Robust Pass-Through Estimation in Discrete Choice Models

Eugenio J. Miravete†  Katja Seim‡  Jeff Thurk§

March 2022

Abstract

We address how discrete choice demand models might offer robust pass-through estimates suitable for the analysis of the incidence of taxation, cost shifts, exchange rate, and mergers in industries with market power. This addresses the current state of the art of discrete choice demand estimation where apparently innocuous modeling decisions critically determine the range of estimable curvature at each elasticity. We characterize the demand manifold of a BLP-type model to show that in addition to allowing for flexible substitution patterns, a unit-demand, mixed-logit, discrete choice, model can also accommodate a wide range of demand curvatures. Demographic interactions, a skewed distribution of price random coefficients, and/or a flexible treatment of income effects allows a unit-demand mixed-logit model to accommodate any curvature at each elasticity estimate, including those of the CES. We further assume that Marshall’s Second Law of Demand holds at the estimated parameter values so that the equilibrium model generates well-behaved comparative statics.

Keywords: Pass-Through, Demand Curvature, Demand Manifold, Subconvexity.

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

---

* We thank Jeremy Fox for his discussion of an early version of this work. 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, 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

We show that in addition to allowing for flexible substitution patterns, a unit-demand, mixed-logit, discrete choice model can also accommodate a very wide range of demand curvatures. The estimated pass-through rate is not independent from the estimated markup as demand curvature and elasticity are connected to each other through demand specification. We argue that the differentiated effect of cost shifts at each price allows us to identify the distribution of the idiosyncratic price responsiveness that determines if the estimated demand is log-concave or log-convex.

Our analysis indicates that the unit-demand mixed-logit model is capable of accommodating any curvature at each elasticity estimate, including those of the CES discrete choice model. The other element that affects the range of attainable demand curvatures in addition to the distribution of price random coefficients is income and how it enters the utility function. Thus, modeling decisions such as to whether to contemplate income effects or not might critically determine the range of estimable curvatures at each elasticity value and predetermine the predicted pass-through rates of the counterfactuals of the model. Having isolated which elements are behind demand curvature, we advocate for additional flexibility in the specification of the distribution of price random coefficients and/or the functional form of other expenses and show how our proposed model performs with simulated and real data.

Flexible models avoid restricting the set of predictions they can generate. Bulow and Pfleiderer (1983) famously highlighted the importance of functional forms for hypothesis testing. Sumner (1981) concluded that the tobacco industry enjoyed substantial market power because only a fraction of state taxes were passed to final consumers. However, this was the consequence of assuming a linear demand and Bulow and Pfleiderer (1983) noticed that, using the same data, Sumner’s conclusions are reversed if econometricians assume a log-convex demand function.

Reducing the likelihood of misspecification explains the nearly universal adoption of the mixed-logit (ML) model among IO economists after Berry (1994) and Berry, Levinsohn and Pakes (1995), BLP hereafter. Berry and Haile (2021) argue that ML offers a sufficiently flexible specification to avoid strong a priori restrictions of results while being sufficiently parsimonious to permit practical applications. ML is computationally feasible and allows for flexible substitution patterns that circumvent the Independence of Irrelevant Alternatives (IIA) and reduces the risk of misspecification relative to multinomial logit (MNL) or CES.

However, demand flexibility does not end at substitution patterns. Demand curvature determines pass-through and is directly responsible for the predictions of structural IO models on matters such as tax incidence in noncompetitive industries or the price effects of cost savings in horizontal mergers.1 Despite these being central questions in industrial organization, little attention

---

1 Many results in economics depend on accurate demand curvature estimates, including the effect of market size on pricing and markup behavior (Krugman 1979; Zhelobodko, Kokoven, Parenti and Thisse 2012), welfare effects of third degree price discrimination (Aguirre, Cowan and Vickers 2010), taxation (Weyl and Fabinger 2013), or the existence of a Laffer curve in oligopoly (Miravete, Seim and Thurst 2018).
has been paid to study the determinants of pass-through in discrete choice models. When hinting at future work, Berry and Haile (2021) argue:

These substitution 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.

Compiani (2022) acknowledges that the BLP model likely produces biased estimates of demand curvature and suggests using nonparametric methods to overcome this potentially important misspecification. Antitrust economists have also documented large discrepancies for predicted price effects of mergers when demands have different curvature but the same price elasticity. Simulating the post-merger outcome to account for cost synergies and multilateral price effects requires considering the concavity of the profit function, which is determined by demand curvature. In the absence of detailed micro-data, practitioners conduct merger simulations repeatedly for several demand specifications as to ensure robustness of their conclusions.

Despite its flexibility with substitution patterns, ML still shares many common features with MNL. One of them is a tendency to predict incomplete pass-through rates. We argue that this is just the consequence of practitioners not being aware of which components of the model specification influences demand curvature rather than any intrinsic limitation of unit-demand ML models to generate log-concave as well as log-convex demands. Nakamura and Zerom (2010) were the first to hint at the role of demographics, in particular income, as potential drivers of demand curvature in ML models. We show that their intuition is correct as long as demographics are interacted with own-price effects, which modulate the empirical distribution of price random coefficients. We show why demographic interactions with product attributes are however not capable of expanding the range of estimable demand curvatures in any meaningful way.

We adopt the demand manifold approach of Mrázová and Neary (2017) to frame our current understanding of how own-price elasticity and inverse demand curvature relate to each other. We use demand manifolds to show how the different elements of the typical BLP specification (product attributes, random coefficients and demographic interactions) expands or restricts the set of estimable elasticity and curvature combinations. Demand manifolds help us isolate two key elements driving curvature in discrete choice models: the moments of the mixture choice probability distribution and the assumed functional form of expenditures in other products. We also use demand manifolds to evaluate the set of attainable elasticity-curvature combinations for different mixing distributions of the price random coefficient.

---


3 Alternatively, Jaffe and Weyl (2013) suggest using pass-through rates to infer the second derivatives of the profit function. This approach is valid for any demand specification as long as merger-triggered price increases are sufficiently small. To accommodate large price changes, Miller, Remer, Ryan and Shen (2015) suggest conducting the merger simulation using a demand system mimicking the pass-through behavior observed in the data.
Economic theory offers some guidance on the range of admissible elasticities but says little about demand curvature. For instance, inelastic demands estimates are deemed the result of poorly specified models because they are not consistent with profit maximizing behavior when firms enjoy market power. But how does market power restrict pass-through in any meaningful way? Measuring pass-through rates is frequently the outcome of the estimation, not the input that helps us select our preferred specification. The shape of demand basically determines the outcome. Therefore, we favor demand specifications that can accommodate the widest range of estimates, including pass-through rates in excess of 100% that MNL cannot generate.

Mixture Distributions. The mixture distribution of ML choice probabilities has two elements: the component multivariate logit and the mixing distribution behind the stochastic behavior of random coefficients. There are both, theoretical and statistical considerations to choose the mixing distribution that will, in the end, determine the shape and properties of the resulting demand.

Caplin and Nalebuff (1991b) specifically address the case of ML and show that the mixture distribution preserves the curvature properties of its components. In particular, since the logit and the commonly assumed normal mixing distribution are both log-concave, the resulting demand is also log-concave. This result concerns the idiosyncratic valuation of product attributes and not prices, which affect the budget constraint. Thus, a ML model that only includes random coefficients on product characteristics still restricts demand to be log-concave with an incomplete pass-through rate. In their companion paper, Caplin and Nalebuff (1991a, Proposition 11) show that even when demand is log-convex there is an oligopoly equilibrium with horizontally differentiated, single product firms. At this stage, and after decades of empirical work with the BLP model, what we lack is a clear understanding of what might push our demand specifications to fall over the log-convex region.

The properties of the mixture distribution are connected to those of the component and mixing distributions. It is well-known that the mixture probability has larger variance than the component distribution but the resulting properties of the ML demand that are of interest for the analysis of oligopoly equilibrium are difficult to characterize in general as they depend on the value of parameter estimates, preferences and the particular mixing distribution of any given specification. We then evaluate how adopting different mixing distributions expands or restricts the range of estimable elasticity-curvature combinations in the subconvex region of demand and argue in favor of estimating a one-parameter price responsiveness Pareto distribution in conjunction with the rest of the parameters of the BLP model.

Most, if not all, BLP-style models customarily assume normally distributed random coefficients both for attributes and price. A two-tail normally distributed vehicle size random coefficient allows us to accommodate that some individuals prefer large sedans while other favor compact

---

4 This is the result of applying the multivariate version of log-concavity preservation of Prékopa (1973). Nocke and Schuetz (2018) prove existence of multi-product oligopoly equilibrium in an aggregative oligopoly without random coefficients but for a wide range of demand curvatures including all log-convex demands between MNL and CES.
vehicles. The associated individual parameter estimate can be positive for some individuals and negative for others. By the same logic a normally distributed price random coefficient opens the door to upward sloping demand estimates for some individuals in the sample (McFadden and Train, 2000). Our choice of a one-tail Pareto distribution ensures that the sign of price response is the same for all individuals. Furthermore, we find that a normal distribution is only compatible with very restrictive sets of log-convex demands, particularly for low elasticity values where market power is a more pressing issue. We compared against other skewed distribution (exponential, lognormal, Rayleigh) we find that the Pareto distribution extends the range of estimable elasticity and curvature combinations the most, to cover nearly all the subconvex region up to the CES.

Our analysis of the choice of the price mixing distribution hints at another potential source of misspecification of BLP models that might have been overlooked and that appear to be related to the second derivative of demand. Imposing the “wrong” mixing distribution risks obtaining upward biased estimates for demand elasticity resulting in an excessively competitive characterization of the market. Fortunately, addressing this issue is quite simple and we suggest using a flexible asymmetric mixing distribution for price random coefficients where its indexing parameter is also estimated. It is precisely the heterogeneous incidence of a tax or a cost shift for high and low valuation consumers what allows us to identify a flexible distribution of the price random coefficient that determines the trade-off between demand elasticity and curvature estimates.

**Income Effects.** While the assumption of quasi-linear demands might be appropriate to model the demand for cereal and other consumer products, income most likely conditions the demand for automobiles and other large household purchases. We suggest a flexible, one-parameter, Box-Cox transformation of other expenses, \((y_i - p_j)^{\lambda}\) that encompasses the quasi-linear model of Nevo (2001) as well as the logarithmic specification of BLP. We show that in a specification without price random coefficients, the estimate of \(\lambda\) determines the set of estimable elasticity-curvature combinations and is identified using the same arguments. Smaller values of \(\lambda\) substantially expand the range of demand curvatures compatible with low elasticity estimates, approaching the CES.

