DISCUSSION PAPER SERIES

DP16981

Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation

Cameron Birchall and Frank Verboven

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation

Cameron Birchall and Frank Verboven

Discussion Paper DP16981
Published 31 January 2022
Submitted 27 January 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Cameron Birchall and Frank Verboven

Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation

Abstract

We extend BLP's aggregate discrete-choice model of product differentiation to create more flexibility in the price functional form. We apply a Box-Cox specification, which relaxes the typical unit demand assumption and creates flexibility on demand curvature. The model provides a unifying framework for mixed logit and mixed CES models. Our illustrative application to the ready-to-eat cereals market shows that the cross-sectional relation between price elasticities and average prices per product is more in line with descriptive elasticity patterns. Furthermore, it suggests lower cross-price elasticities between similarly priced products than in more restrictive specifications.

JEL Classification: N/A

Keywords: BLP, demand curvature, Mixed logit, mixed CES

Cameron Birchall - cameron.birchall@kuleuven.be
KU Leuven

Frank Verboven - frank.verboven@kuleuven.be
KU Leuven and CEPR

Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation*

Cameron Birchall† Frank Verboven‡

January 27, 2022

Abstract

We extend BLP's aggregate discrete-choice model of product differentiation to create more flexibility in the price functional form. We apply a Box-Cox specification, which relaxes the typical unit demand assumption and creates flexibility on demand curvature. The model provides a unifying framework for mixed logit and mixed CES models. Our illustrative application to the ready-to-eat cereals market shows that the cross-sectional relation between price elasticities and average prices per product is more in line with descriptive elasticity patterns. Furthermore, it suggests lower cross-price elasticities between similarly priced products than in more restrictive specifications.

Keywords: BLP, demand curvature, mixed logit, mixed CES

*Acknowledgements: We thank Jan De Loecker, Jo Van Biesebroeck and participants at an IO-Leuven workshop for useful comments. Financial support from a Methusalem grant is gratefully acknowledged. The conclusions drawn from the IRI data do not reflect the views of IRI. IRI is not responsible for and was not involved in analyzing and preparing the results reported herein.

†KU Leuven, email: cp.birchall@gmail.com

‡KU Leuven, email: frank.verboven@kuleuven.be

1 Introduction

Substitution patterns between differentiated products are crucial to understand many important economic questions in industrial organization, international trade, public economics and other fields. In their pioneering contributions, Berry (1994) and Berry, Levinsohn, and Pakes (1995, hereafter BLP) developed a discrete-choice random coefficients logit model to account for unobserved consumer heterogeneity in the valuation of product characteristics. The popularity of the BLP model stems from its ability to generate rich substitution patterns using only market-level sales data and a limited number of parameters.

The literature has paid considerable attention to account for unobserved consumer heterogeneity through flexible specifications for the random coefficients. However, it has largely neglected the role of demand curvature, i.e. the functional form through which a product's price enters the consumers' indirect utility. Most of the discrete-choice literature using market-level sales data has assumed that utility is linear in price or, more generally, additive in income and price, so that utility-maximizing consumers purchase a single unit of their preferred product. This functional form implies a tendency for price elasticities to be increasing in price. This is most evident for logit and nested logit models, where price elastictities are essentially linearly increasing with prices. Nevertheless, the typical random coefficients logit models also contain restrictions on demand curvature, and it remains an open question how this may bias parameter estimates. Björnerstedt and Verboven (2016) consider an alternative utility specification where utility is linear in the logarithm of both income and price. In this specification consumers have unit-elastic demand for their preferred products, which implies a tendency for price elasticities to be independent of price.[1] From a different angle, Adao, Costinot, and Donaldson (2017) and Dubé, Hortaçsu, and Joo (2021) posit essentially the same empirical model by directly incorporating random coefficients in a representative consumer CES demand model. Adao et al. (2017) label this a mixed CES, as opposed to BLP's mixed logit model.

Against this background, we relax the demand curvature restrictions that are implicit in aggregate discrete choice demand models by introducing a simple yet flexible Box-Cox transformation of price and income (from Box and Cox, 1964). This joint Box-Cox and BLP

[1] Nair, Dubé, and Chintagunta (2005) take a related approach, with a more complicated income term.

model implies that consumers do not necessarily have perfectly inelastic or unit-elastic demand for their preferred products. Our approach is attractive for at least three reasons. First, the joint model permits richer substitution patterns by allowing for a more flexible price functional form in addition to unobserved consumer heterogeneity. This breaks the mechanical link between price elasticities and prices, which may have been responsible for biased elasticity estimates even under rich consumer heterogeneity. Second, the Box-Cox specification nests both BLP's mixed logit model and the mixed CES model as special cases, and hence provides a unifying framework for existing models in various fields. Third, our specification is tractable because it requires only a single additional parameter.

To illustrate our demand framework, we apply it to the "Ready-to-Eat" cereal market, which several papers explain is particularly well suited for estimating demand in differentiated product markets (Nevo, 2000, 2001; Backus, Conlon, and Sinkinson, 2021). We observe product-level sales data from Dutch supermarkets at a weekly frequency during 2011-2013. A preliminary descriptive analysis reveals two stylized facts. First, there is substantial price variation between different cereal products: the most expensive cereal is priced an order of magnitude higher than the cheapest one. Second, descriptive log-log regressions per product suggest that product-level elasticities are roughly independent of average product prices. These findings indicate the importance of allowing for sufficient flexibility in either unobserved consumer heterogeneity or demand curvature, or both.

Given this motivating evidence, we next assess the ability of our joint Box-Cox and BLP model to recover more plausible elasticities (and markups) compared to several popular but more restricted models. The estimates of the joint model show that there is significant heterogeneity in price sensitivity, and that price enters utility somewhere in between the linear form of BLP's mixed logit and the log-linear form of the mixed CES. These findings imply that restricting either the Box-Cox or the price heterogeneity parameter may entail biased estimates, and hence restrict substitution patterns.

We illustrate the implications of greater flexibility by plotting the own-price elasticities against prices for the various models. Using the descriptive estimates as guidance, we find that the joint model successfully recovers own-price elasticities that are roughly independent of price. The simple Box-Cox model, which abstracts from consumer heterogeneity, also recovers this pattern, but has the drawback of restricting cross-price elasticities. By contrast, the simple logit model has own-price elasticities that are linearly increasing in price (as is well-known), which is inconsistent with our descriptive estimates. Finally, the BLP model (with substantial heterogeneity) entails a U-shaped profile of own-price elasticities against price. This reflects the outcome of two opposing effects. First, because price enters utility linearly, the own-price elasticities scale linearly with price (as in the simple logit). Second, consumer heterogeneity means less price sensitive consumers are more likely to purchase higher-priced products and vice versa. At low prices, the first effect dominates, while at high prices the second effect is more important. Our application also suggests that the estimated cross-price elasticities between similarly priced products are lower in the joint Box-Cox and BLP model than in the standard BLP model.

We draw two implications for estimating differentiated products demand systems with aggregate sales data. First, to uncover adequate substitution patterns, it is not sufficient to focus on flexible random coefficients to account for consumer heterogeneity. It is also important to incorporate a flexible functional form for price. This conclusion is particularly relevant for applications that hinge on the demand curvature, such as the pass-through of a tax, tariff or exchange rate. A second conclusion is more pragmatic. The simple Box-Cox without random coefficients suggests that the CES model fits the data significantly better than the logit model. Practitioners who make use of logit or nested logit models because of data limitations or computational simplicity may therefore also consider the CES or nested CES as simple alternatives.

