Estimation and Counterfactual Analysis in Differentiated Product Markets Without Price Data

Victor Aguirregabiria [*] Robert Clark [†] Eduardo Souza-Rodrigues [‡]

June 28, 2026

Abstract

We develop a methodological framework for structural estimation and counterfactual analysis in differentiated-product markets that does not require observing prices. Empirical analyses of oligopolistic competition typically rely on price data to identify demand, recover marginal costs, and evaluate market power and merger effects. In many important industries, however, prices are unavailable, confidential, aggregated, or measured with error. Our approach builds on the standard random-coefficients BLP framework with Bertrand competition. By substituting the equilibrium pricing equations into the demand system, we derive equilibrium conditions that eliminate prices, leaving market shares as the only endogenous variables. The key insight is that these conditions identify the structural parameters determining product-level social surpluses and demand substitution patterns using only market shares and exogenous product characteristics, without requiring price data. Importantly, these objects are sufficient to identify markups, profits, consumer surplus, diversion ratios, and merger effects. We illustrate the approach using data from the European automobile industry and show that it is robust to price aggregation and measurement error.

Keywords: Differentiated products; BLP; Welfare analysis; Mergers; Oligopoly; Market shares; Noisy prices; Identification; Counterfactuals.

JEL Codes: L13, L21, L40, D43, C51.

[*]University of Toronto and CEPR. victor.aguirregabiria@utoronto.ca.
[†]University of Toronto rob.clark@utoronto.ca.
[‡]University of Toronto e.souzarodrigues@utoronto.ca.

1 Introduction

Structural models of differentiated product demand and oligopolistic competition have become a central tool in empirical industrial organization. Since the seminal contributions of Berry (1994), Berry et al. (1995), Nevo (2001), and Petrin (2002), the combination of discrete-choice demand systems with Bertrand pricing has been widely used to analyze market power, merger effects, and welfare consequences of policy interventions across a broad range of industries. A defining feature of this literature is its reliance on price data. Prices are used to identify consumers' marginal utility of income, to recover marginal costs, and to discipline firms' strategic interactions. In many empirical applications, however, price data are either unavailable, or of questionable quality, or observed only at a highly aggregated level that ignores price dispersion or discrimination across markets or groups of consumers.

For instance, in health insurance and hospital markets, many studies do not observe the negotiated prices between hospitals and insurers, which are typically confidential (Ho, 2009; Gowrisankaran et al., 2015; Ho and Lee, 2017). In studies of bank competition, interest rates on loan and deposit products are often observed only at a highly aggregated level –typically at the national bank level and pooled across multiple products– masking substantial variation across locations, customers, and products (Egan et al., 2017; Aguirregabiria et al., 2025). Similarly, in media markets such as radio broadcasting, advertisers commonly negotiate discounts from posted rates. As a result, researchers typically observe only posted advertising rates rather than the actual transaction prices at the station-advertiser level –or even at the station level (Berry and Waldfogel, 1999; Sweeting, 2013). Even in industries that have been extensively studied in empirical IO, such as automobiles or airlines, most applications rely on average prices that obscure substantial heterogeneity in transaction-level prices across consumers, products, and markets.

This paper develops a new methodological framework that enables researchers to estimate structural models and conduct welfare analysis in differentiated product markets without observing prices, or when the available price data are noisy or highly aggregated. Our approach also provides a test of whether observed price data are contaminated by economically significant measurement error.

In equilibrium, market shares must simultaneously satisfy the demand system implied by consumer preferences and the pricing equations implied by firms' oligopolistic behavior. Substituting the pricing equations into the demand system eliminates prices and yields a system of equations in which market shares are the only endogenous variables. This system links market shares directly to the model's exogenous variables and structural parameters, establishing a mapping between market shares and product-level unit social surpluses, defined as the difference between consumers' average willingness to pay and marginal cost. This result relies critically on two assumptions: the indirect utility function is quasi-linear in prices, and the price coefficient (the marginal utility of money) does not vary with unobserved consumer heterogeneity.

This structure has two important implications. First, it identifies the structural parameters governing product-level social surpluses and the distribution of random coefficients using only market shares and exogenous product characteristics. Instrumental variables remain necessary to address the endogeneity of market shares, but they are no longer required to identify the price coefficient. Consequently, the approach weakens the instrumental variable power requirements in demand estimation. Second, equilibrium market shares can be recovered for any hypothetical vector of social surpluses. This feature enables a wide range of counterfactual analyses using the estimated model, including product entry and exit, changes in product characteristics, taxes, and mergers. These results apply to a broad class of equilibrium models in differentiated product markets, including models with random coefficients, multi-product firms, and alternative forms of oligopolistic competition.

We illustrate the empirical relevance of our estimation approach using data from the European automobile industry, a canonical setting in the demand estimation literature. The empirical application is designed to assess whether our method, implemented without using price data, can recover structural parameters of social surpluses and demand substitution that are economically and statistically indistinguishable from those obtained when prices are correctly measured. It also evaluates the robustness of our approach to price measurement error by comparing its performance with conventional estimation based on contaminated or highly aggregated prices. The automobile data are particularly well suited for this purpose because prices can be constructed at different levels of aggregation (e.g., model-country-year, model-year, or brand-year), providing a natural environment to assess the consequences of price aggregation and measurement error for structural estimation.

The results show a remarkably close correspondence between the estimates obtained with and without price data. The two nesting parameters, which govern substitution patterns, differ by only 2.3% and 6.7%, respectively, while the estimated effects of product characteristics on social surpluses are similarly stable across the two approaches. Although the large sample size makes many of these differences statistically significant, they are economically small and imply virtually identical substitution patterns and product rankings. These findings provide direct evidence that equilibrium market shares contain sufficient information to identify the demand-side objects required for welfare analysis, even when prices are unavailable.

The estimated substitution parameters also generate nearly identical measures of market power and counterfactual outcomes. After normalizing for the scale indeterminacy inherent in the no-price approach, the estimated price-cost margins obtained with and without price data exhibit an almost one-to-one relationship, with an R^2 exceeding 0.99, and the corresponding merger simulations produce nearly identical predictions for changes in market shares and price-cost margins. By contrast, when conventional estimation is based on geographically aggregated or otherwise contaminated prices, the resulting substitution parameters, price-cost margins, and merger predictions differ substantially from those obtained using correctly measured prices. These comparisons show that the proposed approach not only reproduces the results of conventional structural estimation when reliable price data are available, but can also outperform standard methods when observed prices are noisy or highly aggregated, while providing a practical specification test for economically important price measurement error.

We further evaluate the proposed approach by replicating the merger simulation in Björnerstedt and Verboven (2014) for the acquisition of General Motors by Volkswagen. The merger counterfactuals obtained without price data closely match those from the conventional price-based approach, predicting nearly identical changes in firm market shares and identifying the same competitors as the primary beneficiaries of the merger. Both methods imply that the merging firms lose market share as higher post-merger markups induce consumers to substitute toward rival products. These results demonstrate that our approach accurately recovers the demand substitution patterns that govern merger outcomes, enabling reliable counterfactual analysis without observing prices.

Related literature: A growing literature has developed methods to recover competitive and welfare-relevant quantities from alternative or incomplete data sources. One group of papers addresses settings where observed product characteristics are too coarse or incomplete to discipline substitution patterns. Magnolfi et al. (2025) augment conventional demand estimation with crowd-sourced "triplet" survey data to construct an embedding of the latent product space that replaces or supplements observed characteristics in mixed logit models. Armona et al. (2021) use search data, jointly recovering latent product characteristics and consumer preferences from web browsing histories and combining them with price and quantity data to estimate demand and simulate a hotel merger. Conlon et al. (2026) propose a semi-parametric estimator defined in product space that requires neither price variation nor characteristics to explain substitution patterns, recovering substitution from first- and second-choice probabilities by restricting the rank of the substitution matrix. Einav et al. (2026) show that when estimating a full demand system is prohibitive, customer overlap provides a tractable proxy for diversion ratios readable directly from panel data. A distinct, but related missing-data problem in the BLP framework arises when the outside good share is unobserved. Bugni et al. (2026) show that structural parameters are then only partially identified and derive sharp identified sets using moment inequality methods.

An alternative approach to the missing-price problem is conjoint analysis, which bypasses price data by eliciting preferences through surveys in which respondents choose among hypothetical products with experimentally varied prices and attributes (Allenby et al., 2019). Because prices are randomized in the survey design, conjoint avoids the endogeneity problems that motivate instrumental variable estimation in standard demand models, and it has been widely used in antitrust contexts to estimate willingness to pay and simulate merger effects when market data are limited. The main limitation is that stated-preference estimates are subject to hypothetical bias: respondents do not face real budget constraints, and willingness-to-pay estimates can differ substantially from those recovered from observed market behavior.

Other solutions have been proposed for markets in which prices are individually negotiated and the researcher observes only the transaction price of the chosen product, not the full set of prices offered to the buyer. Allen et al. (2014, 2019) study mortgage markets, where negotiated prices are not systematically recorded across competing offers. Merger effects are recovered through a search and bargaining model identified from choice patterns rather than price menus.[1]

This paper also contributes to the large literature on merger simulation (Hausman et al., 1994; Werden and Froeb, 1994; Nevo, 2000; Werden and Froeb, 2002). Standard implementations require price data to estimate demand and recover marginal costs; prominent empirical applications include Miller and Weinberg (2017) and the retrospective evaluations in Peters (2006), Weinberg (2011), and Weinberg and Hosken (2013). A related literature develops sufficient statistics for evaluating merger effects without solving for a full post-merger equilibrium, including compensating marginal cost reductions (Werden, 1996), upward pricing pressure (Farrell and Shapiro, 2010), and pass-through based approaches (Miller et al., 2013; Miller and Sheu, 2021).

Closest to the present paper is Koh (2025), who demonstrates that merger effects can be

[1] See also Miller (2014) for a related approach in procurement settings.

evaluated without observing prices if one observes revenues, margins, and revenue diversion ratios. Our contribution is fundamentally different. We show that equilibrium market shares themselves identify the substitution patterns and social-surplus objects needed to evaluate mergers, and to recover diversion ratios and markups. Thus, we show that the sufficient statistics that Koh takes as inputs are endogenous outcomes of the equilibrium system and can be recovered without price data.

Outline: The remainder of the paper is organized as follows. Section 2 presents the model, derives the equilibrium mapping between market shares and social surpluses, and provides closed-form examples. Section 3 discusses identification and estimation of structural parameters and counterfactual analysis. Section 4 presents results from an empirical application. Section 5 concludes.

2 Model

2.1 Consumers and demand

Consider a market with J differentiated products indexed by $j \in \mathcal{J} = \{1, 2, \ldots, J\}$, and an outside option $j = 0$. A continuum of consumers choose the product that maximizes utility. Utility from product j for consumer i is

$$u_{ij} \;=\; \boldsymbol{x}_j' \boldsymbol{\beta}_i - \alpha\, p_j + \xi_j + \varepsilon_{ij}, \tag{1}$$

where $\boldsymbol{x}_j$ is a K dimension vector of observed product characteristics, $\boldsymbol{\beta}_i$ represents the marginal utilities of these product attributes for consumer i, p_j is the product price, ξ_j captures unobserved product quality, α is the marginal utility of money that is assumed constant across consumers, and ε_{ij} is consumer i's idiosyncratic taste for product j which is unobservable to the researcher.

The utility specification in equation (1) follows the canonical formulation of Berry et al.

(1995), with the benchmark assumption that the marginal utility of income is common across consumers, i.e., $\alpha_i = \alpha$ for every consumer i. Although this assumption is convenient for the exposition and plays a central role in the derivation of our main results, it should not be interpreted as a fundamental limitation of the framework. Rather, it is a sufficient condition for our identification argument. The framework extends to richer forms of consumer heterogeneity whenever the available quantity data permit the separate identification of the distribution of price sensitivity. For example, if market shares are observed separately for different demographic groups, our approach allows the price coefficient to vary freely across groups, with the identification results applying group by group.

The demand model is completed by specifying the distribution of consumer heterogeneity in $\boldsymbol{\beta}_i$ and ε_{ij}. The vector of random coefficients $\boldsymbol{\beta}_i$ has a known distribution up to the vector of means $\boldsymbol{\beta}'$ and the variance-covariance matrix $\boldsymbol{\Sigma}$. That is, $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\nu}_i$, where $\boldsymbol{\nu}_i$ is a K-dimension vector of i.i.d. random variables with zero-mean and known distribution F up to the variance-covariance matrix $\boldsymbol{\Sigma}$.

