Elasticity and Curvature of Discrete Choice Demand Models *

Eugenio J. Miravete† Katja Seim‡ Jeff Thurk§

July 2023

Abstract

We explore the determinants of demand curvature and pass-through in aggregate, unit-demand, discrete choice mixed logit models. Accurate pass-through estimates are at the heart of analyses of mergers, taxation, tariffs, cost shocks, and exchange rates when firms have market power. To overcome the inherent curvature restrictions in multinomial logit models, we highlight the need to incorporate heterogeneity in both price responsiveness and preferences for product characteristics. A flexible and parsimonious specification of preference heterogeneity expands the feasible range of elasticity-curvature pairs up to those of the constant elasticity of substitution (*CES*) demand. We demonstrate empirically significant differences in estimated elasticity and curvature compared to simpler models and highlight their economic relevance in the context of price discrimination.

Keywords: Market Power, Pass-Through, Demand Curvature, Demand Manifold.

JEL Codes: C51, D43, L13, L41, L66

* We thank Simon Anderson, Christopher Conlon, Jeremy Fox, Monika Mrázová, Philipp Schmidt-Dengler, Stephen Ryan, Paulo Somaini, and Alan Sorensen for valuable discussions. All remaining errors are only our own.

† The University of Texas at Austin, Department of Economics, Austin, TX 78712; and CEPR. E–mail: *eugenio@utexas.edu*; http://www.eugeniomiravete.com

‡ Yale University, Department of Economics and School of Management, New Haven, CT 06520; CEPR and NBER. E–mail: *katja.seim@yale.edu*

§ University of Georgia, Department of Economics, Athens, GA 30602. E–mail: *jeff.thurk@uga.edu*; http://www.jeffthurk.com

1 Introduction

Demand curvature and cost pass-through drive the conclusions to many substantive questions in industrial organization (IO), including the ability of digital platforms such as Amazon.com to affect the division of surplus between third-party sellers and consumers (Gutierrez, 2022), the welfare implications of uniform pricing observed in settings ranging from consumer packaged goods (DellaVigna and Gentzkow, 2019) to consumer financial products (Cuesta and Sepúlveda, 2021), or the predicted price effects of horizontal mergers that generate cost efficiencies.[1] Demand curvature is also central to the incidence of taxes and exchange rates in non-competitive industries (Weyl and Fabinger, 2013) and the role of regulation in controlling externalities (Fabra and Reguant, 2014; Miller, Osborne and Sheu, 2017).

These examples highlight the importance of a flexible demand specification in avoiding restricting the set of model predictions. This point was underscored by Bulow and Pfleiderer (1983), who demonstrated, in the context of the tobacco industry, that functional form assumptions on demand can strongly bias statistical tests examining the existence of market power. Similarly, Froeb, Tschantz and Werden (2005) document that the predicted pass-through rates of cost efficiencies of the WorldCom–Sprint merger are seven times larger with a *CES* than with a linear demand system.

We focus on the mixed-logit (*ML*) model of discrete choice demand. This workhorse framework in applied economics is able to capture realistic substitution patterns among heterogeneous consumers and hence, aid in predicting substitution in response to a price change after a merger or in identifying collusion among firms. Less attention has been paid to the determinants of pass-through in discrete choice models, however. Berry and Haile (2021), for example, state:

> ...[S]ubstitution patterns drive answers to many questions of interest—e.g., the sizes of markups or outcomes under a counterfactual merger. However, other kinds of counterfactuals can require flexibility in other dimensions. For example, "pass-through" (e.g., of a tariff, tax, or technologically driven reduction in marginal cost) depends critically on second derivatives of demand. It is not clear that a mixed-logit model is very flexible in this dimension.

This paper examines the connection between preference specification and the range of feasible estimates for elasticity and curvature in discrete choice demand models. We highlight the implications of modeling choices of representing consumer preference heterogeneity for answering questions such as: When do assumptions on preference heterogeneity restrict feasible curvature estimates? How can we model preference heterogeneity flexibly to simultaneously allow for the estimation of realistic estimates of demand elasticity (market power) and curvature (pass-through)?

[1] Such price effects depend on the concavity of the profit function and thus demand curvature. Jaffe and Weyl (2013) suggest that for small merger-induced price increases, observed pass-through rates allow inference of the concavity of profit. For large price changes, Miller, Remer, Ryan and Sheu (2015) suggest conducting a merger simulation with a demand system constrained to mimic observed pass-through.

Figure 1: Breakfast Cereal: Elasticity and Curvature Estimates

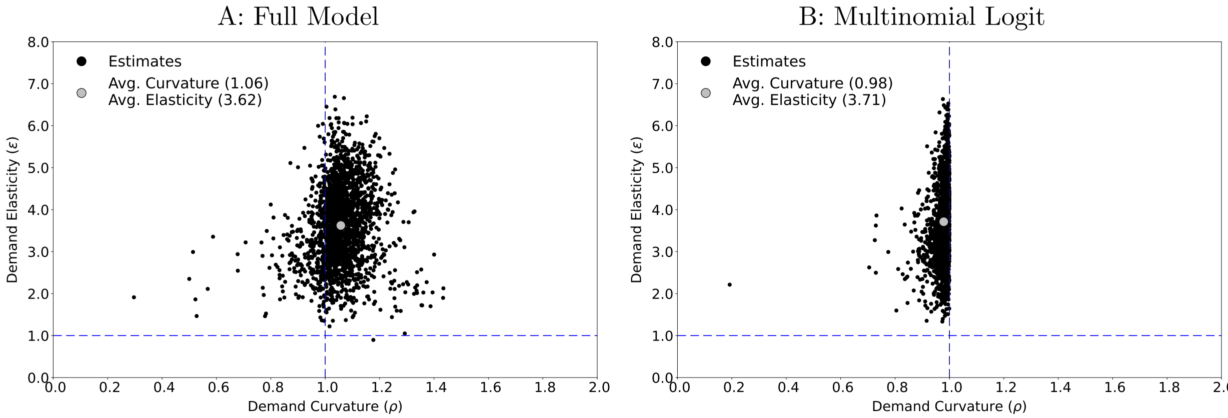


Figure Notes: Dots represent the estimated own-price elasticity and curvature for a product in the sample with the gray dot corresponding to the average elasticity and curvature.

A Motivating Example. We begin by illustrating the consequences of preference heterogeneity on elasticity and curvature estimates using the well-known aggregate data for ready-to-eat cereal from Nevo (2000). Figure 1 compares estimated own-price elasticity (ε) and curvature (ρ) for the mixed-logit model estimated in Nevo (2000) that accommodates heterogeneity in both price sensitivity and valuation of product attributes and for the simpler multinomial logit model (*MNL*). The *ML* model (Panel A) nests the *MNL* model, which abstracts from preference heterogeneity for product attributes and price. We estimate each model using Nevo's (2000) original set of Hausman-style price instruments.[2] In each panel, a dot represents an elasticity-curvature pair for a cereal product in the data, evaluated at the observed prices.

We find that the distribution of estimated elasticity-curvature pairs in the *MNL* model (Panel A) looks very different from the one in the *ML* model (Panel B). This difference suggests that how the researcher specifies demand plays an important role in estimating market power (elasticity) and pass-through (curvature). Interestingly, while the two specifications deliver similar average elasticities – a statistic often reported by researchers in the literature – their average curvatures differ substantially. Hence, alternative model specifications can deliver identical estimates of market power but have very different predictions for pass-through.

We observe that estimated demand curvatures in Panel A exceed one for the majority of the products. Hence, estimates from the *ML* model indicate that a one-dollar increase in cost results in more than a one-dollar increase in price for these products. In contrast, the *MNL* indicates that the firm price-response for all products is less than one dollar; the *MNL* specification in Panel B truncates demand curvature and hence pass-through at one.

[2] We also considered other demand specifications that rely on only an idiosyncratic price random coefficient or on only price-demographic interactions to represent heterogeneity in price sensitivity. The full description of the different specifications and the relevant estimates are reported in Appendix A.

Contributions. The motivating example of cereal demand highlights the first-order importance of relying on empirical demand models that imply robust estimates of not only substitution but also pass-through. Our objective is to highlight the implications for pass-through of modeling choices for representing consumer preference and to provide guidance for how best to estimate flexibly the shape of demand to limit the role of modeling restrictions in answers to important empirical questions. We make several theoretical and empirical contributions toward achieving this objective.

First, we identify how different components of mixed-logit demand influence demand curvature. We use the "demand manifold" approach of Mrázová and Neary (2017). While they address the behavior of elasticity and curvature for different continuous demand systems (e.g., *CES*, Pollak, translog) in a single-product monopoly model, we evaluate how components of mixed logit demand influence the relationship between elasticity and curvature in a discrete choice framework suitable for differentiated products oligopoly models. We thus connect features of the *ML* model, such as the distribution of consumer preference heterogeneity, with demand elasticity and curvature mathematically, while at the same time providing a useful representation of this relationship. For empirical applications, we demonstrate – as in the motivating example above – that depicting estimated product-level demand elasticity and curvature is a useful approach for visualizing heterogeneous demand in the industry, or equivalently, for analyzing possible restrictions of the demand specification placed on the set of feasible elasticity-curvature pairs.

It is well understood that *MNL* demand – or the related nested logit demand, which accommodates more reasonable substitution patterns at minor computational burden – is always log-concave. Hence, the curvature of demand for both models is restricted to be less than one as in Panel B of our motivating example. Using a simple single-product monopoly model, we illustrate how allowing for preference heterogeneity may alleviate this restriction. We demonstrate that curvature is determined by a tug-of-war between heterogeneity of preference over product attributes and heterogeneity in price sensitivity. When consumers have heterogeneous tastes over product attributes, demand curvature decreases at all prices relative to that of the simple *MNL* model. Thus, by incorporating heterogeneity in tastes for product attributes *only* pass-through remains at most complete. This result builds on Caplin and Nalebuff (1991b) who show that a *MNL* model with heterogeneous valuations of product attributes preserves the curvature properties of the *MNL* model. In particular, since the logit and the common normal distribution of preference heterogeneity are both log-concave functions, the resulting demand is also log-concave. Intuitively, while consumers have heterogeneous tastes over characteristics, their demand response to a change in price is uniform, which leads firms with market power to absorb some changes in marginal cost.

We show that heterogeneity in price sensitivity enables log-convex demand and that the shape of the price mixing distribution plays a vital role. Idiosyncratic price responsiveness is thus a key determinant of demand curvature, accommodating (but not imposing) more than complete pass-through. We consider three ways of specifying idiosyncratic price responsiveness – distributional assumptions for unobserved heterogeneity in price sensitivity; observable consumer heterogeneity via demographic-price interactions; and heterogeneity in income effects.

Our second contribution lies in demonstrating that unit-demand discrete choice models are capable of mimicking the shape of *CES* demand. Using *CES* demand as the starting point, the trade literature has exploited demand manifolds to illustrate the implications of relaxing some of its assumptions for pass-through and hence markups. We complement this literature by starting from the unit-demand discrete choice models commonly used in IO and illustrating that they are sufficiently flexible to accommodate any curvature at a given elasticity estimate, up to those of *CES* demand. Our result complements Anderson, de Palma and Thisse (1992) without requiring the continuous demand that may be unappealing in many empirical contexts. We further show that for a given demand curvature, competition reduces pass-through relative to the single-product, monopoly case. This stands in contrast to *CES* models where pass-through does not vary with the number of competitors. Hence, *ML* and *CES* demand models may generate identical estimates of demand elasticity and curvature, but the *CES* model will predict larger counterfactual price responses to cost changes.

Our third contribution is to evaluate the empirical implications of popular functional forms for preference heterogeneity. We first consider quasi-linear preferences and address the role of the distribution of unobserved heterogeneity in price sensitivity in generating equilibrium elasticities and curvatures, abstracting from observed heterogeneity. We show that the choice of price mixing distribution is a key modeling decision with important implications not only when using individual data (McFadden and Train, 2000), but also when using aggregate market share information. Assuming a one-tailed distribution for price random coefficients, for example, not only ensures that the demand of all simulated consumers is downward sloping (Train, 2009), but *also* expands the support of feasible demand curvatures for each possible elasticity estimate.

We then address the role of observed heterogeneity in the form of price-demographic interactions. The literature has recognized that how the demographic attributes of consumers enter the empirical specification of price sensitivity plays an important role in generating economically meaningful consumer responses to price. For example, Nevo (1998) comments that the estimated price sensitivity distribution is:

> achieved mainly by freeing the restrictive linear form in which log income influenced the price coefficient; once we allow for log income to enter in a non-linear fashion, by introducing a quadratic term, we achieve a reasonable distribution of price sensitivity.

How should we specify observed heterogeneity in sensitivity to price? We offer an easy and parsimonious way of modulating how demographics and price interact to increase the range of feasible demand curvatures. A nice feature of our approach is that it nests a *ML* model with standard distributional assumptions.

Empirically, we suggest an instrumentation strategy that recovers the distribution of price sensitivity by connecting the price responses to cost shifts at different price points to the empirical distribution of demographics through a simple and parsimonious formulation. We demonstrate identification and the consequences of misspecification via Monte Carlo simulation of a discrete choice model of demand with income effects. Consistent with our motivating example, we show that pass-through mis-specification bias can be large even when demand elasticities are identical.

We then employ this identification strategy, together with panel micro-moments, to evaluate the welfare implications of uniform pricing – a long-standing question in economics dating back to Robinson (1933) and addressed more recently by Aguirre, Cowan and Vickers (2010) and DellaVigna and Gentzkow (2019). Using scanner and individual-level data for ready-to-eat cereals, we show that our flexible specification of price sensitivity generates demand elasticity and curvature estimates that differ substantially from those derived under standard distributional assumptions because these constrained models cannot match the observed consumption patterns in the micro data. These differences in demand elasticity and curvature estimates translate to substantive differences in the evaluation of the distributional consequences of uniform pricing – a question that has received considerable policy attention.[3]

Alternative Approaches. Our paper aims to add flexibility to the *ML* model in order to reduce the impact of model specification on pass-through behavior. We focus on empirical, unit-demand, discrete choice applications only. Alternative approaches have extended the range of feasible curvatures by adopting a discrete-continuous choice framework where heterogeneous consumers choose either a budget allocation for a given product (Adão, Costinot and Donaldson, 2017; Björnerstedt and Verboven, 2016) or several units of the same product (Anderson and de Palma, 2020; Birchall and Verboven, 2022). There are many environments where modeling the budget share allocated to a product might be more appropriate than the assumption that consumers have unit demand for a product, but this flexibility also comes at a cost. For example, markups are invariant to sales and prices in the *CES* model.

Following much of the empirical literature, we focus on parametric specifications. Alternatively, Compiani (2022) uses a non-parametric approach to estimate demand flexibly. This attractive solution places few restrictions on the shape of demand in answering empirical questions of interest, such as predicting post-merger prices. The additional flexibility and steep computational cost limit this approach to empirically rare settings with at most a handful of products.

Outline. We first provide a brief mathematical introduction for the demand manifold framework in Section 2. In Section 3 we characterize the demand manifold of a general *ML* model. We evaluate the implications of different quasi-linear preference specifications for curvature and elasticity in Section 4, before we extend the analysis to environments with income effects in Section 5. Section 6 addresses the estimation and identification of heterogeneity in price sensitivity and hence demand curvature. We describe our proposed instrumentation strategy and investigate its properties in Monte Carlo analyses before turning to our empirical application of uniform pricing from the

[3] See, for example, the 2015 report by the Council of Economic Advisors on "Big Data and Differential Pricing," available at `https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf`, accessed on 5/29/2023.

ready-to-eat cereal industry. Section 7 concludes. Additional results and derivations are reported in the Appendices.

2 A Primer on Demand Manifolds

In this section, we introduce the concept of a demand manifold, a smooth relationship between demand elasticity and curvature consistent with profit maximization. Mrázová and Neary (2017) provide an excellent formal derivation of demand manifolds and their properties for a wide range of continuous demand specifications. We employ demand manifolds to assess the flexibility of alternative preference specifications in the context of discrete-choice demand, highlighting relevant issues that relate to the estimation of mixed-logit demand from an applied perspective.

We begin with a discussion of the demand manifold for a single-product monopolist, as we rely on this setup in Sections 3-5 to illustrate graphically the properties of common discrete-choice demand specifications. Next, we discuss demand sub-convexity, which we impose on the demand systems in these analyses to ensure the existence of well-behaved equilibria and comparative statics. Demand sub-convexity weakly limits the feasible elasticity and curvature combinations by ensuring that demand becomes more elastic at higher prices.

2.1 Single-Product Monopoly

Consider the case of a single-product monopolist with constant marginal cost c. The monopolist sets the price p that maximizes profits $\Pi(p) = (p-c)\cdot q(p)$ and the following necessary condition holds:

$$\Pi_p(p) = q(p) + (p-c)\cdot q_p(q) = 1 - \frac{p-c}{p}\varepsilon(p) = 0 \quad \Longleftrightarrow \quad \varepsilon(p) \equiv -\frac{p\cdot q_p(p)}{q(p)} > 1\,, \tag{1}$$

where ε denotes the elasticity of demand. Similarly, the sufficient condition for price p to maximize monopoly profits is:

$$\Pi_{pp}(p) = 2q_p(p) + (p-c)\cdot q_{pp}(q) < 0 \quad \Longleftrightarrow \quad \rho(p) \equiv \frac{q(p)\cdot q_{pp}(p)}{[q_p(p)]^2} < 2\,, \tag{2}$$

letting ρ denote the curvature of demand. While demand can be concave ($\rho < 0$), linear ($\rho = 0$), or convex ($\rho > 0$), concavity of the profit function rules out excessively convex demands.

Mrázová and Neary (2017) prove that a well-defined smooth equilibrium relationship connecting elasticity ε and curvature ρ exists for continuous demands that are decreasing ($q_p(p) < 0$ and $p_q(q) < 0$) and three times differentiable. This allows us to invert the elasticity in Equation (1), and substituting into Equation (2), we obtain the demand manifold:

$$\rho[\varepsilon(p)] = \frac{p^2\cdot q_{pp}(p)}{\varepsilon^2(p)\cdot q(p)}\,. \tag{3}$$

The slope of demand plays a central role in the profit maximization necessary condition (1); in equilibrium, demand must be elastic whenever firms have market power. Economists frequently rewrite the necessary profit maximization condition in terms of markups, or the Lerner Index.

The sufficient condition for profit maximization further requires that at the equilibrium price, the monopolist's marginal revenue function is non-increasing, which we rewrite in turn in Equation (2) as a constraint on the equilibrium curvature of demand. Cournot (1838) first established the connection between demand curvature and pass-through for a monopolist with constant marginal costs:

$$\frac{dp}{dc} = \frac{1}{2-\rho} > 0\,, \tag{4}$$

Hence, when the monopolist faces log-concave demand with $\rho < 1$, its pass-through of cost shocks is incomplete, while it is more than complete in the case of log-convex demand with $\rho > 1$. Complete pass-through occurs when $\rho = 1$.[4] Our representation of the manifold in terms of (ε, ρ) therefore directly relates to economic outcomes of interest, namely markups and pass-through.

2.2 Demand sub-convexity

Demand is said to be "sub-convex" ("super-convex") if $\log[q(p)]$ is concave (convex) in $\log(p)$. In the monopoly manifold examples we consider in Sections 3- 5, we focus attention on sub-convex demand. Sub-convexity of demand is equivalent to a demand elasticity that increases with price; i.e.,

$$\varepsilon_p(p) = \frac{\varepsilon^2(p)}{p} \cdot \left[1 + \frac{1}{\varepsilon(p)} - \rho(p)\right] > 0 \quad \Longleftrightarrow \quad \rho(p) < 1 + \frac{1}{\varepsilon(p)} = \rho(p)^{CES}. \tag{5}$$

Equation (5) also establishes a cutoff condition for the curvature of sub-convex demand. For a given elasticity ε, this cutoff is the curvature of the Constant Elasticity of Substitution (*CES*) demand, $q(p) = \beta q^{-1/\sigma}$. *CES* demand is the only demand system where a single parameter determines both elasticity and curvature: $\varepsilon^{CES} = \sigma$ and $\rho^{CES} = (\sigma+1)\sigma^{-1} > 1$. Thus, $\varepsilon_p(p) = 0$, which implies the well-known result that *CES* markups and pass-through are invariant to price.

There is widespread empirical evidence supporting the so-called *Marshall's (1920) Second Law of Demand* of demand becoming more elastic as prices rise.[5] More importantly, the equilibrium existence results for oligopoly models with differentiated products in Caplin and Nalebuff (1991a)

---

[4] In oligopoly markets, the pass-through rate also depends on substitution between products affected by a common cost shock. Weyl and Fabinger (2013) focus on the symmetric single-product oligopoly version of equation (4):

$$\frac{dp}{dc} = \frac{1}{1+\theta(1-\rho)} > 0\,, \tag{4'}$$

where θ is a conduct parameter ranging from $\theta = 0$ for perfectly competitive to $\theta = 1$ for monopoly. We evaluate the difference between (4) and (4') in the context of our Monte Carlo study of a non-symmetric oligopoly setting in Section 6.3.

