Estimating substitution patterns and demand curvature in Discrete-Choice models of product differentiation

Cameron Birchall (KU Leuven, cameron.birchall@kuleuven.be)
Debashrita Mohapatra (University of Connecticut, debashrita.mohapatra@uconn.edu)
Frank Verboven (KU Leuven, frank.verboven@kuleuven.be)

Selected Paper prepared for presentation at the 2024 Agricultural & Applied Economics Association Annual Meeting, New Orleans, LA; July 28-30, 2024

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

Cameron Birchall† Debashrita Mohapatra‡ Frank Verboven§

March 1, 2024

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, while remaining computationally tractable. We provide an illustrative application to the ready-to-eat cereals market. This shows that the cross-sectional relation between price elasticities and average prices per product is more in line with descriptive elasticity patterns, and that substitution between product pairs may be affected to some extent.

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

*Acknowledgements: We thank Jan De Loecker, Mario Samano, Jo Van Biesebroeck and participants at an IO-Leuven workshop for useful comments. We are also grateful to two referees and the Editor for very helpful comments. FWO (grant G0C8821N) and Methusalem provided financial support, NERA provided support to Cameron Birchall during part of this project. 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

‡University of Connecticut, debashrita.mohapatra@uconn.edu

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

1 Introduction

Substitution patterns between differentiated products are crucial to understanding 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

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

Box-Cox transformation of price and income (from Box and Cox, 1964). This joint Box-Cox and BLP 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 be 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 relating to the functional form for price. To identify this parameter, we suggest using transformations of the available price instruments.

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 demand models. Using the descriptive estimates as guidance, we find that the joint model successfully recovers own-price elasticities that are roughly independent of prices across products. 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, if there is sufficient heterogeneity, the BLP model may entail 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 may become more important (at least under sufficient consumer heterogeneity).

Finally, we find that our extended model may alter the estimated substitution between product pairs (as measured by the diversion ratio) to some extent. But it does not appear to affect relative patterns across products. As such, our model mainly affects the pattern of own-price elasticities (and markups) across products, and the implied cost pass-through.

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 in our application the CES model is not rejected by the data, in contrast with 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 part of a robustness analysis (micro-founded in the same discrete-choice setting).

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). Their conceptual framework shows how consumer heterogeneity in price sensitivity allows for more flexible demand curvature, including the possibility of log-convex demand. They use consumer-level data and focus on flexibly modelling income heterogeneity to obtain economically meaningful differences in the pass-through of a tax. In a recent paper Miravete, Seim, and Thurk (2023) also stress the importance of consumer heterogeneity, considering in addition the role of distributional assumptions of unobservable heterogeneity in price sensitivity. Our approach shows how additional flexibility on demand curvature can also be obtained in the absence of consumer heterogeneity in price sensitivity, by relaxing the usual assumption of perfectly inelastic conditional demand.

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.

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 (i.e. region and week), there exist L consumers, $i = 1, \ldots, L$. Each consumer chooses one alternative from $J+1$ differentiated products, $j = 0, \ldots, J$, 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 (in real terms); 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. $\alpha_i = \alpha + \sigma \nu_i$, where ν_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 power function or 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. In statistics, the Box-Cox transformation is used to transform non-normal random variables with a skewed distribution into a normal shape, and covers the normal and log-normal distribution as special cases. We also use the transformation to cover the linear and logarithmic form of price and income, but the economic intuition instead relates to the implied consumer's demand, as we discuss below.[2] Note that the budget share parameter γ is a potential market size variable, similar to the number of consumers L. In our application below, we will not estimate them, but make assumptions on both γ and L. 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 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 $\lambda \longrightarrow 0$, 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 ($q_{ij}(y_i, p_j) = \frac{\gamma y_i}{p_j}$).

[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 $\lambda_y = 1$.

[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.

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).

One may also specify other functional forms for $f(y_i, p_j)$ than the Box-Cox specification (2), such as polynomials in y_i and p_j (with suitable parameter restrictions). Other empirical work has restricted attention to functional forms of the type $f(y_i - p_j)$, implying unit demand (Griffith et al., 2018; Miravete et al., 2023).[4]

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:

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

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

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

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

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

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

[4] Miravete et al. (2023) also allow for alternative distributions of α_i.

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 have the same income (within a region), $\bar{y}$, so $\int y_i^{1-\lambda} dP_y(y) = \bar{y}^{1-\lambda}$. 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:

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

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 an average choice probability or market share variable, which is less than one if L and $\gamma\bar{y}$ are sufficiently large. 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, $\frac{p_j q_j}{\gamma \bar{y} L}$.

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.1.3 for details):

$$\ln\left(\frac{p_j^{1-\lambda} q_j}{L(\gamma\bar{y})^{1-\lambda} - \sum_{k=1}^{J} p_k^{1-\lambda} q_k}\right) = x_j\beta - \alpha \frac{p_j^{\lambda} - 1}{\lambda} + \xi_j. \tag{8}$$

2.3 Implications for price elasticities and curvature

As shown in Appendix A.1.4, the own- and cross-price elasticities for the joint Box-Cox and BLP model can be written as:

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

where $s_j \equiv \int s_{ij} dP_\nu(\nu)$. The first line is the own-elasticity, η_{jj}, which separates into the typical average choice probability elasticity and a conditional demand elasticity. The second line is the cross-elasticity, η_{jk}. Equation (9) clarifies the role of both the Box-Cox parameter and consumer heterogeneity.[5]

Intuitively, the Box-Cox parameter λ relaxes the typical unit demand assumption, and creates greater flexibility on demand curvature. Specifically, equation (9) 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 $\lambda < 1$, 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.

More formally, a measure of demand curvature is the second derivative of log demand:

$$\begin{aligned} \frac{\partial^2 \ln q_j}{\partial p_j^2} = & \int \left(\frac{s_{ij}}{s_j} \frac{\partial^2 \ln s_{ij}}{\partial p_j^2} \right) dP_\nu \\ & + \left[\int \left(\frac{s_{ij}}{s_j} \left(\frac{\partial \ln s_{ij}}{\partial p_j} \right)^2 \right) dP_\nu - \left(\int \frac{s_{ij}}{s_j} \frac{\partial \ln s_{ij}}{\partial p_j} dP_\nu \right)^2 \right] \\ & + (1 - \lambda) \frac{1}{p_j^2}, \end{aligned} \tag{10}$$

[5] Table A.1 in Appendix A.1.4 presents the elasticities in several special cases: $\lambda = 1$ or $\lambda = 0$, with and without heterogeneity in α_i

as shown in Appendix A.1.5. This generalizes the expression obtained by Griffith et al. (2018). The first and second terms have a similar structure, i.e. respectively the weighted average of the second derivative of log individual demand and the weighted variance of the slope of log individual demand. The third term is new and is the absolute value of the conditional individual elasticity, divided by p_j^2.

If expression (10) is positive (negative), demand is log-concave (log-convex). As discussed by Griffith et al. (2018), the first term of (10) is negative if individual demand is log-concave, and the second term is zero without consumer heterogeneity ($\sigma = 0$) and strictly positive otherwise. The new third term is zero with perfectly inelastic conditional demand ($\lambda = 1$) and strictly positive otherwise. Hence, unlike Griffith et al. (2018) demand may be log-convex even in the absence of consumer heterogeneity because of elastic individual demand ($\lambda < 1$). Appendix A.1.5 provides a more detailed formal discussion, including the relationship between (10) and the superelasticity and pass-through. For example, the price elasticity (in absolute value) increases with price (positive superelasticity) for $\lambda \in (0, 1)$, whereas it may decrease with price if $\lambda < 0$. Furthermore, lower values of λ translate into a higher pass-through rate.[6]

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).