**Approach.** To convince the reader that choosing the mixing distribution of price random coefficient is an integral part of economic modeling, we first describe how our results apply to the discrete choice model with quasi-linear preferences of Nevo (2001). Once we present our analysis of the mixing distribution of price random coefficients we address how to incorporate price-income interactions. Quasi-linear preferences, e.g., utility depending on \((y_i - p_j)\), eliminate the need to specify any direct effect of income on consumer choice probabilities. We later show how adopting a utility subfunction \(\ln(y_i - p_j)\) as in BLP allows for a wider range of demand curvature estimates. We further generalize this functional to further expand the set of estimable elasticity-curvature combinations. The rest of the paper characterizes the demand manifold for the ML choice model, discusses the choice of the mixing distribution within the manifold framework and shows that the indexing parameter of this distribution is identified.
Findings. The paper clarifies potential sources of curvature misspecification of BLP-style models. We show that while average demand elasticities are quite robust, results involving demand curvature vary substantially depending how we model idiosyncratic price responsiveness. We illustrate the role of product attributes, random coefficients and demographic interactions on the distribution of demand curvature. We show that MNL demand is log-concave and its pass-through rate is necessarily incomplete, something that random coefficients on product attributes cannot relax. Instead, demographic traits are very effective in expanding the set of estimable demand curvatures when interacted with price. We also show that asymmetric mixing distributions to model idiosyncratic price responsiveness extend the set estimable elasticity-curvature combinations substantially, particularly for low elasticity values, with the Pareto distribution basically covering the whole region of log-convex demands up to the CES. Finally, we show that cost shifts interacted with product prices instrument for the differentiated pass-through rate at different price ranges of demand identify the one-parameter distribution of price random coefficients or the effect of a one-parameter transformation of other expenses.

Related Literature. There are some recent attempts that try to add flexibility to the original ML model. Compiani (2022) suggests a nonparametric approach suitable for demand estimation of a small basket of products. Our model, closer to BLP, can accommodate a large set of product alternatives and our parametric approach incorporates the intrinsic relationship between demand elasticity and curvature. Other approaches vindicate the use of the CES discrete choice model (Adao, Costinot and Donaldson 2017; Björnerstedt and Verboven 2016). These models circumvent the limitations of IIA by including random coefficients. However, the CES discrete choice model (as well as Birchall and Verboven 2022) uses a discrete/continuous choice framework (Hanemann 1984) rather than the typical unit-demand discrete choice specification. There are many environments where modeling the budget share allocated to a product might be more appropriate than using a unit-demand discrete choice approach but the added substitution flexibility comes also at a cost: markups are invariant to sales and prices in the CES model. This might have important consequences. Instead, we do not need to reformulate market shares into budget shares and using the common single-unit ML we allow for the same range of demand curvature that CES imposes.

Organization. Section 2 illustrates how different elements of the typical BLP specification constraint the behavior of demand elasticity and curvature using simulated data from Nevo’s the breakfast cereal papers. Section 3 introduces the basic elements of the demand manifold framework. Section 4 characterizes the demand manifold of a general ML model, relate its features to the properties of the selected price mixing distribution and discusses the identification of demand curvature. Section 4.1 justifies the use of the AGN mixing distribution to flexibly model idiosyncratic price responsiveness. Section 6 repeats the analysis for nonlinear transformations of expenditures in

For instance, if Miravete et al. (2018) had used the CES model, retail price will not respond to changes in commodity taxation even thought firms enjoy substantial market power. This negates the possibility of a Laffer curve in commodity taxation and will erroneously predict that tax collection increases monotonically with unit tax rates.
the outside good and show that a model with price-income interactions and no idiosyncratic price responsiveness can generate the same demand curvature that a model with flexibly distributed price random coefficients with quasi-linear preferences. Section 7 concludes. Additional results and derivations are reported in the Appendices.
2 Elasticity and Curvature of Demand for Breakfast Cereal

Ever since the works of [Nevo (2000, 2001)], BLP models became ubiquitous in the economics and marketing literatures. The breakfast cereal is an ideal application for the estimation of equilibrium models of horizontally differentiated products with discrete demands. The breakfast cereal is a noncompetitive industry where firms set prices strategically as they offer a portfolio of items located across a large product space defined by calories per serving, sugar, fiber and many other quality dimensions. Since in addition, the average consumer spends a very limited budget on breakfast items, econometricians can safely ignore income effects and assume quasi-linear preferences to model choice utilities. To study mergers in a horizontally differentiated industry, Nevo (2000) specify preferences as follows (ignoring possible location and time indices):

\[ u_{ij} = x_j \beta_i^* + \alpha_i^* p_j + \xi_j + \epsilon_{ij}, \quad i \in I, \quad j \in J, \quad \epsilon_{ij} \sim \text{EV1}, \]  
\[ (\alpha_i^*, \beta_i^*) = (\alpha, \beta) + \Pi D_i + \Sigma \nu_i, \quad \nu_i \sim N(0, I_{n+1}), \]

where \( x_j \) denotes the \( n \times 1 \) vector of observed product characteristics while \( p_j \) is the price of product \( j \) for the set of (inside) products available in each market, \( J \), with \( J = |J| \). If consumers prefer the outside good they obtain a payoff \( u_{io} = \epsilon_{io} \). Random coefficients of product characteristics, \( \beta_i^* \), or price responsiveness, \( \alpha_i^* \), accommodate idiosyncratic individual preferences. Preferences might be correlated to a \( d \)-vector of demographic traits \( D_i \) so that the \( (n + 1) \times d \) matrix \( \Pi \) of coefficients measures the effect of observable demographics on the valuation of product attributes, including price, also allows for cross-price elasticity to vary across markets with different demographic composition. To further account for individual preferences over unobservable product attributes, \( \nu_i \) captures mean-zero, unobserved preference shifters with a diagonal variance-covariance matrix \( \Sigma \). Lastly, the idiosyncratic unobserved preference by consumer \( i \) for product \( j \), \( \epsilon_{ij} \), follows the Type-I extreme value distribution across all products in \( J \).

We estimate Nevo’s full specification using the simulated breakfast cereal data of [Conlon and Gortmaker (2020)]. We also estimate three other specifications where we modify how idiosyncratic preferences affect price responsiveness. The four specifications are:

[A] \[ \alpha_i^* = \alpha + \sum_{k=1}^{d} \pi_{ak} D_i + \sigma_{ai} \nu_i, \] (Nevo – Full Model)  
(B) \[ \alpha_i^* = \alpha + \sigma_{ai} \nu_i, \] (Only Price Random Coefficient)  
[C] \[ \alpha_i^* = \alpha + \sum_{k=1}^{d} \pi_{ak} D_i, \] (Only Demographic Price Interactions)  
[D] \[ \alpha_i^* = \alpha, \] (No Price Interactions)
Estimates of the price related parameters of these four specifications are reported in Table I. Figure 1 highlights the importance of modeling heterogeneous price effects flexibly. Panel plots the elasticity and curvature estimates for all products for each specification. Notice that very few products violate the elastic demand constraint under any specification. We do not show areas where none of our demand estimates fall, such as the regions of concave ($\rho < 0$), very elastic ($\varepsilon > 8$) or very convex demands ($\rho > 2$). All four specifications generate estimated curvatures sufficiently lower than 2 and thus equating marginal revenue to marginal cost ensures profit maximization. We therefore limit our attention mostly to subconvex demands between the linear ($\rho = 0$) and the $CES$ loci $\rho = 1 + 1/\varepsilon$. Simple visual inspection shows a stark difference between right and left panels resulting exclusively from different model specifications, and particularly the flexibility of price responsiveness. The obvious conclusion is that the estimated demand is bounded to be log-concave as in the $MNL$ case unless the price random coefficient or demographic price interactions are significant and large enough relative to the estimated mean slope of demand.

The first (top-left) panel corresponds to the specification of Nevo (2000). The average elasticity and curvature estimates of Model A are $\bar{\varepsilon}_A = 3.62$ and $\bar{\rho}_A = 1.06$, respectively. There are products with both log-concave and log-convex demands in Nevo’s specification. In some few cases
Table 1: Breakfast Cereal: Price Related Estimates

<table>
<thead>
<tr>
<th>Specification</th>
<th>Means</th>
<th>Std. Dev.</th>
<th>Demographic Interactions (π’s)</th>
<th>Manifold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(α’s)</td>
<td>(σ’ s)</td>
<td>log(INCOME)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>log(INCOME)^2</td>
<td>CHILD</td>
</tr>
<tr>
<td></td>
<td>(14.8032)</td>
<td>(1.3402)</td>
<td>(270.4410)</td>
<td>(14.1012)</td>
</tr>
<tr>
<td>[B]</td>
<td>-30.9982</td>
<td>2.0216</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.9674)</td>
<td>(0.9367)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>[C]</td>
<td>-53.1367</td>
<td>—</td>
<td>444.7281</td>
<td>-22.3987</td>
</tr>
<tr>
<td></td>
<td>(12.1023)</td>
<td>—</td>
<td>(209.6548)</td>
<td>(10.7282)</td>
</tr>
<tr>
<td>[D]</td>
<td>-30.8902</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.9944)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Notes: GMM estimates of parameters related to price sensitivity using simulated breakfast cereal data as reported by Conlon and Gortmaker (2020). Robust standard errors are given in parentheses.

The estimates fall to the right of the CES dividing curve, on the superconvex region of demand where elasticity decreases (markup increases) as price increases. The curvature of a CES demand with elasticity ε = 3.62 is ρ_A^CES = 1.28. However, the bulk of the observations fall into the subconvex region where demand becomes more elastic (markup decreases) as price increases. Here is where the so-called Marshall’s Second Law of Demand holds, comprising both log-concave and log-convex demands to the left of the CES loci. Consequently, the pass-through rate at the average curvature estimate is 106%, more than complete.

For the other three cases we use the same demand specification of Nevo (2000) including random coefficient of product attributes, their demographic interaction and even the same instrumentation. We only modify the heterogeneity of price responsiveness according to (2b)-(2d). Parameter estimates involving price are reported in Table 1 for each specification. The second (top-right) panel corresponds to Model B where price heterogeneity is limited to the price random coefficient, which is first, very small relative to the slope of demand and second, not significant.

In an influential paper, Hellerstein and Goldberg (2013, §4.1.1) argue that ML is a demand system that is flexible enough to evaluate pass-through as it can accommodate a wide range of elasticities and superelasticities. Figure 1 shows that simply considering random coefficients might not be enough. We need to include significant price-demographic interactions (as these authors do) or estimate a spread enough distribution of idiosyncratic price sensitivity to approximate the behavior of a CES demand. It is obvious that not considering demographic price interactions seriously restricts demand curvature estimates to the log-concave region. If model A is correctly specified, then Model B generates a slightly upward biased average demand elasticity estimate, ě_B = 3.74, and a downward biased average curvature estimate, ě_B = 0.96 now predicting a qualitatively very different incomplete pass-through rate of 96%. Log-concave demands (ρ < 1) mitigate the effect of cost shifts leading to incomplete pass-through rates while log-convex demands (ρ > 1)

---

6 See Marshall (1920, Book III, Chapter IV, §2) and the discussion in Mrázová and Neary (2017, §I.B).
7 We discuss the relationship between curvature and superelasticity in Section 3.2 and later in Appendix A.4.
amplifies the cost shift effects with more than 100% pass-through rates. Model D, at the fourth (right-bottom) panel, ignores all idiosyncratic responses to price and produces results that are undistinguishable from Model B.

