THEMA Working Paper n°2021-04
CY Cergy Paris Université, France

The Inverse Product Differentiation Logit Model

Mogens Fosgerau, Julien Monardo, André de Palma

January 2021

The Inverse Product Differentiation Logit Model*

Mogens Fosgerau† Julien Monardo‡ André de Palma§

January 27, 2021

Abstract

We propose the Inverse Product Differentiation Logit (IPDL) model, a structural (inverse) demand model for differentiated products that captures market segmentation with segments that may overlap in any way. The IPDL model generalizes the nested logit model to allow richer substitution patterns, including complementarity in demand, and can be estimated by linear instrumental variable regression using aggregate data. We use the IPDL model to estimate the demand for cereals in Chicago. We then extend it to a general demand model that is consistent with a utility model of heterogeneous, utility-maximizing consumers. (JEL: C26, D11, D12, L)

*First Draft: March, 2016. Authors are in alphabetical order. We would especially like to thank Steven Berry who guided us to better highlight our main contributions. We are also grateful for comments from Xavier D'Haultfoeuille, Laurent Linnemer, Moshe Ben-Akiva, Debopam Bhattacharya, Alfred Galichon, Laura Grigolon, Alessandro Iaria, Jonas Lieber, Hugo Molina, Yurii Nesterov, Aureo de Paula, John Rehbeck, Bernard Salanié, Jesper R.-V. Sørensen, Michelle Sovinsky, Kenneth Small and Thibaud Vergé as well as from seminar participants in Paris, Leuven, Mannheim, Amsterdam, Copenhagen, Chicago, Montréal, and Toulouse. Mogens Fosgerau has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 740369). André de Palma has received financial support from the ARN project (Elitisme). This paper is a continuation of the working paper that was first circulated in 2016 with the title "Demand system for market shares" and later with the title "Demand Models for Differentiated Products with Complementarity and Substitutability".

†University of Copenhagen; mogens.fosgerau@econ.ku.dk.

‡Telecom Paris; julien.monardo@telecom-paris.fr.

§CY Cergy Paris Université; andre.de-palma@cyu.fr.

1 Introduction

We propose the Inverse Product Differentiation Logit (IPDL) model of demand for differentiated products. The IPDL model is a structural inverse demand model that captures market segmentation by several characteristics with segments that may overlap in any way. It improves on the nested logit model by accommodating richer substitution patterns but has the same attractive features: first, following Berry (1994), the IPDL model can easily be estimated by linear instrumental variables regression to deal with the endogeneity of prices and market shares while allowing for unobserved product characteristics; second, it is consistent with a model of heterogeneous, utility-maximizing consumers.

The nested logit model is commonly used to estimate the aggregate demand in differentiated products markets that exhibit product segmentation (e.g. Björnerstedt and Verboven, 2016; Berry et al., 2016). Its inverse demand function has a closed-form expression, which makes standard linear instrumental variable regression techniques easily applicable. However, it imposes severe restrictions on the substitution patterns. In particular, the substitution between products is implied by a nesting structure that groups products into nests according to one or more characteristics. The nested logit model restricts this nesting structure to be hierarchical, nests at one level of segmentation are divided into subnests at the next level and so on. Moreover, the sequence in the hierarchy is not unique and often not obvious.[1]

In many applications, a non-hierarchical structure is more appropriate. A plethora of GEV models (McFadden, 1978; Fosgerau et al., 2013) based on different non-hierarchical structures have been proposed.[2] However, none of the GEV models beyond the nested logit model yield inverse demand in closed form. Like these GEV

[1] Hellerstein (2008) writes, concerning the beers market, "[D]emand models such as the multi-stage budgeting model or the nested logit model do not fit this market particularly well. It is difficult to define clear nests or stages in beer consumption because of the high cross-price elasticities between domestic light beers and foreign light and regular beers. When a consumer chooses to drink a light beer that also is an import, it is not clear if he categorized beers primarily as domestic or imported and secondarily as light or regular, or vice versa."

[2] These models include, e.g. the ordered logit (Small, 1987), the ordered nested logit (Grigolon, 2020), and the FC-MNL (Davis and Schiraldi, 2014). Closest in spirit to the IPDL model is the product differentiation logit model (Bresnahan et al., 1997), which defines a non-hierarchical structure by grouping products along several characteristics.

models, the IPDL model relies on a general non-hierarchical structure to control the substitution patterns, where products that share more overlapping groups are closer substitutes. The crucial difference is that the non-hierarchical structure retains the IPDL inverse demand function in closed form.

A large literature employs the BLP method (Berry et al., 1995), which uses a random coefficient logit (RCL) model with structural error terms to allow for unobserved product characteristics, and which handles endogeneity of prices and market shares. Substitution patterns are determined by a random coefficients specification of the distribution of unobserved preference heterogeneity, which means that complementarity is ruled out. By contrast, the substitution patterns in the IPDL model are determined by a segmentation of the differentiated products, which leads to a model that allows complementarity. The BLP method involves a non-linear, non-convex optimization problem as well as the simulation and numerical inversion of the demand function.[3] By contrast, the IPDL model can be estimated by linear instrumental variable regression to deal with endogeneity issues due to the presence of unobserved product characteristics. Indeed, with the IPDL model, we depart from the BLP method by directly specifying the inverse demand function so that numerical demand inversion is not required.

The IPDL model is consistent with utility maximization. We introduce the Generalized Logit (GL) inverse demand model, which has the IPDL model as a special case, and show that any GL model is consistent with a representative consumer who chooses a vector of market shares to maximize her utility function subject to a budget constraint, trading off variety against quality.[4] Furthermore, the GL model is consistent with the model of utility-maximizing, heterogeneous consumers of Allen and Rehbeck (2019b). Finally, we show that any Additive Random Utility Model (ARUM) (McFadden, 1981), is a GL model, but not vice-versa.

Important economic questions hinge on whether products are substitutes or

[3]Conlon and Gortmaker (2020) consolidate best estimation practices in a python package. Note also that other approaches that implement the BLP estimator have been proposed (Dubé et al., 2012; Lee and Seo, 2015).

[4]In the GL model, the taste for variety is given by a generalization of entropy. When the model corresponds to an additive random utility model, it is the convex conjugate (Rockafellar, 1970) of the surplus function (Hofbauer and Sandholm, 2002; Galichon and Salanié, 2020; Fosgerau et al., 2020).

complements in demand, i.e. on the sign of cross-price elasticities of demand. In particular, this directly affects the incentives to introduce a new product on the market, to bundle products, to merge, etc. (Gentzkow, 2007; Iaria and Wang, 2020). In contrast to the ARUM (and hence also to the RCL model), products can be complements in the IPDL and GL models. As a by-product, we therefore establish a new demand inversion result for the GL model, which extends Berry (1994) and supplements Berry et al. (2013) by allowing complementarity in unit-demand models.

Other approaches exist for estimating the demand for differentiated products (see, e.g. Barnett and Serletis, 2008; Nevo, 2011). One is the flexible functional form approach (e.g. the AIDS model of Deaton and Muellbauer, 1980), which has been successfully applied to many economic questions. However, in this approach, the econometric error terms have no immediate structural interpretation, there are very many parameters to estimate, and the introduction of new products cannot be addressed. Other authors (e.g., Pinkse and Slade, 2004; Haag et al., 2009; Blundell et al., 2012) propose to estimate demand functions semi- or non-parametrically. Closest to our approach is Compiani (2020) who suggests to non-parametrically estimate inverse demand functions for differentiated products based on aggregate data. This last approach faces the trade-off between functional form restrictions and the need for large datasets.

The paper is organized as follows. Section 2 presents our general setting and discuss the role of demand inversion. Section 3 introduces the IPDL model, discusses estimation with aggregate data, and presents simulations suggesting that the IPDL model offers a good compromise between computational simplicity and richness of the substitution patterns. Section 4 uses the IPDL model to estimate the demand for ready-to-eat cereals in Chicago, finding evidence that complementarity exists between some products from different market segments. Section 5 introduces the Generalized Logit model. Section 6 relates the GL model to the ARUM and the representative consumer model. Section 7 concludes. A supplement provides simulations results on the IPDL model as well as general methods and examples for building other GL models than the IPDL model.

2 General Setting

We begin by introducing our setting and discussing the role of demand inversion for estimation. Consider a population of consumers making choices among a set of $J+1$ differentiated products, indexed by $\mathcal{J} = \{0, 1, \ldots, J\}$, where product $j = 0$ is the outside good. We consider aggregate data on market shares s_{jt}, prices p_{jt} and K product/market characteristics $\mathbf{x}_{jt}$ for each product $j = 1, \ldots, J$ in each market $t = 1, \ldots, T$ (Berry, 1994; Berry et al., 1995; Nevo, 2001). For each market t, the market shares s_{jt} are positive and sum to 1, i.e. $\mathbf{s}_t = (s_{0t}, \ldots, s_{Jt}) \in \text{relint}(\Delta)$, where $\text{relint}(\Delta)$ is the relative interior of the unit simplex in $\mathbb{R}^{J+1}$.

Following Berry and Haile (2014), let $\delta_{jt} \in \mathbb{R}$ be an index given by

$$\delta_{jt} = \delta\left(p_{jt}, \mathbf{x}_{jt}, \xi_{jt}; \boldsymbol{\theta}_1\right), \quad j \in \mathcal{J}, \quad t = 1, \ldots, T,$$

where $\xi_{jt} \in \mathbb{R}$ is an unobserved characteristics term for product/market jt and $\boldsymbol{\theta}_1$ is a vector of parameters. Consider the system of demand equations

$$s_{jt} = \sigma_j\left(\boldsymbol{\delta}_t; \boldsymbol{\theta}_2\right), \quad j \in \mathcal{J}, \quad t = 1, \ldots, T, \tag{1}$$

which relates the vector of observed market shares, $\mathbf{s}_t = (s_{0t}, \ldots, s_{Jt})^\intercal$, to the vector of product indexes $\boldsymbol{\delta}_t = (\delta_{0t}, \ldots, \delta_{Jt})^\intercal$, through the demand function $\boldsymbol{\sigma} = (\sigma_0, \ldots, \sigma_J)$, where $\boldsymbol{\theta}_2$ is a vector of parameters.

Normalize the index of the outside good by setting $\delta_{0t} = 0$ in each market t so that $\boldsymbol{\delta}_t \in \mathcal{D} \equiv \{\boldsymbol{\delta}_t \in \mathbb{R}^{J+1} : \delta_{0t} = 0\}$, and assume that the function $\boldsymbol{\sigma}(\cdot; \boldsymbol{\theta}_2) : \mathcal{D} \to \text{relint}(\Delta)$ is invertible. Then the inverse demand function, denoted by σ_j^{-1}, maps from market shares $\mathbf{s}_t$ to each index δ_{jt} with

$$\delta_{jt} = \sigma_j^{-1}\left(\mathbf{s}_t; \boldsymbol{\theta}_2\right), \quad j \in \mathcal{J}, \quad t = 1, \ldots, T. \tag{2}$$

In addition, assume a linear index,

$$\delta_{jt} = \mathbf{x}_{jt}\boldsymbol{\beta} - \alpha p_{jt} + \xi_{jt}, \quad j \in \mathcal{J}, \quad t = 1, \ldots, T,$$

where the vector of parameters $\boldsymbol{\beta} \in \mathbb{R}^K$ capture the consumers' taste for characteristics $\mathbf{x}_{jt}$ and the parameter $\alpha > 0$ is the consumers' marginal utility of income. Then the unobserved product characteristics terms, ξ_{jt}, can be written as a function of the data and parameters $\boldsymbol{\theta}_1 = (\alpha, \boldsymbol{\beta})$ and $\boldsymbol{\theta}_2$ to be estimated,

$$\xi_{jt} = \sigma_j^{-1}(\mathbf{s}_t; \boldsymbol{\theta}_2) + \alpha p_{jt} - \mathbf{x}_{jt}\boldsymbol{\beta}, \quad j \in \mathcal{J}, \quad t = 1, \ldots, T. \tag{3}$$

The product characteristics terms, ξ_{jt}, are the structural error terms of the model, as they are observed by consumers and firms but not by the modeller. Prices are likely to be endogenous since firms may consider both observed and unobserved product characteristics when setting prices. Market shares are endogenous by construction since they are defined by the system of Equations (1), where the demand function of each product depends on the entire vector of endogenous prices and unobserved product characteristics. Then, following Berry (1994), we can estimate the demand function $\boldsymbol{\sigma}$ based on the conditional moment restrictions $\mathbb{E}[\xi_{jt}|\mathbf{z}_t] = 0$ for all $j \in \mathcal{J}$ and $t = 1, \ldots, J$, provided that there exist appropriate instruments $\mathbf{z}_t$ for prices and market shares.

Since the seminal papers by Berry (1994) and Berry et al. (1995), the standard practice of the demand estimation literature with aggregate data has been to specify an ARUM or an RCL model and to compute the corresponding demand function, which, except for the logit and nested logit models, must then be inverted numerically during estimation.[5] In this paper, we instead directly specify closed-form inverse demand functions of the form

$$\sigma_j^{-1}(\mathbf{s}_t; \boldsymbol{\theta}_2) = \ln G_j(\mathbf{s}_t; \boldsymbol{\theta}_2) + c_t = \delta_{jt}, \quad j \in \mathcal{J}, \tag{4}$$

where the vector function $\mathbf{G} = (G_0, \ldots, G_J)$ is invertible as a function of $\mathbf{s}_t \in \mathrm{relint}(\Delta)$, and where $c_t \in \mathbb{R}$ is a market-specific constant that is determined by the normalization of the vector $\boldsymbol{\delta}_t$.

Importantly, when G_j is linear in parameters $\boldsymbol{\theta}_2$, estimation amounts to a linear regression. To see this, consider the three-level nested logit model, which partitions

[5]The logit and nested logit models have inverse demand in closed form.

the choice set into nests and further partitions nests into subnests. In this model,

$$\ln G_j\left(\mathbf{s}_t;\mu_1,\mu_2\right)=\left(1-\sum_{d=1}^{2}\mu_d\right)\ln\left(s_{jt}\right)+\mu_1\ln\left(\sum_{k\in\mathcal{G}_1(j)}s_{kt}\right)+\mu_2\ln\left(\sum_{k\in\mathcal{G}_2(j)}s_{kt}\right),$$

where the nesting parameters $\mu_1,\mu_2\geq 0$ satisfy $\sum_{d=1}^{2}\mu_d<1$ and where $\mathcal{G}_1(j)$ and $\mathcal{G}_2(j)$ are the sets of products belonging the same nest and to the same subnest as product j, respectively.[6] The three-level nested logit model corresponds to the logit when $\mu_1=0$ and $\mu_2=0$ and to the two-level nested logit when $\mu_1=0$ or $\mu_2=0$.

Assume that the outside good is in a nest by itself, such that $\ln G_0(\mathbf{s}_t;\mu_1,\mu_2)=\ln(s_0)$. Then, the three-level nested logit model boils down to the linear regression model (Verboven, 1996a)

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right)=\mathbf{x}_{jt}\boldsymbol{\beta}-\alpha p_{jt}+\mu_1\ln\left(\frac{s_{jt}}{\sum_{k\in\mathcal{G}_1(j)}s_{kt}}\right)+\mu_2\ln\left(\frac{s_{jt}}{\sum_{k\in\mathcal{G}_2(j)}s_{kt}}\right)+\xi_{jt}\quad(5)$$

for all products $j=1,\ldots,J$ in each market $t=1,\ldots,T$, which requires one instrument for price and two for the endogenous log-share terms for identification.