3.1 Descriptive Evidence

Our data set on the "Ready-to-Eat" cereal market comes from IRI. A detailed discussion of the data and summary statistics is provided in Appendix A.3. The unit of observation is a product j (barcode), region r (i.e. 6 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

[6]One may in principle target a reduced-form estimate of the pass-through rate as an additional moment if λ is otherwise difficult to identify in practice. But this would require a careful interpretation of the reduced-form pass-through estimate (including whether it incorporates rival responses and/or delayed adjustment).

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:

$$\ln(q_{jrt}) = \eta_j \ln(p_{jrt}) + w_{jrt}\theta_j + u_{jrt}, \tag{11}$$

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. To 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).

Our main interest at this point 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). Our purpose is to obtain an idea of the cross-sectional relationship between the elasticities and prices, without imposing a detailed underlying structure on how η_j varies across products. However, we caution that the estimated η_j do not measure the true elasticities (i.e., they may be biased) because equation (11) assumes there are no cross-price effects.

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 a constant pattern of elasticities across products 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.

3.2 Estimation

We discuss, in turn, the demand specification, identification and instruments, and 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. As we discussed in Subsection 2.1, this includes both the total number of potential consumers L (as usual) 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 parameter estimates are robust to assumptions regarding L (e.g., 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, Box-Cox parameter, and consumer heterogeneity $(\alpha, \lambda, \sigma)$. Berry and Haile (2014) establish identification in a setting that is more general than ours, and we focus here on the role of specific instruments to identify the three main parameters.

To identify the price coefficient α, we use the average prices of the same product in other markets, i.e., Hausman instruments from Hausman (1996). These instruments mainly help to identify the price coefficient (α). As discussed in more detail in Nevo (2000), their validity is based on two assumptions: marginal costs are correlated across markets and demand shocks are uncorrelated (conditional on the included year-region fixed effects). We use Hausman instruments mainly for simplicity here. In other settings, natural price instruments can be cost shifters, such as exchange rates or input prices interacted with product characteristics (e.g., Reynaert and Verboven, 2014).

Note that price enters the demand model non-linearly through the Box-Cox parameter λ. To identify the shape of the demand curve, it is therefore natural to consider different functional forms of the price instruments. Appendix A.2 shows how further guidance can be obtained from constructing (approximately) optimal instruments in a simplified setting without consumer heterogeneity (see also Reynaert and Verboven, 2014; Conlon and Gortmaker, 2020). Concretely, we use the Hausman instruments in their original linear form, and also in square-root and logarithmic form. These are approximately optimal instruments for α if respectively $\lambda = \{1, 0.5, 0\}$. The logarithmic form also resembles the optimal instrument for λ.

Finally, to identify the heterogeneity parameter σ, we use functions of the other product characteristics, 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 mainly help identify distributional parameters (i.e., σ).

In summary, for all demand models our base specification includes the following instruments (apart from the fixed effects): Hausman instruments in linear, square root and logarithmic form, and BLP instruments. In a sensitivity analysis, we also use different sets of instruments: (i) we omit the BLP instruments, as these may be less relevant for demand models without heterogeneity; (ii) we add Hausman instruments per price group (low, medium and high-priced), as this may help identify the heterogeneity coefficient, in particular if elasticities may show a non-linear relationship with price; and (iii) we add approximately optimal instruments for λ assuming market shares are small (based on equation (A.22) in Appendix A.2 for $\lambda = \{1, 0.5, 0\}$).

Estimation: Estimation of the joint model can proceed based on GMM, using the well-documented BLP algorithm for inverting the market shares functions as outlined in Berry et al. (1995). The market shares (i.e., the left-hand side of (7)) now depend on λ; see Appendix A.1.3 for details. We estimate λ together with the other parameters (σ, α, β), and as discussed above additional instruments can be helpful to identify λ.

Note that one may in principle extend our approach to incorporate micro data.

Such data could be used to validate whether the estimate of λ based on aggregate demand data is consistent with the underlying individual demand equation (3). Alternatively, one may use the micro data to impose moment conditions to help identify λ or have possibly more flexible functional forms than (3). See Appendix A.1.2 for further discussion.

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 (and the corresponding 90% confidence intervals show overlap). In sharp contrast, the average implied cost pass-through rates vary widely across the specifications: it is close to one for the logit model, much above one for the simple Box-Cox and BLP models, and in between these extremes for the joint Box-Cox and BLP model. The pass-through elasticity is close to one for the Box-Cox (where $\lambda \approx 0$), while it is less than one in the joint Box-Cox and BLP model. These differences stem from the curvature parameter λ and heterogeneity parameter σ.

In the simple model without consumer heterogeneity (panel (A)), we estimate a Box-Cox parameter $\lambda = -0.039$. 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 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 $\sigma = 0.42$, compared with a mean price sensitivity parameter $\alpha = 1.12$. In the joint Box-Cox and BLP model, we appear to estimate even larger heterogeneity, $\sigma = 0.56$, but the mean price sensitivity parameter also increases to $\alpha = 1.52$ (so σ actually becomes somewhat less important in relative terms). The Box-Cox parameter λ equals 0.592, which is significantly different from both the traditional BLP model ($\lambda = 1$) and the mixed CES model ($\lambda = 0$). The estimate of $\lambda = 0.592$ under consumer heterogeneity contrasts with our earlier estimate of $\lambda = -0.039$ 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.

Table A.3 reports the estimation results from a sensitivity analysis with our three alternative instrument sets discussed above. We find that the results remain very similar to those using the base instruments.

3.4 Implications for price elasticities and markups

In Figure 2, we show how the own-price price elasticities vary across the price distribution by plotting the own-elasticity against own-price for all 73 products. This plot 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 blue dots 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.

The red dots denote 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 almost -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 the own-price elasticity equals roughly -1.2 for the most expensive product (so that the own-elasticity of the most expensive and the cheapest products are similar). While the range of these elasticities is certainly more plausible than the simple logit, the specific U-shape profile of elasticities may not necessarily be realistic.[7] In particular, this pattern implies that elasticities almost triple from the cheapest products to the middle priced products and then implausibly fall by a factor of almost three.

We next present the results for the simple Box-Cox and the joint Box-Cox and BLP model. The orange dots report the simple Box-Cox own-price elasticities, which are roughly independent of 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 black dots represent 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 -2.3 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 and σ has lower relative importance. While the difference in patterns should be clear, this comparison can be further appreciated by referring to Figure A.2 in the appendix. This figure drops the simple logit elasticities, which changes the y-scale, so the differences between the joint model and BLP model are clearer.

These findings are reflected in the pattern of percentage and absolute (euro) markups (reported in Figure A.3 in the appendix). Both the logit and BLP model entail considerably higher percentage 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 percentage markups over the price range. Absolute markups are nearly

[7]For a recent other illustration of a U-shape pattern stemming from price heterogeneity, see Xue (2021). We note, however, that the U-shape profile in the BLP model does not need to hold generally, but only under sufficient heterogeneity for the price valuation parameter.

constant for the logit model. For the BLP model, they increase less than proportionally with price at low price levels and more than proportionally at high price levels (with an extremely high markup for the most expensive product).