Related Literature: This paper contributes to the growing literature on estimating models of demand in differentiated products markets; for two recent surveys, see Berry and Haile (2021) and Gandhi and Nevo (2021). Berry and Haile (2014) obtain non-parametric identification results for differentiated products demand systems with market-level data. Their framework allows for flexible specifications for price. Compiani (2021) builds on their theoretical results to estimate a non-parametric analogue of the BLP model. While certainly allowing for additional flexibility, a non-parametric approach presents at least two practical challenges: the number of estimated parameters grows exponentially with the number of products, and estimation requires sufficiently rich price variation. Compiani therefore illustrates his framework to the market for fresh strawberries, which consists of only two products and exhibits large seasonal price movements. Our more targeted approach strikes a more pragmatic balance between flexibility and tractability. The Box-

Cox approach easily accommodates many products and can be estimated using standard levels of price variation. Moreover, it nests several popular but more restricted models, so can pragmatically guide applied demand analysis.

A number of papers focus on obtaining more flexibility by using either micro-moments (Berry, Levinsohn, and Pakes, 2004) or consumer-level data (Griffith, Nesheim, and O'Connell, 2018). Most relevant to our paper is Griffith et al. (2018), who use consumer-level data to demonstrate how estimating correlations between consumer income, purchase patterns, and demographics lead to economically meaningful differences in the pass-through of a tax. Our approach shows how additional flexibility on demand curvature can also be obtained in models with market-level data.

Finally, we contribute to the microfoundations of aggregate demand systems. Head and Mayer (2021) analyze the ability of the CES model to generate predictions in line with the BLP model. We show how our joint Box-Cox and BLP model can guide the choice of functional form in empirical applications, as it provides a unifying framework that is microfounded in a discrete choice theory. Our framework strikes a balance between incorporating heterogeneity with linear price, and alternative functional forms such as logarithmic price without heterogeneity. In an independent recent paper, Anderson and De Palma (2020) use a variant of our specification, but their focus is different. They do not provide an empirical framework, but instead analyze theoretical relationships between equilibrium distributions of productivity, output, etc.

Overview: The paper proceeds as follows. Section 2 introduces the joint Box-Cox and BLP demand model, develops the estimating equations and discusses the model's implied substitution patterns. Section 3 illustrates our framework with an empirical application. Subsection 3.1 describes the data and provides motivating stylized facts, while next subsections discuss the empirical results of the joint model, and compares them with more restrictive models. Last, section 4 concludes.

2 Demand Model and Elasticities

This section derives the joint Box-Cox and BLP model, which allows for both flexibility in the price functional form and for unobserved consumer heterogeneity. Subsection 2.1

formulates the theoretical framework and subsection 2.2 derives the estimating equations. Next, subsection 2.3 outlines implications for own- and cross-elasticities.

2.1 Utility and Demand

Consumers choose both their preferred product, and how many units to purchase of it. In this subsection, we first specify utility and conditional demand for the preferred product, then the choice probability of each product, and finally aggregate demand.

Utility: In each market, there exist L consumers, $i = 1, \dots, L$. Each consumer chooses one alternative from $J+1$ differentiated products, $j = 0, \dots, J$, where $j = 0$ is the outside good. Conditional on purchasing j, consumer i has the following indirect utility function:

$$u_{ij} = x_j\beta + \alpha_i f(y_i, p_j) + \xi_j + \varepsilon_{ij}, \tag{1}$$

where x_j is a vector of observed product characteristics; y_i and p_j denote consumer income and price; and $f(y_i, p_j)$ specifies how price and income enter indirect utility. For simplicity, the taste parameter vector for the product characteristics, β, is common across consumers. The price sensitivity parameter, α_i, is a normally distributed random coefficient with mean α and standard deviation σ, i.e. $\alpha_i = \alpha + \sigma v_i$ where v_i is a standard normal variable. Last, ξ_j captures unobserved product characteristics, which are common to all consumers, and ε_{ij} is a consumer-specific taste term for good j.

Box-Cox specification and conditional demand: We specify that price and income enter utility through a Box-Cox transformation (Box and Cox, 1964), i.e.:

$$f(y_i, p_j) = \gamma^{\lambda-1}\frac{y_i^\lambda - 1}{\lambda} - \frac{p_j^\lambda - 1}{\lambda}, \tag{2}$$

where $\lambda \leq 1$ represents the Box-Cox parameter and γ is the fraction of income a consumer allocates to the cereal category.[2] Conditional on selecting product j, the demand of consumer i follows from Roy's identity, $q_{ij}(y_i, p_j) = -\frac{\partial u_{ij}/\partial p_j}{\partial u_{ij}/\partial y_i}$. Using (1) and (2), this is given

[2]One can in principle have a separate Box-Cox parameter for price (λ_p) and income (λ_y), but it is less obvious how to identify λ_y from aggregate sales data. In their theoretical contribution, Anderson and de Palma (2020) essentially specified $\lambda_y = 1$.

by:

$$q_{ij}(y_i, p_j) = \left(\frac{\gamma y_i}{p_j}\right)^{1-\lambda}, \tag{3}$$

The Box-Cox parameter λ allows for flexibility, but also nests two existing specifications in the literature. First, with $\lambda = 1$, price and income enter utility linearly, i.e., $f(y_i, p_j) = y_i - p_j$, and a consumer purchases one unit of her preferred product ($q_{ij}(y_i, p_j) = 1$). This linear price specification with unit demand is often adopted in the traditional BLP model.[3] Second, as $\lambda \to 0$, price and income enter utility logarithmically (from l'Hôpital's rule), i.e., $f(y_i, p_j) = \gamma^{-1}\ln(y_i) - \ln(p_j)$). In this case, a consumer spends a constant fraction of her income to her preferred product ($q_{ij}(y_i, p_j) = \frac{\gamma y_i}{p_j}$). This is essentially a CES specification derived from a discrete choice model. Björnerstedt and Verboven (2016) refer to it as a constant expenditures specification, and it is increasingly adopted in applied work (e.g., see Fang, 2019; Eizenberg, Lach, and Oren-Yiftach, 2021; Hatan, Fleischer, and Tchetchik, 2021).

Choice Probability: We can write utility more compactly as:

$$u_{ij} = K_i + \delta_j + \mu_{ij} + \varepsilon_{ij}, \tag{4}$$

where $K_i = \alpha_i \gamma^{\lambda-1} \frac{y_i^\lambda - 1}{\lambda}$ is constant for each consumer over products; $\delta_j = x_j\beta - \alpha \frac{p_j^\lambda - 1}{\lambda} + \xi_j$ is the mean valuation for product j shared by all consumers; and $\mu_{ij} = \sigma v_i \frac{p_j^\lambda - 1}{\lambda}$ is a consumer-specific valuation for product j.

Each consumer i chooses the product j that maximizes her random utility U_{ij}. Assuming the random taste parameter, ε_{ij}, follows an extreme value distribution and normalizing $\delta_0 = 0$, the probability a consumer i chooses product j takes the form:

$$s_{ij}(\boldsymbol{\delta}, \sigma, \lambda) \equiv \frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_{k=1}^{J} \exp(\delta_k + \mu_{ik})}, \tag{5}$$

where the separable term K_i cancels out from the choice probabilities.

[3] BLP consider an alternative specification where price and income enter through the term $\alpha \ln(y_i - p_j)$. Since both variables enter additively, this also results in unit demand. As we will see below, their specification creates flexibility only through heterogeneity in price sensitivity.

Aggregate Demand: Assuming that v_i, y_i and ε_{ij} are independent, aggregate demand for product j is given by:

$$q_j = \int s_{ij}\left(\boldsymbol{\delta}, \sigma, \lambda\right) q_{ij}\left(y_i, p_j\right) dP_\nu(\nu) dP_y(y) L \tag{6a}$$

$$= \int s_{ij}\left(\boldsymbol{\delta}, \sigma, \lambda\right) dP_\nu(\nu) \int q_{ij}(y_i, p_j) dP_y(y) L \tag{6b}$$

$$= \int s_{ij}\left(\boldsymbol{\delta}, \sigma, \lambda\right) dP_\nu(\nu) \int \left(\frac{\gamma y_i}{p_j}\right)^{1-\lambda} dP_y(y) L \tag{6c}$$

where (p_y, p_v) are the income and price sensitivity distributions.