Nevo’s full model specification has a very small and non-significant price random coefficient, as reported in the first row of Table 1. Most of the idiosyncratic heterogeneity regarding price responses are thus captured by demographic interactions rather than by a pure price random coefficient. Notice that the size of the estimates of demographic interactions in Table 1 are several times larger than the corresponding estimate of the average slope of demand, $\hat{\alpha}$. This is the reason why Model C produces results so similar to Model A: nearly the same average demand elasticity, $\hat{\varepsilon}_C = 3.6$, and a slightly more convex demand, $\hat{\rho}_C = 1.08$, now leading to a 109% pass-through rate. Of course this is true only at the average elasticity and curvature estimates. The third (bottom-left) panel shows that many more observations fall on the superconvex region and that estimates are much closer to the CES case, particularly for those with lower demand elasticities.

As we indicated in the Introduction, Caplin and Nalebuff [1991b, Example 4.1] show that idiosyncratic attribute preferences over linear product attributes coefficients leads to a log-concave demand. This is clearly the case for models B and D. Our results indicate that the power of moving away from the log-concave, incomplete pass-through, region of demand by only considering idiosyncratic preferences for product attributes is negligible, something that our more formal analysis of Section 3 confirms.

Figure 2 shows the disparate predictions of ML when demand curvature is not robustly estimated. We analyze the distribution of price increases and the reduction in quantity sold after a 1% increase in marginal cost. The left panel shows that models A and C, both with demographic price interactions, predict more than complete pass-through rates while Models B and C, without such interactions predict incomplete pass-through rates of the 1% marginal cost increase, as any MNL model would do. This is not only a matter of differences among average pass-through rates
Figure 3: Distribution of Estimated Price Sensitivity

Figure Notes: Empirical distribution of the estimated individual slope of demand, $\hat{\alpha}_i$, for the four alternative specifications.

but also of the distribution of these effects that show little overlap. All models predict quantity reductions ranging between 7.86% and 8.55%, or about a 9% prediction discrepancy.

Figure 3 illustrates the impact of each specific modeling assumption on the distribution of price responsiveness across individuals. Models A (solid blue line) and C (solid orange) produce nearly identical asymmetric distributions of price sensitivity. Incorporating demographic interactions, based on the distribution of demographic traits, results in strongly left skewed distributions. It is this feature, the skewness of the distribution of price sensitivity what leads to a more curved demand and increases pass-through rates closer to the CES loci or past it. Model D assumes away any individual heterogeneity. The slope of demand is the same for all consumers and its distribution is degenerate (dashed black). Model B assumes a symmetric normal distribution for the price random coefficient. The small price random coefficient estimate results on a tight symmetric distribution of price sensitivity around the mean demand slope (solid green).

The remaining two cases do not account for demographic price interactions and they fail to recover a rich variety of individual responses to price variations. Model D assumes away any individual heterogeneity. Thus, by construction, the slope of demand is the same for all consumers and its distribution becomes degenerate (black dashed). Model B assumes a symmetric normal distribution for the price random coefficient. The small price random coefficient estimate results on a tight symmetric distribution of price sensitivity around the mean demand slope (solid green).

It is striking how different the mixing distribution of idiosyncratic price responsiveness across these four specifications are. The distribution of consumer demographics combined with
the price response depicts a situation where consumers appear to be wildly different, many of them not responding at all to prices. This results in a reversed lognormal distribution of price sensitivity. At the same time it is shocking that the price random coefficient of Model B is unable to capture such idiosyncratic heterogeneity of price responsiveness when we omit these demographic price interactions. A possible explanation is that the assumed symmetric normal distribution is incapable of accommodating the real distribution of idiosyncratic price sensitivity, which Models A and C clearly hint at being strongly skewed.

Before concluding this section it is worth mentioning that many of the features of the estimates of the demand for breakfast cereal are common across the empirical IO literature. For instance, although there is significant curvature dispersion, most products correspond to log-convex demands with $1 \leq \hat{p}_A \leq 1.2$. Price elasticity is increasing in price for subconvex demands. Subconvexity of demand estimates for spirits is also behind the Laffer curve of Miravete et al. (2018), the markup adjustments of Indian firms after India’s 1991 trade liberalization in De Loecker, Goldberg, Khandelwal and Pavcnik (2016) and of coffee roasters in Nakamura and Zerom (2010). In the next section we point at the skewness of the choice probability is the key driver of the curvature of a discrete choice demand model.

---

8 Notice that the demographics interacted with price by Nevo (2001) do not enter directly into the specification of preferences. Compiani and Kitamura (2016) §2.2 show that this exclusion restriction, demographics entering the mixing distribution not entering the component distribution, helps identify the parameters of this mixture model.

9 Their static coffee demand estimates are consistent with incomplete pass-through: Only their price-income interaction is significant, e.g., see Nakamura and Zerom (2010 Table 5), but not very large, just 18% relative to the average slope of demand. This depicts a situation very similar to the left panels of Figure 1 with most observations falling on the log-concave demand region with incomplete pass-through. However, they assume a lognormal distribution of income possibly expanding the range of demand curvatures to the log-convex regions due to the large effect on price sensitivity of individuals on the long tail of the distribution of income.
3 A Primer on Demand Manifolds

Demand elasticity and curvature are closely related to each other through necessary and sufficient conditions for profit maximization. We first explore this link and describe the features of demand manifolds for different discrete choice models, including the important property of demand subconvexity that conditions pricing and markup behavior in response to cost shocks. This conceptual framework helps us illustrate how adding flexibility to the treatment of idiosyncratic price responses expands the range of attainable elasticity-curvature combinations.

3.1 Basics of Demand Manifold Analysis

Let \( q(p) \) denote the direct demand function and \( p(q) \) the corresponding inverse demand function. Both are positive, continuous, decreasing (\( q(p) < 0 \) and \( p(q) < 0 \)) and three times differentiable. Mrázová and Neary (2017) prove that with the exception of the CES these conditions are sufficient to characterize a well-defined smooth equilibrium relationship connecting elasticity \( \varepsilon(q) \) and curvature \( \rho(q) \), the so-called demand manifold, for any profit maximizing firm.\(^{10}\)

A profit maximizing monopolist with constant returns to scale equates marginal revenue and marginal cost. Demand must be elastic in equilibrium for any firm with market power.\(^{11}\)

\[
p(q) + q \cdot p_q(q) = p(q) \left[ 1 - \frac{1}{\varepsilon(q)} \right] = c > 0 \quad \iff \quad \varepsilon(q) = -\frac{p(q)}{q \cdot p_q(q)} > 1. \tag{3}
\]

Similarly, concavity of the profit function excludes the possibility of excessively convex demands. This sufficient condition requires that the marginal revenue function is not increasing at the equilibrium prices:\(^{12}\)

\[
2p_q(q) + q \cdot p_{qq}(q) = p_q(q) \left[ 2 - \rho(q) \right] < 0 \quad \iff \quad \rho(q) < 2, \tag{4}
\]

where curvature of the inverse demand function is defined as follows:

\[
\rho(q) = -q \cdot \frac{p_{qq}(q)}{p_q(q)} \quad \iff \quad \rho(p) = q(p) \cdot \frac{q_{pp}(p)}{[q(p)]^2}. \tag{5}
\]

The slope of demand plays a key role in the profit maximization necessary condition, which we rewrite in terms of the demand elasticity in equation (3). The sufficient profit maximization condition further restricts the slope of the marginal revenue function to be negative, which in turn

\(^{10}\)Mrázová and Neary (2017) also find conditions for this elasticity-curvature relation to be invariant to demand shifts, thus opening the door for elasticity and curvature to become sufficient statistics.

\(^{11}\)It is straightforward to write the equivalent first order condition for an oligopoly. Accounting for the number of inelastic demands at the observed equilibrium prices has become a customary specification test.

\(^{12}\)This is a sufficient condition for the monopoly case only. Caplin and Nalebuff (1991a, Proposition 11) prove the existence of single-product equilibria in the subconvex region of demand (to left of the CES loci). We are not aware of any equilibrium result for oligopoly industries on the superconvex region of demand. See Section 3.2 below.
we rewrite as a constraint on the curvature of demand in equation (4). Thus, if $\rho < 0$ demand is concave, linear when $\rho = 0$, and convex if $\rho > 0$. The sufficient condition for profit maximization is seldom used in empirical work and yet, it is the key to recover the equilibrium estimate of demand curvature, an unobservable magnitude of economic interest. The connection between demand curvature and pass-through rate has long been known (Cournot, 1838, §5):

$$\frac{dp}{dc} = \frac{1}{2 - \rho(q)} > 0. \quad (6)$$

For $\rho < 1$ demand is log-concave with an incomplete pass-through (“sub-pass-through”) while if demand is log-convex the pass-through rate exceeds 100% when $\rho > 1$ (“super-pass-through”). Complete pass-through occurs when $\rho = 1$.

### 3.2 Demand Subconvexity

There is a very specific one-to-one relationship between demand elasticity and curvature for the constant elasticity CES demand as already shown in Figure 1:

$$\rho = 1 + \frac{1}{\varepsilon} \iff \varepsilon = \frac{1}{\rho - 1}. \quad (7)$$

For the CES, markup, curvature and pass-through rate are invariant to price and sales. Furthermore, this is the only demand specification with this invariance property as the CES is the limiting case separating superconvex and subconvex demands, which we now define. A function $f(x)$ is subconvex if log[$f(x)$] is concave in log($x$). Mrázová and Neary (2017, Online Appendix B) show that:

$$\frac{d^2 \log q(p)}{d(\log p)^2} = -p \cdot \frac{d\varepsilon(p)}{dp} = -\varepsilon(p) \cdot \left[1 + \varepsilon(p) - \varepsilon(p)\rho(p)\right], \quad (8)$$

which is positive for superconvex and negative for subconvex demands, and zero for the CES because of (7). Thus, any demand function is superconvex if, for a given elasticity $\varepsilon$, its curvature $\rho$ exceeds the curvature of CES demand, $\rho > 1 + 1/\varepsilon$. If $\rho < 1 + 1/\varepsilon$, demand is subconvex.

Markups increase with price as demand becomes less elastic at higher prices if demand is superconvex. Conversely, markups decrease with price as demand becomes more elastic at higher prices for subconvex demands, i.e., Marshall’s Second Law of Demand. As for curvature, other than for the CES demand, it varies with price and sales thus providing a key identification argument for demand curvature: the effect of cost shifts are different for different price levels. The next section explores how typical preference specifications in discrete choice demand estimation conditions the shape and position of demand manifolds linking elasticity and curvature estimates.

---

13Superconvexity of $f(x)$ requires log[$f(x)$] to be convex in log($x$). Subconvexity of the direct demand function is identical to quasi-concavity of the firm’s profit function in own price where profits are strictly positive. See Caplin and Nalebuff (1991a, Proposition 11). Notice also that the last term in between brackets in equation (8) is precisely the definition of superelasticity, $S$ of Kimball (1995), frequently used in macroeconomics to measure demand-side real rigidities as price markups adjust when $S > 0$. 

---
4 Discrete Choice Demand Specification and Curvature

The choice of mixing distribution is an integral part of model specification. McFadden and Train (2000, §4.1) and Train (2009, §6) argue that a one-tail mixing distribution ensures that all individuals have the same sign response of the estimated mean effect in discrete choice models with micro data. Choosing a two-tail mixing distribution allows more heterogeneity up to the point that different individuals might have opposite valuations of the same attribute. Some might prefer large vehicles while others favor small ones. One-tail distributions are suggested to ensure, for instance, that all individuals in the sample have downward sloping demands.