We finally ask what the different models imply for the estimated substitution between product pairs. This can be measured by the diversion ratio, i.e. the fraction of sales that goes from product j (row) to product k (column) after a price increase by product j. To illustrate, Table A.4 of Appendix A.4 presents the diversion ratios for the five top selling products. This table shows that including a Box-Cox parameter may alter the estimated substitution to some extent, but not the relative patterns across products. More specifically, in the standard logit (top panel), the diversion ratios are approximately symmetric, and this remains the case in the simple Box-Cox (where $\lambda = -0.039$). In the BLP model (third panel), the diversion ratios are asymmetric (because of the estimated consumer heterogeneity in price sensitivity σ), and these relative magnitudes remain comparable in the joint Box-Cox and BLP model.[8]

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

[8]We caution that the relatively low diversion ratios do not allow one to conclude that, say, Kellogg and Quaker are in different markets. The estimated demand model was designed to assess how curvature affects the price/elasticity relationship rather than provide the most realistic cross-product substitution patterns (since we do not account for non-price product characteristics). The low diversions ratios in part follow from our conservative assumptions regarding the market size, and would also be higher if calculated at the firm level instead of product level.

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. Alternatively, researchers may seek for other forms of increased flexibility on the price parameter, for example through intermediate specifications between the standard BLP model and the non-parametric approach of Compiani (2021), or by modeling consumer heterogeneity in price sensitivity more flexibly as in Griffith et al. (2018) and Miravete et al. (2023).

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), and we also illustrated in our application. 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.

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," .

Conlon, C. and J. Gortmaker (2020): "Best practices for differentiated products demand estimation with pyblp," *The RAND Journal of Economics*, 51, 1108–1161.

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," .

Miravete, E., K. Seim, and J. Thurk (2023): "Elasticity and Curvature of Discrete Choice Demand Models," Tech. rep., Working paper.

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.

Reynaert, M. and F. Verboven (2014): "Improving the performance of random coefficients demand models: the role of optimal instruments," *Journal of Econometrics*, 179, 83–98.

Xue, Q. (2021): "Vertical Relations, Demand Risk, and Upstream Concentration: the Case of the US Automobile Industry," .

5 Table and Figures

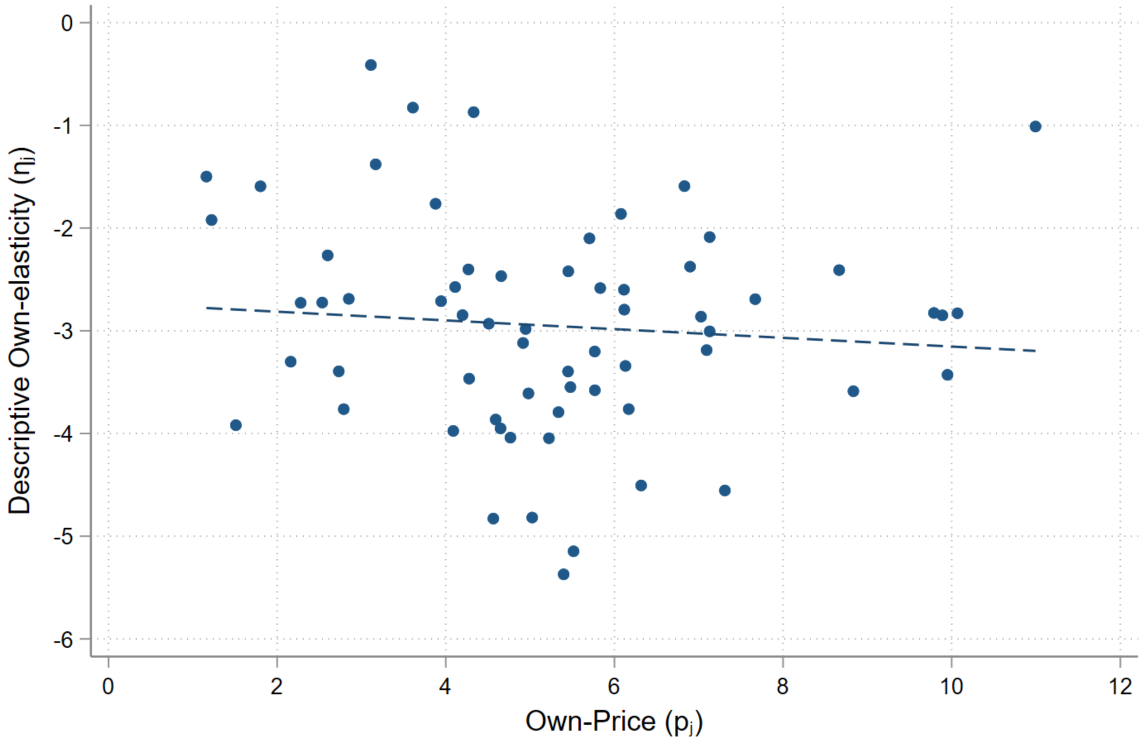

Figure 1: Descriptive own-price elasticity vs. own-price

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 (11) 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.

Table 1: Demand Parameter Estimates

	(A) Simple		(B) Random Coefficient	
	Logit	Box-Cox	Logit	Box-Cox
Price ($-\alpha$)	-0.46	-1.55	-1.12	-1.52
	(0.004)	(0.058)	(0.04)	(0.06)
Price Heterogeneity (σ)	0.00	0.00	0.42	0.56
	–	–	(0.01)	(0.02)
Box-Cox (λ)	1.00	-0.039	1.00	0.592
	–	(0.040)	–	(0.057)
Own-elasticity (η_{jj})	-2.47	-2.50	-2.45	-2.48
Confidence Interval	[-2.51, -2.43]	[-2.56, -2.43]	[-2.56, -2.37]	[-2.58, -2.38]
Pass-through rate	0.99	1.67	1.80	1.20
Pass-through elasticity	0.56	1.00	1.29	0.72

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). "Own-elasticity" refers to the average own-price elasticity across products and markets. A 90% confidence interval for the own-elasticity of the average product is computed using a parametric bootstrap procedure (100 draws of the parameters from their asymptotic distribution). The pass-through rate (elasticity) of product j is the absolute (percentage) price increase after a unit (percentage) cost increase of product j, holding other products' costs constant and assuming single product Bertrand competition (see also Appendix A.1.5). The reported numbers refer to averages across products and markets. 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.

