MKT927: INTRO TO QUANTITATIVE MARKETING

Prof. Andrey Fradkin

Lecture 5: Modeling an Industry

WHY MODEL AN INDUSTRY VS JUST DEMAND?

WHAT ELSE DO WE NEED OTHER THAN DEMAND TO MODEL AN INDUSTRY?

- Model of other sides of the market.
  - Sellers (costs)
  - Platform (objective function, fee structure and other policy instruments)
  - Government (objective function, taxes and other policy instruments)
- Equilibrium Concept

SUPPLY SIDE

- Can be useful to help identify the demand side, and also needed for counterfactuals.
- For example, for a single-product monopolist, the profit function is $(p - c)q$ and the first order condition is $q'(p - c) + q = 0$.
- We can solve for price: $p = -\frac{q}{q'} + c$.
- We can estimate this equation and variations (multi-product monopolist, oligopoly, etc.). Especially helps pin down demand 'reasonable' demand elasticities.
- Risk: Mis-specification of the supply side may lead to biased demand estimates. Argument for separate estimation of demand and supply.

THE EVOLUTION OF MARKET POWER IN THE US AUTOMOBILE INDUSTRY

GRIECO, MURRY, YURUKOGLU (GMY)

WHY GMY?

- It is a recent and good paper, so close to the frontier of papers.
- It addresses an important industry and a question of broad interest to society.
- It nicely combines all the essential parts of a good structural paper.
- Nonetheless, it is far from perfect.
- Marketing angle: Importance of product improvements for demand!

SCOPE OF THE PAPER

- Models the US Auto Industry from 1980 to 2018.
- Why not just model 1 year of data?
  - How much have products improved over time?
  - How much have costs changed over time?
  - How has competition changed over time?
  - Counterfactuals about industry evolution.

MOTIVATION GMY

From 1980 to 2018, the U.S. automobile industry experienced numerous technological and regulatory changes and its market structure changed dramatically. The goal of this paper is to examine whether these changes led to discernible changes in industry performance. This work complements a recent academic and policy literature analyzing long-term trends in market power and sales concentration from a macroeconomic perspective (De Loecker et al., 2020; Autor et al., 2020) with an industry-specific approach. Several papers and commentators point to a competition problem where price-cost margins and industry concentration have increased during this time period (Economist, 2016; Covarrubias et al., 2020). Our estimates indicate a significant decline in markups over the past

CONTRAST WITH DE LOECKER ET AL. (2020)

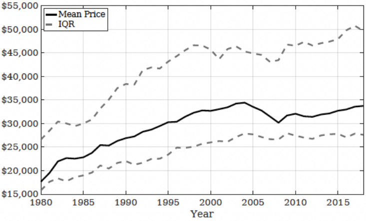
> We document the evolution of market power based on firm-level data for the U.S. economy since 1955. We measure both markups and profitability. In 1980, aggregate markups start to rise from 21% above marginal cost to 61% now. The increase is driven mainly by the upper tail of the markup distribution: the upper percentiles have increased sharply. Quite strikingly, the median is unchanged. In addition to the fattening upper tail of the markup distribution, there is reallocation of market share from low- to high-markup firms. This rise occurs mostly within industry. We also find an increase in the average profit rate from 1% to 8%. Although there is also an increase in overhead costs, the markup increase is in excess of overhead. We discuss the macroeconomic implications of an increase in average market power, which can account for a number of secular trends in the past four decades, most notably the declining labor and capital shares as well as the decrease in labor market dynamism.

DATA SOURCES

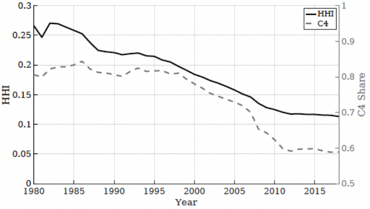
- **Ward's Automotive** (1980-2018)
  - Sales, manufacturer suggested retail prices (MSRP)
  - Vehicle characteristics and specifications
- **Consumer Surveys**
  - Consumer Expenditure Survey (CEX): 1980-2005
    - Average 1,014 observations per year
    - Individual purchases and demographics
  - MRI Survey of American Consumer: 1992-2018
    - Average 2,005 observations per year
    - Individual purchases and demographics
  - MaritzCX Second Choice Survey: 1991, 1999, 2005, 2015
    - Data on consumers' alternative vehicle preferences

INDUSTRY STATS

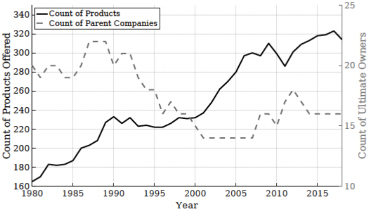
Figure II: Prices and Market Structure, 1980-2018