2.2 Estimating equations

Rearranging the aggregate demand from equation (6c) into an estimating equation requires two steps. A first step specifies the distribution of income. For simplicity, assume that all consumers within a region share the average region income, $\bar{y}_r$, so $\int y_i^{1-\lambda} dP_y(y) = \bar{y}_r^{1-\lambda}$. Appendix A.1.1 shows how one may incorporate income heterogeneity using two approaches: income draws from an empirical distribution, or a Taylor expansion. The second step approximates the integral over unobserved consumer heterogeneity ν. We follow the BLP methodology by taking n simulated draws of a standard normal distribution (see Berry et al., 1995). Combining these two steps and rearranging leads to the following estimating equation for the joint Box-Cox and BLP model:

$$\frac{p_j^{1-\lambda} q_j}{L(\gamma \bar{y}_r)^{1-\lambda}} = \frac{1}{n} \sum_{i=1}^{n} \frac{\exp\left(\delta_j + \mu_{ij}\right)}{1 + \sum_k \exp\left(\delta_k + \mu_{ik}\right)}. \tag{7}$$

The right-hand side has the usual interpretation as averaging over consumer choice probabilities (where the Box-Cox parameter implicitly enters through δ_j and μ_{ij}). The left-hand side of equation (7) may be interpreted as a market share variable. For instance, a linear price ($\lambda = 1$) implies unit demand, so the market share variable simplifies to a product's aggregate demand relative to the total number of consumers, $\frac{q_j}{L}$; a log price ($\lambda = 0$) implies constant expenditures demand, so the market share variable simplifies to a product's aggregate revenue relative to the total budget, $\frac{p_j q_j}{L\gamma y}$.

Following BLP's contraction mapping, the market share system (7), for $j = 1, \cdots, J$, can be inverted to solve for the mean utilities δ_j. Without unobserved heterogeneity, one can follow the analytical inversion approach from Berry (1994) (see Appendix A.1.2 for details):

$$\ln\left(\frac{p_j^{1-\lambda} q_j}{L(\gamma \overline{y}_r)^{1-\lambda} - \sum_J p_j^{1-\lambda} q_j}\right) = x_j \beta - \alpha \frac{p_j^{\lambda} - 1}{\lambda} + \xi_j. \tag{8}$$

2.3 Implications for own- and cross-elasticities

As shown in Appendix A.1.3, the own- and cross-price elasticities for the joint Box-Cox and BLP model can be written as:

$$\eta_{jk} = \begin{cases} -\dfrac{p_j^{\lambda}}{s_j} \displaystyle\int \alpha_i s_{ij} \left(1 - s_{ij}\right) dP_{\nu}(\nu) - (1 - \lambda) & \text{if } j = k \\ \dfrac{p_k^{\lambda}}{s_j} \displaystyle\int \alpha_i s_{ij} s_{ik} dP_{\nu}(\nu) & \text{if } j \neq k. \end{cases} \tag{9}$$

The first line is the own-elasticity, η_{jj}, which separates into a typical choice probability elasticity and a conditional demand elasticity, and the second line is the cross-elasticity, η_{jk}. Equation (9) clarifies the role of both the Box-Cox parameter and consumer heterogeneity.[4]

The Box-Cox parameter λ relaxes the typical unit demand assumption, and creates greater flexibility on demand curvature. Specifically, it reveals a relationship between elasticities and prices across products j, as seen from the terms p_j^{λ} and p_k^{λ} in front of the integral. With $\lambda = 1$, the own- and cross-price elasticities scale quasi-linearly with own- and cross-prices. We say quasi-linearly because price also enters the choice probability, s_{ij} in the integral term. A key insight is that with a more concave price functional form, $\lambda < 1$, this scaling becomes less than linear. For a log price specification, $\lambda = 0$, there is no scaling between elasticities and price, while $\lambda < 0$ would imply a decreasing relationship.

Consumer heterogeneity affects substitution patterns in two ways. First, it affects the link between price elasticities and price. Intuitively, price insensitive consumers tend to purchase higher priced products. This counterbalances the increasing relationship between price elasticities and price that arises for $\lambda \in (0, 1)$. Second, heterogeneity in price sensitivities affects the cross-price elasticities, as it implies stronger substitution between

[4] Table 2 in Appendix A.1.3 presents the elasticities in several special cases: $\lambda = 1$ or $\lambda = 0$, with and without heterogeneity in α_i

similarly price products (regardless of their price level).

3 Illustrative Application

To illustrate our demand framework, we apply it to the "Ready-to-Eat" cereal market, consistent with a prior literature using the cereal category to demonstrate the performance of different demand models (e.g., Nevo, 2000, 2001; Backus et al., 2021). Subsection 3.1 provides descriptive evidence to motivate our analysis. Subsection 3.2 discusses the specification and estimation of the joint Box-Cox and BLP demand model. Subsection 3.3 and 3.4 present the demand parameter estimates and the implications for price elasticities. A detailed discussion of the data and summary statistics is provided in Appendix A.2.

3.1 Descriptive Evidence

Our data set on the "Ready-to-Eat" cereal market comes from IRI. The unit of observation is a product j (barcode), market or region r (6 regions, i.e. provinces in the North of the Netherlands) and week t (156 weeks during 2011-2013). The total number of observations is 50,836, amounting to an average number of products of 54.31 per market and week (and 73 distinct products over the entire period). We have information on the following variables: quantity sold (kg), revenues and price (€), and size (kg).

For each product j, we estimate the following descriptive regression

$$\ln(q_{jrt}) = \eta_j \ln(p_{jrt}) + w_{jrt}\theta_j + u_{jrt}, \tag{10}$$

where $\ln(q_{jrt})$ and $\ln(p_{jrt})$ are the log quantity and log price of product j in market r for week t. The vector w_{jrt} includes market, year and month-of-year fixed effects. Our main interest is in the price coefficient η_j, which we interpret as a descriptive estimate of the own-price elasticity of product j (i.e., without directly modeling substitution between different products). The account for possible endogeneity issues, we include a standard set of Hausman and BLP instruments (as in our structural demand model, explained in more detail below).

Figure 1 presents these estimates by plotting the estimated own-elasticity against the average price per product to make two points. First, cereal prices vary widely, with the most expensive cereal price an order of magnitude higher than the cheapest one. Second, the own-price elasticities are roughly independent of price. The joint findings of wide price variation and constant elasticities motivate a joint Box-Cox and BLP demand model. Specifically, this model can evaluate to which extent a traditional BLP model with a linear price variable can generate this constant pattern, or whether a more flexible functional form through the Box-cox parameter λ is required.