3 The IPDL Model

The linear structure of the nested logit regression in Equation (5) makes standard linear instrumental variable regression techniques (e.g. two-stage least squares) easily applicable and empirical identification clear. Due to its parsimony, it is also able to handle very large choice sets. However, the nested logit model imposes strong restrictions on the substitution patterns that can be accommodated. In this section, we introduce the Inverse Product Differentiation Logit (IPDL) model, which generalizes the inverse demand function of the nested logit model, while maintaining its desirable features.

[6] Indeed, setting $\gamma_1=\mu_1+\mu_2$ and $\gamma_2=\mu_1$, we recover Equation (10) of Verboven (1996a) and the model satisfies the constraint $0\leq\gamma_2\leq\gamma_1<1$ that makes it consistent with random utility maximization.

Setting Suppose that each market exhibits product segmentation along D discrete product characteristics, indexed by d. Each characteristic d defines a finite number of *groups* of products, such that each product belongs to exactly one group by each characteristic. The grouping structure is assumed to be exogenous and common across markets. For example, cars may be grouped by characteristics such as brand, size and fuel type.

Let $\boldsymbol{\theta}_2 = (\mu_1, \ldots, \mu_D)$, with $\sum_{d=1}^{D} \mu_d < 1$ and $\mu_d \geq 0$, $d = 1, \ldots, D$, and let $\mathcal{G}_d(j)$ be the set of products grouped with product j on characteristic d. The IPDL model has an inverse demand function of the form of Equation (4), where $\ln G_j$ is defined as

$$\ln G_j(\mathbf{s}_t; \boldsymbol{\theta}_2) = \left(1 - \sum_{d=1}^{D} \mu_d\right) \ln(s_{jt}) + \sum_{d=1}^{D} \mu_d \ln\left(\sum_{k \in \mathcal{G}_d(j)} s_{kt}\right). \tag{6}$$

Two products are of the same *type* if they belong to the same group for all characteristics d. We assume that the outside good is the only product of its type, i.e.

$$\ln G_0(\mathbf{s}_t; \boldsymbol{\theta}_2) = \ln(s_{0t}). \tag{7}$$

The IPDL model extends the (inverse demand of the) nested logit model by allowing arbitrary, non-hierarchical grouping structures, i.e. any partitioning of the choice set for each characteristic. This extension works in the same way as the product differentiation logit (PDL) model of Bresnahan et al. (1997). Therefore, the IPDL model has the nested logit model and the logit model as a special cases: the logit is obtained when product segmentation does not matter, and the nested logit model is obtained when the grouping structure is hierarchical. The freedom in defining grouping structures can also be used to build inverse demand models that are similar in spirit to GEV models based on grouping, which have proved useful for demand estimation purposes (Train, 2009, Chap. 4).[7]

Several remarks are in order. First, the vector function $\ln \mathbf{G}$ with elements given

[7]For example, as in Small (1987) and Grigolon (2020), it is possible to define grouping structures that describe markets where products are naturally ordered.

by Equation (6) can be shown to be invertible.[8] That is, any observed vector of market shares $\mathbf{s}$ is rationalized by a unique vector of product indexes $\boldsymbol{\delta} \in \mathcal{D}$. Then, in the IPDL model, the Independence from Irrelevant Alternatives (IIA) property holds for products of the same type, but does not in general for product of different types. In practice, this implies that there are as many cross-price elasticities per product as there are different types. Furthermore, the parametrization of the IPDL model does not depend on the number of products, but instead on the number of grouping characteristics. This implies that the IPDL model can handle markets involving very many products. Lastly, as it it the case for the nested logit model, the IPDL model can be extended to allow group parameters μ_d to be group-specific (see the empirical application of Section 4).

Microfoundation In Section 6, we show that the IPDL model is consistent with a representative consumer model with taste for variety whose utility function belongs to the class of utilities studied by Allen and Rehbeck (2019b), which may be interpreted as representing the behavior of heterogeneous, utility-maximizing consumers. Specifically, it is consistent with a representative consumer, endowed with income y, who chooses a vector $\mathbf{s}_t \in \operatorname{relint}(\Delta)$ of nonzero market shares in market t so as to maximize her utility function given by

$$\alpha y + \sum_{j \in \mathcal{J}} \delta_{jt} s_{jt} - \left(1 - \sum_{d=1}^{D} \mu_d\right) \sum_{j \in \mathcal{J}} s_{jt} \ln\left(s_{jt}\right) - \sum_{d=1}^{D} \mu_d \left[\sum_{g \in \mathcal{G}_d} s_{gt} \ln\left(s_{gt}\right)\right], \quad (8)$$

where $s_{gt} = \sum_{k \in g} s_{kt}$, and $\mathcal{G}_d$ is the set of groups for characteristic d. The second term in Equation (8) captures the net utility derived from the consumption of $\mathbf{s}_t$ in the absence of interaction among products and the remaining terms express taste for variety. Specifically, the quantity $(1 - \sum_{d=1}^{D} \mu_d)$ measures taste for variety over all products of the choice set; and each parameter μ_d measures taste for variety across groups by characteristic d: higher μ_d means that variety at the level of groups matters more, meaning that products in the same group by characteristic d are more

[8] The key assumption that ensures invertibility is that $\sum_{d=1}^{D} \mu_d < 1$. Proposition 1 below establishes invertibility for the Generalized Logit model introduced in Section 5, which has the IPDL model as a special case.

similar (see Verboven, 1996b, for a similar interpretation of the nesting parameter in the nested logit model).

Complementarity We now consider a simple example which shows how complementarity may arise in the IPDL model due to taste for variety at the group level. Suppose there are $J = 3$ products and one outside good. Products are grouped by two characteristics: the grouping is $\{1\}, \{2, 3\}$ for the first characteristic and $\{1, 2\}, \{3\}$ for the second characteristic. This grouping structure induces substitutability between products 1 and 2 as well as between products 2 and 3. However, depending on parameters, products 1 and 3 may be substitutes or complements. Complementarity occurs if and only if $(1 - \mu_1 - \mu_2)(s_1 + s_2)(s_2 + s_3) - \mu_1\mu_2 s_0 s_2 < 0$, which simplifies to $4/3 < \mu_1\mu_2/(1 - \mu_1 - \mu_2)$ for $s_0 = 1/2$ and $s_1 = s_2 = s_3 = 1/6$. In particular, products 1 and 3 are substitutes for $\mu_1 = 1/4$ and $\mu_2 = 1/3$, but are complements for $\mu_1 = 2/5$ and $\mu_2 = 1/2$. See Proposition 6 in Appendix A.3 for details.

In the supplement, we provide simulation results investigating the patterns of substitution of the IPDL model. We find that products of the same type are always substitutes, while products of different types may be substitutes or complements, and that closer products into the characteristics space used to form product types (i.e. higher values of μ_d and/or whether products belong to the same groups or not) have higher cross-price elasticities.

As shown by Cardell (1997) and further studied by Galichon (2021), the (two-level) nested logit model is an RCL model for which the dummy variables that form the nesting structure receive a random coefficient with a certain distribution. This observation motivates the question of whether an IPDL model is also equivalent to some RCL model. We can immediately rule out IPDL models exhibiting complementarity, since products can only be substitutes in the RCL model. Furthermore, there are IPDL models that violate the sign condition that holds for the higher order mixed partial derivatives of the demand function of the RCL model. So we can conclude that the IPDL model, even without complementarity, accommodates behavior that cannot be accommodated by any RCL model.

Identification and Estimation Combining Equations (6) and (7) and using that $\delta_{0t} = 0$ for all $t = 1, \ldots, T$, the IPDL model boils down to the linear regression of market shares on product characteristics and log-share terms

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \mathbf{x}_{jt}\boldsymbol{\beta} - \alpha p_{jt} + \sum_{d=1}^{D} \mu_d \ln\left(\frac{s_{jt}}{\sum_{k\in\mathcal{G}_d(j)} s_{kt}}\right) + \xi_{jt}, \tag{9}$$

for all products $j = 1, \ldots, J$ in each market $t = 1, \ldots, T$.

Equation (9) has the same form as the logit and nested logit equations, except for the log-share terms. Following the literature, we assume that product characteristics $\mathbf{x}_{jt}$ are exogenous and that prices and log-share terms are endogenous. As a consequence, the IPDL model reduces to a linear IV regression, where identification requires at least one instrument for price and one for each of the log-share terms. As it is well known, instruments for prices include cost shifters and markup shifters (see e.g. Berry and Haile, 2014, 2016). In particular, the second set of instruments include the BLP instruments (Berry et al., 1995; Gandhi and Houde, 2020). Following Verboven (1996a) and Bresnahan et al. (1997), for the IPDL model, these instruments include, for each grouping characteristic, the sums of characteristics of other products of the same group or the corresponding differences in those characteristics. The same instruments can also be computed for products of the same type.

Furthermore, identification of group parameters μ_d requires exogenous variation in the relative share $s_{jt}/\sum_{k\in\mathcal{G}_d(j)} s_{kt}$. Intuitively, since they drive substitution patterns among products, identification requires instruments that provide exogenous variation in the choice set, including changes in prices. Both cost shifters and markup shifters are therefore good candidates for instruments for log-share terms.

Comparison to Existing Models We now consider three experiments that compare the IPDL model to three existing models, the three-level nested logit model, the PDL model of Bresnahan et al. (1997), and the RCL model. For each experiment, one dataset consists of $T = 200$ markets with $J = 45$ products, where markets exhibit product segmentation along two characteristics that form four types of products.[9] In each experiment, we generate a fully structural model of demand and supply, where the supply side is a static price competition model with multi-product firms. This allows us to compare models both in terms of estimated elasticities and markups. See Appendix B for details.

The IPDL model and the (three-level) nested logit model coincide when the former has a hierarchical grouping structure; both are estimated by linear instrumental variable regression. The first experiment assesses the extent to which imposing a hierarchical grouping structure biases the estimated substitution patterns, when the true grouping structure is non-hierarchical. We first simulate two IPDL models with different grouping parameters such that complementarity occurs in the second model but not in the first. Then, we estimate the two possible nested logit models. The results presented in Table 1 indicate that the nested logit model leads to biased estimates of the price elasticities and markups, especially when some products are complements. It also shows that the nested logit model shrinks negative cross-price elasticities towards zero and that the hierarchy of nests substantially affects the estimated substitution patterns.

The PDL model avoids the hierarchical grouping structure of the nested logit model in a way that is similar to the IPDL model, but requires the BLP method for estimation. The second experiment assesses the ability of the IPDL model to fit the elasticities generated by the PDL model and the implied markups. We first simulate two PDL models with different grouping parameters ρ_1 and ρ_2 that control the substitution between products. Then, we estimate the corresponding IPDL models. The results presented in Table 2 suggest that the IPDL model is quite well able to fit the elasticities of the PDL model and the implied markups.

[9] This approximately corresponds to the amount of data we use in the empirical application.

Table 1: IPDL MODEL VS. THREE-LEVEL NESTED LOGIT MODEL

	IPDL model with $\mu_1 = 0.1$ and $\mu_2 = 0.3$						IPDL model with $\mu_1 = 0.2$ and $\mu_2 = 0.7$					
	Price Elasticities					Markup	Price Elasticities					Markup
	Own	Type 1	Type 2	Type 3	Type 4		Own	Type 1	Type 2	Type 3	Type 4	
True IPDL												
Type 1	-2.066	0.051	0.023	0.042	0.012	0.565	-6.582	0.339	0.079	0.203	-0.056	0.278
Type 2	-2.062	0.022	0.057	0.010	0.047	0.559	-6.572	0.083	0.383	-0.102	0.250	0.254
Type 3	-2.061	0.040	0.010	0.061	0.029	0.557	-6.505	0.201	0.112	0.452	0.172	0.264
Type 4	-2.072	0.011	0.045	0.028	0.065	0.564	-6.474	-0.055	0.192	0.147	0.478	0.298
Estimated NL: First Hierarchical Structure												
Type 1	-2.545	0.046	0.040	0.021	0.023	0.447	-7.826	0.467	0.034	0.019	0.019	0.203
Type 2	-2.548	0.041	0.046	0.021	0.023	0.443	-7.680	0.030	0.654	0.019	0.019	0.195
Type 3	-2.539	0.021	0.020	0.058	0.055	0.443	-7.438	0.014	0.017	0.899	0.044	0.200
Type 4	-2.553	0.021	0.020	0.050	0.062	0.445	-7.610	0.014	0.017	0.043	0.722	0.226
Estimated NL: Second Hierarchical Structure												
Type 1	-2.519	0.062	0.019	0.044	0.021	0.459	-7.708	0.412	0.008	0.214	0.010	0.227
Type 2	-2.508	0.019	0.075	0.020	0.048	0.455	-7.653	0.008	0.508	0.009	0.252	0.208
Type 3	-2.501	0.041	0.019	0.085	0.021	0.453	-7.510	0.188	0.008	0.653	0.010	0.212
Type 4	-2.523	0.019	0.045	0.020	0.082	0.459	-7.558	0.008	0.190	0.009	0.600	0.242

Notes: The top panels give the true demand elasticities; and the middle and bottom panels give the estimated demand elasticities. In each panel, entries $j1$ involve averages of own-price elasticities of demand of products of type j; and entries $i, j+1$ involve averages of cross elasticities of products of type i with respect to the price of products of type j. Each average elasticity represents the average elasticity across markets and products of the given types.

Table 2: IPDL MODEL VS. PDL MODEL

	True PDL						Estimated IPDL					
	Price Elasticities					Markup	Price Elasticities					Markup
	Own	Type 1	Type 2	Type 3	Type 4		Own	Type 1	Type 2	Type 3	Type 4	
Dataset is PDL model with $\rho_1 = \rho_2 = 0.5$												
Type 1	-4.089	0.115	0.079	0.068	0.032	0.297	-4.436	0.123	0.085	0.073	0.027	0.272
Type 2	-4.069	0.070	0.143	0.029	0.096	0.293	-4.406	0.076	0.162	0.020	0.100	0.269
Type 3	-4.082	0.066	0.032	0.124	0.086	0.293	-4.425	0.071	0.022	0.137	0.091	0.269
Type 4	-4.084	0.028	0.096	0.079	0.145	0.297	-4.423	0.024	0.010	0.084	0.163	0.274
Dataset is PDL model with $\rho_1 = 0.9$ and $\rho_2 = 0.5$												
Type 1	-3.523	0.085	0.087	0.037	0.036	0.341	-3.709	0.079	0.051	0.072	0.042	0.320
Type 2	-3.205	0.076	0.090	0.032	0.043	0.369	-3.664	0.044	0.105	0.037	0.094	0.320
Type 3	-3.662	0.036	0.037	0.097	0.096	0.326	-3.711	0.072	0.042	0.081	0.051	0.317
Type 4	-3.340	0.031	0.043	0.085	0.098	0.359	-3.682	0.037	0.096	0.045	0.103	0.322

Notes: The left panels give the true demand elasticities; and the right panels give the estimated demand elasticities. In each panel, entries $j1$ involve averages of own-price elasticities of demand of products of type j; and entries $i, j+1$ involve averages of cross elasticities of products of type i with respect to the price of products of type j. Each average elasticity represents the average elasticity across markets and products of the given types.

