The Flexible Inverse Logit Model*

Julien Monardo†

December 24, 2020
PRELIMINARY AND INCOMPLETE

Abstract

This paper develops the flexible inverse logit model, a structural inverse demand model which is able to describe the behavior of heterogeneous, utility-maximizing consumers choosing from a choice set of possibly many products that are differentiated in a way that is both observed and unobserved by the modeller. The FIL model is easy to estimate by linear instrumental variables regression to deal with the endogeneity issues of prices and market shates due to the modelling of unobserved product differentiation. Furthermore, the FIL model accommodates rich substitution patterns, including complementarity in demand. In particular, simulations show that it is able to match the substitution patterns of the random coefficient logit pretty well and to get quite right predictions of a merger's price and share effects.

*First Draft: January, 2019. I would like to thank Steven Berry, Debopam Bhattacharya, Philippe Choné, Olivier De Groote, André de Palma, Xavier D'Haultfoeuille, Roxana Fernandez, Alfred Galichon, Gautam Gowrisankaran, Laura Grigolon, Ulrich Kaiser, Laurent Linnemer, Hugo Molina, and Pasquale Schiraldi as well as participants at the internal seminar in microeconomics at CREST, at the Paris Young Economists Seminar on the Economics of Digitization (Telecom Paris), at the 2nd Workshop on Platforms, E-commerce and Digital Economics (CREST- Telecom Paris), at the external ALISS Seminar (INRAE), at the 10th conference on Industrial Organization and the Food Industry (TSE), at the 46th EARIE Annual Conference, 12th Econometric Society World Congress, at the DSE WINTER SCHOOL 2020. This paper is based on Chapter 2 of my doctoral dissertation at École Normale Supérieure Paris-Saclay.

†CREST, ENS Paris-Saclay, University Paris-Saclay; jmonardo@ens-paris-saclay.fr.

1 Introduction

How consumers substitute across differentiated products plays an important role in economics, since substitution patterns (i.e., the own- and cross-price demand elasticities), when combined with a supply model, serve as key input to many important economic questions. Prominent examples span a wide range of topics in industrial organization, such as market power (Berry et al., 1995; Nevo, 2001), mergers (Nevo, 2000) and products' entry (Petrin, 2002; Gentzkow, 2007), as well as regulatory changes in taxes and trade policy (Goldberg, 1995; Verboven, 1996a; Berry et al., 1999; Griffith et al., 2019).[1] Therefore, it is important to obtain good estimates of substitution patterns, since any error will propagate into the supply side estimates and potentially lead to wrong conclusions.

The literature following Berry (1994) and Berry et al. (1995) has shown how to flexibly estimate substitution patterns by incorporating rich heterogeneity in consumer preferences while handling endogeneity issues due to the modelling of unobserved product differentiation. The state-of-the art is the BLP approach (Berry et al., 1995), which uses the random coefficient logit model to accommodate rich substitution patterns. However, this flexibility complicates the (empirical) identification and estimation.[2] In contrast, another widely used approach employs the nested logit models, which are easy to estimate by linear instrumental variables (IV) regression, but have been criticized for restricting substitution patterns.

Motivated by these observations, this paper proposes the Flexible Inverse Logit (FIL) model, a novel (inverse) demand model between the random coefficient logit and the nested logit models that (i) flexibly describes the utility-maximizing behavior of a population of heterogeneous consumers choosing among products that are differentiated in a way that is both observed and unobserved by the modeller; and (ii) is easy to estimate by linear IV regression.

The FIL model can be motivated as a member of Fosgerau et al. (2020)'s nesting-based class of inverse demand models where there is a nest for each pair of products with its

[1]See Berry and Haile (2014, Footnote 1) and Berry and Haile (2016, Table 1) for additional examples.

[2]Flexibility of the RCL model can be difficult to obtain in practice, since, theoretically, it requires many random coefficients (McFadden and Train, 2000), which are not easily identified in applications (Reynaert and Verboven, 2014; Gandhi and Houde, 2020). Estimation can be painful and time-consuming because it requires non-linear, non-convex optimization, simulation and numerical inversion of the demand function. This also implies dealing with the associated issues of local optima and choice of starting values, accuracy of the simulation and numerical inversion (see, e.g., Knittel and Metaxoglou, 2014, and references therein). See Conlon and Gortmaker (2020) for the current best practices in the estimation of structural demand models using BLP approach. Other approaches to solve BLP's problem have been proposed (Dubé et al., 2012; Lee and Seo, 2015; Salanié and Wolak, 2019).

corresponding nesting parameter.[3] Noting that the FIL model boils down to the logit model when all nesting parameters equal zero, this implies that the FIL model follows the fruitful approach in demand estimation that deviates from the logit model and its counterintuitive substitution patterns thanks to its nesting parameters.[4] It is also consistent with a specific instance of the large class of models of heterogeneous, utility-maximizing consumers studied by Allen and Rehbeck (2019), where its nesting parameters capture consumer heterogeneity in preferences.

Furthermore, the FIL model accommodates rich substitution patterns among products that are governed by its nesting parameters. Theoretically, it is Diewert (1974)'s flexible in a large class of well-defined inverse demand models, i.e., it can match the matrix of own- and cross-price elasticities of demand implied by any model of this class. Besides, in contrast to the random coefficient logit model, the FIL model does not restrict the products to be substitutes in demand, as defined by a negative cross-price derivative of demand.[5]

The FIL model is easy to estimate by a nested logit-type linear IV regression of shares on prices, product characteristics and log-shares of the product into its nests. This linear IV regression allows to deal with the endogeneity of prices and market shares due to the modelling of unobserved product differentiation and makes the estimation easy and fast. It also helps clarifying its identification, which amounts to identify its parameters thanks to instruments, i.e., variables that generates exogenous variation in each of the endogeneous variables. Instruments for the FIL model include conventional instruments, which implies identification does not requires any unconventional source of variation.

Lastly, this paper uses simulations to compare the FIL model to the BLP approach. For this purpose, using insights from Pinkse et al. (2002), the nesting parameters are projected into characteristics space to make, as in the BLP approach, the substitution patterns of the FIL model depend on the similarity of products into that characteristics directly and a function of a small number of parameters. Simulations demonstrate the ability of the FIL model, projected into characteristics space, to match the substitution patterns and the markups implied by the BLP approach as well as to obtain good predictions of the price and share effects of two counterfactuals, a merger (Nevo, 2000) and product's entry (Petrin, 2002; Gentzkow, 2007).

This paper is linked to two strands of literature on demand estimation. First, it is in line

[3] The idea of using such a nesting structure is not new. See Chu (1989); Koppelman and Wen (2000); Davis and Schiraldi (2014).

[4] See discussion in Section 2.

[5] As it is common in the literature, this paper rules out income effects.

with the structural approach that follows Berry (1994) and Berry et al. (1995)'s method, which models unobserved product differentiation through the inclusion of structural errors. These latter terms therefore have an economic interpretation that can be used for identification purposes. However, this approach typically leads to inverse demand equations for which an explicit formula does not exist, which must then be inverted numerically during estimation. This paper follows this approach by allowing for structural errors but contrasts with it by just requiring linear regression for estimation. Closest to this paper are Davis and Schiraldi (2014), Compiani (2020) and Fosgerau et al. (2020). Davis and Schiraldi (2014) also builds a Diewert (1974)'s flexible structural demand model. However, by contrast to the FIL model, the FC-MNL restricts the product to be substitutes in demand and requires the numerical inversion for estimation. Furthermore, like Compiani (2020) and Fosgerau et al. (2020), this paper estimates closed-form inverse demand functions that allow for unobserved product differentiation by regression techniques.[6]

Furthermore, this paper relates to the flexible functional form approach (see Barnett and Serletis, 2008, and references therein). This line of research builds flexible models in the sense of Diewert (1974) to derive demand equations that only comprise observables and require the addition of additive error terms to serve as a basis for estimation. This means that the error terms have no immediate structural interpretation, which prevents the use of the standard economic arguments for identification used by the structural approach.

The remainder of the paper is organized as follows. Section 2 introduces the FIL model. Section 3 develops the methods to estimate the FIL model with data on market shares, prices and product characteristics and discusses about its identification. Section 4 compares the FIL model to the BLP approach. Section 5 concludes.

2 A Flexible Demand Model

Consider a population of consumers choosing from a choice set of $J+1$ differentiated products, where products $j=1,\dots,J$ are the inside products and product $j=0$ is the

[6] Compiani (2020) develops a non-parametric approach to estimate inverse demands in differentiated products markets based on aggregate data. As this paper, his approach does not make any distributional assumptions on unobservables and imposes minimal functional form restrictions based on economic theory. However, by contrast to this paper, he imposes the connected substitutes structure of Berry et al. (2013), which rules out complementarity defined by negative cross-price derivative of *unit* demand), and because he uses a non-parametric estimator, his approach is subject to a curse of dimensionality that may constrain its feasibility to settings with small choice sets. Note also that he allows for non-unit demand, whereas I assume unit demand.

outside good. Each inside product j is defined by $(\mathbf{x}_j, p_j, \xi_j)$, where $\mathbf{x}_j \in \mathbb{R}^K$ is a vector of K observed characteristics, $p_j \in \mathbb{R}$ is its price, and $\xi_j \in \mathbb{R}$ is an unobserved characteristics term. Let $\mathbf{x} \equiv (\mathbf{x}_1, \ldots, \mathbf{x}_J)$, $\mathbf{p} \equiv (p_1, \ldots, p_J)$ and $\boldsymbol{\xi} \equiv (\xi_1, \ldots, \xi_J)$, and defined $\Delta_J^+ \equiv \left\{\mathbf{s} \in (0, \infty)^J : \sum_{j=1}^J s_j < 1\right\}$.