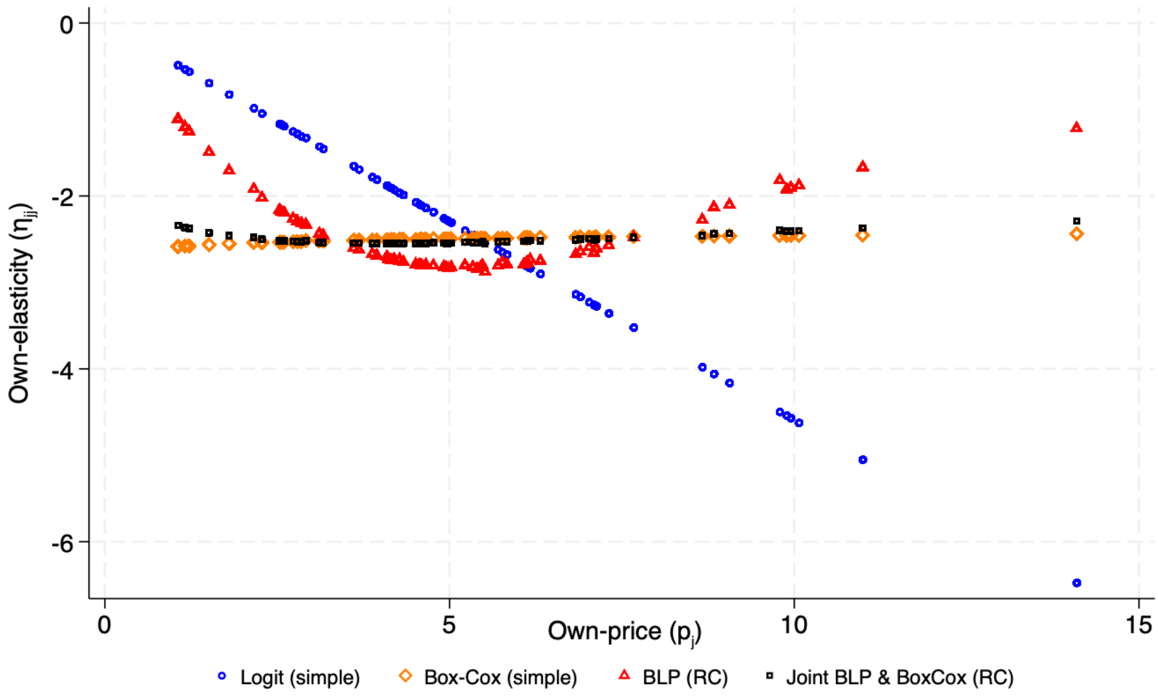

Figure 2: Own-elasticity vs. own-price

Notes: Each dot represents a pair of own-price elasticity and own-price for a particular product and model. Own-elasticity for each model is calculated using equations listed in Table A.1 of Appendix A.1.4. Own-price is the average price per product. The sample consists of 73 products in the cereal category.

A Appendix

This online Appendix provide additional details relating to the model and data for the paper *"Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation"* by Cameron Birchall, Debashrita Mohapatra and Frank Verboven.

A.1 Model

Section A.1.1 discusses the extension of the model to incorporate income heterogeneity. Section A.1.2 discusses the model's implied relationship between expenditure, income and prices, which may be used if one has available micro data. Section A.1.3 derives the analytical Berry inversion in the case without consumer heterogeneity. Section A.1.4 derives the model's own- and cross-price elasticities. Section A.1.5 derives the model's implied curvature, superelasticity and pass-through rate.

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:

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

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 $\int y_i^{1-\lambda} dP_y(y) = \int (\overline{y} + (y_i - \overline{y}))^{1-\lambda}\, dP_y(y)$. Taking the second-order Taylor expansion of the bracketed term gives:

$$(\overline{y} + (y_i - \overline{y}))^{1-\lambda} \approx (\bar{y})^{1-\lambda} + (1-\lambda)(\bar{y})^{-\lambda}(y_i - \bar{y}) - \frac{\lambda(1-\lambda)(\bar{y})^{-\lambda-1}}{2}(y_i - \bar{y})^2 \tag{A.2}$$

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 write $\int y_i^{1-\lambda} dP_y(y)$ as:

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

One may feasibly substitute this approximation 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 $(\lambda, \overline{y}, \sigma_y^2)$.

To improve the approximation, one may also add a third-order term to (A.3):

$$\frac{(1+\lambda)\lambda(1-\lambda)(\bar{y})^{-\lambda-2}}{6} \int (y_i - \bar{y})^3\, dP_y(y), \tag{A.4}$$

which includes the third central moment of the income distribution. Note that this term is zero if the income distribution is normal, but it is positive if it is skewed (as empirically observed, e.g. under a log normal income distribution) and $\lambda \in (-1, 1)$.

A.1.2 Implications for the relationship between expenditure, income, and prices

In this section, we show how different values of the Box-Cox parameter λ affect the relationship between conditional demand, income, and price. The purpose is twofold: (i) show how relaxing the elasticity-price relationship affects relationships between consumption and price/income; (ii) show how one could use these relationships as additional micro moments if one had access to micro data.

The conditional demand equation (3) implies that an individual i's consumption share for product j is

$$\frac{p_j q_{ij}}{y_i} = \gamma \left(\frac{p_j}{\gamma y_i}\right)^{\lambda}$$

Hence, its consumption share is independent of income if $\lambda = 0$, it is decreasing in income for $\lambda \in (0, 1)$, and it is increasing in income for $\lambda < 0$. With micro data, one may assess this to validate the aggregate demand model. Taking the log of the demand equation (3) gives

$$\ln(q_{ij}) = (1-\lambda)\ln(\gamma) + (1-\lambda)\ln(y_i) - (1-\lambda)\ln(p_j). \tag{A.5}$$

Using micro data, the constant identifies γ, and variation in price and income identifies λ. The equation constrains the price and income elasticities to be equal, but this can be made more flexible with separate λ_p and λ_y as discussed in footnote 2:

$$\ln(q_{ij}) = (1-\lambda_y)\ln(\gamma) + (1-\lambda_y)\ln(y_i) - (1-\lambda_p)\ln(p_j). \tag{A.6}$$

Hence, with micro data one can validate the plausibility of λ from the aggregate demand model. If this turns out to be too restrictive, one may allow for a separate income coefficient λ_y or consider even more flexible conditional demands, such as in Griffith et al. (2018).

A validation analysis may in principle also be conducted with aggregate data but in our application variation in budgets ($\gamma\bar{y}_r$) is limited to only six regions. To make this concrete, Figure A.1 plots the relationship between the total (weekly) quantity of cereal and income per potential consumer, for each region.[9] This clarifies there is no clear pattern, so in our application we identify the curvature parameter λ from price rather than income variation.

A.1.3 Market Share System Inversion

Estimating the model requires solving the demand system (7), i.e. $s_j = s_j(\delta)$ for $j = 1, \ldots, J$, for the mean valuation vector δ. For a fixed value of $\lambda \leq 1$, the "market shares" (or aggregate choice probabilities) s_j are defined by

$$s_j \equiv \left(\frac{p_j}{\gamma\bar{y}}\right)^{1-\lambda} \frac{q_j}{L}, \tag{A.7}$$

which are less than one if $q_j < L$ and $p_j < \gamma\bar{y}$. The market share functions $s_j(\delta)$ in (7) are the same as in Berry et al. (1995), hence satifying their conditions for a contraction

[9]Note that the budget is calculated using the methodology in subsection 3.2, but we obtain a very similar picture if we instead use a direct income measure.

