogeneous, it's harder to estimate the right thing
5. Monetization is hard
   - WTP may exceed expenditures
   - It's hard to capture all the relevant components that reflect WTP for something, including non-health expenditures and impacts (e.g. amenities, cognitive/productivity effects, etc.), behavioral change, etc.

Chay and Greenstone (2003): The Impact of Air Pollution on Infant Mortality

Quick summary: This paper is an early example of bringing tools from labor (FE and IV) to a question in environment/health. Use shocks to air pollution from a recession to identify the effect of air pollution on infant mortality.
Methods: FE: regress TSP levels on infant mortality levels with county and year FEs. IV: instrument first differences in TSPs and infant mortality using lagged TSPs as an instrument.
Results: Generally find that TSPs increase infant mortality, results are similar across FE and IV specs, larger magnitudes for IV. The methods are not up to modern empirical standards (concerns with both approaches, specifically OVB for FE, instrument weakness and exogeneity concerns)

4.3 Hedonics

Greenstone and Gallagher (2008): Does Hazardous Waste Matter?

Quick summary: Evaluate the effect of superfund site cleanup on home values, using a (fuzzy) RD comparing places that got a cleanup to places that were under consideration but didn't get cleaned up. Use hedonic model to map this to welfare.
Data: Census tract level data on home prices, migration, demographics. Only available every 10 years.
Adapting the hedonic method for a non-marginal change: One of the things this paper does is derive empirical tests that allow it to investigate what the standard hedonic

estimates could be missing since the change in environmental quality due to superfund site cleanup is non-marginal. The authors identify four effects of a non-marginal change:

- **Price:** Prices will increase, unless if housing supply is perfectly elastic (in which case price will stay the same).
- **Sorting:** The change in environmental quality will cause sorting, so in the new equilibrium people living in the cleaned-up area will have higher WTP for environmental quality tahn people who left.
- **Housing supply:** Will increase, in response to the increase in demand.
- **Gains will accrue to landowners (not renters).**

Follow-up by Gamper-Rabindran and Timmins (2011) : Using census tract data is a high level of aggregation. Find much higher level of capitalization using census block data, as well as higher population density and housing unit density, sorting (increased income and college educated), and increased minorities

4.4 The costs of environmental regulation

Effects of economic regulation (Joskow and Rose (n.d.)):

1. Prices: average price level, structure of prices
2. Static costs of production (direct regulatory costs, input prices, reduced competition)
3. Dynamic efficiency (rate and direction of innovation, productivity, dynamic aspects of market structure)
4. Product quality and variety
5. Distribution of income and rents (via profitability, rent-sharing with factors fo production)
6. Transition costs of capital and/or labor

Ryan (2012): "The Costs of Environmental Regulation in a Concentrated Industry"

Research question: How does regulation of a concentrated industry affect firm entry, investment, and market power? What are the implications of these effects for the welfare effects of the regulation?
Methods: Estimates a dynamic oligopoly model. Using a two-step estimation procedure, recovers the cost structure of the industry, including sunk entry costs and capacity adjustment costs.
Counterfactuals: Looks at counterfactuals of a new market and a fully established market. The idea is that a new market will be most susceptible to costs due to changing fixed costs, and thus represents an upper bound on welfare costs. Conversely, a mature/fully established market is least susceptible to these and thus represents a lower bound.
Findings: CAA Amendments (CAAA) doubled sunk entry costs, reducing net entry and increasing market power. CAAA led to higher investment by incumbents, but lower aggregate market capacity. Taken together, these decreased consumer welfare by 25%.

Cicala (2015): When does regulation distort costs?

Research question: How did fuel procurement practices by coal- and gas-fired power plants change after deregulation? (In this setting, deregulation refers to the end of cost-of-service regulation.)

Theories of regulatory distortion: This paper tests different theories of how regulation can distort costs proposed in the literature, include:

- Capital bias when rate of return exceeds the cost of capital
- Principal-agent problem in which regulator cannot observe cost (or cost of effort), and thus cannot induce cost-reducing effort
- Regulatory capture: interest groups (e.g. coal producers) influence the regulator's decision which costs to allow

Methods: DiD with matching to identify control group. Matching helps deal with confounds including heterogeneous transport costs and railroad regulatory environments, correlated time-varying shocks, and SUTVA (regulated and deregulated plants participate in the same coal markets)

Results: Deregulated plants reduce the price they paid for coal, employed less capital-intensive production techniques (procuring low-sulfur coal instead of installing scrubbers), and shifted toward more productive coal mines.