Following Berry and Haile (2014), assume a linear index restriction. That is, partition the vector of product characteristics as $\mathbf{x} = \left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}\right)$, with $\mathbf{x}^{(1)} \in \mathbb{R}^{K_1}$ and $\mathbf{x}^{(2)} \in \mathbb{R}^{K_2}$ with a support denoted by $\mathcal{X}_2$, and define linear indexes as

$$\delta_j \equiv \mathbf{x}_j^{(1)}\boldsymbol{\beta} - \alpha p_j + \xi_j, \quad j = 1, \ldots, J, \tag{1}$$

Furthermore, consider the inverse demand function

$$\boldsymbol{\sigma}^{-1} = \left(\sigma_1^{-1}, \ldots, \sigma_J^{-1}\right) : \Delta_J^+ \times \mathcal{X}_2 \to \mathbb{R}^J,$$

which gives the vector $\boldsymbol{\delta} \equiv (\delta_1, \ldots, \delta_J)$ of products indexes as a function of the vector of nonzero market shares $\mathbf{s} \equiv (s_1, \ldots, s_J) \in \Delta_J^+$ and product characteristics $\mathbf{x}^{(2)}$,

$$\boldsymbol{\sigma}^{-1}\left(\mathbf{s}, \mathbf{x}^{(2)}; \boldsymbol{\mu}\right) = \boldsymbol{\delta} \tag{2}$$

where $\boldsymbol{\mu}$ is a parameter vector to be estimated.[7] With the linear index restriction, characteristics $\mathbf{x}^{(1)}$ and $\boldsymbol{\xi}$, and prices $\mathbf{p}$ enter the inverse demand equations (2) only through the product indexes $\boldsymbol{\delta}$, whereas $\mathbf{x}^{(2)}$ can enter in an unrestrictive way.[8] This implies that $\mathbf{x}^{(1)}$ and $\boldsymbol{\xi}$ are perfect substitutes and that consumer behavior is arbitrarily affected by $\mathbf{x}^{(2)}$.

Lastly, define the market share of the outside good by the identity $s_0 = 1 - \sum_{j=1}^J s_j$ and normalize the vector $\boldsymbol{\delta}$ by setting $\delta_0 = 0$.

The remainder of this section introduces the Flexible Inverse Logit model, studies its microfoundation, its implied patterns of substitution among products. Proofs for this section are provided in Appendix A.

[7] As it is common in the literature, this setting relies on three implicit assumption: the unobserved characteristics terms are scalars, there is no income effect, zero demand is ruled out.

[8] In the logit model, there is no $\mathbf{x}^{(2)}$ and in the random coefficient logit model, $\mathbf{x}^{(2)}$ are the characteristics that have a random coefficient.

2.1 The Flexible Inverse Logit (FIL) Model

Let the FIL model be the inverse demand function $\boldsymbol{\sigma}^{-1}: \Delta_J^+ \to \mathbb{R}^J$ defined by

$$\sigma_j^{-1}(\mathbf{s}; \boldsymbol{\mu}) \equiv \ln\left(\frac{s_j}{1 - \sum_{k=1}^J s_k}\right) - \sum_{i \neq j} \mu_{ij} \ln\left(\frac{s_j}{s_i + s_j}\right) = \delta_j, \quad j = 1, \ldots, J, \tag{3}$$

where δ_j is defined by Equation (1) and where $\boldsymbol{\mu} \equiv (\mu_{ij})_{i,j=1,\ldots,J}$ is a parameter vector.

The FIL model reduces to the logit model when all μ_{ij} equal zero. As is well known, the logit model leads to counterintuitive substitution patterns whereby the decrease in the price of product j decreases the demand of any other product $k \neq j$ by the same percentage.[9] This restrictive pattern is a manifestation of the independence from irrelevant alternatives (IIA) property of the logit model,[10] and has led the literature to develop demand models that allow for deviations from IIA: (i) by introducing unobserved consumer heterogeneity in preferences through different random coefficients specifications (e.g, Berry et al. (1995) uses the random coefficient logit model); (ii) by using different nesting structures in the GEV framework developed by McFadden (1978).[11] However, with the exception of the nested logit model, these richer models complicate the estimation, since they do not have an explicit inversion formula such as Equation (3) and, in turn, prevent the use of regression for estimation.[12]

Recently, Fosgerau et al. (2020) have proposed a class of closed-form inverse demand functions that generalizes that of the nested logit model by allowing any possible nesting structure. The FIL model can be viewed, under some restrictions on $\boldsymbol{\mu}$, as a member of this class, where the nesting structure has a nest for each pair (i, j) of inside products $i \neq j$ and a nest for the outside good alone. These restrictions on $\boldsymbol{\mu}$ are

[9] For example, if, following a 2 percent decrease in the price of product 1, the demand the demand of product 2 drops by 1 percent, then the demand of any other product $j > 2$ also drops by 1 percent.

[10] As the FIL model reduces to the logit model when all μ_{ij}'s equal zero, the IIA property can be tested using standard Wald tests.

[11] Prominent examples include the nested logit model (Ben-Akiva, 1973), the ordered logit model (Small, 1987), the PDL model (Bresnahan et al., 1997), the FC-MNL model (Davis and Schiraldi, 2014), the ordered nested logit model (Grigolon, 2019), etc. See also Chapter 4 in Train (2009).

[12] Following (Berry, 1994), Equations of the form of (3) can be rearranged to obtain inverse demand equations in which unobserved characteristics terms ξ_j, which play the role of structural error terms, are functions of the data (i.e., $\mathbf{s}$, $\mathbf{p}$ and $\mathbf{x}$) and the parameters (i.e., α, $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$) to be estimated. He then suggests to use them as a basis for demand estimation. When the inverse demand function σ_j^{-1} has an explicit formula, then one can use standard regression techniques for estimation. Otherwise, one can implement Berry et al. (1995)'s estimator using either the nested-fixed point algorithm of Berry et al. (1995) or the mathematical program with equilibrium constraints algorithm proposed by Dubé et al. (2012).

(R1) $\sum_{i \neq j} \mu_{ij} < 1$ for all $j = 1, \ldots, J$,

(R2) $\mu_{ij} = \mu_{ji}$ for all $i, j = 1, \ldots, J$, $i \neq j$.

As it will be made clear below, restrictions (R1) and (R2) are key for the FIL model to be consistent with utility maximization and have implications in terms of substitution patterns. Furthermore, they also imply that the FIL model is invertible in $\mathbf{s}$, i.e., that the specified inverse demand (3) defines a demand function (rather than a demand correspondence).[13]

Proposition 1. Let the FIL model satisfy Restrictions (R1) and (R2). Consider any vector $\boldsymbol{\delta} \in \mathbb{R}^J$ of product indexes. Then, there exists a unique vector $\mathbf{s} \in \Delta_J^+$ of nonzero market shares such that $\boldsymbol{\delta} = \boldsymbol{\sigma}^{-1}(\mathbf{s}; \boldsymbol{\mu})$.

Proposition 1 implies there exists a demand function $\boldsymbol{\sigma} : \mathbb{R}^J \rightarrow \Delta_J^+$ that gives the vector of market shares as a function of product indexes, $\mathbf{s} = \boldsymbol{\sigma}(\boldsymbol{\delta}; \boldsymbol{\mu})$.

Microfoundation. The FIL model can also be derived as specific instance of the large class of utility models of consumer heterogeneity studied by Allen and Rehbeck (2019). To see this, observe first that the FIL model is consistent with a (representative) consumer choosing a vector $\mathbf{s} \in \Delta_J^+$ of shares so as to maximize her utility function given by

$$\sum_{j=1}^{J} \delta_j s_j - \left[\sum_{j=1}^{J} s_j \ln \left(\frac{s_j}{1 - \sum_{k=1}^{J} s_k} \prod_{i \neq j} \left(\frac{s_j}{s_i + s_j} \right)^{\mu_{ij}} \right) + \ln \left(1 - \sum_{k=1}^{J} s_k \right) \right], \tag{4}$$

where δ_j is defined by Equation (1), thereby implying that $\alpha > 0$ is the consumer' price sensitivity (i.e., its marginal utility of income) and $\boldsymbol{\beta}$ captures the consumer' taste for characteristics $\mathbf{x}^{(1)}$.