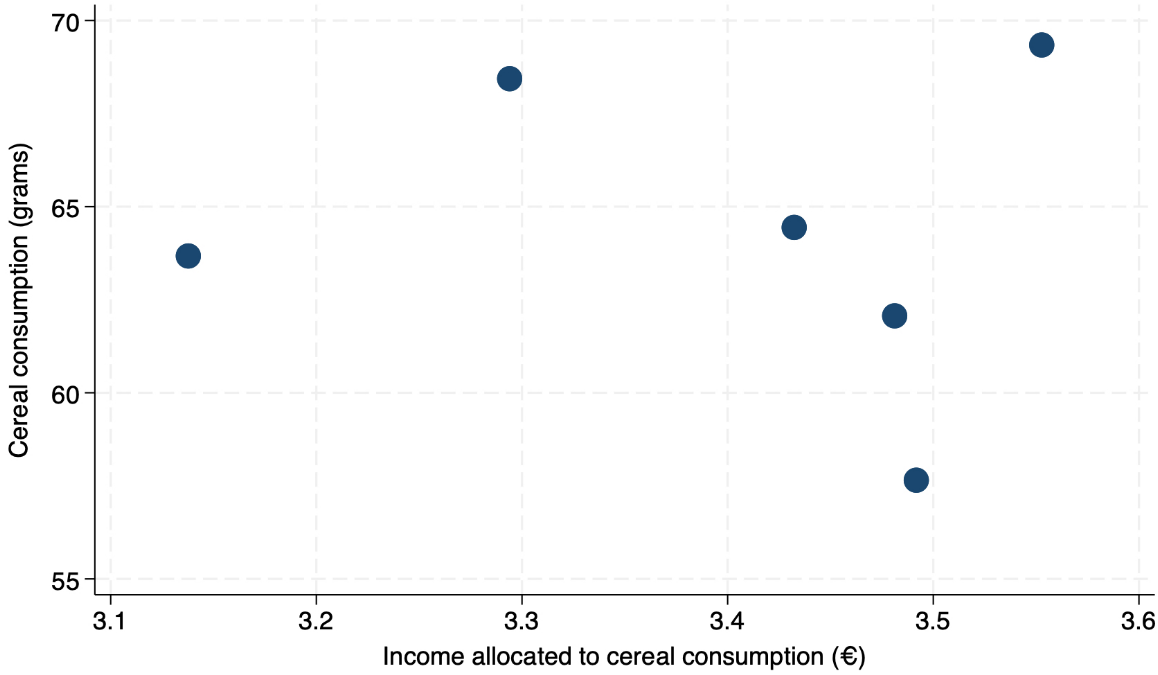

Figure A.1: Cereal consumption per potential consumer vs. income

Explanation: Each dot represents a pair of total (weekly) cereal consumption per potential consumer and income ($\gamma\bar{y}_r$), for a particular province r. The budget is calculated using the methodology explained in subsection 3.2.

mapping:

$$f(\delta) = \delta + \ln s - \ln s(\delta) \tag{A.8}$$

To obtain a solution to $f(\delta) = \delta$, one may apply fixed point iteration on (A.8), or use a fixed point acceleration method, as discussed in Conlon and Gortmaker (2020). In practice, simple fixed point iteration worked well with our data, and we also did not encounter convergence difficulties when choosing alternative values of λ to minimize the GMM objective function.

Abstracting from consumer heterogeneity we can obtain an analytic solution for the market share inversion following similar steps as in Berry (1994). The estimating equation (7) simplifies to:

$$\frac{p_j^{1-\lambda} q_j}{L\left(\gamma \bar{y}\right)^{1-\lambda}} = \frac{\exp\left(\delta_j\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_k\right)}. \tag{A.9}$$

Summing over $j = 1, ..., J$, and rearranging, we may write the following choice probability for good 0:

$$\frac{L(\gamma\bar{y})^{1-\lambda} - \sum_{k=1}^{J} p_k^{1-\lambda} q_k}{L(\gamma\bar{y})^{1-\lambda}} = \frac{1}{1 + \sum_{k=1}^{J} \exp(\delta_k)}. \tag{A.10}$$

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:

$$\frac{p_j^{1-\lambda} q_j}{L(\gamma\bar{y})^{1-\lambda} - \sum_{k=1}^{J} p_k^{1-\lambda} q_k} = \exp\left(\delta_j\right). \tag{A.11}$$

Taking logs arrives at equation (8).

A.1.4 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:

$$\eta_{jj} = \frac{\partial q_j}{\partial p_j}\frac{p_j}{q_j} \tag{A.12}$$

$$= \left(\int \frac{\partial s_{ij}}{\partial p_j} dP_\nu(\nu) \left(\frac{\gamma\bar{y}}{p_j}\right)^{1-\lambda} L + \int s_{ij} dP_\nu(\nu) \frac{\partial}{\partial p_j}\left(\frac{\gamma\bar{y}}{p_j}\right)^{1-\lambda} L\right) \frac{p_j}{q_j} \tag{A.13}$$

$$= \left[-\int \alpha_i p_j^{\lambda} s_{ij}\left(1 - s_{ij}\right) dP_\nu(\nu) - \int s_{ij} dP_\nu(\nu)(1-\lambda)\right]\frac{1}{s_j} \tag{A.14}$$

$$= -\frac{p_j^{\lambda}}{s_j}\int \alpha_i s_{ij}\left(1 - s_{ij}\right) dP_\nu(\nu) - (1-\lambda). \tag{A.15}$$

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:

$$\eta_{jk} = \frac{\partial q_j}{\partial p_k} \frac{p_k}{q_j} \tag{A.16}$$
$$= \left(\int \frac{\partial s_{ij}}{\partial p_k} dP_\nu(\nu) \left(\frac{\gamma \bar{y}}{p_j} \right)^{1-\lambda} L \right) \frac{p_k}{q_j} \tag{A.17}$$
$$= \frac{p_k^\lambda}{s_j} \int \alpha_i s_{ij} s_{ik} dP_\nu(\nu), \tag{A.18}$$

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 A.1 list the own- and cross-elasticities for each model.

Table A.1: Own- and Cross-elasticities for each model

RC	Model	Own-elasticity, η_{jj}	Cross-elasticity, η_{jk}
Yes	Box-Cox	$-\frac{p_j^\lambda}{s_j} \int \alpha_i s_{ij} \left(1 - s_{ij}\right) dP_\nu(\nu) - (1 - \lambda)$	$\frac{p_k^\lambda}{s_j} \int \alpha_i s_{ij} s_{ik} dP_\nu(\nu)$
	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	$-p_j^\lambda \alpha \left(1 - s_j\right) - (1 - \lambda)$	$p_k^\lambda \alpha s_k$
	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.1.5 Curvature, superelasticity and pass-through

Log-concavity versus log-convexity: To gain a further understanding on the role of the curvature parameter λ, we first calculate the second derivative of the log of aggregate demand with respect to price p_j. If this is negative (positive), demand is log-concave (log-convex). Tedious calculations show that:

$$\begin{aligned}
\frac{\partial^2 \ln q_j}{\partial p_j^2} &= \frac{1}{s_j}\int \frac{\partial^2 s_{ij}}{\partial p_j^2} dP_v - \left(\frac{1}{s_j}\int \frac{\partial s_{ij}}{\partial p_j} dP_v\right)^2 + (1-\lambda)\frac{1}{p_j^2} \\
&= \int \left(\frac{s_{ij}}{s_j}\frac{\partial^2 \ln s_{ij}}{\partial p_j^2}\right) dP_v \\
&\quad + \left[\int \left(\frac{s_{ij}}{s_j}\left(\frac{\partial \ln s_{ij}}{\partial p_j}\right)^2\right) dP_v - \left(\int \frac{s_{ij}}{s_j}\frac{\partial \ln s_{ij}}{\partial p_j} dP_v\right)^2\right] \\
&\quad + (1-\lambda)\frac{1}{p_j^2},
\end{aligned}$$

where $s_j \equiv \int s_{ij} dP_v = q_j p_j^{1-\lambda} / \left(\int (\gamma y_i)^{1-\lambda} \, dP_y L\right)$. The second equality neatly generalizes the expression obtained by Griffith et al. (2018) to our joint Box-Cox and BLP model. The first and second terms are similar (although now a function of λ), i.e. respectively the weighted average of the second derivative of log individual demand and the weighted variance of the slope of log individual demand. The third term is new and is the absolute value of the conditional individual elasticity, divided by p_j^2.

