

Identification and Estimation of Demand Models with Endogenous Product Entry and Exit

Victor Aguirregabiria[†]

Alessandro Iaria[‡]

Senay Sokullu[§]

July 27, 2026

Abstract

Firms introduce products in markets where they anticipate stronger demand, using information unobserved by researchers. This creates endogenous selection in demand estimation. In differentiated-product oligopolies, multidimensional demand unobservables and strategic entry can render the ordinary propensity score insufficient for selection correction. Existing approaches either restrict firms' information at entry or jointly estimate demand, pricing, and entry under strong supply-side assumptions. We derive a new mixture representation of the selection-bias function using latent propensity scores: entry probabilities conditional on observables and a latent market state generating dependence across entry decisions. This representation yields a convenient two-step semiparametric estimator that corrects for selection and price endogeneity while accommodating richer information at entry. The approach makes weaker supply-side assumptions and is simpler to implement because it avoids repeatedly solving the full model. Applied to the US airline industry, the method yields more elastic demand and less market power than estimates ignoring endogenous entry.

Keywords: Demand estimation; Product entry; Selection bias; Airline markets.

JEL codes: C14, C34, C35, C57, D22, L13, L93.

Acknowledgments. We are grateful for helpful comments from Roy Allen, Gaurab Aryal, Giovanni Compiani, Andreea Enache, Christos Genakos, Ying Fan, Michele Fioretti, Philip Haile, Nail Kashaev, Mathieu Marcoux, Mateusz Mysliwski, Anders Munk-Nielsen, David Pacini, Bertel Schjerning, Philipp Schmidt-Dengler, Jesse Shapiro, Yutec Sun, Yuanyuan Wan, Ao Wang, and Christine Zulehner; as well as from seminar participants at the Universities of Bolzano, Bonn, Copenhagen, CREST, Duke, ENSAI-Rennes, Georgia, Glasgow, Harvard, Helsinki GSE, Cambridge (Judge), Mannheim, NHH-Bergen, Penn State, Rochester, Sciences Po, TSE, and Vienna; and from participants at the Advances in Demand Analysis Workshop (2024), BSE Summer Forum on Structural Microeconometrics (2024), CEPR IO Meeting (2024), Cowles Conference on Models & Measurement (2024), IAAE (2023), MaCCI Summer Institute (2022), Midwest Econometrics Group Meeting (2023), and Oxford Econometrics Workshop (2025).

[†]University of Toronto and CEPR, victor.aguirregabiria@utoronto.ca. [‡]University of Bristol and

CEPR, alessandro.iaria@bristol.ac.uk. [§]University of Bristol, senay.sokullu@bristol.ac.uk.

1 Introduction

Demand systems for differentiated products are commonly estimated using data across markets and periods in which product availability varies. Because firms tend to introduce products where they anticipate stronger demand, partly on the basis of information unobserved by researchers, this variation reflects endogenous product selection and can substantially bias demand estimates. The problem arises in many industries, including airlines ([Berry et al., 2006](#); [Berry and Jia, 2010](#); [Aguirregabiria and Ho, 2012](#)), supermarket chains ([Smith, 2004](#)), radio stations ([Sweeting, 2013](#)), personal computers ([Eizenberg, 2014](#)), and ice cream ([Draganska et al., 2009](#)).

The selection problem in this model differs from conventional cases because demand unobservables are multi-dimensional and enter firms' expected profits non-additively. The resulting entry rule need not be monotone or admit the single-index representation under which the ordinary propensity score—the probability of product entry given exogenous observables—is sufficient to control for selection bias ([Angrist, 1997](#)). Multiple equilibria in the entry and pricing games create an additional source of non-monotonicity when different equilibria are selected across markets. Consequently, standard identification results and two-step estimation methods based on the ordinary propensity score do not generally apply (e.g., [Ahn and Powell, 1993](#); [Das et al., 2003](#); [Aradillas-Lopez et al., 2007](#); [Newey, 2009](#)).

Existing approaches to this selection problem either rule out endogenous selection by assumption or accommodate it through joint estimation of demand, pricing, and entry under strong supply-side assumptions. In this paper, we establish more general conditions for the sequential identification and estimation of demand parameters when product entry is endogenous. Our approach leverages cross-product dependence in firms' market-entry decisions to identify entry probabilities that condition not only on observable characteristics but also on a latent market state capturing unobserved interdependencies among firms' choices. We refer to these probabilities as latent propensity scores. Unlike the ordinary propensity score, which averages over this latent state, latent propensity scores retain the information needed to characterize the selection-bias function in the demand equation.

Our identification result proceeds in two steps. First, we use the joint distribution of entry decisions to recover the latent propensity scores that summarize how this latent state shifts firms' entry probabilities. This step exploits a key feature of

the model: conditional on observable characteristics and the latent market state, firms' entry decisions are independent, so the dependence across products observed in the data is generated by this latent state. Second, we show that the selection-bias function in the demand equation can be represented in terms of these latent propensity scores and the conditional means of the demand shocks given this latent state. We develop two complementary approaches that make this representation useful for identification: one based on a finite-mixture representation and one based on continuous normal heterogeneity with index restrictions.