For the variables ε_{i1}, ε_{i2}, ..., ε_{iJ} capturing consumers' idiosyncratic taste for each product, we consider a nested logit model, which encompasses as a particular case the standard logit model. The $J+1$ products are partitioned into $G+1$ mutually exclusive groups indexed by $g \in \{0, 1, ..., G\}$. We denote by g_j the group to which product j belongs. The outside good is the single element of group $g = 0$. Then,

$$\varepsilon_{ij} = \varepsilon_{i,g_j}^{(1)} + (1-\sigma)\, \varepsilon_{ij}^{(2)}, \tag{2}$$

where variables $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$ are independently distributed, with ε_{ij}^{2} being i.i.d. Type I extreme value, $\sigma \in [0,1]$ is a parameter, and ε_{i,g_j}^{1} is i.i.d. with a $C(\sigma)$ distribution as defined in Cardell (1997). This specification becomes the standard logit model when $\sigma = 0$.

We use $\boldsymbol{\sigma}$ to represent the vector of parameters in variance matrix $\boldsymbol{\Sigma}$ together with parameter σ in the nested logit specification. Following the standard notation, let δ_j represent the indirect utility of product j for the average consumer in the market.

$$\delta_j \;\equiv\; \mathbb{E}_{\nu_i,\varepsilon_i}\left(u_{ij}\right) \;=\; \boldsymbol{x}_j'\boldsymbol{\beta} \;-\; \alpha\, p_j \;+\; \xi_j. \tag{3}$$

We refer to δ_j as the *average utility* of product j. The vector of average utilities across the J products is denoted by $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_J)'$. Similarly, $\boldsymbol{x}$ represents the vector with the observable characteristics of all the products, $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2', \ldots, \boldsymbol{x}_J')'$.

Each consumer purchases at most one unit of the differentiated product (unit demand) and chooses whether to buy a product and, if so, which product to purchase in order to maximize utility. Aggregating these unit-demand choices over the distribution of consumer heterogeneity $(\boldsymbol{\nu}_i, \boldsymbol{\varepsilon}_i)$ determines the market shares, that is, the proportions of consumers purchasing each product. The market share of product j is denoted by s_j.

Proposition 1 establishes a fundamental property of this demand system that plays a central role in the results developed in this paper: market shares depend on prices only through the vector of average utilities $\boldsymbol{\delta}$.

PROPOSITION 1. *The demand system can be represented using the system of equations, for $j \in \mathcal{J}$:*

$$s_j = \mathfrak{s}_j\left(\boldsymbol{\delta},\, \boldsymbol{x},\, \boldsymbol{\sigma}\right) = \int \frac{\exp\left(\frac{\delta_j + \boldsymbol{x}_j'\boldsymbol{\nu}_i}{1-\sigma}\right)}{\sum_{k\in g_j}\exp\left(\frac{\delta_k + \boldsymbol{x}_k'\boldsymbol{\nu}_i}{1-\sigma}\right)}\; \frac{\left[\sum_{k\in g_j}\exp\left(\frac{\delta_k + \boldsymbol{x}_k'\boldsymbol{\nu}_i}{1-\sigma}\right)\right]^{1-\sigma}}{1 + \sum_{g=1}^{G}\left[\sum_{k\in g}\exp\left(\frac{\delta_k + \boldsymbol{x}_k'\boldsymbol{\nu}_i}{1-\sigma}\right)\right]^{1-\sigma}}\; dF(\boldsymbol{\nu}_i). \tag{4}$$

Proof. It follows directly from the model assumptions and the restriction that $\alpha_i = \alpha$ for every consumer i. □

Remark 1.1: In this demand system, observable product characteristics in $\boldsymbol{x}$ enter the demand system in two distinct ways. First, they affect mean utilities directly. Second, they interact with the random coefficients that capture consumer heterogeneity. For this reason, $\boldsymbol{x}$ appears as a separate argument in the demand function $\mathfrak{s}_j(\cdot)$. Prices, by contrast, enter the demand system only through mean utilities, δ. This property is key for the main results in this paper, and specifically, for our Proposition 2.

Remark 1.2: The absence of random coefficients on price is sufficient for prices to enter the demand system only through mean utilities, but it is not necessary. In section **??**, we present a simple variation of the model that allows for consumer heterogeneity in α_i and satisfies the structure $s_j = \mathfrak{s}_j(\delta, x, \sigma)$ in Proposition 1.

2.2 Inverse demand system

Let $\boldsymbol{s} = (s_1, s_2, \ldots, s_J)' \in \mathcal{S}$ be the vector of the J market shares, where $\mathcal{S}$ is J-dimensional simplex. The system of J demand equations in (4) can be represented in vector form as:

$$\boldsymbol{s} \;=\; \mathfrak{s}\left(\boldsymbol{\delta},\, \boldsymbol{x},\, \boldsymbol{\sigma}\right), \tag{5}$$

where $\mathfrak{s}$ is the vector-valued mapping that stacks the J demand functions $\mathfrak{s}_j$. As established in Berry (1994), the demand system in (5) is invertible in δ such that we can obtain the inverse demand equations, for $j \in \mathcal{J}$:

$$\delta_j = \mathfrak{s}_j^{-1}\left(\boldsymbol{s},\, \boldsymbol{x},\, \boldsymbol{\sigma}\right), \tag{6}$$

where $\mathfrak{s}_j^{-1}$ represents the $j-th$ element of the vector-valued inverse mapping, $\mathfrak{s}^{-1}(\boldsymbol{s},\, \boldsymbol{x},\, \boldsymbol{\sigma})$.

By Proposition 1, the inverse mapping, $\mathfrak{s}_j^{-1}(\boldsymbol{s},\, \boldsymbol{x},\, \boldsymbol{\sigma})$, does not depend on prices. Therefore, given the definition of δ_j as $\boldsymbol{x}_j'\boldsymbol{\beta} + \xi_j - \alpha\, p_j$, and solving for price, we have the following system of inverse demand equations:

$$p_j \;=\; \frac{\boldsymbol{x}_j'\boldsymbol{\beta} + \xi_j}{\alpha} - \frac{1}{\alpha}\, \mathfrak{s}_j^{-1}\left(\boldsymbol{s},\, \boldsymbol{x},\, \boldsymbol{\sigma}\right). \tag{7}$$

Let v_j represent the average consumer willingness to pay for a unit of product j, in dollar amounts. By definition,

$$v_j \;=\; \frac{\boldsymbol{x}_j'\boldsymbol{\beta} + \xi_j}{\alpha}. \tag{8}$$

Let $\boldsymbol{v}$ be the vector of average consumer willingness to pay across the J products, $(v_1, v_2, \ldots, v_J)'$. Combining equations (7) and (8), we can write the inserve demand system in vector form as:

$$\boldsymbol{p} \;=\; \boldsymbol{v} - \frac{1}{\alpha}\,\mathfrak{s}^{-1}\left(\boldsymbol{s},\, \boldsymbol{x},\, \boldsymbol{\sigma}\right). \tag{9}$$

2.3 Firms and competition

Products are supplied by multiproduct firms competing within a market. Our approach applies to a broad class of models of competition with differentiated products, including Bertrand and Cournot competition, as well as collusive or cooperative behavior when firms compete in prices or quantities. The framework can also accommodate increasing marginal costs and economies of scope across products. For now, we focus on the standard case in this literature: Bertrand competition with constant marginal costs.

We index firms by f and use $\mathcal{J}_f \subset \mathcal{J}$ to represent the set of products supplied by firm f. A firm's profit is:

$$\pi_f \;=\; \sum_{j \in \mathcal{J}_f} \left(p_j - c_j\right)\; \mathfrak{s}_j(\boldsymbol{p}), \tag{10}$$

where c_j denotes the unit cost of product j, which is an unrestricted function of observable product characteristics $\boldsymbol{x}_j$ and unobserved cost components represented by ω_j: $c_j = c(\boldsymbol{x}_j, \omega_j)$. The first-order condition for profit maximization with respect to the price of product j is:

$$s_j \;+\; \sum_{k \in \mathcal{J}_f} \left(p_k - c_k\right)\, \frac{\partial \mathfrak{s}_k}{\partial p_j} \;=\; 0. \tag{11}$$

We can represent these first-order conditions for every product in matrix form as:

$$\boldsymbol{s} + \left(\boldsymbol{\Omega} \circ \frac{\partial \boldsymbol{\mathfrak{s}}'}{\partial \boldsymbol{p}} \right) (\boldsymbol{p} - \boldsymbol{c}) \; = \; \boldsymbol{0}, \tag{12}$$

where $\boldsymbol{s}$, $\boldsymbol{p}$, and $\boldsymbol{c}$ are $J \times 1$ vectors of shares, prices, and unit costs, respectively; $\frac{\partial \boldsymbol{\mathfrak{s}}'}{\partial \boldsymbol{p}}$ is the $J \times J$ Jacobian matrix of the demand system with respect to prices, $\boldsymbol{\Omega}$ is the ownership matrix, with element $\Omega_{jk} = 1$ is products j and k are produced by the same firm, and zero otherwise; and $\circ$ denotes the Hadamard (element-by-element) product.

Under mild regularity conditions –closely related to those ensuring invertibility of the demand system with respect to the vector of average utilities δ– the Jacobian matrix of market shares with respect to prices and the corresponding ownership-adjusted matrix, $\boldsymbol{\Omega} \circ \frac{\partial \boldsymbol{\mathfrak{s}}'}{\partial \boldsymbol{p}}$ are generically non-singular. As a result, the first-order conditions for Bertrand equilibrium can be rewritten as the following system of pricing equations:

$$\boldsymbol{p} - \boldsymbol{c} \; = \; - \left(\boldsymbol{\Omega} \circ \frac{\partial \boldsymbol{\mathfrak{s}}'}{\partial \boldsymbol{p}} \right)^{-1} \boldsymbol{s}. \tag{13}$$

DEFINITION 1. *A Bertrand equilibrium in this model is a pair of market shares and prices,* $(\boldsymbol{s}, \boldsymbol{p})$ *that satisfies the systems of inverse demand equations in* (9) *and Bertrand pricing equations in* (13).

$$\begin{cases} \textit{Inverse demand system:} & \boldsymbol{p} - \boldsymbol{v} \; = \; -\dfrac{1}{\alpha} \, \mathfrak{s}^{-1} \left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma} \right), \\ \textit{Bertrand pricing equations:} & \boldsymbol{p} - \boldsymbol{c} \; = \; - \left(\boldsymbol{\Omega} \circ \dfrac{\partial \boldsymbol{\mathfrak{s}} \left(\boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{\sigma} \right)'}{\partial \boldsymbol{p}} \right)^{-1} \boldsymbol{s}. \end{cases} \quad \blacksquare \tag{14}$$

The available results on existence of Bertrand–Nash price equilibria in differentiated-product models depend on the demand specification and the ownership structure of firms' product portfolios. For single-product firms, Caplin and Nalebuff (1991) establish existence under conditions ensuring profit quasiconcavity, while Mizuno (2003) uses supermodularity arguments for logit and nested-logit demand. For multiproduct firms, Nocke and Schutz

(2018) establish existence and provide conditions for uniqueness for a class of demand systems that includes logit and CES, but excludes the general mixed-logit model. More recently, Garrido (2024) extends these results to general nested-logit and nested-CES demand systems, establishing existence and proposing a simple criterion to assess uniqueness.

2.4 Equilibrium fixed-point mapping in market shares only

We now present a representation of the vector of equilibrium market shares as the solution to a fixed-point mapping that does not explicitly include prices.

By Proposition 1, the demand system is such that prices affect market shares only through the vector of average utilities δ. By the chain rule, and applying Berry's inversion, this implies that the Jacobian demand-price derivatives $\frac{\partial \boldsymbol{\mathfrak{s}}}{\partial \boldsymbol{p}'}$ have the following structure:

$$\begin{aligned} \frac{\partial \boldsymbol{\mathfrak{s}}\left(\boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{\sigma}\right)}{\partial \boldsymbol{p}'} &= \frac{\partial \boldsymbol{\mathfrak{s}}\left(\boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{\sigma}\right)}{\partial \boldsymbol{\delta}'} \frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{p}'} = \frac{\partial \boldsymbol{\mathfrak{s}}\left(\boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{\sigma}\right)}{\partial \boldsymbol{\delta}'} \left(-\alpha\right) \\ &= -\alpha \, \frac{\partial \boldsymbol{\mathfrak{s}}\left(\mathfrak{s}^{-1}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right), \boldsymbol{x}, \boldsymbol{\sigma}\right)}{\partial \boldsymbol{\delta}'} = -\alpha \, \boldsymbol{\Delta}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right). \end{aligned} \tag{15}$$

That is, the demand-price derivatives are proportional to functions that depend only on $(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})$. We use $\boldsymbol{\Delta}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)$ to denote the $J \times J$ matrix of demand derivatives, with generic element $\Delta_{kj} \equiv \frac{\partial \mathfrak{s}_k}{\partial \delta_j}$.