As discussed by Griffith et al. (2018), the first term is negative if individual demand is log-concave. The second term is zero without consumer heterogeneity ($\sigma = 0$), and strictly positive otherwise. The third term is strictly positive except if conditional individual demand is perfectly inelastic ($\lambda = 1$). In sum, despite possibly log-concave individual demands (first term), aggregate demand may be log-convex because of either consumer heterogeneity (second term) or elastic individual demand (third term). Put differently, unlike Griffith et al. (2018) demand may be log-convex even in the absence of consumer heterogeneity.

To see this more concretely, without consumer heterogeneity ($\sigma = 0$) we obtain:

$$\begin{aligned}\frac{\partial^2 \ln q_j}{\partial p_j^2} &= \frac{\partial^2 \ln s_j}{\partial p_j^2} + (1-\lambda)\frac{1}{p_j^2} \\ &= -\alpha^2 p_j^{2(\lambda-1)} s_j(1-s_j) \\ &\quad +(1-\lambda)\frac{1}{p_j^2}\left(\alpha(1-s_j)p_j^\lambda + 1\right), \end{aligned} \tag{A.19}$$

since the second term vanishes and the first term simplifies because $s_{ij} = s_j$. In the usual unit demand case ($\lambda = 1$), this is negative, so demand is log-concave. However, if $\lambda < 1$, this becomes positive provided that s_j is sufficiently small.

Superelasticity: Further insights can be obtained from the superelasticity E_j, i.e. the elasticity of the own-price elasticity. In general, we can calculate this as

$$\begin{aligned} E_j &\equiv \frac{\partial \eta_{jj}}{\partial p_j}\frac{p_j}{\eta_{jj}} \\ &= \left(\frac{\partial^2 q_j}{\partial p_j^2}\frac{p_j}{q_j} - \frac{\partial q_j}{\partial p_j}\frac{p_j}{q_j^2}\frac{\partial q_j}{\partial p_j} + \frac{\partial q_j}{\partial p_j}\frac{1}{q_j}\right)\frac{p_j}{\eta_{jj}} \\ &= \frac{\partial^2 \ln q_j}{\partial p_j^2}\frac{p_j^2}{\eta_{jj}} + 1. \end{aligned}$$

According to the last equality, $E_j > 1$ if and only if demand is log concave ($\frac{\partial^2 \ln q_j}{\partial p_j^2} < 0$).

If there is no consumer heterogeneity ($\sigma = 0$), we can insert (A.19) and use $\eta_{jj} = -\alpha p_j^\lambda \left(1 - s_j\right) - (1-\lambda)$ from Table A.1 to obtain:

$$E_j = \alpha p_j^\lambda \left(\lambda + \alpha p_j^\lambda s_j\right)(1-s_j)\frac{-1}{\eta_{jj}}.$$

In the usual unit demand case ($\lambda = 1$),

$$E_j = \alpha s_j p_j + 1,$$

so the superelasticity E_j strictly exceeds 1 and approaches 1 as s_j becomes small. In the constant expenditures case ($\lambda = 0$),

$$E_j = \frac{\alpha^2(1-s_j)s_j}{\alpha(1-s_j)+1},$$

so the superelasticity E_j may be below 1 and approaches 0 as s_j becomes small. For intermediate cases, i.e. $\lambda \in (0,1)$, we have that $E_j \in (0,1)$ as s_j becomes small (because then $E_j = \lambda \frac{\alpha p_j^\lambda}{\alpha p_j^\lambda + 1 - \lambda}$). Finally, for $\lambda < 0$, it is possible that $E_j < 0$ provided that s_j is sufficiently small.

In sum, without consumer heterogeneity and with sufficiently small market shares the own-price elasticity increases nearly proportionally with price for $\lambda = 1$, it is nearly independent of price if $\lambda = 0$, and it may be decreasing in price if $\lambda < 01$. Adding consumer heterogeneity complicates the relationship, and can give rise to a U-shaped pattern, as documented in the application, especially if λ is close to one (traditional BLP model).

Pass-through: It is also instructive to consider simplified analytical expressions of the cost pass-through rate. Assume single product price-setting firms with constant marginal costs c_j for firm j and consider the own-cost pass-through rate, holding constant the other firms' prices, i.e. abstracting from equilibrium responses by other firms.[10] Firm j's first-order condition defines the optimal price $p_j^*(c_j)$. Implicit differentiation gives the following relationship between the pass-through rate and the superelasticity:

$$\begin{aligned} \frac{\partial p_j^*}{\partial c_j} &= \frac{(-\eta_{jj})}{(-\eta_{jj} - 1 + E_j)} \\ &= \frac{1}{1 - \frac{\partial^2 \ln q_j}{\partial p_j^2} \frac{p_j^2}{\eta_{jj}^2}}, \end{aligned}$$

which confirms the well-known result that pass-through is incomplete if and only if $E_j > 1$, or equivalently, if and only if demand is log-concave.

For our demand model without consumer heterogeneity ($\sigma = 0$), it can be verified

[10]The reported pass-through rates in our application account for equilibrium responses.

that

$$\frac{\partial p_j^*}{\partial c_j} = \frac{\left(\alpha(1-s_j)p_j^\lambda + (1-\lambda)\right)^2}{\alpha^2(1-s_j)p_j^{2\lambda} - (1-\lambda)\,\lambda + \alpha(1-s_j)p_j^\lambda\,(1-\lambda)}$$

In the unit demand case ($\lambda = 1$), this simplifies to:

$$\frac{\partial p_j^*}{\partial c_j} = 1 - s_j.$$

So the pass-through rate is strictly less than 1, but it approaches 1 as s_j becomes small.

In the constant expenditures case ($\lambda = 0$), this becomes

$$\frac{\partial p_j^*}{\partial c_j} = \frac{\left(\alpha(1-s_j)+1\right)^2}{\alpha\,(\alpha+1)\,(1-s_j)}.$$

So the pass-through rate exceeds 1 if s_j is sufficiently small. The pass-through elasticity $\frac{\partial p_j^*}{\partial c_j}\frac{\partial c_j}{\partial p_j^*} = \frac{\alpha(1-s_j)+1}{\alpha+1}$ is less than 1, but it approaches 1 as s_j becomes small.

In sum, without consumer heterogeneity and with small s_j there is constant absolute pass-through for $\lambda = 1$ while there is constant percentage pass-through for $\lambda = 0$. The parameter λ thus provides extra flexibility on the pass-through rates compared with the unit demand or constant expenditures models.