Lastly, we compare our approach to the BLP method. In the third experiment, the demand side is an RCL model with two normally dependent random coefficients on dummies for groups. Results are presented in Table 3. They show that the IPDL model fits well the own-price elasticities and the implied markups of the RCL model, and that it yields estimated cross-price elasticites that are reasonably close to the true ones (even though not as close as in the second experiment). The IPDL model is thus able, at least in this example, to match the rich substitution patterns of the RCL model while entailing low computational cost.[10]

Table 3: IPDL MODEL VS. RCL MODEL

	True PDL						Estimated IPDL					
	Price Elasticities					Markup	Price Elasticities					Markup
	Own	Type 1	Type 2	Type 3	Type 4		Own	Type 1	Type 2	Type 3	Type 4	
Type 1	-4.109	0.129	0.083	0.147	0.101	0.300	-4.220	0.161	0.076	0.104	0.019	0.294
Type 2	-5.116	0.081	0.105	0.102	0.142	0.242	-5.274	0.078	0.121	0.021	0.064	0.236
Type 3	-5.063	0.061	0.043	0.236	0.181	0.245	-5.149	0.250	0.049	0.327	0.126	0.242
Type 4	-6.026	0.036	0.052	0.156	0.242	0.204	-6.209	0.053	0.177	0.146	0.269	0.199

Notes: The left panel gives the true demand elasticities; and the right panel gives the estimated demand elasticities. In each panel, entries $j1$ involve averages of own-price elasticities of demand of products of type j; and entries $i, j+1$ involve averages of cross elasticities of products of type i with respect to the price of products of type j. Each average elasticity represents the average elasticity across markets and products of the given types.

Recall that the IPDL model is capable of accommodating complementarity (see more in the supplement). The reader may wonder whether the estimates from the IPDL model may indicate complementarity even when the data generating process does not exhibit complementarity. We verify that this is not the case in the simulation experiments, where we estimate the IPDL model on data generated by the PDL and the RCL models, which rule out complementarity. We find that the IPDL model does not falsely indicate complementarity in any of our experiments. This observation suggests that when we find complementarity between cereals for kids and cereals for adults in the empirical application, this is not a model artifact.

4 Empirical Application

In this section, we use the IPDL model to study the type of relationships between products of different market segments, i.e. whether they are substitutes, independent or complements. As an illustration, we consider the market for ready-to-eat

[10]We have also estimated the two possible nested logit models. The parameter estimates do not satisfy the restrictions that make them consistent with random utility maximization, which indicates that there is no nested logit model that can rationalize the simulated RCL model.

(RTE) cereals in Chicago, which has been extensively studied since Nevo (2001), and investigate the relationship between cereals for kids and cereals for adults.

4.1 Data

Data Sources We use data from the Dominick's Dataset that are made publicly available by the James M. Kilts Center, University of Chicago Booth School of Business. This is weekly store-level scanner data, comprising information on 30 categories of packaged products at the UPC level for all Dominick's Finer Foods chain stores in the Chicago metropolitan area over the period 1989-1997. The data are supplemented by store-specific information, including average household size and daily store traffic.

For our analysis, we consider the RTE cereals category during the period 1991 –1996. We aggregate data from 62 Dominick's stores into 3 pricing zones defined by Dominick's and we aggregate UPCs into products, where a product is a cereal (e.g. Special K) associated to its brand (e.g. Kellogg's).[11] We define a market as a zone-month pair. We select 46 products from 6 national manufacturers (General Mills, Kellogg's, Nabisco, Post, Quaker and Ralston), so that they represent around 75% of each manufacturer total sales on the period.[12] We define three market segments, namely Adults, Kids and All-family, according to the classification provided by the website cerealfacts.org.

Prices are retail prices calculated as the volume-weighted average price per ounce of the UPCs that form the product, deflated by the monthly Consumer Price Index for All Urban Consumers in the Chicago-Naperville-Elgin area from the U.S. Bureau of Labor Statistics. We compute the potential market size by multiplying the total number of persons in a market by the monthly per capita consumption of cereals.[13] We compute the total volume of a product sold in a market, which we di-

[11] Only package sizes between 10 and 32 ounces are included.

[12] The 46 products account for around 60% of the national market (see e.g. Corts, 1996).

[13] For each store in a month, the total number of persons is computed as the weekly average number of households who visited that store in that given month, times the average household size. The weekly average number of households is computed using information on the daily traffic store and assuming that consumers visit stores twice a week. The total number of persons in a market is then obtained by summing over stores of a given zone. The monthly per capita consumption of

vide by the potential market size to obtain the product's market share. The market share of the outside good is then the difference between one and the sum of the 46 products' market shares.

We supplement Dominick's Dataset with information on the nutrient content (fiber, sugar and calories) of the cereals from the USDA Nutrient Database for Standard Reference (release SR11, year 1996), and on the type of grains (rice, wheat, corn and oats) using manufacturers' websites and different websites collecting nutritional information. We also use monthly input prices from the websites indexmundi.com (corn, rice, sugar and wheat) and macrotrends.net (oats) to construct cost-based instruments.

Descriptive Statistics Table 4 presents descriptive statistics on market shares and retail prices of cereals, by firm and market segment.

Table 4: SHARES AND PRICES BY FIRM AND MARKET SEGMENT

	All-family		Adults		Kids		Total	
	shares	prices	shares	prices	shares	prices	shares	prices
General Mills	3.52	20.03	2.17	20.14	3.35	20.93	9.04	20.39
Kellogg's	1.37	17.06	6.36	16.79	6.47	18.33	14.2	17.52
Nabisco			0.91	17.70			0.91	17.70
Post	0.90	16.28	2.53	15.90	1.01	21.84	4.44	17.63
Quaker	2.16	15.74	1.27	14.34			3.43	15.22
Ralston	0.79	20.79			0.20	24.75	0.99	21.57
Total	8.74	18.20	13.24	17.08	11.03	19.55	33.01	18.21
Outside good							66.99	

Notes: The number of observations is 9,246. Shares and prices refer to average (across markets) market shares in percent and retail prices (in cents) per ounce, respectively

Kellogg's and General Mills are the largest two firms, and are active in all market segments. Market segments have about equal market shares and cereals for kids have higher prices on average.

cereals is computed using the information from the USDA's Economic Research Service that per capita US consumption of cereals was equal to 13.4 pounds in 1991, 13.9 in 1992, 14.6 in 1993, 14.8 in 1994, 14.6 in 1995 and 14.3 in 1996.

4.2 Specification and Identification

Specification For the purpose of our empirical analysis, we specify an IPDL model with three grouping characteristics: i) whether or not cereals are adults-friendly (*A* for adults-friendly, *NA* for non-adults-friendly), ii) whether or not they are kids-friendly (*K* for kids-friendly, *NK* for non-kids-friendly), and iii) which brand the cereals belong to (G for General Mills, K for Kellogg's, N for Nabisco, P for Post, Q for Quaker, and R for Ralston). All-family cereals are both kids-friendly and adults-friendly, cereals for adults are adults-friendly but not kids-friendly, and cereals for kids are kids-friendly but not adults-friendly. Cereals for kids contain more sugar and less fiber than cereals for adults and for all-family; Nabisco offers cereals with less sugar and less calories, while Quaker and Ralston offer cereals with less fiber and more calories. Therefore, the grouping characteristics also proxy, at least partially, for the nutrient content of the cereals.

The corresponding IPDL model is estimated using the linear IV regression

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \beta_0 + \mathbf{x}_j\boldsymbol{\beta} - \alpha p_{jt} + \mu_{\mathcal{G}_1(j)} \ln\left(\frac{s_{jt}}{\sum_{k\in\mathcal{G}_1(j)} s_{kt}}\right) + \mu_{\mathcal{G}_2(j)} \ln\left(\frac{s_{jt}}{\sum_{k\in\mathcal{G}_2(j)} s_{kt}}\right) + \mu_{\mathcal{G}_3(j)} \ln\left(\frac{s_{jt}}{\sum_{k\in\mathcal{G}_3(j)} s_{kt}}\right) + \xi_{jt}, \tag{10}$$

where $\mathcal{G}_1(j) \in \{A, NA\}$, $\mathcal{G}_2(j) \in \{K, NK\}$ and $\mathcal{G}_3(j) \in \{G, K, N, P, Q, R\}$ and where $\mathbf{x}_j$ are market-invariant product characteristics (fiber, sugar, calories, corn, oats, rice, wheat). As in Bresnahan et al. (1997), the error term is specified as $\xi_{jt} = \xi_s + \xi_b + \xi_m + \xi_z + u_{jt}$, where ξ_s, ξ_b, ξ_m and ξ_z are fixed effects for segments, brands, months and zones, respectively and where u_{jt} is the remaining structural error.

The advantages provided by the three grouping characteristics are therefore parametrized by the fixed effects ξ_s and ξ_b, which measure the extent to which belonging to a group shifts the demand for the cereal, as well as the parameters for groups $\mu_{\mathcal{G}_1(j)}$, $\mu_{\mathcal{G}_2(j)}$, and $\mu_{\mathcal{G}_3(j)}$, which measure the extent to which cereals within a group are protected from substitution from cereals in other groups by each grouping characteristic.

We estimate two specifications of the model (10): first, a *restricted* specification where parameters for groups are equal across groups for a given grouping characteristic; second, a *flexible* specification in which they are allowed to vary across groups. For the sake of parsimony, brands are divided into three groups according to their popularity (measured in terms of market shares): General Mills and Kellogg's, Post and Nabisco, Quaker and Ralston. We use the two-step efficient generalized method of moments (GMM) estimator, with instruments described below.

Identification To identify the substitution patterns, we rely on three sets of instruments. The first set comprises the cost shifters. We use input prices (sugar, corn, oats, rice and wheat) multiplied by the corresponding characteristics, which we interact with firm's fixed effects to generate instruments that vary by cereals, by firms and across time.

The second set consists of BLP instruments. We construct several sums of the differences in sugar: sums over competing cereals belonging to the same group for each grouping characteristic, the same sums over cereals of the same firm and over cereals of rival firms, respectively. For the flexible specification, we interact these instruments with the corresponding groups' fixed effects.

Lastly, in line with Miller and Weinberg (2017), we use Post's acquisition of the Nabisco cereal line that occurred in January 1993 as a markup shifter. To examine the effects of the Post-Nabisco merger on prices and market shares, we consider the following regressions

$$\ln(y_{jt}) = a_i + b_i \mathrm{PostMerger}_t + c_i(\mathbf{w}_{jt}) + \phi_j + \phi_z + \phi_m + \varepsilon_{ijt},$$

where $y_{jt} \in \{p_{jt}, s_{jt}\}$ for a product j in market segment i, where $\mathrm{PostMerger}_t$ is a post-merger indicator, c_i controls for firm-specific cost changes through time, $\mathbf{w}_{jt}$ are cost shifters, and ϕ_j, ϕ_z, ϕ_m are fixed effects for products, zones and months, respectively (see, e.g. Björnerstedt and Verboven, 2016; Miller and Weinberg, 2017).

Table 5 presents the results and shows that the Post-Nabisco merger affected market segments differently. This is not surprising since the merger directly involved the segment adults, in which both Post and Nabisco were present before the merger, but not (at least not directly) the segments kids and all-family in which only Nabisco was active. The result that cereals for kids experienced a price decrease and a demand increase after the merger is less obvious. However, these results show that the merger can be used as an instrument for prices and log-shares. In practice, we include a post-merger indicator which we interact with fixed effects for firms and for segments, respectively, and we interact the cost-based instruments with the post-merger indicator.

Table 5: PRICE AND MARKET SHARE EFFECTS OF THE MERGER

	$\ln(p_{jt})$		$\ln(s_{jt})$	
Fixed Effects for Segments (a_i)				
Adults	-1.514	(0.0443)	-4.666	(0.234)
Kids	-1.911	(0.0657)	-4.820	(0.322)
All-family	-2.037	(0.0735)	-5.653	(0.225)
Interaction Segment – Post-Merger Indicator (b_i)				
Adults × PostMerger	0.0138	(0.00398)	-0.117	(0.0309)
Kids × PostMerger	-0.0285	(0.00659)	0.0877	(0.0381)
All-family × PostMerger	0.0108	(0.00746)	-0.00110	(0.0340)
RMSE	0.127		0.468	

Notes: The number of observations is 9,246. Standard errors are clustered at the market (zones-month) level and shown in parentheses. Fixed effects for products, months, and zones, as well as controls for costs are included.

A potential problem is weak identification, which occurs when instruments are only weakly correlated with the endogenous variables. In both specifications, the Sanderson and Windmeijer (2016)'s F-statistics to test whether each endogenous variable is weakly identified are far above 10, the rule-of-thumb usually used for linear IV regressions, thereby suggesting that instruments are not weak.

4.3 Results

Demand Parameters Table 6 presents the parameter estimates from the IPDL model. Columns (1) and (2) provides the results for the restricted and the flexible specifications, respectively. As expected, the estimated parameter on the negative of price (α) is significantly positive for both specifications. The estimated parameters for groups have magnitude and sign that satisfy the assumptions of the IPDL model.[14]

The flexible specification nests the restricted specification; the restriction can then be tested by a simple Wald test. The test rejects the restricted specification at any conventional level of significance, indicating that the grouping parameters are statistically different from each other in a given dimension.

In both specifications, the estimated fixed effects suggest that the brand reputation of the cereals confers a significant advantage to products from General Mills and Kellogg's and that cereals for family have a significant advantage.

Furthermore, looking at the estimated grouping parameters, we find that the market segments confer more protection from substitution than brand reputation does (cereals of the same market segment are more protected from cereals from different market segments than cereals of the same brand are from cereals of different brands). Overall, this implies that cereals of the same type are closer substitutes.

Price Elasticities Table 7 presents the estimated own- and cross-price elasticities of demand for both specifications, averaged across markets (month-zone pairs) and product types (all-family, adults, kids). We obtain own-price elasticities in line with the literature (see e.g. Nevo, 2001).

Both specifications give qualitatively similar results regarding the relationships between cereals of different market segments: all-family cereals are more substitutable with cereals for adults than with cereals for kids; and cereals for kids and for adults are complements.

To verify that complementarity is not a model artifact, we compute the price elasticities using different values of the grouping parameters from the estimated values. We find that parameter values exist for the model structure does not impose complementarity.

We also compute markups assuming a static oligopolistic price competition between the firms. We find, for General Mills, Kellogg's, Post and Nabisco in 1994, that the average combined retailer-manufacturer markup is equal to 42% and 32% for the restricted and the flexible specifications, respectively. These results are in

[14] For the restricted specification, $\mu_1 \geq 0$, $\mu_2 \geq 0$, $\mu_3 \geq 0$ and $1 - \mu_1 - \mu_2 - \mu_3 > 0$; similarly, for the flexible specification. No constraints were imposed on the parameters during the estimation.