[5] This includes evidence on the relationship between markups and the scale of production in macroeconomics (see Mrázová, Neary and Parenti, 2021, and references therein), markup adjustments after trade liberalization (De Loecker, Goldberg, Khandelwal and Pavcnik, 2016), pass-through of exchange rates for coffee and beer in trade (Nakamura and Zerom, 2010; Hellerstein and Goldberg, 2013), as well as tax pass-through in the legal marijuana

for single-product firms and in Nocke and Schutz (2018) for multi-product aggregative games rely heavily on sub-convexity of demand as it is interchangeable with quasi-concavity of the firm's profit function in own price where profits are strictly positive. Our analysis below also shows that sub-convexity is helpful in generating well-behaved comparative statics and equilibria: as price rises, the firm no longer has the incentive to continue raising price and garner increasing markups as it would when demand becomes increasingly inelastic in price.

3 Demand Elasticity and Curvature for Discrete Choice Models

In this section we rely on demand manifolds to explore the relationship between curvature and elasticity in the context of the discrete choice demand model that forms the backbone of much empirical work in IO: mixed logit demand. We characterize the demand manifold in general, for arbitrary specifications of preference heterogeneity, which we refine in the following sections. We define the indirect utility of consumer i from purchasing product j as:

$$u_{ij} = x_j \beta_i^\star + f_i(y_i, p_j) + \xi_j + \epsilon_{ij}\,, \qquad i \in \mathcal{I},\ j \in \mathcal{J},\ \epsilon_{ij} \sim \text{EV1}\,, \tag{6}$$

where (x_j, ξ_j) denote observed and unobserved characteristics of product j, respectively, p_j its price, and y_i consumer i's income. Mixed logit allows for heterogeneity in consumers' valuation of the product characteristics x via $\beta_i^\star$. We normalize the value of the outside option to zero.

The sub-function f_i represents how spending on the outside good, $y_i - p_j$, affects indirect utility. The effect of outside good spending varies by individual i, both because income varies across consumers and because consumers differ in their price sensitivities. To simplify notation, we write:

$$f'_{ij} = \frac{\partial f_i(y_i, p_j)}{\partial p_j}\,, \qquad \text{and} \qquad f''_{ij} = \frac{\partial^2 f_i(y_i, p_j)}{\partial p_j^2}\,. \tag{7}$$

Thus, f'_{ij} represents the marginal effect of price p_j on consumer i's indirect utility while f''_{ij} represents how this marginal effect changes with price.

Individual i purchases product j if $u_{ij} \geq u_{ik}\,,\ \forall k \in \{0, 1, \ldots, J\}$. Because of the additive i.i.d. type-I extreme value distribution of ϵ_{ij}, individual i's choice probability of product j is:

$$\mathbb{P}_{ij}(p) = \frac{\exp\left(x_j \beta_i^\star + f_i(y_i, p_j) + \xi_j\right)}{1 + \sum_{k=1}^{J} \exp\left(x_k \beta_i^\star + f_i(y_i, p_k) + \xi_k\right)}\,. \tag{8}$$

Notice that individual i makes a dichotomous decision about the purchase of product j (i.e., "Buy j" vs. "Buy Something Else"). The purchase decision is the outcome of a Bernoulli random process with a probability of success $\mathbb{P}_{ij}$, which varies with the vector of prices and characteristics

market (Hollenbeck and Uetake, 2021) and markup adjustments to changes in commodity taxation in IO (Miravete, Seim and Thurk, 2018).

of the different alternatives. The Bernoulli random variable has mean $\mu_{ij} = \mathbb{P}_{ij}$, variance $\sigma^2_{ij} = \mathbb{P}_{ij}(1-\mathbb{P}_{ij})$, and (non-standardized) skewness of $sk_{ij} = \sigma^2_{ij}(1-2\mathbb{P}_{ij})$. Aggregating over the measure of heterogeneous individuals summarized by $G(i)$, total demand for product j becomes:

$$Q_j(p) = \int\limits_{i\in\mathcal{I}} \mathbb{P}_{ij}(p)\, dG(i)\,. \tag{9}$$

We can now write the elements defining the demand manifold, elasticity and curvature of product j, relegating the detailed derivations to Appendix B. The own-price demand elasticity of product j amounts to a scale-free measure that aggregates individual price responses (demand slopes) weighted by their choice variance:

$$\varepsilon_j(p) = -\frac{p_j}{Q_j(p)} \int\limits_{i\in\mathcal{I}} f'_{ij} \cdot \sigma^2_{ij}\, dG(i)\,, \tag{10}$$

Similarly, the demand curvature of our discrete choice model is:

$$\rho_j(p) = \int\limits_{i\in\mathcal{I}} \mu_{ij}\, dG(i) \times \frac{\int f''_{ij} \cdot \sigma^2_{ij}\, dG(i) \;+ \int \left(f'_{ij}\right)^2 \cdot sk_{ij}\, dG(i)}{\left[\int f'_{ij} \cdot \sigma^2_{ij}\, dG(i)\right]^2}\,. \tag{11}$$

Combining elasticity (10) and curvature (11), we obtain the general expression for the mixed logit demand manifold:

$$\rho_j[\varepsilon_j(p)] = \frac{p_j^2}{\varepsilon_j^2(p)\cdot Q_j(p)} \cdot \left[\int f''_{ij} \cdot \sigma^2_{ij}\, dG(i) \;+ \int \left(f'_{ij}\right)^2 \cdot sk_{ij}\, dG(i)\right]. \tag{12}$$

How the researcher defines the pricing sub-function $f(\cdot)$ plays a fundamental role in determining both demand elasticity and curvature. A common empirical sub-function is simply the linear function of outside good spending, i.e., $f_{ij}(y_i, p_j) = \alpha^\star_i(y_i - p_j)$, resulting in quasi-linear demand. For a given elasticity, the curvature is now driven by heterogeneity in the idiosyncratic price sensitivity $\alpha^\star_i$. We consider this case in Section 4. Alternatively, the researcher could impose a non-linear sub-function (Griffith, Nesheim and O'Connell, 2018), with different implications for curvature and pass-through, which we discuss in Section 5.

While we illustrate graphically the manifold properties for the monopoly case in the following sections, it is important to note that the above demand manifold definition extends to multi-product oligopoly settings by including both direct own-price effects and indirect cross-price effects through the dependence of choice probabilities in Equation (8) on the vector of all prices p. Equation (12) is thus the manifold of residual demand for product j. How the introduction of competition affects the link between curvature and pass-through in practice depends on the specific substitution effects across products. We return to this issue in Section 6.3.

4 Quasi-Linear Preferences

In this section we consider quasi-linear preferences, which researchers commonly rely on for inexpensive products like the cereal varieties considered in Nevo (2000). Quasi-linear preferences also imply simpler curvature derivations since $f''_{ij}=0$. We thus consider the following variant of equation (6):

$$u_{ij} = x_j\beta_i^\star + \alpha_i^\star\big(y_i - p_j\big) + \xi_j + \epsilon_{ij}\,, \quad i \in \mathcal{I},\ j \in \mathcal{J},\ \epsilon_{ij} \sim \text{EV1}\,, \tag{13}$$

where we include $\alpha_i^\star$ to capture consumers' heterogeneous price sensitivity which we model as $\alpha_i^\star = \alpha + \sigma_p\phi_i$. The distribution of price sensitivity, therefore, is a mixture between a degenerate mean utility price coefficient α and a mean-zero distribution $\phi_i \sim \Phi$ of deviations. The relative magnitude of α and σ_p, as well as the shape of Φ, determine the shape of the distribution of $\alpha_i^\star$.

We follow the literature in specifying heterogeneity in the valuation of product characteristic x by decomposing $\beta_i^\star$ into $\beta_i^\star = \beta + \sigma_x\nu_i$, where β similarly denotes the mean valuation while ν_i captures the idiosyncratic heterogeneity in the valuation of the observed product characteristic, which we assume to take the form of a standard normal random variable scaled by σ_x.

Note that purchase decisions based on indirect utility comparisons do not depend on individual income y_i, which shifts the indirect utility of all products by $\alpha_i^\star y_i$. There are thus no income effects. Furthermore, with $f_i(y_i, p_j)$ linear in price and income, $f'_{ij} = -\alpha_i^\star$ and $f''_{ij} = 0$. Hence, demand curvature is:

$$\rho_j(p) = \int_{i\in\mathcal{I}} \mu_{ij}\, dG(i) \times \frac{\int (\alpha_i^\star)^2 \cdot sk_{ij}\, dG(i)}{\left[\int -\alpha_i^\star \cdot \sigma_{ij}^2\, dG(i)\right]^2}\,, \tag{14}$$

and the demand manifold simplifies to:

$$\rho_j[\varepsilon_j(p)] = \frac{p_j^2}{\varepsilon_j^2(p)\cdot Q_j(p)} \cdot \int (\alpha_i^\star)^2 \cdot sk_{ij}\, dG(i)\,. \tag{15}$$

Curvature and elasticity are thus inversely related for any price-quantity pair as long as $sk_{ij} = \sigma_{ij}^2(1 - 2\mathbb{P}_{ij}) > 0$, when the probability $\mathbb{P}_{ij}$ of choosing any single product is sufficiently small, e.g., the common case of choosing among hundreds of consumer products.

4.1 Demand Manifolds of Common Discrete Choice Demand Specifications

We now employ Equation (15) to explore the demand manifolds of several workhorse discrete choice specifications from the empirical literature: *MNL*, *CES*, *ML* with random coefficients on product attributes, and *ML* with a random coefficient on price. The extent and manner in which these specifications introduce flexibility in the preference specification vary, enabling us to demonstrate how the demand model's capacity to accommodate feasible combinations of elasticity and curvature changes as we relax these restrictions.

Multinomial Logit (*MNL*). In the *MNL* model, there is no unobserved heterogeneity, so $\sigma_p = \sigma_x = 0$ and $\alpha_i^\star = \alpha$ and $\beta_i^\star = \beta$. Hence, $\mathbb{P}_{ij} = \mathbb{P}_j = s_j(p)$ is the market share of product j. Elasticity and curvature reduce to:

$$\varepsilon_j(p) = \alpha p_j \left(1 - \mathbb{P}_j\right), \tag{16a}$$

$$\rho_j(p) = \frac{1 - 2\mathbb{P}_j}{1 - \mathbb{P}_j} < 1\,. \tag{16b}$$

Combining Expressions (16a)-(16b), we obtain the *MNL* demand manifold:

$$\rho_j[\varepsilon_j(p)] = \frac{\alpha p_j \left(1 - 2\mathbb{P}_j\right)}{\varepsilon_j(p)}\,. \tag{17}$$

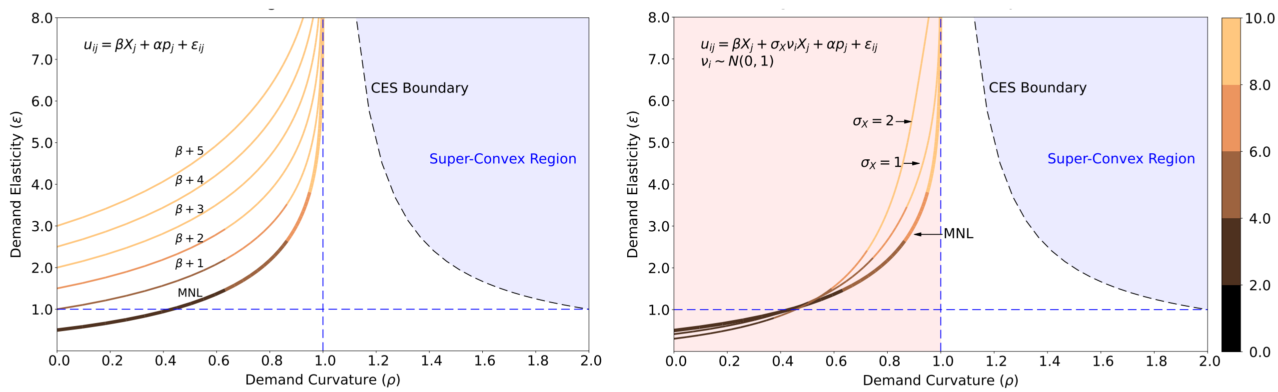
Equation (16b) shows that *MNL* demand is concave, $\rho_j(p) < 0$, only in very concentrated markets where the share of a single product exceeds 50% of sales. For less concentrated industry configurations demand is convex, $\rho_j(p) > 0$, but *MNL* also restricts demand to be log-concave, as $\rho_j(p) < 1$ for all possible prices. Thus, pass-through in *any MNL* demand model is necessarily incomplete regardless of setting and identification strategy. Furthermore, since *MNL* demand curvature (16b) decreases in $\mathbb{P}_j$, pass-through grows arbitrarily close to complete for settings with a multitude of atomistic products – a common feature of the multi-product oligopolies studied in practice, such as automobiles, breakfast cereals, spirits, etc.

The left panel of Figure 2 depicts several demand manifolds for a single-product monopoly *MNL* model. We fix the product attribute to take a value of $X = 1$ and allow consumer valuations for the attribute β to range from $\{\beta, \beta+1\ldots, \beta+5\}$, with $\beta = 1$. We set the price response coefficient $\alpha = 0.5$ and consider elasticity-curvature combinations at different price levels. Each manifold is color-coded by level of price, ranging from $p_j = 0$ (darkest) to $p_j = 10$ (lightest). Note that higher prices always correspond to more elastic demands (because *MNL* demand is sub-convex) and hence, lower equilibrium markups.

Increasing the average valuation of the product attribute, β, to $\beta+1$, $\beta+2$,$\ldots$, shifts the demand manifold upwards from the base *MNL* demand manifold in Figure 2. Increasing mean demand for a product thus decreases both demand curvature and price elasticity for a given price, consistent with Equation (17).

Constant Elasticity of Substitution. The decreasing and convex black dashed curve in Figure 2 represents the (ε, ρ) combinations for *CES* demand under alternative values for the elasticity of substitution. Anderson, de Palma and Thisse (1987, 1992) were the first to show that a discrete choice model where individuals spend a fraction of their income on a continuous quantity of a single

Figure 2: Multinomial and Mixed Logit Manifolds



Notes: The left panel shows six alternative *MNL* demand manifolds with one inside good assuming $\alpha = 0.5$, $X = 1$, and $\beta \in \{1, ..., 6\}$. The right panel shows manifolds for a *ML* model with a random coefficient on the product characteristic under alternative standard deviations σ_x and $\beta = 1$.

product can generate the *CES* utility function of the representative consumer model (Dixit and Stiglitz, 1977). Thus, *CES* arises naturally in the context of discrete-continuous models (Hanemann, 1984), while *MNL* is most appropriate when consumers have unit demand. However, like the *MNL* model, *CES* choice probabilities suffer from the *IIA* property in producing unrealistic substitution patterns.

Since *MNL* is log-concave, its manifolds always fall to the left of the *CES* boundary delimiting super-convex demands. Thus, while the *CES* and *MNL* models can both accommodate a wide range of elasticities, their demand curvatures (and pass-through) are different. The researcher's choice of one of these two demand models thus restricts pass-through in stark ways, accommodating either only under- or over-shifted pass-through, respectively, which may not be consistent with the underlying data.

ML with Characteristic Random Coefficients. It is well-known and a primary motivation for empirical research that accounting for idiosyncratic preferences for product attributes can relax the restricted substitution patterns generated by *MNL* demand. We thus consider introducing individual heterogeneity in the valuation of the product attribute, continuing to assume that all consumers have the same price responsivenessi.e., $\alpha_i = \alpha$. Will adding this flexibility also address the limitations of *MNL* to achieve greater degrees of curvature?

The right panel of Figure 2 shows several demand manifolds for such a *ML* model, allowing the standard deviation of the random coefficient on the product attribute to increase from $\sigma_x = 1$ to $\sigma_x = 2$, while holding fixed the mean product valuation at $\beta = 1$. Adding individual preference heterogeneity "rotates" manifolds: for a given demand elasticity, preference heterogeneity reduces demand curvature and hence, pass-through. The firm now faces a segment of consumers with high valuations for its attribute over whom it has market power locally, and it reduces its pass-through relative to the case of uniform preferences.

The light-red shaded area denotes the combinations of elasticity and curvature that a *ML* model with heterogeneity in the valuation of the product characteristic can generate for mean valuations of $\beta \geq 1$. The figure illustrates that the *ML* model with normally distributed attribute preferences continues to generate log-concave demand. Caplin and Nalebuff (1991b) show that *ML* demand remains log-concave under any other log-concave distribution of idiosyncratic preferences, comprising the vast majority of distributions used in economics (Bagnoli and Bergstrom, 2005). Further, this result extends naturally to the nested logit – a demand system commonly employed in antitrust economics – because it provides for more reasonable substitution patterns with small computational burden.[6] Mathematically, equation 14 demonstrates that curvature can only come through the shape of the choice probability distribution ($\mathbb{P}_{ij}$), particularly the skew. Achieving greater curvature with product characteristics therefore would require mixing distributions for idiosyncratic preferences that feature a large tail.[7]

It is evident that this version of a *ML* model has inherent limitations when used to empirically study pass-through in non-competitive environments: pass-through is necessarily restricted to be incomplete.[8] In empirical settings with log-convex demand, firms with market power aim to over-shift cost shocks, however. Employing a *MNL* or a *ML* model with idiosyncratic preferences over attributes in such instances would result in biased preference estimates that generate the closest demand curvature to the true data-generating process that these models can produce, a curvature of effectively one. Figure 2 illustrates that to exhibit such demand curvature, the estimated model would either understate the true degree of idiosyncratic product attribute preferences or overstate consumers' true price sensitivity, generating the appearance of a competitive environment with full pass-through.

ML with Price Random Coefficients. How can we expand the range of curvatures that the *ML* estimates can accommodate to allow for log-convex demand and thus over-shifting of pass-through? The only element of preferences that remains to be considered is idiosyncratic price responsiveness. Substituting $\alpha_i^\star = \alpha + \sigma_p \phi_i$ into the demand manifold for quasi-linear preferences (15) results in:

$$\rho_j[\varepsilon_j(p)] = \frac{p_j^2}{\varepsilon_j^2(p) \cdot Q_j(p)} \cdot \int (\alpha + \sigma_p \phi)^2 \cdot sk_{ij} \, d\Phi(i) \tag{18}$$

In the absence of idiosyncratic price heterogeneity, $\sigma_p = 0$, this demand manifold coincides with the manifold of the *MNL* in Equation (17). Thus, for any given demand elasticity and price-quantity

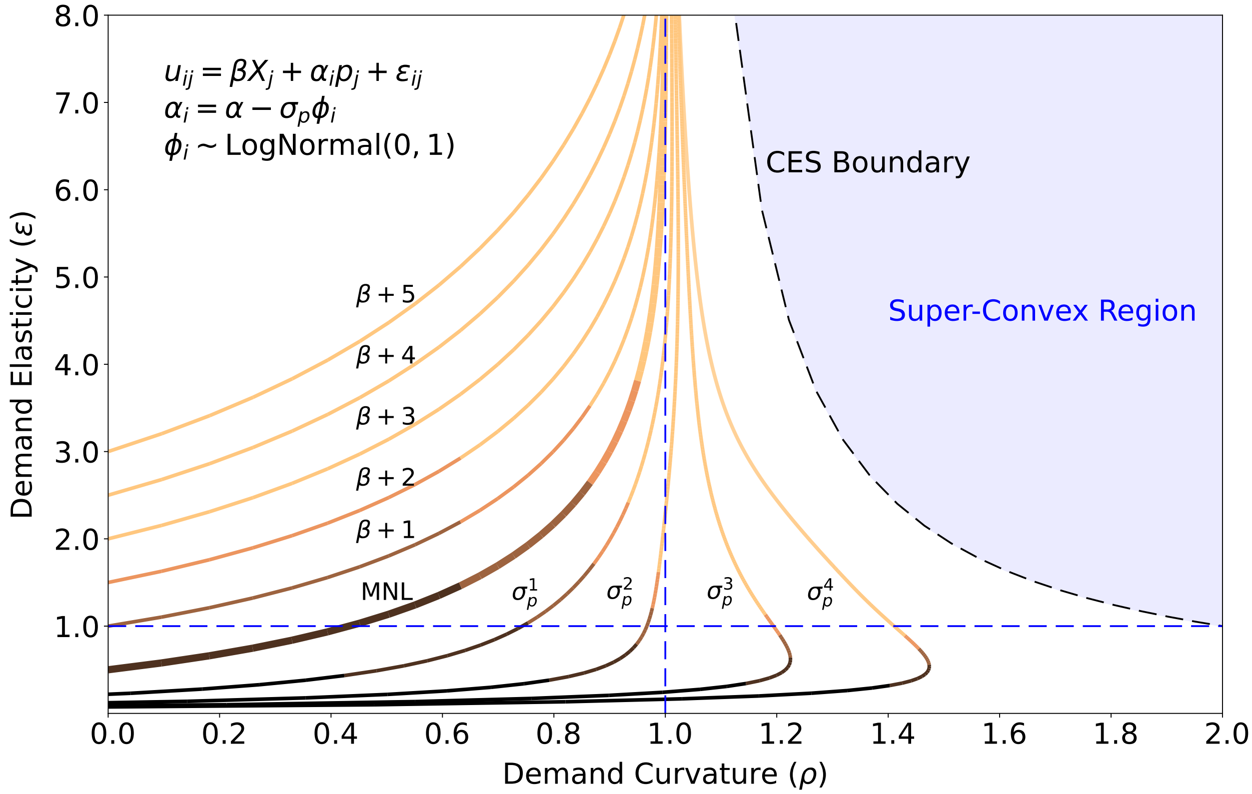
---

[6] McFadden and Train (2000) demonstrate that a *ML* specification with random coefficients on product characteristics can generate equivalent substitution patterns to the nested logit model.