Cicala (2022): Imperfect Markets vs. Imperfect Regulation is US Electricity Generation

Research question: What are the welfare gains from moving to deregulated wholesale markets?

Methods: Staggered DiD

Results: Deregulation let to substantial net improvements in allocative efficiency, including due to reduced out-of-merit order operation and gains from trade coming from purchasing electricity that is lower on the merit order from another power control area (PCA).

Shapiro and Walker (2021): Is air pollution regulation too stringent?

Research question: How do the costs of air pollution regulation compare to their benefits?

Motivation: Regulation has substantially reduced air pollution in the last 50 years. If marginal abatement costs are convex, we could be reaching the point where MC > MB.

Methods: Use a previously unexplored feature of the clean air act that created decentralized markets for local pollutants. For nonattainment areas, new entrants must offset their emissions by paying incumbents to reduce their emissions. Thus, permit transaction prices should equal marginal abatement costs. Combine this with estimates of marginal benefits of abatement from air quality models to determine for each location whether the regulation is too stringent or not stringent enough. Plotting the ordered value of the offset prices against the cumulative abatement also gives and estimate of the marginal abatement cost curve.

Results: For most regions and pollutants, marginal benefits of pollution abatement exceed marginal cost 10-fold. In at least one market (Houston, VOCs), MC > MB. Furthermore, the abatement cost curve appears to be steeper than previously believed, so price instruments may be preferred over quantity instruments in these markets.

Walker (2013): Transitional costs of sectoral reallocation

Research question: How large are the costs of sectoral reallocation for workers who lose their jobs due to environmental regulation?
Methods: DDD estimator (pre/post, attainment/nonattainment, and sectoral polluter status). Identifying assumption: no time-varying county-level economic shock to polluting industries in the years after the policy.
Results: total foregone wage bill is about \$5.4 B. Mostly driven by nonemployment and lower earnings after changing jobs. Most switchers switch to a different industry in the same county. Overall, earnings patterns similar across different types of switchers (same/diff industry $\times$ same/diff county).
Interpretation: If workers are paid their marginal product, wage losses are a social cost. If workers make rents, wage losses overstate cost of reallocation. Also need to consider DWL from additional social assistance (unemployment insurance). Regardless, these costs are dwarfed by the estimated benefits of the CAAA (\$1.6 T)

4.5 Production Functions

4.5.1 General setup

Output function taking the form

$$Y_{it} = \underbrace{\exp(\omega_{it})}_{\substack{\text{Hicks-neutral}\\ \text{productivity}\\ \text{term}}} F(\underbrace{L_{it}, K_{it}, M_{it}, \ldots}_{\text{Factors of production}})$$

Can also think about a value added production function that is net of the inputs, $Y - M$. Common functional forms include:

- Cobb-Douglas
- Translog. 3-factor example:

$$\begin{aligned}\ln Y = & \ln A + a_L \ln L + a_k \ln K + a_M \ln M + b_{LL} \ln(L)^2 + b_{KK} \ln(K)^2 + b_{MM} \ln(M)^2 + \\ & b_{LK} \ln L \ln K + b_{LM} \ln L \ln M + b_{KM} \ln K \ln M\end{aligned}$$

- Leontief. 2-factor example: $Y = \min(aL, bK)$
- Perfect substitutes. 2-factor example: $Y = aL + bK$

4.5.2 Challenges to estimation

Two main issues

1. "Transmission bias"/simultaneity: factor allocations depend on unobserved productivity

- A simple Cobb-Douglas example:

$$\ln Y_t = \alpha \ln K_t + \beta \ln L_t + \underbrace{\ln A_t}_{\text{unobserved}}$$