Now, using the equilibrium conditions in (14), obtaining the product-by-product difference between the Bertrand pricing equation and the inverse demand equation, and applying the definition of matrix $\boldsymbol{\Delta}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)$, we obtain the following system of equations:

$$\alpha\left(\boldsymbol{v} - \boldsymbol{c}\right) = \mathfrak{s}^{-1}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right) + \Big[\boldsymbol{\Omega} \circ \boldsymbol{\Delta}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)\Big]^{-1} \boldsymbol{s}. \tag{16}$$

The vector $\boldsymbol{v} - \boldsymbol{c} = (v_1 - c_1, v_2 - c_2, \ldots v_J - c_J)$ contains the unit *social surplus*, in dollars, associated with each product j. Specifically, $v_j - c_j$ is the difference between the average consumer willingness to pay for product j and its unit cost. We refer to this exogenous object as the unit social surplus of product j. Since α is the marginal utility of money, $\alpha(v_j - c_j)$ represents unit social surplus in utils.

Proposition 2 establishes that the system of equations in (16) provides a complete characterization of equilibrium market shares.

PROPOSITION 2. *A vector of prices and market shares* $(\boldsymbol{p}^*, \boldsymbol{s}^*)$ *satisfies the equilibrium equations in* (14) *if and only if the vector of market shares* $\boldsymbol{s}^*$ *solves the system of equations in* (16). ■

Proof. See Appendix A.1. □

The system of equilibrium equations in (16) has two properties that are central to the approach developed in this paper.

i. Given the primitives of the model, the vector of equilibrium market shares can be characterized without explicitly solving for equilibrium prices. As a result, the model delivers a system of equilibrium equations that can be taken directly to data on market shares and product characteristics, even in environments where price data are unavailable. In Section 3, we show that these equilibrium equations can identify important structural parameters, most notably product social surpluses up to scale.

ii. Average consumer willingness to pay, $\boldsymbol{v}$, and unit costs, $\boldsymbol{c}$, enter the equilibrium system only through their difference, $\boldsymbol{v} - \boldsymbol{c}$, which we refer to as the vector of unit social surpluses. Therefore, conditional on observable product characteristics, the only primitives needed to characterize equilibrium market shares are the random-coefficients parameters σ and the vector of social surpluses $\alpha(\boldsymbol{v} - \boldsymbol{c})$. In Section 3.4, we show that this property allows us to implement a rich set of counterfactual experiments using only data on market shares, without requiring price data.

This model may admit multiple equilibria for a given value of the primitives (Nocke and Schutz (2018), Garrido (2024)). This is reflected in the existence of multiple vectors of market shares that solve the system of equations in (16). This equilibrium multiplicity does not affect our identification results for the structural parameters, for essentially the same reason that multiplicity does not undermine identification of demand and cost parameters in standard applications of the BLP model. However, as in the standard BLP applications, multiplicity does matter for the implementation and interpretation of counterfactual experiments, since different equilibria may imply different counterfactual outcomes.

2.5 Examples

2.5.1 Standard logit with single product firms

The standard logit model corresponds to the following restrictions on the general model: $(\boldsymbol{\nu}_i, \sigma) = (\mathbf{0}, 0)$. For this logit model, the equilibrium conditions are:

$$\begin{cases} \textit{Inverse demand:} & p_j - v_j = -\dfrac{1}{\alpha}\left[\ln(s_j) - \ln(s_0)\right], \\ \textit{Bertrand pricing equation:} & p_j - c_j = \dfrac{1}{\alpha}\dfrac{1}{1-s_j}. \end{cases} \tag{17}$$

Combining these equations, in this case product by product, we obtain the following version of the system of equilibrium equations for market shares in (16):

$$\alpha(v_j - c_j) = \ln(s_j) - \ln\left(1 - \sum_{k=1}^{J} s_k\right) + \frac{1}{1-s_j}. \tag{18}$$

This equation shows that equilibrium market shares, combined according to the expression on the right-hand side, identify the unit social surpluses, measured in utils.

We show in Appendix A.2 that the system of equilibrium equations in (18) has a unique equilibrium. Therefore, in the standard logit model, counterfactual experiments are not affected by equilibrium multiplicity.

2.5.2 Nested Logit with single product firms

The nested logit model corresponds to the following restrictions on the general model: $\boldsymbol{\nu}_i = \mathbf{0}$. For this model, the equilibrium equations are:

$$\begin{cases} \textit{Inverse demand:} & p_j - v_j = -\dfrac{1}{\alpha}\left[\ln(s_j) - \ln(s_0) - \sigma \ln(s^w_{j|g})\right], \\ \textit{Bertrand pricing equation:} & p_j - c_j = \dfrac{1}{\alpha}\dfrac{1}{1-(1-\sigma)\,s_j - \sigma\, s^w_{j|g}}, \end{cases} \tag{19}$$

where $s^w_{j|g}$ represents the within-group market share of product j. Combining these equations, product by product, we obtain the following version of the system of equilibrium equations for market shares in (16):

$$\alpha(v_j - c_j) = \ln(s_j) - \ln\left(1 - \sum_{k=1}^{J} s_k\right) - \sigma \ln(s^w_{j|g}) + \frac{1}{1-(1-\sigma)\,s_j - \sigma\, s^w_{j|g}}. \tag{20}$$

For the nested logit model, the system of equilibrium equations for market shares depends on two sets of parameters: the vector of unit social surpluses (measured in utils) and the nesting parameter σ. In contrast to the standard logit model, identification and estimation of these $J+1$ parameters from J market-share equations require additional structure, typically through a specification of social surpluses as functions of observable and unobservable product characteristics. As we show in Section 3, these structural parameters are identified under standard assumptions commonly used in this literature.

2.5.3 Random Coefficients Logit with single product firms

In this model, the random coefficients $\boldsymbol{\nu}_i$ are unrestricted, but nesting parameter σ is zero. For this model, the equilibrium equations are:

$$\left\{ \begin{array}{rrcl} \textit{Inverse demand:} & p_j - v_j & = & -\dfrac{1}{\alpha}\, \mathfrak{s}_j^{-1}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}), \\ \textit{Bertrand pricing equation:} & p_j - c_j & = & \dfrac{1}{\alpha}\, \dfrac{s_j}{\Delta_{jj}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})}, \end{array} \right. \tag{21}$$

with $\Delta_{jj}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}) \equiv \dfrac{\partial \mathfrak{s}_j(\mathfrak{s}^{-1}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}), \boldsymbol{x}, \boldsymbol{\sigma})}{\partial \delta_j}$. Combining these equations, product by product, we obtain the following version of the system of equilibrium equations for market shares in (16):

$$\alpha(v_j - c_j) \;=\; \mathfrak{s}_j^{-1}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}) \;+\; \frac{s_j}{\Delta_{jj}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})}. \tag{22}$$

For this random coefficient model, similarly to the nested logit, the system of equilibrium equations depends on two sets of parameters: the vector of unit social surpluses (measured in utils) and the random coefficient parameter $\boldsymbol{\sigma}$. As we show in Section 3, these structural parameters are identified under standard assumptions in the literature.

2.6 Equilibrium Outcomes as Functions of Social Surpluses

We now show that the information contained in the vector of unit social surpluses, together with product characteristics and the parameters in $(\boldsymbol{x}, \boldsymbol{\sigma})$, is sufficient to identify not only equilibrium market shares, but also other important equilibrium outcomes, including price-cost margins, profits, consumer welfare, and diversion ratios.

2.6.1 Price-cost margins

Define the price-cost margin of the product j, $PCM_j = p_j - c_j$, and let ($\boldsymbol{PCM}$ be the vector of price-cost margins for the J products. Combining the Bertrand pricing equations in (14) with the expression in (15), we have:

$$\alpha\,\boldsymbol{PCM} \;=\; \Big[\boldsymbol{\Omega} \circ \boldsymbol{\Delta}\,(\boldsymbol{s},\,\boldsymbol{x},\,\boldsymbol{\sigma})\Big]^{-1}\boldsymbol{s}. \tag{23}$$

The left-hand side of this equation is the vector of margins measured in utils or, equivalently, in dollars up to the scale parameter α. The expression on the right-hand side is a known function of $(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})$. Because equilibrium market shares are determined solely by $\alpha(\boldsymbol{v} - \boldsymbol{c})$, $\boldsymbol{x}$, and $\boldsymbol{\sigma}$), it follows that price-cost margins, up to scale, are also completely characterized by these same arguments.

2.6.2 Firms' profits

Let π_f be firm f's profit, in dollars, as defined in equation (10). Using its definition, we can write the following expression:

$$\alpha\,\pi_f \;=\; \sum_{j\in\mathcal{J}_f} [\alpha\, PCM_j]\; s_j. \tag{24}$$

The left-hand side of this equation is the firm's profit measured in utils or, equivalently, in dollars up to the scale parameter α. As shown above, the expression on the right-hand side is solely determined solely by $(\alpha(\boldsymbol{v} - \boldsymbol{c}), \boldsymbol{x}, \boldsymbol{\sigma})$. Therefore, firms' profits, up to scale, are only determined by these same arguments.

2.6.3 Consumer welfare

Let CW denote aggregate consumer welfare in equilibrium, measured in dollars. In this model, we have that:

$$\alpha\; CW \;=\; \int \ln\left(1 + \sum_{g=1}^{G}\left[\sum_{k\in g}\exp\left(\frac{\delta_k + \boldsymbol{x}_k'\,\boldsymbol{\nu}_i}{1-\sigma}\right)\right]^{1-\sigma}\right) dF(\boldsymbol{\nu}_i) \;=\; cw\,(\boldsymbol{s},\boldsymbol{x},\boldsymbol{\sigma},) \tag{25}$$

where the second equality follows from the representation of the mean utilities, δ_k, as functions of $(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})$ (Berry's invertibility), implying that consumer welfare can be written as the function $cw\,(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})$. Then, equation (16) and Proposition 2 imply that consumer welfare in equilibrium is determined by $(\alpha(\boldsymbol{v} - \boldsymbol{c}), \boldsymbol{x}, \boldsymbol{\sigma})$.

2.6.4 Diversion ratios

A diversion ratio measures how much demand lost by one product is diverted to another product when the first product becomes less attractive, typically because its price increases or the product disappears. It is a central concept in merger analysis because it captures the closeness of substitution between products. Let D_{jk} denote the diversion ratio from product j to k. It is defined as:

$$D_{jk} \;=\; -\frac{\partial s_k(\boldsymbol{p})/\partial p_j}{\partial s_j(\boldsymbol{p})/\partial p_j}. \tag{26}$$

A high diversion ratio between j and k means the products are close substitutes. A low diversion ratio means consumers leaving j mostly switch elsewhere (other products or the outside option).

Taking into account equation (15), we can rewrite the diversion ratios as:

$$D_{jk} \;=\; -\frac{\Delta_{kj}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})}{\Delta_{jj}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})}. \tag{27}$$

Based on the same argument as above, diversion ratios are determined by $(\alpha(\boldsymbol{v} - \boldsymbol{c}), \boldsymbol{x}, \boldsymbol{\sigma})$.

2.7 Geographic Markets: Price Discrimination vs. Uniform Pricing

There are many industries in which competition –through firm and product entry, capacity choices, and pricing or quantity decisions– takes place at a much narrower geographic level than the industry as a whole, such as a region, province, city, or even a smaller local area. Prominent examples include retail, insurance, banking, airlines, and many other industries. In these industries, many firms operate across multiple geographic markets.

Suppose the researcher observes product-level market shares across M geographic markets indexed by $m \in \mathcal{M} = 1, 2, \ldots, M$. Let s_{jm} denote the market share of product j in local market m. However, price information is available only at a more aggregate geographic level. For simplicity, suppose that prices are observed only at the product level, $\overline{p}_j$. The researcher knows that $\overline{p}_j$ does not represent a uniform price that applies to all local markets but instead it is an average price obtained by aggregating some information (potentially limited and noisy) of product prices across the multiple local markets.

The results derived above apply to each local market separately. In particular, for market m, equation (16) becomes:

$$\alpha_m \left(\boldsymbol{v}_m - \boldsymbol{c}_m\right) \;=\; \mathfrak{s}^{-1}\left(\boldsymbol{s}_m, \boldsymbol{x}_m, \boldsymbol{\sigma}_m\right) \;+\; \left[\boldsymbol{\Omega} \circ \boldsymbol{\Delta}\left(\boldsymbol{s}_m,\, \boldsymbol{x}_m,\, \boldsymbol{\sigma}_m\right)\right]^{-1} \boldsymbol{s}_m, \tag{28}$$

where parameters and exogenous variables can vary freely across markets.