Furthermore, restrictions (R1) and (R2) imply that the second term of utility (4) is a strictly concave function of $\mathbf{s}$ that do not depend on $\boldsymbol{\delta}$. Then, by Allen and Rehbeck (2019), utility (4) can be derived, after an aggregation across consumers, from a utility model of

[13] Berry et al. (2013) show that their "connected substitutes" structure is sufficient for invertibility. The connected substitutes structure requires that (i) products be weak substitutes, i.e., everything else equal, an increase in δ_j weakly decreases demand σ_i for all other products; and (ii) the "connected strict substitution" condition hold, i.e., there is sufficient strict substitution between products to treat them in one demand system. Proposition 1 accommodates some substitution patterns that are not allowed by Berry et al. (2013), including some form of complementarity. See also Proposition 4 in Appendix A that extends this result to a large class of inverse demand functions.

heterogeneous, utility-maximizing consumers. Therefore, the FIL model allows for consumer heterogeneity modelled by the second term of utility (4);[14] and the parameter vector $\boldsymbol{\mu}$ has an obvious interpretation: it is a vector of structural parameters as it comprises key parameters describing consumers' preferences (it controls for the distribution of preferences in the population of consumers) and is invariant to changes in economic policy, such as taxes, or in firms' strategy, such as pricing strategies, product characteristics, new products (see Hurwicz, 1966).[15]

Substitution Patterns. The matrix $\boldsymbol{\eta} = [\eta_{ij}] = [(\partial\sigma_i/\partial p_j)\,(p_j/\sigma_j)]$ of own- and cross-price elasticities captures the patterns of substitution among products: own-price elasticities express how much a price increase would cause a drop in demand, whereas cross-price elasticities indicate where the lost demand goes following the price increase. Formally, this matrix involves the derivatives of the demand $\boldsymbol{\sigma}$ with respect to the prices $\mathbf{p}$. However, by the implicit function theorem, it can also be expressed in terms of the derivatives of the inverse demand function $\boldsymbol{\sigma}^{-1}$ with respect to the shares as $\boldsymbol{\eta} = -\alpha\left(\left[\left(\partial\sigma_i^{-1}/\partial s_j\right)(\sigma_i/p_j)\right]\right)^{-1}$. For the FIL model,

$$\frac{\partial\sigma_i^{-1}}{\partial s_j} = \begin{cases} \dfrac{1}{1-\sum_{k=1}^{J} s_k} + \dfrac{1-\sum_{i\neq j}\mu_{ij}}{s_j} + \sum_{i\neq j}\dfrac{\mu_{ij}}{s_i+s_j} & \text{if } i=j \\ \\ \dfrac{1}{1-\sum_{k=1}^{J} s_k} + \dfrac{\mu_{ij}}{s_i+s_j} & \text{if } i\neq j \end{cases} \tag{5}$$

Equation (5) shows that restrictions (R1) and (R2) imply that the FIL model yields a demand function with a derivative matrix that is negative definite and symmetric.[16] This, in turn, has two implications in terms of substitution patterns: (i) the demand of each product is strictly in its own price (equivalently, strictly increasing in its own index); (ii) as explained below, the FIL model does not restrict products to be substitutes in demand, as defined by a positive cross-price derivative of demand.

[14]The FIL model, as it stands, does not allow for observed heterogeneity in preferences related to observed individual characteristics as well as unobserved heterogeneity in preferences through random coefficients

[15]Following Fosgerau et al. (2020), the FIL model is also consistent with a representative consumer model choosing a positive quantity of every product while trading-off variety against quantity. With this representation, the parameters μ_{ij} measure taste for variety over products i and j, since $\mu_{ij} > 0$ makes the representative consumer more willing to choose products i and j (see Verboven, 1996b, for the nested logit model).

[16]See Monardo (2020) for further details. Also, since symmetry is implied by restriction (R2), symmetry can be tested, rather than imposed, using standard Wald tests. Note, however, if symmetry is rejected by the data, any study based on the demand estimates would not rely on an underlying structural utility model.

Furthermore, it is worth noting, by contrast to the nested logit model, that the nesting structure of the FIL model does not require the modeller to take a stand on the relevant dimensions along which nests can be defined, i.e., its implied substitution patterns are not constrained by a predetermined segmentation of the market.[17] Rather, Equation (5) emphasizes that the nesting structure is a way to fully parametrize the matrix of inverse demand derivatives (and thus its inverse). By contrast, the logit model fully sparsifies the off-diagonal entries of this matrix, which just reduce to $1/(1-\sum_{k=1}^{J} s_k)$.

Formally, the FIL model can be motivated as being flexible in the sense of Diewert (1974) in a large class of well-defined inverse demand functions, i.e., it can match the vector of market shares as well as any matrix of own- and cross-price elasticities implied by an inverse demand function of this class.[18]

The proof of the flexibility can be sketched in two steps. The first step uses by Proposition 1 to show that there always exists a vector $\boldsymbol{\delta}$ of indexes that equate the vector $\mathbf{s}$ of observed shares to the vector $\boldsymbol{\sigma}$ of predicted share. The second step shows the ability to the FIL model to match any own- and cross-price elasticities. Intuitively, one can match the cross-price elasticity η_{ij} by appropriately choosing the value of the nesting parameter μ_{ij}. Once this is done, all own-price elasticity η_{jj} are automatically matched because, using that the FIL model exhibits unit demand, they can be expressed as $\eta_{jj} = -\sum_{i\neq j} \eta_{ij}(s_i/s_j)$. The following proposition summarizes this discussion.

Proposition 2. The FIL model is flexible in the sense of Diewert (1974) in the class of inverse demand functions of the form of $\sigma_j^{-1}(\mathbf{s};\boldsymbol{\mu}) = \ln G_j(\mathbf{s};\boldsymbol{\mu}) - \ln(s_0)$, where $\ln \mathbf{G}$ is homogeneous of degree one and has a matrix of derivatives that is symmetric and positive definite on Δ_J^+.

Closest to the FIL model is the FC-MNL model developed by Davis and Schiraldi (2014), which (to my knowledge) is the only existing Diewert (1974) flexible demand model that allows for unobserved product differentiation. This is in contrast to the flexible

[17] The choice of the nesting structure can be problematic in applications. Consider for example the market for cars, where cars are assumed to belong to five segments: subcompact, compact, standard, intermediate, and luxury. Grigolon (2019) suggests a natural ordering of cars from subcompact to luxury, while Brenkers and Verboven (2006) consider a nested structure without prior ordering. Determining which of the two nesting structures best describes the market is not obvious.

[18] A demand system is said to be flexible in the sense of Diewert (1974) if it is able to provide a first-order approximation to any theoretically grounded demand system at a point in price space. Equivalently, flexibility can also be viewed as the ability of the (direct or indirect) utility function to provide second-order approximations to any utility function. This is because the partial derivatives of the demand function can be uniquely derived from the second partial derivatives of the utility function.

functional form approach which also develops Diewert (1974) flexible demand models, but does not model unobserved product differentiation.

The sketch of the proof of the flexibility of the FIL model has highlighted the role of the nesting parameters in driving the substitution patterns. To get further intuition, consider the following two stylized examples.

Example 1. Let $J = 3$ and assume that $s_1 = s_2 = s_3 = s < 1/3$ and $p_1 = p_2 = p_3 = 1$. Then, the cross-price demand elasticity for product i with respect to product j is given by

$$\varepsilon_{ij} = \alpha \left[s - \frac{\mu_{12}\mu_{13} + \mu_{12}\mu_{23} + \mu_{13}\mu_{23} - 2\mu_{ij}}{D} \right], \tag{6}$$

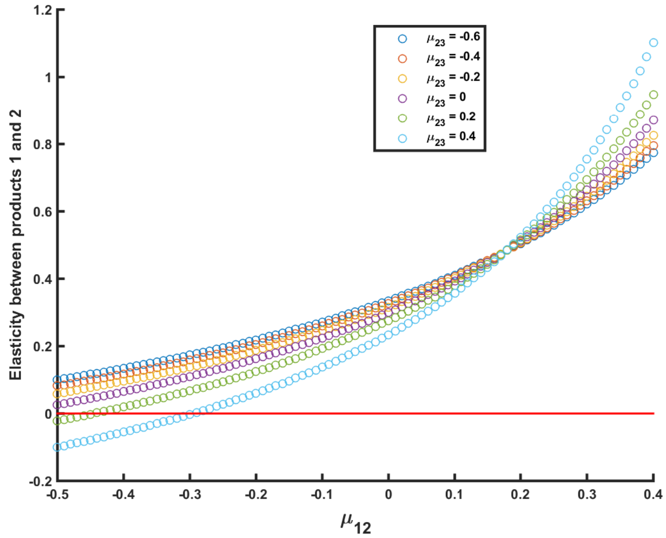
where $D \equiv 4(1 - \mu_{12} - \mu_{13} - \mu_{23}) + 3(\mu_{12}\mu_{13} + \mu_{12}\mu_{23} + \mu_{13}\mu_{23}) > 0$,[19] so that

$$\varepsilon_{12} - \varepsilon_{13} = \frac{\alpha(\mu_{12} - \mu_{13})}{D} \gtrless 0 \quad \Leftrightarrow \quad \mu_{12} \gtrless \mu_{13}, \tag{7}$$

i.e., higher cross-price derivatives are associated with higher values for the associated nesting parameters.