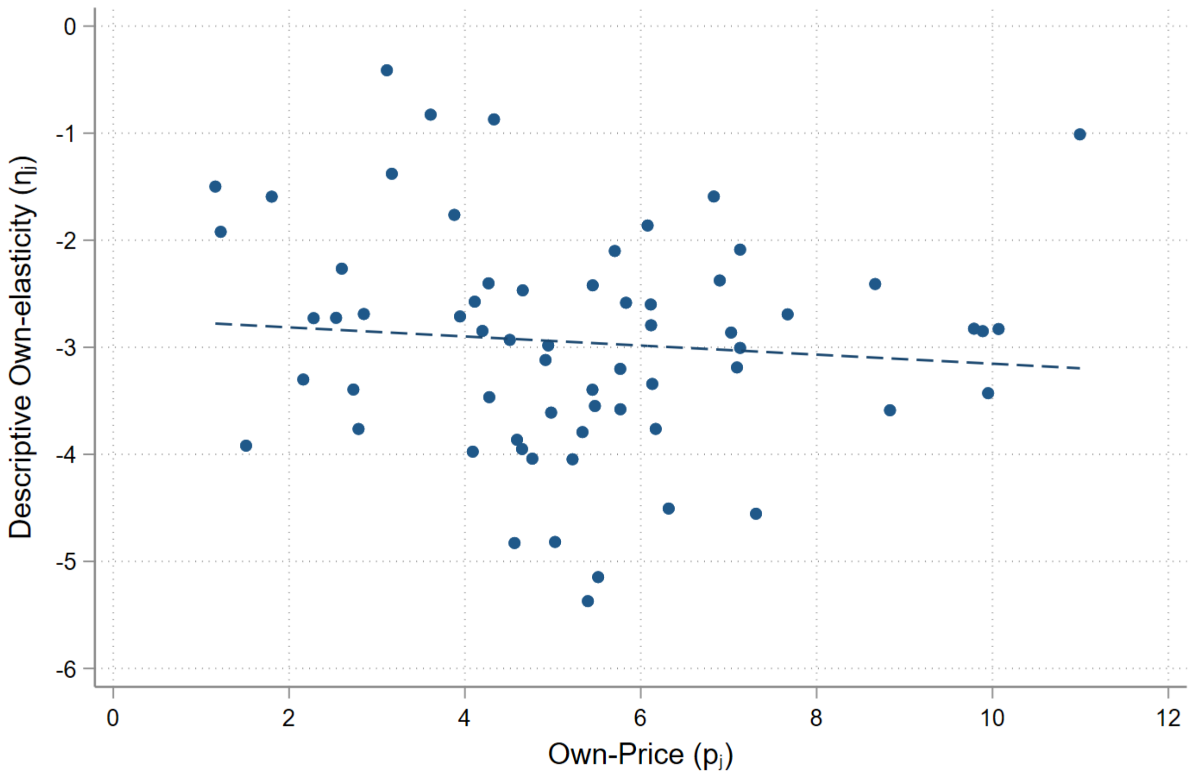

Figure 1: Descriptive own-price elasticity vs. own-price

Explanation: Scatter plot of descriptive estimate of own-price elasticity against average price of product. We estimate an own-elasticity separately per product using equation (10) for 73 products in the Cereal category. The figure excludes 8 observations because the estimate is not statistically significant, or the estimated elasticity is positive. The navy dashed line represents the estimated relationship (i.e., fitted values) between own-price elasticity and price.

3.2 Estimation

We first provide details on the demand specification, then discuss our set of instruments and identification of the Box-Cox parameter, and finally explain the estimating algorithm.

Specification: Our unit of observation is the product j in region r in week t, so we can add subscripts r and t to all variables in our (inverted) estimating equation (7), which includes the unobserved quality or error term ξ_{jrt}. We exploit the long weekly panel to estimate a fixed effect for each product j to account for time-invariant unobserved product characteristics affecting mean utility. We further include a fixed effect per year-region, and per month-of-year to capture unobserved demand shocks both across time and between markets. The richness of these fixed effects, and in particular the product fixed effects, enable us to focus attention on the joint role of the price functional form and consumer heterogeneity in determining substitution patterns.

Defining the market share variable requires us to determine the size of the potential market. In our setting this includes both the total number of potential consumers, L, and the consumers' total potential budget allocated to the cereal category, $\gamma \bar{y}_r$. To obtain both variables, we first calculate the total quantity and sales per market and week, we then take the maximum of each variable across regions and weeks, and conservatively multiply each variable by a factor of ten. Previous research typically finds the demand estimates are robust to these assumptions (e.g., see Nevo, 2000).

Identification and instruments: We start from the commonly used identification assumption that the non-price product characteristics (in this case the various fixed effects) are uncorrelated with the error term ξ_{jrt}. Under this assumption, the fixed effects are instruments for themselves. We require additional instruments to identify the price coefficient, consumer heterogeneity, and Box-Cox parameter $(\alpha, \sigma, \lambda)$. Following the literature, we firstly use functions of the other product characteristics as instruments, i.e., BLP instruments from Berry et al. (1995). The set consists of counts of own- and other-brand products for the following segmentations: entire cereal category, broad product description (e.g. cornflakes or children's cereal), detailed product description (e.g., standard or organic muesli), packaging type and package size. As explained in Berry and Haile (2014), BLP instruments help identify distributional parameters. Secondly, we use the average prices of the same product in other markets, i.e., Hausman instruments from Hausman (1996). We extend the traditional Hausman instruments by also including the log and square-root of these average prices. Intuitively, the different functional forms pin down the Box-Cox parameter by providing information on the shape of the demand curve (e.g., see: Amemiya and Powell, 1981; Powell, 1996).

Berry and Haile (2014) establish identification in a setting that is more general than ours. We fix attention on explaining how the Box-Cox parameter λ is separately identified from the price heterogeneity parameter σ. Different values of λ and σ imply differences in how the own- and cross-elasticities vary with prices. In particular, λ mainly determines how elasticities scale with prices. By contrast, σ affects this same link (but in a different way) and additionally determines the extent to which consumers more readily substitute to more similarly priced products. The first mechanism is identification from the shape of the demand curve. The second mechanism relies on the covariance between the cross-price elasticity and prices.

Estimation: Estimation of the joint model can proceed based on the well-documented BLP algorithm as outlined in Berry et al. (1995) (where additional instruments can be helpful to identify the shape parameter λ). To facilitate estimation, we estimate the model for fixed candidate values of λ to obtain the conditional utility parameters $((\sigma, \alpha, \beta)|\lambda)$. We perform a grid search to select the λ with the lowest GMM criterion value. We calculate Newey and West (1987) likelihood ratio or distance test statistics for the optimal λ against alternative values to construct a confidence interval for λ as in e.g. Moreira (2003).

3.3 Demand parameter estimates

Table 1 summarizes the parameter estimates relating to the price variable: the mean price coefficient α, the standard deviation σ and the Box-Cox parameter λ. Panel (A) reports estimates for the simple models without consumer heterogeneity ($\sigma = 0$). Panel (B) reports estimates for the random coefficient models.

In all specifications, α enters with the expected sign and significantly, indicating that consumers on average dislike paying higher prices. Taken together, α, σ and λ show quite some variation across specifications, because the free parameters partly take over the imposed restrictions on the fixed parameters. Nevertheless, the resulting average own-price elasticities are remarkably similar across all models.

In the simple model without consumer heterogeneity (panel (A)), we estimate a Box-Cox parameter $\lambda = -0.068$. Interestingly, this is not significantly different from zero and significantly less than 1. Hence, we cannot reject the CES model with log-linear price and market shares in value terms, while we can reject the logit model with linear price and

Table 1: Demand Parameter Estimates

	(A) Simple		(B) Random Coefficient	
	Logit	Box-Cox	Logit	Box-Cox
Price ($-\alpha$)	-0.46	-1.52	-0.91	-1.34
	(0.004)	(0.058)	(0.005)	(0.013)
Price Heterogeneity (σ)	0.00	0.00	0.30	0.43
	–	–	(0.008)	(0.020)
Box-Cox (λ)	1.00	-0.068	1.00	0.49
	–	(0.043)	–	(0.105)
Own-elasticity (η_{jj})	-2.50	-2.43	-2.62	-2.54

Notes: Simple refers to a model imposing zero consumer heterogeneity ($\sigma = 0$), while Random Coefficient refers to a model estimating consumer heterogeneity. Logit refers to the traditional linear price and unit demand model (i.e., $\lambda = 1$, while Box-Cox estimates the Box-Cox parameter as derived in equation (2). Robust standard errors reported in parentheses (and "–" denotes an imposed values e.g., a linear price or zero consumer heterogeneity). The standard error for the joint Box-Cox and BLP model is estimated using an inverted distance test statistic. The parameters are estimated using a sample of 50,836 observations for 2011–2013, where an observation represents a product-province-week. The demand specification includes a fixed effect for each product, year-market combination, and month.

market shares in volume terms.