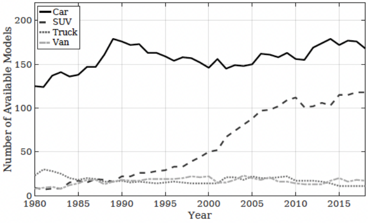
(a) Prices



(b) Measures of Concentration



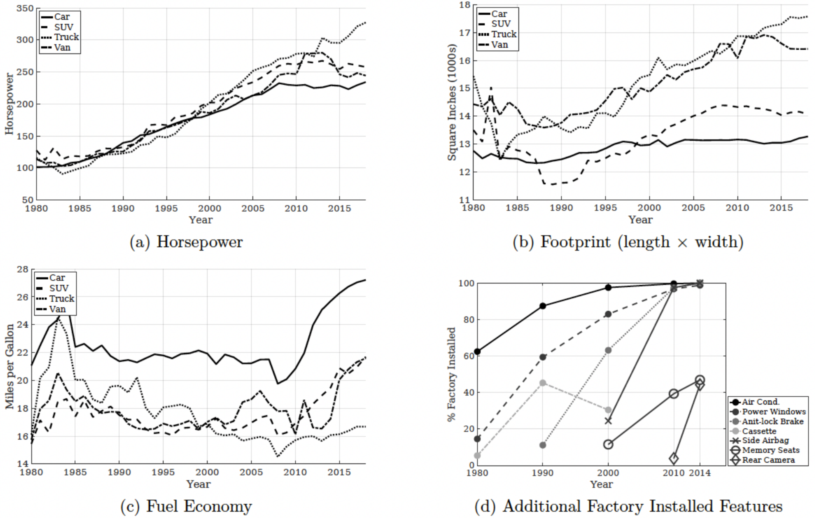
(c) Products and Manufacturers



(d) Count of Products by Styles

PRODUCT STATS

Figure III: Physical Vehicle Characteristics, 1980-2018



(a) Horsepower

(b) Footprint (length × width)

(c) Fuel Economy

(d) Additional Factory Installed Features

Notes: Panels (a)-(c) display average characteristics for available models in our sample. Panel (d) is the percent of each feature installed on total "cars" sold (i.e. not trucks, SUVs, or vans). Factory installed features were compiled

DEMAND MODEL: BASIC SETUP

- **Consumer Utility**

  $u_{ijt} = \beta_{it} x_{jt} + \alpha_{it} p_{jt} + \xi_{jt} + \epsilon_{ijt}$

- **Components:**
  - x_{jt}: observable vehicle attributes (size, HP, MPG)
  - p_{jt}: price
  - ξ_{jt}: unobserved product quality
  - ϵ_{ijt}: i.i.d. Gumbel shocks

- **Outside Option**
  - $u_{i0t} = \gamma_t + \epsilon_{i0t}$
  - γ_t captures time-varying value of not purchasing
  - Reflects business cycles, urbanization, used car market changes

IDENTIFICATION CHALLENGE: QUALITY VS OUTSIDE OPTION

- **Key Issue**
  - Models only identify utility relative to outside good
  - Need to separate:
    - τ_t: unobserved quality of new cars
    - γ_t: value of outside option
- **Solution: Continuing Products Strategy**
  - For vehicles produced in consecutive years ($j \in \mathcal{C}t$):
    $E[\xi jt - \xi_{jt-1}] = E[(\tau_t - \tau_{t-1}) + (\tilde{\xi} jt - \tilde{\xi} jt - 1)] = 0$
  - Intuition: Within model generation, average quality changes are zero
- **Two-Step Identification**
  - Step 1: Identify $\tau_t - \gamma_t$ from choice probabilities
  - Step 2: Use continuing products condition to identify τ_t

DECOMPOSING PRODUCT QUALITY

- **Product Quality Structure**
  - Unobserved quality decomposed as: $\xi_{jt} = \tau_t + \tilde{\xi}_{jt}$
  - τ_t: mean quality in year t relative to base year
  - $\tilde{\xi}jt$: product-specific deviation
- **Identification Requirements**
  - Zero conditional mean: $E[\tilde{\xi}_{jt}|z_{jt}] = 0$
  - Instruments z_{jt} include:
    - Observable characteristics (x_{jt})
    - Year dummies
    - Real exchange rate (RXR) (Key price instrument)

CONSUMER PREFERENCE HETEROGENEITY

- **Price Sensitivity**

$$\alpha_{it} = \bar{\alpha} + \sum_h \alpha_h D_{it}^h$$

- **Characteristic Preferences**

$$\beta_{ik} = \bar{\beta}_k + \sum_h \beta_{kh} D_{it}^h + \sigma_k \nu_{ik}$$

- **Key Components**
  - D_{it}^h: demographic characteristics (income, etc.)
    - Distribution from Current Population Survey
  - ν_{ik}: unobserved taste variation
    - i.i.d. standard normal draws
  - σ_k: magnitude of random taste variation

MARKET SHARE EQUATION

$$s_{jt} = \int_i \frac{\exp(\beta_{it} x_{jt} + \alpha_{it} p_{jt} + \xi_{jt})}{\exp(\gamma_t) + \sum_{l \in \mathcal{J}_t} \exp(\beta_{it} x_{lt} + \alpha_{it} p_{lt} + \xi_{lt})} dF(i)$$

ESTIMATION OVERVIEW

- **Three-Step GMM Estimation Procedure: Appendix A**
  - Step 1: Joint estimation of consumer heterogeneity and mean valuations
  - Step 2: Mean taste parameters using instrumental variables
  - Step 3: Separate quality from outside option value
- **Estimation Inputs**
  - Product-level data: shares, prices, characteristics
  - Consumer-level data: demographics, purchases
  - Second-choice data from MaritzCX
  - Real exchange rates as cost instruments

STEP 1: CONSUMER HETEROGENEITY

- **Demographic Moments**

$$g_1(\theta) = E[x_{y_i}|i \in \mathcal{H}] - \int \sum_j x_j s_{ij}(\theta) dF(i|i \in \mathcal{H})$$

  - $\mathcal{H}$: demographic group (e.g., high income households)
  - $E[x_(y_i)|i \in \mathcal{H}]$: mean characteristics of products purchased for group $\mathcal{H}$
  - Match characteristics by demographic groups
  - Example: average price by income quintile