Example 2. Let $J = 3$ and assume that $s_1 = 0.15$, $s_2 = 0.25$ and $s_3 = 0.20$ and $p_1 = p_2 = p_3 = 1$. Furthermore, set $\mu_{23} = 0.2$ and let μ_{12} and μ_{13} vary. The scatter plot below shows that higher value for μ_{12} implies higher value for the elasticity between products 1 and 2. By varying the value of μ_{13}, it also shows that the relationship (in terms of elasticity) between products 1 and 2 is affected by product 3.

[19] D is positive as it can be shown to be proportional to the determinant of a positive definite matrix.



The scatter plot shows the cross-price demand elasticity of product 1 with respect to product 2 as a function of μ_{12} for different values of μ_{13}. The red line correspond to the threshold between complementarity and substitutability.

Examples 1 and 2 show how the nesting parameter μ_{ij} governs the substitution patterns between products i and j. First, the higher the value for μ_{ij}, the higher the cross-price elasticity η_{ij}. Then, the FIL model allows for complementarity in demand, i.e., $\eta_{ij} < 0$. Lastly, whether products i and j are complements or substitutes does not uniquely depend on μ_{ij} but also on μ_{ik}, $k \neq j$, $k \neq i$. This is consistent with theoretical result whereby whether two products are complements or substitutes depends on the relation of the two products to the other products (Samuelson, 1974; Ogaki, 1990).[20]

2.2 Projection into Characteristics Space

As it stands, the FIL model is defined into product space, rather than into product characteristics space. This implies, by contrast to the RCL model, that many parameters need to be estimated and that substitution patterns do not depend on product characteristics directly (they depend on product characteristics only through the indexes).

To deal with these issues, I apply the distance-metric approach of Pinkse et al. (2002): based on the intuitive idea that similarities (or distances) between products into characteristics space should drive substitutions, that is, that closer products into characteristics space

[20] A future version of this paper will discuss how the nesting parameters μ_{ij} relate to different definitions of complementarity used in the empirical literature.

are likely to be more substitutable, I suggest to project the nesting parameters mu_{ij} into characteristics space $\mathbf{x}^{(2)}$.[21]

The projection is as follows

$$\mu_{ij} = \mu\left(\mathbf{d}_{ij}^{(2)}; \boldsymbol{\gamma}\right), \tag{8}$$

where $\mathbf{d}_{ij}^{(2)} \equiv (d_{ij1}^{(2)}, \ldots, d_{ijK_2}^{(2)})$ is a measure of similarity between products i and j into the characteristics space formed by $\mathbf{x}^{(2)} = \left(x_1^{(2)}, \ldots, x_{K_2}^{(2)}\right)$ and $\boldsymbol{\gamma}$ is a parameter vector to be estimated.

The projection (8) maps the FIL model from product space into the characteristics space formed by $\mathbf{x}^{(2)}$ by writing each parameter μ_{ij} as a function of a measure of similarity between products i and j.

Furthermore, a question is whether the parameter vector $\boldsymbol{\mu}$ is structural after the projection. Clearly, it is no longer structural, since it is not invariant to changes in product characteristics by firms. However, the parameter vector $\boldsymbol{\gamma}$, which parametrizes the function μ, is structural and has a clear interpretation. As the random coefficients in a RCL model, it controls for the distribution of valuation for product characteristics $\mathbf{x}^{(2)}$ in the population of consumers.

Lastly, with the projection, the FIL model (3) can be rewritten as

$$\sigma_j^{-1}\left(y_j, \mathbf{y}_{-j}; \boldsymbol{\gamma}\right) = \ln\left(\frac{s_j}{1 - \sum_{k=1}^J s_k}\right) - \sum_{i=1}^J \mu\left(\mathbf{d}_{ij}^{(2)}; \boldsymbol{\gamma}\right) \ln\left(\frac{s_j}{s_i + s_j}\right) + C, \tag{9}$$

where $C \in \mathbb{R}$ is a market-specific constant.

Examining Equation (9) shows the implications of the projection.

Proposition 3. Let $y_j \equiv (s_j, \mathbf{x}_j^{(2)})$ and $\mathbf{y}_{-j} \equiv (y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_J)$. The FIL model projected into characteristics space (9) is

(i) symmetric: $\sigma_j^{-1}\left(y_j, \mathbf{y}_{-j}; \boldsymbol{\gamma}\right) = \sigma_k^{-1}\left(y_j, \mathbf{y}_{-j}; \boldsymbol{\gamma}\right) = \sigma^{-1}\left(y_j, \mathbf{y}_{-j}; \boldsymbol{\gamma}\right)$ for any $j \neq k$,

[21]This strategy has been successfully applied by Pinkse and Slade (2004), Slade (2004) and Rojas (2008) for demand estimation purposes. See also Pinkse and Slade (1998). Note that I do not implement the semi-parametric estimator of Pinkse et al. (2002). In my model, their method would use a series expansion to approximate μ, and in turn, this would introduce an additional source of endogeneity. Indeed, in addition to the structural error $\boldsymbol{\xi}$, their method adds an approximation error, due neglected expansion errors, that is a function of characteristics $\mathbf{x}^{(2)}$.

(ii) anonymous: $\sigma^{-1}\left(y_j, \mathbf{y}_{-j}; \boldsymbol{\gamma}\right) = \sigma^{-1}\left(y_j, \mathbf{y}_{\rho(-j)}; \boldsymbol{\gamma}\right)$, where $\rho(-j)$ is any permutation of the product indexes $-j$,

(iii) invariant to translation in $\mathbf{x}^{(2)}$: $\sigma^{-1}\left(y_j + (0, c), \mathbf{y}_{-j} + (\mathbf{0}, c\mathbf{1}); \boldsymbol{\gamma}\right) = \sigma^{-1}\left(y_j, \mathbf{y}_{-j}; \boldsymbol{\gamma}\right)$ for all $c \in \mathbb{R}$.

Overall this implies that the identity of products does not matter and that only differences in characteristics $\mathbf{x}^{(2)}$ do. Then, the FIL model, projected into product characteristics, places the same restrictions as the RCL model on its inverse demand function (Gandhi and Houde, 2020). The projection therefore relates to Compiani (2020) who uses anonymity to reduce the number of parameters of the inverse demand to be estimated in a nonparametric setting and on Gandhi and Houde (2020) who use anonymity and symmetry of any inverse demand function consistent with an additive random utility linear in $\mathbf{x}^{(2)}$ to construct new approximations of the optimal instruments.

3 Estimation and Identification of the FIL Model

3.1 Estimation

Consider having data on market shares s_{jt}, prices p_{jt} and product characteristics $\mathbf{x}_{jt}$ for T markets, indexed by t, and J products per market (see, e.g., Berry et al., 1995; Nevo, 2001).

Rearranging Equations (3) shows that the FIL model is estimated by the following nested logit-like regression of market shares on product characteristics, prices and log-shares terms related to its nesting structure

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \mathbf{x}_{jt}^{(1)}\boldsymbol{\beta} - \alpha p_{jt} + \sum_{i \neq j} \mu_{ij} \ln\left(\frac{s_{jt}}{s_{it} + s_{jt}}\right) + \xi_{jt}, \tag{10}$$

where $s_{0t} = 1 - \sum_{j=1}^{J} s_{jt}$ for all $t = 1, \ldots, T$, and, with the projection, $\mu_{ij} = \mu\left(\mathbf{d}_{ijt}^{(2)}; \boldsymbol{\gamma}\right)$.

Following the literature, the unobserved product/market characteristic terms ξ_{jt} are the structural error terms. They summarize all the product/market characteristics that observed by consumers and firms but not by the modeller. Furthermore, product characteristics $\mathbf{x}$ are assumed to be exogenous (i.e., to be uncorrelated with the structural error terms $\boldsymbol{\xi}$), whereas prices and the log-share terms in the right-hand side of Equation (10) are considered as endogenous. Prices are endogenous because, as it is typically assumed in price competition models with differentiated products, firms consider both observed and unobserved characteristics when they set their prices. Market shares are endogenous since they are determined a full system of demand equations involving the entire vectors of endogenous prices and of unobserved characteristics, and because consumers choose products while potentially considering the unobserved characteristics.

Therefore, the FIL model boils down to a linear instrumental variable (IV) regression. The IV regression allows to deal with the endogeneity of prices and market shares due to the modelling of unobserved product differentiation through the inclusion of structural error terms ξ_{jt}. Furthermore, the linear structure helps clarifying the empirical identification (in terms of instruments needed and tests for weak identification) and also eases and accelerates the estimation (e.g., by application of the two-stage least squares estimator).

Consider now the case where the nesting parameters are projected into characteristics space. Each nesting parameter μ_{ij} is therefore an unknown function of a measure of similarity $\mathbf{d}_{ij}^{(2)}$ between products i and j into the characteristic space formed by the characteristics $\mathbf{x}^{(2)}$. In practice, the modeller will face several choices regarding the characteristics space to consider (i.e., which $\mathbf{x}^{(2)}$ to choose), the measure of similarity or distance to use (e.g., $\mathbf{d}_{ij}^{(2)}$ is an absolute value, an Euclidian distance, etc.) and the functional form, known up to some parameters $\boldsymbol{\gamma}$ to be estimated, to give to μ (e.g., μ a polynomial of degree 6 in $\mathbf{d}_{ij}^{(2)}$). The following example consider the set of choices that are made in the simulations of the next section.