- If input use and TFP are correlated, we have an omitted variable.
  - One solution: assume $A_{it} = \mu_i + \varepsilon_{it}$, where ε_{it} is uncorrelated with input choices. Then use firm fixed effects to absorb μ_i.
  - But, this leads to more problems:
    * FEs exacerbate measurement error, if factors of production do not have much variation and are measured with error, coefficients will be attenuated
    * Firm-specific productivity μ_i might be time-varying; this might be exactly what we are hoping to identify (e.g., the impact of some shock on productivity)
  - Other possible solutions include:
    * IV (but finding good instruments is hard)
    * Control functions (strong assumptions about the DGP for productivity)
    * Dynamic panel methods (weak instruments, poor instruments, strong identifying assumptions)
  - A final approach is to use "index" methods: pick a production function, measure output at the factors of production, and estimate the parameters.
    * Example: Cobb-Douglas production with CRS and no markups. Estimate the output elasticities and cost/revenue shares. Then

$$\ln A_t = \ln Y_t - \alpha \ln K_t - \beta \ln L_t$$

2. Productive firms are less likely to exit
   - This means you don't observe the full distribution of TFPs (the left tail will be truncated)

Other challenges include measurement error, misspecification of production function, multicollinearity of inputs

Greenstone et al. (2012): The effects of environmental regulation on the competitiveness of US manufacturing

Research question: What is the impact of environmental regulation on manufacturing plants' TFP levels?
Framework: Production is Cobb-Douglas: $Q = A\tilde{L}^{\alpha}\tilde{K}^{1-\alpha}$, where the tildes denote "production effective" inputs (actual inputs i scaled down by some factor $\lambda_i < 1$, reflecting the fact that compliance with regulation requires some inputs not used directly in production, such as emissions-abating capital or a regulation compliance officer). $TFP = \frac{Q}{L^{\alpha}K^{1-\alpha}} =$

$A\lambda_L^\alpha \lambda_K^{1-\alpha}$, so as compliance costs go up ($\uparrow \lambda_i$), TFP goes down.
Can regress the estimated TFPs on non-attainment status in a DiD framework to get the effect of regulation on TFP.

Empirical concerns and some solutions:

- Price effects: What if regulation affects the price?
  - Look at the concrete industry (homogenous product with one price which is observed)
  - Convert TFPR into TFPQ given estimated price changes[a]
- Entry and exit could be part of the story
  - Can look at these as an outcome

Results: Regulation on all pollutants except for carbon monoxide regulation reduces TFP. The annual lost output is \$11B.

[a] Recall the difference between TFRP and TFPQ. Let production be Cobb-Douglas in capital and labor.

$$TFPQ = \frac{AK^\alpha L^{1-\alpha}}{K^\alpha L^{1-\alpha}} = A$$
$$TFPR = \frac{p \cdot AK^\alpha L^{1-\alpha}}{(rK)^\alpha (wL)^{1-\alpha}} = \frac{p}{r^\alpha w^{1-\alpha}} A$$

We typically can only observe TFPR, but are interested in TFPQ.

4.6 Measuring markups

4.6.1 Production function approach

Given a general production function with Hicks-neutral productivity Ω, variable inputs V and dynamic inputs K, $Q = F(V, K)\Omega$, can estimate markups using the firm's FOC. The firm's Lagrangian is

$$\sum_{v=1}^{V} p^v V^v + rK + \lambda[Q - Q(V, K, \Omega)]$$

Rearrange FOCs to get:

$$\frac{\partial L}{\partial V^v} = p^v - \lambda \frac{\partial Q}{\partial V^v} = 0 \tag{11}$$
$$\underbrace{\frac{\partial Q}{\partial V^v} \frac{V^v}{Q}}_{\text{Output elasticity}} = \underbrace{\frac{P}{\lambda}}_{\text{markup}} \times \underbrace{\frac{P^v V^v}{PQ}}_{\text{Input } v\text{'s revenue share}} \tag{12}$$

Thus can estimate markups using an estimated output elasticity and data on revenue shares. As a bonus, can use price data to estimate marginal costs, because MC = price - markup.

4.6.2 Demand side approach

Given assumptions on a utility function and a market structure, you get a rule for markups. Then can estimate marginal costs based on the identity $\ln MC = \ln P - \ln MU$.

Ganapati et al. (2020): Energy cost pass-through in US manufacturing

Research question: In imperfectly competitive markets, how much of an externality-correcting tax is passed through to consumers? What does this imply for the incidence of the tax?

