BLP Applications

Meredith Fowlie
UC Berkeley

October 2021

Demand for differentiated products: Refresher

- j indexes mutually exclusive, exhaustive, and finite set of choices.
- In general, consumers are assumed to choose the option that yields the highest utility.
- Assume the consumer chooses at most one of the differentiated products.
- Outside good? Consumers choose between $J + 1$ options, where $j = 0$ is usually reserved for the outside option.
- The inclusion of the outside good allows us to use these models to study aggregate demand because we do not condition on purchasing a new car.

RUM (basic set up)

- Mixed logit provides the foundation for the demand model.
- Consumer n derives utility from choice j in market m.
- Assume that the consumer chooses the option that maximizes utility.
- Allow taste parameters to vary with observable demographics.
- Utility is a function of the expenditures on other goods and services and the attributes of the differentiated good:

$$u_{njt} = \alpha p_{jt} + \sum_k x_{jtk} \beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{1}$$

Random utility model (basic set up)

$$u_{njt} = \alpha p_{jt} + \sum_k x_{jtk}\beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{2}$$

- The price of product j in market t is p_{jt}. Assume all consumers in market t are offered the same price.
- Let k denote the observable non-price attributes. Assume all consumers in market t face the same choice set.
- The X_{jt} matrix includes observable non-price product attributes that can vary across markets and products.
- The standard specification assumes all consumers in a market face the same product characteristics and the same prices.

How to model preference heterogeneity?

Decompose taste parameters into a deterministic and a stochastic component:

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

- d_n is a vector of observed demographics. For expositional simplicity, I assume that all preference parameters are allowed to vary systematically with all household characteristics.
- Let r index the consumer characteristics we observe: $d_i = \{d_{i1}...d_{ir}\}$.
- Demographic variables typically normalized to have mean zero so the β can be interpreted as average values.

How to model preference heterogeneity?

Decompose taste parameters into a deterministic and a stochastic component:

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

- The η_{ik} capture the effects of unobserved characteristics that we assume are randomly distributed in the population.
- Because taste parameters vary systematically with demographics, we capture more of the variation in tastes in this systematic component, which reduces our reliance on the parametric assumptions that identify the model.

Decomposition of indirect utility

Taken together, indirect utility can be expressed as a sum of four terms:

$$
\begin{aligned}
u_{njt} &= \underbrace{\alpha p_{jt} + X_{jt}\beta + \xi_{jt}} + \underbrace{(\sum_r \mu_{kr} d_{nr} + \eta_{nk})' X_{jt} - (\mu_\alpha + \eta_{i\alpha})\frac{p_{jt}}{y_n}} + \varepsilon_{njt} \\
&= \delta_{jt}(X_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \upsilon_{njt}(X_{jt}.p_{jt}, \eta_n, d_{nr}; \theta_2) + \varepsilon_{njt} \\
&\quad \eta_n | d_n \tilde{\ } N(0, \Sigma) \\
&\quad d_n \tilde{\ } f(d_n) \\
&\quad \varepsilon_{njt} \tilde{\ } iid\ EV1
\end{aligned}
$$

- $\alpha p_{jt} + X_{jt}\beta + \xi_{jt}$ captures the mean utility term that is common to all consumers in market t.
- The model-specific constants (aka mean utlility) are summarized as δ_{jt}.

Some applications

A good way to understand this machinery is to study its applications!

- Grigolon et al (2018): Demand-side focus: Underscore importance of accounting for preference heterogeneity.
- Ito and Zhang (2018) Hedonic Valuation of Air Quality
- Depro et a. (2015) Sorting versus Siting...

Consumer valuation of fuel costs and tax policy

Research question: To what extent do car buyers undervalue future fuel costs? And what does this apply for the effectiveness and welfare of alternative policies?

Approach: Use detailed data from a long panel of European countries and exploit variation in fuel costs by engine type.

Contribution: Underscore the importance of accounting for consumer mileage heterogeneity. Focus on mileage variation (across customers) and fuel cost variation (across engines).

Punchline: Despite modest undervaluation of fuel costs, fuel taxes better target high mileage customers.

Larger policy context

Policies to reduce CO_2 emissions

- Taxes on fuel:
  - Direct effect: increase use cost and thus impact the intensive margin
  - Indirect effect: impact vehicle choice
- Standard or product taxes based on fuel economy
  - Direct effect: impact vehicle choice
  - Indirect effect: Could impact utilization rate via rebound? (they will assume this away..)

Which one is more efficient with respect to reducing GHG reductions?

- Fuel taxes:
  - if driving behavior is elastic
  - If consumers account for future fuel costs, taxes do a better job of targeting the consumers with high mileage.
- Product taxes:
  - if investment inefficiency results from consumer myopia (or inattention?)

Methodology

- Build on aggregate RP Logit demand model (BLP):
  - Account for heterogeneous responses to fuel costs and relate this explicitly to consumer mileage heterogeneity.
  - Conservatively assume driving behavior is inelastic w.r.t. fuel price
- Rich data:
  - Detailed data at the level of car model j and engine variant k
  - 7 European countries 1998-2011
- Identification:
  - In Europe, for most car models, consumers can choose between diesel vs gasoline: options. Diesel more expensive (30%), higher fuel economy (20%), cheaper fuel price (20 cents per liter)
  - This generates within-model variation in operating costs: gasoline/diesel fuel price (country/time) $*$ fuel economy (products)

Model set up

- There are T markets (defined as country/year combinations)
- There are I_t potential consumers in each market t.
- Car choices defined in terms of model j and engine variant k.
- Includes outside option (don't buy a car)
- Use aggregate random coefficient logit demand model of BLP, adding discounted expected fuel costs as a non-price attribute.
- Include car model FE (exploiting fuel cost variation between engine types w/i same car model).
- Assume each consumer chooses at most one vehicle to maximize her utility