This topic has attracted no attention in the aggregate discrete choice literature. Random coefficients in BLP models are customarily assumed to be normally distributed. The moments of the mixture choice probability integrate out the unobserved heterogeneity of the conditional moments of the component distribution, which the mixing distribution might enhance or soften (Lindsay, 1988, §2). It is difficult, however, to envision how the result of this integration translates into properties of the resulting demand, and in particular demand curvature.

In this section we explore the connection between the properties of the mixing distribution and demand curvature in discrete choice models with aggregate market data by making use of the demand manifold framework. We characterize the demand manifold in general, for any mixing distribution and functional form of expenses in the outside good. We then explore how changes in the mixing distribution alter the shape of demand manifolds and the range of estimable \((\varepsilon, \rho)\) combinations and use these insights to discuss the identification of demand curvature.

4.1 Consumer Preferences with a General Price Distribution

Consider the quasi-linear preference framework of Nevo (2000) with an additional object of interest: a mean-zero distribution of the price random coefficient, \(\Phi\), that is not necessarily symmetric:

\[
\begin{align*}
    u_{ij} &= x_j \beta^*_i + f_i(y_i, p_j) + \xi_j + \epsilon_{ij}, \\ 
    \beta^*_i &= \beta + \sigma_x \nu_i, \\ 
    f_i(y_i, p_j) &= \alpha^*_i(y_i - p_j) = (\alpha + \sigma_p \phi_i) \times (y_i - p_j), \\ 
    \nu_i &\sim N(0, I_n), \\ 
    \phi_i &\sim \Phi(0, 1),
\end{align*}
\]

where we normalized the outside option to zero and \(\alpha^*_i = (\alpha + \sigma_p \phi_i)\) denotes the price random coefficient. Given the very limited role we showed they play in Section 2, we assume that random coefficients of attributes, \(\nu_i\) are a mean zero, normally distributed variables. Notice that in this streamlined model we do not consider demographic interactions. The idea is to study the shape of demand manifolds as it changes with the properties of a mean-zero distribution of idiosyncratic
price responsiveness, Φ. The goal is to ensure the widest range of curvature as possible by avoiding distributions that unnecessarily restricts it.

4.2 Choice Distribution Moments and the Shape of Demand

Individual \( i \) purchases her preferred product \( j \) if:

\[
q_{ij}(p) = 1(u_{ij} \geq u_{ik}, \forall k \in \{0, 1, \ldots, J\}),
\]

which, for quasi-linear preferences, is equivalent to:

\[
(x_j - x_k)\beta_i^* - \alpha_i^*(p_j - p_k) + (\xi_j - \xi_k) + (\epsilon_{ij} - \epsilon_{ik}) \geq 0, \quad \forall k \in \{0, 1, \ldots, J\}.
\]

Because of the additive i.i.d. type-I extreme value distribution of \( \epsilon_{ij} \), individual \( i \)'s choice probability of product \( j \) is:

\[
\Pr_{ij}(p) = \frac{\exp \left( x_j\beta_i^* - \alpha_i^*p_j + \xi_j \right)}{\sum_{k=0}^{J} \exp \left( x_k\beta_i^* - \alpha_i^*p_k + \xi_k \right)},
\]

which follows from Holman and Marley’s theorem (Anderson, de Palma and Thisse 1992, §2.6.1).

Individual \( i \) makes a dichotomous decision regarding the purchase of product \( j \) with probability \( \Pr_{ij} \). This is the mean of an individual-specific Bernoulli distribution that varies with the vector of prices of the different alternatives. Aggregating over the measure of heterogeneous individuals (including random coefficients and demographic interactions) summarized by \( G(i) \), total demand for product \( j \) becomes:

\[
Q_j(p) = \int_{i \in \mathcal{I}} \Pr_{ij}(p) \, dG(i).
\]

We can now write the elements defining the demand manifold, i.e., elasticity and curvature of product \( j \) as defined in (3) and (5), respectively but referred to the choice model with preferences given by (9a)-(9c). For this purpose, let’s briefly consider a general price subfunction \( f_i(y_i, p_j) \). To simplify notation we write:

\[
f_{ij}' = \frac{\partial f_i(y_i, p_j)}{\partial p_j}, \quad \text{and} \quad f_{ij}'' = \frac{\partial^2 f_i(y_i, p_j)}{\partial p_j^2}.
\]

\[14\] Notice that individual income, \( y_i \) cancels out since \( f_i(y_i, p_j) \) is linear in income as written in (9c). Indeed, (Nevo 2000, 2001) does not include income in his preference specification and income appears only in demographic interactions with the random coefficients of price and product attributes. We prefer to include income in (9c) for consistency of notation and to ease the comparison of this quasi-linear preference specification to the specification in logs on other expenditures of BLP later in the paper.
The own-price demand elasticity of product $j$ can then be written as follows:

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

(15)

which amounts to a scale-free measure of aggregating individual price responses (demand slopes) weighted by their choice variance. Similarly, the demand curvature of our discrete choice model becomes:

$$
\rho_j(p) = \int_{i \in I} \mu_{ij} dG(i) \times \frac{\int (f''_{ij} \cdot \sigma^2_{ij} dG(i) + \int (f'_{ij})^2 \cdot \mu_{ij,3} dG(i))}{\left[ \int f'_{ij} \cdot \sigma^2_{ij} dG(i) \right]^2}.
$$

(16)

For the case of quasi-linear preferences, $f_i(y_i, p_j)$ in (9c) is linear in price $p_j$ and thus, $f''_{ij} = 0$, which simplifies demand curvature to:

$$
\rho_j(p) = \int_{i \in I} \mu_{ij} dG(i) \times \frac{\int (f'_{ij})^2 \cdot \mu_{ij,3} dG(i)}{\left[ \int f'_{ij} \cdot \sigma^2_{ij} dG(i) \right]^2}.
$$

(17)

The demand for product $Q_j(p)$ is the result of aggregating over the individual purchase probabilities $\mu_{ij} = P_{ij}$. Notice that $f'_{ij}$ represents the marginal effect of price $p_j$ on consumer $i$’s utility while $\sigma^2_{ij} = P_{ij}(1 - P_{ij})$ is the variance of her Bernoulli discrete choice distribution while $\mu_{ij,3}$ is its non-standardized skewness. Derivations are detailed in Appendix A.

Next, combining elasticity (15) and curvature (16) we obtain the general expression for the demand manifold:

$$
\rho_j(p) = p_j^2 \cdot \frac{Q_j(p)}{\varepsilon_j^2(p)} \cdot \left[ \int f''_{ij} \cdot \sigma^2_{ij} dG(i) + \int (f'_{ij})^2 \cdot \mu_{ij,3} dG(i) \right].
$$

(18)

Notice that for quasi-linear preferences, $f''_{ij} = 0$, and elasticity and curvature are inversely related as long as $\mu_{ij,3} > 0$, i.e., with positively skewed choice distributions. Demand specification, both through the concavity or convexity of price subfunction $f_i(p_j)$ and the moments of the choice distribution, affect both the slope of demand (elasticity) and the slope of the demand manifold (curvature). Thus, the skewness of the distribution of idiosyncratic price sensitivity plays a major role in determining the range of attainable curvatures, confirming the empirical results documented in Figures 2 and 3.

In the following subsections we first characterize the demand manifolds of the MNL and ML models. We then discuss how the commonly assumed symmetric normal distribution of price responsiveness restricts the set of estimable elasticity and curvature combinations and argue that we can expand this set by considering an asymmetric mixing distribution. In particular, we explore

---

15 We postpone a detailed discussion of the effect of nonlinearity of $f_i(y_i, p_j)$ on demand curvature to Section 6.
how the moments of the mixing distribution of price random coefficients, $\Phi(0, \sigma_p^2)$, affects demand elasticity and curvature as they influence the distribution of individual choices.

### 4.3 Demand Manifolds of Discrete Choice Models

We now explore the demand manifolds of the following discrete choice models: MNL, CES, ML with random coefficients on product attributes, and lastly, ML with price random coefficients.

**Multinomial Logit.** Since there is no unobserved heterogeneity, $\mathbb{P}_{ij} = \mathbb{P}_j = s_j(p)$, the market share of product $j$, for all individuals. Elasticity and curvature now become:

$$
\varepsilon_j(p) = \alpha p_j \left(1 - \mathbb{P}_j\right), \quad (19a)
$$

$$
\rho_j(p) = \frac{1 - 2\mathbb{P}_j}{1 - \mathbb{P}_j} < 1. \quad (19b)
$$

The expression of MNL demand elasticity is well-known in the literature. Equation (19b) shows that MNL demand might be concave, $\rho_j(p) < 0$, only in very concentrated markets where a single product exceeds 50% of sales. Furthermore, if there are only two products in a market, they can split it equally, $\mathbb{P}_1 = \mathbb{P}_2$, only if $\rho_j(p) = 0$, i.e., if both demands are linear.\(^\text{16}\) Thus, for less concentrated industry configurations demand is convex, $\rho_j(p) > 0$. The log-concavity of the MNL demand is obvious as $\rho_j(p) < 1$ for all possible prices and thus, the associated pass-through rate of any MNL estimate is necessarily incomplete.

Since MNL demand curvature (19b) is decreasing in $\mathbb{P}_j$, the pass-through rate increases to become arbitrarily close to complete in an atomized industry because $\mathbb{P}_j$ become smaller as the number of products grows.\(^\text{17}\) Oligopolies with multiple firms offering tens if not hundreds of products are common in practice. Such is the case of automobiles, breakfast cereal, spirits, and others frequently studied in IO. It should thus not be surprising that applied economists and practitioners systematically document near complete pass-through rates when estimating MNL or even ML discrete choice models of demand with limited idiosyncratic price sensitivity.

Combining the expressions (19a)-(19b) for MNL own-price elasticity and curvature we obtain the MNL demand manifold:

$$
\rho_j(p) = \frac{\alpha p_j \left(1 - 2\mathbb{P}_j\right)}{\varepsilon_j(p)}. \quad (20)
$$

Therefore, for any given curvature, demand is less elastic the larger the market share of the product and conversely, for any demand elasticity, a larger market share of a product reduces demand curvature. The left panel of Figure [4] shows one set of combinations of price elasticities

\(^\text{16}\) The impossibility result of Jaffe and Weyl (2010) is then immediate: if there are more than two products, all market shares cannot add up to 50% and demands cannot all be linear.

\(^\text{17}\) The effect of market size on curvature is well known in the literature. See Mrázová and Neary (2017, Appendix B.5).
The left panel shows one MNL demand manifold) while the right panel shows few other manifolds for a ML model with random coefficients on product attributes. Stars identify the combination of elasticity and curvature for an arbitrarily high equilibrium price $p_j = 8$.

and curvatures for the MNL model. The manifold is color coded to indicate the value of the equilibrium price induced by any particular combination of demand features. Since the log-convex MNL is also subconvex, higher equilibrium prices are always associated to more elastic demands, which necessarily implies lower equilibrium markups at higher prices.