What happens if firms set uniform prices across geographic local markets, while demand remains segmented across those markets? We show in Appendix **??** that, in this case, the equilibrium equations for market shares take the following form:

$$\sum_{m=1}^{M} \left[\boldsymbol{s}_m \;+\; \Big(\boldsymbol{\Omega} \circ \boldsymbol{\Delta}\left(\boldsymbol{s}_m,\, \boldsymbol{x}_m,\, \boldsymbol{\sigma}_m\right)\Big)\Big(\alpha_m \left(\boldsymbol{v}_m - \boldsymbol{c}_m\right) + \mathfrak{s}_m^{-1}\left(\boldsymbol{s}_m, \boldsymbol{x}_m, \boldsymbol{\sigma}_m\right)\Big)\right] \;=\; \mathbf{0}. \tag{29}$$

It is straightforward to see that the system of J equilibrium equations in (29) is the aggregated counterpart of the M systems of J equations in (28). Therefore, any vector of market shares (for every product and market) that is an equilibrium under price discrimination (i.e., solves (28)) also satisfies the equilibrium conditions under uniform pricing in (29). However, the converse is not generally true.

In some industries, there is limited evidence about the geographic level at which firms set prices. Do firms set a uniform product price nationwide? Do they engage in third-degree price discrimination, and if so, at what geographic level? Or do they instead use a hybrid system that combines uniform posted prices with some discretion for local managers to offer promotions or discounts? These questions are difficult to answer when price information is limited or available only at a highly aggregated national level. In Section (3), we show that, even in the absence of price data, the equilibrium conditions in (28) and (29) imply testable restrictions that can be used to distinguish between price discrimination and uniform pricing.

2.8 Other Forms of Competition

We now show that the main result in Proposition 2 extends to other forms of competition. We focus on two models: price competition with conduct parameters; and quantity competition.

2.8.1 Price Competition with Conduct Parameters

Following the literature on price competition with common ownership and its collusive interpretation, we can write firm f's objective function as:[2]

$$\pi_f + \sum_{f' \neq f} \kappa_{ff'} \, \pi_{f'} \; = \; \sum_{g=1}^{N} \kappa_{ff'} \left[\sum_{j \in \mathcal{J}_{f'}} \left(p_j - c_j\right) \; s_j(\boldsymbol{p}) \right], \tag{30}$$

where $\kappa_{ff'} \in [0, 1]$ represents the weight that firm f assigns to firm f''s profit in its objective function, or equivalently, the value to firm f of one dollar of profit earned by competitor f'. By definition, $\kappa_{ff} = 1$.

In the standard model in which firms care only about their own profits, $\kappa_{ff'} = 0$ for every $f' \neq f$. Under perfect collusion among all firms in the market, $\kappa_{ff'} = 1$ for every pair (f, f'). Intermediate values of $\kappa_{ff'}$ capture different degrees or forms of coordination among firms.

Given these objective functions, and under the assumption that firms compete in prices and apply Nash conjecture when making profit maximizing decisions, the first order condition

[2] See the seminal paper by Bresnahan and Salop (1986), the survey of the literature by Backus et al. (2019), and recent work by Backus et al. (2021).

of optimality for firm f and product j is:

$$s_j \;+\; \sum_{f'=1}^{N} \kappa_{ff'} \left[\sum_{k \in \mathcal{J}_{f'}} (p_k - c_k) \; \frac{\partial \mathfrak{s}_k}{\partial p_j} \right] \;=\; 0. \tag{31}$$

We can represent these first-order conditions for every product in vector form as:

$$\boldsymbol{s} \;+\; \left(\boldsymbol{A}\, \boldsymbol{\mathcal{K}}\, \boldsymbol{A}' \circ \frac{\partial \boldsymbol{\mathfrak{s}}'}{\partial \boldsymbol{p}} \right) (\boldsymbol{p} - \boldsymbol{c}) \;=\; \boldsymbol{0}, \tag{32}$$

where $\boldsymbol{A}\boldsymbol{\mathcal{K}}\boldsymbol{A}'$ is a $J \times J$ matrix whose (j,k) element is equal to $\kappa_{ff'}$, where f is the firm that owns product j, and f' is the firm that owns product k.[3]

Using the fact that $\frac{\partial \boldsymbol{\mathfrak{s}}'}{\partial \boldsymbol{p}} = -\alpha\, \boldsymbol{\Delta}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})$, and combining the system of pricing equations in (32) with the inverse demand system, we obtain the following system of equations for market shares:

$$\alpha\, (\boldsymbol{v} - \boldsymbol{c}) \;=\; \mathfrak{s}^{-1}\, (\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}) \;+\; \Big[\boldsymbol{A}\, \boldsymbol{\mathcal{K}}\, \boldsymbol{A}' \circ \boldsymbol{\Delta}\, (\boldsymbol{s},\, \boldsymbol{x},\, \boldsymbol{\sigma}) \Big]^{-1} \boldsymbol{s}. \tag{33}$$

This system of equations has essentially the same structure as the system in (16) for the Bertrand equilibrium. The only difference is that the ownership matrix, $\boldsymbol{\Omega} = \boldsymbol{A}\, \boldsymbol{A}'$, is replaced by the ownership-and-conduct matrix, $\boldsymbol{A}\, \boldsymbol{\mathcal{K}}\, \boldsymbol{A}'$. Equivalently, the difference is that the conduct matrix $\boldsymbol{\mathcal{K}}$ is not restricted to be the identity matrix.

[3] More specifically, $\boldsymbol{A}$ is a $J \times N$ matrix with element $A_{jf} = 1$ if product j belongs to firm f, and 0 otherwise. The matrix $\boldsymbol{\mathcal{K}}$ is an $N \times N$ conduct matrix with element $\mathcal{K}_{ff'} = \kappa_{ff'}$. Therefore, the (j,k) element of $\boldsymbol{A}\boldsymbol{\mathcal{K}}\boldsymbol{A}'$ is $\sum_{f=1}^{N} \sum_{f'=1}^{N} D_{jf}\, \kappa_{ff'}\, D_{kf'}$, which simplifies to $\kappa_{ff'}$, where f is the owner of product j and f' is the owner of product k. The ownership matrix $\boldsymbol{\Omega}$ is a special case: when $\boldsymbol{\mathcal{K}}$ is the identity matrix, we have $\boldsymbol{A}\boldsymbol{\mathcal{K}}\boldsymbol{A}' = \boldsymbol{A}\boldsymbol{A}' = \boldsymbol{\Omega}$.

2.8.2 Competition in Quantities

Suppose that firms compete in quantities and apply Nash conjecture when making profit maximizing decisions. Then, the first order condition of optimality for firm f and product j is:

$$p_j - c_j \ + \sum_{k \in \mathcal{J}_f} \frac{\partial p_k}{\partial s_j} \, s_k \ = \ 0. \tag{34}$$

Remember that the inverse demand function for product j is $p_j(\boldsymbol{s}) = v_j - \frac{1}{\alpha} \, \mathfrak{s}_j^{-1} \left(\boldsymbol{s}, \, \boldsymbol{x}, \, \boldsymbol{\sigma}\right)$, such that $\frac{\partial p_k}{\partial s_j} = -\frac{1}{\alpha} \, \frac{\partial \mathfrak{s}_k^{-1}}{\partial s_j}$. Therefore, we can write the first order conditions as,

$$p_j - c_j \ = \ \frac{1}{\alpha} \sum_{k \in \mathcal{J}_f} \frac{\partial \mathfrak{s}_k^{-1}}{\partial s_j} \, s_k, \tag{35}$$

or in vector form as:

$$\boldsymbol{p} - \boldsymbol{c} \ = \ \frac{1}{\alpha} \, \boldsymbol{\Omega} \circ \frac{\partial \boldsymbol{\mathfrak{s}}^{-1} \left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)'}{\partial \boldsymbol{s}} \, \boldsymbol{s}. \tag{36}$$

By the Inverse Function Theorem:

$$\frac{\partial \boldsymbol{\mathfrak{s}}^{-1} \left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)'}{\partial \boldsymbol{s}} \ = \ \left(\left. \frac{\partial \boldsymbol{\mathfrak{s}} \left(\boldsymbol{\delta}\right)'}{\partial \boldsymbol{\delta}} \right|_{\boldsymbol{\delta} = \boldsymbol{\mathfrak{s}}^{-1}(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma})} \right)^{-1} \ = \ \boldsymbol{\Delta} \left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)^{-1}. \tag{37}$$

Combining this system of pricing equations with the inverse demand equation, we obtain the following system of equilibrium equations for market shares:

$$\alpha \left(\boldsymbol{v} - \boldsymbol{c}\right) \ = \ \mathfrak{s}^{-1} \left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right) \ + \ \boldsymbol{\Omega} \circ \boldsymbol{\Delta} \left(\boldsymbol{s}, \, \boldsymbol{x}, \, \boldsymbol{\sigma}\right)^{-1} \boldsymbol{s}. \tag{38}$$

3 Estimation and Counterfactual Analysis

3.1 Econometric Model

The dataset consists of a cross section of J products with information on $(s_j, \boldsymbol{x}_j)$ for each product. To complete the econometric specification, we model unit costs as a linear index of observed product characteristics plus an additive mean-zero unobservable:

$$c_j \;=\; \boldsymbol{x}_j' \boldsymbol{\gamma} + \omega_j. \tag{39}$$

Combining this specification with the expression for average consumer willingness to pay, $v_j = (\boldsymbol{x}_j' \boldsymbol{\beta} + \xi_j)/\alpha$, implies that unit social surplus also has a linear-index representation with an additive error term:

$$\alpha(v_j - c_j) \;=\; \alpha\left(\frac{\boldsymbol{x}_j' \boldsymbol{\beta} + \xi_j}{\alpha} - \boldsymbol{x}_j' \boldsymbol{\gamma} - \omega_j\right) \;=\; \boldsymbol{x}_j' \boldsymbol{\theta} + \eta_j, \tag{40}$$

where $\boldsymbol{\theta} = \boldsymbol{\beta} - \alpha \boldsymbol{\gamma}$ is the vector of parameters with the marginal social surpluses associated with the characteristics in $\boldsymbol{x}_j$, and the mean-zero error term $\eta_j = \xi_j - \alpha \omega_j$ is the mean-zero unobservable component of unit social surplus.

Plugging this specification of the unit social surpluses into the equilibrium conditions for markets shares in equation (16), we obtain the following regression model which, in general, is partially linear-in-parameters:

$$\Phi_j\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right) \;=\; \boldsymbol{x}_j' \boldsymbol{\theta} + \eta_j, \tag{41}$$

where $\Phi_j\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)$ is a scalar function with the following structure:

$$\Phi_j\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right) \;=\; \mathfrak{s}_j^{-1}\left(\boldsymbol{s}, \boldsymbol{x}_j, \boldsymbol{\sigma}\right) \;+\; \sum_{k \in \mathcal{J}_f} \Delta_{jk}^{-1}\left(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{\sigma}\right)\; s_k, \tag{42}$$

where $\Delta_{jk}^{-1}(\boldsymbol{s},\boldsymbol{x},\boldsymbol{\sigma})$ represents the (j,k)-element in the inverse matrix $\boldsymbol{\Delta}(\boldsymbol{s},\boldsymbol{x},\boldsymbol{\sigma})^{-1}$.

The first term in Φ_j is the inverse demand function, $\mathfrak{s}_j^{-1}(\boldsymbol{s},\boldsymbol{x}_j,\boldsymbol{\sigma})$. This would be the only component of Φ_j if we were estimating the demand equation. In that case, the right-hand side of equation (41) would be $\boldsymbol{x}_j'\boldsymbol{\beta}+\xi_j-\alpha p_j$. The second term in Φ_j, $\sum_{k\in\mathcal{J}_f}\Delta_{jk}^{-1}(\boldsymbol{s},\boldsymbol{x},\boldsymbol{\sigma})\,s_k$, represents the price-cost margin. This would be the only component of Φ_j if we were estimating the pricing equation. In that case, the right-hand side of equation (41) would be $\alpha(p_j-\boldsymbol{x}_j'\boldsymbol{\gamma}-\omega_j)$.

Table 1 illustrates the structure of the function Φ_j under several standard specifications of consumer heterogeneity and firms' product ownership.