Building on our constructive proof of identification, we propose a transparent and computationally simple two-step estimator that jointly corrects for endogenous product selection and price endogeneity in demand estimation. In the first step, we estimate the entry-probability components needed to construct the selection-bias function by semiparametric Maximum Likelihood Estimation (MLE). In the second step, we estimate demand by a Generalized Method of Moments (GMM) that controls for selection bias. This yields consistent estimates without specifying and jointly estimating a full equilibrium model of demand, pricing, and entry.

We illustrate the method using data from the airline industry. The results demonstrate the importance of accounting for endogenous product entry and highlight the limitations of conventional selection-correction approaches. Specifically, demand estimates that do not correct for endogenous entry substantially underestimate the magnitude of own- and cross-price elasticities and overstate market power. We also detect selection bias—though weaker than for demand—in the estimation of marginal-cost functions derived from Bertrand–Nash pricing equations. Moreover, our reduced-form estimation of entry probabilities—capturing rich correlations in firms' entry decisions—provides economically meaningful insights. In particular, specifications that ignore correlated unobservables in market-entry decisions substantially understate the deterrent effect of competitors' presence on entry—a bias that would overstate market contestability in the evaluation of mergers.

Our approach contributes to two related strands of literature on differentiated-product demand with endogenous product entry. One strand estimates structural oligopoly models sequentially, beginning with demand ([Draganska et al., 2009](#); [Aguirregabiria and Ho, 2012](#); [Fan, 2013](#); [Sweeting, 2013](#); [Eizenberg, 2014](#); [Schau-mans and Verboven, 2015](#); [Fan and Yang, 2020](#); [Bontemps et al., 2023](#); [Caoui and Steck, 2026](#); [Liu and Luo, 2025](#)). To avoid selection bias in this first step, these studies

impose informational restrictions that rule out selection on demand unobservables, at the cost of potential misspecification. A second strand allows for such selection by jointly estimating demand, marginal costs, and entry costs using nested fixed-point algorithms (Ciliberto et al., 2021; Li et al., 2022). While powerful, this approach makes demand identification depend on strong functional-form and distributional assumptions and is computationally demanding. We show that endogenous selection in demand estimation can instead be accommodated within a sequential approach: cross-firm dependence in entry decisions identifies the selection-bias controls without requiring the joint estimation of the full model or solving the entry game inside the demand estimator.

Our method contributes to the literature on semiparametric estimation of sample selection models (Das et al., 2003; Newey, 2009; Powell, 2001; Aradillas-Lopez et al., 2007). We extend two-step propensity-score selection-function approaches to settings where the unobservables in the selection equation violate the standard monotonicity condition. When selection decisions arise within a system of simultaneous equations and non-monotonic unobservables generate dependence across them, we show that it is still possible to identify a selection function correcting for selection bias. Our approach applies in several contexts with this feature, such as labor market models with two-sided matching (Choo and Siow, 2006; Galichon and Salanié, 2022), joint household decisions (Browning et al., 2014), or peer effects models with endogenous network formation (Graham, 2017; De Paula et al., 2018).

The remainder of the paper is organized as follows. Sections 2–4 introduce the model, characterize the selection problem, and present our identification and estimation methods. Section 5 applies our methods to the US airline industry, and Section 6 concludes. Additional details and results appear in the Online Appendix.

2 Model

We study endogenous product entry in the canonical Berry–Levinsohn–Pakes (BLP) framework (Berry et al., 1995). For notational simplicity, we focus on single-product firms. In Online Appendix B, however, we discuss how our arguments can be adapted to multi-product firms. There are J firms indexed by $j \in \mathcal{J} = \{1, \dots, J\}$ and T markets indexed by $t \in \{1, \dots, T\}$. A market may be a geographic location, a period, or a combination of both. Consumers in market t can buy only the products

available in that market. Market-entry decisions, prices, and quantities arise from the equilibrium of a two-stage game. In the first stage, each firm chooses whether to enter the market to maximize its expected profit. In the second stage, active firms set prices in a Bertrand–Nash equilibrium, with quantities determined by demand. The game is played separately across markets.¹ The baseline model is static; however, Online Appendix C discusses how the same logic can be extended to dynamic games of firms’ product entry and exit. To help keep track of the notation, Online Appendix Tables A.1 and A.2 summarize the main symbols used in the paper.