Demand model overview

Utility consumer i buying new car j engine k:

$$u_{ijk} = x_{jk}\beta_i^x - \alpha_i(p_{jk} + \gamma G_{ijk}) + \xi_{jk} + \varepsilon_{ijk}$$

Definition

x_{jk} vector of observed model and engine characteristics
β_i^x individual-specific taste parameters
ξ_{jk} unobserved product characteristic
ε_{ijk} individual-specific valuation for car jk
Total expected expenses on the vehicle: $(p_{jk} + G_{ijk})$
p_{jk} total price or capital costs of the vehicle
G_{ijk} present discounted value of expected driving costs
α_i marginal utility of income
γ future valuation: $\gamma = 1$: correct trade off p_{jk} vs G_{ijk}; $\gamma < 1$: undervalue; $\gamma > 1$ overvalue

Present Discounted Value of Expected Driving Costs

- G_{ijk} consumer-specific: heterogeneity in annual mileage (β_i^m)

- G_{ijk} depends on expected fuel prices (g_{ks}), on the relevant time horizon (S), and interest rate (r)

$$G_{ijk} = E\left[\sum_{s=1}^{S}(1+r)^{-s}\beta_i^m e_{jk} \cdot g_{ks}\right]$$

Definition

e_{jk} inverse of the fuel economy of car j with engine k (liter per km)
g_{ks} fuel price of engine type k (gasoline or diesel) at time s (per liter)

So this is just the NPV of km / year $\cdot$ liter/km $\cdot$ \$/liter= \$/year.

Assumptions on Present Discounted Value of Expected Driving Costs

- Annual mileage (β_i^m) varies across consumers. Estimated using the country-specific empirical distribution of mileage

- Annual mileage β_i^m perfectly inelastic (independent of the fuel price)

- Fuel prices for each engine type k (g_{ks}) follow a random walk: $E\left[g_{ks}\right] = g_k$. Thank you Anderson, Kellogg, and Sallee (2013)

$$G_{ijk} = \rho \beta_i^m e_{jk} g_k,$$

Definition

ρ is the capitalization coefficient: converts the annual fuel cost $\beta_i^m e_{jk} g_k$ into net present value

Model

- Substitute $G_{ijk} = \rho\beta_i^m e_{jk} g_k$ in utility:

$$u_{ijk} = x_{jk}\beta_i^x - \alpha_i(p_{jk} + \gamma\rho\beta_i^m e_{jk} g_k) + \xi_{jk} + \varepsilon_{ijk}$$

- $\gamma\rho$ identified information on annual mileage distribution
- They directly estimate and report $\alpha\gamma\rho$.
- They can also retrieve the future valuation (or attention weight) parameter γ by dividing $\alpha\gamma\rho$ by the estimate of the price parameter α and a value of the capitalization coeffcient ρ (assuming an interest rate and time horizon S).

$$\rho = \frac{1}{r}[1 - (1+r)^{-S}]$$

Predicted market share for model j with engine k:

Remember our conceptual mapping of choice probabilities to market shares...

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

Find the parameter values that minimize the distance between observed shares and predicted share (approximated by simulation using R draws rfom assumed distributions).

Data

- Rich panel: 7 European countries from 1998-2011
- Includes sales, prices, and product characteristics for every new passenger car sold
- Unit of observation: car variant jk, i.e. the combination of a car model j equipped with engine k.
  - Car model j: brand/model/body type combination (Volkswagen Golf hatchback)
  - Engine k: fuel engine type (gasoline or diesel), size, height, displacement horsepower (gasoline, 1,390cc, 59kW)
- 80,000 observations
- Add data on population, income, gasoline prices
- Add data on distribution of vehicle miles traveled (UK travel survey):
  - average annual mileage 14,700 km/year
  - distribution of mileage skewed to the right:

Specification of the taste parameters

$\beta_i = (\beta_i^x, \alpha_i, \beta_i^m)$ with distribution $F_\beta(\beta; \theta)$

- β_i^m follows the observed empirical distribution of mileage
- β_i^x normally distributed: estimate means and standard deviations of β_i^x :

$$\beta_i^x = \overline{\beta}^x + \Sigma^x \nu_i^x$$

- α_i is constant across individuals and inversely proportional to income y_t in market t

Final utility specification:

$$u_{ijkt} = x_{jkt}\beta_i^x - \alpha p_{jkt}/y_t - \alpha\gamma\rho\beta_i^m e_{jkt} g_{kt}/y_t + \xi_{jkt} + \varepsilon_{ijkt}$$

More details

- The vector of product characteristics includes horsepower, size, height, foreign, and a diesel indicator.
- Only preferences for horsepower, size, and foreign vary randomly in the population.
- The use a contraction mapping to invert the market share system.
- Operating cost variation comes from temporal variation in fuel prices, differences between gasoline and diesel engines, differences within gasoline and diesel engines. Most variation comes from the last source.
- To deal with the potential endogeneity price, they use the BLP instruments
- Also consider cost shifters as alternative instruments (e.g. labor costs in the country of product).

Empirical results

To assess the importance of heterogeneity, the authors estimate three models:

1. Simple logit model - all consumers assigned average mileage.
2. Allow for heterogeneity in mileage (and thus heterogeneity in fuel costs) using the empirical mileage distribution.
3. Allow for mileage heterogeneity and heterogeneity in valuation of other car attributes (e.g. horsepower).

Estimation overview (from last time)

- Collect data on product attributes X and market shares y.
- Specify latent utility function (which includes parametric assumptions about how random taste parameters are distributed in the population.
- Simulate drawing from these assumed distributions (generate a set of R draws) using an initial guess for the non-liner parameters.
- Estimate the implied market shares given these draws and an initial guess of the mean utilities δ and an initial guess of the non-linear parameters θ_2.
- Calibrate the simulated log likelihood function (SLL) or moment conditions and objective function (GMM).
- Use numerical optimization algorithms to identify the estimates of non-linear parameters and the δ parameters.
- Once converged, step outside and use estimated δ to recover the linear parameters.