Background on incidence: *Incidence* is the ration of change in consumer surplus to the change in producer surplus caused by a marginal change in the tax rate: $I = \frac{dCS/dt}{dPS/dt}$. The pass-through rate is the change in the price consumer face with respect to a change in the tax rate: $\rho = \frac{dp}{dt}$. Under perfect competition, the change in CS is ρ, and the change in PS is the change in marginal costs γ minus the pass-through rate: $\gamma - \rho$. Thus incidence under perfect competition is $I^{\text{Competition}} = \frac{\rho}{\gamma - \rho}$. Under monopoly, incidence if $I^{\text{Monopoly}} = \frac{\rho}{\gamma}$. The authors derive a general formula for incidence under oligopoly: $I^{\text{Oligopoly}} = \frac{\rho}{1-(1-L \times \epsilon_D)\rho}$ where $L = \frac{p-MC}{p}$ is the Lerner index.

Methods: Take a sufficient statistics approach to estimating incidence for arbitrary forms of imperfect competition. The four key statistics are:

1. Pass-through rate
   - Estimate this by regressing price on marginal cost:
     $$p_{it} = \rho \text{MC}_{it} + X_{it}'\gamma + \eta_i + \varepsilon_{it},$$
     where ρ is the pass-through rate.
   - Instrument for marginal cost with a shift-share instrument where the exogenous component is changes in fuel prices and the share is electricity generation mix
2. Markup: Estimate a production function to get markup, as in equation 12, then subtract this from the price to get MC
   - Will use a shift-share instrument for energy prices to estimate how markups, prices, and marginal costs respond to a change in energy prices
3. Demand elasticity
   - Estimate this in a "standard" way, regressing quantity on price and using a TFP index as an instrument for price.
4. Cost-shift rate (the marginal effect of the input tax rate on marginal costs)
   - Same estimation strategy as for markups

Results: Standard methods (that don't take into account market power and assume perfect pass-through) overstate share of burden to consumers.

4.7 Principal-agent problems

Duflo et al. (2018): The value of regulatory discretion

Research question: How do discretionary inspections compare to randomly assigned ones in terms of detecting non-compliance?
Methods: RCT adding random environmental inspections. Combine this with a structural model of a dynamic game where regulator has discretion over what it does.
Results: Treatment plants increase compliance and pollution, both by small amounts. This is explained by structural model where plants generally find is not worth it to abate when they know probability of a random inspection is low.

5 Things to come back to

- Meredith
  - Deryugina et al. (2019)
- Joe
  - Cases in Rosen (1974)
  - Disutility of sickness and VSL
- Jim
  - Incidence supplemental lecture
  - Fullerton and Heutel extension of Harberger model (Lectures 6, 7)
    - Maybe also go back to those nice Hoynes notes on Harberger
  - Jacobsen et al. (2020)
  - Could be useful to try working through a Sandmo model with multiple externalities

References

Ai, Chunrong and Edward C. Norton, "Interaction Terms in Logit and Probit Models," *Economics Letters*, July 2003, *80* (1), 123–129.

Allcott, Hunt, Benjamin Lockwood, and Dmitry Taubinsky, "Ramsey Strikes Back: Optimal Commodity Taxes and Redistribution in the Presence of Salience Effects," *AEA Papers and Proceedings*, May 2018, *108*, 88–92.

Banzhaf, H. Spencer, "Difference-in-Differences Hedonics," *Journal of Political Economy*, August 2021, *129* (8), 2385–2414.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *journal of political economy*, 2007, p. 51.

Black, Sandra, "Do Better Schools Matter," *Quarterly Journal of Economics*, 1999.

Burgess, Robin, Michael Greenstone, Nicholas Ryan, and Anant Sudarshan, "Demand for Electricity on the Global Electrification Frontier," 2020, p. 78.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel, "Global Non-Linear Effect of Temperature on Economic Production," *Nature*, November 2015, *527* (7577), 235–239.

Chay, K. Y. and M. Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *The Quarterly Journal of Economics*, August 2003, *118* (3), 1121–1167.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li, "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy," *Proceedings of the National Academy of Sciences*, August 2013, *110* (32), 12936–12941.

Cicala, Steve, "When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation," *American Economic Review*, January 2015, *105* (1), 411–444.

_ , "Imperfect Markets versus Imperfect Regulation in US Electricity Generation," *American Economic Review*, February 2022, *112* (2), 409–441.

Currie, Janet, John Voorheis, and Reed Walker, "What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality," Technical Report w26659, National Bureau of Economic Research, Cambridge, MA January 2020.

Davis, Lucas W, "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster," *American Economic Review*, November 2004, *94* (5), 1693–1704.

_ , "What Matters for Electrification? Evidence from 70 Years of U.S. Home Heating Choices," 2021, p. 60.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken, "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, July 2012, *4* (3), 66–95.