[7] Our own experiments show that while it is possible to extend demand curvature to be greater than one with idiosyncratic product characteristics alone, it takes a significant skew to achieve small increases in demand curvature beyond one.

[8] This is at odds with the mounting evidence of pass-through rates exceeding 100% in horizontally differentiated products industries such as groceries (Besley and Rosen, 1999); clothing and personal care items (Poterba, 1996); branded retail products (Besanko, Dubé and Gupta, 2005); gasoline and diesel fuel (Marion and Muehlegger, 2011); as well as beer, wine, and spirits (Kenkel, 2005) among others.

Figure 3: Multinomial and Mixed Logit Manifolds



Notes: Starting with the demand manifold of the *MNL* model, $\beta+1, \beta+2, \ldots$ indicate the demand manifolds of *MNL* models for higher valuations of the inside good. The other manifolds refer to the *ML* model with price random coefficients where $\sigma_p^1 < \sigma_p^2 < \sigma_p^3 < \sigma_p^4$. The random component of the slope of demand is more important for large values of σ_p.

pair, an increase in the spread of the distribution of idiosyncratic price heterogeneity via σ_p *expands* the range of demand curvatures that the model can generate.[9] With a sufficiently large σ_p, we show the manifolds will cross the unit curvature threshold, allowing discrete choice demand to accommodate pass-through rates in excess of 100%.

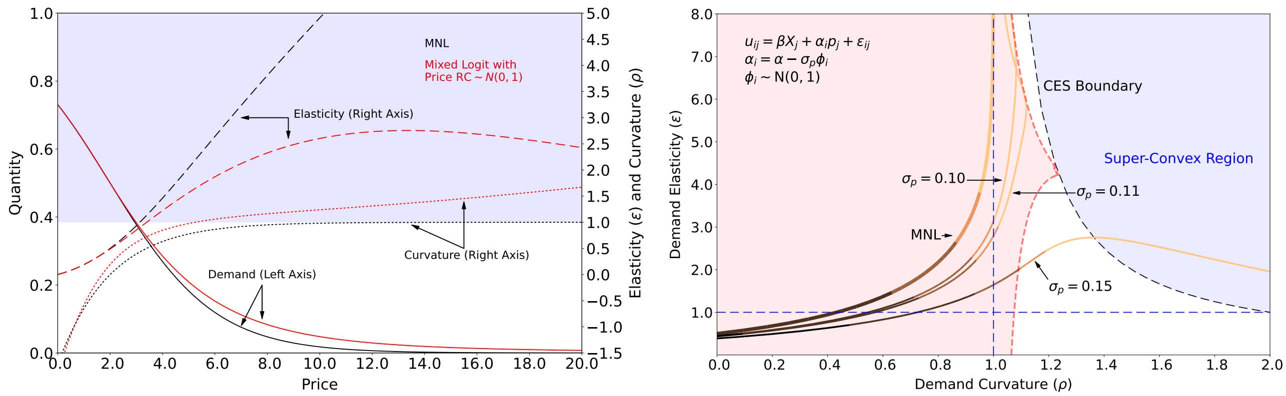
To illustrate this argument, we assume a log-normal distribution for the idiosyncratic price responsiveness, which ensures that individual demands are all downward sloping (Train, 2009); $\phi_i \sim \Phi(0,1) = \text{log-normal}(0,1)$. In Figure 3, we start from the *MNL* manifold and depict how the shape of the manifold changes as we increase the standard deviation of the log-normal mixing distribution σ_p from 0 to 1 for a fixed mean price sensitivity α.[10] We find that manifolds now cross into the log-convex region of demand with more than complete pass-through, a result that is consistent with many of the $(\hat{\varepsilon}, \hat{\rho})$ estimates for the breakfast cereal products in Figure 1's first panel.

[9] Indeed the shift of each manifold to the right is proportional to the second order moment of the distribution Φ_i.

[10] Note that to say σ_p is "large" is only meaningful in its relation to the mean price coefficient (α).

Figure 4: Demand Manifolds: Standard Normal Price Mixing Distribution

(a) Elasticity & Curvature (b) Demand Manifold



Notes: Panel (a) contrasts quantity, elasticity, and curvature under *MNL* in black and *ML* in red. Panel (b) represents demand manifolds in the (ε, ρ) plane. Light-shaded regions represent all feasible (ε, ρ) pairs conditional on the price-mixing distribution.

4.2 The Shape of Price Mixing Distribution

Figure 3 uses the example of the log-normal distribution to show that increasing the variation in idiosyncratic price responsiveness increases the feasible curvatures the *ML* model can accommodate for a given elasticity value. We now consider the choice of price mixing distribution, focusing on the range of feasible elasticity and curvature combinations up to the *CES* boundary that a candidate price mixing distribution can generate. We limit attention to two price mixing distributions common in empirical work: normal and log-normal distributions.

4.2.1 Demand under a Normal Price Random Coefficient

The left panel of Figure 4 represents demand, elasticity, and curvature at different prices for a *ML* model with a normal price random coefficient. We illustrate the case where the product attribute X and consumer valuations for the attribute β both take on values of one; mean price responsiveness remains at $\alpha = 0.5$ and the standard deviation of the price random coefficient $\sigma_p = 0.15$ so that we can address demand behavior both in the sub-convex and super-convex regions. We measure quantity on the left axis and elasticity and curvature on the right axis. The top shaded area identifies the log-convex region of demand, corresponding to curvatures greater than one.

Starting from the black solid line representing the *MNL* case when $\sigma_p = 0$, allowing for heterogeneous price sensitivity rotates demand up to the red solid line as some consumers' price sensitivity is now lower. Demand elasticity increases monotonically in price for the *MNL* model (black dashed lined), but the inclusion of the random price coefficient dampens this pattern (red dashed lines). Indeed, the *ML*'s demand elasticity reaches a maximum.

In the right panel, we depict, among others, the demand manifold corresponding to this particular demand specification with $\sigma_p = 0.15$; the manifold depiction illustrates that the maximum elasticity is reached precisely at the price level where the demand manifold crosses the *CES* locus.[11] For higher prices, elasticity decreases in price and demand becomes super-convex, violating Marshall's Second Law. We also observe that in order for demand to be sub-convex, we require less heterogeneity in price-sensitivity among consumers (i.e., smaller values of σ_p).

4.2.2 Demand Under a Log-normal Price Random Coefficient

The left panel of Figure 5 depicts the equivalent demand system under the assumption that idiosyncratic price sensitivity is distributed log-normal with $\sigma_p = 0.3$. Relative to the prior case, a log-normally distributed price random coefficient induces a larger rotation of demand as the mass of consumers with low price responsiveness is larger. The *ML*'s price elasticity now grows quasi-linearly in price for a larger price range. Thus, with a log-normal price random coefficient, it is less likely that the manifold crosses into the super-convex region of demand. As with a normally distributed price random coefficient, log-normally distributed idiosyncratic price responsiveness accommodates a curvature above one, but the curvature quickly reaches a maximum at low prices and converges asymptotically to $\rho = 1$ as the price increases.[12]

4.2.3 Normal vs. Log-normal Price Random Coefficients

The right panels of Figures 4 and 5 depict the demand manifolds when price random coefficients are normally and log-normally distributed, respectively, for alternative values of σ_p. The light-red shaded area identifies all combinations of (ε, ρ) within the sub-convex region of demand that are feasible under each model for any combination of the structural parameters $(\alpha, \sigma_p, \beta)$.

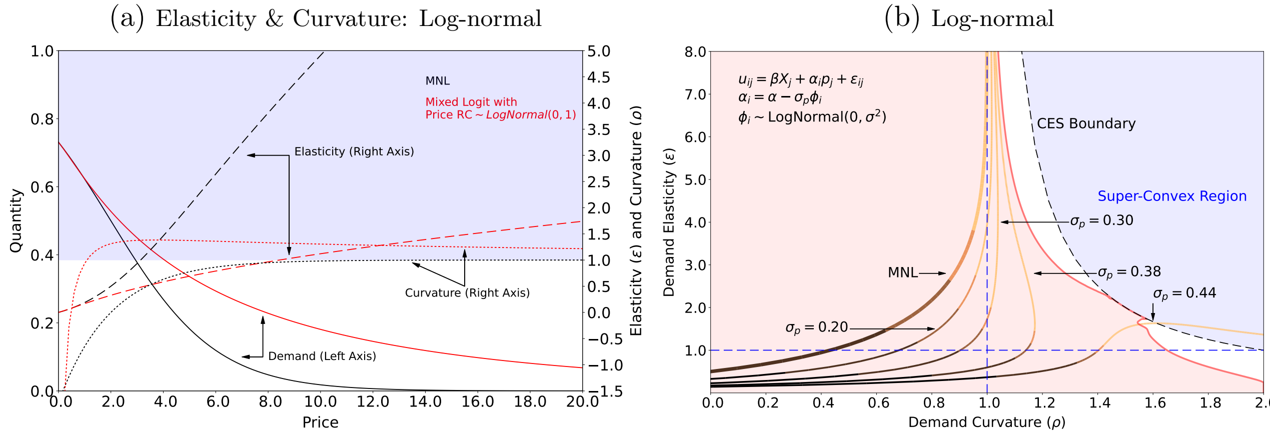
The right panel of Figure 4 illustrates that while a normal price random coefficient accommodates some log-convex demands, the range of log-convex (ε, ρ) combinations is limited. For large values of σ_p, demand manifolds are upward sloping until they cross the *CES* locus. Constraining demand to be sub-convex limits the role of idiosyncratic price responses in preferences, as the admissible values of σ_p are small. The symmetric normal distribution includes both positive and negative deviations from the mean price sensitivity α. Thus, unless the extent of heterogeneity in price sensitivity is limited, the model has to accommodate an increasingly large share of individuals with upward-sloping demands. The feasible log-convex elasticity-curvature combinations are thus frequently characterized by a high elasticity of demand, effectively minimizing instances of upward-sloping demand.

The utilization of a one-tailed log-normal distribution introduces skewness (Equation 18) and expands the scope for more prominent differences in price sensitivity and curvature; the right

[11] Davis (2005) first addressed this behavior of demand elasticity estimates in discrete choice models. Chintagunta (2002) documented empirically that demand elasticity is quasi-linearly increasing in price in *ML* models while Björnerstedt and Verboven (2016) attributed this property to the linearity of conditional utility in price.

[12] This hints at demand manifolds becoming downward sloping at some price level in the log-convex region for the log-normal mixture but not necessarily so for the normal distribution.

Figure 5: Demand Manifolds: Log-normal Mixing Distribution



Notes: Panel (a) contrasts quantity, elasticity, and curvature under *MNL* in black and *ML* in red. Panel (b) represents demand manifolds in the (ε, ρ) plane. Light-shaded regions represent all feasible (ε, ρ) pairs conditional on the price-mixing distribution.

panel in Figure 5 shows larger values of σ_p continue to generate sub-convex demand. This results in a much larger range of feasible curvatures for a given demand elasticity, in particular for less elastic demands where firms enjoy more market power. Figure 5 hence shows that a model with a log-normal price random coefficient can admit most well-behaved curvature-elasticity pairs in the sub-convex region of demand, with the exception of a small set of (ε, ρ) combinations close to the *CES* locus. As these are part of the feasible region of the specification with a normal price random coefficient, we explore the ability of the generalized normal distribution as a mixture between a normal and log-normal distribution to extend the set of (ε, ρ) pairs; see Appendix C.

4.3 Demographic Interactions and Demand Curvature Estimates

In empirical applications, researchers rely on the fact that idiosyncratic price responsiveness is correlated with demographics. Rather than imposing a distribution on idiosyncratic price sensitivities, as we did above, one might therefore specify the idiosyncratic price sensitivity α_i as a function of an observable demographic d_i, i.e., $\alpha_i^\star = \alpha + \pi_d d_i$. The equivalence to the analysis of Section 3 is apparent: it is now the empirical distribution of demographic d_i that underlies measure $G(i)$ in the manifold expression (3) and that determines the feasible combinations of (ε, ρ) pairs that the demand system can accommodate. In Section 6.4, we consider how to relax the assumption that d_i linearly shifts price sensitivity by allowing the data to determine a flexible relationship between the demographic attribute and price sensitivity.

4.4 Summary

The analysis in this section indicates that a quasi-linear discrete choice demand model that incorporates flexible heterogeneity in consumer preferences for product attributes and, notably, in price sensitivities does not impose substantial ex-ante restrictions on the curvatures and elasticities the model can accommodate. In particular, the quasi-linear model is capable of accommodating curvature-elasticity pairs all the way up to, and including, those observed in the *CES* demand model.

5 Beyond Quasi-Linear Preferences

Quasi-linear preferences may be appropriate for representing the demand for low-priced products where income effects are likely small. For products like the original car application in *BLP*, the purchase price accounts for a substantial portion of consumer income, however. The indirect utility specification in *BLP* accommodates income effects by incorporating a nonlinear function of outside good spending into preferences. In this section, we explore the implications of this specification for the demand model's ability to encompass the full range of curvatures and elasticities associated with sub-convex demand.

5.1 Flexible Income Effects

In contrast to the quasi-linear case where outside good spending enters consumers' indirect utility linearly, *BLP* specifies the preferences in Equation (6) with the following price sub-function:

$$f_i(y_i, p_j) = \alpha \ln \left(y_i - p_j\right). \tag{19}$$

Both the quasi-linear price sub-function and *BLP*'s alternative are, however, special cases of a Box-Cox power transformation (Box and Cox, 1964) of outside good spending, which is consistent with utility maximization in discrete choice contexts for any value of parameter $\lambda \in \mathbb{R}$ driving the convexity or concavity of the transformation. We, therefore, specify the generalized price sub-function,

$$f_i(y_i, p_j) = \alpha_i^\star \left(y_i - p_j\right)^{(\lambda)} = \begin{cases} \alpha_i^\star \dfrac{\left(y_i - p_j\right)^\lambda - 1}{\lambda}, & \text{if } \lambda \neq 0\,, \\ \alpha_i^\star \ln\left(y_i - p_j\right), & \text{if } \lambda = 0\,, \end{cases} \tag{20}$$

and explore how the value of the power parameter λ affects demand elasticity (10), curvature (11), and the shape and position of the manifold (12) through its effect on f'_{ij} and f''_{ij} in Equation (7). In line with the *BLP* specification, we abstract from heterogeneity in price sensitivity and consider the case of $\alpha_i^\star = \alpha$. A power parameter of $\lambda = 0$ thus yields the *BLP* model, while a power parameter of $\lambda = 1$ results in a *MNL* model. This means that the income distribution captures any idiosyncratic price responsiveness across individuals, modulated byλ.[13]

[13] It is worth comparing our setup to the multiunit demand model of Birchall and Verboven (2022) who rely on a different Box-Cox transformation in the price subfunction, $f(y_i, p_j) = \gamma^{\lambda-1}\left(y_i^\lambda - 1\right)\lambda - \left(p_j^\lambda - 1\right)\lambda$. The associated conditional demand function of $q_{ij} = \left(\gamma y_i / p_j\right)^{1-\lambda}$ is a nonlinear function of the fraction of the allocated share of

As in Berry, Levinsohn and Pakes (1999), we adopt a first-order Maclaurin series approximation (at $p_j = 0$) of the Box-Cox transformation:[14]

$$f_i(y_i, p_j) = \alpha\big(y_i - p_j\big)^{(\lambda)} \simeq \alpha y_i^{(\lambda)} - \frac{\alpha p_j}{y_i^{1-\lambda}}\,. \tag{21}$$

As the first term in this sub-function does not vary across products j, the marginal effect of price p_j on indirect utility is again constant with $f'_{ij} = -\alpha(y_i)^{\lambda-1}$ and $f''_{ij} = 0$. The resulting demand elasticity and curvature are:

$$\varepsilon_j(p) = -\frac{p_j}{Q_j(p)} \int\limits_{i\in\mathcal{I}} -\frac{\alpha}{y_i^{1-\lambda}} \cdot \sigma_{ij}^2 \, dG(i)\,, \tag{22a}$$

$$\rho_j(p) = \int\limits_{i\in\mathcal{I}} \mu_{ij}\, dG(i) \times \frac{\int \frac{(1-\lambda)\alpha^2}{y_i^{2-\lambda}} \cdot \sigma_{ij}^2 \, dG(i) \; + \int \frac{\alpha^2}{y_i^{2(1-\lambda)}} \cdot sk_{ij}\, dG(i)}{\left[\int -\frac{\alpha}{y_i^{1-\lambda}} \cdot \sigma_{ij}^2 \, dG(i)\right]^2}\,, \tag{22b}$$

yielding a demand manifold of:

$$\rho_j[\varepsilon_j(p)] = \frac{p_j^2}{\varepsilon_j^2(p)\cdot Q_j(p)} \cdot \int\limits_{i\in\mathcal{I}} \frac{\alpha^2 \cdot \big[(1-\lambda) y_i^{-\lambda} \cdot \sigma_{ij}^2 \, + sk_{ij}\big]}{y_i^{2(1-\lambda)}}\, dG(i)\,. \tag{23}$$

In Figure 6, we plot the demand manifold under various power parameters λ, assuming as above that the valuation of the product attribute, βX_j, equals one and price sensitivity λ equals 0.5. We rely on a log-normal approximation to the U.S. income distribution in our representation of y_i. The figure illustrates that, as in the case of the quasi-linear utility with flexible idiosyncratic price sensitivities, accommodating income effects via the approximate Box-Cox transformation of outside good spending yields, for a power parameter between zero and one, preferences that can accommodate curvatures close to those of the *CES* boundary.[15]

To provide some initial empirical context for the role of the Box-Cox power transform in shaping elasticity and curvature of *ML* demand, we conduct a similar analysis to the one in the

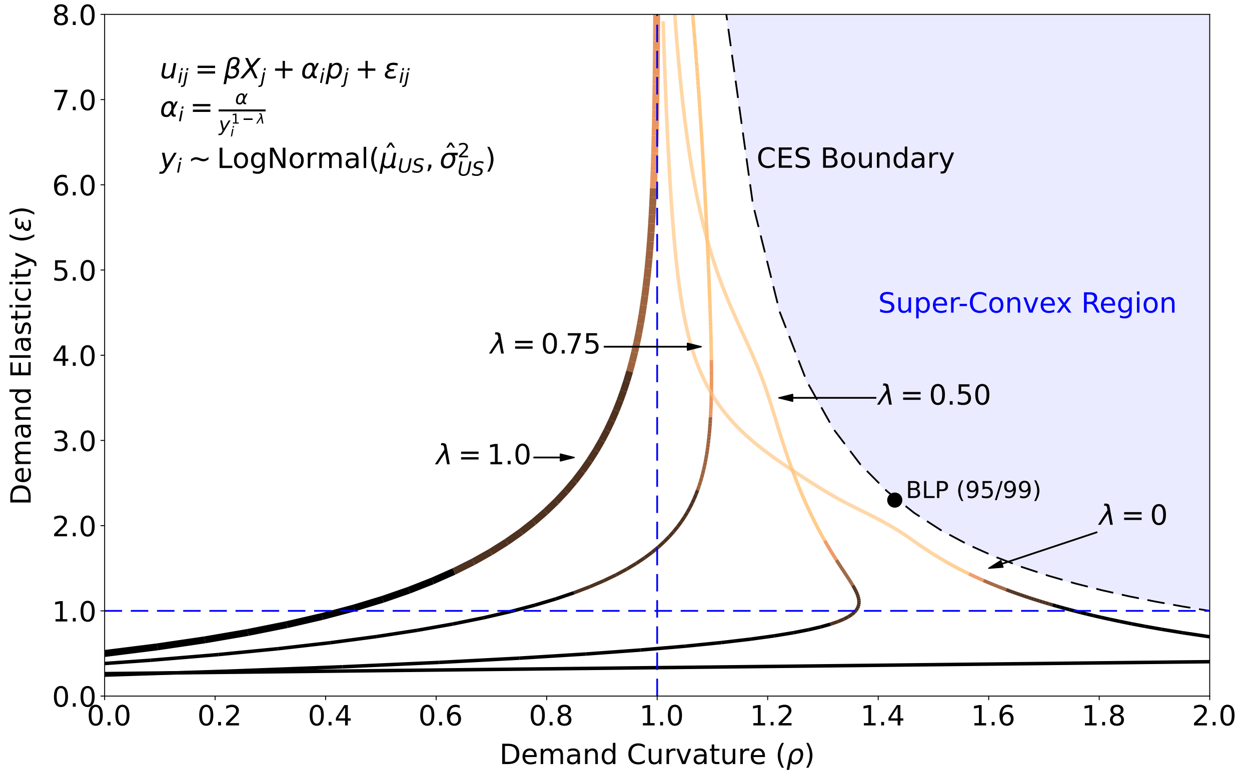
---

income, γ, spent on a chosen product. Their transformation is an *h-function* bridging *MNL* and *CES* demands, e.g., Nocke and Schutz (2018, Proposition VII, Appendix VI.1), which Anderson and de Palma (2020, §5.4) show to be well-defined for $\lambda \in (0, 1)$. Curvature flexibility thus results from a hybrid combination of these two demand models, but disappears when the specification reduces to the quasi-linear unit-demand case when $\lambda = 1$. Our goal in specifying subfunction (20) is instead to allow for curvature flexibility within the confines of a unit-demand setup consistent with utility maximization (e.g., Roy's identity holds for $q_{ij} = 1$). Our Box-Cox transformation parameter can take any real value, accommodating stronger or weaker income effects and hence curvature flexibility.