Example 3 (Projection into a One-dimensional Space)**.** Consider the case of $K_2 = 1$ product characteristic $x^{(2)}$ taking values on the interval $[0, 1]$. Then, $d_{ijt}^{(2)} = 1 - |x_{it} - x_{jt}|$ is one measure of similarity ranging from 0 (minimal similarity) to 1 (maximimal similarity).

The function μ can be specified as a polynomial function in $d_{ijt}^{(2)}$ for it to be a flexible function

$$\mu\left(d_{ijt}^{(2)}; \boldsymbol{\gamma}\right) = \sum_{k=0}^{M} \gamma_k \left(d_{ijt}^{(2)}\right)^k, \tag{11}$$

where M is the degree of the polynomial function to be chosen by the modeller, and where $\boldsymbol{\gamma} \equiv (\gamma_0, \ldots, \gamma_M)$ is the vector of parameters to be estimated.

3.2 Identification

Identification of the FIL model amounts to identifying its parameters. The FIL model boils down to the linear IV regression (10), where prices and log-share terms are endogenous.

Therefore, the main identification assumption is the existence of instruments **z** , that is, the existence of variables that induce enough independent exogenous variation in each of these endogenous variables. The discussion about identification thus reduces to the discussion about what sources of empirical variation will help identifying demand parameters.

Consider first the vector of utility indexes δ. It is easy to identify, since higher market shares implies higher utility indexes. Formally, as shown in Proposition 1, given $\boldsymbol{\mu}$, there is a one-to-one mapping between the vector of utility indexes $\boldsymbol{\delta}$ and the vector of market shares **s**.

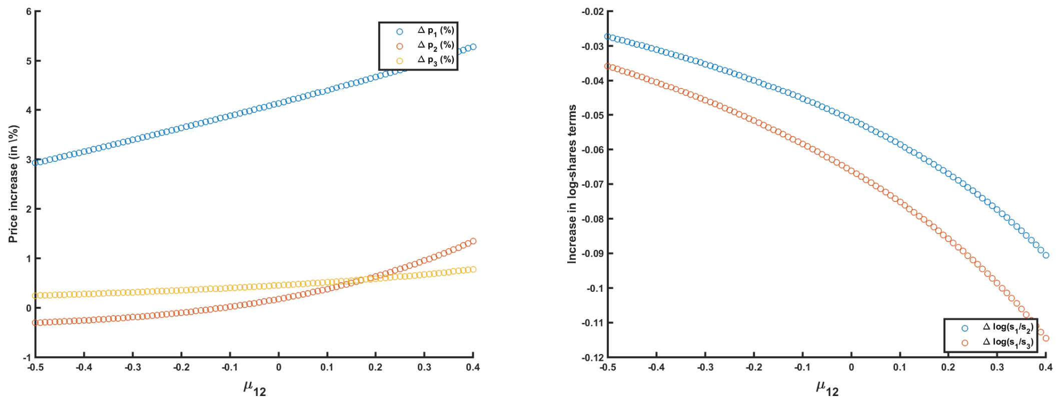
Regarding the parameters α and β parametrizing the utility indexes, their identification is fairly simple. As is well known in the literature (see e.g., Berry, 1994; Berry et al., 1995; Berry and Haile, 2014, etc.), this requires dealing with price endogeneity. This is done by finding valid supply-side instruments, i.e., cost shifters and/or markup shifters. The first set of instruments includes the Hausman instruments (Hausman et al., 1994; Nevo, 2001); the second set involves the BLP instruments (Berry et al., 1995; Gandhi and Houde, 2020) as well as market shocks such mergers Miller and Weinberg (2017).[22]

Turn now to the nesting parameters μ_{ij}, which, as mentioned above (cf. Example 2), govern the substitution between products i and j. Their identification is more tricky since it requires exogenous variation in the relative popularity of product j with respect to product i. In other words, one needs instruments that reveal about the substitution patterns among products. Variables that generate exogenous variation in the choice set (including changes in prices, product characteristics and number of products) are therefore good candidates as instruments.

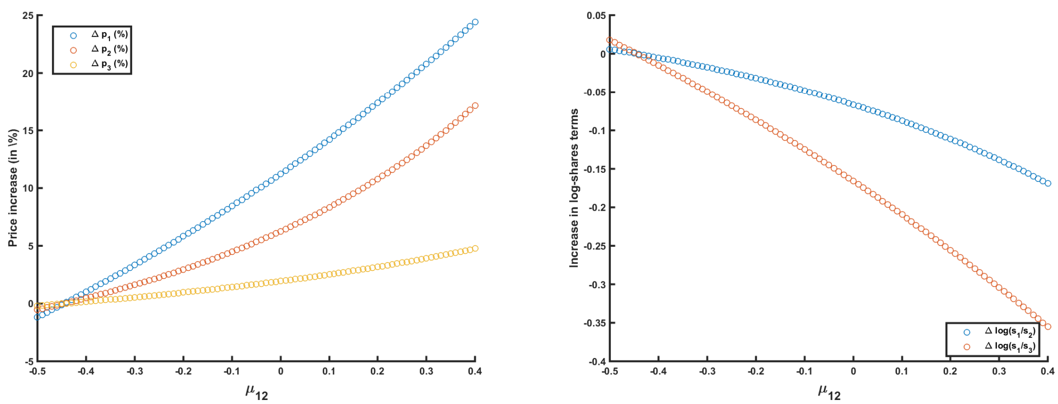
To get further insights on the identification of $\boldsymbol{\mu}$, consider the following two sets of instruments that help learning about the substitution patterns: the first set comprises cost shifters of own and competing products and the second set is a merger. For concreteness, consider a stylized $J = 3$ case (similar to Example 2), where the demand model is a FIL model and the supply model is a price competition model with three single-product firms. Let again $s_1 = 0.15$, $s_2 = 0.25$, $s_3 = 0.20$ and $p_1 = p_2 = p_3 = 1$, but $\mu_{13} = \mu_{23} = 0.2$.

Consider two sets of simulations that shows how the two sources of data variation help identifying $\boldsymbol{\mu}$. First, assume that product 1s cost c_1 increase by 10%. The following scatter plots show how the prices (left panel) and the relative shares of the products (right panel) are affected depending on the level of μ_{12}.

[22]Hausman instruments are prices in other markets. BLP instruments are functions of the characteristics of competing products).



Second, let firms 1 and 2 merge. The following scatter plots show how prices and relative shares are affected depending on the level of μ_{12}.



In both sets of simulations, one observes monotonic relationships between the between the variation in prices and μ_{12} on one hand, and between the variation in relative shares and μ_{12} on other hand. This shows that the way prices and relative shares change with product 1s cost increase or with the marger drives the estimates of the $\boldsymbol{\mu}$s.

Lastly, consider the case of the projection. Identification of the parameters $\boldsymbol{\gamma}$ parametrizing the projection μ requires finding a unique vector of parameters $\boldsymbol{\gamma}$ such that $\mu_{ij} = \mu\left(\mathbf{d}_{ij}^{(2)}\right)$ for all $i \neq j$. In the one-dimensional projection case of Example 3, $\boldsymbol{\gamma}$ can be viewed as the OLS estimates of a regression of μ_{ij} on $\{d_{ij}^{(2)k}\}_{k=0,\ldots,M}$. The identification assumptions are therefore the same as in the OLS setting. This result easily extends to the multi-dimensional case.

4 Comparison to the BLP Approach

This sections uses simulations to compare the FIL model to the BLP approach in terms of post-estimation outputs (namely, price elasticities of demand and markups) and two counterfactuals (namely, merger simulation and new product).

4.1 Simulations

Setting. I simulate a fully structural model of demand and supply. The demand side is a standard static RCL model with a single normally distributed random coefficient on an exogenous continuous characteristics. In the RCL model, the conditional indirect utility of a consumer n in market t from choosing an inside product j is given by

$$u_{njt} = \beta_0 + \beta_n x_{jt} - \alpha p_{jt} + \xi_{jt} + \varepsilon_{njt}, \tag{12}$$

where the utility from choosing the outside good $j = 0$ is normalized to $u_{n0t} = \varepsilon_{n0t}$, for all markets $t = 1, \ldots, T$, where the ε_{njt}'s are assumed to be distributed i.i.d. type I extreme value. Each consumer n chooses one unit of the product that provides her the highest utility. Then, the market share of product j in market t is computed as the probability that product j provides the highest utility across all products in market t.

The supply side is a static oligopolistic price competition model with multiproduct firms,[23] where the marginal cost c_{jt} is parametrized as follows

$$c_{jt} = \gamma_0 + \gamma_x x_{jt} + \gamma_w w_{jt} + \omega_{jt},$$

where x_{jt} is the product characteristic, which affects utility and cost, w_{jt} is a variable which only affects cost, and ω_{jt} is an unobserved cost component.