Depro, Brooks, Christopher Timmins, and Maggie O'Neil, "White Flight and Coming to the Nuisance: Can Residential Mobility Explain Environmental Injustice?," *Journal of the Association of Environmental and Resource Economists*, September 2015, *2* (3), 439–468.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif, "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction," *American Economic Review*, December 2019, *109* (12), 4178–4219.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India," *Econometrica*, 2018, *86* (6), 2123–2160.

Fowlie, Meredith, Edward Rubin, and Reed Walker, "Bringing Satellite-Based Air Quality Estimates Down to Earth," *AEA Papers and Proceedings*, May 2019, *109*, 283–288.

Fowlie, Meredith L, "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage," *American Economic Journal: Economic Policy*, July 2009, *1* (2), 72–112.

Gamper-Rabindran, Shanti and Christopher Timmins, "Hazardous Waste Cleanup, Neighborhood Gentrification, and Environmental Justice: Evidence from Restricted Access Census Block Data," *American Economic Review*, May 2011, *101* (3), 620–624.

Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker, "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes," *American Economic Journal: Applied Economics*, April 2020, *12* (2), 303–342.

Goulder, Lawrence H. and Ian W. H. Parry, "Instrument Choice in Environmental Policy," *Review of Environmental Economics and Policy*, July 2008, *2* (2), 152–174.

Greenstone, Michael, "The Continuing Impact of Sherwin Rosen's "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition"," *journal of political economy*, p. 12.

_ **and Justin Gallagher**, "DOES HAZARDOUS WASTE MATTER? EVIDENCE FROM THE HOUSING MARKET AND THE SUPERFUND PROGRAM," *QUARTERLY JOURNAL OF ECONOMICS*, 2008, p. 53.

_ **, John List, and Chad Syverson**, "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing," Technical Report w18392, National Bureau of Economic Research, Cambridge, MA September 2012.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, "Consumer Valuation of Fuel Costs and Tax Policy: Evidence from the European Car Market," *American Economic Journal: Economic Policy*, August 2018, *10* (3), 193–225.

Hernandez-Cortes, Danae and Kyle Meng, "Do Environmental Markets Cause Environmental Injustice? Evidence from California's Carbon Market," Technical Report w27205, National Bureau of Economic Research, Cambridge, MA May 2020.

Ito, Koichiro and Shuang Zhang, "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China," *journal of political economy*, p. 46.

Jacobsen, Mark R., Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem, "The Use of Regression Statistics to Analyze Imperfect Pricing Policies," *Journal of Political Economy*, May 2020, *128* (5), 1826–1876.

Joskow, Paul L and Nancy L Rose, "THE EFFECTS OF ECONOMIC," p. 58.

Kaplow, Louis, "On the (Ir)Relevance of Distribution and Labor Supply Distortion to Government Policy," 2004, p. 21.

Keiser, David A and Joseph S Shapiro, "Consequences of the Clean Water Act and the Demand for Water Quality*," *The Quarterly Journal of Economics*, February 2019, *134* (1), 349–396.

Kline, Patrick and Enrico Moretti, "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority*," *The Quarterly Journal of Economics*, February 2014, *129* (1), 275–331.

Pizer, William A. and Brian C. Prest, "Prices versus Quantities with Policy Updating," *Journal of the Association of Environmental and Resource Economists*, May 2020, *7* (3), 483–518.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, December 1982, *90* (6), 1257–1278.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, January 1974, *82* (1), 34–55.

Ryan, Stephen, "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 2012, *80* (3), 1019–1061.

Sallee, James, "Pigou Creates Losers: On the Implausibility of Achieving Pareto Improvements from Efficiency-Enhancing Policies," Technical Report w25831, National Bureau of Economic Research, Cambridge, MA May 2019.

Segerson, Kathleen, "Uncertainty and Incentives for Nonpoint Pollution Control," *Journal of Environmental Economics and Management*, March 1988, *15* (1), 87–98.

Shapiro, Joseph S and Reed Walker, "Is Air Pollution Regulation Too Stringent?," 2021, p. 42.

Walker, W. Reed, "The Transitional Costs of Sectoral Reallocation: Evidence From the Clean Air Act and the Workforce*," *The Quarterly Journal of Economics*, November 2013, *128* (4), 1787–1835.