DEMOGRAPHIC MOMENTS IN DETAIL

		MRI			CEX		
Demographic	Car Attribute	Data	Model	Only Demos	Data	Model	Only Demos
$\mathbb{E}[x\|Income\,Q_5] - \mathbb{E}[x\|Income\,Q_1]$	Price	0.215	0.383	0.363	0.603	0.400	0.421
$\mathbb{E}[x\|Income\,Q_4] - \mathbb{E}[x\|Income\,Q_1]$	Price	0.016	0.229	0.215	0.356	0.266	0.268
$\mathbb{E}[x\|Income\,Q_3] - \mathbb{E}[x\|Income\,Q_1]$	Price	-0.080	0.131	0.121	0.189	0.149	0.150
$\mathbb{E}[x\|Income\,Q_2] - \mathbb{E}[x\|Income\,Q_1]$	Price	-0.146	0.068	0.062	0.069	0.078	0.075
$\mathbb{E}[x\|Age > 60] - E[x\|Age < 30]$	Price	0.518	0.369	0.328	0.257	0.344	0.342
$\mathbb{E}[x\|Age50-60] - E[x\|Age < 30]$	Price	0.388	0.340	0.330	0.239	0.291	0.316
$\mathbb{E}[x\|Age40-50] - E[x\|Age < 30]$	Price	0.329	0.294	0.295	0.265	0.236	0.246
$\mathbb{E}[x\|Age30-40] - E[x\|Age < 30]$	Price	0.266	0.146	0.148	0.265	0.135	0.139
$\mathbb{E}[x\|Family = 2] - \mathbb{E}[x\|Family = 1]$	Van	0.025	0.022	0.022	0.020	0.021	0.021
$\mathbb{E}[x\|Family = 3/4] - \mathbb{E}[x\|Family = 1]$	Van	0.063	0.062	0.061	0.058	0.059	0.059
$\mathbb{E}[x\|Family = 5+] - \mathbb{E}[x\|Family = 1]$	Van	0.132	0.131	0.126	0.120	0.121	0.122
$\mathbb{E}[x\|Family = 2] - \mathbb{E}[x\|Family = 1]$	Footprint	0.036	0.027	0.026	0.026	0.027	0.027
$\mathbb{E}[x\|Family = 3/4] - \mathbb{E}[x\|Family = 1]$	Footprint	0.018	0.024	0.023	0.026	0.025	0.025
$\mathbb{E}[x\|Family = 5+] - \mathbb{E}[x\|Family = 1]$	Footprint	0.030	0.035	0.033	0.036	0.036	0.036
$\mathbb{E}[x\|Rural] - \mathbb{E}[x\|Not\,Rural]$	Truck	-0.103	-0.098	-0.101	–	–	–
$\mathbb{E}[x\|Income\,Q_2]\,/\,\mathbb{E}[x\|Income\,Q_1]$	PurchaseProb	–	–	–	7.813	9.239	9.091
$\mathbb{E}[x\|Income\,Q_3]\,/\,\mathbb{E}[x\|Income\,Q_1]$	PurchaseProb	–	–	–	5.466	5.077	5.246
$\mathbb{E}[x\|Income\,Q_4]\,/\,\mathbb{E}[x\|Income\,Q_1]$	PurchaseProb	–	–	–	3.641	2.706	2.866
$\mathbb{E}[x\|Income\,Q_5]\,/\,\mathbb{E}[x\|Income\,Q_1]$	PurchaseProb	–	–	–	2.265	1.660	1.684

SECOND CHOICE MOMENTS

$$g_2(\theta) = E[x(y_i) \circ x(z_i)] - \int \sum_{j,k} (x_j \circ x_k) s_{i(j,k)}(\theta) dF(i)$$

- Match correlations between first and second choices
- Example: size correlation between chosen and second-choice vehicle

STEPS 2 & 3: MEAN PARAMETERS AND TIME EFFECTS

- **Step 2: Mean Preferences**

  $\delta_{jt} = \bar{\beta}x_{jt} + \bar{\alpha}p_{jt} + \iota_t + \tilde{\xi}_{jt}$

  - Back out δ_{jt} using nested fixed point procedure.
  - IV regression using real exchange rates.
  - $\iota_t = \tau_t - \gamma_t$ absorbs mean value of outside option.

- **Step 3: Separate Time Effects**

  $\hat{\tau}_t = \hat{\tau}_{t-1} + \sum_{j\in\mathcal{C}_t}(\hat{\tilde{\xi}}_{jt-1} - \hat{\tilde{\xi}}_{jt})$

  - Uses continuing products condition
  - Separates quality (τ_t) from outside option (γ_t)
  - Normalizes $\tau_0 = 0$

ESTIMATION DETAILS

- **Implementation**
  - Simulated GMM (so we take draws from a distribution and approximate an integral)
  - Weight matrix: inverse variance of data moments
  - Bootstrap standard errors
- **Bootstrap Procedure**
  - Resample micro data (CEX, MRI, MaritzCX)
  - Re-estimate full three-step procedure
  - Account for sampling variation in:
    - Consumer surveys
    - Second choices
    - Product-level unobservables (ξ_{jt})