Now consider the random coefficients models of panel (B). In the standard random coefficients logit of BLP (with $\lambda = 1$), we estimate significant heterogeneity in price sensitivity: the standard deviation $\sigma = 0.30$, compared with a mean price sensitivity parameter $\alpha = 0.91$. In the joint Box-Cox and BLP model, we appear to estimate even larger heterogeneity, $\sigma = 0.43$, but the mean price sensitivity parameter also increases to $\alpha = 1.34$. The Box-Cox parameter λ equals 0.49, which is halfway between a linear and log form. This contrasts with our earlier estimate of $\lambda = -0.068$ in the simple model without consumer heterogeneity. In that model, λ entirely captured the earlier documented independence between price elasticities and prices (Figure 1), while in the random coefficients model both σ and λ take this role.

3.4 Implications for price elasticities

Figure 2 shows how the own-price price elasticities vary across the price distribution. This allows us to evaluate the various demand models against the descriptive evidence on price elasticities in subsection 3.1 (Figure 1).

The simple logit with linear price (denoted using the yellow line on Figure 2) serves as a reference model to explain how more flexible models may generate more plausible elasticity patterns. Average product prices vary from €1.06 to €14.01 This implies, through the logit structure, that own-price elasticities also vary by an order of magnitude from -0.49 to -6.50. Beyond the implausibly large variation of own-price elasticities, it is particularly striking that the highest priced products are also the most price elastic.

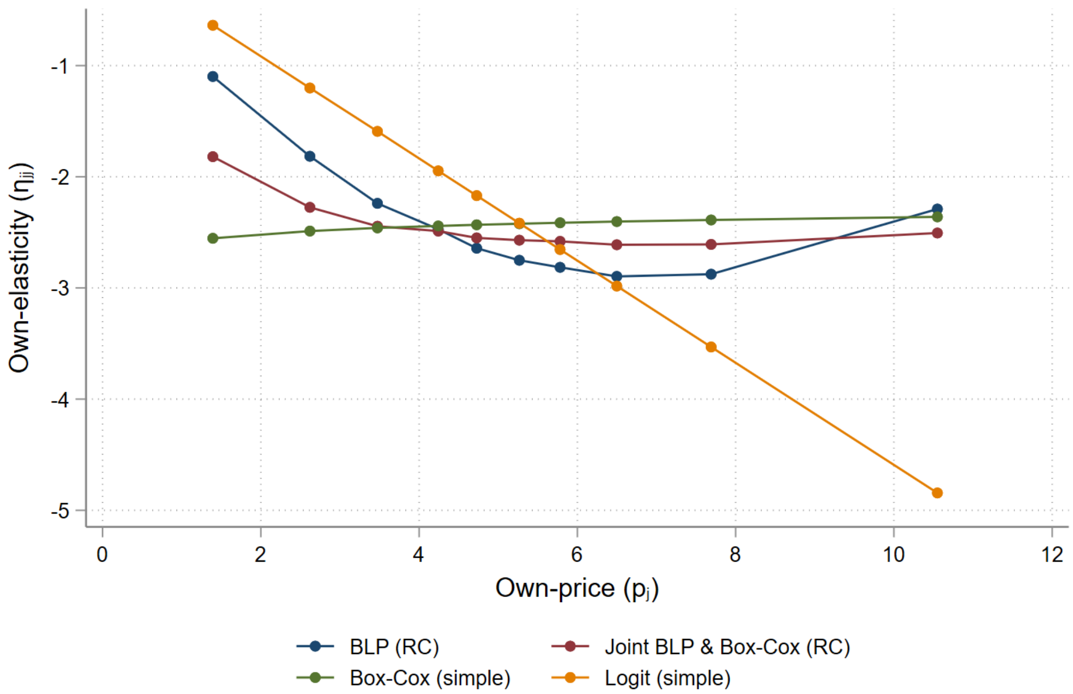

Figure 2: Own-elasticity vs. own-price

Explanation: Figure is a binned scatter to smooth across randomness introduced in the BLP simulation algorithm. We define ten equal sized bins and plot the average own-elasticity against the average price for each bin. We calculate the own-elasticity for each model using equations listed in Table 2 of Appendix A.1.3. Own-price is the average price per product. The sample consists of 73 products in the cereal category.

The blue line denotes the own-price elasticities from the BLP model. The BLP model firstly shows a significantly flatter profile compared to the simple logit, and secondly reveals an interesting U-shaped profile. The U-shaped profile reflects the outcome of two opposing channels. First, the own-price elasticities tend to scale linearly with price because price enters utility linearly (i.e., the same mechanism as in the simple logit). Second, because of heterogeneity in the price coefficient, less price sensitive consumers are more likely to purchase higher-priced products and vice versa. From the lowest price of roughly €1 up to roughly €6, the first channel of linear price dominates. As a result, own-price elasticities increase from roughly -1 for the lowest priced product to roughly -3 for the products priced around €6. At a price of around €6, the second channel of consumer heterogeneity starts to dominate. Accordingly, demand becomes less elastic and own-price elasticities become roughly -2.2 for the most expensive products. While the range of these elasticities is certainly more plausible than the simple logit, the largest elasticities still implausibly exceed the smallest elasticities by a factor of three. Further, the specific U-shape profile of elasticities may not be realistic.

We next present the results for the simple Box-Cox and the joint Box-Cox and BLP model. The green line reports the simple Box-Cox demand elasticities, which are roughly independent of own-price. This constant elasticity pattern results from the Box-Cox parameter λ being close to zero, which breaks the mechanical linear scaling between elasticities and prices. The red line represents the joint Box-Cox and BLP model, which generates a flatter profile of elasticities when compared to the BLP model. More specifically, own-price elasticities are roughly -2 for the cheapest products, and settle to roughly -2.5 from around €4. Compared to the BLP model, the flatter U-profile arises because the estimated Box-Cox parameter is less than one.

These findings are confirmed in the pattern of markups (reported in Figure A.1 in the appendix). Both the logit and BLP model entail considerably higher markups for the cheapest products; the BLP model has a similar U-shaped profile, while the Box-Cox and joint Box-Cox and BLP model imply more stable markups over the price range.

We finally ask what the different models imply for the estimated pattern of cross-price elasticities. We relate the cross-price elasticities between pairs of products to their absolute price differences. As expected, in the simple logit and Box-Cox models without heterogeneity, the cross-price elasticities between two products are approximately independent of their price differences. In contrast, in the BLP model there is a strong decreasing relationship: the cross-price elasticities are about three times higher for products with similar prices than for products with the the largest price differences. This reflects the importance of consumer heterogeneity in their price valuation. In the joint Box-Cox and BLP model there is also a decreasing relationship, though it is less pronounced, as we illustrate in Figure A.2 of the Appendix. This provides some suggestive evidence that the standard

BLP model may overestimate localized substitution patterns between similar products. This conclusion may not be generally true, but it suggests the importance of incorporating demand curvature to get the substitution patterns correct.

Summary: Taken collectively, these empirical patterns lead to four conclusions. First, the linear scaling between elasticities and prices is especially restrictive in the presence of large price variation. This point is worth emphasizing, as it may be a source of misspecification even if cross-price elasticities would be relatively symmetric. Second, consumer heterogeneity in the BLP model breaks this link in a very specific way, i.e. through a U-shaped profile of both elasticities and markups. While not impossible, this pattern follows directly from the assumed price functional form. Third, the joint Box-Cox and BLP model generates a more plausible pattern for elasticities (and markups). Specifically, the additional flexibility in the price functional form breaks the linear scaling between elasticities and prices. By requiring a smaller role for consumer heterogeneity, the U-shaped profile still exists but is significantly less pronounced. Fourth, the simple Box-Cox model recovers roughly constant elasticities, which closely resembles the results from the joint model. While the simple Box-Cox model may represent a useful approximation for many applications, we caution it does not recover rich patterns of cross-elasticities.