Estimation via simulation

$$\widehat{s}_{it}(\delta_{it};\Theta_2) = \int\int \frac{\exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha_n})p_{it})}{1 + \sum_j \exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha})p_{jt})} df(\alpha_n, \beta_{in}).$$

- Given distributional assumptions, we can predict the market share of each product in each market as a function of the observed product characteristics, prices, random tastes, etc.
- To simulate this multi-dimensional integral, we are taking a weighted sum across choice probabilities of all individuals, where the weights are given by:
  - The probability distribution of the α_n and β_n in the population.
  - The distributions of the α_n and β_n in the population are, in turn, determined by the distribution of the η_{nk} and d_{nr} in the population.
- Non-linear parameter estimates Θ_2 are those that minimize the distance between estimated and observed market shares.

Empirical results

	Logit		RC Logit I		RC Logit II	
	Est.	St.Err.	Est.	St.Err.	Est.	St.Err.
			Mean valuations			
Price/Inc. (α)	-4.52	0.19	-6.22	0.22	-5.33	0.21
Fuel Costs/Inc. ($\alpha\gamma\rho$)	-39.03	1.41	-46.48	0.94	-47.11	9.22
Power (kW/100)	2.28	0.14	2.60	0.17	0.25	0.61
Size (cm^2/10,000)	13.25	0.44	16.69	0.48	16.77	2.02
Height (cm/100)	3.00	0.30	4.45	0.32	5.19	0.33
Foreign	-0.83	0.02	-0.75	0.02	-0.89	0.04
			Standard Deviations of valuations			
Power (kW/100)	-	-	-	-	1.95	0.25
Size	-	-	-	-	4.31	2.04
Foreign	-	-	-	-	0.49	0.43
Mileage distribution	No		Yes		Yes	
			Valuations of Future Fuel Costs			
Fuel Costs/Price ($\gamma\rho$)	8.63	0.55	7.47	0.24	8.84	1.77
Future Valuation γ ($r = 6\%$)	0.89	0.06	0.77	0.02	0.91	0.18
Consumer Loss from Misoptim. (€)	73.07		328.13		39.71	
Implicit Interes rate ($T = 10$)	2.77		5.69		2.31	
Implicit Interes rate ($T = 15$)	7.87		10.32		7.48	

Empirical results

- To impute γ (future valuation parameter), set r $=$ 6% and assume a time horizon S $=$ 10 or 15.
- The logit model implies that $\gamma\rho = 8.63$. Using the above values of r and S, the implied attention weight parameter is 0.89 (moderate undervaluation).
- The random coefficients logit model with only mileage heterogeneity implies slightly more undervaluation: the required payback time (attention weight parameter of 0:77)
- Full model attention weight parameter is 0.91. WTP 0.91 for 1 Euro reduction in discounted fuel costs (insignificant undervaluation)

Consumer mis-optimization?

- They can estimate consumer surplus loss from mis-optimization by re-estimating vehicle choice assuming full valuation (setting attention parameter $=1$).
- Welfare loss will only manifest if consumers make different choices under full valuation.
- Estimate consumer surplus losses per vehicle of 40 Euros.
- These estimates are qualitatively consistent with others in the literature.
- BUT more structural approach allows them to explore policy implications in more detail...

Reaping the benefits of a structural demand (and supply) model

"Since our approach is based on the estimates of a structural demand model, we can report interesting additional information. First. we can compare the impact of a fuel tax on the market shares with the impact of a revenue-neutral vehicle tax. Second,. we are able to compute policy impacts under alternative scenarios with full forward looking behavior."

Policy Counterfactuals

- First use first-order conditions to recover firms' marginal costs.
- Then simulate market impacts of a fuel tax and a (revenue neutral) vehicle tax.
- Total welfare is the sum of tax revenues, reduction in environmental externalities, and change in consumer surplus.
- When consumers undervalue future fuel costs, consumer surplus is the sum of the decision consumer surplus and the belief error.
- Impact of the taxes on tax revenues and consumer surplus can be directly computed from demand estimates.
- Emissions impacts require estimates of externality per unit of fuel.

Key Findings

- Using the simple logit with no heterogeneity, the fuel tax and product tax have the same impact on sales-weighted fuel economy and fuel consumption.
- In all models where mileage heterogeneity is accounted for, the fuel tax is more effective than the vehicle tax at reducing total fuel usage because it targets high mileage customers.
- Using the RCL model, the authors show that the fuel tax welfare dominate an emissions equivalent vehicle tax (because a lower fuel tax required to achieve the same emissions reduction).

Research design trade offs

- Suppose you are not so excited about using this BLP framework and all of it's structural assumptions.
- How might you use the same kind of detailed information about vehicle attributes, gas prices, vehicle prices, vehicle sales to answer similar questions with less structure?
- What other empirical strategy might you use?

How do consumers' vehicle choices respond to gas prices?

- Busse, Knittel, and Zettelmeyer (2013) explore how consumers account for future fuel costs in their vehicle purchases.
- They do not attempt to estimate the deep structural parameters of the demand system (or the supply side).
- Instead they estimate the 'reduced form' effects of fuel operating cost variation on equilibrium vehicle prices and quantities.
- They then invoke additional structure post-estimation to explore what their reduced form estimates imply for consumer myopia and policy design.

Busse, Knittel, Zettelmeyer(2013): Data overview

- individual new and used car transactions over 1999-2008;
- transaction prices;
- Monthly, station-level gas prices (aggregate to designated market areas);
- vehicle (non-price) attributes
- consumer demographics.

They use these data to estimate a reduced form model of the short-run equilibrium effects of changes in gasoline prices on the transaction prices, market shares, and unit sales of cars of different fuel economics.