Table 6: PARAMETER ESTIMATES OF DEMAND

	(1) Restricted		(2) Flexible	
Constant (β_0)	-0.776	(0.0398)	-0.776	(0.0413)
Price ($-\alpha$)	-1.047	(0.131)	-1.464	(0.173)
Fixed Effects for Segments				
Kids	-0.503	(0.0247)	-0.457	(0.0211)
All-family	0.0486	(0.00621)	0.165	(0.0147)
Fixed Effects for Brands				
Kellogg's	0.00754	(0.00380)	0.00939	(0.00449)
Nabisco	-0.159	(0.0276)	-0.241	(0.0339)
Post	-0.0880	(0.0135)	-0.163	(0.0213)
Quaker	-0.101	(0.0156)	-0.118	(0.0158)
Ralston	-0.157	(0.0280)	-0.183	(0.0290)
Grouping parameters for adults-friendly (μ_1)				
A	0.806	(0.0234)	0.774	(0.0201)
NA	same		0.742	(0.0240)
Grouping parameters for kids-friendly (μ_2)				
K	0.106	(0.0130)	0.139	(0.0122)
NK	same		0.101	(0.0114)
Grouping parameters for brands (μ_3)				
G – K	0.0511	(0.0104)	0.0663	(0.0114)
N – P	same		0.0446	(0.0098)
Q – R	same		0.0713	(0.0123)

Notes: The number of observations is 9,246. Standard errors are clustered at the market (zones-month) level and shown in parentheses. Fixed effects for products, months, and zones, as well as characteristics (fiber, sugar, calories, corn, oats, rice and wheat) are included.

line with Corts (1996) who finds an average markup of 37% using accounting data.

5 The Generalized Logit Model

Having shown that the IPDL model is econometrically convenient, we now show that it is consistent with utility maximization. For this purpose, we introduce the Generalized Logit (GL) model, which has the IPDL model as a special case. Then in Section 6, we establish the GL model, and thus the IPDL model, as a utility-maximizing model and relate it to the representative consumer model and to the additive random utility model. Proofs for this section are provided in Appendix A.4. To ease exposition, we omit notation for parameters $\boldsymbol{\theta}_2$ and markets t.

Table 7: SUBSTITUTION PATTERNS

		Cross elasticities		
	Own elasticity	All-family	Adults	Kids
Restricted Model				
All-family	-4.3354	0.2595	0.0474	0.0089
	[-4.4437 ; -4.2271]	[0.2415 ; 0.2775]	[0.0399 ; 0.0550]	[0.0064 ; 0.0113]
Adults	-4.3412	0.0466	0.1789	-0.0012
	[-4.4512 ; -4.2312]	[0.0391 ; 0.0541]	[0.1707 ; 0.1870]	[-0.0016 ; -0.0008]
Kids	-4.9885	0.0043	-0.0005	0.3180
	[-5.1138 ; -4.8632]	[0.0029 ; 0.0057]	[-0.0008 ; -0.0002]	[0.3079 ; 0.3282]
Flexible Model				
All-family	-8.8607	0.6009	0.0665	0.0169
	[-9.4625 ; -8.2590]	[0.5417 ; 0.6602]	[0.0557 ; 0.0773]	[0.0127 ; 0.0211]
Adults	-3.5719	0.0567	0.1305	-0.0014
	[-3.6731 ; -3.4707]	[0.0471 ; 0.0663]	[0.1236 ; 0.1374]	[-0.0019 ; -0.0010]
Kids	-4.2946	0.0075	-0.0005	0.2660
	[-4.4254 ; -4.1638]	[0.0054 ; 0.0095]	[-0.0009 ; -0.0001]	[0.2542 ; 0.2777]

Notes: This table gives the estimated own- and cross-price elasticities of demands, averaged across markets (month-zone) and product types (all-family, adults, kids), using a parametric bootstrap. We draw repeatedly from the estimated joint distribution of parameters. For each draw, we compute the average elasticities, thus generating a bootstrap distribution; 1,500 draws are taken. The middle number is the average over draws; lower numbers in brackets are the bounds of the 95% confidence interval.

Definition 1. A generalized logit (GL) inverse demand function is a function $\ln \mathbf{G}$, where $\mathbf{G} : (0, \infty)^{J+1} \to (0, \infty)^{J+1}$ is linearly homogeneous and where the Jacobian $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is positive definite and symmetric.

In a GL model, the vector of product indexes is given up to an additive market-specific constant $c \in \mathbb{R}$ by the GL inverse demand function, i.e.

$$\ln G_j(\mathbf{s}) = \delta_j - c, \quad j \in \mathcal{J}. \tag{11}$$

This definition implies that the IPDL model is a GL model. In the supplement, we provide a range of general methods for building other GL inverse demand functions along with illustrative examples. As stated in the following proposition, a GL inverse demand function is injective and hence invertible on its range.

Proposition 1. A GL inverse demand function is injective on $\mathrm{relint}(\Delta)$. Furthermore, if $|\ln \mathbf{G}(\mathbf{s})| \to \infty$ whenever $\mathbf{s} \to \mathbf{s}^0$, where $\mathbf{s}^0$ is on the boundary of Δ, then the range of $\ln \mathbf{G}$ is $\mathbb{R}^{J+1}$.

Invertibility of the inverse demand function is equivalent to invertibility of the demand function itself. The demand function has domain equal to the range of the inverse demand function. The proposition provides a simple condition that ensures this range is equal to all of $\mathbb{R}^J$; this condition is satisfied by the IPDL model.

Consider any vector of market shares $\mathbf{s} \in \operatorname{relint}(\Delta)$. Then, holding $\delta_0 = 0$, the injectivity of the GL inverse demand function ensures that there exists a unique vector of indexes $\boldsymbol{\delta} \in \mathcal{D}$ that rationalizes demand, i.e. $\mathbf{s} = \boldsymbol{\sigma}(\boldsymbol{\delta})$.

Let $\mathbf{H} = \mathbf{G}^{-1}$ denote the inverse of $\mathbf{G}$. Inverting Equation (11) and using that the demand vector sums to one together with the linear homogeneity of $\mathbf{G}$ leads to the demand function. Then the consumer surplus function can be computed using Roy's identity.

Proposition 2. Let $\ln \mathbf{G}$ be a GL inverse demand function. Then, the corresponding demand function is

$$s_j = \sigma_j(\boldsymbol{\delta}) = \frac{H_j\left(e^{\boldsymbol{\delta}}\right)}{\sum_{k \in \mathcal{J}} H_k\left(e^{\boldsymbol{\delta}}\right)}, \quad j \in \mathcal{J}, \tag{12}$$

and the consumer surplus is given, up to an additive constant, by the convex function

$$CS(\boldsymbol{\delta}) = \ln\left(\sum_{k \in \mathcal{J}} H_k\left(e^{\boldsymbol{\delta}}\right)\right).$$

Expression (12) extends the logit demand function in a non-trivial way through the presence of the function $\mathbf{H}$. The consumer surplus is simply the logarithm of the denominator of the demand in Equation (12), just as for the logit model.[15]

Furthermore, using that the demand vector sums to one, we obtain that the market-specific constant in (11) is equal to the consumer surplus $c = CS(\boldsymbol{\delta})$. Thus, differentiating Equation (11) with respect to $\boldsymbol{\delta}$ and rearranging terms leads to the matrix of demand derivatives $\partial\sigma_j(\boldsymbol{\delta})/\partial\delta_i$.

Proposition 3. Let $\ln \mathbf{G}$ be a GL inverse demand function and let $\mathbf{s} = \boldsymbol{\sigma}(\boldsymbol{\delta})$. Then the matrix of demand derivatives is $\mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta}) = \left[\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})\right]^{-1} - \mathbf{s}\mathbf{s}^{\intercal}$.

[15] Nocke and Schutz (2018) employ a similar form, but impose $H_j(e^{\boldsymbol{\delta}}) = H_j(e^{\delta_j})$, which implies the restrictive IIA property. At the same time, in contrast to the present paper, they do not assume unit-demand.

In the absence of income effects, the matrix of demand derivatives is the Slutsky matrix. It is symmetric and positive semi-definite, which implies that GL demand functions are non-decreasing in their own index δ_j, $\partial\sigma_j(\boldsymbol{\delta})/\partial\delta_j \geq 0$.

The GL model accommodates substitution patterns that go beyond those of standard ARUM. In particular, it allows for complementarity in demand: this is for example the case of the IPDL model. Our invertibility result in Proposition 1 therefore extends Berry (1994)'s invertibility result, which restricts the products to be strict substitutes. Proposition 1 also supplements Berry et al. (2013), who show invertibility for demand functions that satisfy their "connected substitutes" conditions, which, in turn, rules out complementarity when the demand is unitary.[16]

6 Relationships between Models

This section puts the GL and the IPDL models into perspective by showing how they relate to the representative consumer (RC) model and to the additive random utility model (ARUM).[17]

6.1 Representative Consumer Model

Consider a representative consumer facing the choice set of differentiated products, $\mathcal{J}$, and a homogeneous numéraire good, with demands for the differentiated products summing to one. Let p_j and v_j be the price and the quality of product $j \in \mathcal{J}$, respectively. The price of the numéraire good is normalized to 1 and the representative consumer's income y is sufficiently high ($y > \max_{j\in\mathcal{J}} p_j$) to guarantee that consumption of the numéraire good is positive.

[16] The connected substitutes structure requires two conditions: (i) products are weak gross substitutes, i.e. everything else equal, an increase in δ_i weakly decreases demand σ_j for all other products; and (ii) the "connected strict substitution" condition holds, i.e. there is sufficient strict substitution between products to treat them in one demand system. In contrast to ours, Berry et al. (2013)'s result does not require that the demand function $\boldsymbol{\sigma}$ is differentiable. Demand systems with some form of complementarity can be covered by Berry et al. (2013)'s result in cases where a suitable transformation of demand can be found such that the transformed demand satisfies their conditions. Our result allows complementarity without requiring such a transformation to be found.

[17] See Anderson et al. (1992) of a comprehensive treatment of the ARUM and RC model

In this subsection, we show that the GL inverse demand function $\ln \mathbf{G}$ is consistent with a representative consumer who chooses a vector $\mathbf{s} \in \Delta$ of market shares of the differentiated products and a quantity $z \geq 0$ of the numéraire good, so as to maximize her direct utility function

$$\alpha z + \sum_{j \in \mathcal{J}} v_j s_j - \sum_{j \in \mathcal{J}} s_j \ln G_j(\mathbf{s}) \tag{13}$$

subject to the budget constraint and the constraint that the demand vector sums to one,

$$\sum_{j \in \mathcal{J}} p_j s_j + z \leq y \quad \text{and} \quad \sum_{j \in \mathcal{J}} s_j = 1, \tag{14}$$

where $\alpha > 0$ is the marginal utility of income. The first two terms of the direct utility (13) describe the utility that the representative consumer derives from the consumption $(\mathbf{s}, z)$ of the differentiated products and the numéraire in the absence of interaction among them. The third term is a strictly concave function of $\mathbf{s}$ that expresses her taste for variety (see Lemma 5 in Appendix A.5.1).

Let $\delta_j = v_j - \alpha p_j$ be the net utility that the consumer derives from consuming one unit of product $j \in \mathcal{J}$. The utility maximization program (13) – (14) leads to first-order conditions for interior solution, which have the form of Equation (11) defining the GL inverse demand function. We state this observation as a proposition and relegate its proof to Appendix A.5.1.

Proposition 4. The GL model (11) is consistent with a representative consumer who maximizes utility (13) subject to constraints (14).

Anderson et al. (1988) and Verboven (1996b) show that the logit and the nested logit models are consistent with a utility-maximizing representative consumer endowed with a direct utility function of the form of Equation (13). Proposition 4 extends these results to the GL model.

Furthermore, as shown by Allen and Rehbeck (2019b), utility (13) can be obtained, by aggregating across heterogeneous and utility-maximizing consumers, from the class of latent utility models with additively separable unobservable heterogeneity called perturbed utility (PU).[18] This implies that any GL model embodies consumer heterogeneity and can be rationalized by a PU model. However, the converse does not hold. For example, when taste for variety is modelled by the strictly concave function $\sum_{j\in\mathcal{J}}\ln(s_j)$, the corresponding candidate GL inverse demand function is $\ln G_j(\mathbf{s}) = \frac{1}{s_j}\ln(s_j)$, which is not linearly homogeneous.

6.2 Additive Random Utility Model

We now turn to the additive random utility model. A population of consumers face the choice set of differentiated products, $\mathcal{J}$, and associate a deterministic utility component $\delta_j = v_j - \alpha p_j$ to each product $j \in \mathcal{J}$. Each individual consumer chooses the product that maximizes her indirect utility given by

$$u_j = \delta_j + \varepsilon_j, \quad j \in \mathcal{J}, \tag{15}$$

where the vector of random utility components $\boldsymbol{\varepsilon} = (\varepsilon_0, \ldots, \varepsilon_J)$ follows a joint distribution with finite means that is absolutely continuous, fully supported on $\mathbb{R}^{J+1}$ and independent of $\boldsymbol{\delta}$. These assumptions are standard in the discrete choice literature. They imply that utility ties occur with probability 0, that the choice probabilities are all everywhere positive, and that random coefficients are ruled out. Specific distributional assumptions for $\boldsymbol{\varepsilon}$ lead to specific models such as the logit, the nested logit or the probit models.[19]

The probability that a consumer chooses product j is

$$P_j(\boldsymbol{\delta}) = \Pr(u_j \geq u_i,\ \forall i \neq j), \quad j \in \mathcal{J}.$$

[18] See Hofbauer and Sandholm (2002), McFadden and Fosgerau (2012) and Fudenberg et al. (2015) for more details on PU models, which have been used to model optimization with effort (Mattsson and Weibull, 2002), stochastic choices (Swait and Marley, 2013; Fudenberg et al., 2015), and rational inattention (Matejka and McKay, 2015; Fosgerau et al., 2020). Allen and Rehbeck (2019a) show that some PU models allow for complementarity.

[19] Note that income does not enter utility (15), which means that there is no income effect. This is equivalent to the case in which income enters linearly. The deterministic utilities, δ_j, are common across all consumers, which rules out heterogeneity in preferences apart from the random utility components, ϵ_j.

Let $\overline{CS} : \mathbb{R}^{J+1} \to \mathbb{R}$ be the consumer surplus, i.e. the expected maximum utility

$$\overline{CS}(\boldsymbol{\delta}) = \mathbb{E}\left(\max_{j \in \mathcal{J}} u_j\right).$$

By the Williams-Daly-Zachary theorem (McFadden, 1981), the conditional choice probabilities are equal to the derivatives of the consumer surplus, i.e. $P_j(\boldsymbol{\delta}) = \partial \overline{CS}(\boldsymbol{\delta}) / \partial \delta_j$. Define a function $\overline{\mathbf{H}} = \left(\overline{H}_0, \ldots, \overline{H}_J\right)$, with $\overline{H}_j : (0, \infty)^{J+1} \to (0, \infty)$ as the derivative of the exponentiated surplus with respect to its jth component, i.e.

$$\overline{H}_j\left(e^{\boldsymbol{\delta}}\right) = \frac{\partial e^{\overline{CS}(\boldsymbol{\delta})}}{\partial \delta_j} = P_j(\boldsymbol{\delta})\, e^{\overline{CS}(\boldsymbol{\delta})}, \quad j \in \mathcal{J}.$$

Summing over $j \in \mathcal{J}$ and using that probabilities sum to one, we can write the choice probabilities implied by the ARUM and the corresponding consumer surplus in terms of $\overline{\mathbf{H}}$ as

$$P_j(\boldsymbol{\delta}) = \frac{\overline{H}_j\left(e^{\boldsymbol{\delta}}\right)}{\sum_{k \in \mathcal{J}} \overline{H}_k\left(e^{\boldsymbol{\delta}}\right)}, \quad j \in \mathcal{J}, \tag{16}$$

and

$$\overline{CS}(\boldsymbol{\delta}) = \ln\left(\sum_{k \in \mathcal{J}} \overline{H}_k\left(e^{\boldsymbol{\delta}}\right)\right).$$

Lemma 7 in Appendix A.5.2 shows that $\overline{\mathbf{H}}$ is invertible, with inverse $\overline{\mathbf{G}} = \overline{\mathbf{H}}^{-1}$, and that $\ln \overline{\mathbf{G}}$ is a GL inverse demand function. Then we can invert Equations (16) to obtain the inverse demand functions implied by the ARUM, which coincide with the GL inverse demand functions (11) when $\mathbf{G} = \overline{\mathbf{G}}$, that is, $\ln \overline{G}_j(\mathbf{s}) + c = \delta_j$ with $c = \overline{CS}(\boldsymbol{\delta})$, for all $j \in \mathcal{J}$.