[14] Note that for $\lambda = 1$, Equation (21) again leads to the *MNL* model, but for $\lambda = 0$, the price sub-function becomes $\alpha \ln y_i - \alpha p_j / y_i$, which only coincides with (20) for $y_i = 1$. The preference specification based on Equation (21) is hence only approximately consistent with utility maximization.

[15] While we consider a power parameter $\lambda \in [0, 1]$, in line with the empirical literature, Box and Cox (1964) consider $\lambda \in [-5, 5]$, which would expand the range of feasible curvature elasticity pairs beyond the ones depicted in Figure 6.

Figure 6: Box-Cox Transformation and Demand Manifolds



Notes: Demand manifolds for different values of the Box-Cox transform parameter λ using the U.S. income distribution and the rest of the model specification of Berry et al. (1999). The dot identified as "BLP (95/99" corresponds to the average estimated curvature and elasticity value using the *BLP* automobile data and estimation best practices as outlined in Conlon and Gortmaker (2020).

introduction, where we display the elasticity and curvature combinations of two alternative models of demand for breakfast cereal in the spirit of Nevo (2001). Now, we rely on the automobile data from Berry, Levinsohn and Pakes (1995) to illustrate the elasticity and curvature properties of a *ML* model with income effects modulated by the power parameter λ, contrasting the original *BLP* specification ($\lambda = 0$) with a quasi-linear specification with a common price sensitivity ($\lambda = 1$) and two in-between cases ($\lambda = 0.5$ and $\lambda = 0.7$). We estimate four separate sets of preferences holding λ fixed at each value, otherwise following *BLP* in choice of specification and identification strategy. Figure 7 shows the scatter plots of $(\hat{\varepsilon}, \hat{\rho})$ for each automobile model in the *BLP* data under these four alternative specifications.

The top left panel represents the quasi-linear case. The average estimated automobile demand elasticity is $\hat{\varepsilon} = 2.75$ with nearly full (single-product) pass-through, $\hat{\rho} = 0.99$, as any mixed *MNL* without idiosyncratic price sensitivity is necessarily log-concave, as shown in Section 4.1. Note also the sorting of automobiles by price: the estimated demand is substantially more elastic for the most expensive vehicles.

The demand estimates are log-convex for all automobile models whenever we allow for some income effects, as shown in the other three panels of Figure 7. Reducing λ increases the

Figure 7: Income Effects and Demand Manifolds

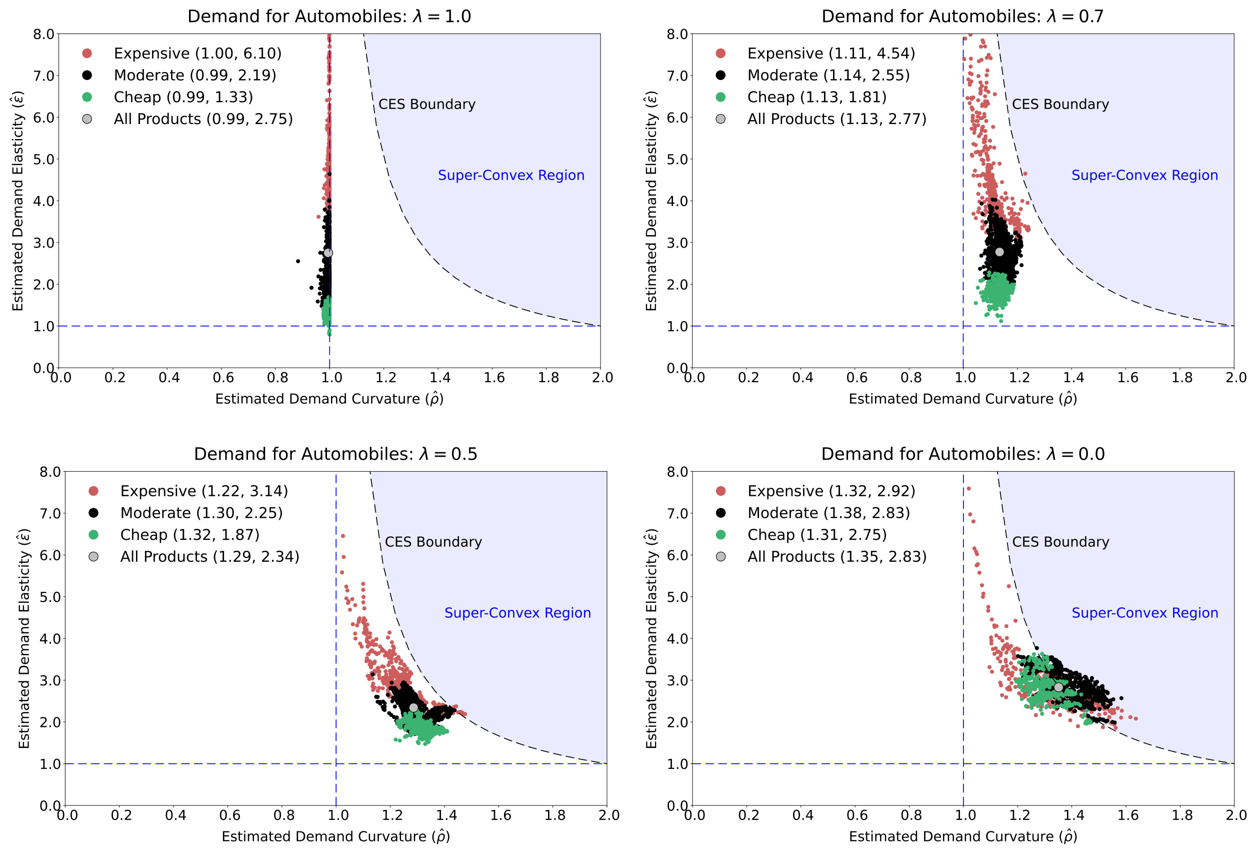


Figure Notes: Each dot represents the point elasticity and curvature estimates for each observation in the *BLP* automobile data, while the gray dot corresponds to the average elasticity and curvature estimates. "Expensive" and " Cheap" products are defined as vehicles with average prices in the top 20% and bottom 20%, respectively. We define the remaining vehicles "Moderate."

importance of income effects through smaller price responses by higher-income households. Moving from quasi-linear demand to demand with income effects does not change the average estimated elasticity significantly, reaching $\hat{\varepsilon}_{\text{BLP}} = 2.83$ when $\lambda = 0$. Despite the similar average price elasticity, the distribution of curvature (passthrough) varies substantially across specifications. This is similar to what we observed in the motivating RTE cereal case. Relative to the quasi-linear specification, the expensive (cheap) market segment is much less (more) competitive under the *BLP* model.

Curvatures decrease monotonically with λ, with $\hat{\rho} = 0.99$ when $\lambda = 1$ to $\hat{\rho}_{\text{BLP}} = 1.35$ when $\lambda = 0$ (which, in this case, coincides with the curvature of the *CES* model evaluated at the average elasticity: $\hat{\rho}_{CES} = 1 + 1/2.83 = 1.35$). Average pass-through rates thus increase from 99% in the quasi-linear specification without income effects to 179% with the strong income effect specification of *BLP* demand – dramatically different predictions. We report estimates of average elasticity, curvature, price markup, and pass-through rate for each scenario in Table D.1 in Appendix D. The intermediate cases of $\lambda = 0.5$ and $\lambda = 0.7$ make clear that income effects broadly not only restrict the range of demand elasticity (and markup) estimates but also expand the range of demand curvature (and pass-through rate) estimates that a discrete choice model of demand can deliver.

5.2 Summary

The preceding sections demonstrate that the *ML* model exhibits significant flexibility, not only in capturing realistic substitution patterns but also in generating a wide range of cost pass-through when we allow for heterogeneity in both consumer valuations for product attributes and sensitivity to price. Idiosyncratic attribute valuations give firms localized market power, leading to under-shifted pass-through, while consumer heterogeneity in price sensitivity entails over-shifted pass-through. The combined effect of these two forces drives a given product's pass-through. The above examples also demonstrate, however, that the exact specification of heterogeneity in price sensitivity has important consequences for the economic outcomes of interest. This raises the question of identifying a flexible specification for price sensitivities and, consequently, demand curvature in an empirical setting, the topic of the next section.

6 Empirical Applications

We now discuss the estimation and identification of flexible demand curvature in an empirical setting. We begin by providing broad intuition for the link between the distribution of price sensitivity and demand curvature. We rely on this intuition to propose an identification strategy that exploits heterogeneous consumer responses to exogenous price changes and show its success at identifying flexible price sensitivities in *ML* models with and without income effects using Monte Carlo evidence. As our price sensitivity specification nests the standard *ML* models, we also show the consequences of model mis-specification through Monte Carlo simulation and an empirical application from ready-to-eat cereal.

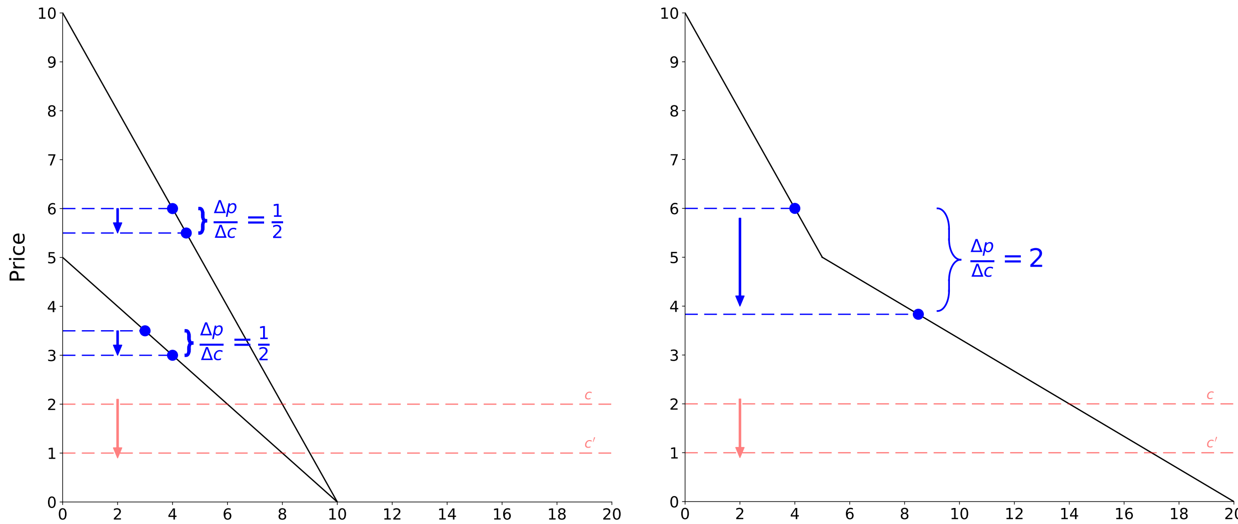
6.1 Heterogeneous Price Sensitivity and Pass-Through Over-shifting: Intuition

We have thus far taken heterogeneity in price responsiveness as a demand primitive that implies both over-shifted pass-through and a wide range of feasible elasticity-curvature combinations. Product demand is, however, not only a function of consumer preferences but also of market definition – which is often controlled by the firm. We, therefore, start this section by presenting a straightforward example of pricing by a monopolist who caters to two consumers with different demand profiles. The purpose of this illustration is twofold: first, to illustrate scenarios wherein firms may choose more than complete cost pass-through, and second, to provide intuition for the instrumentation strategy we employ to identify empirically a flexible distribution of price sensitivity.

Figure 8 depicts, in the left panel, the demands of the two consumers with simple linear demands of different slopes. Consider first the case where the monopolist can set individual prices for each consumer. The left panel illustrates the well-known result of monopoly under-shifting: in response to a drop in marginal cost from \$2 to \$1, the monopolist lowers each consumer's price by \$0.5.

Figure 8: Pass-through implications of Targeted and Uniform Monopoly Pricing

(a) Price Discrimination (b) Uniform Pricing



In many empirical settings, including the ready-to-eat cereal context we consider below, firms do not practice such perfect price discrimination. Consider therefore the case where the monopolist charges a single price to both customers. The right panel shows that in setting the uniform price, the monopolist now faces a kinked demand.[16] At the initial marginal cost of \$2, the monopolist's optimal price excludes the low-valuation consumer from the market: serving only the high-valuation consumer is more profitable. Once marginal cost drops to \$1 though, setting a price that induces both consumers to buy becomes more profitable; the optimal price drops from \$6 to \$4, and pass-through is over-shifted to induce the price-sensitive consumer to buy.

More generally, in responding to a drop in cost, a firm serving heterogeneous consumers with a uniform price trades off the standard incentive to remain on the elastic portion of demand and the benefits of drawing in a larger, more price-sensitive, customer-base when costs fall. The right panel illustrates that, for a given change in cost, the firm's choice of cost pass-through depends on the initial price level: had we started from a higher initial marginal cost and hence, a higher initial uniform price, the firm would continue to serve only the high-valuation consumer when its cost drops by \$1, reducing the price by \$0.5 as in the left panel. Our identification strategy exploits this idea that for demands outside the family of iso-convex demand (which includes the linear demand as a special case; see Mrázová and Neary, 2017), the pass-through of cost shocks differs at different price levels and/or across different demographic groups.

[16] In an influential paper, Kimball (1995) first suggests a smooth differentiable version of this kinked demand to ensure subconvexity and markups increasing with the scale of production in macro models.

6.2 Instruments to Connect Demand Manifolds and price sensitivity

The previous discussion suggests that an interaction between cost shifts and price levels can serve as an instrument to recover the shape of the distribution of consumers' price sensitivities and hence, the curvature of a unit demand function. Empirically, one could represent such price sensitivities with a flexible function of consumers' demographics, a flexibly distributed price random coefficient, or nonlinear income effects through the Box-Cox power parameter.[17]

A challenge, of course, that arises in taking this idea to data, is the endogeneity of prices in an oligopoly equilibrium: unobserved demand shocks ξ may confound the response in price to a change in cost ω. We address this issue here by constructing exogenous price predictions via a reduced-form hedonic price regression based on exogenous characteristics x_t and cost shocks ω_t:[18]

$$p_t = \gamma_0 + \gamma_1 x_t + \gamma_2 \omega_t + u_t \,. \tag{24}$$

We run the above regression and use the results to construct the vector of predicted (exogenous) prices $\hat{p}_t$. We then follow Gandhi and Houde (2020) and construct differences in price-space between product j and its competitors:

$$Z^p_{jt} = \sum_r \left(\hat{p}_{rt} - \hat{p}_{jt} \right)^2 . \tag{25}$$

Equation (24) enables us to construct exogenous prices by separating price effects due to changes in demand (via ξ) from changes in cost (via ω). It is also a simple pass-through regression. Cost pass-through, therefore, informs the identification of demand primitives related to curvature using $\hat{\gamma}_2$ via the substitution patterns captured in equation (25). Since curvature in quasi-linear and income-effect discrete choice models comes through heterogeneity in price-sensitivity, equation (25) identifies the price random coefficient (σ_p) in a quasi-linear utility model and our proposed Box-Cox income transformation (λ) in a consumer model with income effects. The instrument traces the demand manifolds using cost shocks, holding constant exogenous demand shifters at different price levels. Interacting this instrument with observable demographics identifies the case when price-sensitivity is correlated with the same demographics.

6.3 Flexible Manifold Estimation: Monte Carlo Analysis

We now conduct a Monte Carlo analysis to demonstrate the validity of our identification strategy and evaluate the potential for mis-specified demand systems to introduce biases in economic outcomes of interest – elasticity and curvature. We focus on a specification of preferences with income effects specification but also apply the instrument in an empirical application with quasi-linear

[17] This is the same argument used long ago in the field of transportation to account for decreasing marginal utility of travel and compare the benefits of a given reduction time for commuting trips of very different length (Gaudry and Wills, 1978; Koppelman, 1981).

[18] Alternatively, one could construct prices non-linearly using firm first-order conditions as in Berry et al. (1999).

demand below. Consider a setting with $J=20$ differentiated products sold by single-product firms competing in prices for $T=50$ periods. Consumer indirect utility takes the following form:

$$u_{jlt} = \underbrace{\beta_0 + \beta_1 x_{jt}^1}_{\substack{\text{Common Across} \\ \text{Consumers}}} + \underbrace{\sum_{k=1}^{K} \left(\beta_{2,k} + \sigma_{X,k}\nu_{ik}\right) x_{jt,k}^2}_{\substack{\text{Idiosyncratic} \\ \text{Characteristic Tastes}}} - \underbrace{\alpha \cdot p_{jt} \cdot y_{it}^{\lambda-1}}_{\substack{\text{Idiosyncratic} \\ \text{Price Sensitivities}}} + \xi_{jt} + \epsilon_{ijt}\,, \tag{26}$$

where, as above, income effects decrease as λ moves from zero to one. In this specification, some product characteristics are observed by the researcher ($\{x_{jt}^1, x_{jt}^2\}$) while others are only observed by consumers and firms (ξ_{jt}). Valuation of the product attribute x_{jt}^1 is common across individuals and we draw x^1 independently from a uniform distribution. We model consumer preference heterogeneity in product characteristics via x_{jt}^2 with two elements ($K=2$) including a constant and a uniformly-distributed product characteristic. As in Gandhi and Houde (2020), product attributes (other than the constant) vary across time.[19] Consumers, therefore, have preference heterogeneity over the J inside goods as a category, via the constant random coefficient, and over variation in the observable product characteristic across the J products and T time periods. We set $\beta_2=1$ and $\sigma_X = 5$ for $k=1,2$. We assume that the unobservable characteristic ξ_{jt} is distributed standard normal. We model heterogeneous price sensitivity using the above approximation to the Box-Cox transformation of outside good spending modulated by parameter λ. We assume that consumer income y_{it} is drawn from a log-normal distribution and parameterize these draws following Andrews, Gentzkow and Shapiro (2017), generating market (e.g., time) variation in these draws by allowing the variance of income to vary.

Single-product firms choose prices simultaneously each period given their constant marginal costs c_{jt}. In the static oligopoly Bertrand-Nash equilibrium, period t equilibrium prices $p_t^\star$, satisfy the set of J first-order conditions for the firms:

$$p_{jt}^\star = c_{jt} - s_j(\delta_t, p_t^\star; \sigma_X, \sigma_p) \times \left[\frac{\partial s_j(\delta_t, p_t^\star; \sigma_X, \sigma_p)}{\partial p_{jt}^\star}\right]^{-1}. \tag{27}$$

Marginal costs are a function of product characteristics and cost shocks:

$$\log c_{jt} = \gamma_0 + \gamma_1 \log x_{jt}^1 + \gamma_2 \log x_{jt}^2 + \omega_{jt} + \zeta_{jt} \tag{28}$$

We set all γ parameters equal to 1 and draw cost shocks $\{\omega_t, \zeta_t\}$ from standard normal distributions. The researcher observes ω_t which provides identification for the distribution of price sensitivity. We generate pricing equilibria in the true data-generating processes by selecting α and β_0 so that the average own-price elasticity is 2.5 with a 20% aggregate inside share for each simulation.

[19] In empirical applications, such as automobiles, this is due to product remodels, which the researcher treats as exogenous to unobserved variation in demand via ξ. For our purposes, this is equivalent to allowing for exogenous product entry and exit – a common assumption in the empirical literature.