BKZ reduced form estimating equation

The specification they estimate can be summarized as:

$$P_{ijrt} = \lambda_0 + \lambda_1 (G_{it} * MPGQuartile_j) + \lambda_2 DEM_{it} + \delta_j + \tau_{rt} + \upsilon_{rt} + \varepsilon_{ijrt}$$

- The subscript i, t, r, j denote transaction, date, region, and car model, respectively.
- They include region-year and region-month-of-year fixed effects.
- Having included region-year and region-month-of-year fixed effects, they argue they do not need any supply-side cost shifters... why?

BKZ reduced form estimating equation

$$P_{ijrt} = \lambda_0 + \lambda_1 (G_{it} * MPGQuartile_j) + \lambda_2 DEM_{it} + \delta_j + \tau_{rt} + \upsilon_{rt} + \varepsilon_{ijrt}$$

Why do they allow the effect of gasoline prices to vary with fuel economy classes?

BKZ reduced form estimating equation

$$P_{ijrt} = \lambda_0 + \lambda_1(G_{it} * MPGQuartile_j) + \lambda_2 DEM_{it} + \delta_j + \tau_{rt} + \upsilon_{rt} + \varepsilon_{ijrt}$$

Why do they allow the effect of gasoline prices to vary with fuel economy classes?

- Consider a decision between two cars–a current vehicle and a new vehicle that is more efficient. Which improvement will save the most gas over 10,000 miles?
  1. An improvement from 10 to 11 MPG
  2. An improvement from 16.5 to 20 MPG
  3. An improvement from 33 to 50 MPG

  All save about the same amount of gas over 10,000 miles!
- The fuel cost advantage of an incremental improvement in fuel efficiency varies with baseline MPG.
- We should expect equilibrium prices for more efficient cars to rise and less efficient cars to fall with an increase in gas prices.

Variation in gas prices?

- Cross-sectional variation in gas prices driven by: differences across locations in gasoline transportation costs (or transportation capacity), variation in the degree of market power, and differences in regional and local policies governing gasoline formulation, production, etc.
- Intertemporal variation in gasoline prices arises mostly from differences in the world price of oil.
- What parts of this variation do they use to identify the λ_1 coefficient?

Identifying variation?

$$P_{ijrt} = \lambda_0 + \lambda_1(G_{it} * MPG_Quartile_j) + \lambda_2 DEM_{it} + \delta_j + \tau_{rt} + \upsilon_{rt} + \varepsilon_{ijrt}$$

- Intertemporal variation: Within a year and region variation in gasoline prices that differs from the average pattern of seasonal variation within that region.
- Cross-sectional variation: Persistent differences across markets *within* a region in factors such as transportation costs, producer restrictions, or market power, as well as month-to-month fluctuations in the gasoline price differentials between markets.
- They pool these sources of cross-sectional and intertemporal variation.

Estimated gasoline price coefficients by quartile

Fuel economy (low to high)	New cars coefficient	Used cars coefficient
1	-$250	-$1,182
2	-$96	-$101
3	-$11	$468
4	$104	$763

A $1 increase in gasoline reduces negotiated price of cars in the lowest fuel economy quartile by $250.

How should we interpret these findings?

How should we interpret these findings?

These are *equilibrium price effects*. These capture the effect of gas prices on the equilibrium vehicle price P once demand and supply responses are both taken into account.

These are *short-run* effects. These capture the change in vehicle prices associated with a gas price change over the time horizon in which manufacturers would be unable to change the configurations of cars they choose to produce.

Isn't this a paper about myopia?

How do we use these reduced form coefficients to test for myopia??

- In order to interpret these parameters with respect to consumer myopia, we need to impose more structure!
- The advantage of this approach is that a reader can create his or her own estimate of consumer myopia using alternative assumptions.

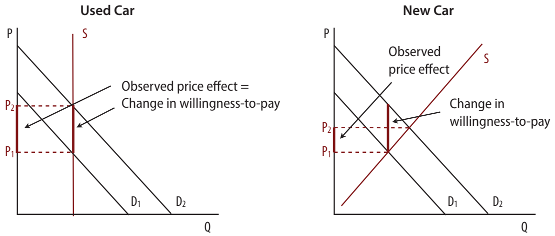
Isn't this a paper about myopia?

They use a standard model of exponential discounting and make assumptions about the following parameters:

- Vehicle miles traveled
- Investment time horizon
- Expectations about future gas prices given current prices.
- Demand elasticity.

A little more structure...

Figure 4: Effects of gasoline price change on hypothetical used and new cars



In the new car market, car dealers can to respond to changes in demand by altering prices, quantities, or both. The equilibrium price effect will be less than the change in the willingness-to-pay.

Authors assume a fixed supply curve in the used vehicle market (because the stock of used cars is predetermined). If true, the change in equilibrium prices is driven by the demand effect.

Upshot

Table 7: New and Used Cars: Implicit Discount Rates

	Market	Assumed Demand Elasticity	NHTSA VMT, NHTSA Survival Rates	VMT from Used Car Transactions, NHTSA Survival Rates	VMT from Tradeins, NHTSA Survival Rates
Q1 vs. Q4	Used	NA	11.8%	4.4%	7.3%
	New	-2	-4.0%	-6.8%	-6.2%
	New	-3	1.0%	-3.0%	-1.9%
	New	-4	5.5%	0.5%	2.1%
	New	-5	9.8%	3.7%	5.8%
Q1 vs. Q3	Used	NA	5.9%	0.1%	1.9%
	New	-2	-3.6%	-6.6%	-5.9%
	New	-3	1.5%	-2.6%	-1.5%
	New	-4	6.1%	0.9%	2.5%
	New	-5	10.4%	4.2%	6.3%
Q2 vs. Q4	Used	NA	20.9%	11.0%	16.2%
	New	-2	0.3%	-3.5%	-2.5%
	New	-3	6.7%	1.4%	3.1%
	New	-4	12.6%	5.8%	8.3%
	New	-5	18.3%	10.0%	13.2%

- Over a range of assumed values, implicit discount rates are not that large.
- Most are near or below typical interest rates for car loans. People appear to be using discount rates near or below the interest rate they would pay on a car loan.