4 Concluding Remarks

We extend the frontier approach to estimating demand in differentiated product markets — the BLP approach — to relax functional form restrictions on price through a simple yet flexible Box-Cox transformation. This extension breaks built-in links between elasticities and prices.

We provide an illustrative application of our joint Box-Cox and BLP model to the market for ready-to-eat cereals to draw two broad conclusions. First, our joint model creates more flexibility to break the link between elasticities and prices across products. The BLP model relies exclusively on consumer heterogeneity to break this link. This creates a U-shaped profile between elasticities and prices. We also make a second, more pragmatic contribution. Applied researchers often abstract from incorporating unobserved consumer heterogeneity, or incorporate it in a simple way through a nested logit demand structure.

These models can be easily modified to include a Box-Cox parameter, or alternatively a sensitivity analysis to (nested) CES models (with $\lambda = 0$) should be considered more widely. This also provides guidance to practitioners in other fields such as trade, macro and labor.

We see several avenues for future research. First, it would be interesting to confirm whether our findings generalize to a broad set of product categories and industries beyond our illustrative application.

Second, our more flexible functional form relies on extending the typical assumption that consumers have unit demand for the preferred products to allow for elastic conditional demand. While such extension is realistic in many consumer goods markets, it may seem less intuitive in durable goods industries such as automobiles (as in BLP's original application), where consumers purchase a single product on a purchase occasion. Nevertheless, similar flexibility may arise by modeling elastic conditional demand over the durable goods' life-cycle, and exploring this would be interesting.

Finally, our model provides increased flexibility to account for demand curvature. In the presence of market power, curvature plays a key role in the extent of pass-through (e.g. Bulow and Pfleiderer, 1983). Applied research that relies on demand estimation to study the pass-through of taxes, tariffs and exchange rates would thus especially benefit from this increased flexibility. Nevertheless, we caution that our model mainly captures curvature through the relationship between elasticities and prices in the cross-section of products. Further extensions to model yet greater flexibility would also be very interesting in future research.

References

ADAO, R., A. COSTINOT, AND D. DONALDSON (2017): "Nonparametric counterfactual predictions in neoclassical models of international trade," *American Economic Review*, 107, 633–89.

AMEMIYA, T. AND J. L. POWELL (1981): "A comparison of the Box-Cox maximum likelihood estimator and the non-linear two-stage least squares estimator," *Journal of Econometrics*, 17, 351–381.

ANDERSON, S. P. AND A. DE PALMA (2020): "Decoupling the CES distribution circle with quality and beyond: equilibrium distributions and the CES-Logit nexus," *The Economic Journal*, 130, 911–936.

BACKUS, M., C. CONLON, AND M. SINKINSON (2021): "Common Ownership and Competition in the Ready-To-Eat Cereal Industry," .

BERRY AND HAILE (2014): "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 82, 1749–1797.

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

BERRY, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 25, 242–262.

BERRY, S. T. AND P. A. HAILE (2021): "Foundations of Demand Estimation," *NBER*.

BERRY, S. T., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

BJÖRNERSTEDT, J. AND F. VERBOVEN (2016): "Does merger simulation work? Evidence from the swedish analgesics market," *American Economic Journal: Applied Economics*, 8, 125–164.

BOX, G. AND D. COX (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society: Series B (Methodological)*, 26, 211–243.

Compiani, G. (2021): "Market Counterfactuals and the Specification of Multi-Product Demand : A Nonparametric Approach," .

Dubé, J.-P., A. Hortaçsu, and J. Joo (2021): "Random-coefficients logit demand estimation with zero-valued market shares," *Marketing Science*.

Einav, L., E. Leibtag, and A. Nevo (2010): "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?" *Quantitative Marketing and Economics*, 8, 207–239.

Eizenberg, A., S. Lach, and M. Oren-Yiftach (2021): "Retail Prices in a City," *American Economic Journal: Economic Policy*, 13, 175–206.

Fang, L. (2019): "The Effects of Online Review Platforms on Restaurant Revenue, Survival Rate, Consumer Learning and Welfare," .

Gandhi, A. and A. Nevo (2021): "Empirical Models of Demand and Supply in Differentiated Products Industries," *NBER*.

Griffith, R., L. Nesheim, and M. O'Connell (2018): "Income effects and the welfare consequences of tax in differentiated product oligopoly," *Quantitative Economics*, 9, 305–341.

Hatan, S., A. Fleischer, and A. Tchetchik (2021): "Economic valuation of cultural ecosystem services: The case of landscape aesthetics in the agritourism market," *Ecological Economics*, 184.

Hausman, J. (1996): "Valuation of new goods under perfect and imperfect competitioin. I," *The Economics of New Goods*.

Head, K. and T. Mayer (2021): "Poor Substitutes? Counterfactual methods in IO and Trade compared," .

Moreira, M. J. (2003): "A conditional likelihood ratio test for structural models," *Econometrica*, 71, 1027–1048.

Nair, H., J.-P. Dubé, and P. Chintagunta (2005): "Accounting for primary and secondary demand effects with aggregate data," *Marketing Science*, 24, 444–460.

Nevo, A. (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513–548.

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

Newey, W. K. and K. D. West (1987): "Hypothesis testing with efficient method of moments estimation," *International Economic Review*, 777–787.

Powell, J. L. (1996): "Rescaled methods-of-moments estimation for the Box-Cox regression model," *Economics Letters*, 51, 259–265.

A Appendix

This Appendix provide additional details relating to the model and data.

A.1 Model

Section A.1.1 discusses the extension of the model to incorporate income heterogeneity. Section A.1.2 derives the analytical Berry inversion in the case without consumer heterogeneity. Section A.1.3 derives the model's own- and cross-price elasticities.

A.1.1 Income heterogeneity extension

Two methods can incorporate heterogeneous income per market into $\int y_i^{1-\lambda} dP_y(y)$. A first method uses income tables or random draws, while a second method uses a Taylor Expansion.

1) Income table or random draws: An income table lists the fraction of consumers,Φ_g with income y_g per group g, such that $\sum_{g=1}^{G} \Phi_g = 1$. Substitute this definition into the joint Box-Cox and BLP estimating equation:

$$\frac{p_j^{1-\lambda} q_j}{L\gamma^{1-\lambda} \sum_{g=1}^{G} \Phi_g y_{rg}^{1-\lambda}} = \int s_{ij}\left(\boldsymbol{\delta}, \sigma\right) dP_\nu(\nu) \tag{A.1}$$

Alternatively one may take simulated income draws e.g., by assuming the data is normally distributed and one knows the mean and standard deviation.

2) Taylor Expansion: We may approximate the income integral, $\int y_i^{1-\lambda} dP_y(y)$ using a Taylor expansion. Write y_i as the mean income plus a deviation from the mean, so $\int y_i^{1-\lambda} dP_y(y) = \int \left(\overline{y_i} + (y_i - \overline{y_i})\right)^{1-\lambda} dP_y(y)$. Taking the second-order Taylor expansion of the bracketed term gives:

$$\left(\overline{y} + (y_i - \overline{y_i})\right)^{1-\lambda} = (\bar{y})^{1-\lambda} + (1-\lambda)(\bar{y})^{-\lambda}\left(y_i - \bar{y}\right) - \frac{\lambda(1-\lambda)(\bar{y})^{-\lambda-1}}{2}\left[(y_i - \bar{y})^2\right] \tag{A.2}$$

Noting that $\int (y_i - \bar{y})\, dP_y(y) = 0$ and $\int (y_i - \bar{y})^2\, dP_y(y) = \sigma_y^2$, we rewrite equation (A.2):

$$\int y_i^{1-\lambda} dP_y(y) = (\bar{y})^{1-\lambda} - \frac{\lambda(1-\lambda)\sigma_y^2}{\bar{y}^{1+\lambda}}. \tag{A.3}$$

One may feasibly substitute this into the estimating equation. The Taylor expansion also allows us to sign the bias from ignoring income heterogeneity. For example, unit- or constant expenditures-demand $\lambda = 0, 1$ implies no bias. Otherwise, for intermediate λ values, bias depends on the combination of $\lambda, \overline{y}, \sigma_y^2$.