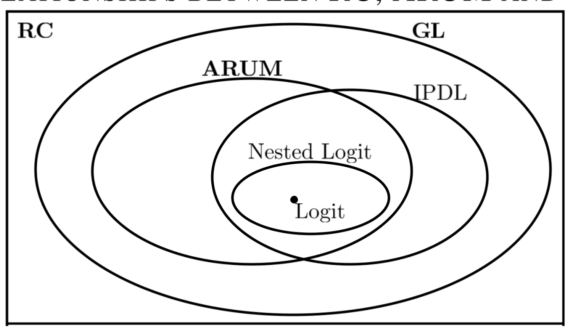
Products are always substitutes in an ARUM. By contrast, as the IPDL model, a GL model may allow for complementarity and cannot therefore be rationalized by any ARUM. We summarize the results as follows.

Proposition 5. The choice probabilities (16) implied by the ARUM coincide with the GL inverse demand functions (12) when $\mathbf{G} = \overline{\mathbf{G}} = \mathbf{H}^{-1} = \overline{\mathbf{H}}^{-1}$.

Furthermore, any ARUM is consistent with some GL model. However, the converse does not hold: some GL models are not consistent with any ARUM.

Proposition 5 shows that the choice probabilities generated by any ARUM can be derived from some GL model. As any GL model is an RC model, but not vice versa, we have therefore strengthened Hofbauer and Sandholm (2002)'s result that the demand functions generated by any ARUM can be derived from some RC models by showing that the GL structure is sufficient to recover any ARUM.

Figure 1: RELATIONSHIPS BETWEEN RC, ARUM AND GL MODELS



Overall, as illustrated in Figure 1, the class of GL models is strictly larger than the class of ARUM, but strictly smaller than the class of RC models.

7 Conclusion

This paper has introduced the IPDL model, which is a structural inverse demand model for differentiated products that captures market segmentation according to several product characteristics. The IPDL model generalizes the nested logit model by allowing arbitrary, non-hierarchical grouping structures, thus accommodating richer substitution patterns, which may include complementarity. Like the nested logit model, it can be estimated by linear instrumental variable regression using aggregate data on market shares, prices, and product characteristics, and it is consistent with a model of heterogeneous, utility-maximizing consumers. Thus, the IPDL model can be used for understanding consumer behaviour, and in turn, for analysing a range of economic questions, including market power, product entry, and regulatory changes in taxes and trade policy.[20]

[20]The supplement to this paper provides general constructions that can be exploited for building models with structures that are tailored to specific applications.

We have introduced the GL model and shown that any additive random utility model, without income effect and without heterogeneity in preferences apart from the random utility term, is a GL model. We have also shown that the GL model is consistent with the model of utility-maximizing, heterogeneous consumers of Allen and Rehbeck (2019b), which admits a representative consumer representation.

Going forward, there are a number of items on the research agenda developing the IPDL (and GL) models further. First, we could allow for income effects and for unobserved heterogeneity in preferences through random coefficients, in analogy with what has been done with the logit and nested logit models. Second, the IPDL model comprises parameters that control the relative importance of the different groups in determining demand. Still, the grouping structure must be defined by the researcher. Hortacsu et al. (2020) propose a lasso-type estimator for IPDL models that include all possible or a very large number of groups, thereby estimating the grouping structure that best fits the data. Third, this paper focuses on demand estimation with aggregate data; a natural next step would be to develop methods of estimation for the IPDL model using individual-level data. Finally, it would be useful to develop dynamic discrete choice (Rust, 1987) versions of the IPDL (and GL) models for settings in which forward-looking behavior is important.

Appendix

A Proofs and Additional Results

A.1 Mathematical Notation

We use italics for scalar variables and real-valued functions, boldface for vectors, matrices and vector-valued functions, and calligraphic for sets. By default, vectors are column vectors: $\mathbf{s} = (s_0, \ldots, s_J)^\intercal \in \mathbb{R}^{J+1}$.

$\Delta \subset \mathbb{R}^{J+1}$ is the unit simplex : $\Delta = \left\{\mathbf{s} \in [0, \infty)^{J+1} : \sum_{j\in\mathcal{J}} s_j = 1\right\}$, and $\text{relint}(\Delta) = \left\{\mathbf{s} \in (0, \infty)^{J+1} : \sum_{j\in\mathcal{J}} s_j = 1\right\}$ is its relative interior.

Let $\mathbf{G} = (G_0, \ldots, G_J) : \mathbb{R}^{J+1} \to \mathbb{R}^{J+1}$ be a vector function composed of functions $G_j : \mathbb{R}^{J+1} \to \mathbb{R}$. Its Jacobian matrix $\mathbf{J_G}(\mathbf{s})$ at $\mathbf{s}$ has entries ij given by $\frac{\partial G_i(\mathbf{s})}{\partial s_j}$.

A univariate function $\mathbb{R} \to \mathbb{R}$ applied to a vector is a coordinate-wise application of the function, e.g. $\ln(\mathbf{s}) = (\ln(s_0), \ldots, \ln(s_J))$. $\mathbf{1} = (1, \ldots, 1)^\intercal \in \mathbb{R}^{J+1}$ is a vector consisting of ones and $\mathbf{I} \in \mathbb{R}^{(J+1)\times(J+1)}$ denotes the identity matrix.

A.2 Preliminary Results

This section states some preliminary mathematical results used in the proofs below.

Lemma 1 (Euler equation)**.** Suppose that $\phi : (0, \infty)^{J+1} \to \mathbb{R}$ is linearly homogeneous. Then $\phi(\mathbf{s}) = \sum_{i=0}^{J} \frac{\partial \phi(\mathbf{s})}{\partial s_i} s_i$ for all $\mathbf{s} \in (0, \infty)^{J+1}$.

Definition 2. A matrix $\mathbf{A} \in \mathbb{R}^{(J+1)\times(J+1)}$ is positive quasi-definite if its symmetric part, defined by $\frac{1}{2}(\mathbf{A} + \mathbf{A}^\intercal)$, is positive definite.

It follows that a symmetric and positive definite matrix is positive quasi-definite.

Lemma 2 (Gale and Nikaido 1965, Theorem 6)**.** If a differentiable mapping $\mathbf{F} : \Theta \to \mathbb{R}^{J+1}$, where Θ is a convex region (either closed or non-closed) of $\mathbb{R}^{J+1}$, has a Jacobian matrix that is everywhere quasi-definite in Θ, then $\mathbf{F}$ is injective on Θ.

Lemma 3 (Simon and Blume, 1994, Theorem 14.4)**.** Let $\mathbf{F} : \mathbb{R}^{J+1} \to \mathbb{R}^{J+1}$ and $\mathbf{G} : \mathbb{R}^{J+1} \to \mathbb{R}^{J+1}$ be continuously differentiable functions. Let $\mathbf{y} \in \mathbb{R}^{J+1}$ and $\mathbf{x} = \mathbf{G}(\mathbf{y}) \in \mathbb{R}^{J+1}$. Then, the composite function $\mathbf{C} = \mathbf{F} \circ \mathbf{G} : \mathbb{R}^{J+1} \to \mathbb{R}^{J+1}$ has a Jacobian matrix $\mathbf{J_C}(\mathbf{y})$ given by $\mathbf{J_C}(\mathbf{y}) = \mathbf{J_{F\circ G}}(\mathbf{y}) = \mathbf{J_F}(\mathbf{x})\mathbf{J_G}(\mathbf{y})$.

A.3 Properties of the IPDL Model

Recall that $\mathcal{G}_d(j)$ is the set of products that are grouped with product j by characteristic d and let $s_{\mathcal{G}_d(j)} = \sum_{k\in\mathcal{G}_d(j)} s_k$ denote the market share of group $\mathcal{G}_d(j)$.

Proposition 6. The IPDL model has the following properties.

1. The IIA property holds for products of the same type; but does not hold in general for products of different types.

2. The matrix of price derivatives of demand is equal to $-\alpha\left([\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1} - \mathbf{s}\mathbf{s}^{\intercal}\right)$, with $\mathbf{s} = \boldsymbol{\sigma}(\boldsymbol{\delta})$ and where $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ has entries ij given by

$$\frac{\partial \ln G_i(\mathbf{s})}{\partial s_j} = \frac{1 - \sum_{d=1}^{D} \mu_d}{s_i} \mathbf{1}\{i = j\} + \sum_{d=1}^{D} \frac{\mu_d}{s_{\mathcal{G}_d(i)}} \mathbf{1}\{j \in \mathcal{G}_d(i)\}. \tag{17}$$

3. Products can be substitutes or complements.

Proof of Proposition 6.

1. Using Equation (6), we obtain for any pair of products j and k that

$$\frac{\sigma_j(\boldsymbol{\delta})}{\sigma_k(\boldsymbol{\delta})} = \exp\left(\frac{\delta_j - \delta_k}{1 - \sum_{d=1}^{D} \mu_d} + \sum_{d=1}^{D} \frac{\mu_d}{1 - \sum_{d=1}^{D} \mu_d} \ln\left(\frac{\sigma_{\mathcal{G}_d(k)}(\boldsymbol{\delta})}{\sigma_{\mathcal{G}_d(j)}(\boldsymbol{\delta})}\right)\right). \tag{18}$$

For products j and k of the same type (i.e. with $\mathcal{G}_d(k) = \mathcal{G}_d(j)$ for all d), Equation (18) reduces to $\frac{\sigma_j(\boldsymbol{\delta})}{\sigma_k(\boldsymbol{\delta})} = \exp\left(\frac{\delta_j - \delta_k}{1-\sum_{d=1}^{D}\mu_d}\right)$, which is independent of the characteristics or existence of all other products, i.e. IIA holds for products of the same type. When products are of different types, the ratio can depend on the characteristics of other products, which means that IIA does not hold in general.

2. This follows from Proposition 3 below applied using Equation (6).

3. Suppose there are $J = 3$ products and one outside good. Products are grouped by two characteristics: the grouping is $\{1\}, \{2, 3\}$ for the first characteristic and $\{1, 2\}, \{3\}$ for the second characteristic.

Let $\boldsymbol{\sigma}(\boldsymbol{\delta}) = \mathbf{s}$. Using Equation (17), we show that $\frac{\partial \sigma_1(\boldsymbol{\delta})}{\partial p_3} = \alpha s_1 s_3 \left[1 + \frac{\mu_1 \mu_2 s_2}{D}\right]$, where $D = -(1 - \mu_1 - \mu_2)(s_1 + s_2)(s_2 + s_3) - \mu_1 \mu_2 s_2 (1 - s_0) < 0$. Products 1 and 3 are complements if and only if $\frac{\partial \sigma_1(\boldsymbol{\delta})}{\partial p_3} < 0$, that is, if and only if $(1 - \mu_1 - \mu_2)(s_1 + s_2)(s_2 + s_3) - \mu_1 \mu_2 s_0 s_2 > 0$. □

A.4 Results for Section 5

Proof of Proposition 1. The function $\ln \mathbf{G}$ is differentiable on the convex region $\text{relint}(\Delta)$ of $\mathbb{R}^{J+1}$. In addition, $\mathbf{J}_{\ln \mathbf{G}}$ is positive quasi-definite on $\text{relint}(\Delta)$, since by assumption it is symmetric and positive definite on $\text{relint}(\Delta)$. Then $\ln \mathbf{G}$ is injective by Lemma 2.

By Definition 1, the function $\mathbf{s} \to \sum_{j\in\mathcal{J}} s_j \ln G_j(\mathbf{s})$ is strictly convex. Hence for any $\boldsymbol{\delta} \in \mathbb{R}^{J+1}$, the maximization problem $\sup_{\mathbf{s}\in\Delta}\{\sum_{j\in\mathcal{J}} s_j \ln G_j(\mathbf{s})\}$ has a unique solution. The requirement that at least one component of $\ln \mathbf{G}$ tends to infinity as $\mathbf{s}$ approaches the boundary ensures that the solution is interior. The first-order conditions are $\delta_j = \ln G_j(\mathbf{s}) + 1, j \in \mathcal{J}$. $\square$

Lemma 4. Consider the GL model defined by Equation (11).

1. The market-specific constant c is equal to

$$c = \ln\left(\sum_{k\in\mathcal{J}} H_k\left(e^{\boldsymbol{\delta}}\right)\right), \tag{19}$$

   where $\mathbf{H}(e^{\boldsymbol{\delta}}) = (H_0(e^{\boldsymbol{\delta}}), \ldots, H_J(e^{\boldsymbol{\delta}})) = \mathbf{G}^{-1}(e^{\boldsymbol{\delta}})$.

2. The Euler-type equation

$$\sum_{j\in\mathcal{J}} s_j \frac{\partial \ln G_j\left(\mathbf{s}\right)}{\partial s_k} = 1, \quad k \in \mathcal{J}, \quad \mathbf{s} \in \text{relint}(\Delta) \tag{20}$$

   holds and can be written in matrix form as $\mathbf{J}_{\ln \mathbf{G}}\left(\mathbf{s}\right)\mathbf{s} = \mathbf{1}$ for $\mathbf{s} \in \text{relint}(\Delta)$.

Proof of Lemma 4.

1. Exponentiating and applying $\mathbf{H}$ on both sides of Equation (11) leads to

$$\mathbf{s} = \mathbf{H}(e^{\boldsymbol{\delta}}e^{-c}) = \mathbf{H}(e^{\boldsymbol{\delta}})e^{-c}, \tag{21}$$

where the last equality uses the homogeneity of $\mathbf{H}$. Using that demands sum to 1 leads to Equation (19).

2. Note that

$$\sum_{j\in\mathcal{J}} s_j \frac{\partial \ln G_j\left(\mathbf{s}\right)}{\partial s_k} = \sum_{j\in\mathcal{J}} s_j \frac{\partial \ln G_k\left(\mathbf{s}\right)}{\partial s_j} = \frac{\sum_{j\in\mathcal{J}} s_j \frac{\partial G_k(\mathbf{s})}{\partial s_j}}{G_k\left(\mathbf{s}\right)} = \frac{G_k\left(\mathbf{s}\right)}{G_k\left(\mathbf{s}\right)} = 1,$$

where the first equality relies on the symmetry of the Jacobian of $\ln \mathbf{G}$ and the third equality uses the Euler equation for the homogeneous function $\mathbf{G}$. $\square$

Proof of Proposition 2. Combine Equations (19) and (21) and use $\sigma_j(\boldsymbol{\delta}) = s_j$ to obtain Equation (12).