A reduced form approach (BKZ)

We choose to estimate reduced form parameters. In order to interpret these parameters with respect to consumer myopia, we have to make assumptions similar to what must be assumed in the structural approach; namely, how many miles the owner will drive each year, how long the car will last, and what the buyer's expectation of future gasoline price is.

The advantage of this approach is that a reader of this article can create his or her own estimate of consumer myopia using alternative assumptions about driving behavior, gasoline prices, or vehicle life.

The disadvantage is that reduced form parameters cannot be used in policy simulations or counterfactuals the way structural parameters can.

More (versus less) structural research design

There are trade offs. Discuss!

More (versus less) structural research design

There are trade offs. Discuss!

- We can estimate unobserved economic or behavioural parameters that could not be otherwise inferred from non-experimental data.
- We can estimate distributions of preference parameters in the population (to explore the implications of preference heterogeneity).
- Welfare analysis! We can model counterfactual policy outcomes and infer welfare implications of a discrete change or policy intervention.
- Out of sample predictions, especially wrt counterfactual policy settings (e.g. how would a gas tax impact fleet composition in equilibrium?)

What are the costs?

What are the costs?

- Imposing more structure often requires more assumptions... are these plausible?
- The more complex/involved your structural assumptions, the less transparent the analysis, and the more difficult it can be to see how assumptions affect results.
- The more complex and non-linear the model, the greater the chance that you will fail to find the global extrema.

The benefits generated by the imposed structure should exceed the costs (but this is a subjective call!)

Some applications

A good way to understand this machinery is to study its applications!

- Grigolon et al (2018): Demand-side focus: Underscore importance of accounting for preference heterogeneity.
- Ito and Zhang (2018) Hedonic Valuation of Air Quality
- Depro et a. (2015) Sorting versus Siting

Simulated choices with inattention

- Jim uses a discrete choice model to simulate how vehicle choices- and associated welfare- would be affected if consumers are inattentive to fuel costs.
- Idea: calculate how consumer utility would change under different assumptions about information costs and choice rubrics.
- Difference between utility values with complete information and utility derived from choice under incomplete information represents the benefit of making an informed choice.
- These can be compared against effort costs.

Simulated choices with rational attention

Jim extends work by Langer(2012) who estimates a mixed logit model of vehicle demand that is highly saturated with model efects (2013).

$$U_{ij} = \delta_j + \alpha p_j \boldsymbol{z}_i + \sum_k \beta_k x_{jk} \boldsymbol{z}_i + \tilde{\alpha} \nu_{ip} p_j + \sum_k \tilde{\beta}_k \nu_{ik} x_{jk} \boldsymbol{z}_i + \varepsilon_{ij}, \tag{6}$$

where

$$\delta_j = \bar{\alpha} p_j + \sum_k x_{jk} \bar{\beta}_k + \xi_j \qquad \text{for each } j = 1, 2, \ . \ . \ . \ , J,$$

- 213 vehicle models each have common utility value term δ_j.
- ξ captures unobserved attributes.
- Heterogeneity is allowed via interactions with consumer characteristics and attributes AND random variation in taste parameters.
- Importantly, fuel consumption is captured fully in the ξ term of δ.

Simulated choices with rational attention

1. Take the demographics of the 13,454 consumers in Langer's data.
2. Draw i x k random coefficient components (one kx1 vector for each hh).
3. Given random draws + coefficient estimates reported by Langer, compute consumer surplus for each vehicle for each customer.
4. Identify vehicle with highest utility for each consumer.
5. Repeat this process, but now substitute perceived fuel consumption for true fuel consumption using three different choice rubrics.

Simulated choices under different assumptions about fuel efficiency perceptions

Table 2

Estimated Welfare Impacts of Choice under Three Information Scenarios

	Global Mean	Class Mean	Attribute-Predicted Mean
Average welfare loss ($ per vehicle)	552	217	89
	(33)	(18)	(11)
Percentage who change choice of vehicle	19	12	7
	(.8)	(.7)	(.5)
Average welfare loss conditional on changing vehicle ($)	2,957	1,800	1,262
	(141)	(112)	(121)
SD of average welfare loss ($)	1,725	909	533
	(96)	(73)	(62)

Note. Statistics are averages over 500 simulations based on the data and coefficients from Langer (2012). Standard deviations of the test statistics over the 500 trials are in parentheses. In the global mean scenario, choices are determined as if all vehicles have the fleet average fuel economy. In the class mean scenario, choices are determined as if all vehicles have the average fuel economy for their class (luxury car, sports car, all other cars, pickup trucks, small sport utility vehicles [SUVs], medium SUVs, large SUVs, and vans). Class predicts 55 percent of the variation in fuel consumption. In the attribute-predicted scenario, vehicles are assigned their predicted fuel consumption from a regression of fuel consumption on class dummies, horsepower, curb weight, passenger capacity, and a dummy for being domestically produced. These variables predict 84 percent of the variation in fuel consumption.

Key take-aways

- Empirical analysis/simulaton provides evidence to suggest that returns to paying attention may be modest. In expectation, returns to effortful information collection are not large.
- Rational attention quite plausible in vehicle choice context (and many others!)
- Rational inattention has implications for empirical research and policy.
- Why should a firm invest in making products if more energy efficiency if customers are rationally inattentive to this attribute?

Willingness to Pay for Clean Air:
Evidence from Air Purifier Markets in China

Koichiro Ito and Shuang Zhang

Motivation

- Valuation of non-market goods (and air quality in particular) is challenging but important.
- Most hedonic papers focus on housing markets.
- Ito and Zhang (2019) take a different approach: analyze WTP for defensive investments in air purifiers in China.
- Provide the first (?) revealed preference estimates of WTP for clean air in developing countries
- Empirical framework based on a random utility model.