Table 1: Form of Function Φ_j for Different Models

Model	**Function Φ_j**
Standard Logit, Single-Product	$\ln(s_j)-\ln(s_0)+\dfrac{1}{1-s_j}$
Standard Logit, Multi-Product	$\ln(s_j)-\ln(s_0)+\dfrac{1}{1-\sum_{k\in\mathcal{J}_f}s_k}$
Nested Logit, Single-Product	$\ln(s_j)-\ln(s_0)-\sigma\ \ln(s_{j\|g}^{w})+\dfrac{1-\sigma}{1-(1-\sigma)\,s_j-\sigma\,s_{j\|g}^{w}}$
Nested Logit, Multi-Product (g indexes groups)	$\ln(s_j)-\ln(s_0)-\sigma\ \ln(s_{j\|g}^{w})+\dfrac{1-\sigma}{1-(1-\sigma)\,S_f-\sigma\,S_{f\|g}}$ where: $S_f=\sum_{k\in\mathcal{J}_f}s_k$, and $S_{f\|g}=\sum_{k\in\mathcal{J}_f\cap g}s_{k\|g}^{w}$
Two-Tier Nested Logit, Single-Product (g indexes groups, h subgroups)	$\ln(s_j)-\ln(s_0)-\sigma_1\ \ln(s_{j\|hg})-\sigma_2\ \ln(s_{h\|g})$ $+\dfrac{(1-\sigma_1)(1-\sigma_2)}{(1-\sigma_2)-(1-\sigma_1)(1-\sigma_2)s_j-(\sigma_1-\sigma_2)s_{j\|hg}-\sigma_2(1-\sigma_1)s_{j\|hg}s_{h\|g}}$
Two-Tier Nested Logit, Multi-Product (g indexes groups, h subgroups)	$\ln(s_j)-\ln(s_0)-\sigma_1\ \ln(s_{j\|hg})-\sigma_2\ \ln(s_{h\|g})$ $+\dfrac{(1-\sigma_1)(1-\sigma_2)}{(1-\sigma_2)-(1-\sigma_1)(1-\sigma_2)S_f-(\sigma_1-\sigma_2)S_{f\|hg}-\sigma_2(1-\sigma_1)S_{f\|hg}s_{h\|g}}$ where: $S_f=\sum_{k\in\mathcal{J}_f}s_k$, and $S_{f\|hg}=\sum_{k\in\mathcal{J}_f\cap g}s_{k\|hg}^{w}$

3.2 GMM Estimation

In versions of the model that include parameters $\boldsymbol{\sigma}$ –such as nested logit or random-coefficients models– these parameters enter the regression through market shares. This creates an endogeneity problem because market shares are correlated with the error term η_j. Intuitively, products with higher unobserved social surplus tend to have larger own market shares and smaller market shares for competing products. To address this endogeneity problem, we use the standard instrumental-variables approach in this literature, namely the BLP instruments.

A standard assumption in this literature is that demand and cost unobservables, (ξ_j, ω_j), and therefore η_j, are mean independent of the observed product characteristics $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_J)$. Under this assumption, the characteristics of products other than j can be used as instruments for market shares in the regression equation for product j.

For instance, consider the regression equation in the Nested Logit model with multi-product firms:

$$\ln(s_j) - \ln(s_0) \;=\; \sigma \ln(s^w_{j|g}) - \frac{1}{1 - (1-\sigma) \sum_{k \in \mathcal{J}_f} s_k - \sigma \sum_{k \in \mathcal{J}_f \cap g} s^w_{k|g}} + \boldsymbol{x}_j' \boldsymbol{\theta} + \eta_j. \tag{43}$$

This equation includes three endogenous regressors: the log-within market shares, $\ln(s^w_{j|g})$; the sum of the market shares for all the firm's products, $\sum_{k \in \mathcal{J}_f} s_k$; and the the sum of the within-group market shares for all the firm's products in the same group as product j, $\sum_{k \in \mathcal{J}_f \cap g} s^w_{k|g}$. In our empirical application in Section 4, we estimate this regression using the following sets of instruments (in addition to the own product characteristics $\boldsymbol{x}_j$. For every product characteristic, say x_{1j}:

i. The ratio $\dfrac{x_{1j}}{\sum_{k \in g} x_{1k}}$, intuitively as an instrument for $\ln(s^w_{j|g})$.

ii. The ratio $\dfrac{\sum_{k \in \mathcal{J}_f} x_{1k}}{\sum_{k=1}^{J} x_{1k}}$, intuitively as an instrument for $\sum_{k \in \mathcal{J}_f} s_k$.

iii. The ratio $\dfrac{\sum_{k \in \mathcal{J}_f \cap g} x_{1k}}{\sum_{k \in g} x_{1k}}$, intuitively as an instrument for $\sum_{k \in \mathcal{J}_f \cap g} s^w_{k|g}$.

The GMM estimation of this regression model closely parallels the standard BLP estimation of demand. The key difference is that the price term, αp_j, in the demand equation is replaced by a parametric function of market shares and parameters $\boldsymbol{\sigma}$ that captures the equilibrium price-cost margin of the product. The parameters $\boldsymbol{\sigma}$ already enter the regression through the inverse-demand term. Therefore, replacing prices with the equilibrium margin term does not introduce any additional structural parameters to be estimated.

This feature of our approach may substantially facilitate the identification and estimation of the structural parameters $(\boldsymbol{\sigma}, \boldsymbol{\theta})$. Although the data do not include prices, our approach does not attempt to separately identify demand and cost parameters, but only the parameters governing social surplus. As a result, identification does not require separate sources of instrumental-variables variation to disentangle $\boldsymbol{\sigma}$ from the price-sensitivity parameter α. As shown by Gandhi and Houde (2019), when price instruments are weak, this joint identification fails. Avoiding this difficult separation problem can significantly strengthen identification.

3.3 Identification of Equilibrium Outcomes

Once the structural parameters $(\boldsymbol{\sigma}, \boldsymbol{\theta})$ have been consistently estimated, the unobserved component of product j's unit social surplus, η_j, can be recovered as the residual from the regression equation. Consequently, we obtain consistent estimates of each product's unit social surplus: $\alpha\widehat{(v_j - c_j)} = \boldsymbol{x}_j'\widehat{\boldsymbol{\theta}} + \widehat{\eta}_j$.

As established in Section 2.6, the equilibrium of the model implies that several economically important outcomes are identified solely by the vector of unit social surpluses $\alpha(\boldsymbol{v} - \boldsymbol{c})$, product characteristics $\boldsymbol{x}$, and the demand parameters $\boldsymbol{\sigma}$). Therefore, once these objects have been identified, the following equilibrium outcomes are also identified:

a. Price-cost margins, up to scale: αPCM_j for every product j.

b. Firms' profits, up to scale: $\alpha\pi_f$ for every firm f.

c. Consumer welfare, up to scale: αCW.

d. Diversion ratios: D_{jk} for every pair of products (j,k).

Thus, our approach identifies key equilibrium objects required for welfare analysis and counterfactual policy evaluation, all without observing prices.

3.4 Counterfactual analysis

The model allows for a rich set of post-estimation counterfactual experiments including product entry and exit, changes in observable or unobservable product characteristics, changes in parameters, or changes in ownership structure (e.g., mergers). Furthermore, these counterfactuals identify not only changes in equilibrium market shares and total social welfare but also percentage changes in price-cost margins, firms' profits, and consumer welfare.

Given estimates $(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\sigma}})$ and recovered unobservables $\widehat{\boldsymbol{\eta}}$, counterfactual equilibria can be computed by specifying a hypothetical market environment $(\mathcal{J}^*, \boldsymbol{x}^*, \boldsymbol{\theta}^*, \boldsymbol{\sigma}^*, \boldsymbol{\eta}^*, \boldsymbol{\Omega}^*)$. Some components may be held fixed at their estimated or observed values, while others may be changed to represent the counterfactual scenario. The set of counterfactual equilibrium market shares is the set of vectors $\boldsymbol{s}^* \in \Delta^{J^*}$ satisfying

$$\boldsymbol{x}^*\boldsymbol{\theta}^* + \boldsymbol{\eta}^* = \mathfrak{s}^{-1}\left(\boldsymbol{s}^*, \boldsymbol{x}^*, \boldsymbol{\sigma}^*\right) + \left[\boldsymbol{\Omega}^* \circ \boldsymbol{\Delta}\left(\boldsymbol{s}^*, \boldsymbol{x}^*, \boldsymbol{\sigma}^*\right)\right]^{-1} \boldsymbol{s}^*. \tag{44}$$

If this system has a unique solution, then the counterfactual market-share vector $\boldsymbol{s}^*$ is identified for the specified hypothetical change in exogenous variables or parameters. If the system has multiple solutions, the model admits multiple counterfactual Bertrand equilibria, and the counterfactual set of market shares is identified rather than a unique vector.

Once $\boldsymbol{s}^*$ has been recovered, remaining equilibrium outcomes can also be recovered. Market shares and diversion ratios are recovered in levels, while price–cost margins, firms' profits, and consumer welfare are recovered up to scale. For the latter objects, proportional changes relative to the baseline equilibrium are nevertheless recovered exactly.

4 Empirical application

The purpose of this empirical application is to illustrate that our model and estimation method, implemented without price data, deliver estimates of the structural parameters $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ that are economically and statistically indistinguishable from those obtained using correctly measured prices. At the same time, estimates based on our method differ markedly from those produced when price data are contaminated by substantial measurement error. In this sense, the application highlights the precision and robustness of our approach to price measurement error.

To this end, we use data from the European automobile industry, which have been widely employed in several influential studies. A key feature of these data is that demand and cost parameters can be estimated using prices aggregated at different levels, such as model-country-year, model-year, or brand-year, providing a natural setting to assess the impact of price aggregation and measurement error.

The dataset comes from Goldberg and Verboven (2001) and Björnerstedt and Verboven (2014).[4] It is organized at the level of a vehicle model observed in a given country and year. It contains a total of 11, 483 observations, spanning a 30-year period from 1970 to 1999 across five European markets: Belgium, France, Germany, Italy, and the United Kingdom. To facilitate a transparent comparison of estimation results and counterfactual analyses with and without price data, we adopt the same sample selection, model specification, and variable construction (e.g., market size) as in Björnerstedt and Verboven (2014).

[4]This dataset was collected by Pinilope Goldberg and Frank Verboven, and maintained on their webpages. The datafile is publicly available in Frank Verboven's website here https://sites.google.com/site/frankverbo/data-and-software/data-set-on-the-european-car-market.

4.1 Model

Following the standard assumption in equilibrium models of the automobile industry, we define markets at the country-year level. On average, each country–year market includes approximately 77 distinct car models. Quantities, market shares, and prices are observed at the model-country-year level, and we assume that firms compete in prices at this level of aggregation.

Based on Goldberg and Verboven (2001) and Björnerstedt and Verboven (2014), we consider a two-tier nested logit demand model. The car market is divided into five upper nests, or groups, corresponding to the market segments: subcompact, compact, intermediate, standard, and luxury. Each segment is further subdivided into two lower nests, or subgroups, according to the origin of the car model in a given country: domestic or foreign. For example, BMW is classified as domestic in Germany and as foreign in all other countries. We index groups by g and subgroups by h.

The observable product attributes in the vector x_{jmt}, affecting demand and marginal costs, are: engine power (`horsepower`), measured in kilowatts; fuel consumption (`fuel`), in liters per 100 kilometers; vehicle width (`width`), in centimeters; vehicle height (`height`), in centimeters; domestic-brand dummy (`domestic`); country dummies; and year dummies. Prices are expressed in thousands of euros, adjusted to constant 1999 purchasing power. Quantities are measured using new vehicle registrations, which proxy for unit sales.

Following Björnerstedt and Verboven (2014), we measure potential market size at the country-year level as the total country population divided by four, which serves as a rough proxy for the number of households. To account for potential misspecification of market size, the demand equation includes controls for the country-year logarithms of population and GDP, as well as country and year fixed effects.

In this Two-tier Nested Logit model with multiproduct firms, the regression demand equation for product j, in country m, and year t has the following form:

$$\ln(s_{jmt}/s_{0mt}) \;=\; \boldsymbol{x}'_{jmt}\boldsymbol{\beta} \;-\; \alpha\, p_{jmt} \;+\; \sigma_1\, \ln(s_{jmt|hg}) \;+\; \sigma_2\, \ln(s_{h,mt|g}) \;+\; \xi_{jmt}, \tag{45}$$

where $s_{jmt|hg}$ is product j share within subgroup h of group g, and $s_{h,mt|g}$ represents subgroup h share within group g. For this demand model, under Bertrand competition, multiproduct firms, and prices set at the country level, the equilibrium pricing equation for product j is:

$$\begin{aligned} p_{jmt} \;&=\; \boldsymbol{x}'_{jmt}\boldsymbol{\gamma} + \omega_{jmt} \\ &+\; \frac{1}{\alpha}\left(\frac{(1-\sigma_1)(1-\sigma_2)}{(1-\sigma_2)-(1-\sigma_1)(1-\sigma_2)S_{fmt}-(\sigma_1-\sigma_2)S_{fmt|hg}-\sigma_2(1-\sigma_1)S_{fmt|hg}s_{h,mt|g}}\right), \end{aligned} \tag{46}$$

where: $S_{fmt} = \sum_{k\in\mathcal{J}_f} s_{kmt}$, and $S_{fmt|hg} = \sum_{k\in\mathcal{J}_f\cap g} s^{w}_{kmt|hg}$. Solving the pricing equation (46) into the demand equation (45), we have our social surplus regression equation:

$$\begin{aligned} \ln(s_{jmt}/s_{0mt}) \;&=\; \boldsymbol{x}'_{jmt}\boldsymbol{\theta} \;+\; \eta_{jmt} \;+\; \sigma_1\, \ln(s_{jmt|hg}) \;+\; \sigma_2\, \ln(s_{h,mt|g}) \\ &-\; \frac{(1-\sigma_1)(1-\sigma_2)}{(1-\sigma_2)-(1-\sigma_1)(1-\sigma_2)S_{fmt}-(\sigma_1-\sigma_2)S_{fmt|hg}-\sigma_2(1-\sigma_1)S_{fmt|hg}s_{h,mt|g}}. \end{aligned} \tag{47}$$

To facilitate comparison with the results in Björnerstedt and Verboven (2014), we begin by following their empirical approach to address the endogeneity of prices and market shares, which relies on a rich set of car-model and country fixed effects rather than instrumental variables. We subsequently present estimates based on the standard instrumental-variable approach in the differentiated-products demand literature, namely *BLP instruments*.