A.2 Optimal instruments for Box Cox model without heterogeneity

To gain intuition on choosing functional forms of specific instruments, this Appendix provides approximate expressions of optimal instruments for the Box Cox model without heterogeneity. The error term in the Box Cox demand model is:

$$\xi_j = \ln\left(\frac{q_j p_j^{1-\lambda}}{(\gamma y)^{1-\lambda} L - \sum_k q_k p_k^{1-\lambda}}\right) - x_j\beta - \alpha\left(\frac{p_j^\lambda - 1}{\lambda}\right) \tag{A.20}$$

The optimal (i.e "most efficient") instrument for a parameter is the expected value of the derivative of that parameter.

A.2.1 Optimal instrument for β

The optimal instrument for β is the expected value of

$$E\left(\frac{\partial \xi_j}{\partial \beta}\right) = -E\left(x_j\right) = -x_j.$$

So the optimal instrument for β is simply the variable itself.

A.2.2 The optimal instrument for α

The optimal instrument for α is the expected value of:

$$E\left(\frac{\partial \xi_j}{\partial \alpha}\right) = -E\left(\frac{p_j^\lambda - 1}{\lambda}\right) = -\frac{E\left(p_j^\lambda\right) - 1}{\lambda}.$$

A first-order Taylor expansion around the mean $\widehat{p}_j \equiv E\left(p_j\right)$ gives

$$E\left(\frac{\partial \xi_j}{\partial \alpha}\right) \approx -\frac{\widehat{p}_j^\lambda - 1}{\lambda}.$$

One may obtain $\widehat{p}_j$ from a regression on instruments w_j, and then compute $\widehat{p}_j^\lambda$ (removing the scaling and constant). This instrument is however not feasible because it depends on λ. As a solution one may use inefficient instruments in a first stage. For example, one may take three instruments $\widehat{p}_j, \widehat{p}_j^{0.5}$ and $\ln \widehat{p}_j$ as implied by $\lambda = \{1, 0.5, 0\}$ to obtain a first-stage estimate $\widehat{\lambda}^{FS}$ and then use this value in an (approximately) efficient second stage.

A.2.3 Optimal instrument for λ

Calculating the derivative of (A.20) with respect to λ is much more involved because λ enters in both the share term and the price term in (A.20). Tedious calculations show that

$$\begin{aligned}
\frac{\partial \xi_j}{\partial \lambda} &= -\ln p_j + \frac{(y\gamma)^{1-\lambda} L \ln(y\gamma) - \sum_k q_k p_k^{1-\lambda} \ln p_k}{(\gamma y)^{1-\lambda} L - \sum_k q_k p_k^{1-\lambda}} - \alpha \frac{\lambda p_j^\lambda \ln p_j - p_j^\lambda + 1}{\lambda^2} \\
&= -\ln p_j + \frac{\ln(y\gamma) - \sum_k s_k \ln p_k}{1 - \sum_k s_k} - \alpha \frac{\lambda p_j^\lambda \ln p_j - p_j^\lambda + 1}{\lambda^2}, \qquad \text{(A.21)}
\end{aligned}$$

where the second equality substitutes the shares $s_k = \frac{p_k^{1-\lambda} q_k}{(\gamma y)^{1-\lambda} L}$.

One may again substitute predicted values $\widehat{p}_j$ and $\widehat{q}_j$ (or $\widehat{s}_k$) and apply a two-stage approach. In a first stage, one may calculate three instruments by substituting $\lambda = \{1, 0.5, 0\}$ in (A.21). This is straightforward for $\lambda = 1$ and $\lambda = 0.5$. For $\lambda = 0$, apply l'Hospital's rule twice on last term to obtain:

$$-\ln p_j + \frac{\ln(y\gamma) - \sum_k s_k \ln p_k}{1 - \sum_j s_k} - \alpha \frac{(\ln p_j)^2}{2}$$

If s_k and $\sum_k s_k$ are small, we can simplify (A.21) to

$$\frac{\partial \xi_j}{\partial \lambda} \approx -\ln p_j + \ln(y\gamma) - \alpha \frac{\lambda p_j^{\lambda} \ln p_j - p_j^{\lambda} + 1}{\lambda^2}. \qquad \text{(A.22)}$$

Overall, these formulas provide guidance to indicate that one may use different functional forms of $\widehat{p}_j$ as instruments.

A.3 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 A.2 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 A.2: 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.4 Additional Tables and Figures

This section reports additional tables and figures referred to in the text. Table A.3 reports the demand parameter estimates using alternative instrument sets, as a sensitivity analysis to Table 1 in the text. Table A.4 reports the diversion ratios from each demand model for the 5 most-selling products, computed as $-\frac{\partial s_k/\partial p_j}{\partial s_j/\partial p_j}$. Figure A.3 shows how implied product markups vary with own price.

Table A.3: Demand Parameter Estimates – Alternative Instruments

	(i)	(ii)	(iii)	(iv)
		Logit		
Price ($-\alpha$)	-0.46	-0.46	-0.45	-0.46
	(0.004)	(0.004)	(0.004)	(0.004)
		Simple Box-Cox		
Price ($-\alpha$)	-1.55	-1.56	-1.54	-1.54
	(0.06)	(0.06)	(0.06)	(0.06)
Boxcox (λ)	-0.039	-0.043	-0.044	-0.031
	(0.040)	(0.040)	(0.041)	(0.041)
		BLP		
Price ($-\alpha$)	-1.12	-1.15	-1.12	-1.13
	(0.04)	(0.03)	(0.03)	(0.03)
Price heterogeneity (σ)	0.42	0.44	0.43	0.43
	(0.01)	(0.01)	(0.01)	(0.01)
		BLP & Box-Cox		
Price ($-\alpha$)	-1.52	-1.52	-1.49	-1.52
	(0.06)	(0.07)	(0.07)	(0.07)
Price heterogeneity (σ)	0.56	0.57	0.55	0.56
	(0.02)	(0.03)	(0.03)	(0.03)
Box-Cox (λ)	0.592	0.645	0.631	0.595
	(0.057)	(0.053)	(0.055)	(0.059)

Notes: Robust standard errors reported in parentheses. 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. (i) refers to base IV, as presented in Table 1; (ii): base IV without BLP instruments; (iii) base IV plus Hausman instruments per price group; (iv) base IV plus approximately optimal instruments for λ under small market shares.