A.1.2 Berry Inversion

Abstracting from consumer heterogeneity implies that the estimating equation (7) becomes:

$$\frac{p_j^{1-\lambda} q_j}{L\left(\gamma \bar{y}_r\right)^{1-\lambda}} = \frac{\exp\left(\delta_j\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_k\right)}. \tag{A.4}$$

Because the market share of the outside good 0 equals the total budget minus the budget allocated to all inside goods, we may write the following choice probability for good 0:

$$\frac{L(\gamma y)^{1-\lambda} - \sum_J p_j^{1-\lambda} q_j}{L(\gamma y)^{1-\lambda}} = \frac{1}{1 + \sum_{k=1}^{J} \exp\left(\delta_k\right)}. \tag{A.5}$$

Dividing the choice probability for each product j by the choice probability of the outside good 0 leads to the following ratio of choice probabilities:

$$\frac{p_j^{1-\lambda} q_j}{L(\gamma y)^{1-\lambda} - \sum_{k=1}^{J} p_j^{1-\lambda} q_j} = \exp\left(\delta_j\right). \tag{A.6}$$

Taking logs arrives at equation (8).

A.1.3 Elasticities

We derive the own- and cross-elasticities for the joint Box-Cox and BLP model and then list the elasticities for special cases.

Own-elasticity:

$$
\begin{aligned}
\eta_{jj} &= \frac{\partial q_j}{\partial p_j}\frac{p_j}{q_j} && \text{(A.7)}\\
&= \left(\int \frac{\partial s_{ij}}{\partial p_j} dP_\nu(\nu) \left(\frac{\gamma \bar{y}}{p_j}\right)^{1-\lambda} L + \int s_{ij} dP_\nu(\nu) \frac{\partial}{\partial p_j}\left(\frac{\gamma \bar{y}}{p_j}\right)^{1-\lambda} L\right)\frac{p_j}{q_j} && \text{(A.8)}\\
&= \left[-\int \alpha_i p_j^\lambda s_{ij}\left(1-s_{ij}\right) dP_\nu(\nu) - \int s_{ij} dP_\nu(\nu)(1-\lambda)\right]\frac{1}{s_j} && \text{(A.9)}\\
&= -\frac{p_j^\lambda}{s_j}\int \alpha_i s_{ij}\left(1-s_{ij}\right) dP_\nu(\nu) - (1-\lambda). && \text{(A.10)}
\end{aligned}
$$

The second line follows from the product rule, since price enters both through the choice probability and conditional demand. The third line firstly inserts the share derivative of $\frac{\partial s_{ij}}{\partial p_j} = -\alpha p_j^{\lambda-1} s_{ij}\left(1-s_{ij}\right)$, secondly inserts the conditional demand derivative of $\frac{\partial}{\partial p_j} p_j^{\lambda-1} = (\lambda-1)p_j^{\lambda-2}$, and thirdly cancels one conditional demand term. Last, the fourth line recognizes that $\int s_{ij} dP_\nu(\nu) = s_j$.

Cross-elasticity:

$$
\begin{aligned}
\eta_{jk} &= \frac{\partial q_j}{\partial p_k}\frac{p_k}{q_j} && \text{(A.11)}\\
&= \left(\int \frac{\partial s_{ij}}{\partial p_k} dP_\nu(\nu)\left(\frac{\gamma \bar{y}}{p_j}\right)^{1-\lambda} L\right)\frac{p_k}{q_j} && \text{(A.12)}\\
&= \frac{p_k^\lambda}{s_j}\int \alpha_i s_{ij} s_{ik} dP_\nu(\nu), && \text{(A.13)}
\end{aligned}
$$

where line three follows by inserting $\frac{\partial s_{ij}}{\partial p_k} = s_{ij} s_{ik} \alpha_i p_k^{\lambda-1}$ and rearranging.

Table of Elasticities: Table 2 list the own- and cross-elasticities for each model.

Table 2: Own- and Cross-elasticities for each model

RC	Model	Own-elasticity, η_{jj}	Cross-elasticity, η_{jk}
Yes	Box-Cox	$\frac{p_j^\lambda}{s_j}\int \alpha_i s_{ij}\left(1-s_{ij}\right)dP_\nu(\nu)+(\lambda-1)$	$\frac{p_k^\lambda}{s_j}\int \alpha_i s_{ij}s_{ik}dP_\nu(\nu)$
	Unit	$\frac{p_j}{s_j}\int \alpha_i s_{ij}\left(1-s_{ij}\right)dP_\nu(\nu)$	$\frac{p_k}{s_j}\int \alpha_i s_{ij}s_{ik}dP_\nu(\nu)$
	Const Exp	$\frac{1}{s_j}\int \alpha_i s_{ij}\left(1-s_{ij}\right)dP_\nu(\nu)$ - 1	$\frac{1}{s_j}\int \alpha_i s_{ij}s_{ik}dP_\nu(\nu)$
No	Box-Cox	$p_j^\lambda\alpha\left(1-s_j\right)+(\lambda-1)$	$p_k^\lambda\alpha s_k$
	Unit	$p_j\alpha\left(1-s_j\right)$	$p_k\alpha s_k$
	Const Exp	$\alpha\left(1-s_j\right)-1$	αs_k

Notes: RC refers to Random Coefficient for price. Unit and Const exp refers to unit-demand and constant expenditures. For instance, the standard BLP is RC = yes and Model = Unit.

A.2 Data

The data on the "Ready-to-Eat" cereal market come from IRI, who provide scanning technology for supermarkets. The IRI data records weekly sales revenue and quantities per barcode for over 1,262 supermarkets in the Northern Netherlands from January 2011 to December 2013. A barcode defines a product, which is a distinct combination of a brand, flavor, packaging, and size. This product definition implies a 375-gram box of Kellogg's Special K is a different product than a 550-gram box or a 375-gram box of the dark chocolate flavor. The data is largely comparable to the widely used Nielsen data covering US retailers, as summarized in Einav, Leibtag, and Nevo (2010).

We refer to Statistics Netherlands to define a geographic market as a province. We have information on the following six provinces: Noord-Holland, Friesland, Groningen, Drenthe, Utrecht, and Flevoland.

We aggregate across all supermarkets within a province, so an observation of product j in geographic market m in week t is the total revenue, r_{jmt}, and total quantity, q_{jmt}. Using these two variables, we calculate the price by dividing the total revenue by total quantity, so $p_{jmt} = \frac{r_{jmt}}{q_{jmt}}$. This step only aggregates over the 903 supermarkets for which we have complete data (i.e., open for the full three-year sample, meaning we drop supermarkets, which open or close partway through the sample period). As is common in the discrete choice literature, we normalize prices to a common size, in our case to a price per kg.

For tractability, we trim the data to keep only economically meaningful products. We select these products by first ranking them by total sales revenue, and second dropping the long tail of products, which make up the bottom 30% of revenue. After dropping these products, the final data set covers 73 products, 6 geographic markets, and 156 weeks. This aggregates to 50,836 total product-market-week observations. Table 3 provides context by summarizing the main variables. The mean revenue is €1,641. The standard deviation is large, as some products have considerably larger sales than other products. The median price per serving equals €5,39/kg, but product prices vary widely. For further context, the average size equals 0.50 kg, so half a kilogram. Last, the average market-week contains observations for 54 products.