PARAMETERS

Table IV: Coefficient Estimates

	β	σ	Demographic Interactions						
			Income	Inc. Sq.	Age	Rural	Fam. Size 2	FS 3-4	FS 5+
Price	-3.112	–	0.094	-0.462	2.065	–	–	–	–
	(1.124)		(0.010)	(0.133)	(0.122)				
Van	-7.614	5.538	–	–	–	–	1.737	3.681	5.840
	(0.598)	(0.133)					(0.165)	(0.176)	(0.176)
SUV	-0.079	3.617	–	–	–	–	–	–	–
	(0.339)	(0.087)							
Truck	-7.463	6.309	–	–	–	3.007	–	–	–
	(0.898)	(0.310)				(0.340)			
Footprint	0.534	1.873	–	–	–	–	0.481	0.459	0.636
	(0.261)	(0.118)					(0.053)	(0.054)	(0.054)
Horsepower	1.018	1.246	–	–	–	–	–	–	–
	(0.954)	(0.361)							
Miles/Gal.	-0.965	1.645	–	–	–	–	–	–	–
	(0.211)	(0.151)							
Luxury	–	2.624	–	–	–	–	–	–	–
		(0.047)							
Sport	-3.046	2.617	–	–	–	–	–	–	–
	(0.549)	(0.075)							
EV	-5.549	3.798	–	–	–	–	–	–	–
	(1.406)	(0.511)							
Euro. Brand	–	1.921	–	–	–	–	–	–	–
		(0.054)							
US Brand	–	2.141	–	–	–	–	–	–	–
		(0.048)							
Constant	–	–	0.362	–	–	–	–	–	–
			(0.034)						

COMMENTS ON PARAMETERS

- Price coefficient looks ‘good.’ Consider the ratio between price coefficient and Van. Interaction with age, etc... Authors also compute elasticities and they look ‘reasonable’ (very few smaller than 1 in magnitude).
- Weird to include some demographic interactions but not others. Why did they do this? If they include everything it would be too imprecisely estimated.
- Nonetheless, interpreting these interactions is a bit tricky.
- Their key point, lots of heterogeneity (most people don’t want trucks, but some really do!).

DOES THE MODEL FIT THE DATA?

Table VI: Attribute Correlation between First and Second Choice

			Alternative Specifications		
	Data	Model	Only Dem. & Footprint RC	Only Demographics	Logit
Van	0.720	0.727	0.048	0.008	-0.008
SUV	0.642	0.640	0.018	-0.007	-0.010
Truck	0.843	0.798	0.246	-0.013	-0.024
Footprint	0.710	0.693	0.665	-0.002	-0.018
Horsepower	0.599	0.588	0.384	0.009	-0.012
MPG	0.647	0.657	0.362	0.003	-0.013
Luxury	0.484	0.493	0.031	0.005	-0.005
Sport	0.277	0.291	0.001	-0.004	-0.004
Electric	0.373	0.192	0.002	-0.001	-0.001
Euro Brand	0.336	0.353	0.018	0.000	-0.003
US Brand	0.479	0.472	0.121	-0.010	-0.012

Notes: Data from MaritzCX survey, 1991, 1999, 2005, 2015. The numbers are the average across these four years. "Model" column represents the predictions from the model presented in Table IV, and column 1 of Tables VIII and IX. The "Logit" column contains model predictions from a simple logit demand specification, with no observed or unobserved heterogeneity. The "Only Demographics" column contains model predictions from a model with the same demographic interactions as our main specification, but without any unobserved heterogeneity. "Logit" and "Only Demographics" are estimated without moments on second choices.

SUPPLY SIDE: THE BIG PICTURE

- **What are we trying to recover?**
  - Marginal costs c_{jt} for each product-year
  - Markup patterns over time
  - Nature of competition
- **Key Assumptions**
  - Nash-Bertrand price competition
  - Static pricing game each year
  - Multi-product firms
  - Full information
- **Why these assumptions?**
  - Standard in IO literature
  - Tractable way to model oligopoly competition

SUPPLY SIDE: FIRST ORDER CONDITIONS

- **Firm's Problem**

$$\max_{p_j : j \in \mathcal{J}_m} \sum_{j \in \mathcal{J}_m} (p_j - c_j) s_j(p)$$

- **First Order Condition**

$$s_{jt} + \sum_{k \in \mathcal{J}_t^m} (p_{jt} - c_{jt}) \frac{\partial s_{jt}}{\partial p_{kt}} = 0$$

- **Key Insight**
  - Given demand estimates, this is a system of equations
  - Can solve for marginal costs
  - Need estimates of demand derivatives $\frac{\partial s_{jt}}{\partial p_{kt}}$

SUPPLY SIDE: RECOVERING MARGINAL COSTS

- **Matrix Form**

$$\underbrace{s_t}_{\text{shares}} + \underbrace{\Omega_t}_{\text{ownership}} \odot \underbrace{\Delta_t}_{\text{derivatives}} (p_t - c_t) = 0$$

- **Solving for Costs**

$$c_t = p_t - (\Omega_t \odot \Delta_t)^{-1} s_t$$

- **Components**
  - Ω_t: ownership matrix (which firm owns which products)
  - Δ_t: matrix of share derivatives
  - $\odot$: element-wise multiplication

SUPPLY SIDE: MARKUPS

- **Lerner Index**

$$L_{jt} = \frac{p_{jt} - c_{jt}}{p_{jt}}$$

- **Properties**
  - Higher markups → more market power
  - Related to elasticity in single-product case:

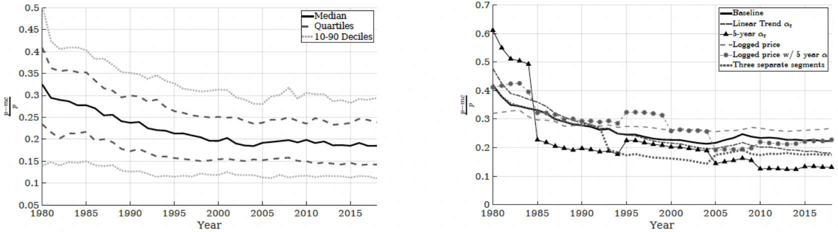
$$L_{jt} = -\frac{1}{\text{elasticity}_{jt}}$$

  - More complex with multi-product firms
- **Analysis**
  - Calculate for each product-year
  - Look at trends over time
  - Compare across segments/firms