**MNL vs. CES.** The decreasing and convex blue line of Figure 4 represents the CES manifold of equation (7). Both, MNL and CES, can easily accommodate any elasticity but they produce estimates of demand curvature that are qualitatively very different and that are imposed on the data by mere demand specification. This is perhaps the most striking implication of choosing the CES over the MNL model rather than their ability to produce some price elasticity estimates or generate more or less flexible substitution patterns in the absence of unobserved individual heterogeneity.

Both models are consistent with utility maximization. The MNL case is well-known and does not need further justification. Anderson et al. (1992, §3.7) and (1987) first show that the CES utility function of the representative consumer model (Dixit and Stiglitz, 1977) can be generated by a discrete choice model where individuals devote a fraction of their income to purchase a variable amount of one of these products. Thus, CES arises naturally in the context of discrete-continuous models (Hanemann, 1984) where consumers allocate a share of their budget to purchase several unite of the preferred alternative. On the contrary, MNL is most appropriate when consumers have unit demands. However, both models suffer from IIA and impose strong restrictions on the substitution patterns that they can generate unless they allow for random coefficients either with a ML model for unit demands or a mixed-CES model. But regardless of whether they allow for idiosyncratic preferences or not, MNL and CES still produce very different results because of their strikingly different demand curvatures.

18For manifold figures we assume one inside good, $|J| = 1$; one thousand consumers $|I| = 1,000$; normalize the average valuation of product attribute “quality” to $\beta = 1$; and an average price responsiveness of $\alpha = 0.5$. 
**ML with Attribute Random Coefficients.** Accounting for idiosyncratic individual preferences on product attributes is sufficient to accommodate unrestricted substitution patterns. However, while \( ML \) might dominate \( CES \) because of the flexibility of cross-price elasticity estimates, \( ML \) fails to expand the range of attainable demand curvatures. This is because both \( MNL \) and the normal price mixing distributions are both log-concave, thus resulting in a log-concave demand. Caplin and Nalebuff (1991b, Proposition 2) prove that the mixture distribution preserves the curvature properties of its components. In particular, since the logit and the commonly assumed normal mixing distribution are both log-concave, the resulting demand is also log-concave.\(^{19}\)

The right panel of Figure 4 shows that idiosyncratic attribute preferences fail to turn demand log-convex, with several demand manifolds for \( ML \) with different values of \( \sigma \), the dispersion of the distribution of random coefficients of product attributes, all falling in the log-concave region. This graphic analysis highlights the serious limitations of using the \( MNL \) or \( ML \) without price random coefficients to study pass-through in noncompetitive environments. While these models are capable of capturing a wide range of elasticities, their demands are also necessarily log-concave.\(^{20}\) If the data is generated by a log-convex demand, a misspecified \( MNL \) or \( ML \) without price random coefficients will produce upward biased elasticity estimates. In order to generate a demand curvature as close as possible to the true value of \( \rho > 1 \) the only possibility is to generate predictions of more elastic demands. Figure 4 clearly illustrates that flexible substitution patterns are only part of the problem and that we should seek models that can also expand the range of demand curvatures into the log-convex region of demand.

**ML with Price Random Coefficients.** The previous discussion might give the false impression that in order to expand the range of possible curvatures researchers have the false choice between a unit, log-concave discrete choice demand model, or the continuous log-convex \( CES \) demand. Our analysis in Section 4.2, and a main result of this paper, shows that is possible to extend demand curvature estimates of the \( ML \) model over the log-convex demand region if we allow for sufficiently flexible unobserved heterogeneity of price sensitivity in a unit demand model.

How can we expand the range of curvatures that our \( ML \) estimates can accommodate? The log-concave region of Figure 5 shows that changing the value of any attribute, e.g., “quality” \( X + 1, X + 2, \ldots \) of the inside good, shifts the demand manifold upwards therefore expanding the set of log-concave combinations \((\varepsilon, \rho)\) that we can jointly accommodate although pass-through is still incomplete. Quality or any other desirable attribute shifts demand upwards, thus decreasing

\(^{19}\)Their shape results apply the theorems of Prékopa (1971) and Borell (1975) to the preservation of \( \rho \)-concavity (Avriel, 1972) under integration. Log-concave functions such as the logit component, which is linear in parameters, corresponds to \( \rho = 0 \). The product of log-concave component and mixing distribution is also log-concave as well as its integral, the mixture distribution characterizing demand.

\(^{20}\)This is at odds with the mounting evidence of pass-through rates exceeding 100% in industries with horizontally differentiated products including groceries (Besley and Rosen, 1999), clothing and personal care items (Poterba, 1996); broad categories of branded retail products (Besanko, Dubé and Gupta, 2005); gasoline and diesel fuel (Marion and Muehlegger, 2011); as well as beer, wine, and spirits (Cook, 1981; Young and Agnieszka, 2002; Kenkel, 2005) among others.
both price elasticity and demand curvature for a given price. Adding idiosyncratic preferences on product attributes will generate “deviations” around each manifold similar to those of the left panel of Figure 4 but still leading to log-concave demand estimates.

To explore how the choice of the mixing distribution affects the properties of the mixture ML estimates of elasticity and curvature we now ignore the random coefficients on product attributes, $\sigma_x = 0$ in equation (9b), and assume that the price random coefficient of equation (9c) is lognormally distributed: $\phi_i \sim \Phi(0, 1) = \text{Lognormal}(0, 1)$. Thus, as $\sigma_p \to 0$ the role of the price mixing distribution vanishes and price responsiveness is identical for all individuals, $\alpha^*_i = \alpha$. Thus, for $\sigma_p = 0$ the demand manifold coincides with the MNL of Figure 5. But as the variance of the mixing distribution increases with $\sigma_p$, the random component of price responsiveness is larger (and more skewed) relative to the mean demand slope $\alpha$. Thus, manifolds now cross into the log-convex region thus confirming our findings on the left panels of Figure 1: sizeable asymmetric heterogeneity of price responsiveness in a model with quasi-linear preferences is key to generate more demand curvature and allow (rather than predetermine) whether the estimated pass-through rate is incomplete or more than complete.
To highlight how the choice of mixing distribution shifts the demand manifold over the log-convex region, we write its general expression (18) for this particular case without attribute random coefficients and lognormally distributed price responses:

$$\rho_j(p) = p_j^2 \cdot \frac{Q_j(p)}{\epsilon^2_j(p)} \cdot \left[ \int f''_{ij}(\phi) \cdot \sigma^2_{ij}(\phi) d\Phi + \int (f'_{ij}(\phi))^2 \cdot \mu_{ij,3}(\phi) d\Phi \right] \times L,$$

(21)

where we substitute the measure function $G(i)$ for $\Phi(0,1) \times L$, where $L$ is the number of potential customers. Notice that $\sigma^2_{ij}$ and $\mu_{ij,3}$ are moments of the component distribution that are integrated out over the mixing lognormal distribution of the price random coefficients. They depend on the value of $\sigma_p$ through $\phi$ and together they define the corresponding moments of the resulting mixture $P_j$, the predicted market share of product $j$. This equation therefore indicates that the statistical properties of the mixing distribution affects the position and shape of the discrete choice demand manifold.

Equation (21) hints at the determinants of the range of estimable $(\epsilon, \text{curv})$. More increasing and convex utility price subfunctions allow for a larger range of demand curvatures for any given elasticity. For quasi-linear preferences, larger slope of demands allow for a wider range of estimable demand curvatures. This range of curvatures also expands for more spread and skewed the mixture component distribution for any given utility subfunction. Similarly, for any curvature, these same arguments make demand less elastic. The combined effect of using a skewed mixing distribution is to extend the coverage of the demand manifold towards CES into the lower-right corner of the log-convex demand region, a result that survives even when preferences are quasi-linear and $f''_{ij} = 0$. We thus turn our attention to the selection of a flexible skewed mixing distribution with the goal of maximizing the set of $(\epsilon, \rho)$ combinations on the log-convex region, thus allowing more than complete pass-through rates rather than imposing it as the CES does.

Figure 5 shows that an asymmetric mixing distribution expands the range of estimable curvatures for each possible value of the elasticity. The skewness of the lognormal distribution interacted with the moments of the mixture component distribution appears to play a major role in expanding the range of estimable $(\epsilon, \rho)$. This is the topic of Section 5 but first we need to address the issue of identification in order to evaluate whether we can credibly recover the mixing distribution.

### 4.4 Identification

Researchers price-quantity data that need to be rationalized by a model. The same transaction could be the result of competition or monopoly since marginal costs are not observable. Bresnahan (1982) famously suggested the use of demand rotations to separate demand shifters from other regressors that result in changes in the sensitivity of demand to prices, such as the introduction of substitute products in the market.
Figure 6 shows the interconnection between demand specification, random coefficients and curvature. The pair of data points \((p_1, q_1)\) and \((p_2, q_2)\) can be explained by black dashed logit demand specification, either through demand shifts captured by random coefficients (on attribute or prices). Alternatively, the specification of the red demand function, with different curvature could explain that same data without the need of idiosyncratic price responsiveness. In essence, any estimator is tasked with figuring out whether variation in price-quantity pairs are due to changes in demand \((\Delta \xi)\) or changes in cost which generate exogenous variation in prices. When the researcher uses a theory of demand which is sufficiently flexible to accommodate these price-quantity pairs via curvature, Figure 6 demonstrates the estimator gains flexibility to connect the dots.

The intuition of our identification strategy is to focus on the different pass-through rate a cost shift evaluated at different price levels. In the unlikely event of a constant pass-through demand, the price effect is the same regardless of the price at which this cost shift is evaluated.\(^{21}\) For any other demand function the effect of this marginal cost increase will be different at high and low prices. Any interaction between cost shifts and price levels should thus serve as a valid instrument to recover the curvature of a unit demand function through the indexing parameter of the distribution of price random coefficients that, as we have argued extensively, influences the range of estimable demand elasticities and curvature.

Identification of demand curvature in the discrete choice framework we propose follows the argument put forward by Berry and Haile (2014) and aligns closest with Gandhi and Houde (2020).\(^{21}\) This is the Bulow-Pfleiderer’s iso-convex demand function, which includes the linear demand as a particular case, studied by Mrázová and Neary (2017 Appendix B.9).
Consider the following inverse-demand $\psi^{-1}$, with exogenous characteristics $x_j$, and endogenous prices $p_j$:

$$
\psi_j^{-1}(s_t, x_t, p_t; \theta) = h(\omega_t; \theta) + C_t(\theta),
$$

where $s_t$ are market shares (purchase probabilities), $x_t$ are exogenous product characteristics, $p_t$ is the vector of endogenous prices, and $\theta$ is a demand-side parameter we seek to estimate. The term $C_t(\theta)$ is a market-specific constant. The term $\omega_t$ is the market state vector which summarizes competition among the product set. Gandhi and Houde (2020) shows that we can write element $k$ of $\omega_t$ as

$$
\{ s_{kt}, d^x_{j,k,t}, d^p_{j,k,t} \}
$$

where $d^x_{j,k,t} = x_{jt} - x_{kt}$ is the distance in characteristic space between product $k$ and its rival product $j$. Similarly, we can describe distance in terms of price, i.e., $d^p_{j,k,t} = p_{jt} - p_{kt}$, though now we face a problem of endogeneity as prices depend upon not changes in marginal cost but also demand shocks via changes in $\xi$. With endogenous prices, the conditional expectation of the inverse demand is taken with respect to the joint distribution of shares and prices $(s_t; p_t)$ given $(x_t; \omega_t)$ but of course these are determined endogenously via market power within the industry.