Summary

- China is a very important setting to understand: High levels of air pollution and relatively low levels of investment in abatement.
- Primary challenge: Two key variables (pollution and price) are likely endogenous!
- Exploit a discontinuity in heating policy along China's Huai River to isolate quasi-random variation in air quality.
- Examine purchases of HEPA air filters north and south of river
- Estimate MWTP for reducing air pollution by 1 unit of PM10 for 1 year is about 1.34 USD per household (interpret this as a lower bound)

Great (market level) data!!

- At the retail store level, collect product-level information on monthly sales, monthly average price, and detailed product attributes over January 2006 through December 2014.
- The product attributes include the information on each purifier's effectiveness to reduce indoor air pollution.
- Comprehensive transaction records of 690 air purifier products sold by 45 firms. Aggregate to the product-city level.
- Pollution data from air pollution monitors 2006-2014.
- Demographic data from the Chinese census (income, education, home size, et).

Endogeneity concerns??

Endogeneity concerns??

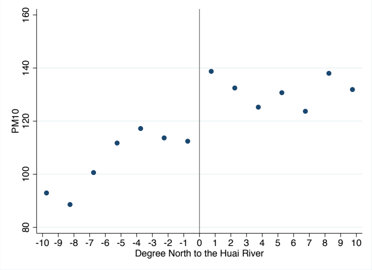
- Air pollution!: Use a spatial regression discontinuity design which exploits discontinuous valuation in air pollution created by a policy-induced natural experiment at the Huai River boundary.
  - Provided city-wide coal-based heating for cities north of the river, which generated substantially higher pollution levels in the northern cities (Almond et al., 2009; Chen et al., 2013).
  - The policy-induced variation in air pollution has existed since the 1950s: long-lasting variation in pollution.

Figure 1: Huai River Boundary and City Locations

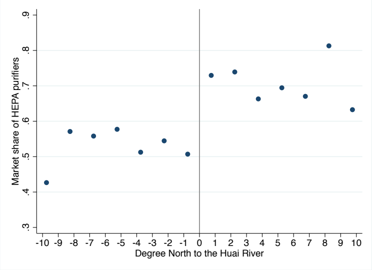
Notes: The line in the middle of the map is the Huai River-Qinling boundary. Each dot represents one city. There are 81 cities in our sample.

Figure 2: Regression Discontinuity Design at the Huai River Boundary

(a) PM10 in Winter



(b) Market Share of HEPA Purifiers in Winter



Note: Figure 2a plots the average PM_{10} during winter (December to March) in 2006 to 2012 by 1.5 degrees of latitude north of the Huai River boundary. The vertical line at 0 indicates the location of the Huai river. Each dot represents cities at 1.5 degrees of latitude and corresponds to the middle point of the range on the x-axis. For example, the dot at 0.75 on the x-axis represents cities between 0 and 1.5 degrees of latitude north of the river line. The y-axis indicates the average PM_{10} level of cities within 1.5 degrees of latitude. Figure 2b shows the market share of HEPA purifiers by 1.5 degrees of latitude north of the Huai River line.

Empirical strategy

- **Air pollution**: Spatial discontinuity because of the Huai River Policy. Observe the same product sold in regions with high versus low air pollution.
- **Purifier prices** are likely to be endogeneous.
- Include product and city fixed effects. But we might still worry if product-city unobservables are correlated with product-city prices
- IV: Distance from each product's manufacturing plant (or its port if the product is imported) to each market captures variation in the transportation cost (supply-side price shifter).

Demand for air purifiers

Consider that consumer i in city c has ambient air pollution x_c (particulate matter). The consumer can purchase air purifier j at price p_{jc} to reduce indoor air pollution by $x_{jc} = x_c \cdot e_j$. We denote purifier j's effectiveness to reduce indoor particulate matters by $e_j \in [0, 1]$. We observe markets for $c = 1, ..., C$ cities with $i = 1, ..., I_c$ consumers. The conditional indirect utility of consumer i from purchasing air purifier j at city c is:

$$u_{ijc} = \beta_i x_{jc} + \alpha_i p_{jc} + \eta_j + \lambda_c + \xi_{jc} + \epsilon_{ijc}, \tag{1}$$

where x_{jc} is the improvement in indoor air quality conditional on the purchase of product j, p_{jc} is the price of product j in market c, η_j is product fixed effects that capture utility gains from unobserved and observed product characteristics, λ_c is city fixed effects, ξ_{jc} is a product-city specific demand shock, and ϵ_{ijc} is a mean-zero stochastic term. β_i indicates the marginal utility for clean air, and α_i indicates the marginal disutility of price. The functional form for the utility function assumes that each variable, including the error term, enter the utility function linearly.

Market definition?

- Air purifiers last for 5 years and require periodic filter replacement.
- They have panel data, but the identifying variation is in the cross-section.
- So... they define markets in the cross-section (city)
- Market size = number of households in a city.
- Sum sales over 9 years of data and multiply by 5/9.
- Market share of outside option: $s_{0c} = 1 - \sum_j s_{jc}$.

Standard logit

We begin with a standard logit model. Suppose that $\beta_i = \beta$ and $\alpha_i = \alpha$ for all consumer i and that the error term ϵ_{ijc} is distributed as a Type I extreme-value function. Consumer i purchases purifier j if $u_{ijc} > u_{ikc}$ for $\forall k \neq j$. Then, the market share for product j in city c can be characterized by[8]

$$s_{jc} = \frac{\exp(\beta x_{jc} + \alpha p_{jc} + \eta_j + \lambda_c + \xi_{jc})}{\sum_{k=0}^{J} \exp(\beta x_{kc} + \alpha p_{kc} + \eta_k + \lambda_c + \xi_{kc})}. \tag{2}$$

The outside option ($j = 0$) is not to buy an air purifier. We make a few assumptions to construct

Note how air quality improvements are modeled here:

- x_{jc} is the air quality improvement associated with choice j in city c.
- HEPA filters reduce PM by 99%. Thus, $x_{jc} = x_c \cdot H$.