SUPPLY SIDE: IMPORTANT CONSIDERATIONS

- **Identification**
  - Marginal costs identified from FOCs
  - But relies heavily on demand estimates
  - Quality of cost estimates depends on quality of demand estimates
- **Potential Issues**
  - Static competition assumption
  - No capacity constraints
  - No dealer relationships
  - No dynamic pricing considerations (learning by doing, etc...)
- **Robustness**
  - Test alternative conduct assumptions
  - Compare to accounting data
  - Examine reasonableness of implied costs

Figure VI: Markups

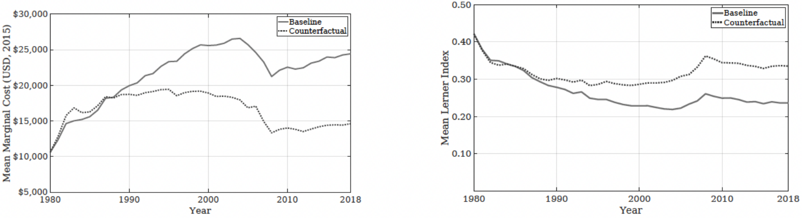


(a) Distribution of Markups

(b) Average Markups, Additional Specifications

Notes: Panel (a) displays the median, 10th, 25th, 75th, and 90th percentiles of markups over time. Panel (b) displays share-weighted markups for our baseline specification and alternative specifications described in the text. Refer to Appendix B for a detailed description of robustness specifications.

Figure IX: Counterfactual Markups, 1980 Distribution of Characteristics



(a) Marginal Cost: Baseline vs. Counterfactual

(b) Markups: Baseline vs. Counterfactual

Note: For each vehicle in each year, we assign the same percentile from the 1980 distribution of each characteristic, recompute marginal costs, which are plotted in panel (a), and recompute the pricing equilibrium and share weighted mean markups which are plotted in panel (b).

Figure X: Markups: Alternative Conduct Assumptions

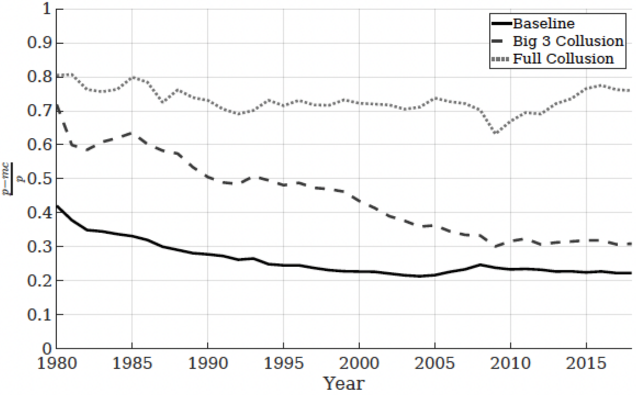
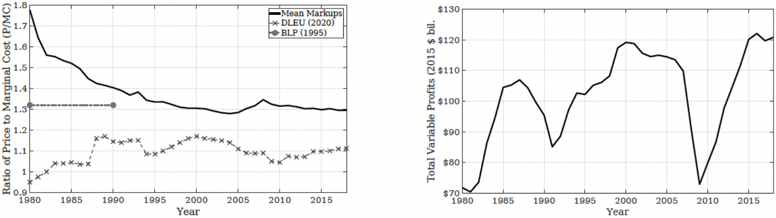


Figure XI: Comparison to De Loecker et al. (2020)



(a) Price over Marginal Cost

(b) Total Variable Profits

Figure XII: Consumer Surplus, Producer Surplus, and Deadweight Loss

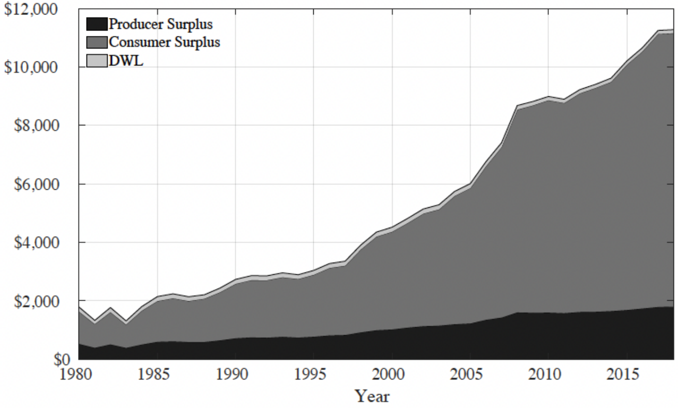
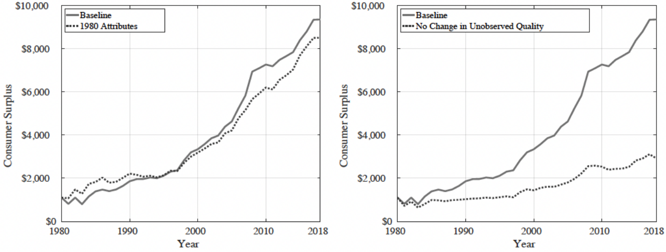


Figure XVI: Consumer Welfare, Changes in Attributes



Notes: Vertical axis represents consumer surplus per U.S. household in 2015 dollars. In the left panel, we simulate the market equilibrium if, in each year, we re-scale the distribution of footprint, horsepower, MPG, curbweight, and height, to match the 1980 distribution. In the right panel we eliminate the improvements to average unobserved quality, ξ_{jt} over time.

EQUILIBRIUM EFFECTS OF FOOD LABELING POLICIES

BARAHONA, OTERO, OTERO (BOO)

ECONOMETRICA 2023

WHY BOO?

- Recent and good paper.
- Information interventions are important in marketing and policy!
  - Labels
  - Badges
  - Notifications
- Supply side response is necessary to design effective policies.