Table 3: Summary Statistics for Cereal 2011-2013

	Mean	Median	P25.	P75.	St Dev	Min	Max
Revenue (€)	1,641	700	333	1,833	2,495	18.93	34,325
Quantity (kg)	424	151	63	418	804	1.72	24,155
Price (€/kg)	5.39	5.12	3.96	6.57	2.50	0.97	15.12
Size (kg)	0.52	0.50	0.40	0.50	0.23	0.10	1.00
Products	54.31	54.00	52.00	57.00	2.84	48.00	59.00

Notes: Observation is product-market-week. Data covers 73 products, 6 geographic markets, and 156 weeks, which aggregates to over 50,836 total product-market-week observations. Products is number of products per category-market-week.

A.3 Additional Figures

This section reports the following two figures: First, Figure A.1 shows how implied product markups vary with own price. Second, Figure A.2 plots how cross-elasticities vary with price differences.

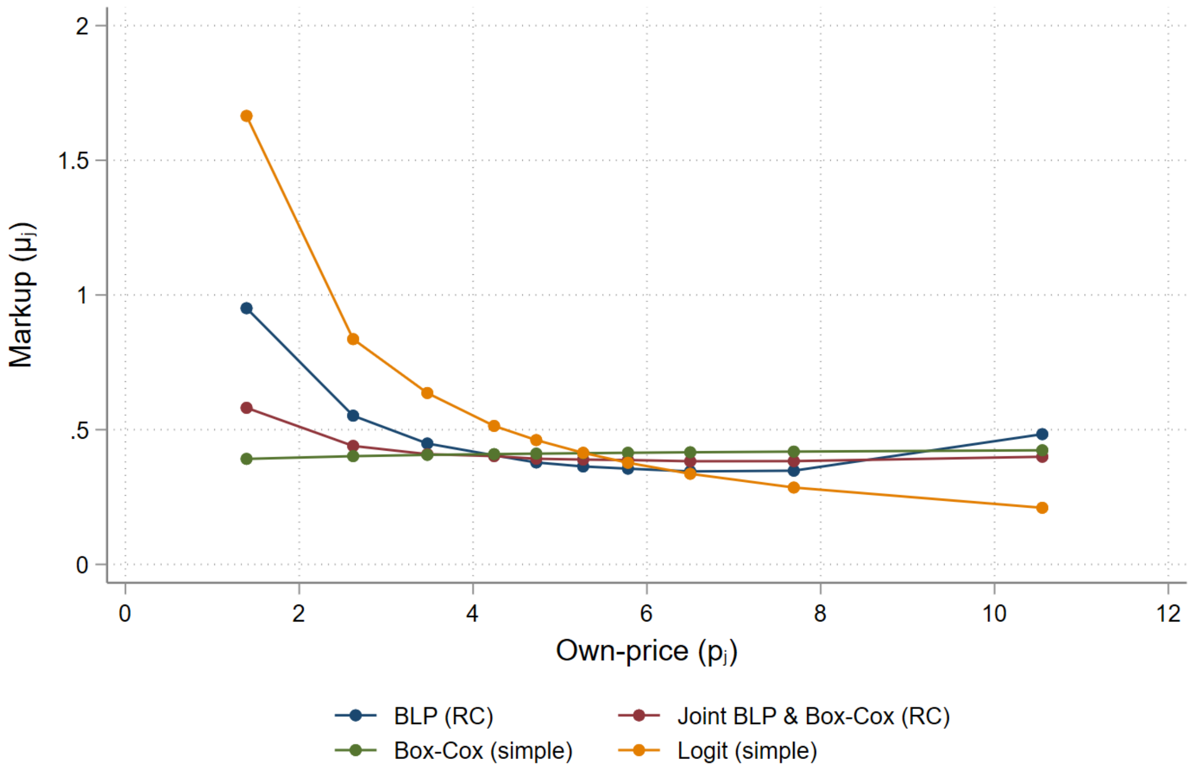

Figure A.1: Markups vs. own-price

Explanation: Figure is a binned scatter to smooth across randomness introduced in the BLP simulation algorithm. We define ten equal sized bins and plot the average markup against the average price for each bin. We calculate the product markups as inverse price elasticities using the elasticity for each model using equations listed in Table 2. Own-price is the average price per product. The sample consists of 73 products in the cereal category.

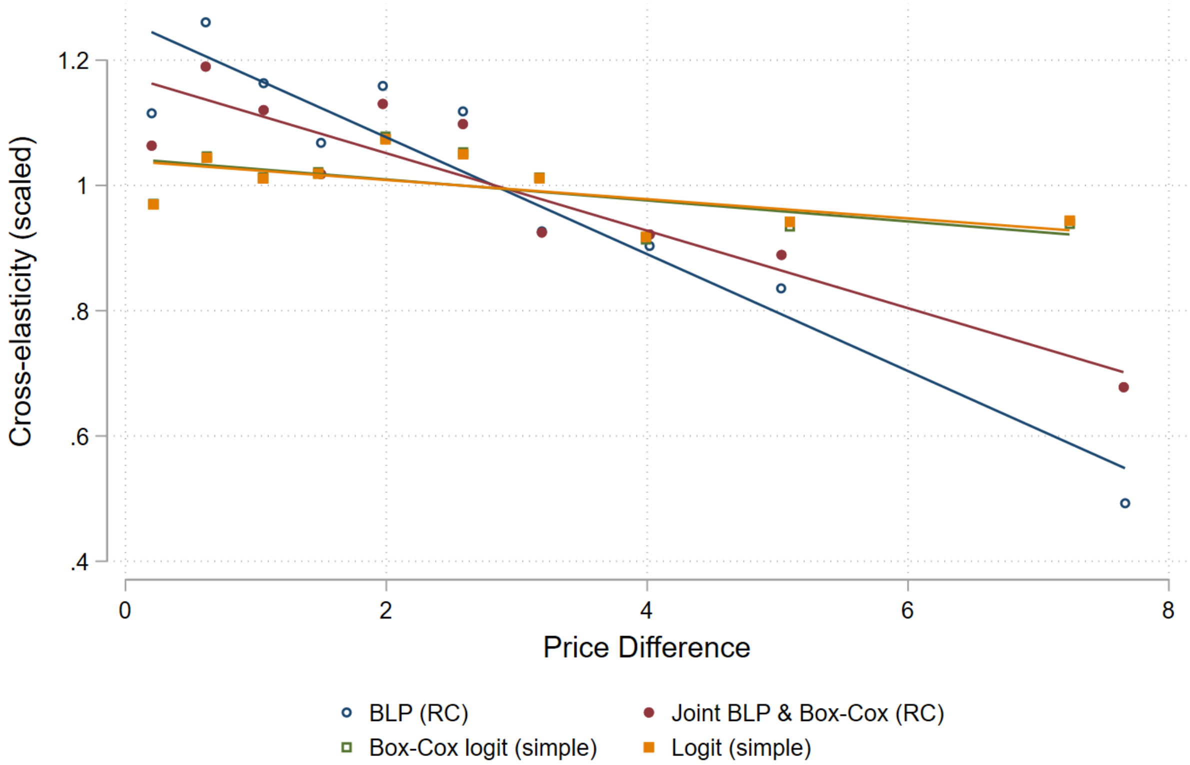

Figure A.2: Cross-elasticity vs. price-difference

Explanation: Figure is a binned scatter to smooth across randomness introduced in the BLP simulation algorithm and variation caused by differences in product shares driven by popularity (i.e., cross-elasticity is greater to a popular product independently of the price difference). We define ten equal sized bins and plot the average cross-elasticity against average price-difference for each bin. We calculate the cross-elasticity between each product j and k for each model using equations listed in Table 2. For interpretability, we scale these cross-elasticities by dividing by the mean cross-elasticity for each model. The X-axis is the absolute price difference, which we calculate as $|p_j - p_k|$. The sample consists of 73 products in the cereal category.