The Berry transformation..

Within the simple logit, share equations are easily manipulated (and estimating equations are now linear equations!)

Because $\ln s_{0c} = -\ln\left(\sum_{k=0}^{J} \exp(\beta x_{kc} + \alpha p_{kc} + \eta_k + \lambda_c + \xi_{kc})\right)$, the difference between the log market share for product j and the log market share for the outside options is $\ln s_{jc} - \ln s_{0c} = \beta x_{jc} + \alpha p_{jc} + \eta_j + \lambda_c + \xi_{jc}.$, as shown by Berry (1994). Since $\ln s_{0c}$ is absorbed by city fixed effects, this equation is simplified to:

$$\ln s_{jc} = \beta x_{jc} + \alpha p_{jc} + \eta_j + \lambda_c + \xi_{jc}, \tag{3}$$

where β is the marginal utility for improvement in air quality, and α is the marginal disutility for price. The marginal willingness to pay (MWTP) for one unit of indoor air pollution reduction can be obtained by $-\beta/\alpha$.

Identification recap (logit model)

$$lns_{jc} = \beta x_c H_j + \alpha p_{jc} + \eta_j +_c + \xi_{jc}$$

- Product FE absorb observed/unobserved non-price attributes
- City FE absorb city-level demand shocks.
- We can identify β off of the cross-city variation in air pollution.
- Identify α because prices vary across cities.
- Instrument for pollution and prices.

Random parameter logit model

- Simple logit leads to a linear estimating equation. Great!
- But CL does not accommodate random variation in preferences/tastes.
- Random coefficient logit supports this preference heterogeneity...but involves non-linear estimation.
- Allow β and α to vary in the population.

$$\beta_i = \beta_0 + \beta_1 y_i + u_i$$

$$\alpha_i = \alpha_0 + \alpha_1 y_i + \epsilon_i$$

Predicted market shares now more complicated

$\mu_{jci} = (\beta_1 y_i + u_i)x_{jc} + (\alpha_1 y_i + e_i)p_{jc}$. Then, the market share for product j in city c can be evaluated using Monte Carlo integration assuming a number n_c of individuals for city c by:[10]

$$s_{jc} = \frac{1}{n_c}\sum_{i=1}^{n_c} s_{jci} = \frac{1}{n_c}\sum_{i=1}^{n_c} \frac{\exp(\delta_{jc} + \mu_{jci})}{\sum_{k=0}^{J}\exp(\delta_{kc} + \mu_{jki})}. \quad (5)$$

The important difference between equations (2) and (5) is that equation (5) now includes elements that vary by i. Therefore, the market share and δ_{jc} has to be calculated numerically by the fixed point iterations: $\delta_{.c}^{h+1} = \delta_{.c}^{h} + \ln S_{.c} - \ln s_{.c}$ for $h = 0, ..., H$ in which $s_{.c}$ is the predicted market share by equation (5) and $S_{.c}$ is the observed market share from the data. Once δ is obtained, ξ_{jc}

Some methodological notes

- Random coefficient model requires non-linear estimation based on numerical optimization.
- Good practice to assess robustness to different starting values and search algorithms. These authors estimate 600 alternatives.
- Report tolerance levels for NFP iterations etc.

WTP for clean air?

- Estimating equations support the estimation of β and α.
- The estimate of $\frac{-\beta}{\alpha}$ provides a lower bound on WTP...why?

WTP for clean air?

- Estimating equations support the estimation of β and α.
- The estimate of $\frac{-\beta}{\alpha}$ provides a lower bound on WTP...why?
- Limited understanding of the costs of air pollution exposure?
- Indoor air quality improvements are not the same as outdoor air quality improvements.
- Other avoidance behaviors possible.

Logit results (2SLS)

Panel B: Second stage of the RD design

	Dependent variable: ln(market share)	
	(1)	(2)
PM10*HEPA (β)	0.0299***	0.0302***
	(0.0030)	(0.0032)
Price (α)	-0.0048***	-0.0048***
	(0.0001)	(0.0001)
Observations	7,359	7,359
First-stage F-Stat	285.16	292.01
Control function for running variable	Linear*North	Quadratic
MWTP for 5 years (-β/α)	6.2077***	6.3100***
	(0.6649)	(0.7130)
MWTP per year	1.2415***	1.2620***
	(0.1330)	(0.1426)

Note: Panel A shows results for the reduced-form estimation in equation (8). All regressions include product fixed effects, city fixed effects, and longitude quartile fixed effects interacted with HEPA. Price is instrumented with the distance variables discussed in the text. Panel B shows results for the second-stage estimation in equation (9). PM10*HEPA and Price are instrumented with North*HEPA and the distance variables discussed in the text. We use the two-step linear GMM estimation with the optimal weight matrix. Standard errors are clustered at the city level. * significant at 10% level; ** significant at 5% level; *** significant at 1% level. We also report the Kleibergen-Paap rk Wald F-statistic. The Stock-Yogo weak identification test critical value for one endogenous variable (10% maximal IV size) is 16.38, and for two endogenous variables (10% maximal IV size) it is 7.03.