We consider the objective of a researcher who estimates consumer demand given observed prices, quantities, and ω cost shocks following the best practices outlined in Conlon and Gortmaker (2020). The researcher also specifies the supply side as in Berry et al. (1999) and correctly specifies the outside option as well as the distribution generating the random coefficients for product characteristics ν_i. The researcher, however, *may incorrectly model* income effects and hence, the distribution of price-sensitivities, as in Section 4.3. The goal of this Monte Carlo analysis is to investigate the success of an empirical demand model with a flexible Box-Cox power transformation of outside good spending at identifying and recovering the true demand curvature underlying the data-generating process.

We consider three data-generating processes: we simulate demand and cost data assuming that (1) $\lambda = 0$, as in the original *BLP* specification; (2) $\lambda = 1$, resulting in quasi-linear demand; and (3) $\lambda = 0.7$, an in-between case with weaker income effects than case (1): the distribution of α_i is compressed, with a coefficient of variation of only 0.56, relative to 3.57 for the case of $\lambda = 0$. In the following, we denote case (1) as 'log'; case (2) as 'linear'; and case (3) as 'box-cox' or 'bc'.

With these three data sets, we then estimate seven specifications. In scenarios (1)-(3), we specify the demand model correctly and verify that we can recover the underlying preferences. In scenarios (4) and (5), we specify general 'box-cox' preferences to recover the simpler 'log' and 'linear' preferences. Lastly, in scenarios (6) and (7) we investigate model mis-specification by using either a 'log' or a 'linear' demand model in estimation to recover 'box-cox' preferences.

Identification. We identify the characteristic random coefficients (σ_X) using the Differentiation IVs of Gandhi and Houde (2020). In Scenarios (3) through (5), where the researcher estimates demand allowing for flexibility in the price sensitivity distribution via λ, we include the interactions of the price differentiation instrument of Gandhi and Houde (2020) with moments of the income distribution:

$$Z_{jt}^P = \sum_r \left(\hat{p}_{rt} - \hat{p}_{jt} \right)^2, \tag{29a}$$

$$Z_{jt}^D = Z_{jt}^P \otimes \left\{ \text{inc}_t^{10\%}, \text{inc}_t^{50\%}, \text{inc}_t^{90\%} \right\}. \tag{29b}$$

Discussion of Results. We present the parameter estimates in Table 1 for seven distinct scenarios. In general, across curvature targets, the estimation succeeds at recovering the underlying parameters when the researcher's preference specification coincides with the true underlying data-generating process, i.e., Scenarios (1)-(3), consistent with Gandhi and Houde (2020) and Conlon and Gortmaker (2020). The estimates of elasticity (market power), curvature (pass-through), and their correlation are consistent with the true quantities in the data.

In Scenarios (4) and (5), the researcher models consumer price-sensitivities flexibly using a Box-Cox transformation of outside expenditure and estimates the income parameter λ. The

Table 1: Monte-Carlo: Parameter Estimates

Scenario	α (varies)		λ (varies)		$\sigma_x = 5$		$\sigma_0 = 5$		*Coeff.Var*		*MAB*		*Corr.*	
True-Specification	*A.Bias*	*RMSE*	*A.Bias*	*RMSE*	*A.Bias*	*RMSE*	*A.Bias*	*RMSE*	σ_α/α	$\hat{\sigma}_\alpha/\hat{\alpha}$	ε	ρ	(ε,ρ)	$(\hat{\varepsilon},\hat{\rho})$
1: log–log	0.003	0.161	0.000	0.000	-0.006	0.072	-0.012	0.231	-3.81	-3.79	0.00	0.00	0.66	0.66
2: linear–linear	0.001	0.011	-	-	0.015	0.090	-0.082	0.947	0.00	0.00	0.00	0.00	0.66	0.66
3: bc–bc	0.000	0.037	-0.001	0.024	0.006	0.079	-0.001	0.735	-0.57	-0.57	0.00	0.00	-0.47	-0.47
4: log–bc	0.331	0.379	0.005	0.006	-0.012	0.070	0.025	0.121	-3.81	-3.77	0.00	0.00	-0.47	-0.47
5: linear–bc	-0.031	0.048	-0.060	0.085	0.006	0.091	0.093	1.109	0.00	-0.11	0.00	-0.01	-0.44	-0.43
6: bc–log	-15.514	15.612	-	-	0.851	0.947	-2.211	2.218	-0.57	-3.77	0.55	-0.69	-0.44	0.63
7: bc-linear	0.248	0.248	-	-	0.015	0.091	-0.272	0.987	-0.57	0.00	-0.16	0.22	-0.44	0.43

Notes: The first column indicate the true data-generating process and the researcher's assumed specification of the price-income interactions. The next three (double) columns report the average bias (*A.Bias*) and root mean standard error (*RMSE*) of the income parameter λ and drivers of the idiosyncratic characteristics tastes using $1,000$ replications for each scenario. The price coefficient, α, varies for each replication to ensure that $\varepsilon = 2.5$. The attribute random coefficients σ_x and σ_0 (constant) are both set to 5. Column "*Coeff.Var*" reports the coefficient of variation of the distribution of price responsiveness of the data-generating process as well as of the estimated model. The remaining set of columns report the coefficient of variation for idiosyncratic prices-sensitivity parameters (α_i), the median average bias (*MAB*) for average product elasticity and curvature (ε, ρ), and the average correlation between product-level elasticity and curvature (corr(ε_j, ρ_j)).

estimates of the Box-Cox model accurately estimate λ and the random coefficients of product attributes when the underlying preferences include a logarithmic function of income, although it overestimates the average price responsiveness α. We also observe that Box-Cox model accurately recovers the distribution of price-sensitivity (columns labeled 'Coeff. Var'), as well as the elasticity-curvature pairs.

Scenarios (6) and (7) address mis-specification biases of imposing particular price-income interactions when the true data-generating process is Box-Cox. Scenario (6) assumes the logarithmic transformation of outside good spending, while Scenario (7) assumes quasi-linear preferences of Nevo (2001). The assumed logarithmic specification leads to a particularly large mis-specification bias in all estimated parameters. The large positive average bias for the random coefficients on the characteristic, σ_x, leads to greater substitution within inside products than the true data, while the average bias of -2.2 for the constant random coefficient indicates greater substitution to the outside option than the true data. Not surprisingly, the economic implications are significant as the average estimated elasticity is -1.95, or 0.55 points less elastic than the true data-generating process, while he average estimated curvature is 0.69 points above the true data-generating process. The researcher, therefore, would tend over-estimate both market power and pass-through. Moreover, specifying log preferences ex ante amounts to imposing a different rate of change of the demand elasticity with income from the true relationship under Box-Cox preferences, leading to much greater heterogeneity in price sensitivity than the underlying data. Such a bias has consequences for welfare calculations, especially since solving for changes in consumer surplus requires accounting for income effects. If the researcher assumes that preferences are quasi-linear, instead, as in scenario (7), the estimated elasticity of -2.66 understates firms' true market power while the estimated curvature is 0.22 points below the true data indicating the estimated model will under-predict the firm pass-through.

The final two columns of Table 1 demonstrate that mis-specification impacts the distribution of estimated elasticity-curvature pairs among products. Looking across the different data-generating processes we observe that the shape of the distribution of price sensitivities, via the income distribution, determines the relationship between demand elasticities and curvature, i.e., the demand manifold. Imposing specific distributions of price sensitivities – Scenarios (6)–(7) – results in a flipped sign of the correlation between product-level elasticities and curvatures, or the slope of the manifold, leading to a mischaracterization of the relationship between market power and pass-through among the products. This could have large consequences for the evaluation of the economic effects of mergers, cost changes, taxation, or tariffs, particularly for different consumer and firm types.

Competition, Demand Curvature, and Pass-through. Our graphical illustrations of the demand manifold relied on the monopoly case, where the connection between demand curvature and pass-through is straightforward. This connection is less clear, however, in empirical settings where firms offer asymmetrically differentiated products, as in our empirical application below.

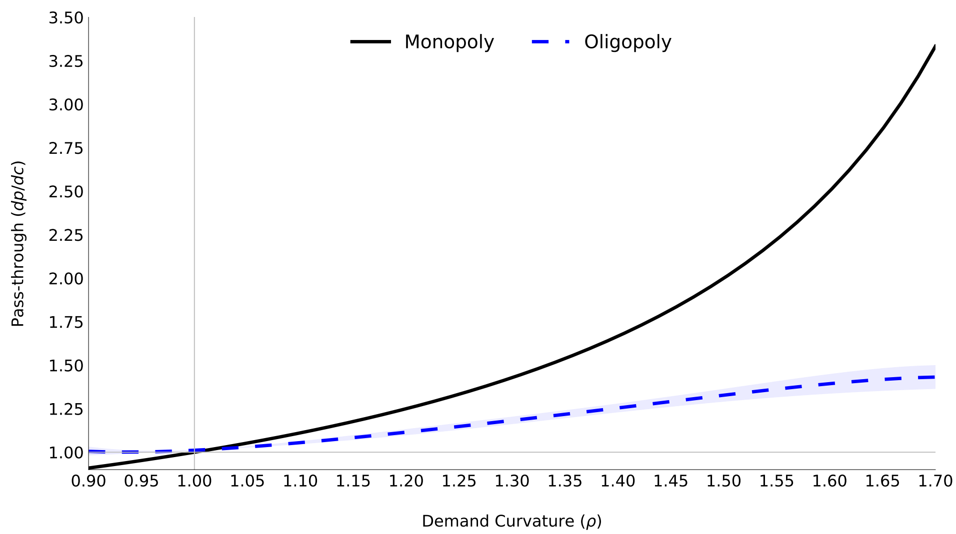
In this section, we use the Monte Carlo environment to compare monopoly pass-through (i.e., assuming, for each product, that $\frac{dp}{dc} = \frac{1}{2-\rho}$) with pass-through in our simulated 20-firm single-product environment. We use the Box-Cox indirect utility and vary λ to generate equilibria of varying degrees of demand curvature. For each simulated equilibrium, we calculate the average pass-through two ways: First, under the assumption that each firm's pass-through rate is that of a monopolist and under the actual market structure, and second by solving for equilibrium pass-through rates due to the introduction of a common 10% increase in marginal costs. We construct the average "oligopoly" pass-through rate as the simple average of equilibrium product pass-through rates. We then illustrate the effect of competition on the connection between pass-through and demand curvature by plotting the "monopoly" and "oligopoly" pass-through conditional on demand curvature (Figure 9).

Competition pushes equilibrium pass-through towards one thereby muting the upward pricing pressure generated by the change in marginal costs. The increase in the common cost leads to both direct and indirect pass-through effects. The price of a product always increases with its own cost. This is the direct effect captured by monopoly pass-through. The indirect effect collects substitution effects induced by price changes of other products similarly affected by the cost increase. The net effect thus depends on "how far" a particular product is from its closest substitutes in product space.[20]

The Monte Carlo evidence thus points to the advantages of employing a flexible specification of consumers' price sensitivity in recovering unbiased elasticities and curvatures from the data and illustrates the continued link between curvature and pass-through in settings beyond the

[20] See Footnote 4. Häckner and Herzing (2016) use these same arguments to show the difference in incidence in a monopoly and in a multi-product oligopoly. Frieberg and Romahn (2018) illustrate how the wedge between the two lines of Figure 9 varies with the ownership structure using the Swedish beer industry as a case study.

Figure 9: Competition and Pass-Through Rates



Notes: Figure presents Monte Carlo results across equilibria of median demand curvature. We generate each equilibrium following the environment discussed in Section 6.3 for the Box-Cox utility specification where $\lambda \in [0, 1]$. For each market t in each equilibrium, we solve for the median (across 20 products) demand curvature. "Monopoly" represents the pass-through rate of a single-product monopolist, e.g., (4). "Oligopoly" is the median pass-through rate for each market t in each equilibrium generated by a 10% increase in marginal costs. The shaded region reflects the 95% confidence interval.

monopolist cases we considered above. As competition attenuates pass-through towards complete pass-through, however, the importance of allowing for flexible pass-through estimation in assessing policy outcomes of interest is ultimately an empirical question. We, therefore, conclude this section with an application.

6.4 Flexible Manifold Estimation: Ready-To-Eat Cereal

In this section, we investigate the implications of specification bias by evaluating the consumer welfare implications of uniform pricing in the ready-to-eat cereal market. Recent empirical work (Adams and Williams, 2019; DellaVigna and Gentzkow, 2019; Hitsch, Hortaçsu and Lin, 2021) highlights the infrequent use of fine market segmentation strategies by retailers in similar consumer packaged goods despite the increasing availability of detailed data that might facilitate such practices. This work focuses on evaluating explanations for the lack of customized pricing, including frictions, such as managerial costs to optimizing pricing, and the more limited profit gains to segmentation under oligopoly.

Our objective here is different: we aim to assess the contribution of demand specification to conclusions about the consumer welfare consequences of targeted pricing, and the role of the estimated curvature therein. Aguirre et al. (2010) establish the connection between aggregate welfare gains from third-degree price discrimination and curvature, building on work by Robinson (1933) and others which showed that third-degree price discrimination enhances aggregate welfare only if it increases aggregate output.

Consider the case of a single-product monopolist who sells to two different markets but produces with a common cost function. Let $\overline{p}$ denote the uniform price and $\{p_w, p_s\}$ the profit-maximizing market-specific prices for the "weak" and "strong" markets, respectively, where $p_w < \overline{p} < p_s$. Aguirre et al. (2010) show that uniform pricing increases consumer welfare if the demand function in the strong market is at least as convex as that in the weak market at the uniform price, or, in terms, of curvature, $\rho_w(\overline{p}) < \rho_s(\overline{p})$. Moreover, the marginal welfare effect of uniform pricing is decreasing in the difference between demand curvature in the "weak" and "strong" markets. This indicates that the distribution of demand curvature, particularly its variance, is an important ingredient in the evaluation of the welfare implications of uniform pricing.

Data. We use scanner data from the marketing company IRI for the period of 2007 to 2011. For a set of cities, we observe cereal revenue and price at the universal product code, store, and week, together with brand name, parent company, and package size, as well as product characteristics such as the types of grain used to produce the cereal. For two markets, Eau Claire, WI and Pittsfield, MA, we observe consumer-level panel data on weekly grocery shopping trips and record cereal purchases and prices paid by an average of 3,700 consumers each week.

We restrict attention to products packaged in cardboard boxes (approximately 94% of total revenue) and a sales rank in the top 20% of products (approximately 95% of total revenue, allowing us to significantly reduce the product set from the original 1,022). We define a cereal product j as the combination of brand and flavor (e.g., Honey Nut Cheerios). After aggregating across product sizes, we obtain 41 products and construct a price per one-ounce serving for each by dividing total revenue by the total number of servings.

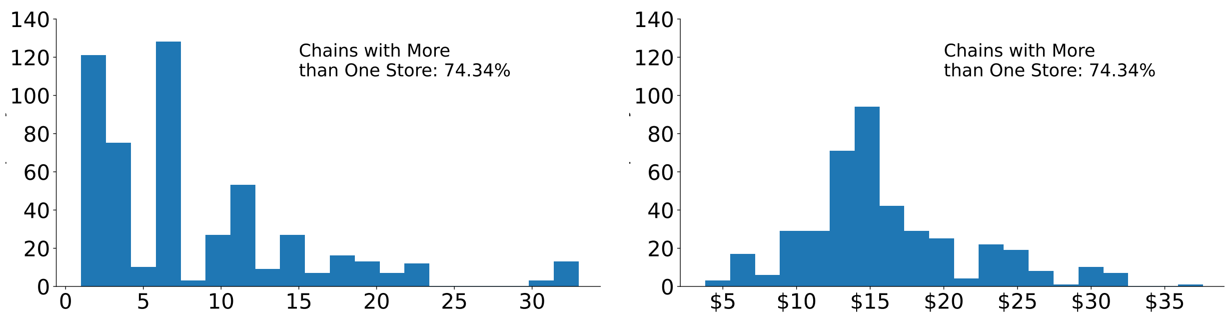
To reduce computational complexity, we focus on stores in large markets with significant geographic variation: Boston (5.2% of total revenue), Philadelphia (4.5%), Chicago (4.2%), San Francisco (3.0%), Seattle (2.5%), Houston (2.5%), and St Louis (2.4%), in addition to the micro data from Eau Claire and Pittsfield. Finally, we append to these data nutritional information (i.e., content of added sugar, calories, protein, fat, sodium, fiber, carbohydrate, potassium, and vitamins) attained via web-scrape. We also obtain time-series data for commodity costs of corn, oats, rice, wheat, and sugar-sweeteners (e.g., high-fructose corn syrup) from Quandl and the Federal Reserve Economic Data.

The raw store-level data include demographic information for customers living within a two-mile radius of each store location. We focus on the presence of children in the household and income. The demographic data for income is binned in discrete categories (e.g., the number of customers with annual incomes between \$25,000 to \$29,999). We fit the binned empirical distributions to beta distributions of the second kind to establish continuous income distributions in each market; Appendix E illustrates the ability of the assumed beta distribution to fit the observed categorical data. We use these fitted distributions, together with the share of households with children, to construct simulated consumers who vary in income and the presence of children in the household. We similarly generate a continuous income level from the recorded income categories for the households in the micro data; we observe the presence of children in each participating household directly.[21]

Motivating Evidence. We begin by addressing the prevalence of multi-store chains in the data. In Figure 10, Panel (a) we demonstrate that 74% of retailers in the data have more than one store and that there is significant heterogeneity in the size of chains as measured by number of stores within each chain. In Panel (b) we explore income variation across the stores. We exploit the store-specific income data and compare income across stores by calculating the standard deviation in average income across stores within a chain. For example, a chain that has two stores where each store is located in a geographic area with an average income of $50,000 will have a standard deviation equal to zero, while a chain with stores in low-income and high-income locations will have a positive standard deviation. We observe that chains do not appear to select locations of similar incomes. If consumer price sensitivity varied systematically with income, these differences in locations' incomes suggest heterogeneity in price sensitivity among consumers shopping at the chain. Moreover, this provides suggestive evidence that, conditional on uniform pricing, demand curvatures are likely to exceed one.

Figure 10: Evidence of Multi-Store Chains

(a) Number of Stores within a Chain (b) Income Variation within a Chain ($000)



DellaVigna and Gentzkow (2019) found that products in many consumer packaged good categories, including ready-to-eat cereal, are priced uniformly across stores within a chain. We see similar behavior in our sample. We first explore product selection in our sample and note that 49.1% of stores sell at least one unit of each of the 41 products at some point during each year, and 90.9% of stores carry (sell) at least 38 of the 41 products. This indicates that chains are not separating consumer types across stores by using different product selections.

We test for uniform pricing using the share of variation in prices explained by chain fixed effects – a similar test employed by Nakamura (2008), Hitsch et al. (2021), and DellaVigna and Gentzkow (2019). We do so by looping through products and for each product regressing the

[21] We ignore correlations between demographics, as the data do not report conditional distributions based on demographics.

average price over weeks in store s for product j, $\overline{p}_{sj}$, on chain and city fixed effects. In our data, we find median R^2 values, across products, for chain and city fixed effects of 0.72 and 0.31, respectively. As chain fixed effects explain a large share of the variation in the data, this suggests the presence of uniform pricing whereas the relatively minor role for city fixed effects suggests less importance pertaining to local market factors such as competition and consumer preferences.

Demand Specification. We abstract from store choice and represent consumer i's choice of which product j to purchase at store l in week t using the quasi-linear indirect utility in Equation (13). We include, as a product characteristic, the sugar content of product j and allow preferences for sugar content to vary with the consumer's observed demographics D_{il} (the presence of children) and income y_{il}, as well as an unobserved preference shifter ν_{il} that we assume to be distributed standard normal. We allow for the same heterogeneity in the valuation of the outside good to capture systematic differences across consumers in the overall taste for cereal, part of which may correlate with household demographics.

Given the role of heterogeneity in price sensitivity in shaping curvature, we introduce additional flexibility in how demographics enter the price coefficient $\alpha_i^\star$. We allow the presence of kids to shift $\alpha_i^\star$ linearly, but follow Nevo (2001) in accommodating a non-linear effect of household income on price sensitivity, as prior work has found sizable differences in price elasticities across low- and high-income consumers. Note, however, that in a quasi-linear model, such patterns do not represent income effects; they simply capture differences in purchase behavior by consumers of different income levels. There are a number of ways of introducing such flexibility in $\alpha_i^\star$. For example, one might allow price sensitivity to differ by income bin, or one could employ more complex methods such as sieve estimation (e.g., Wang, 2022). To keep the specification parsimonious, we instead leverage the Box-Cox transformation again. As in Equation (20), we thus allow the power parameter λ to reflect differences in price sensitivity between low- and high-income consumers:

$$y_{il}^{(\lambda)} = \begin{cases} \dfrac{y_{il}^{\lambda} - 1}{\lambda}\,, & \text{if } \lambda > 0\,, \\[2ex] \ln\left(y_{il}\right), & \text{if } \lambda = 0\,. \end{cases} \tag{30}$$

A nice feature of the Box-Cox transformation is that it nests common empirical applications. A power parameter of $\lambda = 1$ corresponds to a linear effect of income on price sensitivity and $\lambda = 0$ denotes the case of log income, but the transform can also accommodate a convex relationship between income and price sensitivity with $\lambda > 1$. The final price coefficient that we specify is

$$\alpha_i \star = -\exp(\alpha + \pi^p y_i^{(\lambda)} + \pi^k \mathbb{1}_i^{\text{kids}}). \tag{31}$$

where the exponential operator is useful to guarantee downward-sloping demand for all consumers.