Simulations. The simulations, which are strongly based on those of Armstrong (2016), consider three data generating processes (DGP) by varying J and T. Simulations and estimations of the RCL model make use of the Python package `PyBLP` by Conlon and

[23] In a static oligopolistic price competition model between F firms, each firm $f =, 1, \ldots, F$ producing the set of products $\mathcal{J}_f$, with $\cup_{f=1}^{F} \mathcal{J}_f = \mathcal{J}_0$ and $\cap_{f=1}^{F} \mathcal{J}_f = \emptyset$, chooses the prices p_j of its products $j \in \mathcal{J}_f$ to maximize its profit function given by $\Pi_f = \sum_{j \in \mathcal{J}_f} (p_j - c_j)\, \sigma_j(\boldsymbol{\delta}, \mathbf{x}^{(2)}; \boldsymbol{\mu})$, where c_j is the marginal cost of product j. Furthermore, assuming that a pure-strategy Nash equilibrium exists, prices $\mathbf{p}$ and market shares $\mathbf{s}$ are therefore determined by the associated first-order conditions.

Gortmaker (2020), which implements the best practices for estimating RCL models using the BLP method, and use the approximate method to compute (an approximation of the) optimal instruments.

For each DGP, 100 Monte Carlo datasets are constructed, and for each of them, T markets are simulated, where each one consists of J products and $F = 5$ firms producing each one $J/5$ products. Each product j in market t is characterized by the vector $(s_{jt}, p_{jt}, x_{jt}, \xi_{jt}, w_{jt}, \omega_{jt})$, where x_{jt} and w_{jt} are drawn from two independent standard uniform distributions, and where the vector of structural error terms, $(\xi_{jt}, \omega_{jt}) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = \begin{bmatrix} 0.50 & 0.25 \\ 0.25 & 0.50 \end{bmatrix}$. Lastly, set $\beta_n \sim \mathcal{N}(6, 3)$, $\alpha = 1$, $\gamma_0 = 2$ and $\gamma_x = \gamma_w = 1$.

4.2 Results

Elasticities and Markups (Berry et al., 1995; Nevo, 2001). The following table shows, for all the DGPs, that the FIL model fits pretty well the true substitution patterns and the true markups. In comparison with the BLP approach, the FIL model does a good job. The best fits are obtained for the second DGP which has a higher number of markets.

Table 1: Post-Estimation Outputs

	Own-Elasticities	Cross-Elasticities	Markups
DGP with $J = 25$ and $T = 100$			
True	-4.065	0.161	0.335
	[-4.095 ; -4.035]	[0.160 ; 0.163]	[0.329 ; 0.341]
FIL	-3.869	0.159	0.363
	[-4.437 ; -3.300]	[0.136 ; 0.182]	[0.303 ; 0.424]
BLP	-4.076	0.162	0.335
	[-4.471 ; -3.681]	[0.146 ; 0.178]	[0.302 ; 0.368]
DGP with $J = 50$ and $T = 200$			
True	-4.157	0.081	0.329
	[-4.173 ; -4.141]	[0.080 ; 0.082]	[0.325 ; 0.332]
FIL	-4.009	0.080	0.341
	[-4.287 ; -3.731]	[0.074 ; 0.085]	[0.318 ; 0.365]
BLP	-4.138	0.080	0.330
	[-4.333 ; -3.942]	[0.076 ; 0.084]	[0.314 ; 0.347]
DGP with $J = 100$ and $T = 20$			
True	-4.207	0.0401	0.325
	[-4.242 ; -4.173]	[0.040 ; 0.042]	[0.318 ; 0.333]
FIL	-4.3410	0.042	0.356
	[-4.889 ; -3.794]	[0.037 ; 0.048]	[0.308 ; 0.403]
BLP	-4.242	0.040	0.324
	[-4.705 ; -3.779]	[0.036 ; 0.045]	[0.288 ; 0.360]

Notes: Summary statistics across 100 Monte Carlo replications. For each replication, I compute the average. The middle number is the average over replications; lower numbers in brackets are the bounds of the 95% confidence interval.

Merger (Nevo, 2000). Consider now a merger between 2 firms and observe the effects on prices and market shares. The following table shows the results. It appears, for all the DGPs, that the FIL model is able to obtain good predictions of the change in prices and in market shares, even when merged firms are distinguished from the firms not concerned by the merger.

Table 2: Counterfactual Results – Merger Simulation

	Price Effect $\Delta p\%$			Share Effect Δs		
	All Firms	Merging Firms	Others	All Firms	Merging Firms	Others
DGP with $J = 25$ and $T = 100$						
True	3.349	7.170	0.775	-0.149	-12.447	8.135
	[3.300 ; 3.400]	[7.082 ; 7.258]	[0.766 ; 0.784]	[-0.150 ; -0.147]	[-12.531 ;-12.363]	[8.008 ; 8.263]
FIL	3.611	7.680	0.872	-0.152	-12.425	8.114
	[3.531 ; 3.691]	[7.528 ; 7.831]	[0.848 ; 0.896]	[-0.155 ; -0.150]	[-12.501 ;-12.348]	[7.991 ; 8.236]
BLP	3.310	7.126	0.739	-0.147	-12.446	8.138
	[3.241 ; 3.377]	[7.009 ; 7.243]	[0.708 ; 0.770]	[-0.150 ; -0.145]	[-12.531 ;-12.362]	[8.010 ; 8.265]
DGP with $J = 50$ and $T = 200$						
True	3.266	7.009	0.777	-0.072	-13.026	8.543
	[3.246 ; 3.286]	[6.976 ; 7.042]	[0.773 ; 0.780]	[-0.072 ; -0.071]	[-13.062 ; -12.990]	[8.488 ; 8.597]
FIL	3.389	7.277	0.804	-0.074	-13.077	8.574
	[3.361 ; 3.417]	[7.222 ; 7.331]	[0.797 ; 0.810]	[-0.074 ; -0.073]	[-13.109 ; -13.045]	[8.520 ; 8.628]
BLP	3.284	7.046	0.782	-0.072	-13.026	8.543
	[3.258 ; 3.309]	[6.999 ; 7.093]	[0.775 ; 0.789]	[-0.072 ; -0.071]	[-13.062 ; -12.990]	[8.488 ; 8.598]
DGP with $J = 100$ and $T = 20$						
True	3.207	6.890	0.774	-0.035	-13.367	8.774
	[3.151 ; 3.263]	[6.806 ; 6.973]	[0.764 ; 0.785]	[-0.036 ; -0.035]	[-13.458 ; -13.275]	[8.606 ; 8.942]
FIL	3.531	7.410	0.968	-0.032	-13.099	8.602
	[3.454 ; 3.607]	[7.278 ; 7.543]	[0.949 ; 0.987]	[-0.033 ; -0.031]	[-13.185 ; -13.014]	[8.439 ; 8.764]
BLP	3.182	6.846	0.762	-0.035	-13.368	8.774
	[3.116 ; 3.248]	[6.733 ; 6.958]	[0.744 ; 0.779]	[-0.036 ; -0.035]	[-13.459 ; -13.277]	[8.606 ; 8.942]

Notes: Summary statistics across 100 Monte Carlo replications. For each replication, I compute the average. The middle number is the average over replications; lower numbers in brackets are the bounds of the 95% confidence interval.

New product (Petrin, 2002). Lastly, consider the exit of one firm and again observe the effects on prices and market shares. The following table shows the results. For all the DGPs, the FIL model is capable of obtaining good predictions of the change in prices and in market shares. It is also interesting to note that for all DGPs, except the last one, the FIL model does better than the BLP method.

Table 3: Counterfactual Results – New Product

	Price Effect $\Delta p\%$	Share Effect Δs
DGP with $J = 25$ and $T = 100$		
True	3.253	24.954
	[3.200 ; 3.305]	[24.664 ; 25.244]
FIL	3.158	24.421
	[3.042 ; 3.274]	[24.138 ; 24.704]
BLP	2.602	24.852
	[2.542 ; 2.662]	[24.561 ; 25.142]
DGP with $J = 50$ and $T = 200$		
True	2.883	24.996
	[2.860 ; 2.905]	[24.844 ; 25.147]
FIL	2.584	24.688
	[2.563 ; 2.605]	[24.538 ; 24.838]
BLP	2.430	24.934
	[2.404 ; 2.457]	[24.782 ; 25.086]
DGP with $J = 100$ and $T = 20$		
True	2.602	24.690
	[2.548 ; 2.656]	[24.323 ; 25.057]
FIL	3.342	24.535
	[3.272 ; 3.412]	[24.170 ; 24.901]
BLP	2.251	24.657
	[2.189 ; 2.312]	[24.289 ; 25.025]

Notes: Summary statistics across 100 Monte Carlo replications. For each replication, I compute the average. The middle number is the average over replications; lower numbers in brackets are the bounds of the 95% confidence interval.

5 Conclusion

This paper has developed the FIL model, a structural inverse demand model for differentiated products that accommodates rich substitution patterns thanks to a simple linear IV regression with data on market shares, prices and product characteristics.

The FIL model uses a flexible nesting structure with a nest for each pair of products. By contrast with the nested logit model, nesting in the FIL model is just a way to fully parametrize the matrix of price elasticities of demand and does not require the modeller to choose the nesting structure before estimation. This means that the implied cross-price elasticities of demand are not constrained by the model, but instead are driven by the data.

The FIL model is then mapped into characteristics space to make the price elasticities depending on product characteristics directly, as it is the case of the RCL model. Besides, simulation results show that the FIL model is able to mimic the substitution patterns from the RCL model pretty well and to get quite right prediction of a merger's price effects.