To obtain the expression for the consumer surplus, we verify that Roy's identity holds. Set $\boldsymbol{\delta} = \ln \mathbf{G}(\mathbf{s})$. Then $(\ln \mathbf{G})^{-1}(\boldsymbol{\delta}) = \mathbf{H} \circ \exp(\boldsymbol{\delta}) = \mathbf{s}$, and by Lemma 3,

$$\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s}) = \left[\mathbf{J}_{(\ln \mathbf{G})^{-1}}(\ln \mathbf{G}(\mathbf{s}))\right]^{-1} = \left[\mathbf{J}_{\mathbf{H}\circ\exp}(\boldsymbol{\delta})\right]^{-1}.$$

By assumption, the Jacobian $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is positive definite and symmetric. Then its inverse $\mathbf{J}_{\mathbf{H}\circ\exp}(\boldsymbol{\delta})$ exists and is symmetric as well, i.e. $\frac{\partial H_i(e^{\boldsymbol{\delta}})}{\partial \delta_j} = \frac{\partial H_j(e^{\boldsymbol{\delta}})}{\partial \delta_i}$.

Then Roy's identity can be verified via

$$\begin{aligned}\frac{\partial CS(e^{\boldsymbol{\delta}})}{\partial \delta_i} &= \frac{\sum_{k\in\mathcal{J}} \frac{\partial H_k(e^{\boldsymbol{\delta}})}{\partial \delta_i}}{\sum_{j\in\mathcal{J}} H_j(e^{\boldsymbol{\delta}})} = \frac{\sum_{k\in\mathcal{J}} \frac{\partial H_i(e^{\boldsymbol{\delta}})}{\partial \delta_k}}{\sum_{j\in\mathcal{J}} H_j(e^{\boldsymbol{\delta}})}, \\ &= \frac{\sum_{k\in\mathcal{J}} \frac{\partial H_i(e^{\boldsymbol{\delta}})}{\partial e^{\delta_k}} e^{\delta_k}}{\sum_{j\in\mathcal{J}} H_j(e^{\boldsymbol{\delta}})} = \frac{H_i(e^{\boldsymbol{\delta}})}{\sum_{j\in\mathcal{J}} H_j(e^{\boldsymbol{\delta}})} = \sigma_i(\boldsymbol{\delta}),\end{aligned}$$

where the second equality uses symmetry of $\mathbf{J}_{\mathbf{H}\circ\exp}(\boldsymbol{\delta})$ and the fourth equality uses the Euler equation for the homogeneous function $\mathbf{H}$.

The Hessian of the consumer surplus is $\mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta})$, which by Proposition 3 is positive semidefinite. Convexity of the consumer surplus then follows. $\square$

Proof of Proposition 3. Differentiate $\delta_j = \ln G_j(\mathbf{s}) + CS(\boldsymbol{\delta})$ with respect to $\boldsymbol{\delta}$, then $\mathbf{I} = \mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})\, \mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta}) + \mathbf{1}\mathbf{s}^{\intercal}$, where $\mathbf{s} = \boldsymbol{\sigma}(\boldsymbol{\delta})$. $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is invertible, then $\mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta}) = [\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1}[\mathbf{I} - \mathbf{1}\mathbf{s}^{\intercal}] = [\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1} - [\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1}\mathbf{1}\mathbf{s}^{\intercal}$. Lastly, use Equation (20) in matrix form, then $[\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1}\mathbf{1}\mathbf{s}^{\intercal} = \mathbf{s}\mathbf{s}^{\intercal}$.

As a consequence, $\mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta})$ is symmetric. As $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is positive definite, the square-root matrix $[\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{1/2}$ exists and is also positive definite. Then

$$[\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{1/2}\mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta})[\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{1/2} = [\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1/2}(\mathbf{I} - \mathbf{1}\mathbf{s}^{\intercal})[\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{1/2},$$

is symmetric and idempotent and hence positive semidefinite. Then also $\mathbf{J}_{\boldsymbol{\sigma}}(\boldsymbol{\delta})$ is positive semidefinite. □

A.5 Results for Section 6

A.5.1 Representative Consumer Model

Lemma 5. Let $\ln \mathbf{G}$ be a GL inverse demand function. Then the function $\mathbf{s} \to -\mathbf{s}^{\intercal} \ln \mathbf{G}(\mathbf{s}) = -\sum_{j \in \mathcal{J}} s_j \ln G_j(\mathbf{s})$ is strictly concave on $\operatorname{relint}(\Delta)$.

Proof of Lemma 5. Consider $\mathbf{s} \in \operatorname{relint}(\Delta)$. By Part 2. of Lemma 4, the Hessian of $-\mathbf{s}^{\intercal} \ln \mathbf{G}(\mathbf{s})$ is $-\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$, which is negative definite by assumption. □

Proof of Proposition 4. Consider the representative consumer maximizing utility (13) subject to constraints (14). The budget constraint is always binding since $\alpha > 0$ and $y > \max_{j \in \mathcal{J}} p_j$. Substituting the budget constraint into the direct utility (13), the representative consumer then chooses $\mathbf{s} \in \Delta$ to maximize

$$u(\mathbf{s}) = \alpha y + \sum_{j \in \mathcal{J}} \delta_j s_j - \sum_{j \in \mathcal{J}} s_j \ln G_j(\mathbf{s})$$

where $\delta_j = v_j - \alpha p_j$. The Lagrangian of the utility maximization program given by

$$\mathcal{L}(\mathbf{s}, \lambda) = u(\mathbf{s}) + \lambda \left(1 - \sum_{j \in \mathcal{J}} s_j\right),$$

yields $\sum_{j \in \mathcal{J}} s_j = 1$ as well as the first-order conditions

$$\delta_j - \ln G_j(\mathbf{s}) - \sum_{k \in \mathcal{J}} s_k \frac{\partial \ln G_k(\mathbf{s})}{\partial s_j} - \lambda = 0, \quad j \in \mathcal{J},$$

which, by Part 2. of Lemma 4, simplify to $\delta_j - \ln G_j(\mathbf{s}) - 1 - \lambda = 0$, for all $j \in \mathcal{J}$.

The first-order condition for an interior solution has a unique solution, since the objective is strictly concave by Lemma 5, hence the utility maximizing demand exists uniquely. Setting $c = 1 + \lambda$, one obtains $\ln G_j(\mathbf{s}) + c = \delta_j$, which shows that the representative consumer model leads to the GL inverse demand function. □

A.5.2 Additive Random Utility Model

Since shifting all the δ_j's by a constant amount $c \in \mathbb{R}$ shifts the maximum expected utility $\overline{CS}$ by the same amount and does not affect choice probabilities $\mathbf{P}$, we may use the normalization $\sum_{j \in \mathcal{J}} \delta_j = 0$, i.e. we consider at no loss of generality the restrictions of $\overline{G}$ and $\mathbf{P}$ to $\Lambda = \left\{ \boldsymbol{\delta} \in \mathbb{R}^{J+1} : \sum_{j \in \mathcal{J}} \delta_j = 0 \right\}$. The following lemma collects some properties of the expected maximum utility $\overline{CS}$.

Lemma 6. The expected maximum utility $\overline{CS}$ has the following properties.

1. It is twice continuously differentiable, convex and finite everywhere.
2. It satisfies the additivity property $\overline{CS}(\boldsymbol{\delta} + c\mathbf{1}) = \overline{CS}(\boldsymbol{\delta}) + c$ for all $c \in \mathbb{R}$.
3. Its Hessian is positive definite on Λ.
4. It is given by $\overline{CS}(\boldsymbol{\delta}) = \sum_{j \in \mathcal{J}} P_j(\boldsymbol{\delta}) \delta_j + \mathbb{E}(\varepsilon_{j^*} | \boldsymbol{\delta})$, where j^* is the index of the chosen product.

Proof of Lemma 6. McFadden (1981) shows Parts 1. and 2. and Hofbauer and Sandholm (2002) show Part 3. Lastly, Part 4. follows from

$$\begin{aligned}
\overline{CS}(\boldsymbol{\delta}) &= \sum_{j \in \mathcal{J}} \mathbb{E}\left(\max_{j \in \mathcal{J}} \{\delta_j + \varepsilon_j\} | j^* = j, \boldsymbol{\delta} \right) P_j(\boldsymbol{\delta}), \\
&= \sum_{j \in \mathcal{J}} \left(\delta_j + \mathbb{E}(\varepsilon_{j^*} | j^* = j, \boldsymbol{\delta}) P_j(\boldsymbol{\delta})\right) = \sum_{j \in \mathcal{J}} P_j(\boldsymbol{\delta}) \delta_j + \mathbb{E}(\varepsilon_{j^*} | \boldsymbol{\delta}),
\end{aligned}$$

where the first equality uses the law of iterated expectations. □

Lemma 7. The function $\overline{\mathbf{H}}$ is invertible, and its inverse $\overline{\mathbf{G}} = \overline{\mathbf{H}}^{-1}$ is a GL inverse demand function.

Lemma 7 is proved in Fosgerau et al. (2020) in a very similar setting. The proof provided here applies to the exact setting of the current paper and has independent value by being simpler.

Proof of Lemma 7. We first show that $\overline{\mathbf{H}}$ is injective. Note that $\overline{\mathbf{H}}$ is differentiable. Consider the function $\boldsymbol{\delta} \to \overline{\mathbf{H}}\left(e^{\boldsymbol{\delta}}\right)$. Its Jacobian is positive definite on Λ since it has elements ij given by $\left\{e^{\overline{CS}(\boldsymbol{\delta})}\frac{\partial\overline{CS}(\boldsymbol{\delta})}{\partial\delta_i}\frac{\partial\overline{CS}(\boldsymbol{\delta})}{\partial\delta_j}\right\}+\left\{e^{\overline{CS}(\boldsymbol{\delta})}\frac{\partial^2\overline{CS}(\boldsymbol{\delta})}{\partial\delta_i\partial\delta_j}\right\}$,where the first term is a positive semi-definite matrix and where, by Part 3 of Lemma 6, the second term is a positive definite matrix on Λ. As it is also symmetric, it follows that the Jacobian is positive quasi-definite. Then $\overline{\mathbf{H}}$ is invertible by Lemma 2. By Norets and Takahashi (2013), the range of $\overline{\mathbf{H}}$ is $\mathrm{relint}(\Delta)$, which then is the domain of the inverse function $\overline{\mathbf{H}}^{-1}$.

We now show that $\ln\overline{\mathbf{G}}$ is a GL inverse demand function. Note that $\overline{\mathbf{G}}$ is linearly homogeneous and that, as shown above, the Jacobian of $\overline{\mathbf{H}}$ is symmetric and positive definite. Then, by Lemma 3, the same is true for the Jacobian of $\ln\overline{\mathbf{G}}$. □

B Details on the Simulations Experiments

In each experiment, we simulate a fully structural model of demand and supply, where the observed characteristic x_{jt} and the cost-shifter z_{jt} are i.i.d. $\mathcal{U}(0,1)$, where the unobserved product characteristics ξ_{jt} and the unobserved cost component ω_{jt} are such that $(\xi_{jt}, w_{jt}) \sim \mathcal{N}(0,\boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \begin{bmatrix} 0.2^2 & 0.1 \\ 0.1 & 0.2^2 \end{bmatrix}$, and where the grouping structure is simulated using a binomial distribution and is common across markets. Prices and market shares are determined endogenously.

Experiment on Three-Level Nested Logit Model vs. IPDL Model We generate two IPDL models. On the demand side, we set $\delta_j = -1 + 2x_{jt} - 0.2p_{jt} + \xi_{jt}$ in both models; and we set $\mu_1 = 0.1$ in the first model and $\mu_2 = 0.3$ and $\mu_1 = 0.2$ and $\mu_2 = 0.7$ in the second model. On the supply side, we specify the marginal cost function as $c_{jt} = 2+x_{jt}+z_{jt}+w_{jt}$. Estimation of the three-level nested logit models follows Verboven (1996a) and uses Gandhi and Houde (2020)'s instruments.

Experiment on PDL Model vs. IPDL Model The PDL model is a GEV model. Its demand function is given by $\sigma_j(\boldsymbol{\delta}) = e^{\delta_j}(\partial G_j(e^{\boldsymbol{\delta}})/\partial e^{\delta_j})/G(e^{\boldsymbol{\delta}})$, where $G(e^{\boldsymbol{\delta}}) = a_1\left[\sum_{g=1}^{2}\sum_{j\in G_{1g}}\left(e^{\delta_j/\rho_1}\right)^{\rho_1}\right]+a_2\left[\sum_{g=1}^{2}\sum_{j\in G_{2g}}\left(e^{\delta_j/\rho_2}\right)^{\rho_2}\right]$, with $a_1 = (1-\rho_1)/(2-$

$\rho_1 - \rho_2)$ and $a_2 = 1 - a_1$. We generate two PDL models. On the demand side, we set $\delta_j = -1 + 2x_{jt} - 0.5p_{jt} + \xi_{jt}$ in both models, and we set $\rho_1 = \rho_2 = 0.5$ in the first model and $\rho_1 = 0.9$ and $\rho_2 = 0.5$ in the second model. On the supply side, we specify the marginal cost function as $c_{jt} = 2 + x_{jt} + z_{jt} + w_{jt}$. Estimation of the IPDL models follows Section 3.

Experiment on RCL Model vs. IPDL Model On the demand side, we simulate an RCL model with a mean utility of product j given by $\delta_{jt} = 3 - p_{jt} + d_{1j} + d_{2_j} + x_{jt} + \xi_{jt}$, where $d_{kj} = 1$ if product j belongs to group k by characteristic $k = 1, 2$, zero otherwise. (d_{1j}, d_{2_j}) are assumed to have random coefficients that follow $\mathcal{N}(0, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \begin{bmatrix} 1^2 & 0.25 \\ 0.25 & 1.5^2 \end{bmatrix}$. On the supply side, we specify the marginal cost function as $c_{jt} = 2 + +d_{1j} + d_{2j} + x_{jt} + z_{jt} + w_{jt}$. Simulation of the RCL model uses the package pyblp by Conlon and Gortmaker (2020) and estimation of the IPDL model follows Section 3.

References

ALLEN, R. AND J. REHBECK (2019a): "Hicksian complementarity and perturbed utility models," *Economic Theory Bulletin*, 1–17.

——— (2019b): "Identification With Additively Separable Heterogeneity," *Econometrica*, 87, 1021–1054.

ANDERSON, S. P., A. DE PALMA, AND J. F. THISSE (1988): "A representative consumer theory of the logit model," *International Economic Review*, 29, 461–466.

——— (1992): *Discrete choice theory of product differentiation*, Cambridge, MA: MIT press.

BARNETT, W. A. AND A. SERLETIS (2008): "Consumer preferences and demand systems," *Journal of Econometrics*, 147, 210–224.

BERRY, S., A. EIZENBERG, AND J. WALDFOGEL (2016): "Optimal product variety in radio markets," *The RAND Journal of Economics*, 47, 463–497.

BERRY, S., A. GANDHI, AND P. HAILE (2013): "Connected substitutes and invertibility of demand," *Econometrica*, 81, 2087–2111.

BERRY, S. AND P. HAILE (2016): "Identification in differentiated products markets," *Annual review of Economics*, 8, 27–52.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile prices in market equilibrium," *Econometrica*, 63, 841–890.

BERRY, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 25, 242–262.

BERRY, S. T. AND P. A. HAILE (2014): "Identification in differentiated products markets using market level data," *Econometrica*, 82, 1749–1797.

BJÖRNERSTEDT, J. AND F. VERBOVEN (2016): "Does merger simulation work? Evidence from the Swedish analgesics market," *American Economic Journal: Applied Economics*, 8, 125–64.

BLUNDELL, R., J. L. HOROWITZ, AND M. PAREY (2012): "Measuring the price responsiveness of gasoline demand: Economic shape restrictions and nonparametric demand estimation," *Quantitative Economics*, 3, 29–51.