We rely on this specification of preference heterogeneity, together with logit shocks ε_{ijlt} to consumer i's utility from product j, in calculating the probability that consumer i purchases product

j in market l in period t, s_{ijlt}, as in Equation (8). We derive aggregate demand for product j in each store location by integrating over the distributions of observable and unobservable consumer attributes D_{il}, y_{il}, and ν_{il}, denoted by $P_D(D_i)$, $P_y(y_i)$, and $P_\nu(\nu_i)$, respectively, and scaling the market share for product j in market l at time t with the market's size:

$$s_{jlt} = M_{lt} \int_{\nu_l} \int_{D_l} \int_{y_l} s_{ijlt} dP_y(y_i) dP_D(D_i) dP_\nu(\nu_i)\,. \tag{32}$$

In deriving the product's aggregate demand, we follow Backus, Conlon and Sinkinson (2021) in relying on an estimate of weekly store traffic as the potential market size M_{lt}. To capture variation in store traffic across weeks, we rely on milk and paper towel purchases, the two largest product categories in the IRI data, and project weekly cereal sales on weekly milk and paper towel purchases. We then scale predicted cereal purchases such that the predicted average outside option share across stores matches the share of shopping occasions that do not include cereal purchases in the IRI micro-data panel.

Estimation. We employ a standard Generalized Method of Moments (*GMM*) estimator. We partition the parameter space into (θ_1, θ_2) where the first set of parameters govern exogenous variables which enter consumer indirect utility linearly, while the second set, including λ, enter non-linearly. We augment data with the consumption micro-moments (Petrin, 2002; Berry, Levinsohn and Pakes, 2004) similar to Grieco, Murry and Yurukoglu (2021) and follow Backus et al. (2021) in estimating preferences without imposing supply-side aggregate orthogonality conditions. We also include brand and market (city-week) fixed effects to absorb unobserved brand and market characteristics. The structural errors are then brand-market-week demand shocks; e.g., an increase in demand for Cheerios in a given market.

The *GMM* estimator exploits the fact that at the true value of parameters $\theta^\star = (\Sigma^\star, \Pi^\star, \lambda^\star)$, the demand instruments (Z^D) are orthogonal to the demand-side structural errors $\xi(\theta^\star)$, i.e., $E\left[Z^{D'}\xi(\theta^\star)\right] = 0$, so that the *GMM* estimates solve

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \left\{ g(\theta_1, \theta_2)' W g(\theta_1, \theta_2) \right\}, \text{ where } g(\theta_1, \theta_2) = \begin{bmatrix} g^{\text{D}}(\theta_1, \theta_2) \\ g^{\text{M}}(\theta_1, \theta_2) \end{bmatrix}, \tag{33}$$

where $g^{\text{D}}(\theta_1, \theta_2)$ represents the orthogonality conditions of interacting the structural errors with instruments, $g^{\text{M}}(\theta_1, \theta_2)$ represents the squared distance between micro-moments implied by the model at guess (θ_1, θ_2) and the empirical micro-moments, and W is a positive, semi-definite the weighting matrix.[22,23]

Our estimates rely on the following instruments. We identify the characteristic random coefficient for sugar content using the differentiation IVs of Gandhi and Houde (2020). The intuition here is that exogenous variation in the availability of products that are similar in sugar content increases substitution for a given product. We similarly identify the random coefficient on the outside option using the total number of cereal products carried in store l in week t. Both of these instruments derive from the available variation in product set discussed in Section 6.4.

Our instrumental variable for price relies on cost data for different grains and sugar sweeteners to capture time-series cost shocks that vary by brand but are common across geographic markets. This accounts for the fact that a given brand is usually produced in a single factory and shipped to stores.[24] We generate predicted prices by projecting price on these input commodity prices interacted with "type of grain" used in production of the cereal, sugar content (grams of sugar) interacted with the price of sweeteners, and exogenous product characteristics. As in Backus et al. (2021), we found doing this in a Random Forest model more effective than a linear projection as the Random Forest is better able to approximate the nonlinear effects of cost changes reflected in firms' optimal prices. Given predicted price, we generate a differentiation instrumental variable (Z^p) that captures substitutability of products in "price-space".

Identification of the mean price coefficient (α) comes via $\hat{p}$ (i.e., through exogenous cost shocks to the commodity prices, particularly sweeteners) while we identify the coefficient of price interacted with income (π^p) and the Box-Cox transform parameter (λ) in part using the interaction of the price differentiation instrumental variable with moments of market l's income distribution:

$$Z_t^D = Z_t^p \otimes \left\{1, \, \%\text{kids}, \, \text{inc}_{lt}^{X\%}\right\}, \tag{34}$$

where $\text{inc}_{lt}^{X\%}$ corresponds to the Xth-percentile store l's fitted household income distribution y_l; we consider average income quartiles. The interaction of predicted price and percentiles of the income distribution allows us to to identify the shape of the distribution of price sensitivity – a feature generated by the Box-Cox parameter λ.

[22] As the instruments come from different data sources, the weighting matrix is block-diagonal. We compute the weighting matrix for the aggregate orthogonality moment conditions using the standard 2-step process. We construct the optimal weighting matrix for the micro-moments by bootstrapping the IRI panel data with replacement, for each sample constructing the corresponding micro-moments, and inverting the covariance matrix of the bootstrapped sample (Gourieroux, Monfort and Renault, 1993).

[23] We solve (33) using first the aggregate demand instruments (Z^D) to find an initial estimate $\hat{\theta}$ and then generate an approximation to the optimal aggregate instruments following Berry et al. (1999), Reynaert and Verboven (2013) and Conlon and Gortmaker (2020).To increase the likelihood of achieving a global minimum, we employ the Knitro Interior/ Direct algorithm suggested by Dubé, Fox and Su (2012) starting from several different initial conditions.

[24] We found that also including measures of distance between store and factory interacted with diesel fuel prices yielded small and insignificant point estimates in a simple first-stage price regression so we have not included fuel prices in our specification.

As λ regulates the distribution of price-sensitivity across consumers and therefore consumption patterns among low- and high-income consumers, identification comes from the likelihood that consumers buy inexpensive versus expensive varieties conditional on income. For example, when $\lambda = 1$ marginal differences in price sensitivity across income levels are uniform. Hence, the predicted average price of the chosen product changes uniformly across income groups, all-else-equal. When $\lambda = 0$ we observe that small differences in income will yield very different consumption patterns with respect to price at low- income levels. We would therefore observe in the data that the average price paid between consumers across the lowest income groups would look very different while the average price paid among the highest income groups would change little. Just the opposite is true for the case when $\lambda > 1$ as the gradient in the average price paid across low-income consumers is flat while we observe a large gradient across high-income consumers. At the same time, the added flexibility of the Box-Cox transform does not preclude, for example, finding that wealthy consumers are less price sensitive than poor consumers. The sign of the price interaction, π_p governs such relationships.

Lastly, we use the IRI panel data to construct micro-moments that are particularly useful in identifying differences in demand which are correlated with demographics, including income. We include average price paid across the top three income quartiles (relative to the first income quartile) among households that purchase any cereal variety. We also include as micro-moments the correlation between price and whether the family has kids to identify the effect of presence of children on price sensitivity, These moments aid in identifying π^p, π^k, and λ and therefore aid in identifying the price-sensitivity distribution in the underlying consumer population. Finally, we include as moments the expected income conditional on buying cereal to identify the average outside option by income and the probability of having kids in the household conditional on buying cereal to identify variation in the outside option by presence of children.

Results. We consider three versions of the model: one where we estimate λ, a specification where we set $\lambda = 0$ and hence include log-income as in Nevo (2001), and a last specification where we set $\lambda = 1$ and hence include income. The *GMM* estimator and instruments are common across all specifications. We also consider an unreported simple multinomial logit specification to provide a base case. Table (2) presents parameter estimates based on a set of $N = 1000$ simulated agents.

We find that all models generate reasonable parameter estimates in implying intuitive patterns among consumers. For example, we observe downward-sloping demand for all model specifications and find that consumers become less price-sensitive as their income grows. Households with kids are more price-sensitive than households without kids, but are more likely to buy cereal overall, and high-sugar cereal specifically (Kids-Sugar $>$ 0). High-income households are less likely to buy cereal (Income-Constant $<$ 0).

However, matching the price gradient across household income (Table 4) requires a transformed income distribution ($\hat{\lambda} = 3.2$), and our estimates enable us to reject log-transformed income ($\lambda = 0$) which is the prevailing approach for modeling price-income interactions in the literature. In

Table 2: **IRI Ready-To-Eat Estimation Results**

Parameter	Flexible	Income	Log-Income
Box-Cox Transform (λ)	3.1511	1.0000	0.0000
	(1.5118)	-	-
Price (α)	2.1368	2.4980	2.2425
	(0.0617)	(0.0131)	(0.0171)
Random Coefficients (Σ):			
Constant	1.6540	1.7483	2.8162
	(2.357)	(2.0901)	(1.2726)
Sugar	1.6256	1.5683	2.7426
	(1.1222)	(1.119)	(1.0542)
Demographic Interactions (Π):			
Income-Constant	-1.7258	-1.3255	-1.2753
	(0.3391)	(0.2445)	(0.1962)
Income-Price	1.1549	0.4041	0.3202
	(0.1744)	(0.0265)	(0.0235)
Income-Sugar	-0.1013	0.1703	0.5913
	(0.0705)	(0.0472)	(0.0416)
Kids-Constant	0.7176	0.7551	0.8736
	(0.3212)	(0.3251)	(0.2125)
Kids-Price	-0.0375	-0.0259	-0.0023
	(0.0325)	(0.0192)	(0.0306)
Kids-Sugar	0.3332	0.2718	0.5638
	(0.095)	(0.0907)	(0.0954)
Implications:			
- Elasticity	2.21	2.93	2.23
- Curvature	1.06	1.01	1.00

Notes: Estimates (standard errors in parentheses) based on IRI scanner data from 2007 to 2011 for Boston (5.2% of total revenue), Philadelphia (4.5%), Chicago (4.2%), San Francisco (3.0%), Seattle (2.5%), Houston (2.5%), and St Louis (2.4%). Sample amounts to 85,829 brand-chain-week observations. GMM estimates include brand and market (city-week) fixed effects. Estimated models include the same set of identifying GMM instruments and cost instruments (commodity prices) discussed in Section 6.4. All statistics under "Implications" correspond to the brand-store-week average. Source: Authors' calculations.

terms of implications, our flexible model generates demand elasticity estimates which are similar in aggregate but feature less spread in demand elasticity (so different estimates of market power) than standard approaches (i.e., using log-income or income interactions with price). We also find demand curvature estimates which are higher (so greater cost pass-through) than standard approaches.

We present estimated curvature-elasticity pairs in Figure 11 and moments of their distributions in Table 3. We observe that moving from the multinomial logit to mixed logit models generates substantially greater variation in estimated product curvature-elasticity pairs. When using raw income as the price interaction (bottom-left panel), we find substantial variation in estimated demand elasticities. When including log-income in $\alpha_i^\star$, we find lower curvature estimates and less dispersion in the estimated elasticity-curvature pairs, reflecting the lower variance and skewness of the log-adjusted distribution.

Figure 11: IRI Breakfast Cereal: Elasticity and Curvature Estimates

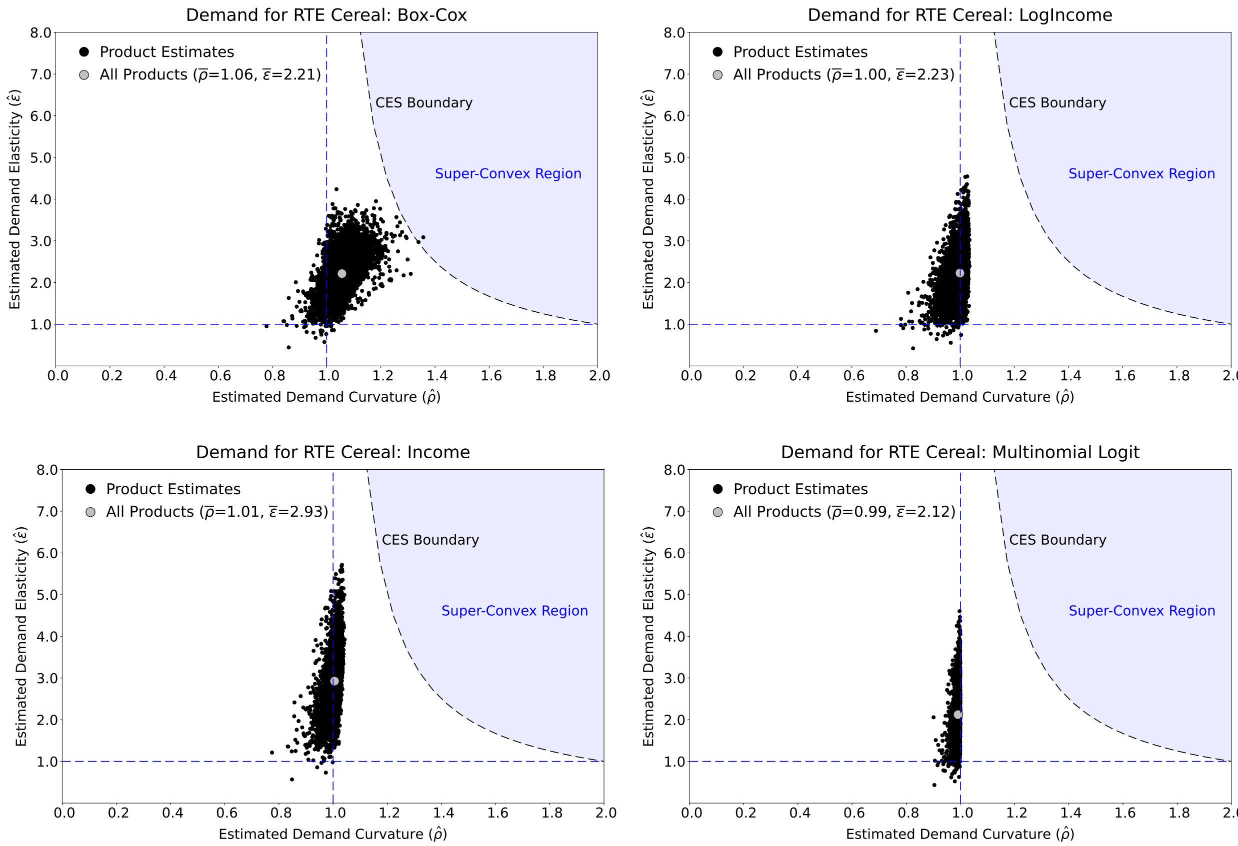


Figure Notes: Dots represent the point elasticity and curvature estimates for each observation in the sample with the silver dot corresponding the the average elasticity and curvature estimates.

In all specifications we find that high-income consumers are less price-sensitive than low-income consumers. The distribution of price sensitivity, however, varies significantly across specifications. The estimated shape parameter ($\hat{\lambda}=3.2$) of the more flexible Box-Cox model implies that low-income consumers are relatively uniform in their price sensitivity while high-income consumers are heterogeneous. This stands in contrast with both the log-income (top-right) and income (bottom-left) models where low-income consumers are more heterogeneous in their price sensitivity.

We observe from the panels that allowing for flexibility in estimation of the distribution of price sensitivity leads to different scatter plots of estimated demand elasticity-demand curvature pairs. We find estimated demands which are less elastic and feature greater curvature than under the restricted models. Hence, this more flexible model implies both greater estimated market power and cost pass-through than commonly used specifications, which may have important implications for in settings where strategic price responses of firms are important; e.g., market power (De Loecker,

Table 3: Elasticity, Curvature, and Flexible Demand

	Flexible	Income	Log-Income	MNL
Elasticity				
- Mean	2.21	2.93	2.23	2.12
- Median	2.21	2.91	2.21	2.10
- Stand. Dev.	0.47	0.67	0.53	0.52
- 90%	2.82	3.79	2.91	2.80
- 10%	1.60	2.06	1.56	1.46
Curvature				
- Mean	1.06	1.01	1.00	0.99
- Median	1.05	1.01	1.01	0.99
- Stand. Dev.	0.05	0.02	0.03	0.01
- 90%	1.11	1.02	1.02	1.00
- 10%	1.01	0.98	0.97	0.98

Eeckhout and Unger, 2020; Grieco et al., 2021), and tariffs (De Loecker et al., 2016; Fajgelbaum, Goldberg, Kennedy and Khandelwal, 2019).

In Table 4 we present the model fit across different specifications. This table provides the link connecting our theoretical results (that the distribution of price-sensitivity drives elasticity-curvature) and empirical results (how demand specification plus identifying moments connect to estimated elasticity-curvature). We find that all demand specifications are able to match the moments not connected to the price-income interactions (lower panel). Only by adding flexibility to the price-income interaction, however, is the model capable of matching the consumption pattern across income quartiles. Further, all demand specifications match the consumption patterns of high-income consumers which suggests that identification of the price-income coefficient (π^p) comes from the right-tail of the income distribution.

Table 4: Matching Consumption Patterns

Moment	Data	Flexible ($\hat{\lambda}=3.2$)	Income ($\lambda=1$)	Log-Income ($\lambda=0$)	MNL
$\mathbb{E}$[Price|IncomeQ_2]/$\mathbb{E}$[Price|IncomeQ_1]	1.0011	1.0019	1.0109	1.0152	1.0000
$\mathbb{E}$[Price|IncomeQ_3]/$\mathbb{E}$[Price|IncomeQ_1]	1.0087	1.0090	1.0235	1.0289	1.0000
$\mathbb{E}$[Price|IncomeQ_4]/$\mathbb{E}$[Price|IncomeQ_1]	1.0492	1.0496	1.0460	1.0421	1.0000
Corr(Price,Kids)	-0.0149	-0.0149	-0.0131	-0.0110	0.0000
E[Income|Buy]	0.9852	0.9852	0.9867	0.9857	1.0000
E[Kids|Buy]	1.2470	1.2470	1.2432	1.2450	1.0000

6.5 Marginal Costs and Pricing

We use the estimated demand models, together with an equilibrium pricing model, to determine marginal costs consistent with the observed prices. In line with the above descriptive evidence, we assume that a single uniform price prevails in all stores affiliated with a given chain in a geographic market. We treat chains as local monopolists. Under uniform pricing, each chain therefore solves:

$$\max_{p_{jt}} \sum_{j \in J} \left[(p_{jt} - c_{jt}) \times \sum_{l=1}^{L} M_{lt} s_{jlt}\Big(p, x, \xi; \theta\Big) \right], \tag{35}$$

where c_{jt} denotes the marginal cost of product j in period t. We assume that differences in marginal cost across stores within a chain due to transportation are negligible. To simplify the notation, we omit the period t subscripts going forward. Define as $s_j(p, x, \xi; \theta)$ the aggregate demand for product j, $\sum_{l=1}^{L} M_l s_{jl}(p, x, \xi; \theta)$. Profit maximization implies the following first-order condition for product j, $\forall j \in J$:

$$s_j\Big(p, x, \xi; \theta\Big) + \sum_{m \in J} \Big(p_m - c_m\Big) \times \frac{\partial s_m}{\partial p_j} = 0\,. \tag{36}$$

The final term $\frac{\partial s_m}{\partial p_j^w}$ is the response in product m's quantity sold to a change in price and, through the pricing rule, the retail price of product j. We transform the first-order condition into vector notation enables us to separate costs from dollar markups:

$$p = c + \underbrace{[\Delta']^{-1} \times s\Big(p, x, \xi; \theta\Big)}_{\text{vector of \$ markups}}, \tag{37}$$

where Δ is the matrix of changes in quantity sold due to changes in retail price with element (k, m) equal to $\frac{\partial s_k}{\partial p_m}$; i.e.,

$$\Delta = \begin{bmatrix} \frac{\partial s_1}{\partial p_1} & \cdots & \frac{\partial s_1}{\partial p_J} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_J}{\partial p_1} & \cdots & \frac{\partial s_J}{\partial p_J} \end{bmatrix}. \tag{38}$$

Given estimates of consumer demand $(\hat{\theta})$ together with price and quantity data, we recover product-level marginal costs $(\hat{c_{jt}})$ for each chain via (37).