Throughout, this paper has considered a mapping into continuous characteristics space. The projection into discrete characteristics (e.g., exploiting product segmentation), into market-level variables (e.g., demographics) and into price spaces is left for future research.

The FIL model can be applied to various topics in industrial organization, international trade, public economics, etc. It can be used to answer relevant policy questions, such as the effect of mergers, products' entry, and changes in regulation. Due to its simplicity of estimation, the likely audience of the FIL model involves researchers as well as antitrust practitioners in consultancies and competition authorities who wish to avoid complex procedures of estimation and/or who are under time pressure.[24]

Appendix A Proofs of Section 2

A.1 Proof of Proposition 1

Let $\Delta_J \equiv \left\{ \mathbf{s} \in [0,\infty)^{J+1} : \sum_{j=1}^{J} s_j < 1 \right\}$. Then, Proposition 1 can be extended as follows.

Proposition 4. Consider the function $\mathbf{G} = (G_1, \ldots, G_J) : [0,\infty)^J \to [0,\infty)^J$. Assume that $\mathbf{G}$ is continuously differentiable and homogeneous of degree one on $\operatorname{int}(\Delta_J)$ and has a matrix of derivatives that is positive definite and symmetric on $\operatorname{int}(\Delta_J)$. Further assume that the 1-norm $|\ln \mathbf{G}(\mathbf{s})|$ approaches infinity as $\mathbf{s}$ approaches $\operatorname{bd}(\Delta_J)$. Let $\mathbf{f} : [0,\infty)^J \to \mathbb{R}^J$ be defined by $\mathbf{f} = (f_1, \ldots, f_J)$ where $f_j(\mathbf{s}) = \ln G_j(\mathbf{s}) - \ln \left(1 - \sum_{k=1}^{J} s_k\right)$. It follows that $\mathbf{f}$ is invertible on $\operatorname{int}(\Delta_J)$.

Proof. The proof is an application of Theorem 26.5 in Rockafellar (1970) to the pair $(\operatorname{int}(\Delta_J), \Omega)$, where Ω is defined by

$$\Omega(\mathbf{s}) = \begin{cases} \sum_{j=1}^{J} s_j f_j(\mathbf{s}) + \ln\left(1 - \sum_{k=1}^{J} s_k\right) & \text{if } \mathbf{s} \in \Delta_J, \\ +\infty & \text{otherwise.} \end{cases}$$

[24] The nested logit models are commonly used by antitrust practitioners and competition authorities (e.g., the European Commission estimated nested logit models to simulate mergers for the Lagardère/Natexis/VUP (2004), TomTom/Tele Atlas (2008), Unilever/Sara Lee (2010) cases; see CCR - Competition Competence Report Autumn 2013/1) and by academics (see e.g., Björnerstedt and Verboven, 2016; Berry et al., 2016, for recent papers that estimate nested logit models with aggregate data). The FIL model possesses the main features that make it appealing for merger evaluation purposes, as highlighted by Pinkse and Slade (2004). It imposes no specific restriction on the price elasticities; it is easily and fastly estimated by linear IV regression using standard computer softwares; and it can handle very large choice sets.

Note that $\nabla\Omega(\mathbf{s}) = \mathbf{f}(\mathbf{s})$. The proof thus consists in showing that the pair $(\text{int}(\Delta_J), \Omega)$ is a convex function of Legendre type.

Ω is strictly convex on $\text{int}(\Delta_J)$, since its Hessian is equal to $\mathbf{J}^{\mathbf{s}}_{\ln\mathbf{G}}(\mathbf{s}) + \mathbf{1}_{JJ}/s_0$ for any $\mathbf{s} \in \text{int}(\Delta_J)$. Ω is essentially smooth, since it is differentiable through the open convex set $\text{int}(\Delta_J)$ with $\lim_{i\to\infty} |\nabla\Omega\,(\mathbf{s}_i)\,| = +\infty$ whenever $\mathbf{s}_1, \mathbf{s}_2, \ldots$ is a sequence in $\text{int}(\Delta_J)$ converging to a point $\mathbf{s} \in \text{bd}(\Delta_J)$. This latter feature is shown by first noting that $\nabla\Omega(\mathbf{s}) = \mathbf{f}(\mathbf{s})$ for $\mathbf{s} \in \text{int}(\Delta_J)$ and then using that $\lim_{\mathbf{s}\to\text{bd}(\Delta_J)} |\ln\mathbf{G}(\mathbf{s})| = +\infty$. □

Proof of Proposition 4. The proof amounts to show that the FIL model satisfies the assumptions of Proposition 4. Let $G_j(\mathbf{s}) = (s_j)^{1-\sum_{i\neq j}\mu_{ij}} \prod_{i\neq j}(s_i + s_j)^{\mu_{ij}}$, then note that f_j corresponds to the FIL model (3). The matrix of derivatives of $\ln\mathbf{G}$ has entries ij given by

$$\left[\frac{1-\sum_{i\neq j}\mu_{ij}}{s_j} + \sum_{i\neq j}\frac{\mu_{ij}}{s_i+s_j}\right]\mathbf{1}\{i=j\} + \left[\frac{\mu_{ij}}{s_i+s_j}\right]\mathbf{1}\{i\neq j\}.$$

It is easy to show that the 1-norm $|\ln\mathbf{G}\,(\mathbf{s})\,|$ that approaches infinity as $\mathbf{s}$ approaches $\text{bd}\,(\Delta_J)$. It then remains to show that the function $\mathbf{G}$ is homogeneous of degree one and that its derivative matrix is positive definite and symmetric.

It is homogeneous of degree one, since for $\lambda > 0$ and $j = 1, \ldots, J$,

$$\begin{aligned}
G_j(\lambda\mathbf{s}) &= (\lambda s_j)^{1-\sum_{i\neq j}\mu_{ij}} \prod_{i\neq j}[\lambda(s_i+s_j)]^{\mu_{ij}} \\
&= \left[\lambda^{1-\sum_{i\neq j}\mu_{ij}} \prod_{i\neq j}\lambda^{\mu_{ij}}\right]\left[(s_j)^{\mu_j}\prod_{i\neq j}(s_i+s_j)^{\mu_{ij}}\right], \\
&= \left[\lambda^{1-\sum_{i\neq j}\mu_{ij}+\sum_{i\neq j}\mu_{ij}}\right] G_j(\mathbf{s}) = \lambda G_j(\mathbf{s}).
\end{aligned}$$

Furthermore, the derivative matrix of $\ln\mathbf{G}$ is symmetric since Restriction (R2) implies that its entry ij, $\mu_{ij}/(s_i+s_j)$, equals its entry ji, $\mu_{ji}/(s_j+s_i)$.

Lastly, the derivative matrix of $\ln\mathbf{G}$ is positive definite. Let $\mathbf{I}_{i,j}$ be a $(J\times J)$ matrix where $\mathbf{I}_{ij}[i,i] = \mathbf{I}_{ij}[j,j] = \mathbf{I}_{ij}[i,j] = \mathbf{I}_{ij}[j,i] = 1$ and zero otherwise. Let $\mathbf{I}_{i,-j}$ be a $(J\times J)$ matrix where $\mathbf{I}_{ij}[i,i] = \mathbf{I}_{ij}[j,j] = 1$, $\mathbf{I}_{ij}[i,j] = \mathbf{I}_{ij}[j,i] = -1$ and zero otherwise. Let $\mathbf{I}_{[}i]$ be a $(J\times J)$ matrix where $\mathbf{I}_{ij}[i,i] = 1$ and zero otherwise. Then, using Restriction (R2), derivative matrix of $\ln \mathbf{G}$ can be written as

$$\sum_{i=1}^{J} \frac{\left(1 - \sum_{i \neq j} \mu_{ij}\right) \mathbf{I}_i}{s_i} + \sum_{0<i<j, \mu_{ij}>0} \frac{|\mu_{ij}| \mathbf{I}_{i,j}}{s_i + s_j} + \sum_{0<i<j, \mu_{ij}<0} \frac{|\mu_{ij}| \mathbf{I}_{i,-j}}{s_i + s_j},$$

which is positive definite since its first term is, by Restriction (R1), a positive definite matrix, and its second and third terms are two sums of positive semi-definite matrices. □

A.2 Proof of Proposition 2

Assume that one observes the vectors of prices and market shares, **p** and **s**. The proof consists in showing that the FIL model can match that vector of market shares as well as the true matrix of own- and cross-price demand elasticities.

Market Shares. Proposition 1 implies that there exists a unique vector of $\boldsymbol{\delta}$ such that $\boldsymbol{\sigma}(\boldsymbol{\delta}; \boldsymbol{\mu}) = \mathbf{s}$, which shows the first requirement for the FIL model to be Diewert (1974) flexible. In particular, given $\boldsymbol{\mu}$ to match the vector of market shares **s**, one can choose $\boldsymbol{\delta}$ such that

$$\delta_j = \ln\left(\frac{s_j}{s_0}\right) - \sum_{i \neq j} \mu_{ij} \ln\left(\frac{s_j}{s_i + s_j}\right).$$

Price Derivatives. Observing prices and market shares, matching price elasticities amounts to matching price derivatives. Since the matrix of price elasticities is positive definite, this is equivalent to matching the inverse of the matrix of price derivatives that has entries ij given by $(1/\alpha)(\partial \sigma_i^{-1}/\partial s_j)$, where $\partial \sigma_i^{-1}/\partial s_j$ is given by Equation (5).