NUTRITIONAL LABELING

- Important policy intervention.
- Lots of policy controversy. For example, in the US individuals cities are considering requiring warning labels for products high in sugar and also sugar taxes.
- Bollinger, Leslie, Jin (2011) study the effects of NYC requiring calorie counts on chain restaurant menus. Show decreased demand, but is there substitution to non-chain restaurants?
- Demand is a function of information AND price.

THE CHILEAN FOOD ACT

- Introduced in stages; Stage 1 implemented in 2016
- Sets thresholds for sugar and calories
  - 22.5g sugar/100g
  - 350 kcal/100g
- Products exceeding thresholds must display warning labels
- Aim: Inform consumers and incentivize product reformulation


ALTO EN AZÚCARES
Ministerio de Salud
ALTO EN GRASAS SATURADAS
Ministerio de Salud
ALTO EN SODIO
Ministerio de Salud
ALTO EN CALORÍAS
Ministerio de Salud


Nestlé
CEREAL INTEGRAL Garantizado
400g
4 CEREALES INTEGRALES
Cheerios
MULTIGRANO
0% Colesterol
Sin Saborizantes Artificiales
Sin Colorantes Artificiales
ALTO EN CALORÍAS
ALTO EN AZÚCARES
SABOR ORIGINAL
Coca-Cola
591 ml
Marco Polo
Caseras
Todas diferentes todas ricas
ALTO EN CALORÍAS
Papas Fritas 38g
McKay
MOROCHA
ALTO EN AZÚCARES
ALTO EN GRASAS SATURADAS
ALTO EN CALORÍAS
Nestlé
SAHNE-NUSS

DATA SOURCES

- **Scanner Data:** Walmart-Chile transactions (2015–2018)
  - Prices, quantities, demographics
- **Nutritional Information:**
  - Pre-policy data from INTA
  - Post-policy digitized nutrition facts
- **Consumer Beliefs:**
  - Survey in Argentina
  - Eliciting perceived sugar and caloric content

EVENT STUDY SPECIFICATION

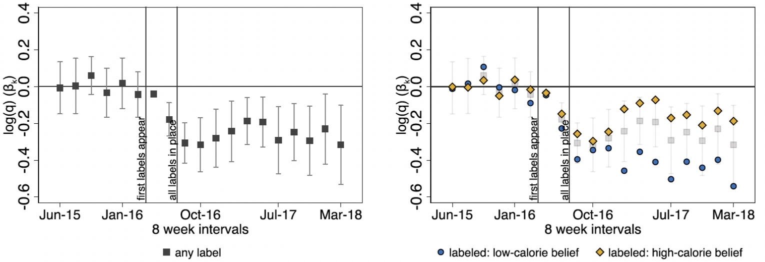
Objective: Capture the impact of the food labeling policy on product sales.

Specification:

$$\log(q_{jst}) = \sum_k \beta_k \cdot L_j \cdot \mathbf{1}\{t = k\} + \gamma \log(p_{jst}) + \delta_{js} + \delta_t + \varepsilon_{jst}$$

- q_{jst}: Quantity of product j sold in store s at period t.
- L_j: Indicator for whether product j is labeled.
- $\mathbf{1}\{t = k\}$: Time dummy for period k.
- p_{jst}: Price of product j in store s at time t.
- δ_{js} and δ_t: Product-store and time fixed effects.
- β_k: Coefficient capturing the relative change in demand post-policy.

EVENT STUDY FIGURE



(a) By labeling status (b) By labeling status and prior beliefs

FIGURE 1.—Relative changes in equilibrium quantities. *Notes*: This figure presents the coefficients of our event study regressions. Panel (a) presents the β_k coefficients from Equation (1). Panel (b) displays the coefficients from Equation (2). Coefficients in circles, diamonds, and light squares denote β^l_k, β^h_k, and β_k estimates, respectively. The vertical segments delimit the 95% confidence intervals. We run the regressions on the sample of 68 ready-to-eat cereals that show up in the pre- and post-policy periods. The sample consists of 27 unlabeled and 41 labeled products for a total of 194,510 observations.

DESCRIPTIVE EVIDENCE: FIRM RESPONSE

- Firms respond by reformulating products to avoid labels
- Evidence of “bunching” at the regulatory thresholds
- Price adjustments:
  - Unlabeled product prices rise
  - Labeled products may fall

DEMAND MODEL OVERVIEW

- Consumers derive utility from taste, price, and long-run health effects
- Utility specification:

$$u_{ijt} = \underbrace{\delta_{ijt}}_{\text{experience/taste}} - \underbrace{\alpha_i p_{jt}}_{\text{price paid}} - \underbrace{w'_{jt}\phi_i}_{\text{health consequences}}$$

- Components:
  - δ_{ijt}: Direct consumption utility (taste, experience)
  - $\alpha_i p_{jt}$: Disutility from price, varies by consumer
  - $w'_{jt}\phi_i$: Long-term health consequences
- Consumers use imperfect information; labels help update beliefs about nutritional content

EXPERIENCE UTILITY COMPONENT (δ_{ijt})

- Experience utility captures product characteristics and individual preferences:

$$\delta_{ijt} = r_j\beta_i + \delta_{jb} + \delta_{T(t)b} + \delta_{S(t)b} + \xi_{jtb} + \epsilon_{ijt}$$

- Components:
  - β_i: Individual preferences for subcategories ($\beta_i \sim N(0, \sigma_\beta)$)
  - δ_{jb}: Product fixed effects
  - $\delta_{T(t)b}, \delta_{S(t)b}$: Period and store fixed effects
  - ξ_{jtb}: Product-market-type demand shock
  - ϵ_{ijt}: Consumer shock (nested logit with correlation ρ)
- Note: Experience utility doesn't vary with nutritional content (w_{jt})
  - Firms maintain taste when reformulating products