We address this issue by constructing exogenous prices via either a reduced-form hedonic price regression based on exogenous characteristics $x_t$, and cost shocks $\omega_t$, where we assume the researcher has knowledge of the latter shocks, or by constructing prices non-linearly using firm first-order condition as in Berry, Levinsohn and Pakes (1999) and Reynaert and Verboven (2013). For heuristics, consider the reduced-form approach:

$$
p_t = \gamma_0 + \gamma_1 x_t + \gamma_2 \omega_t + u_t.
$$

We runs the above regression and use the results to construct the vector of predicted (exogenous) prices $\hat{p}_t$. We then construct the price differentiation IV as in Gandhi and Houde (2020):

$$
Z_{jt} = \sum_r \left( \hat{p}_{rt} - \hat{p}_{jt} \right)^2.
$$

Note that equation (23) not only enables us to construct exogenous prices by separating price effects due to changes in demand (via $\xi$) from changes in cost (via $\omega$), it also amounts to 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 (24). Since curvature in the quasi-linear utility discrete choice model comes primarily through heterogeneity in price-sensitivity, equation (24) is useful in identifying the distribution of idiosyncratic price responsiveness, or alternatively the price random coefficient $\pi$.

A useful implication of the demand manifold analysis is that for subconvex demands, where oligopolistic equilibria may exists, all else equal, increases in price lead to more elastic demands and convergence of demand curvature to one. We have also documented that the shape of the demand manifold, or equivalently the rate at which increases in price make demand more elastic and curvature converge to one, is determined by the shape of distribution of price sensitivities among consumers which we assume to be fixed across time. We can therefore identify the distribution of
price sensitivities by interacting with exogenous product characteristics, including exogenous prices. Intuitively, these interactions identify the distribution of heterogenous price sensitivities by tracing the relationship between demand elasticities and curvature across the demand manifolds in product set using the time-series variation in cost shocks.

5 The Shape of Price Mixing Distribution

In this section we evaluate the importance of the assumed distribution of price random coefficients, and in particular the asymmetry or skewness of the distribution, to expand the range of estimable \((\varepsilon, \rho)\) and thus minimize the risk of misspecification. We first present our arguments by comparing the consequences of assuming either a normal or a lognormal distribution of price random coefficients. We then conduct a Monte-Carlo analysis to evaluate the estimation bias and the ability to conduct counterfactuals under each distributional assumption. We conclude by suggesting a flexible distribution for price random coefficients encompassing both symmetric and asymmetric cases.

5.1 Normal vs. Lognormal Distributions of Price Random Coefficients

We begin by comparing the ability of normal and lognormal mixing distribution to extend the range of demand elasticity and curvatures estimates in a unit-demand \(ML\) model. In choosing our preferred mixing distribution we seek to cover as much as possible of the subconvex region of demand. The implication of this criteria is that increasing prices in response to an increase in a common cost across all products results in all demands becoming more elastic and markups smaller for all products. This is an important regularity condition for the model to produce coherent counterfactual predictions. We want to minimize the likelihood that a common cost increase leads to markup reductions for some products and increases for others, making difficult (if not impossible) to find a new industry equilibrium. The comparison of the normal and lognormal mixing distribution highlights why we favor asymmetric mixing distributions over the commonly used normal distribution to model price random coefficients as their skewness extend the range of attainable curvatures at each elasticity value, and very particularly for less elastic demands, when market power is more prevalent.

The two top panels of Figure 7 represent demand, elasticity and curvature at different prices for both, a model with normal (left) and lognormal random coefficients (right). Starting from the black solid line representing the \(MNL\) case, demand shifts to the red solid line as we add price random coefficients. This shift corresponds to a positive idiosyncratic price effect, \(\sigma_p > 0\). Dashed lines represent how own price elasticity varies with price while dotted lines indicate the demand

\(^{22}\)Demographic price interactions are identified by variations of prices and demographics across time and markets. It is conceivable to obtain a market or time specific distribution estimate for \(\Phi\), whose indexing parameters could be made a function of income or other demographics varying over time and locations. Given a panel data structure we could obtain estimates of demographic interactions in a second stage via minimum distance estimation in a similar way as \(\text{Nevo (2001)}\) projects product-market fixed effects into a smaller space of product attributes.
Figure 7: Normal vs. Lognormal Mixing Distribution

(a) Demand Elasticity & Curvature: Standard Normal

(b) Demand Elasticity & Curvature: Lognormal

(c) Demand Manifold: Standard Normal

(d) Demand Manifold: Lognormal

Top panels represent demand, elasticity and curvature in the price-quantity plane while the bottom panels represent demand manifolds in the \((\varepsilon, \rho)\) plane. Demand specification is always the same except that in the left panels price random coefficients are normally distributed while in the right panels they are lognormally distributed.

curvature at each price. Both, elasticity and curvature are measured on the right axis and the top shaded area identifies the log-convex demand region. Demand curvature approaches \(\rho = 1\) as price increases since the \(MNL\) demand is log-concave. Elasticity increases monotonically in price as log-concave demands are also subconvex.

Own-price elasticity increases less with price when we include random coefficients relative to the \(MNL\) case. Price elasticity is always quasi-linear in price for the lognormal case.\(^{23}\) If price is high enough, the own-price elasticity of the normal mixing distribution model reaches a maximum (precisely as the demand manifold crosses the \(CES\) loci) and then decreases with price as we move into the superconvex region of demand. We can see this in the bottom panels for manifolds corresponding to large values of \(\sigma_p\).

Top panels also show that the addition of price random coefficients allows demand to be log-convex for both mixing distributions. Convexity of demand increases substantially with price

\(^{23}\) Chintagunta (2002) first documented empirically that demand elasticity is quasi-linearly increasing in price in \(ML\) models. Björnerstedt and Verboven (2016) §2A) attributed this property to the linearity of conditional utility in price rather than to the subconvexity of demand.
when the price random coefficient is normally distributed. Log-normal price random coefficients
at the current parameterization appear to quickly reach a maximum for low prices and converge
asymptotically to $\rho = 1$ as the price increases. This hints at demand manifolds that at some
price level become downward sloping in the log-convex region for the lognormal mixture but not
necessarily so for the normal distribution.\footnote{Appendix A.4 discusses how the slope of manifolds relate to the moments of the component distribution.}

Discrete choice models have little difficulty accommodating any elasticity estimate but
cannot deliver large values of demand curvatures, particularly for low demand elasticities. The two
bottom panels of Figure 7 show this rather limited range of log-convex $(\varepsilon, \rho)$ combinations generated
by a normal mixing distribution when restricted to generate counterfactuals on the subconvex region
of demand. For large values of $\sigma_p$, demand manifolds are increasing up to their crossing with the
$CES$ loci and then they decrease. This occurs for a large range of less elastic demands. Remember
that in the superconvex region of demand, increasing prices also increases markups as higher prices
reduce demand elasticity. Thus, data might be consistent with subconvexity, a widespread property
of demand ensuring more than complete pass-through for $1 < \rho < 1 + 1/\varepsilon$. But for products nearby
those manifolds, counterfactuals analysis might result in wildly different predictions. An increase
in price, perhaps motivated by an increase in marginal cost or taxation might thus result in markup
reductions for products on manifolds that fall entirely in the subconvex region while for those that
go across the $CES$, the same price increase requires a markup increase.

In order to evaluate the ability of each mixing distribution to generate the widest $(\varepsilon, \rho)$
combination falling entirely on the subconvex region of demand, we vary all model elements that
determine the position of the manifolds: quality $X$, price effect of idiosyncratic price response
$\sigma_p$, and variance $\omega^2$ of the mixing distribution). This subset is identified by the plum shaded
area in the log-convex region of demand. The subset is particularly small for the normal mixing
distribution and covers very elastic demands only. More importantly, it allows for very limited
curvature and pass-through rates. If the data generating process needs to accommodate large
demand curvatures, a normal mixing distribution will most likely produce upward biased elasticity
estimates, thus perhaps characterizing an industry as excessively competitive only as result of this
particular model specification.

The positive skewness of the lognormal distribution increases the range of attainable curva-
tures very substantially for any demand elasticity as indicated by equation (21). This is particularly
true for less elastic demands where firms enjoy more market power. Notice however that, even after
allowing for any combination of parameters, there are still (fewer) combinations of $(\varepsilon, \rho)$ close to
the $CES$ loci that are not covered by manifolds that always fall on the subconvex region of demand,
including their counterfactuals. However, the lognormal case covers most feasible solutions with an
assumed normal price random coefficient and expands substantially the set of $(\varepsilon, \rho)$ combinations
of well-behaved equilibria, particularly for less competitive market configurations, thus reducing
the chances of misspecification.
5.2 Monte-Carlo Analysis

Eugenio: I would include here the results of the Monte-Carlos comparing the effect of specifying a normal vs. a lognormal distribution. Then offer the generalized normal as a compromise where we can encompass both specifications. Then move on to study if we can recover the specification of the right data generating process.

In this section we evaluate the empirical effects of mis-specification of the mixing distribution via Monte-Carlo simulation. In each simulation we model $J = 10$ differentiated products competing in prices for $T = 100$ periods. Consumer indirect utility is quasi-linear:

$$u_{jt} = \beta_0 + \beta_1 x_{jt}^{(1)} + \sum_{k=1}^{K} (\beta_{2,k} + \sigma_{X,k} \nu_{ik}) x_{jt,k}^{(2)} + \left(\alpha + \sigma_p \phi_i\right)p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

We focus on a single product characteristic random coefficient (i.e., $K = 1$) and set $\sigma_X = 5$. Some product characteristics are observed by the researcher ($\{x_{jt}^{(1)}, x_{jt}^{(2)}\}$) while others are only observed by consumers and firms ($\xi_{jt}$). We draw observable characteristics from a uniform distribution and these characteristics vary across time. Unobservable characteristics ($\xi_{jt}$) are distributed standard normal and also vary across time. The price random coefficient can be generated by either a standard normal distribution or a lognormal distribution with mean zero and unit variance.

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 ($p_{jt}^*$) satisfy the set of $J$ first-order conditions for the firms:

$$p_{jt}^* = c_{jt} - s_j(\delta_t, p_{jt}^*; \sigma_X, \sigma_p) \times \left[ \frac{\partial s_j(\delta_t, p_{jt}^*; \sigma_X, \sigma_p)}{\partial p_{jt}^*} \right]^{-1}.$$  \hspace{1cm}(26)

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

$$\log c_{jt} = \gamma_0 + \gamma_1 \log x_{jt}^{(1)} + \gamma_2 \log x_{jt}^{(2)} + \omega_{jt} + \zeta_{jt}$$  \hspace{1cm}(27)

We set $\gamma = 1$ and draw cost shocks $\{\omega_t, \zeta_t\}$ from standard normal distributions. Importanty, we assume that the researcher observes $\omega_t$ which will provide important identification for the price random coefficient ($\sigma_p$).