Consider matching the off-diagonal entries λ_{ij} of the inverse of the true matrix of price derivatives. Then, given α, one can choose μ_{ij} such that

$$\lambda_{ij} = \frac{1}{\alpha}\left(\frac{1}{1 - \sum_{k=1}^{J} s_k} + \frac{\mu_{ij}}{s_i + s_j}\right), \tag{13}$$

$$\Leftrightarrow \mu_{ij} = \left(\alpha \lambda_{ij} - \frac{1}{1 - \sum_{k=1}^{J} s_k}\right)(s_i + s_j). \tag{14}$$

Consider now matching the diagonal entries λ_{jj}. Differentiating $\sum_{k=1}^{J} \sigma_k(\boldsymbol{\delta}) + \sigma_0(\boldsymbol{\delta}) = 1$ implies, for all $j = 1, \ldots, J$, that $\sum_{k=1}^{J} \frac{\partial \sigma_k(\boldsymbol{\delta})}{\partial p_j} + \frac{\partial \sigma_0(\boldsymbol{\delta})}{\partial p_j} = 0$, which can be rearranged as

$\sum_{k=1}^{J} \frac{\partial \sigma_k(\boldsymbol{\delta})}{\partial p_j} = \alpha s_0 s_j$. This rearrangement uses that the relationship between the outside good and any inside product is of the logit form (and justifies the definition of the class in which the FIL model is flexible) and that market shares are matched. Stacking these equations and inverting, one obtains

$$\frac{\mathbf{1}_{JJ}}{\alpha\left(1-\sum_{k=1}^{J}\sigma_k\right)} = \boldsymbol{\lambda}\mathbf{s}, \tag{15}$$

which implies that all λ_{jj}'s are matched.

Lastly, one can choose a value for α so that restriction (R1) is satisfied.

A.3 Proof of Proposition 3

Set $\mu_{ij} = \mu\left(\mathbf{d}_{ij}^{(2)}\left(\mathbf{x}_i^{(2)}, \mathbf{x}_j^{(2)}\right)\right)$. Then, the FIL model (3) can be rewritten as

$$\begin{aligned}
\sigma_j^{-1}\left(\mathbf{s}, \mathbf{x}^{(2)}\right) &= \sigma_j^{-1}\left(\mathbf{s}, \mathbf{d}^{(2)}\left(\mathbf{x}^{(2)}\right)\right) \\
&= \ln\left(\frac{s_j}{1-\sum_{k=1}^{J} s_k}\right) - \sum_{i \neq j} \mu\left(\mathbf{d}_{ij}^{(2)}\left(\mathbf{x}_i^{(2)}, \mathbf{x}_j^{(2)}\right)\right) \ln\left(\frac{s_j}{s_i + s_j}\right), \qquad (16) \\
&= \ln\left(\frac{s_j}{1-\sum_{k=1}^{J} s_k}\right) - \sum_{i=1}^{J} \mu\left(\mathbf{d}_{ij}^{(2)}\left(\mathbf{x}_i^{(2)}, \mathbf{x}_j^{(2)}\right)\right) \ln\left(\frac{s_j}{s_i + s_j}\right) + C, \qquad (17)
\end{aligned}$$

where $C = \mu(\mathbf{0}) \ln(1/2) \in \mathbb{R}$.

Observe in Equation (17) that the sum is over all products $i = 1, \ldots, J$, including product j itself, and that the constant C is product-invariant. This shows that the FIL model projected into characteristics space is symmetric and anomynous.

Lastly, using Equation (16) shows that σ_j^{-1} is invariant to translation in $\mathbf{x}^{(2)}$ since, for all $c \in \mathbb{R}$,

$$\sigma_j^{-1}\left(\mathbf{s}, \mathbf{x}^{(2)} + c\mathbf{1}\right) = \sigma_j^{-1}\left(\mathbf{s}, \mathbf{d}^{(2)}\left(\mathbf{x}^{(2)} + c\mathbf{1}\right)\right) = \sigma_j^{-1}\left(\mathbf{s}, \mathbf{d}^{(2)}\left(\mathbf{x}^{(2)}\right)\right) = \sigma_j^{-1}\left(\mathbf{s}, \mathbf{x}^{(2)}\right).$$

References

ALLEN, R. AND J. REHBECK (2019): "Identification With Additively Separable Heterogeneity," *Econometrica*, 87, 1021–1054.

ARMSTRONG, T. B. (2016): "Large market asymptotics for differentiated product demand estimators with economic models of supply," *Econometrica*, 84, 1961–1980.

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

BEN-AKIVA, M. E. (1973): "Structure of passenger travel demand models." Ph.D. thesis, Massachusetts Institute of Technology.

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 (2014): "Identification in differentiated products markets using market level data," *Econometrica*, 82, 1749–1797.

——— (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.

——— (1999): "Voluntary export restraints on automobiles: Evaluating a trade policy," *American Economic Review*, 89, 400–430.

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

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

BRENKERS, R. AND F. VERBOVEN (2006): "Liberalizing a distribution system: the European car market," *Journal of the European Economic Association*, 4, 216–251.

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*, S17–S44.

CHU, C. (1989): "A Paired Combinational Logit Model for Travel Demand Analysis," in *Transport Policy, Management and Technology Towards 2001: Selected Proceedings of the Fifth World Conference on Transport Research*, Vol. 4, Western Periodicals, Ventura, CA, 295–309.

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," *Forthcoming at The RAND Journal of Economics*.

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.

DIEWERT, W. (1974): "Application of Duality Theory," *Frontiers of Quantitative Economics*, 2.

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., J. MONARDO, AND A. DE PALMA (2020): "The Inverse Product Differentiation Logit Model," *Available at SSRN: https://ssrn.com/abstract=3141041*.

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.

GOLDBERG, P. K. (1995): "Product differentiation and oligopoly in international markets: The case of the US automobile industry," *Econometrica*, 891–951.

GRIFFITH, R., M. O'CONNELL, AND K. SMITH (2019): "Tax design in the alcohol market," *Journal of Public Economics*, 172, 20–35.

GRIGOLON, L. (2019): "Blurred boundaries: a flexible approach for segmentation applied to the car market," *Unpublished, available on author's website*.

HAUSMAN, J., G. LEONARD, AND J. D. ZONA (1994): "Competitive analysis with differenciated products," *Annales d'Economie et de Statistique*, 159–180.

HURWICZ, L. (1966): "On the structural form of interdependent systems," in *Studies in Logic and the Foundations of Mathematics*, Elsevier, vol. 44, 232–239.

KNITTEL, C. R. AND K. METAXOGLOU (2014): "Estimation of Random-Coefficient Demand Models: Two Empiricists' Perspective," *Review of Economics and Statistics*, 96, 34–59.

KOPPELMAN, F. S. AND C.-H. WEN (2000): "The paired combinatorial logit model: properties, estimation and application," *Transportation Research Part B: Methodological*, 34, 75–89.

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.

MCFADDEN, D. (1978): "Modeling the choice of residential location," *Transportation Research Record*.

MCFADDEN, D. AND K. TRAIN (2000): "Mixed MNL models for discrete response," *Journal of applied Econometrics*, 15, 447–470.

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

MONARDO, J. (2020): "Shape Restrictions for Structural Demand Estimation," *Available at SSRN:https://ssrn.com/abstract=3474330*.

NEVO, A. (2000): "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *The RAND Journal of Economics*, 31, 395–421.

——— (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307–342.

OGAKI, M. (1990): "The indirect and direct substitution effects," *The American Economic Review*, 80, 1271–1275.

PETRIN, A. (2002): "Quantifying the benefits of new products: The case of the minivan," *Journal of political Economy*, 110, 705–729.

PINKSE, J. AND M. E. SLADE (1998): "Contracting in space: An application of spatial statistics to discrete-choice models," *Journal of Econometrics*, 85, 125–154.

——— (2004): "Mergers, brand competition, and the price of a pint," *European Economic Review*, 48, 617–643.

PINKSE, J., M. E. SLADE, AND C. BRETT (2002): "Spatial price competition: a semiparametric approach," *Econometrica*, 70, 1111–1153.

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

ROCKAFELLAR, R. T. (1970): *Convex analysis*, Princeton university press.

ROJAS, C. (2008): "Price competition in US brewing," *The Journal of Industrial Economics*, 56, 1–31.

SALANIÉ, B. AND F. A. WOLAK (2019): "Fast," robust", and approximately correct: estimating mixed demand systems," *National Bureau of Economic Research*.

SAMUELSON, P. A. (1974): "Complementarity: An essay on the 40th anniversary of the Hicks-Allen revolution in demand theory," *Journal of Economic literature*, 12, 1255–1289.

SLADE, M. E. (2004): "Market power and joint dominance in UK brewing," *The Journal of Industrial Economics*, 52, 133–163.

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

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*, 240–268.

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