BRESNAHAN, T. F., S. STERN, AND M. TRAJTENBERG (1997): "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics*, 28, S17–S44.

CARDELL, N. S. (1997): "Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity," *Econometric Theory*, 185–213.

COMPIANI, G. (2020): "Market Counterfactuals and the Specification of Multi-Product Demand: A Nonparametric Approach," Unpublished, available on author's website.

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

CORTS, K. (1996): "The Ready-to-eat Breakfast Cereal Industry in 1994 (A)," *Harvard Business School*.

DAVIS, P. AND P. SCHIRALDI (2014): "The flexible coefficient multinomial logit (FC-MNL) model of demand for differentiated products," *The RAND Journal of Economics*, 45, 32–63.

DEATON, A. AND J. MUELLBAUER (1980): "An almost ideal demand system," *The American economic review*, 70, 312–326.

DUBÉ, J.-P., J. T. FOX, AND C.-L. SU (2012): "Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation," *Econometrica*, 80, 2231–2267.

FOSGERAU, M., D. MCFADDEN, AND M. BIERLAIRE (2013): "Choice probability generating functions," *Journal of Choice Modelling*, 8, 1–18.

FOSGERAU, M., E. MELO, A. DE PALMA, AND M. SHUM (2020): "Discrete choice and rational inattention: A general equivalence result," *International Economic Review*, 61, 1569–1589.

FUDENBERG, D., R. IIJIMA, AND T. STRZALECKI (2015): "Stochastic choice and revealed perturbed utility," *Econometrica*, 83, 2371–2409.

GALE, D. AND H. NIKAIDO (1965): "The Jacobian matrix and global univalence of mappings," *Mathematische Annalen*, 159, 81–93.

GALICHON, A. (2021): "On the representation of the nested logit model," *Econometric Theory*, 1–11.

GALICHON, A. AND B. SALANIÉ (2020): "Cupid's invisible hand: Social surplus and identification in matching models," *Available at SSRN: https://ssrn.com/abstract=1804623*.

GANDHI, A. AND J.-F. HOUDE (2020): "Measuring substitution patterns in differentiated products industries," *Available at SSRN: https://ssrn.com/abstract=3472810*.

GENTZKOW, M. (2007): "Valuing new goods in a model with complementarity: Online newspapers," *The American Economic Review*, 97, 713–744.

GRIGOLON, L. (2020): "Blurred boundaries: a flexible approach for segmentation applied to the car market," Unpublished, Available on author's website.

HAAG, B. R., S. HODERLEIN, AND K. PENDAKUR (2009): "Testing and imposing Slutsky symmetry in nonparametric demand systems," *Journal of Econometrics*, 153, 33–50.

HELLERSTEIN, R. (2008): "Who bears the cost of a change in the exchange rate? Pass-through accounting for the case of beer," *Journal of international economics*, 76, 14–32.

HOFBAUER, J. AND W. H. SANDHOLM (2002): "On the global convergence of stochastic fictitious play," *Econometrica*, 70, 2265–2294.

HORTACSU, A., J. LIEBER, J. MONARDO, AND A. DE PAULA (2020): "Estimating Nesting Structures," *Working Paper*.

IARIA, A. AND A. WANG (2020): "Identification and Estimation of Demand for Bundles," *CEPR Discussion Paper No. DP14363. Available at SSRN: https://ssrn.com/abstract=3535448*.

LEE, J. AND K. SEO (2015): "A computationally fast estimator for random coefficients logit demand models using aggregate data," *The RAND Journal of Economics*, 46, 86–102.

MATEJKA, F. AND A. MCKAY (2015): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *The American Economic Review*, 105, 272–298.

MATTSSON, L.-G. AND J. W. WEIBULL (2002): “Probabilistic choice and procedurally bounded rationality,” *Games and Economic Behavior*, 41, 61–78.

MCFADDEN, D. (1978): “Modelling the choice of residential location,” in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlquist, F. Snickars, and J. W. Weibull, Amsterdam: North Holland, vol. 673, 75–96.

——— (1981): “Econometric models of probabilistic choice,” in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski and D. McFadden. Cambridge, MA: MIT Press.

MCFADDEN, D. L. AND M. FOSGERAU (2012): “A theory of the perturbed consumer with general budgets,” NBER Working Paper No. 17953.

MILLER, N. H. AND M. C. WEINBERG (2017): “Understanding the price effects of the MillerCoors joint venture,” *Econometrica*, 85, 1763–1791.

NEVO, A. (2001): “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 69, 307–342.

——— (2011): “Empirical models of consumer behavior,” *Annu. Rev. Econ.*, 3, 51–75.

NOCKE, V. AND N. SCHUTZ (2018): “Multiproduct-firm oligopoly: An aggregative games approach,” *Econometrica*, 86, 523–557.

NORETS, A. AND S. TAKAHASHI (2013): “On the surjectivity of the mapping between utilities and choice probabilities,” *Quantitative Economics*, 4, 149–155.

PINKSE, J. AND M. E. SLADE (2004): “Mergers, brand competition, and the price of a pint,” *European Economic Review*, 48, 617–643.

ROCKAFELLAR, R. T. (1970): *Convex Analysis*, Princeton, N.J.: Princeton University Press.

RUST, J. (1987): “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 55, 999–1033.

SANDERSON, E. AND F. WINDMEIJER (2016): "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 190, 212–221.

SIMON, C. P. AND L. BLUME (1994): *Mathematics for economists*, New York, NY: W.W. Norton Company.

SMALL, K. A. (1987): "A discrete choice model for ordered alternatives," *Econometrica*, 55, 409–424.

SWAIT, J. AND A. A. MARLEY (2013): "Probabilistic choice (models) as a result of balancing multiple goals," *Journal of Mathematical Psychology*, 57, 1–14.

TRAIN, K. E. (2009): *Discrete choice methods with simulation*, New York, NY: Cambridge University Press.

VERBOVEN, F. (1996a): "International price discrimination in the European car market," *The RAND Journal of Economics*, 27, 240–268.

——— (1996b): "The nested logit model and representative consumer theory," *Economics Letters*, 50, 57–63.

Supplement to "The Inverse Product Differentiation Logit Model" - For Online Publication

Mogens Fosgerau[*] Julien Monardo[†] André de Palma[‡]

January 28, 2021

Abstract

We first present simulations investigating some properties of the Inverse Product Differentiation Logit (IPDL) model. Next, we provide a range of general methods for building Generalized Logit (GL) models along with illustrative examples that go beyond the IPDL model.

Notation We use italics for scalar variables and real-valued functions, boldface for vectors, matrices and vector-valued functions, and calligraphic for sets. By default, vectors are column vectors: $\mathbf{s} = (s_0, \ldots, s_J)^\intercal \in \mathbb{R}^{J+1}$.

$\Delta_J \subset \mathbb{R}^{J+1}$ is the unit simplex: $\Delta_J = \left\{\mathbf{s} \in [0,\infty)^{J+1} : \sum_{j\in\mathcal{J}} s_j = 1\right\}$, and $\operatorname{int}(\Delta_J) = \left\{\mathbf{s} \in (0,\infty)^{J+1} : \sum_{j\in\mathcal{J}} s_j = 1\right\}$ is its interior, where $\mathcal{J} = \{0, 1, \ldots, J\}$.

Let $\mathbf{G} = (G_0, \ldots, G_J)^\intercal : \mathbb{R}^{J+1} \to \mathbb{R}^{J+1}$ be a vector function composed of functions $G_j : \mathbb{R}^{J+1} \to \mathbb{R}$. Its Jacobian matrix $\mathbf{J}_\mathbf{G}(\mathbf{s})$ at $\mathbf{s}$ has entries ij given by $\frac{\partial G_i(\mathbf{s})}{\partial s_j}$.

[*] University of Copenhagen; mogens.fosgerau@econ.ku.dk.
[†] Telecom Paris; julien.monardo@telecom-paris.fr.
[‡] CY Cergy Paris Université; andre.de-palma@cyu.fr.

A univariate function $\mathbb{R} \to \mathbb{R}$ applied to a vector is a coordinate-wise application of the function, e.g., $\ln(\mathbf{s}) = (\ln(s_0), \ldots, \ln(s_J))$.$|\tilde{\mathbf{s}}| = \sum_{j \in \mathcal{J}} |\tilde{s}_j|$ denotes the 1-norm of vector $\tilde{\mathbf{s}}$.

1 Simulation Results for the IPDL Model

Recall that $\mathcal{G}_d(j)$ is the set of products that are grouped with product j by characteristic d and let $s_{\mathcal{G}_d(j)} = \sum_{k \in \mathcal{G}_d(j)} s_k$ denote the market share of group $\mathcal{G}_d(j)$.

In the IPDL model, the matrix of own- and cross-price derivatives of demand is equal to $-\alpha\left([\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})]^{-1} - \mathbf{s}\mathbf{s}^{\intercal}\right)$, with $\mathbf{s} = \boldsymbol{\sigma}(\boldsymbol{\delta})$ and where $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ has entries ij given by

$$\frac{\partial \ln G_i(\mathbf{s})}{\partial s_j} = \frac{1 - \sum_{d=1}^{D} \mu_d}{s_i} \mathbf{1}\{i = j\} + \sum_{d=1}^{D} \frac{\mu_d}{s_{\mathcal{G}_d(i)}} \mathbf{1}\{j \in \mathcal{G}_d(i)\}. \tag{1}$$

We cannot obtain closed-form formulae for the entries of the matrix of own- and cross-price derivatives. We therefore perform simulations to better understand the substitution patterns of the IPDL model.

Simulated Data We simulate

- A market with 20 products and an outside good;
- 20 different grouping structures along 3 dimensions, and with 3 groups per dimension. We obtain a grouping structure by simulating a 20×3 matrix of random numbers following a generalized Bernoulli distribution;
- 20 different vectors of grouping parameters $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_3)$. We obtain a vector of $\boldsymbol{\mu}$ by simulating a 4-vector of uniformly distributed random numbers, where the first element is μ_0, then normalizing so that $\boldsymbol{\mu} \in \text{int}(\Delta_3)$;
- 20 different vectors of market shares $\mathbf{s} = (s_0, \ldots, s_{20})$. We obtain a vector of market shares by simulating a 21-vector of uniformly distributed random numbers, where the first element is s_0, then by normalizing the vector of market shares of products so that $\mathbf{s} \in \text{int}(\Delta_{20})$.

This normalization ensures that we simulate markets with very low and very high values for μ_0 and s_0. Combining the grouping structures, the grouping parameters, and the market shares, we form $8,000$ markets. The following table gives summary statistics on the simulated data.

TABLE 1: SUMMARY STATISTICS ON THE SIMULATED DATA

Variable	Mean	Min	Max
s_0	0.5253	0.0064	0.9906
s_j	0.0158	3e-06	0.0697
μ_0	0.4662	0.0697	0.9532
μ_1	0.2014	0.0135	0.8480
μ_2	0.1420	0.0175	0.4036
μ_3	0.1904	0.0059	0.5212

Grouping Structure Table 2 shows the distribution of the own- and cross-price derivatives according to the number of common groups.

Own-price elasticities are always negative, while cross-price elasticities can be either negative (complementarity) or positive (substitutability). Products of the same type are always substitutes. Products that are very similar (i.e., that are grouped together by all characteristics but one) are also always substitutes. Products that are very different can be either substitutes or complements. Products are less likely to be substitutes as they become more different.

Table 2: DISTRIBUTION OF PRICE DERIVATIVES ACCORDING TO THE NUMBER OF COMMON GROUPS

Common groups	$\mathbf{J}_{\boldsymbol{\sigma}} > 0$	Median	Min	Max	Freq.
Own-price derivatives					
–	0.00%	-0.0222	-0.7781	-3e-06	100.00%
Cross-price derivatives					
0 (None)	45.33%	-7e-07	-0.1539	0.0251	25.09%
1	90.38%	0.0002	-0.1114	0.2082	43.59%
2	100.00%	0.0006	-1e-09	0.2641	26.47%
3 (All)	100.00%	0.0009	-1e-09	0.3100	4.85%
Total	82.09%	0.0002	-0.1539	0.3100	100.00%

Notes: Column "$\mathbf{J}_{\boldsymbol{\sigma}} > 0$" gives the percentage of positive cross-price elasticities according to the number of common groups (e.g., the row "2" concerns products that share two groups). Column "Freq." gives the frequencies (in percentage) of the cross-price elasticities (e.g., 4.85 percent of the cross-price elasticities involve products of the same type).

Grouping Parameters Table 3 shows the distribution of cross-price derivative according to the proximity of products into the characteristics space used to form product types, as measured by $\mu_{jk} = \sum_{d=1}^{3} \mu_d \mathbf{1}\{j \in \mathcal{G}_d(k)\}$ for two products j and k.

As the parameter μ_{jk} becomes larger, we observe that (i) the derivatives increase in values, and that (ii) the share of substitutes increases. This is because higher μ_d means that products of the same group by characteristic d become more similar.

Table 3: PERCENTAGE OF SUBSTITUTES ACCORDING TO THE VALUE OF μ_{jk}

μ_{jk}	$\mathbf{J}_{\boldsymbol{\sigma}} > 0$	Median	Min	Max
$[0, 0.1[$	65.60%	0.0000	-0.1539	0.0286
$[0.1, 0.2[$	96.37%	0.0002	-0.0538	0.1462
$[0.2, 0.3[$	93.52%	0.0003	-0.1114	0.1670
$[0.3, 0.4[$	94.16%	0.0007	-0.0673	0.2082
$[0.4, 0.5[$	93.89%	0.0009	-0.0432	0.2049
$[0.5, 1[$	100.00%	0.0020	1e-08	0.3100

Summary In the IPDL model,

1. (Grouping structure) Products of the same type are always substitutes. Products of different types may be substitutes or complements, depending on the

degree of closeness between products as measured by the value of the parameters μ_d and by the closeness of the products into the characteristics space used to form product types. The closer two products are, the more likely they are to be substitutes.

2. (Grouping parameters) The size of the cross-derivatives depends on the degree of closeness. The closer two products are, the higher is their cross-derivative.

2 Construction of GL Models

In this section, we provide general methods for building GL models, along with illustrative examples. They allow us to obtain alternative models to the logit and nested logit models that have interesting features: some of them can accommodate complementarity, others have closed form for both the demand function and its inverse.

Definition A. A generalized logit (GL) inverse demand function is a function $\ln \mathbf{G}$, where $\mathbf{G} : (0, \infty)^{J+1} \to (0, \infty)^{J+1}$ is linearly homogeneous and where the Jacobian $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is positive definite and symmetric.

Definition B. An almost GL inverse demand is a function that satisfies the requirements for being a GL inverse demand, except that the Jacobian $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is only required to be positive semi-definite rather than positive definite.

2.1 General Methods and Illustrative Examples

The first result in this section shows that averaging an almost GL inverse demand with a GL inverse demand yields a new GL inverse demand.

Proposition A (Averaging)**.** Let $\mathbf{G}_k, k \in \{1, \ldots, K\}$, be almost GL inverse demands with at least one being a GL inverse demand. Let $(\alpha_1, \ldots, \alpha_K) \in \text{int}(\Delta_{K-1})$. Then

$$\ln \mathbf{G}(\mathbf{s}) = \sum_{k=1}^{K} \alpha_k \ln \mathbf{G}_k(\mathbf{s})$$

is a GL inverse demand.