4.2 Parameter estimates

Tables 2 and 3 provide the central empirical validation of our approach. The estimates of the social surplus and nested-logit substitution parameters obtained without price data are economically very similar to those obtained using conventional demand and marginal cost estimation with prices.

Table 2 presents four sets of estimates. The first two columns report the conventional sequential estimation of demand and marginal cost using price data. The third column reports the implied social surplus parameters obtained from these estimates. The fourth column reports estimates of the social surplus equation derived in this paper, which exploits only market shares and product characteristics and does not use price data.

Several results are noteworthy. First, all coefficients in the social surplus specification estimated without price data are highly precise, with standard errors that are often comparable to those obtained in the demand equation estimated with prices. Thus, the absence of price information does not appear to generate a substantial loss of precision.

Second, the estimated coefficients are generally very similar across the two approaches. The most important result concerns the nesting parameters. The estimate of the upper-level nesting parameter increases only from 0.9047 to 0.9263, while the estimate of the lower-level nesting parameter changes from 0.5678 to 0.5311. These differences correspond to only 2.3% and 6.6%, respectively.[5] Given the central role that these parameters play in determining substitution patterns, diversion ratios, markups, and welfare effects, the close correspondence between the two sets of estimates provides strong support for their identification without price data.

More generally, the signs and magnitudes of the coefficients on product characteristics are remarkably stable across methods. The estimates imply very similar rankings of products and very similar relationships between product characteristics and social surplus. While some coefficients differ more substantially than others, the overall picture is one of substantial similarity between the estimates obtained with and without price data.

[5]For percentage differences, we use the absolute symmetric difference: $ASD(x, y) = \dfrac{2|x-y|}{|x|+|y|}$.

Table 2: Estimation of Demand, Marginal Cost (Sequential), and Social Surplus

	With Price Data			**Without Price Data**
	Demand	Marg. Cost	Social Surplus	Social Surplus
Price ($-\alpha$)	-0.0468***			
	(0.0013)			
σ_1	0.9047***		0.9047***	0.9263***
	(0.0041)		(0.0041)	(0.0041)
σ_2	0.5678***		0.5678***	0.5311***
	(0.0085)		(0.0085)	(0.0092)
Horsepower	0.0038***	0.1665***	-0.0039***	-0.0035***
	(0.0006)	(0.0045)	(0.0006)	(0.0005)
Fuel	-0.0271***	-0.1151***	-0.0216***	-0.0213***
	(0.0045)	(0.0372)	(0.0049)	(0.0047)
Width	0.0104***	0.0541***	0.0078***	0.0067***
	(0.0017)	(0.0136)	(0.0019)	(0.0017)
Height	0.0004	0.0175	-0.0003	-0.0010
	(0.0022)	(0.0181)	(0.0025)	(0.0023)
Domestic	0.5231***	-3.0127***	0.6641***	0.6643***
	(0.0124)	(0.0805)	(0.0177)	(0.0140)
Year	0.0017	-0.1887***	0.0105***	0.0121***
	(0.0012)	(0.0094)	(0.0013)	(0.0012)
France	-0.6621***	0.9087***	-0.7047***	-0.7531***
	(0.0139)	(0.0891)	(0.0156)	(0.0151)
Germany	-0.5883***	-0.7089***	-0.5551***	-0.6115***
	(0.0147)	(0.0890)	(0.0181)	(0.0160)
Italy	-0.7129***	1.0527***	-0.7622***	-0.7983***
	(0.0137)	(0.0926)	(0.0164)	(0.0137)
UK	-0.4155***	7.8127***	-0.7815***	-0.8253***
	(0.0167)	(0.0888)	(0.0140)	(0.0142)
Observations	11,483	11,483	11,483	11,483
R^2	0.9431	0.8944		0.9471

Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 3 formally compares the σ and social surplus estimates obtained under the two approaches. Because the sample contains more than 11,000 observations, both sets of estimates are very precise. As a consequence, even relatively small differences are often statistically significant. This distinction between statistical and economic significance is important for interpreting the results.

Table 3: Testing Equality of Subsitution & Social Surplus Estimates With & Without Price Data

	With Price	Without	Abs. Diff.	% Diff.	χ^2	P-value
σ_1	0.9047***	0.9263***	0.0216***	2.36%	93.38	0.0000
σ_2	0.5678***	0.5311***	0.0366***	6.67%	40.59	0.0000
Horsepower	-0.0039***	-0.0035***	0.0004	10.58%	38.84	0.0000
Fuel	-0.0216***	-0.0213***	0.0004	1.84%	1.17	**0.2786**
Width	0.0078***	0.0067***	0.0011***	15.39%	33.66	0.0000
Height	-0.0003	-0.0010	0.0006***	98.98%	11.87	0.0006
Domestic	0.6641***	0.6643***	0.0001	0.02%	0.001	**0.9874**
Year	0.0105***	0.0121***	0.0015***	13.72%	86.72	0.0000
France	-0.7047***	-0.7531***	0.0483***	6.63%	56.19	0.0000
Germany	-0.5551***	-0.6115***	0.0564***	9.68%	57.11	0.0000
Italy	-0.7622***	-0.7983***	0.0361***	4.62%	37.85	0.0000
UK	-0.7815***	-0.8253***	0.0438***	5.45%	60.01	0.0000
Observations	11,483	11,483				

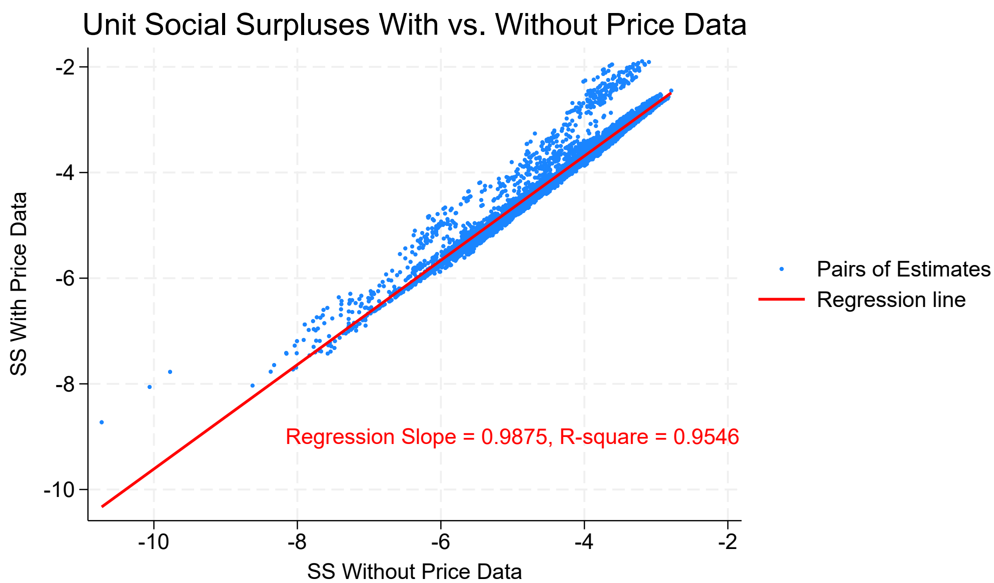
* p<0.10, ** p<0.05, *** p<0.01

Note: *Abs. Diff.* is the absolute value of the difference between the two estimates.

Note: *% Diff.* is the Absolute Symmetric Percentage Difference, $ASD(x, y) = \frac{2|x-y|}{|x|+|y|}$.

The nesting parameters provide the clearest example. The difference in the estimates of σ_1 and σ_2 are statistically significant, with large chi-square statistics, yet the magnitudes of the differences are only $0.02(2.3\%)$ and $0.03(6.6\%)$, respectively. Although the null hypothesis of equality is rejected in both cases, these differences are economically small and imply very similar substitution patterns. The same conclusion applies to all the product-characteristic coefficients. Thus, the statistical tests should not be interpreted as evidence that the two

Figure 1: Scatterplot of Unit Social Surpluses With & Without Price Data



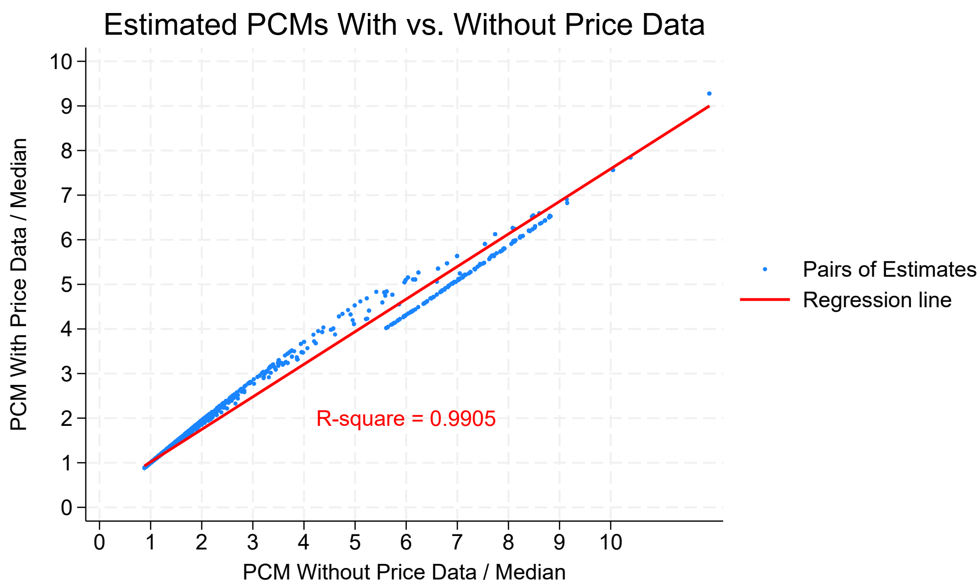
approaches lead to substantially different economic conclusions. Rather, they reflect the fact that both sets of estimates are obtained with very high precision.

The results have two important implications. First, they provide direct evidence that the information contained in equilibrium market shares is sufficiently rich to recover the key demand-side determinants of welfare even when prices are unavailable. Second, and perhaps more importantly, the close correspondence between the estimates suggests that the proposed approach successfully recovers the substitution patterns that govern market power and competitive interactions.

4.3 Social Surpluses

Figure 1 presents a scatterplot of the estimated social surpluses, $xjmt'\widehat{\boldsymbol{\theta}} + \widehat{\eta}jmt$, obtained with and without price data. The two sets of estimates are remarkably similar, as illustrated by the fitted regression line, which has a slope of 0.9875 and R^2 of 0.9546.