6.6 Consumer Welfare Implications of Uniform Pricing

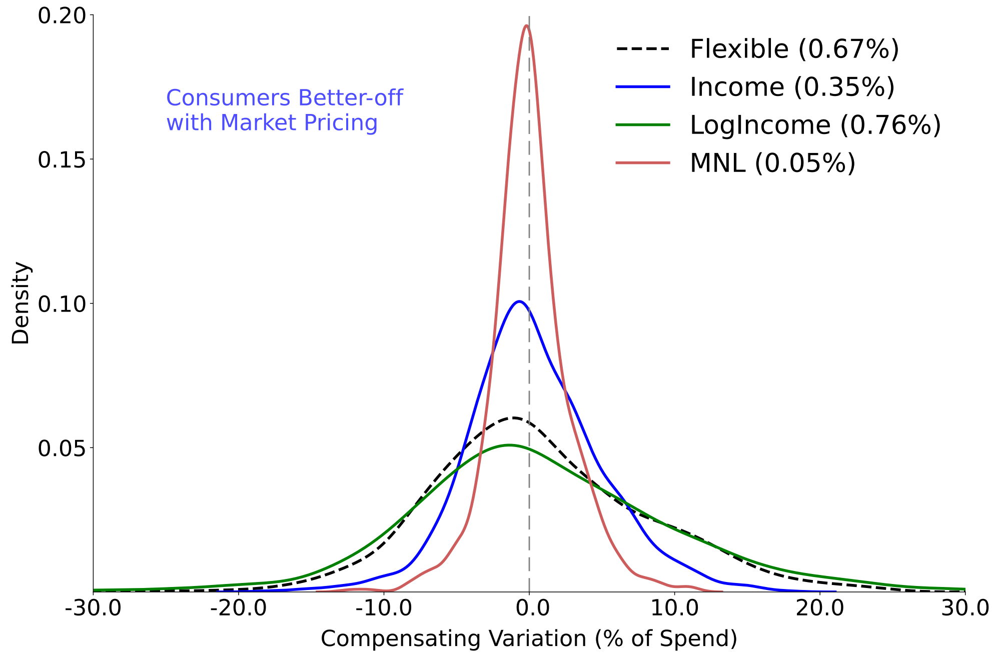
We now rely on the estimated equilibrium model and estimated marginal costs to assess the welfare implications of uniform pricing, and the role of curvature therein. For each of the four demand specifications, we predict the optimal store-specific price, assuming – as in estimation – that consumers are captive to a particular store. These prices therefore are the result of the following profit-maximization problem:

$$\max_{p_{jlt}} \sum_{j \in J} \Big[(p_{jlt} - c_{jlt}) \times M_{lt} s_{jlt}\Big(p, x, \xi; \theta\Big)\Big]\ , \tag{39}$$

We identify beneficiaries of uniform pricing by evaluating changes in consumer welfare in moving from the observed uniform to store-specific pricing via compensating variation, i.e., the amount of income necessary to keep individuals in a given location indifferent between any counterfactual set of prices p' and the uniform ones p. Residents in location l are thus on average better-off under uniform pricing when compensating variation is positive.

We calculate each household's compensating variation following Small and Rosen (1981) and aggregate across household demographics and unobserved preference heterogeneity in deriving aggregate compensating variation for store l consumers. We present the distribution of compensating variation (as a percent of location l cereal spend observed in the data) in Figure 12.

Figure 12: Consumer Welfare Implications of Uniform Pricing



Notes: Figure presents kernel densities across year-store pairs as a percent of cereal spend observed in the data. Average effects are presented in parentheses.

The distribution of compensating variation, particularly the standard deviation, follows directly from each model's ability to match the distribution of price sensitivity (Table 4) and the corresponding distribution of demand curvature (Table 3 and Figure 11). While all four specifications predict that on average, consumers are near indifferent between targeted and uniform pricing, the figure illustrates that the models make different predictions for the distributional implications of uniform pricing: the multinomial logit, which effectively imposes identical curvature across consumers and markets, predicts relatively small distributional effects of uniform pricing. In line with the theoretical results in Aguirre et al. (2010), allowing for more flexibility in price sensitivity and hence curvature, as in the remaining three models, leads to more pronounced distributional implications. Importantly, Figure 12 makes clear that providing the flexibility to match the gradient of price sensitivity in the data is an important ingredient to evaluate the welfare implications of uniform pricing, particularly to different consumers.

In Figure 13 we identify winners and losers of uniform pricing by demographic group. The implications of modelling price-sensitivities flexibly – or equivalently of allowing for flexibility in the estimation of demand curvature – for welfare is made clear by looking at the welfare effects across different model specifications (Panel a). We find that high-income consumers benefit from uniform-pricing at the expense of low-income consumers in all of our mixed-logit specifications. This speaks to the general value of allowing for heterogeneity in price-sensitivity. The differential effects across income groups, however, is governed by the distribution of price-sensitivity across income. Here, we observe that specifying this distribution as proportional to either "income" or "log-income" biases the distributional effects downwards. For example, we find that our flexible model implies that the welfare effect for high-income consumers is 1.3 and 2.5 times that of the models with log-income and income price interactions, respectively.

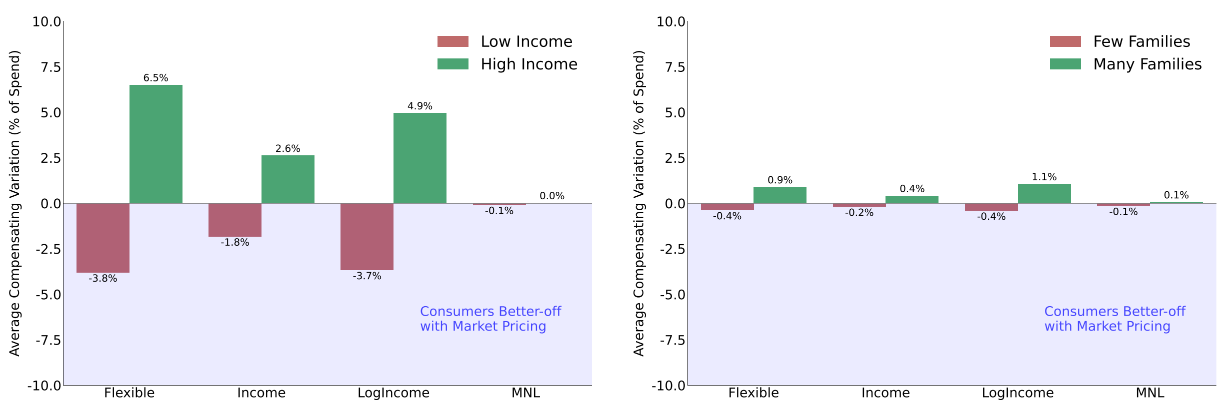
As our model had less scope for variation in price-sensitivity across families of different sizes, we observe little variation in welfare effects for families of different sizes across models (Panel b). If we observed significant variation in consumption patterns across price and family size (or any other observable demographic), adding this level of flexibility would make the model's welfare implications for these groups richer.

Summary. The objective of this empirical application was to demonstrate not only how to flexibly estimate the distribution of price-sensitivity, but also to address the implications of allowing for this kind of flexibility. Figure 13 therefore demonstrates that how the researcher models the distribution of price-sensitivity has important substantive implications, particularly for welfare across individuals. As equity is increasingly the focus of policy debates, these results indicate that allowing for flexibility in the estimation of the distribution of price-sensitivity is of first-order importance to the evaluation of alternative policy solutions.

Lastly, it is important to note that our conclusions regarding the under-estimation of distributional implications are context specific and not the necessary outcome of imposing simple distributions of price-sensitivity *ex ante*. Our results are driven by the fact that we find the patterns

Figure 13: Distributional Implications of Uniform Pricing

(a) Income (b) Kids



Notes: Figure present average CV/spend across markets of similar demographic characteristic where each characteristic is divided into quartiles. "Low Income" ("High Income") reflects markets which are in the bottom (top) 25% of average income in the sample. "Few Families" ("Many Families") reflects markets which are in the bottom (top) 25% of percent of households with a child.

in the data to be consistent with a distribution of price sensitivity that is more skewed than the distribution of either income or log-income can accommodate. In a different empirical setting, however, this may ot be the case.

7 Concluding Remarks

We have shown that the unit-demand mixed-logit model is capable of accommodating a wide array of empirically-relevant elasticity-curvature pairs, thereby providing further evidence of the power of the mixed-logit model as a demand framework and policy tool. We have also demonstrated how different components of the demand specification contribute to expanding the set of attainable elasticity-curvature pairs. This is useful as it both aids in identification of the mixed-logit model and demystifies the mixed-logit model by enabling the researcher to articulate the path from data to model to empirical result and/or policy recommendation for questions centered on cost passthrough. In particular, our theoretical and empirical results highlight the importance of modeling the shape of the price-sensitivity distribution flexibly in order to keep a healthy distance between assumptions and results.

References

Adão, Rodrigo, Arnaud Costinot, and Dave Donaldson (2017): "Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade," *American Economic Review*, Vol. 107, pp. 633–689.

Adams, Brian and Kevin Williams (2019): "Zone Pricing in Retail Oligopoly," *American Economic Journal: Microeconomics*, Vol. 11, pp. 124–56.

Aguirre, Iñaki, Simon Cowan, and John Vickers (2010): "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, Vol. 100, pp. 1601–1615.

Anderson, Simon P. and André de Palma (2020): "Decoupling the CES Distribution Circle with Quality and Beyond: Equilibrium Distributions and the CES-Logit Nexus," *Economic Journal*, Vol. 130, pp. 911–936.

Anderson, Simon P., André de Palma, and Jaques-François Thisse (1987): "The CES Is a Discrete Choice Model?" *Economics Letters*, Vol. 24, pp. 139–140.

——— (1992): *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: The MIT Press.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, Vol. 132, pp. 1553–1592.

Backus, Mathew, Christopher Conlon, and Michael Sinkinson (2021): "Common Ownership and Competition in the Ready-To-Eat Cereal Industry," working paper, University of California at Berkeley.

Bagnoli, Mark and Ted Bergstrom (2005): "Log-Concave Probability and Its Applications," *Economic Theory*, Vol. 26, pp. 445–469.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, pp. 841–890.

——— (1999): "Voluntary Export Restraints on Automobiles: Evaluating Strategic Trade Policy," *American Economic Review*, Vol. 89, pp. 400–430.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, Vol. 112, pp. 68–105.

Berry, Steven T. and Philip A. Haile (2021): "Foundations of Demand Estimation," in Kate Ho, Ali Hortaçsu, and Alessandro Lizzeri eds. *Handbook of Industrial Organization,* Vol. 4, New York, NY: North-Holland.

Besanko, David, Jean-Pierre Dubé, and Sachin Gupta (2005): "Own-Brand and Cross-Brand Retail Pass-Through," *Marketing Science*, Vol. 24, pp. 123–137.

Besley, Timothy and Harvey S. Rosen (1999): "Sales Taxes and Prices: An Empirical Analysis," *National Tax Journal*, Vol. 52, pp. 157–78.

Birchall, Cameron and Frank Verboven (2022): "Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation," Discussion Paper 16981, CEPR.

Björnerstedt, Jonas and Frank Verboven (2016): "Does Merger Simulation Work? Evidence from the Swedish Analgesics Market," *American Economic Journal: Applied Economics*, Vol. 8, pp. 125–164.

Box, George E.P. and David R. Cox (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B*, Vol. 26, pp. 211–252.

Bulow, Jeremy I. and Paul Pfleiderer (1983): "A Note on the Effect of Cost Changes on Prices," *Journal of Political Economy*, Vol. 91, pp. 182–185.

Caplin, Andrew and Barry Nalebuff (1991a): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, Vol. 59, pp. 25–59.

——— (1991b): "Aggregation and Social Choice: A Mean Voter Theorem," *Econometrica*, Vol. 59, pp. 1–23.

Chintagunta, Pradeep K. (2002): "Investigating Category Pricing Behavior at the Retail Chain," *Journal of Marketing Research*, Vol. 39, pp. 141–154.

Compiani, Giovanni (2022): "Market Counterfactuals and the Specification of Multi-Product Demand: A Nonparametric Approach," *Quantitative Economics*, Vol. 13, pp. 545–591.

Conlon, Christopher T. and Jeff Gortmaker (2020): "Best Practices for Demand Estimation with pyBLP," *RAND Journal of Economics*, Vol. 51, pp. 1108–1161.

Cournot, Antoine A. (1838): *Recherches sur les principes mathématiques de la théorie des richesses*, Paris, France: Hachette.

Cuesta, José Ignacio and Alberto Sepúlveda (2021): "Price Regulation in Credit Markets: A Trade-off Between Consumer Protection and Credit Access," working paper, Stanford University.

Davis, Peter (2005): "Constraints on Own- and Cross-Price Elasticities Obtained from Random Coefficient Multinomial-Logit Models," working paper, London School of Economics.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik (2016): "Prices, Markups and Trade Reform," *Econometrica*, Vol. 84, pp. 445–510.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020): "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, Vol. 135, pp. 561–644.

DellaVigna, Stefano and Matthew Gentzkow (2019): "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, Vol. 134, pp. 2011–2084.

Dixit, Avinash and Joseph E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, Vol. 67, pp. 297–308.

Dubé, Jean-Pierre, Jeremy T. Fox, and Che-Lin Su (2012): "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, Vol. 80, pp. 2231–2267.

Fabra, Natalia and Mar Reguant (2014): "Pass-Through of Emissions Costs in Electricity Markets," *American Economic Review*, Vol. 104, pp. 2872–2899.

Fajgelbaum, Pablo D, Pinelopi K Goldberg, Patrick J Kennedy, and Amit K Khandelwal (2019): "The Return to Protectionism," *The Quarterly Journal of Economics*, Vol. 135, pp. 1–55, 11.

Frieberg, Richard and André Romahn (2018): "Pass-Through by Multi-Product Firms," *International Journal of Economics of Business*, Vol. 25, pp. 265–295.

Froeb, Luke M., Steven Tschantz, and Gregory J. Werden (2005): "Pass-Through Rates and the Price Effects of Mergers," *International Journal of Industrial Organization*, Vol. 23, pp. 703–715.

Gandhi, Amit and Jean-François Houde (2020): "Measuring Substitution Patterns in Differentiated Products Industries," working paper, University of Wisconsin-Madison.

Gaudry, Marc J. I. and Michael J. Wills (1978): "Estimating the Functional Form of Travel Demand Models," *Transportation Research*, Vol. 12, pp. 257–289.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, Vol. 8, pp. S85–118.

Grieco, Paul L. E, Charles Murry, and Ali Yurukoglu (2021): "The Evolution of Market Power in the U.S. Auto Industry," Discussion Paper 29013, National Bureau of Economic Research.

Griffith, Rachel, Lars Nesheim, and Martin O'Connell (2018): "Income Effects and the Welfare Consequences of Tax in Differentiated Product Oligopoly," *Quantitative Economics*, Vol. 9, pp. 305–341.

Gutierrez, German (2022): "The Welfare Consequences of Regulating Amazon," working paper, University of Washington.

Häckner, Jonas and Mathias Herzing (2016): "Welfare Effects of Taxation in Oligopolistic Markets," *Journal of Economic Theory*, Vol. 163, pp. 141–166.

Hanemann, W. Michael (1984): "Discrete/Continuous Models of Consumer Demand," *Econometrica*, Vol. 52, pp. 541–562.

Hellerstein, Rebecca and Pinelopi Koujianou Goldberg (2013): "A Structural Approach to Identifying the Sources of Local Currency Price Stability," *Review of Economic Studies*, Vol. 80, pp. 175–210.

Hitsch, Günter, Ali Hortaçsu, and Xiliang Lin (2021): "Prices and Promotions in U.S. Retail Markets," *Quantitative Marketing and Economics*, Vol. 19, 12.

Hollenbeck, Brett and Kosuke Uetake (2021): "Taxation and Market Power in the Legal Marijuana Industry," *RAND Journal of Economics*, Vol. 19, pp. 417–455.

Jaffe, Sonia and E. Glen Weyl (2013): "The First-Order Approach to Merger Analysis," *American Economic Journal: Microeconomics*, Vol. 5, pp. 188–218.

Kenkel, Don (2005): "Are Alcohol Tax Hikes Fully Passed Through to Prices? Evidence from Alaska," *American Economic Review*, Vol. 95, pp. 273–277.

Kimball, Miles S. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, Vol. 27, pp. 1241–1277.

Koppelman, Frank S. (1981): "Non-Linear Utility Functions in Models of Travel Choice Behavior," *Transportation*, Vol. 10, pp. 127–146.

MacGillivray, Helen L. (1986): "Skewness and Asymmetry: Measures and Orderings," *The Annals of Statistics*, Vol. 14, pp. 994–1011.

Marion, Justin and Erich Muehlegger (2011): "Fuel Tax Incidence and Supply Conditions," *Journal of Public Economics*, Vol. 95, pp. 1202–1212.

Marshall, Alfred (1920): *Principles of Economics*, London, UK: Macmillan, 8th edition.

McDonald, James B. (1984): "Some Generalized Functions for the Size Distribution of Income," *Econometrica*, Vol. 52, pp. 647–663.

McFadden, Daniel L. and Kenneth E. Train (2000): "Mixed MNL Models for Discrete Response," *Journal of Applied Econonometrics*, Vol. 15, pp. 447–470.

Miller, Nathan H., Marc Remer, Conor Ryan, and Gloria Sheu (2015): "Pass-Through Rates and the Prediction of Mergers Price Effects," *Journal of Industrial Economics*, Vol. 64, pp. 684–709.

Miller, Nathan H., Matthew Osborne, and Gloria Sheu (2017): "Pass-Through in a Concentrated Industry: Empirical Evidence and Regulatory Implications," *The RAND Journal of Economics*, Vol. 48, pp. 69–93.

Miravete, Eugenio J., Katja Seim, and Jeff Thurk (2018): "Market Power and the Laffer Curve," *Econometrica*, Vol. 85, pp. 1651–1687.

Mrázová, Monika, J. Peter Neary, and Mathieu Parenti (2021): "Sales and Markup Dispersion: Theory and Empirics," *Econometrica*, Vol. 89, pp. 1753–1788.

Mrázová, Monika and J. Peter Neary (2017): "Not So Demanding: Demand Structure and Firm Behavior," *American Economic Review*, Vol. 107, pp. 3835–3874.

Nadarajah, Saralees (2005): "A Generalized Normal Distribution," *Journal of Applied Statistics*, Vol. 32, pp. 685–694.

Nakamura, Emi (2008): "Pass-Through in Retail and Wholesale," *American Economic Review*, Vol. 98, pp. 430–37, May.

Nakamura, Emi and Dawit Zerom (2010): "Accounting for Incomplete Pass-Through," *Review of Economic Studies*, Vol. 77, pp. 1192–1230.

Nevo, Aviv (1998): "Measuring Market Power in the Ready-To-Eat Cereal Industry," Discussion Paper 6387, NBER.

——— (2000): "Mergers With Differentiated Products: The Case of the Ready-To-Eat Cereal Industry," *RAND Journal of Economics*, Vol. 31, pp. 395–421.

——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, Vol. 69, pp. 307–342.

Nocke, Volker and Nicolas Schutz (2018): "Multiproduct-Firm Oligopoly: An Aggregative Games Approach," *Econometrica*, Vol. 86, pp. 523–557.

Petrin, Amil (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, Vol. 110, pp. 705–729.

Poterba, James M. (1996): "Retail Price Reactions to Changes in State and Local Sales Taxes," *National Tax Journal*, Vol. 49, pp. 165–76.

Reynaert, Mathias. and Frank. Verboven (2013): "Improving the Performance of Random Coeffcients Demand Models: the Role of Optimal Instruments," *Journal of Econometrics*, Vol. 178.

Robinson, Joan (1933): *Economics of Imperfect Competition*, London, UK: Macmillan.

Small, Kenneth A. and Harvey S. Rosen (1981): "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, Vol. 49, pp. 105–130.

Train, Kenneth E. (2009): *Discrete Choice Methods with Simulation*, New York, NY: Cambridge University Press, 2nd edition.

Wang, Ao (2022): "Sieve BLP: Sieve Blp: A Semi-nonparametric Model of Demand for Differentiated Products," *Journal of Econometrics*, p. forthcoming.

Weyl, E. Glen and Michal Fabinger (2013): "Pass-Through as an Economic Tool: Principles of Incidence Under Imperfect Competition," *Journal of Political Economy*, Vol. 121, pp. 528–583.