PRICE AND HEALTH COMPONENTS

- **Price Component** ($\alpha_i p_{jt}$):
  - Price sensitivity $\alpha_i \sim \text{logN}(\alpha_b, \sigma_\alpha)$
  - Varies across consumer types
- **Health Component** ($w'_{jt}\phi_i$):
  - $\phi_i \sim \text{logN}(\phi_b, \sigma_\phi)$: perceived health damage
  - Consumers have prior beliefs $\pi_{ij} \sim N(\mu_{jb}, \Sigma_{jb})$
  - Expected utility:

$$E_{\pi_{ij}}[u_{ijt}] = \delta_{ijt} - \alpha_i p_{jt} - E_{\pi_{ij}}[w_{jt}|L_{jt}]\phi_i$$

  - Labels help consumers update beliefs via Bayes' rule

DEMAND MODEL IDENTIFICATION

- Price coefficient: cost of inputs into cereal. Uses commodities pricing data. Instrument for price with predict price as a function of these.
- Nutritional beliefs: match the survey responses to form prior.
- Preference heterogeneity:
  - Competitor cost shifters (set 1)
  - Additional cost shifters (set 2)
  - Product entry timing (τ_{jt})

DEMAND ESTIMATES

TABLE I

ESTIMATED DEMAND PARAMETERS.

Panel A: Preferences for price and healthiness (α_i, ϕ_i)

	First moments				Second moments			
	low-SES		high-SES		low-SES		high-SES	
Price (α_i)	$\bar{\alpha}_l$	0.255 (0.072)	$\bar{\alpha}_h$	0.189 (0.059)	σ_{α_l}	0.152 (0.034)	σ_{α_h}	0.113 (0.036)
Sugar (ϕ_i^s)	$\bar{\phi}_l^s$	0.013 (0.004)	$\bar{\phi}_h^s$	0.013 (0.005)	$\sigma_{\phi_l^s}$	0.054 (0.151)	$\sigma_{\phi_h^s}$	0.055 (0.153)
Calories (ϕ_i^c)	$\bar{\phi}_l^c$	0.026 (0.007)	$\bar{\phi}_h^c$	0.025 (0.008)	$\sigma_{\phi_l^c}$	0.028 (0.019)	$\sigma_{\phi_h^c}$	0.028 (0.017)

Panel B: Individual preferences for different subcategories (Σ_β)

Plain		Sugary		Chocolate		Granola		Oatmeal	
$\sigma_{\beta_{r_1}}$	0.058 (0.145)	$\sigma_{\beta_{r_2}}$	0.195 (0.186)	$\sigma_{\beta_{r_3}}$	0.215 (0.139)	$\sigma_{\beta_{r_4}}$	0.036 (0.167)	$\sigma_{\beta_{r_5}}$	0.295 (0.361)

Panel C: Remaining parameters (ρ, μ)

Nest parameter	ρ	0.959 (0.004)
Beliefs shifter	μ	−0.129 (0.019)

Note: Nutritional content is measured in grams of sugar and kilocalories per gram of cereal, and prices in dollars per 100 grams of cereal. Subscripts l and h correspond to parameters for low- and high-SES consumers, respectively. For random parameters $x_i \in \{\alpha_i, \phi_i, \beta_i\}$, we report their average $\bar{x}$ and standard deviation σ_x. Standard errors are calculated using the delta method and reported in parentheses.

SUPPLY MODEL OVERVIEW

- Firms choose prices and nutritional content to maximize profits
- Reformulation: Firms may adjust product nutrient levels (i.e., sugar/calories) at a cost
- Introduces the concept of a *bliss point* for each product—nutrient level that minimizes production cost

SUPPLY-SIDE MODEL OVERVIEW

Firm's Problem:

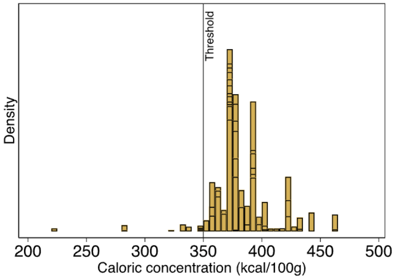
$$\max_{\{p_{jt}, w_{jt}\}_{j \in J_f}} \sum_{j \in J_f} \Big(p_{jt} - c_{jt}(w_{jt}) \Big) \cdot s_{jt}(p_t, E[w_t | L_t])$$

- **Products:** Each firm f produces a bundle J_f of products.
- **Price (p_{jt}) and Nutritional Content (w_{jt}):** Firms choose both to maximize profit.

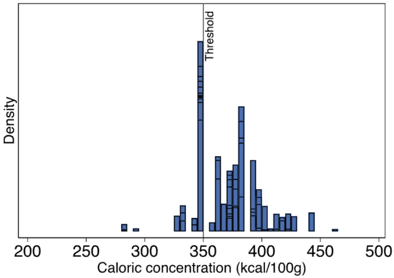
COST FUNCTION:

$$c_{jt}(w_{jt}) = \bar{c}_{jt} + (w_{jt} - \nu_j)'\Omega_j(w_{jt} - \nu_j)$$

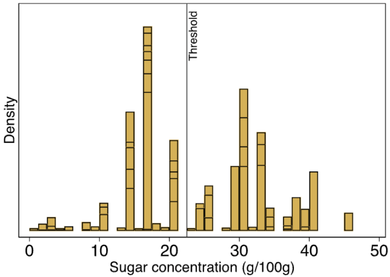
- $\bar{c}_{jt}$: Baseline cost when w_{jt} is at the bliss point ν_j.
- Ω_j: Diagonal matrix capturing marginal reformulation costs for each nutrient.
- **Bunching Behavior:** Firms may adjust w_{jt} to just avoid the regulatory threshold.