Figure 2: Scatterplot of Price-Cost Margins With & Without Price Data

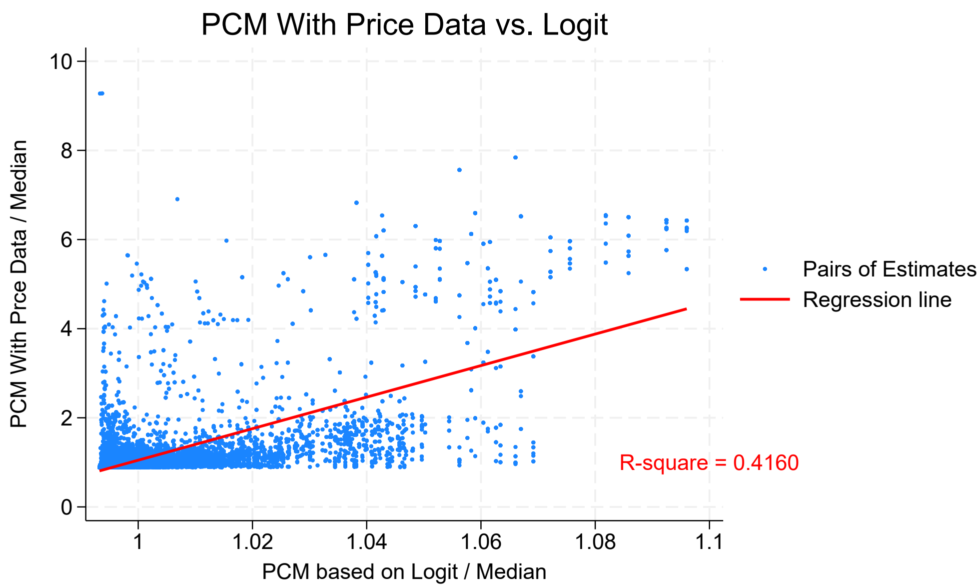


4.4 Price-Cost Margins

Because the approach without price data identifies price-cost margins only up to a scale factor, a meaningful comparison between the two sets of estimates requires a normalization. One possibility would be to normalize both sets of estimates using a benchmark product-country-year (e.g., the BMW 3-Series sold in Germany in 1980). However, comparisons of the levels would then be sensitive to the arbitrary choice of the benchmark product. Instead, we normalize each set of estimates by its average price-cost margin. That is, we divide every estimated margin obtained without price data by the median of those estimates, and we do the same for the margins estimated with price data. Under this normalization, the median normalized margin equals one in each approach, though the product-country-year corresponding to the median margin needs not be the same across the two approaches. This normalization provides a more robust basis for comparing the cross-sectional distributions of estimated margins.

Figure 2 presents a scatterplot of the price-cost margins estimated with and without price data. The correspondence between the two sets of estimates is remarkably strong, with an R^2

Figure 3: Scatterplot of PCMs With Price Data vs. Logit Estimates

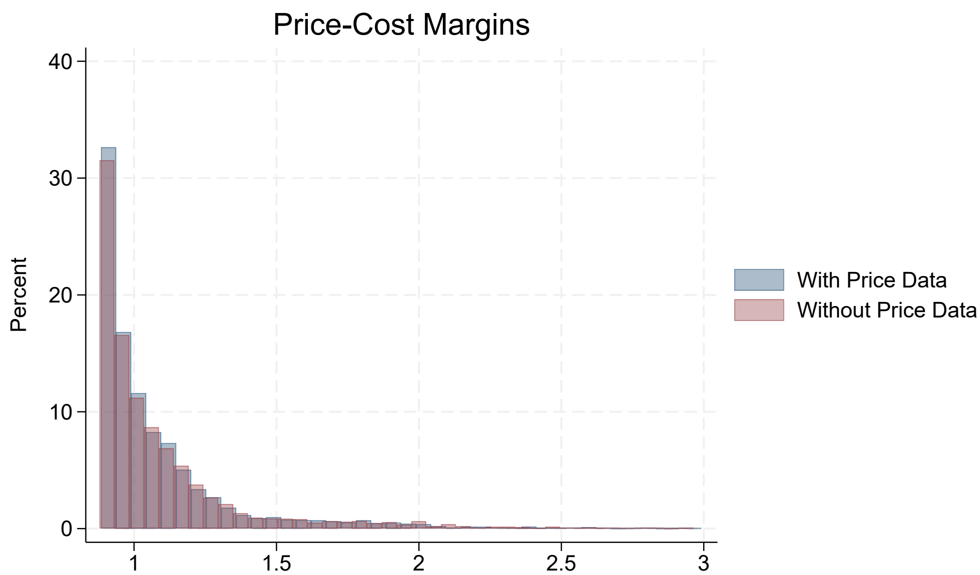


of 0.9905. Both sets of margins are computed from the same nonlinear equilibrium mapping between observed market shares and price-cost margins. The only difference is that this mapping is evaluated using different estimates of the nesting parameters, σ_1 and σ_2. Because these parameter estimates are themselves very similar, the resulting estimates of price-cost margins exhibit an equally close correspondence.

This close agreement should not be interpreted as implying that price-cost margins are generally insensitive to the nesting parameters. On the contrary, inconsistent estimates of these parameters can generate dramatically different estimates of price-cost margins, potentially eliminating even the monotonic relationship with the correctly estimated margins. Figure 3 illustrates this point. It plots the price-cost margins estimated with price data against the margins obtained when the nesting parameters are incorrectly fixed at their standard logit values, $\sigma_1 = \sigma_2 = 0$. The sharp contrast with Figure 2 demonstrates that the close correspondence between the two approaches depends critically on the accurate identification of the demand substitution parameters.

Figure 2 presents the histograms of the estimated price-cost margins with and without

Figure 4: Histograms of Price-Cost Margins With & Without Price Data



price data. The distributions closely overlap, indicating that the approach without price data reproduces not only the cross-sectional ranking of price-cost margins but also their overall distribution.

Tables 4 to 6 provide a more direct assessment of the extent to which the approach without price data reproduces the economically relevant implications of the conventional approach. These tables compare the average equilibrium price-cost margins implied by the two methods across countries, market segments, and firms. The results reveal a remarkable degree of similarity.

Table 4 shows that the average margins obtained under the two approaches are very similar across all five countries. The largest difference is observed in Italy, where the average margin (ratio to median) increases from 1.858 to 2.166, and 15.3% difference. For the the remaining countries the differences are substantially smaller. Most importantly, the ranking of countries by average margins is virtually unchanged.

Table 5 presents a similar comparison across automobile segments. Again, the correspondence between the two approaches is striking. The average margins differ only modestly

Table 4: Average Price-Cost Margins by Country

	With Price	Without Price	Difference	% Diff.
Italy	1.858	2.166	-0.308	15.31
France	1.191	1.210	-0.019	1.58
Germany	1.186	1.220	-0.034	2.83
UK	1.151	1.184	-0.033	2.83
Belgium	1.005	1.004	0.002	0.10
Total	1.257	1.327	-0.071	5.42
N	11483			

Note: % *Diff.* is the Absolute Symmetric Percentage Difference, $ASD(x, y) = \frac{2|x-y|}{|x|+|y|}$.

Table 5: Average Price-Cost Margins by Segment

	With Price	Without Price	Difference	% Diff.
luxury	1.661	1.829	-0.168	9.63
subcompact	1.367	1.476	-0.108	7.67
intermediate	1.180	1.226	-0.045	3.82
standard	1.164	1.209	-0.045	3.79
compact	1.130	1.167	-0.036	3.22
Total	1.257	1.327	-0.071	5.42
N	11483			

Note: % *Diff.* is the Absolute Symmetric Percentage Difference, $ASD(x, y) = \frac{2|x-y|}{|x|+|y|}$.

across the two methods, and the ranking of segments is identical. Luxury vehicles exhibit the highest margins under both methods, followed by subcompact vehicles, while compact, intermediate, and standard segments display very similar margins. The largest discrepancy occurs for luxury vehicles, where the estimated margin increase by 9.6% from 1.661 to 1.829. Nevertheless, the qualitative and quantitative conclusions remain unchanged.

Table 6: Average Price-Cost Margins by Firm

	With Price	Without Price	Difference	% Diff.
Mercedes	2.291	2.418	-0.127	5.39
Fiat	1.963	2.314	-0.352	16.41
AlfaRomeo	1.271	1.386	-0.115	8.66
PSA	1.207	1.225	-0.018	1.48
BMW	1.195	1.220	-0.025	2.07
VW	1.195	1.227	-0.032	2.64
Renault	1.174	1.188	-0.014	1.19
Ford	1.148	1.162	-0.015	1.21
GM	1.144	1.188	-0.044	3.77
Rover	1.138	1.168	-0.030	2.60
Volvo	1.068	1.095	-0.026	2.50
Talbot-Simca	1.023	1.032	-0.009	0.88
Nissan	0.951	0.950	0.001	0.11
Honda	0.937	0.935	0.002	0.21
Toyota	0.935	0.932	0.003	0.32
Saab	0.931	0.928	0.003	0.32
Seat	0.930	0.928	0.003	0.22
DAF	0.929	0.926	0.003	0.32
Mazda	0.924	0.921	0.003	0.33
Mitsubishi	0.921	0.917	0.004	0.44
DeTomaso	0.904	0.900	0.005	0.44
Daewoo	0.902	0.898	0.005	0.44
Suzuki	0.901	0.897	0.005	0.44
Hyundai	0.898	0.893	0.005	0.56
Talbot-Matra	0.898	0.893	0.005	0.56
Kia	0.890	0.885	0.005	0.56
Total	1.257	1.327	-0.071	5.42
N	11483			

Note: % *Diff.* is the Absolute Symmetric Percentage Difference, $ASD(x,y) = \frac{2|x-y|}{|x|+|y|}$.

The firm-level comparison presented in Table 6 provides a more demanding test of the methodology, since it requires matching a much richer cross-sectional pattern. Despite this challenge, the agreement between the two approaches remains remarkably strong. For most firms, the differences between the two methods are smaller than 1%. Firms with high margins under the conventional demand system, such as Mercedes, Fiat, BMW, PSA, Renault, and Volkswagen, are also identified as high-margin firms by the approach without price data. Similarly, firms with relatively low margins, including Hyundai, Kia, Daewoo, and Suzuki, remain among the lowest-margin firms under both methods. While some differences emerge in the exact magnitude of the margins, the overall ranking of firms is largely preserved. In particular, Mercedes remains by far the firm with the highest average margin under both approaches, while the relative positions of most firms change very little.

Taken together, these results provide evidence that the approach without price data accurately recovers the cross-sectional variation in equilibrium price-cost margins.

4.5 Counterfactual Mergers

We replicate the counterfactual merger exercise considered in Björnerstedt and Verboven (2014) that simulates the acquisition of General Motors (GM) by Volkswagen (VW) in 1998. As in the main specification in Björnerstedt and Verboven (2014), we assume that merger generates no efficiency gains for the merging parties and that no coordinated effects arise. Table 7 compares the predicted changes in firm market shares following a hypothetical merger between VW and GM under the standard estimation procedure with price data and our approach without price data. The two methods produce remarkably similar predictions. Both imply that the market shares of the merging firms decline after the merger, while the market shares of close competing firms increase. Moreover, the two approaches identify essentially the same competitors as the main beneficiaries of the merger, with Ford, Mercedes, and BMW experiencing the largest gains in market share. The close agreement between the two sets

Table 7: Pre- and post-merger firm market shares under the two methods

Firm	Pre-Merger	Post-Merger	
		With Price Data	**Without Price Data**
GM	0.166	0.108	0.093
VW	0.300	0.280	0.287
Mercedes	0.100	0.116	0.120
Ford	0.095	0.132	0.137
BMW	0.074	0.079	0.081
Renault	0.051	0.054	0.056
Fiat	0.043	0.045	0.047
PSA	0.034	0.037	0.038
Toyota	0.027	0.029	0.030
Mazda	0.025	0.027	0.028
Nissan	0.025	0.027	0.028
Mitsubishi	0.015	0.017	0.017
Volvo	0.012	0.013	0.014
Honda	0.012	0.012	0.013
Hyundai	0.006	0.006	0.007
Suzuki	0.006	0.006	0.007
Daewoo	0.006	0.007	0.007
Kia	0.003	0.003	0.004

Note: Post-merger equilibrium estimated using a fixed-point iteration method in equation (47).

of counterfactual predictions provides further evidence that market shares alone contain sufficient information to recover the substitution patterns that determine the competitive effects of mergers, even in the absence of price data.

The predicted changes in market shares are also economically intuitive. Under Bertrand competition, the merger internalizes competition (cannibalization) between the products of the merging firms, reducing their incentive to compete aggressively for consumers. As a result, the merged firm raises prices, accepting lower sales in exchange for higher markups. Consumers who no longer purchase GM or VW products substitute toward competing brands, generating the observed increases in rivals' market shares. The distribution of these gains across competitors is determined by the estimated demand substitution patterns: firms whose products are closer substitutes for GM and VW capture a larger fraction of the diverted demand. Thus, although the merger reduces the market shares of the merging firms, it simultaneously increases their market power and profitability through higher equilibrium markups, while reallocating sales toward competing producers.

5 Conclusion

This paper develops a framework for structural estimation and counterfactual analysis in differentiated-product markets that does not require observing prices. By combining the equilibrium pricing equations with the demand system, we derive a system of equilibrium conditions in which market shares are the only endogenous variables. This representation identifies the structural parameters governing product-level social surpluses and demand substitution patterns using only market shares and exogenous product characteristics. These objects are sufficient to recover the equilibrium outcomes required for welfare analysis, including price-cost margins, profits, consumer surplus, diversion ratios, and merger effects, without directly observing prices.