Random Parameter Logit

Table 7: Random-Coefficient Logit Estimation Results

	Dependent variable: ln(market share)	
	(1)	(2)
PM10 · HEPA		
Mean coefficient (β_0)	0.0459***	0.0498***
	(0.0084)	(0.0092)
Interaction household income (β_1)	0.0924***	0.0891***
	(0.0224)	(0.0253)
Standard deviation (σ_β)	0.0323***	0.0570***
	(0.0117)	(0.0119)
Price		
Mean coefficient (α_0)	-0.0069***	-0.0071***
	(0.0007)	(0.0007)
Interaction with household income (α_1)	0.0028**	0.0028**
	(0.0011)	(0.0011)
Standard deviation (σ_α)	0.0006	0.0005
	(0.0007)	(0.0007)
Observations	7,359	7,359
Control function for running variable	Linear*North	Quadratic
GMM objective function value	375.05	378.93
MWTP per year: 5th percentile	0.38	0.07
MWTP per year: 10th percentile	0.49	0.20
MWTP per year: 25th percentile	0.75	0.53
MWTP per year: 50th percentile	1.19	1.10
MWTP per year: mean	1.34	1.41
MWTP per year: 75th percentile	1.90	2.04
MWTP per year: 90th percentile	2.92	3.45
MWTP per year: 95th percentile	3.86	4.69

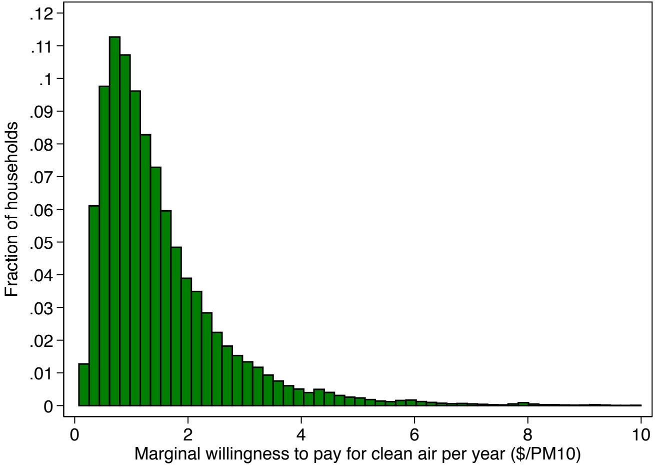
Note: This table shows the results of the random-coefficient logit estimation in equation (6). All regressions include product fixed effects, city fixed effects, and longitude quartile fixed effects interacted with HEPA. Column 1 uses a linear control for the running variable interacted with the North dummy variable, and column 2 uses a quadratic control for the running variable. Asymptotically robust standard errors are given

Implications?

- Use random-coefficient logit estimation results to calculate the mean/median of MWTP: USD \$ 1.19 \$1.34 .
- Estimates imply that a northern household is willing to pay USD \$32.70 per year to avoid the pollution increases induced by the Huai River policy.
- They can also calculate hh level MWTP as $-(\beta_0 + \beta_1 y_i + u_i)/(\alpha_0 + \alpha_1 y_i + \epsilon_i)$

Distribution of MWTP?

Figure 3: Distribution of Marginal WTP for Clean Air



Note: This histogram is based on the random-coefficient logit estimation results in column 1 of Table 7 and household-level annual income from the 2005 census micro data.

Policy implications

Table 8: Policy Implications

Panel A: Policy-relevant MWTP measures ($ per 1 ug/m^3 annual reduction in PM_{10})

	Household-level ($)	Aggregate ($)
In-sample estimate (from Table 7)	1.34	
Seven northern cities	1.62	10.13 million
Nationwide	1.26	0.45 billion

Panel B: Cost-benefit analysis: Heating reform in seven northern cities

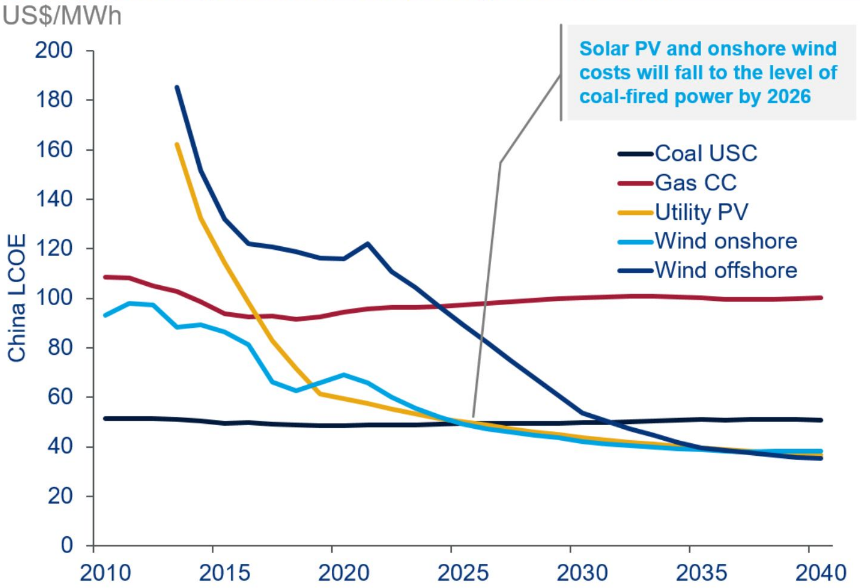
Abatement cost (million $)	2.25
Estimated PM_{10} reduction (ug/m^3)	11.91
Total WTP (million $)	105.07
Benefit-cot ratio	46.70

Panel C: Cost-benefit analysis: Replacement of coal power plants by wind or natural gas

	Wind	Natural gas
Estimated PM_{10} reduction (ug/m^3)	0.56	0.46
Total WTP (billion $)	0.26	0.21
MWTP for replacing coal-based electricity ($/MWh)	17.9	14.5

Note: This table shows policy-relevant MWTP measures and the cost-benefit analysis of two policies discussed in Section 6.

Average power generation cost (LCOE) trend in China



Source: Wood Mackenzie

Some applications

A good way to understand this machinery is to study its applications!

- Grigolon et al (2018): Demand-side focus: Underscore importance of accounting for preference heterogeneity.
- Ito and Zhang (2018) Hedonic Valuation of Air Quality
- Depro et a. (2015) Sorting versus Siting