We generate pricing equilibria in the true data generating processes at different combinations of elasticity and curvature. We do so by selecting mean price sensitivity ($\alpha$) and the price random coefficient ($\sigma_p$) to generate an average (absolute) own-price elasticity of 2.5 and one of four curvatures: 0.9,1.0,1.1,1.2. Figure illustrates these equilibria in the context of elasticity-curvature space. Note that we target average elasticities and curvatures across $J$ products and $T$ periods so
for each equilibrium there exists a cloud of product-period points so our choices of target average curvatures enables space for this cloud to extend to the CES frontier.

For each of the four different data generating processes in a given simulation, we consider the objective of a hypothetical researcher to estimate consumer demand given observed prices, quantities, and \( \omega \) cost shocks. We assume the researcher correctly specifies the outside option, the distribution which generates the random coefficients for product characteristics (\( \nu_i \)), and the firm pricing equation ((26)) but may incorrectly specify the distribution which generates the price random coefficient (\( \phi_i \)). Specifically, the researcher may make the common assumption that this distribution is generated by a standard normal or that it is generated by lognormal distribution. The former case is a standard empirical assumption while the latter represents the researcher including interactions between price and income since income distributions are often lognormal-distributed.

We have then four potential scenarios for each simulation with a specific target elasticity and target curvature which we present in Table (2).

<table>
<thead>
<tr>
<th>Scenario</th>
<th>DGP</th>
<th>Researcher</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Standard Normal</td>
<td>Standard Normal</td>
<td>Researcher correctly specifies symmetric distribution.</td>
</tr>
<tr>
<td>(2)</td>
<td>Lognormal</td>
<td>Lognormal</td>
<td>Researcher correctly specifies asymmetric distribution.</td>
</tr>
<tr>
<td>(3)</td>
<td>Standard Normal</td>
<td>Lognormal</td>
<td>Researcher incorrectly specifies asymmetric distribution.</td>
</tr>
<tr>
<td>(4)</td>
<td>Lognormal</td>
<td>Standard Normal</td>
<td>Researcher incorrectly specifies symmetric distribution.</td>
</tr>
</tbody>
</table>
Scenarios 1 and 2 indicate whether the GMM estimation is capable of recovering the true parameters in the data generating process. Scenarios 3 and 4 provide insight into the consequence of mis-specification of the price random coefficient.

In each estimation the researcher identifies the random coefficient for characteristics using the following Differentiation IV of Gandhi and Houde (2020):

\[ Z_{jt} = \sum_r (x_{rt} - x_{jt})^2. \]  

(28)

The researcher uses the cost shocks (\( \omega \)) to identify the price random coefficient. To do so, the researcher takes advantage of the panel structure of their data by first conducting a flexible hedonic price regression by projecting observed prices onto product fixed effects and product-level interactions with the observed cost shocks. This amounts to a flexible first-stage price regression (23) though the cost interactions allow for different levels of cost pass-through across the products. The price regression generates predicted prices (\( \hat{p}_t \)) which are a function solely of persistent differences across products (via product fixed effects) and variation in exogenous costs.

Define the period \( t \) distance between the exogenous prices of products \( r \) and \( j \) as \( d(\hat{p}_{rt}, \hat{p}_{jt}) = (\hat{p}_{rt} - \hat{p}_{jt}) \). We then construct the identifying instrument using the quadratic differences of products with prices close to product \( j \)

\[ Z^p_{jt} = \sum_{r \in D} d(\hat{p}_{rt}, \hat{p}_{jt})^2. \]  

(29)

where \( D \) is defined as the set of products with exogenous price within 1/2 standard deviation of product \( j \).

5.3 Flexible Price Mixing Distributions

**Eugenio: This is the section that needs more work. We need to decide what generalized distribution we are going to suggest (perhaps both of them) and run a Monte-Carlo analysis to show that we recover a skewed distribution when the DGP is that of a lognormal or similar, and a symmetric distribution when DGP is normal. If we include the Box-Cox Normal-Lognormal distribution we need to derive it in Appendix B.**

We favor using a one parameter asymmetric generalized normal distribution of Hosking (1990) and Nadarajah (2005) where a single parameter \( \kappa \) determines the skewness. The distribution is, by construction, centered around zero so that the population price sensitivity is spread around the mean value of the demand slope, \( \alpha \). The AGN, converges to the normal distribution when \( \kappa \to 0 \), it is right-skewed for \( \kappa < 0 \) and left-skewed for \( \kappa > 0 \). However, for the estimation se can concentrate on either positive or negative values of \( \kappa \) only as its probability density function is symmetric for identical values of \( \kappa \) of opposite sign and \( \pi \) will capture its effect. Appendix B.1 derives the one-parameter AGN distribution that we represent in Figure 9. This AGN distribution can thus
accommodate from symmetric normal to extremely skewed distributions beyond the skewness of other well-know asymmetric distributions.
6 Beyond Quasi-Linear Preferences

We have studied at length how the shape of the distribution of price random coefficient helps expand the range of attainable demand curvatures when estimating a discrete choice model of unit demand for differentiated products. We now study how income might add to a flexible treatment of idiosyncratic price responsiveness. The starting point is the common consensus that own-price demand elasticity is heavily influenced by income.

BLP does not include a price random coefficient and if we extrapolate our findings for the breakfast cereal case, demand for automobiles should be log-concave, as in the left panels of Figure 1. However this is not the case and the estimated demand of BLP is log-convex for most automobiles as shown in the left panel of Figure 10. Preferences in BLP are as in (9a)-(9b) but substituting the price subfunction in equation (9c) as follows:

\[ f_i(y_i, p_j) = \alpha \ln(y_i - p_j), \tag{30} \]

so that instead of estimating a price random coefficient, differences in income account for the heterogeneous individual price response. When other expenses enter linearly, \((y_i - p_j)\) as in Nevo (2000), rather than in logs, the effect of income cancels. The right panel of Figure 10 shows log-concave demand estimates for all automobile models obtained using the data from BLP and Nevo’s quasi-linear preferences. The assumed specification of the utility subfunction \(f_i(y_i, p_j)\) thus have important implications for the estimates of demand curvature and pass-through. Indeed, the logarithmic transformation of other expenses allows a unit demand discrete choice model to nearly maximize the range of estimable demand curvatures in the subconvex demand region.

Nevo’s linear expense and BLP’s log-linear specification are particular cases that might be restricting demand curvature and pass-through estimates. They are two unique cases of the following, more general specification based on the power transform of Box and Cox (1964):

\[ f_i(y_i, p_j) = \alpha (y_i - p_j)^{\lambda} = \begin{cases} \frac{(y_i - p_j)^{\lambda} - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \alpha \ln(y_i - p_j), & \text{if } \lambda = 0. \end{cases} \tag{31} \]

We can therefore study how changing the value of \(\lambda\) allows us to expand the range of estimable \((\varepsilon, \rho)\) from a maximum of \(\rho = 1\) for Nevo’s specification to the CES limit of BLP, both without price random coefficients. In essence, varying \(\lambda\) we can explore how \(f'_{ij}\) and \(f''_{ij}\) in equation (14) influence the estimates of the own-price elasticity (15), curvature (16) and the shape and position of the manifold (18).

The quasi-linear preferences of Nevo (2000) are appropriate for studying demand of products that are inexpensive relative to the average household income, i.e., when income effects are absent, but it implies a maximum curvature \(\rho = 1\) and full pass-through unless we include a price random
Figure 10: Income Effects and Demand Manifolds

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

coefficient. Results from the previous sections show that a unit demand discrete choice model can actually generate pass-through rates larger than 100%. The same identification arguments serve to recover the transformation parameter \( \lambda \) in a generalized \( BLP \) model without price random coefficients that is more appropriate to study demand for products that are expensive relative to the average household income, i.e., when income effects are likely to exist.\(^{25}\) Rather than using the logarithmic specification of \( BLP \), we adopt the widespread specification of Berry et al. (1999):

\[
f_i(y_i, p_j) = -\frac{\alpha_i}{y_i} p_j = -\alpha^*_i p_j .
\] (32)

This specification assumes that higher income individuals are less price responsive. Furthermore, income distribution accounts for the distribution of idiosyncratic price responsiveness. Berry et al. (1999) actually employ the lognormal estimate of the U.S. income distribution over the sample period. They also argue that this specification is a first order approximation to the logarithmic utility subfunction of \( BLP \). The first-order Maclaurin series (at \( p_j = 0 \)) of the Box-Cox transformation specification (31) is:\(^{26}\)

\[
f_i(y_i, p_j) = \alpha (y_i - p_j)^{\lambda} \simeq \alpha y_i^{\lambda} \alpha p_j - \frac{\alpha p_j}{y_i^{1-\lambda}} .
\] (33)

\(^{25}\) Using micro data, Griffith, Nesheim and O’Connell (2018) estimate a \( ML \) with nonlinear income effects with polynomials or other functions that are nonlinear in the outside good but do not encompass quasi-linear and logarithmic preferences as the Box-Cox transformation does and apply it to evaluate the pass-through of a fat tax on consumption of butter and margarine, which arguably induce zero income effects.

\(^{26}\) Notice that for \( \lambda = 1 \) leads exactly to Nevo’s specification \( \alpha (y_i - p_j) \) but for \( \lambda = 0 \) is \( \alpha \ln y_i - \alpha p_j / y_i \), which only coincides with (32) for \( y = 1 \). The consequence of this approximation discrepancy is that a preference model made of (6a) minus (6b) plus (32) is only approximately consistent with utility maximization. Roy’s identity lead to the following conditional individual demand \( d_i(y_i, p_j) = 1 / [1 + (1 - \lambda)p_j / y_i] \), which converges to 1 for quasi-linear preferences when \( \lambda \to 1 \) (Nevo’s case) or when the price of the chosen alternative \( j \) is exceedingly smaller than household \( i \)’s income (not exactly \( BLP \)’s demand for automobiles).
Starting with the demand manifold of the MNL model, \( X + 1, X + 2, \ldots \) indicate the demand manifolds of MNL models for higher quality levels of the inside good. The other manifolds refer to the ML model with price random coefficients. The smaller \( \omega \) is the more important is the random component of the slope of demand.