Our approach has two important practical advantages. First, it substantially expands the range of industries in which structural demand estimation and welfare analysis can be implemented. Many important markets, including healthcare, banking, media, and business-to-business transactions, are characterized by confidential, unavailable, or highly aggregated prices. In these settings, the proposed framework makes structural analysis feasible using information that is often readily available. Second, because the approach does not rely on observed prices, it is naturally robust to economically important price measurement error. The empirical application to the European automobile industry shows that the proposed method closely reproduces the parameter estimates, price-cost margins, and merger counterfactuals obtained under conventional estimation when accurate price data are available, while outperforming conventional methods when prices are contaminated by aggregation or measurement error.

Although the paper focuses on the standard BLP framework with Bertrand competition, the underlying idea is considerably more general. The equilibrium mapping between market shares and product-level social surpluses extends to alternative models of demand and competition whenever equilibrium conditions allow prices to be eliminated from the system. We hope that this perspective will broaden the scope of empirical industrial organization by enabling structural analyses in settings where reliable price data are unavailable and by motivating further research on identification and counterfactual analysis using equilibrium restrictions beyond prices.

A Appendix: Proofs

A.1 Proof of Proposition 2

Necessary condition: Suppose that $(\boldsymbol{p}^*, \boldsymbol{s}^*)$ is an equilibrium such that the following conditions hold:

$$\begin{cases} \boldsymbol{p}^* - \boldsymbol{v} &= -\dfrac{1}{\alpha}\, \mathfrak{s}^{-1}(\boldsymbol{s}^*, \boldsymbol{x}, \boldsymbol{\sigma}) \\[2ex] \boldsymbol{p}^* - \boldsymbol{c} &= \dfrac{1}{\alpha} \left(\boldsymbol{\Omega} \circ \boldsymbol{\Delta}(\boldsymbol{s}^*, \boldsymbol{x}, \boldsymbol{\sigma})'\right)^{-1} \boldsymbol{s}^* \end{cases} \tag{48}$$

Using these equations to calculate the product-by-product difference between the pricing equation and the inverse demand equation, we have:

$$\alpha\,(\boldsymbol{v} - \boldsymbol{c}) \;=\; \mathfrak{s}^{-1}(\boldsymbol{s}^*, \boldsymbol{x}, \boldsymbol{\sigma}) \;+\; \left[\boldsymbol{\Omega} \circ \boldsymbol{\Delta}(\boldsymbol{s}^*, \boldsymbol{x}, \boldsymbol{\sigma})\right]^{-1} \boldsymbol{s}^* \tag{49}$$

which implies that $\boldsymbol{s}^*$ solves the system of equations (16).

Sufficient condition: Suppose that the vector of market shares $\boldsymbol{s}^*$ solves the system of equations in (16). Given $\boldsymbol{s}^*$, we can construct the corresponding vector of prices as: $\boldsymbol{p}^* - \boldsymbol{v} = -\frac{1}{\alpha}\, \mathfrak{s}^{-1}(\boldsymbol{s}^*, \boldsymbol{x}, \boldsymbol{\sigma})$. It is straightforward to verify that the pair $(\boldsymbol{p}^*, \boldsymbol{s}^*)$ satisfies the conditions in (48). Therefore, $(\boldsymbol{p}^*, \boldsymbol{s}^*)$ constitutes a Bertrand equilibrium. $\blacksquare$

A.2 Equilibrium uniqueness in logit model

Define $w_j \equiv \alpha(v_j - c_j)$. We can write equation (18) as: $w_j + \ln(s_0) \;=\; \ln(s_j) + \dfrac{1}{1 - s_j}$. The function in the right-hand-side is strictly increasing in $s_j \in (0,1)$ and ranges from $-\infty$ to $+\infty$. This implies that there is an inverse function $s_j = g_j(w_j, s_0)$ that solves this equation for the market share s_j for a fixed value of s_0. Furthermore, it is straightforward to show that this inverse function $g_j(w_j, s_0)$ is strictly increasing in s_0.

Given that $s_0 = 1 - \sum_{j=1}^{J} s_j$, we have:

$$s_0 \;=\; 1 \;-\; \sum_{j=1}^{J} g_j(w_j, s_0) \tag{50}$$

Since every function g_j is strictly increasing in s_0, it is straightforward to show that this equation has a unique solution for s_0. This solution is a function of the vector of social surpluses $\boldsymbol{w}$: $s_0 = G_0(\boldsymbol{w})$. Finally, we obtain the inside market shares using:

$$s_j \;=\; g_j\left(w_j,\, G_0(\boldsymbol{w})\right) \qquad \blacksquare \tag{51}$$

A.3 Uniform Pricing

Firm f's profit across the M markets is:

$$\pi_f \;=\; \sum_{m=1}^{M} \sum_{j \in \mathcal{J}_f} \left(p_j - c_{jm}\right) \; \mathfrak{s}_{jm}(\boldsymbol{p}), \tag{52}$$

The first-order condition for profit maximization with respect to the price of product j is:

$$\sum_{m=1}^{M} \left[s_{jm} \;+\; \sum_{k \in \mathcal{J}_f} \left(p_k - c_{km}\right) \; \frac{\partial \mathfrak{s}_{km}}{\partial p_j} \right] \;=\; 0. \tag{53}$$

We can represent these first-order conditions in vector form as:

$$\sum_{m=1}^{M} \left[\boldsymbol{s}_m \;-\; \Big(\boldsymbol{\Omega} \circ \boldsymbol{\Delta} \left(\boldsymbol{s}_m,\, \boldsymbol{x}_m,\, \boldsymbol{\sigma}_m\right) \Big) \; \alpha_m \; \left(\boldsymbol{p} - \boldsymbol{c}_m\right) \right] \;=\; \boldsymbol{0}. \tag{54}$$

where we have used the relationship $\frac{\partial \boldsymbol{\mathfrak{s}}_m'}{\partial \boldsymbol{p}} = -\alpha_m \, \boldsymbol{\Delta} \left(\boldsymbol{s}_m,\, \boldsymbol{x}_m,\, \boldsymbol{\sigma}_m\right)$.

From the demand-side, for every market m, the inverse demand equation is:

$$\boldsymbol{p} \;=\; \boldsymbol{v}_m \;-\; \frac{1}{\alpha_m} \; \mathfrak{s}_m^{-1} \left(\boldsymbol{s}_m, \boldsymbol{x}_m, \boldsymbol{\sigma}_m\right) \tag{55}$$

Solving the inverse-demand equations into the system of pricing equations, we get:

$$\sum_{m=1}^{M} \left[\boldsymbol{s}_m \;-\; \Big(\boldsymbol{\Omega} \circ \boldsymbol{\Delta} \left(\boldsymbol{s}_m,\, \boldsymbol{x}_m,\, \boldsymbol{\sigma}_m\right) \Big) \Big(\alpha_m \; \left(\boldsymbol{v}_m - \boldsymbol{c}_m\right) - \mathfrak{s}_m^{-1} \left(\boldsymbol{s}_m, \boldsymbol{x}_m, \boldsymbol{\sigma}_m\right) \Big) \right] \;=\; \boldsymbol{0}. \tag{56}$$

References

Aguirregabiria, V., Clark, R., and Wang, H. (2025). The geographic flow of bank funding and access to credit: Branch networks, synergies, and local competition. *American Economic Review*, 115(6):1818–56.

Allen, J., Clark, R., and Houde, J.-F. (2014). The effect of mergers in search markets: Evidence from the canadian mortgage industry. *American Economic Review*, 104(10):3365–3396.

Allen, J., Clark, R., and Houde, J.-F. (2019). Search frictions and market power in negotiated-price markets. *Journal of Political Economy*, 127(4):1550–1598.

Allenby, G. M., Hardt, N., and Rossi, P. E. (2019). Economic foundations of conjoint analysis. In Dubé, J.-P. and Rossi, P. E., editors, *Handbook of the Economics of Marketing*, volume 1, pages 151–192. North-Holland.

Armona, L., Lewis, G., and Zervas, G. (2021). Learning product characteristics and consumer preferences from search data. Working paper.

Backus, M., Conlon, C., and Sinkinson, M. (2019). The common ownership hypothesis: Theory and evidence. *Economic Studies at Brookings*, 1980.

Backus, M., Conlon, C., and Sinkinson, M. (2021). Common ownership in america: 1980–2017. *American Economic Journal: Microeconomics*, 13(3):273–308.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Berry, S. T. and Waldfogel, J. (1999). Free entry and social inefficiency in radio broadcasting. *The RAND Journal of Economics*, 30(3):397–420.

Björnerstedt, J. and Verboven, F. (2014). Merger simulation with nested logit demand. *The Stata Journal*, 14(3):511–540.

Bresnahan, T. F. and Salop, S. C. (1986). Quantifying the competitive effects of production joint ventures. *International Journal of Industrial Organization*, 4(2):155–175.

Bugni, F. A., Horowitz, J. L., and Zhang, L. (2026). Demand estimation without outside good shares. Working paper, arXiv:2602.19154.

Caplin, A. and Nalebuff, B. (1991). Aggregation and imperfect competition: On the existence of equilibrium. *Econometrica*, 59(1):25–59.

Cardell, S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 13(2):185–213.

Conlon, C., Mortimer, J. H., and Sarkis, P. (2026). Estimating preferences and substitution patterns from second-choice data alone. Working paper.

Egan, M., Hortaçsu, A., and Matvos, G. (2017). Deposit competition and financial fragility: Evidence from the us banking sector. *American Economic Review*, 107(1):169–216.

Einav, L., Guido, M., and Klenow, P. (2026). Customer overlap and diversion ratios. Working paper.

Farrell, J. and Shapiro, C. (2010). Antitrust evaluation of horizontal mergers: An economic alternative to market definition. *The B.E. Journal of Theoretical Economics*, 10(1).

Gandhi, A. and Houde, J.-F. (2019). Measuring substitution patterns in differentiated-products industries. *NBER Working paper*, (w26375).

Garrido, F. (2024). An aggregative approach to pricing equilibrium among multi-product firms with nested demand. *The RAND Journal of Economics*, 55(3):359–374.

Goldberg, P. K. and Verboven, F. (2001). The evolution of price dispersion in the european car market. *The Review of Economic Studies*, 68(4):811–848.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105(1):172–203.

Hausman, J. A., Leonard, G. K., and Zona, J. D. (1994). Competitive analysis with differentiated products. *Annals of Economics and Statistics*, (34):143–157.

Ho, K. (2009). Insurer-provider networks in the medical care market. *The American Economic Review*, 99(1):393–430.

Ho, K. and Lee, R. S. (2017). Insurer competition in health care markets. *Econometrica*, 85(2):379–417.

Koh, P. S. (2025). Merger analysis with unobserved prices. Working paper, arXiv:2404.07684.

Magnolfi, L., McClure, J., and Sorensen, A. (2025). Triplet embeddings for demand estimation. *American Economic Journal: Microeconomics*, 17(1):282–307.

Miller, N. H. (2014). Modeling the effects of mergers in procurement. *International Journal of Industrial Organization*, 37:201–208.

Miller, N. H., Remer, M., and Sheu, G. (2013). Using cost pass-through to calibrate demand. *Economics Letters*, 118(3):451–454.

Miller, N. H. and Sheu, G. (2021). Quantitative methods for evaluating the unilateral effects of mergers. *Review of Industrial Organization*, 58(1):143–177.

Miller, N. H. and Weinberg, M. C. (2017). Understanding the price effects of the MillerCoors joint venture. *Econometrica*, 85(6):1763–1791.

Mizuno, T. (2003). On the existence of a unique price equilibrium for models of product differentiation. *International Journal of Industrial Organization*, 21(6):761–793.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 31(3):395–421.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342.

Nocke, V. and Schutz, N. (2018). Multiproduct-firm oligopoly: An aggregative games approach. *Econometrica*, 86(2):523–557.

Peters, C. (2006). Evaluating the performance of merger simulation: Evidence from the U.S. airline industry. *Journal of Law and Economics*, 49(2):627–649.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Sweeting, A. (2013). Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica*, 81(5):1763–1803.

Weinberg, M. C. (2011). More evidence on the performance of merger simulations. *American Economic Review*, 101(3):51–55.

Weinberg, M. C. and Hosken, D. (2013). Evidence on the accuracy of merger simulations. *Review of Economics and Statistics*, 95(5):1584–1600.

Werden, G. J. (1996). A robust test for consumer welfare enhancing mergers among sellers of differentiated products. *Journal of Industrial Economics*, 44(4):409–413.

Werden, G. J. and Froeb, L. M. (1994). The effects of mergers in differentiated products industries: Logit demand and merger policy. *The Journal of Law, Economics, and Organization*, 10(2):407–426.

Werden, G. J. and Froeb, L. M. (2002). The antitrust logit model for predicting unilateral competitive effects. *Antitrust Law Journal*, 70(1):257–260.