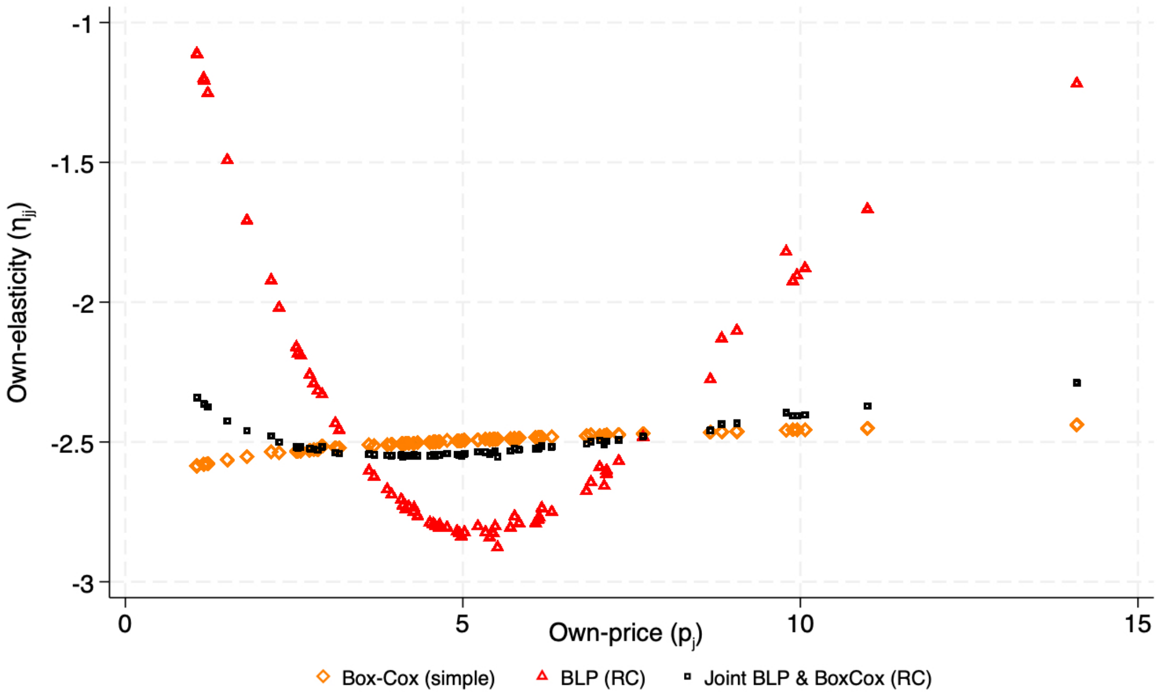

Figure A.2: Own-elasticity vs. own-price (without simple logit)

Explanation: Each dot represents a pair of own-elasticity and own-price for a particular product and model. We calculate the elasticity for each model using equations listed in Table A.1 of Appendix A.1.4. Own-price is the average price per product. The sample consists of 73 products in the cereal category.

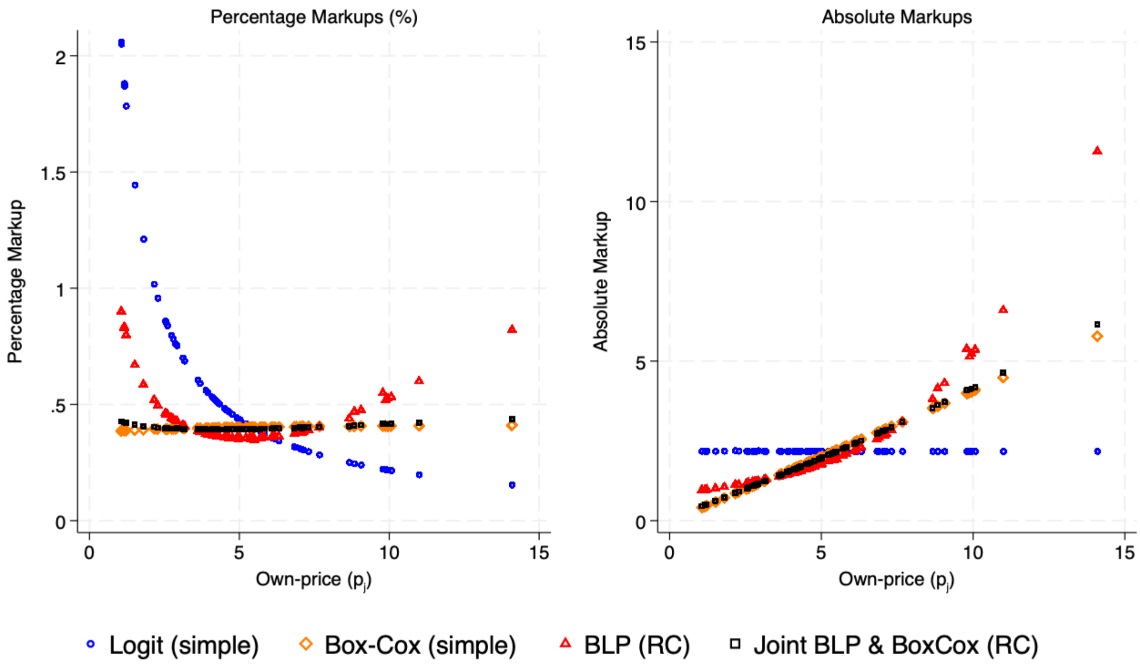

Figure A.3: Markups vs. own-price

Explanation: Both panels report the distribution of the markup vs. own-price for a particular product and model. The left panel reports *percentage* markups, which are $\mu_j = \frac{p_j - c_j}{p_j} = \frac{1}{\epsilon_{jj}}$ and the right panel reports *absolute* (or Euro) markups, which are $\mu_j = p_j - c_j = \frac{p_j}{\epsilon_{jj}}$. For each figure, each dot represents a pair of markups and price for a particular model. The product markups use the elasticity for each model using equations listed in Table A.1 of Appendix A.1.4. Own-price is the average price per product. The sample consists of 73 products in the cereal category.

Table A.4: Diversion ratios of top selling models

Product	K Special	Q Cruesli	K Fruit & Fibre	J Original	K Loops
			Logit		
Kellogg's Special	-	0.0023	0.0013	0.0227	0.0008
Quaker Cruesli	0.0016	-	0.0013	0.0227	0.0008
Kellogg's Fruit & Fibre	0.0016	0.0023	-	0.0227	0.0008
Jordans Original	0.0016	0.0024	0.0013	-	0.0008
Kellogg's Loops	0.0016	0.0023	0.0013	0.0227	-
			Simple Box-Cox		
Kellogg's Special	-	0.0015	0.0014	0.0056	0.0010
Quaker Cruesli	0.0022	-	0.0014	0.0056	0.0010
Kellogg's Fruit & Fibre	0.0022	0.0015	-	0.0056	0.0010
Jordans Original Crunchy	0.0023	0.0016	0.0014	-	0.0010
Kellogg's Loops	0.0022	0.0015	0.0014	0.0056	-
			BLP		
Kellogg's Special	-	0.0062	0.0053	0.0249	0.0035
Quaker Cruesli	0.0019	-	0.0018	0.0239	0.0011
Kellogg's Fruit & Fibre	0.0044	0.0048	-	0.0247	0.0023
Jordans Original	0.0005	0.0017	0.0007	-	0.0004
Kellogg's Loops	0.0053	0.0053	0.0042	0.0247	-
			BLP & Box-Cox		
Kellogg's Special	-	0.0045	0.0041	0.0150	0.0028
Quaker Cruesli	0.0025	-	0.0019	0.0145	0.0012
Kellogg's Fruit & Fibre	0.0044	0.0038	-	0.0150	0.0021
Jordans Original	0.0009	0.0015	0.0008	-	0.0005
Kellogg's Loops	0.0050	0.0041	0.0035	0.0150	-

Explanation: This table reports the diversion ratios for the top selling products. The diversion ratio from product j (row) to product k (column) is defined as the ratio of the cross-price derivative over the absolute value of the own-price derivative, i.e. $\frac{\partial q_k / \partial p_j}{-\partial q_j / \partial p_j}$.