Based on this approximation, we can now obtain the predicted own-demand elasticity, curvature, and the expression of the manifold for this model:

\[
\varepsilon_j(p) = -\frac{p_j}{Q_j(p)} \int_{i \in I} -\frac{\alpha}{y_i^{1-\lambda}} \cdot \sigma_{ij}^2 \, dG(i),
\]

(34a)

\[
\rho_j(p) = \int_{i \in I} \mu_{ij} \, dG(i) \times \frac{\int (1-\lambda) \alpha^2 y_i^{-\lambda} \cdot \sigma_{ij}^2 \, dG(i) + \int \frac{\alpha^2}{y_i^{2(1-\lambda)}} \cdot \mu_{ij,3} \, dG(i)}{\left[ \int -\frac{\alpha}{y_i^{1-\lambda}} \cdot \sigma_{ij}^2 \, dG(i) \right]^2},
\]

(34b)

\[
\rho_j(p) = p_j^2 \cdot \frac{Q_j(p)}{\varepsilon_j^2(p)} \int_{i \in I} \frac{\alpha^2 \cdot [(1-\lambda) y_i^\lambda \cdot \sigma_{ij}^2 + \mu_{ij,3}]}{y_i^{2(1-\lambda)}} \, dG(i).
\]

(34c)

Figure 11 shows ...
7 Concluding Remarks

*To be written.*
References


Appendix

A Choice Probability Distribution and Demand Manifolds

A.1 Moments

Because of the additive i.i.d. type-I extreme value distribution of $\epsilon_{ij}$, the individual $i$’s choice probability of product $j$ given by (12) is also the mean of an individual-specific Bernoulli distribution:

$$\mu_{ij} = P_{ij}, \quad (A.1)$$

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

$$\sigma^2_{ij} = P_{ij}(1 - P_{ij}). \quad (A.2)$$

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

$$\mu_{ij,3} = P_{ij}(1 - P_{ij})(1 - 2P_{ij}), \quad (A.3)$$

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

$$\tilde{\mu}_{ij,3} = \frac{\mu_{ij,3}}{\sigma^3_{ij}} = \frac{P_{ij}(1 - P_{ij})(1 - 2P_{ij})}{\sqrt{P_{ij}(1 - P_{ij})}^3} = \frac{1 - 2P_{ij}}{\sqrt{P_{ij}(1 - P_{ij})}}, \quad (A.4)$$

where $\sigma^3_{ij}$ is the third raw moment of the individual choice probability distribution.

A.2 Moment Derivatives

We use the derivative of the choice probability (12) with respect to price repeatedly. It is straightforward:

$$P'_{ij} = \frac{\partial P_{ij}}{\partial p_{ij}} = f'_{ij} \cdot P_{ij}(1 - P_{ij}). \quad (A.5)$$

The derivative of the variance with respect to price is:

$$\frac{\partial \sigma^2_{ij}}{\partial p_{ij}} = \frac{\partial P_{ij}(1 - P_{ij})}{\partial p_{ij}} = P_{ij}(1 - P_{ij}) - P_{ij}P'_{ij} = f'_{ij} \cdot P_{ij}(1 - P_{ij})(1 - 2P_{ij}) = f'_{ij} \cdot \mu_{ij,3}. \quad (A.6)$$

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

$$\mu'_{ij,3} = \left[(1 - P_{ij})^2 - 4P_{ij}(1 - P_{ij}) + P_{ij}^2\right]P'_{ij} = \left[(1 - 2P_{ij})^2 - 2P_{ij}(1 - P_{ij})\right]f'_{ij} \cdot P_{ij}(1 - P_{ij}). \quad (A.7)$$
A.3 Demand Manifold

We obtain the demand elasticity of product \( j \) by differentiating (13) with respect to \( p \) and substituting (A.5):

\[
\varepsilon_j(p) \equiv -p_j \frac{Q_j(p)}{Q_j(p)} \cdot \frac{\partial Q_j(p)}{\partial p_j} = -p_j \int_{i \in I} f'_{ij} \cdot P_{ij} (1 - P_{ij}) \, dG(i), \tag{A.8}
\]

while the inverse demand curvature of product \( j \) is:

\[
\rho_j(p) \equiv Q_j(p) \cdot \frac{\partial^2 Q_j(p)}{\partial p^2_j} = \int_{i \in I} P_{ij} dG(i) \times \left[ \int f''_{ij} \cdot P_{ij} (1 - P_{ij}) \, dG(i) + \left( \int f'_{ij} \cdot P_{ij} (1 - P_{ij}) \, dG(i) \right)^2 \right]. \tag{A.9}
\]

Equations (15) and (16) follow after substituting (13), (A.2) and (A.3) into these expression. Combining elasticity and curvature we obtain the expression for the demand manifold (18):

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

A.4 Subconvexity, Superelasticity and Manifold Slope

Demand is subconvex to the left of the \( CES \) loci, when own-price elasticity increases with price, which according to (8) requires:

\[
\frac{d\varepsilon_j(p)}{dp} = \frac{\varepsilon_j(p)}{p} [1 + \varepsilon - \varepsilon \rho] = \Psi \left[ p, f'_{ij}, f''_{ij}, P_{ij}, \sigma^2_{ij}, \mu_{ij,3} \right] > 0, \tag{A.11}
\]

which is a nonlinear function depending on shape of the utility price subfunction and the moments of the choice distribution (A.1)–(A.3) as they enter elasticity (A.8) and curvature (A.9).

It is interesting to connect the elements of the utility specification to Kimball’s superelasticity, \( S = 1 + \varepsilon - \varepsilon \rho \) commonly used to evaluate how far the market configuration is from the \( CES \) measured by relative pass-through:

\[
\frac{d \log p}{d \log c} = \frac{\varepsilon - 1}{\varepsilon}, \quad \frac{1}{2 - \rho} = \frac{\varepsilon - 1}{\varepsilon - 1 + S}, \tag{A.12}
\]

which is 1 (100% relative pass-through) for the \( CES \) when \( S = 0 \). If demand is subconvex, \( S > 0 \), elasticity increases with prices forcing firms to reduce their markup so that relative pass-through is less than 100%. Combining again elasticity (A.8), curvature (A.9) and the moments of the choice distribution (A.1)–(A.3) we have:
\[ S = 1 + \varepsilon + p_j \cdot \frac{\int f''_{ij} \cdot \sigma^2_{ij} dG(i) + \int \left( f'_{ij} \right)^2 \cdot \mu_{ij,3} dG(i)}{\int f'_{ij} \cdot \sigma^2_{ij} dG(i)} . \]  

(A.13)

The sum \( 1 + \varepsilon > 0 \) since profit maximizing firms price on the elastic region of demand. It follows that if contemplate quasi-linear preferences, \( f''_{ij} = 0 \), a symmetric or positively skewed choice distribution, \( \mu_{ij,3} \geq 0 \) suffices for \( S > 0 \) and demand to be subconvex. A convex price subfunction, \( f''_{ij} > 0 \), also favors demand subconvexity. It is only if the price subfunction is sufficiently concave, \( f''_{ij} << 0 \) that the superelasticity might become negative, i.e., a superconcave demand were relative pass-through exceeds 100%.

Since the CES loci is downward sloping, demand manifolds become eventually downward sloping for subconvex demands. Since elasticity increases with price, a downward sloping manifold requires that curvature decreases with price in that region. Deriving curvature (16) involves third derivatives of the price subfunction \( f_i(p_j) \) and thus we focus here on quasi-linear preferences and the more restrictive curvature expression (17).

\[
\frac{dp_j(p)}{dp} = \frac{\int \left( f'_{ij} \right)^2 \cdot \mu_{ij,3} dG(i) + \int \mu_{ij,3} dG(i) \times \int \left( f'_{ij} \right)^2 \cdot \mu_{ij,3} dG(i)}{\left[ \int f'_{ij} \cdot \sigma^2_{ij} dG(i) \right]^2} 
\]

(A.14)

\[
- \frac{\int \mu_{ij,3} dG(i) \times 2 \left[ \int \left( f'_{ij} \right)^2 \cdot \mu_{ij,3} dG(i) \right]^2}{\left[ \int f'_{ij} \cdot \sigma^2_{ij} dG(i) \right]^3} .
\]

The sign of this derivative is still difficult to evaluate. If we further assume that the choice distribution is symmetric, \( \mu_{ij,3} = 0 \) and the sign of this derivative depends exclusively on the sign of \( \mu'_{ij,3} \), which we can rewrite from (A.7) as follows after making use of (A.2) and (A.3):

\[
\mu'_{ij,3} = \left[ \frac{\mu_{ij,3}}{\sigma^2_{ij}} \right]^2 - 2f'_{ij} \sigma^2_{ij} \cdot f'_{ij} \cdot \sigma^2_{ij} .
\]

(A.15)

Notice that for symmetric choice distributions \( \mu_{ij,3} = 0 \) and \( \mu'_{ij,3} = -2f'_{ij} \sigma^4_{ij} > 0 \) because \( f'_{ij} = \sigma_i^* < 0 \). In this case \( dp/dp > 0 \) and the demand manifold ends up crossing the CES loci into the superconvex region. We therefore need a skewed choice distribution to ensure that \( \mu'_{ij,3} > 0 \) and \( dp/dp < 0 \) so that demand manifolds eventually become downward sloping for quasi-linear preferences.
B A General Mixing Distribution

B.1 The Asymmetric Generalized Normal Distribution

Idiosyncratic demand sensitivity is modeled as $\alpha_i^* = \alpha + \pi \phi_i$, where $\alpha$ is the mean slope of demand and $\pi$ captures the effect on price heterogeneity of preferences across individuals. We model draws of individual types $\phi_i$ after a three-parameter Asymmetric Generalized Normal distribution defined as follows (Hosking, 1990; Nadarajah, 2005):

\[
\text{Prob}(\phi < x; \iota, \zeta, \kappa) = \Phi_N(y) \text{ where } y = \begin{cases} 
-\frac{1}{\kappa} \log \left(1 - \frac{\kappa(x - \iota)}{\zeta}\right), & \text{if } \kappa \neq 0, \\
\frac{x - \iota}{\zeta}, & \text{if } \kappa = 0,
\end{cases}
\]  

(B.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 $\kappa < 0$ so that the support of the distribution is $(\iota + 1/\kappa, \infty)$. The distribution is right-skewed, mimicking a lognormal distribution for $\kappa = -1$ and converging to a normal distribution as $\kappa \to 0$. Furthermore we center the distribution around the mean slope:

\[
E[\phi] = \iota - \frac{\zeta}{\kappa} \left(e^{\kappa^2/2} - 1\right) = 0,
\]  

(B.2)

so that:

\[
\iota = \frac{1}{\kappa} \left(e^{\kappa^2/2} - 1\right).
\]  

(B.3)

The one-parameter AGN distribution can then be written as:

\[
\text{Prob}(\phi < x; \kappa) = \Phi_N(y) \text{ where } y = \begin{cases} 
-\frac{\log \left(e^{\kappa^2/2} - \kappa x\right)}{\kappa}, & \text{if } \kappa \neq 0, \\
\frac{x - \iota}{\zeta}, & \text{if } \kappa = 0,
\end{cases}
\]  

(B.4)

with mean, variance, and skewness:

\[
\mu[\phi; \kappa] = 0,
\]  

(B.5)

\[
\sigma^2[\phi; \kappa] = \frac{e^{\kappa^2/2}(e^{\kappa^2/2} - 1)}{\kappa^2},
\]  

(B.6)

\[
\tilde{\mu}_3[\phi; \kappa] = \frac{3e^{\kappa^2/2} - 3e^{3\kappa^2/2} - 2}{(e^{\kappa^2/2} - 1)^{3/2}}.
\]  

(B.7)
Eugenio: According to Wikipedia the skewness should have a negative sign for $\kappa < 0$ but I do not see it. Both exponential and Rayleigh have positive skewness and the same shape than the AGN with negative $\kappa$. Why did you not include it? I am a little lost.
Eugenio: I am not sure what to make out of these distributions. Yes, as the variance increases eventually there are some individual upward sloping demands. So? Is this worth mentioning in the text? I do not see the connection with the curls but I might be wrong.

B.2 Distribution Variance and Demand Slope