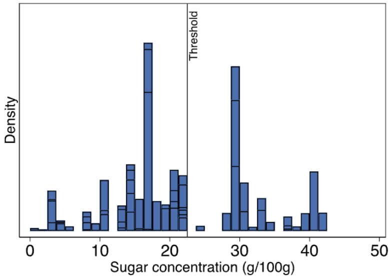
(a) Distribution of calorie content in 2016



(b) Distribution of calorie content in 2018



(c) Distribution of sugar content in 2016



(d) Distribution of sugar content in 2018

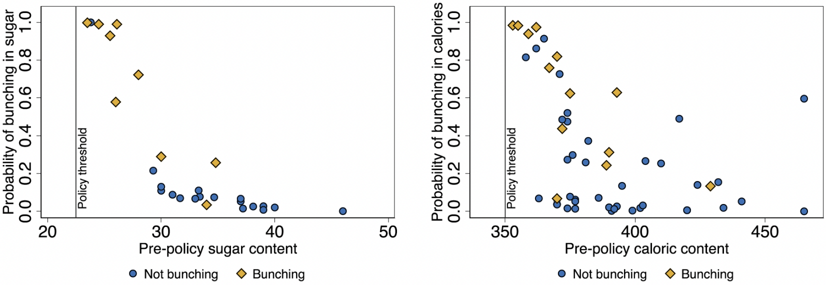
Estimation Strategy:

- Recover the **bliss point** ν_j from product characteristics.
- Estimate changes in marginal cost via a second-order Taylor approximation around ν_j:

$$c_{jt}(w) = \bar{c}_{jt} + (w - \nu_j)'\Omega_j(w - \nu_j)$$

- Use observed “bunching” at the regulatory threshold to infer Ω_j:
  - Products reformulating near the threshold indicate the cost of deviating from ν_j.
- Instruments include variation in pre-policy nutritional content and category-specific cost shifters.

BUNCHING



(a) Sugar (b) Calories

SUPPLY MODEL ESTIMATION

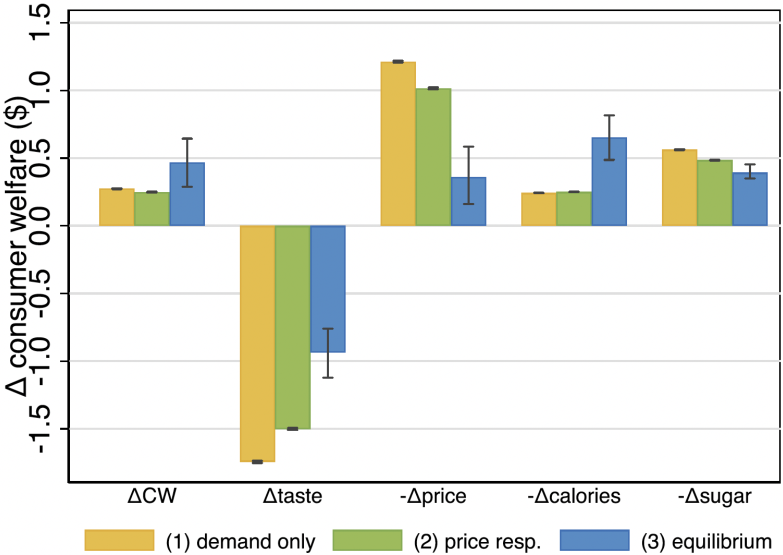
Key Findings:

- Average marginal cost increases by approximately 4–5% of product price for reformulated products.
- Evidence of significant “bunching” suggests firms are sensitive to the cost of crossing thresholds.
- Supply responses (both reformulation and price adjustments) play a crucial role in amplifying consumer welfare gains.

COUNTERFACTUAL ANALYSIS

- Four scenarios analyzed:
  - No intervention
  - Demand-only (labels without firm response)
  - Price response only
  - Full equilibrium (price and nutritional content responses)

COUNTERFACTUAL ANALYSIS



OPTIMAL POLICY DESIGN

- Threshold selection is critical:
  - Ignoring firm response suggests a higher threshold (maximizing informational value)
  - Accounting for reformulation leads to a lower threshold, boosting incentives to reformulate
- Simulation results indicate that an optimally set threshold under equilibrium increases consumer welfare by an additional 20%

FOOD LABELS VS. SUGAR TAXES

- Both policies reduce sugar/calorie intake, but with different mechanisms:
  - **Food Labels:** Target consumer misperceptions; more progressive
  - **Sugar Taxes:** Change prices directly; can be regressive
- Under optimal design, labels generate higher welfare gains when consumers are misinformed

TAKEAWAYS

- Food labeling policies can substantially improve consumer welfare by correcting misperceptions and incentivizing reformulation
- Equilibrium responses on the supply side amplify the benefits of labeling
- Optimal policy design requires accounting for both demand and supply responses
- Low-SES consumers tend to be less informed about nutritional content
- When compared to sugar taxes, food labels offer a more targeted, progressive intervention

OTHER INTERESTING INDUSTRY STUDIES

- Wollman (2018) on trucking.
- Allende et al. on giving parents information about school quality.
- Grundewald et al. (2025) about auto dealer loans.

NEXT TIME: PRICING

- pp 131 - 167, Dube and Misra (2023).
- Moshary et al. (2024).
- pp 1 - 22, Hortacsu et al. (2024).
- June discusses Chevailier et al. (2003). Please read at least the intro.