Proof of Proposition A. The function $\mathbf{G}$ is linearly homogeneous since for $\lambda > 0$,

$$\begin{aligned} \mathbf{G}(\lambda \mathbf{s}) &= \prod_{k=1}^{K} G_k(\lambda \mathbf{s})^{\alpha_k} = \prod_{k=1}^{K} \lambda^{\alpha_k} G_k(\mathbf{s})^{\alpha_k}, \\ &= \left(\prod_{k=1}^{K} \lambda^{\alpha_k}\right)\left(\prod_{k=1}^{K} G_k(\mathbf{s})^{\alpha_k}\right), \\ &= \left(\lambda^{\sum_{k=1}^{K}\alpha_k}\right)\left(\prod_{k=1}^{K} G_k(\mathbf{s})^{\alpha_k}\right) = \lambda \mathbf{G}(\mathbf{s}), \end{aligned}$$

where the second equality uses the homogeneity of the functions G_k and the fourth equality uses the restrictions on parameters $\sum_{k=1}^{K}\alpha_k = 1$.

The Jacobian of $\ln \mathbf{G}$, given by $\mathbf{J}_{\ln \mathbf{G}} = \sum_{k=1}^{K} \alpha_k \mathbf{J}_{\ln \mathbf{G}_k}$, is symmetric as the linear combination of symmetric matrices, and positive definite as the linear combination of at most $K-1$ positive semi-definite matrices and at least one positive definite matrix. □

Proposition A leads to the following corollary.

Corollary A (General grouping)**.** Let $\mathcal{G} \subseteq 2^{\mathcal{J}}$ be a finite set of groups with associated parameters μ_g, where $\mu_{0j} + \sum_{\{g\in\mathcal{G}|j\in g\}} \mu_g = 1$ for all $j \in \mathcal{J}$ with $\mu_g \geq 0$ for all $g \in \mathcal{G}$ and $\mu_{0j} > 0$ for all $j \in \mathcal{J}$. Let $\ln \mathbf{G} = (\ln G_0, \dots, \ln G_J)$ be given by

$$\ln G_j(\mathbf{s}) = \mu_{0j} \ln(s_j) + \sum_{\{g\in\mathcal{G}|j\in g\}} \mu_g \ln\left(\sum_{i\in g} s_i\right).$$

Then $\ln \mathbf{G}$ is a GL inverse demand.

Proof of Corollary A. Let $T_j^0(\mathbf{s}) = s_j$ and for each $g \in \mathcal{G}$, $\mathbf{T}^g = (T_1^g, \dots, T_J^g)$ with $T_j^g(\mathbf{s}) = \left(\sum_{i\in g} s_i\right)^{\mathbf{1}\{j\in g\}}$. The Jacobian of $\ln \mathbf{T}^g$ has elements jk given by $\frac{\mathbf{1}\{j\in g\}\mathbf{1}\{k\in g\}}{\sum_{i\in g} s_i}$, and thus $\mathbf{J}_{\ln \mathbf{T}^g} = \frac{\mathbf{1}_g \mathbf{1}_g^{\intercal}}{\sum_{i\in g} s_i}$ where $\mathbf{1}_g = (\mathbf{1}\{1 \in g\}, \dots, \mathbf{1}\{J \in g\})^{\intercal}$. Each $\mathbf{T}^g$ is an almost GL inverse demand while $\mathbf{T}_0$ is the logit inverse demand. Lastly, $\sum_{\{g\in\mathcal{G}|j\in g\}} \mu_g + \mu_{0j} = 1$. Then the conditions for application of Proposition A are fulfilled. □

The grouping structure in Corollary A is arbitrary and therefore allows the grouping structure that defines the IPDL model. The presence of the inverse logit demand, due to $\mu_0 > 0$, ensures that the Jacobian $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s})$ is always positive definite and hence that the inverse demand is indeed invertible.

If the outside good 0 belongs only to one group and is the only member of that group, then $\ln G_0(\mathbf{s}) = \ln(s_0) = \delta_0 + c$. Setting $\delta_0 = 0$ and assuming a linear index, the model of Corollary A boils down to the linear regression model

$$\ln\left(\frac{s_j}{s_0}\right) = \mathbf{x}_j\boldsymbol{\beta} - \alpha p_j + \sum_{g\in\mathcal{G}(j)} \mu_g \ln\left(\sum_{k\in g} s_k\right) + \xi_j, \quad j = 1, \dots, J.$$

The following proposition shows how a GL inverse demand can be transformed into another GL inverse demand by application of a location shift and a matrix with non-negative elements that sum to one across rows and columns. Let unnormalized demands $\tilde{\mathbf{s}}$ be demands obtained before normalizing their sum to one, i.e., $\mathbf{s} = \tilde{\mathbf{s}}/|\tilde{\mathbf{s}}|$.

Proposition B (Transformation)**.** Let $\mathbf{T}$ be a GL inverse demand and $\mathbf{m} \in \mathbb{R}^{J+1}$ be a location shift vector. Let $\mathbf{A} \in \mathbb{R}^{(J+1)\times(J+1)}$ be an invertible matrix such that $a_{ij} \geq 0$ and $\sum_{i\in\mathcal{J}} a_{ij} = \sum_{j\in\mathcal{J}} a_{ij} = 1$. Then the function $\ln \mathbf{G}$ given by

$$\ln \mathbf{G}(\mathbf{s}) = \mathbf{A}^{\intercal}\left[\ln\left(\mathbf{T}\left(\mathbf{A}\mathbf{s}\right)\right)\right] + \mathbf{m} \tag{2}$$

is a GL inverse demand, and the corresponding unnormalized demands are given by

$$\tilde{\mathbf{s}} = \mathbf{A}^{-1}\mathbf{T}^{-1}\left(\exp\left[\left(\mathbf{A}^{\intercal}\right)^{-1}\left(\boldsymbol{\delta} - \mathbf{m}\right)\right]\right). \tag{3}$$

Proof of Proposition B. The function $\mathbf{G}$ defined by Equation (2) is linearly homogeneous since for $\lambda > 0$,

$$\begin{aligned}
\mathbf{G}\left(\lambda\mathbf{s}\right) &= \exp\left(\mathbf{A}^{\intercal}\ln\mathbf{T}\left(\mathbf{A}\left(\lambda\mathbf{s}\right)\right) + \mathbf{m}\right), \\
&= \exp\left(\mathbf{A}^{\intercal}\ln\lambda + \mathbf{A}^{\intercal}\ln\mathbf{T}\left(\mathbf{A}\mathbf{s}\right) + \mathbf{m}\right), \\
&= \exp\left(\ln\lambda + \mathbf{A}^{\intercal}\ln\mathbf{T}\left(\mathbf{A}\mathbf{s}\right) + \mathbf{m}\right) = \lambda\mathbf{G}\left(\mathbf{s}\right),
\end{aligned}$$

where the second equality uses the homogeneity of $\mathbf{T}$, and the third equality uses the feature that columns of $\mathbf{A}$ sum to 1.

The Jacobian of $\ln \mathbf{G}$ is $\mathbf{J}_{\ln \mathbf{G}}(\mathbf{s}) = \mathbf{A}^{\intercal}\mathbf{J}_{\ln \mathbf{T}}(\mathbf{s})\mathbf{A}$, which is symmetric and positive definite. Unnormalized demands (3) follow from solving $\ln \mathbf{G}(\tilde{\mathbf{s}}) = \boldsymbol{\delta}$. □

Proposition B provides models where both demand and inverse demand have closed form, as it is the case of the logit and nested logit models. We illustrate this proposition with a GL inverse demand that allows for complementarity.

Example A. Let $J+1=3$, $\mathbf{m}=\mathbf{0}$, $\mathbf{T}(\mathbf{s})=\mathbf{s}$, and

$$\mathbf{A} = \begin{pmatrix} p & 1-p & 0 \\ 1-p & p & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

with $p<0.5$. Then we obtain that

$$\tilde{\mathbf{s}} = \mathbf{A}^{-1}\left(\exp\left[\left(\mathbf{A}_{\intercal}\right)^{-1}\boldsymbol{\delta}\right]\right) = \begin{pmatrix} \frac{p}{2p-1}e^{\frac{p}{2p-1}\delta_1-\frac{1-p}{2p-1}\delta_2} - \frac{1-p}{2p-1}e^{\frac{p}{2p-1}\delta_2-\frac{1-p}{2p-1}\delta_1} \\ \frac{p}{2p-1}e^{\frac{p}{2p-1}\delta_2-\frac{1-p}{2p-1}\delta_1} - \frac{1-p}{2p-1}e^{\frac{p}{2p-1}\delta_1-\frac{1-p}{2p-1}\delta_2} \\ e^{\delta_3} \end{pmatrix},$$

so that

$$s_3 = \sigma_3(\boldsymbol{\delta}) = \frac{e^{\delta_3}}{e^{\frac{p}{2p-1}\delta_1-\frac{1-p}{2p-1}\delta_2} + e^{\frac{p}{2p-1}\delta_2-\frac{1-p}{2p-1}\delta_1} + e^{\delta_3}},$$

and $\frac{\partial\sigma_3(\boldsymbol{\delta})}{\partial\delta_1} > 0$ if and only if $\delta_2-\delta_1 > (2p-1)\ln\left(\frac{1-p}{p}\right)$.

2.2 Zero Demands

The constructions above rule out zero demands (this is also the case of the models discussed in the main text). The following proposition shows how we can build models that allow zero demands by slightly modifying Proposition A and applying it to functions $\mathbf{G}$ defined on $[0,\infty)^{J+1}$ instead of just $(0,\infty)^{J+1}$.

Proposition C (Invertible grouping)**.** Let $\mathcal{G} = \{g_0, \ldots, g_J\}$ be a finite set of $J+1$ groups (i.e., the number of groups is equal to the number of products). Let $\mu_g > 0$, for all $g \in \mathcal{G}$, be the associated parameters, where $\sum_{\{g\in\mathcal{G}|j\in g\}} \mu_g = 1$ for all $j \in \mathcal{J}$. Let $\mathbf{G} = (G_0, \ldots, G_J) : [0,\infty)^{J+1} \to (0,\infty)^{J+1}$ be given by

$$\ln G_j(\mathbf{s}) = \sum_{\{g\in\mathcal{G}|j\in g\}} \mu_g \ln\left(\sum_{i\in g} s_i\right). \tag{4}$$

Let $\mathbf{W} = \operatorname{diag}(\mu_{g_0}, \ldots, \mu_{g_J})$ and let $\mathbf{M} \in \mathbb{R}^{(J+1)\times(J+1)}$ with entries $M_{jk} = \mathbf{1}_{\{j\in g_k\}}$ (where rows correspond to products and columns to groups). If $\mathbf{M}$ is invertible, then $\ln \mathbf{G}$ has all the properties of a GL inverse demand, except that it is defined on Δ_J, and the unnormalized demands satisfy

$$\boldsymbol{\delta} = \ln \mathbf{G}(\tilde{\mathbf{s}}) \Leftrightarrow \tilde{\mathbf{s}} = \left(\mathbf{M}^\intercal\right)^{-1} \exp\left(\mathbf{W}^{-1}\mathbf{M}^{-1}\boldsymbol{\delta}\right).$$

Proof of Proposition C. Following the proof of Proposition A, the function $\mathbf{G}$ given by Equation (4) clearly has all the properties of an almost GL inverse demand. Thus, it remains to show that the Jacobian of $\ln \mathbf{G}$ is positive definite if $\mathbf{M}$ is invertible.

Observe that

$$\begin{aligned}\ln G_j(\mathbf{s}) &= \sum_{k\in\mathcal{J}} \mu_{g_k} \mathbf{1}\{j \in g_k\} \ln\left(\sum_{i\in g_k} s_i\right)\\ &= \sum_{k\in\mathcal{J}} \mu_{g_k} \mathbf{1}\{j \in g_k\} \ln\left(\sum_{i\in\mathcal{J}} \mathbf{1}\{i \in g_k\} s_i\right),\end{aligned}$$

and, in turn, that

$$\frac{\partial \ln G_j(\mathbf{s})}{\partial s_l} = \sum_{k\in\mathcal{J}} \mu_{g_k} \frac{\mathbf{1}\{j\in g_k\}\mathbf{1}\{l\in g_k\}}{\sum_{i\in g_k} s_i},$$

which can be expressed in matrix form as

$$\mathbf{J}_{\ln\mathbf{G}}(\mathbf{s}) = \mathbf{M}\mathbf{V}\mathbf{M}^\intercal,$$

with $\mathbf{V} = \operatorname{diag}\left(\frac{\mu_{g_0}}{\sum_{i\in g_0} s_i}, \ldots, \frac{\mu_{g_J}}{\sum_{i\in g_J} s_i}\right)$. This is positive definite since all μ_g are strictly positive and $\mathbf{M}$ is invertible.

Lastly, with $\mathbf{M}$ invertible, unnormalized demands solve the equation $\ln \mathbf{G}(\tilde{\mathbf{s}}) = \mathbf{MW} \ln (\mathbf{M}^\intercal \tilde{\mathbf{s}}) = \boldsymbol{\delta}$ and are given by $\tilde{\mathbf{s}} = (\mathbf{M}^\intercal)^{-1} \exp (\mathbf{W}^{-1}\mathbf{M}^{-1}\boldsymbol{\delta})$. $\square$

As it is illustrated in the following example and as it is the case in ARUM where error terms have bounded support, Proposition C allows for zero demands when there is no group containing only one product. Note that this proposition also allows to build models with closed form for both the demands and their inverses.

Example B. Define groups from the symmetric matrix $\mathbf{M}$ with entries $M_{ij} = \mathbf{1}_{\{i \neq j\}}$, so that each product belongs to $J + 1$ groups. Its inverse, $\mathbf{M}^{-1}$, has entries ij equal to $\frac{1}{J+1} - 1_{\{i=j\}}$.

Let $\mu_g = 1/(J + 1)$ for each group $g = 0, \ldots, J$. Then the unnormalized demands are given by $\tilde{\mathbf{s}} = (\mathbf{M})^{-1} \exp [(J + 1)\mathbf{M}^{-1}\boldsymbol{\delta}]$ and lead to the following demands

$$\sigma_i(\boldsymbol{\delta}) = \frac{\tilde{s}_i}{\sum_{j\in\mathcal{J}} \tilde{s}_j} = \frac{\sum_{j\in\mathcal{J}} e^{-(J+1)\delta_j} - (J+1)e^{-(J+1)\delta_i}}{\sum_{j\in\mathcal{J}} e^{-(J+1)\delta_j}}. \tag{5}$$

Demands (5) are non-negative only for values of $\boldsymbol{\delta}$ within some set. To ensure positive demands, it is sufficient to average with the simple inverse logit demand. Demands (5) are not consistent with any ARUM since they do not exhibit the feature of the ARUM that the mixed partial derivatives of $\sigma_i(\boldsymbol{\delta})$ alternate in sign. Indeed, products are substitutes

$$\frac{\partial \sigma_1(\boldsymbol{\delta})}{\partial \delta_2} = -J^2 e^{-J(\delta_1+\delta_2)} / \left(\sum_{j\in\mathcal{J}} e^{-J\delta_j} \right)^2 < 0,$$

but

$$\frac{\partial^2 \sigma_1(\boldsymbol{\delta})}{\partial \delta_2 \partial \delta_3} = -2J^3 e^{-J(\delta_1+\delta_2+\delta_3)} / \left(\sum_{j\in\mathcal{J}} e^{-J\delta_j} \right)^3 < 0.$$