Appendix

A Elasticity and Curvature of Demand for Breakfast Cereal

Nevo (2000) specifies preferences as follows (ignoring market location and time indices):

$$u_{ij} = x_j\beta_i^\star + \alpha_i^\star p_j + \xi_j + \epsilon_{ij}\,, \qquad i \in \mathcal{I},\ j \in \mathcal{J},\ \epsilon_{ij} \sim \text{EV1}\,, \tag{A.1a}$$

$$\begin{pmatrix}\alpha_i^\star \\ \beta_i^\star\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \Pi D_i + \Sigma\nu_i\,, \qquad \nu_i \sim N(0, I_{n+1})\,, \tag{A.1b}$$

where x_j is the $(n\times1)$ vector of observed product characteristics and p_j is the price of (inside) product j available in each market, $\mathcal{J}$, with $J = |\mathcal{J}|$. Payoff of the outside good is $u_{i0} = \epsilon_{i0}$. There are random coefficients of product characteristics, β_i^* and price responsiveness, α_i^*. Preferences might be correlated to a d-vector of demographic traits D_i through the $(n+1)\times d$ matrix Π of interaction estimates that allows for cross-price elasticity to vary across markets with different demographic composition. To further account for individual preferences over unobservable product attributes, ν_i captures mean-zero, unobserved preference shifters with a diagonal variance-covariance matrix Σ. Lastly, the idiosyncratic unobserved preference by consumer i for product j, ϵ_{ij}, follows the Type-I extreme value distribution across all products in $\mathcal{J}$.

Table A.1: Breakfast Cereal: Price Related Estimates

	Means	Std. Dev.	Demographic Interactions (π_p)			Manifold	
SPECIFICATION	(α)	(σ_p)	log(INCOME)	log(INCOME)^2	CHILD	ε	ρ
[A]	-62.7299	3.3125	588.3252	-30.1920	11.0546	3.62	1.06
	(14.8032)	(1.3402)	(270.4410)	(14.1012)	(4.1226)		
[B]	-30.9982	2.0216	—	—	—	3.74	0.96
	(0.9674)	(0.9367)	—	—	—		
[C]	-53.1367	—	444.7281	-22.3987	16.3664	3.60	1.08
	(12.1023)	—	(209.6548)	(10.7282)	(4.7824)		
[D]	-30.8902	—	—	—	—	3.74	0.96
	(0.9944)	—	—	—	—		

Notes: *GMM* estimates of parameters related to price sensitivity using simulated breakfast cereal data estimated via "best practices" described in Conlon and Gortmaker (2020). Remaining parameters for product characteristics follow Nevo (2001) and are included in each demand specification but are not reported. Robust standard errors in parentheses.

We consider four alternative specifications:

$$\text{[A]}\quad \alpha_i^\star = \alpha + \sum_{k=1}^{d}\pi_{\alpha k}D_i + \sigma_\alpha\nu_i\,, \qquad \text{(Nevo – Full Model)} \tag{A.2a}$$

$$\text{[B]}\quad \alpha_i^\star = \alpha + \sigma_\alpha\nu_i\,, \qquad \text{(Only Price Random Coefficient)} \tag{A.2b}$$

$$\text{[C]}\quad \alpha_i^\star = \alpha + \sum_{k=1}^{d}\pi_{\alpha k}D_i\,, \qquad \text{(Only Demographic Price Interactions)} \tag{A.2c}$$

$$[\text{D}] \quad \alpha_i^\star = \alpha \,, \qquad \text{(No Price Interactions)} \tag{A.2d}$$

The estimation results of Model A is represented graphically in Panel A of Figure 1 in the main text. We contrast it with a variant of Model D in Panel B that removes not only price interactions, but also product characteristic interactions.

B Choice Probability Distribution and Demand Manifolds

B.1 Moments

Because of the additive i.i.d. type-I extreme value distribution of ϵ_{ij}, the individual i's choice probability of product j given by (8) is also the mean of an individual-specific Bernoulli distribution:

$$\mu_{ij} = \mathbb{P}_{ij} \,, \tag{B.1}$$

which are functions of the vector of prices p that we omit to reduce clutter. The variance is:

$$\sigma_{ij}^2 = \mathbb{P}_{ij}(1 - \mathbb{P}_{ij}) \,. \tag{B.2}$$

And finally, the third central moment or non-standardized skewness is:

$$sk_{ij} = \mathbb{P}_{ij}(1 - \mathbb{P}_{ij})^2 - \mathbb{P}_{ij}^2(1 - \mathbb{P}_{ij}) = \mathbb{P}_{ij}(1 - \mathbb{P}_{ij})(1 - 2\mathbb{P}_{ij}) \,, \tag{B.3}$$

from where we obtain standardized moment or *skewness* (MacGillivray, 1986) as:

$$\tilde{\mu}_{ij,3} = \frac{sk_{ij}}{\sigma_{ij}^3} = \frac{\mathbb{P}_{ij}(1 - \mathbb{P}_{ij})(1 - 2\mathbb{P}_{ij})}{\sqrt{[\mathbb{P}_{ij}(1 - \mathbb{P}_{ij})]^3}} = \frac{1 - 2\mathbb{P}_{ij}}{\sqrt{\mathbb{P}_{ij}(1 - \mathbb{P}_{ij})}} \,, \tag{B.4}$$

where σ_{ij}^3 is the third raw moment of the individual choice probability distribution.

B.2 Moment Derivatives

We use the derivative of the choice probability (8) with respect to price repeatedly:

$$\mathbb{P}'_{ij} = \frac{\partial \mathbb{P}_{ij}}{\partial p_j} = f'_{ij} \cdot \mathbb{P}_{ij}(1 - \mathbb{P}_{ij}) \,. \tag{B.5}$$

The derivative of the variance with respect to price is:

$$\frac{\partial \sigma_{ij}^2}{\partial p_j} = \frac{\partial \mathbb{P}_{ij}(1 - \mathbb{P}_{ij})}{\partial p_j} = \mathbb{P}'_{ij}(1 - \mathbb{P}_{ij}) - \mathbb{P}_{ij}\mathbb{P}'_{ij} = f'_{ij} \cdot \mathbb{P}_{ij}(1 - \mathbb{P}_{ij})(1 - 2\mathbb{P}_{ij}) = f'_{ij} \cdot sk_{ij} \,. \tag{B.6}$$

To conclude, we obtain the price derivative of skewness by differentiating the first equality in (B.3):

$$sk'_{ij} = \left[(1 - \mathbb{P}_{ij})^2 - 4\mathbb{P}_{ij}(1 - \mathbb{P}_{ij}) + \mathbb{P}_{ij}^2\right] \cdot \mathbb{P}'_{ij} = \left[(1 - 2\mathbb{P}_{ij})^2 - 2\mathbb{P}_{ij}(1 - \mathbb{P}_{ij})\right] \cdot f'_{ij} \cdot \mathbb{P}_{ij}(1 - \mathbb{P}_{ij}) \,. \tag{B.7}$$

B.3 Demand Manifold

Price differentiate (9) and substitute (B.5) to obtain demand elasticity of product j with respect to p:

$$\varepsilon_j(p) \equiv -\frac{p_j}{Q_j(p)} \cdot \frac{\partial Q_j(p)}{\partial p_j} = -\frac{p_j}{Q_j(p)} \int\limits_{i \in \mathcal{I}} f'_{ij} \cdot \mathbb{P}_{ij} \left(1 - \mathbb{P}_{ij}\right) \, dG(i) \,. \tag{B.8}$$

Similarly, the inverse demand curvature of product j is:

$$\rho_j(p) \equiv Q_j(p) \cdot \frac{\partial^2 Q_j(p)/\partial p_j^2}{[\partial Q_j(p)/\partial p_j]^2} = \int\limits_{i \in \mathcal{I}} \mathbb{P}_{ij} dG(i) \times \frac{\left[\begin{array}{c} \int f''_{ij} \cdot \mathbb{P}_{ij} \left(1 - \mathbb{P}_{ij}\right) dG(i) + \\ \int \left(f'_{ij}\right)^2 \cdot \left[\mathbb{P}_{ij} \left(1 - \mathbb{P}_{ij}\right) \left(1 - 2\mathbb{P}_{ij}\right)\right] dG(i) \end{array} \right]}{\left[\int f'_{ij} \cdot \mathbb{P}_{ij} \left(1 - \mathbb{P}_{ij}\right) dG(i) \right]^2} \,. \tag{B.9}$$

Equations (10) and (11) follow after substituting (9), (B.2) and (B.3) into these expressions. Combining elasticity and curvature we obtain the expression for the demand manifold (12):

$$\rho_j[\varepsilon_j(p)] = \frac{p_j^2}{\varepsilon_j^2(p) \cdot Q_j(p)} \cdot \left[\int f''_{ij} \cdot \mathbb{P}_{ij} \left(1 - \mathbb{P}_{ij}\right) dG(i) \; + \int \left(f'_{ij}\right)^2 \cdot \left[\mathbb{P}_{ij} \left(1 - \mathbb{P}_{ij}\right) \left(1 - 2\mathbb{P}_{ij}\right)\right] dG(i) \right] \,. \tag{B.10}$$

C A General Mixing Distribution

Idiosyncratic demand sensitivity is modeled as $\alpha_i^\star = \alpha + \pi\phi_i$, where α is the mean slope of demand and π captures the effect on price heterogeneity of preferences across individuals. We model draws of individual types ϕ_i after the following three-parameter Asymmetric Generalized Normal distribution (Nadarajah, 2005):

$$\text{Prob}\big(\phi < x; \iota, \zeta, \eta\big) = \Phi_N(y) \text{ where } = \begin{cases} \dfrac{-1}{\eta} \log\left(1 - \dfrac{\eta(x - \iota)}{\zeta}\right) \,, & \text{if } \eta \neq 0 \,, \\ \dfrac{x - \iota}{\zeta} \,, & \text{if } \eta = 0 \,, \end{cases} \tag{C.1}$$

and where $\Phi_N(\cdot)$ denotes the cumulative distribution function of a standard normal. To avoid an overparameterized model, we normalize the scale parameter $\zeta = 1$, and $\eta < 0$ so that the support of the distribution is $(\iota + 1/\eta, \infty)$. The distribution is right-skewed, mimicking a log-normal distribution for $\eta = -1$ and converging to a normal distribution as $\eta \longrightarrow 0$. Furthermore we center the distribution around the mean slope:

$$E[\phi] = \iota - \frac{\zeta}{\eta}\left(e^{\eta^2/2} - 1\right) = 0 \,, \tag{C.2}$$

so that:

$$\iota = \frac{1}{\eta}\left(e^{\eta^2/2} - 1\right). \tag{C.3}$$

The one-parameter Asymmetric Generalized Normal distribution can then be written as:

$$\text{Prob}\big(\phi < x; \eta\big) = \Phi_N(y) \text{ where } = \begin{cases} -\dfrac{\log\left(e^{\eta^2/2} - \eta x\right)}{\eta}\,, & \text{if } \eta \neq 0\,, \\ \dfrac{x-\iota}{\zeta}\,, & \text{if } \eta = 0\,, \end{cases} \tag{C.4}$$

with mean, variance, and skewness:

$$\mu[\phi;\eta] = 0\,, \tag{C.5}$$

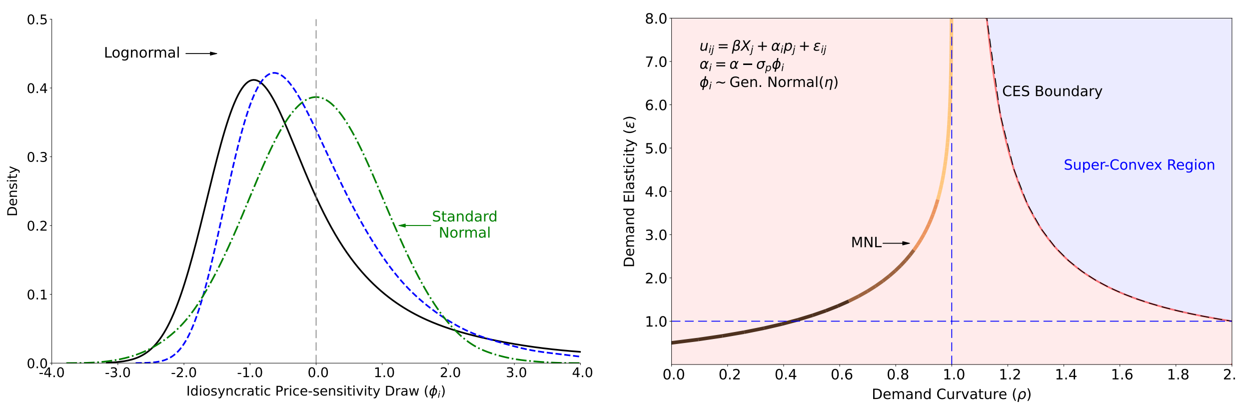
$$\sigma^2[\phi;\eta] = \frac{e^{\eta^2/2}\left(e^{\eta^2/2} - 1\right)}{\eta^2}\,, \tag{C.6}$$

$$\tilde{\mu}_3[\phi;\eta] = \frac{3e^{\eta^2/2} - e^{3\eta^2/2} - 2}{\left(e^{\eta^2/2} - 1\right)^{3/2}}\,. \tag{C.7}$$

Figure C.1: Covering the Space with a Flexible Price Mixing Distribution

(a) Potential Price Mixing Distributions

(b) Generalized Normal Manifolds



Notes: The left panel shows three specifications of the price random coefficient distribution. The right panel shows the combinations of all structural parameters generating well-behaved solutions for (ε, ρ) in the sub-convex region.

Figure C.1 explores the implications of using this flexible mixing distribution for the price random coefficient. In panel (a) we present three different variants of how the price mixing distribution may look: ranging from standard normal to log-normal. We also consider an intermediate case that might represent a particular mixture of these two distributions. In panel (b) we present the implications of this flexibility for covering (ε, ρ) space. As before, we focus our attention to specifications ensuring sub-convexity of demand (light shaded region). Panel (b) shows that allowing for sufficient flexibility in the price mixing distribution expands the support of the parameters of interest and facilitates obtaining robust estimates $(\hat{\varepsilon}, \hat{\rho})$ by relaxing the constraints that other distributions of price random coefficients might impose.

D Nonlinear Income Effects

Table D.1: Income Effects, Markups, and Pass-Through Rates

	$\lambda = 0$		$\lambda = 0.5$		$\lambda = 0.75$		$\lambda = 1$	
Elasticity (ε)	2.83	(0.26)	2.34	(0.48)	2.77	(1.01)	2.75	(2.05)
Curvature (ρ)	1.35	(0.08)	1.19	(0.07)	1.13	(0.05)	0.99	(0.01)
Markup (%)	44.41	(5.26)	46.25	(8.77)	44.48	(13.77)	48.12	(20.55)
Pass-Through (%)	178.99	(18.33)	145.91	(16.38)	117.90	(7.27)	99.41	(0.01)

Notes: Mean and standard deviations (in parentheses) of demand elasticity and curvature plus their implied price markup and pass-through rate.

Figure D.1: Income Effects and Demand Manifolds (by Vehicle Origin)

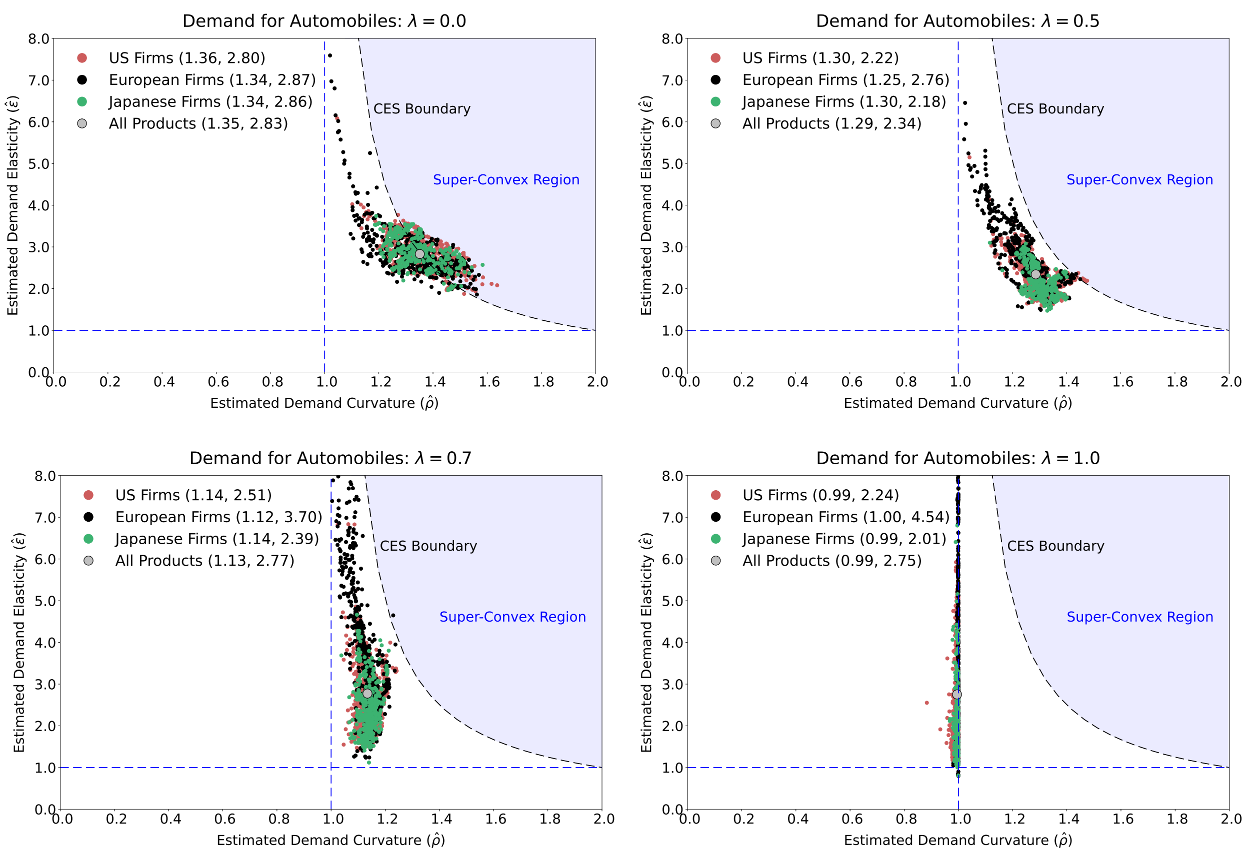


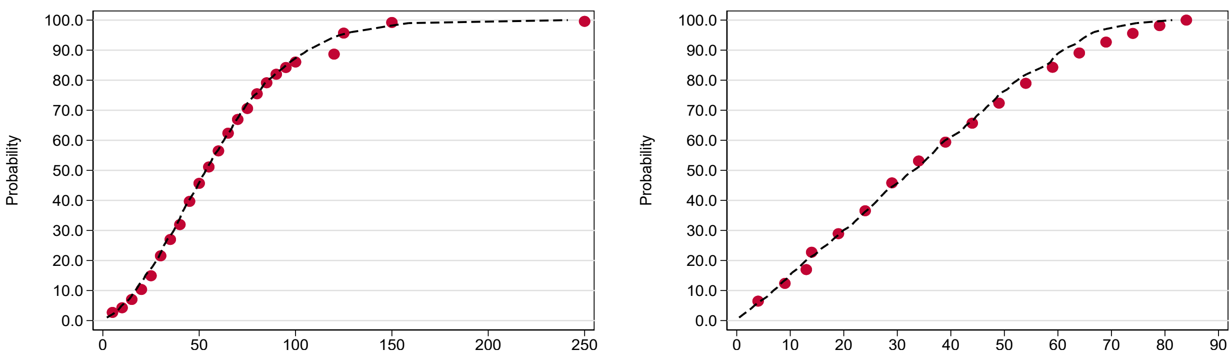
Figure Notes: Dots represent the point elasticity and curvature estimates for each observation in the sample with the red dot corresponding the the average elasticity and curvature estimates.

E Ready-to-Eat Cereal

Simulating Consumers. We construct the sample of simulated consumers for each market by relying on the empirical distributions of the demographic attributes age, presence of children in the household, and income. We use the IRI demographic supplement, which includes demographic statistics for consumers within a two-mile radius of the store. We fit continuous market-specific distributions to the discrete distributions of income and age using generalized beta distributions of the second kind to fit the empirical income and age distributions for each market l. McDonald (1984) highlights that the beta distribution provides a good fit to empirical income data relative to other parametric distributions. In Figure E.1 we compare the estimated cumulative distribution functions (dashed lines) versus the binned data (dots) for a representative store.

Figure E.1: Estimating Demographics

(a) Income (b) Age



Notes: In each panel we compare estimated income and age distribution (dashed lines) and the discrete income and age distributions (dots) in the IRI data.

The IRI data do not have a time dimension so we assume demographics are stable across our time period (2007-2011). Finally, we account for the unobserved preferences for product attributes (ν_{il}) via Halton draws which Train (2009) demonstrated is an efficient method to efficiently cover the space of unobserved preferences (ν_{il}). We then draw 1,000 individuals for each store to derive the predicted probability of choosing product j numerically via monte carlo simulation.