2.1 Demand

The indirect utility of consumer h in market t from buying product j is:

$$U_{hjt} \equiv \delta_{jt} + \left(p_{jt}, \mathbf{x}_{jt}^{D'} \right) \boldsymbol{\Omega}_{\sigma} \mathbf{v}_{ht} + \varepsilon_{hjt}, \quad (1)$$

where p_{jt} is price and δ_{jt} is mean utility. For the baseline static model, let $\mathbf{x}_t$ collect all exogenous market- and product-level variables observed by the researcher, including the characteristics of every potential product regardless of entry, and let $\mathbf{x}_{jt}^D \in \mathbb{R}^K$ denote the subvector entering product j ’s demand. The vector $\mathbf{v}_{ht} \in \mathbb{R}^{K+1}$ has a known mean-zero distribution F_v , and $\boldsymbol{\Omega}_{\sigma}$ is a $(K+1) \times (K+1)$ matrix that is a known, continuously differentiable function of the parameter vector σ . Thus, $\boldsymbol{\Omega}_{\sigma} \mathbf{v}_{ht}$ is the vector of consumer-specific random-coefficient deviations. The shock ε_{hjt} is independent of $\mathbf{v}_{ht}$ and is i.i.d. over (h, j, t) with a type-I extreme-value distribution.

Following the standard linear specification, the mean utility of product j is:

$$\delta_{jt} \equiv \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \xi_{jt}, \quad (2)$$

where $\alpha \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^K$ are parameters, and ξ_{jt} captures product characteristics unobserved by the researcher. Throughout the paper, we normalize $\mathbb{E}(\xi_{jt}) = 0$. The outside option is indexed by $j = 0$, with $U_{h0t} = \varepsilon_{h0t}$.

Let $a_{jt} \in \{0, 1\}$ indicate whether product j is offered in market t , and let $\mathbf{a}_t \equiv$

¹Treating markets as separate games is standard in empirical industrial organization. Important exceptions are structural entry models that allow potential entrants to internalize network externalities across markets (e.g., Bontemps et al., 2023; Jia, 2008; Aguirregabiria and Ho, 2012). These models do not study the endogenous sample-selection problem considered here.

$(a_{jt} : j \in \mathcal{J})$ collect the offer indicators. The outside option $j = 0$ is always offered. Every consumer chooses the product that maximizes their utility. Define $\delta_t \equiv (\delta_{jt} : j \in \mathcal{J})$. The market share s_{jt} of product j in market t , equal to the proportion of consumers choosing that product, is

$$s_{jt} = d_{jt}(\delta_t, \mathbf{a}_t, \sigma) \equiv \int \frac{a_{jt} e^{\delta_{jt} + (p_{jt}, \mathbf{x}_{jt}^{D'}) \Omega_\sigma \mathbf{v}}}{1 + \sum_{i=1}^J a_{it} e^{\delta_{it} + (p_{it}, \mathbf{x}_{it}^D)' \Omega_\sigma \mathbf{v}}} dF_v(\mathbf{v}), \quad (3)$$

where, to simplify notation, the market subscript on d_{jt} is intended to capture the dependence of d_{jt} on prices and observed demand characteristics.

For our analysis, we restrict the demand system to the products that are offered. Define $\mathcal{J}_t^a \equiv \{j \in \mathcal{J} : a_{jt} = 1\}$, $\mathbf{s}_t^a \equiv (s_{jt} : j \in \mathcal{J}_t^a)$, and $\delta_t^a \equiv (\delta_{jt} : j \in \mathcal{J}_t^a)$. We write this subsystem as

$$\mathbf{s}_t^a = \mathbf{d}_t^a(\delta_t^a, \sigma). \quad (4)$$

Proposition 1 establishes that, for any configuration of $\mathbf{a}_t$, this subsystem is invertible with respect to δ_t^a (Berry, 1994).

PROPOSITION 1. *Suppose that the outside option $j = 0$ is always offered. Fix a market t , its prices and observed demand characteristics, and a configuration $\mathbf{a}_t \in \{0, 1\}^J$. Define the set of feasible interior market shares for the offered products as*

$$\mathcal{S}^a \equiv \left\{ \mathbf{s}^a \in (0, 1)^{|\mathcal{J}_t^a|} : \sum_{j \in \mathcal{J}_t^a} s_j < 1 \right\}.$$

Then, for every fixed value of σ , the mapping $\mathbf{d}_t^a(\cdot, \sigma)$ is a bijection from $\mathbb{R}^{|\mathcal{J}_t^a|}$ onto $\mathcal{S}^a$. Hence, for every $\mathbf{s}_t^a \in \mathcal{S}^a$, there exists a unique δ_t^a such that equation (4) holds, and the inverse $(\mathbf{d}_t^a)^{-1}(\cdot, \sigma)$ is well defined on $\mathcal{S}^a$. ■

Proof of Proposition 1: See Online Appendix D.1.

For every product-market observation with $a_{jt} = 1$, let $d_{jt}^{-1}(\cdot, \sigma) \equiv [(\mathbf{d}_t^a)^{-1}(\cdot, \sigma)]_j$ denote the j -th component of the inverse demand system. Then

$$d_{jt}^{-1}(\mathbf{s}_t^a, \sigma) = \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \xi_{jt}. \quad (5)$$

The following example illustrates Proposition 1 in the case of the nested logit model.

EXAMPLE 1 (Nested logit model). Let the products be partitioned into nests $\{\mathcal{J}_g\}_{g=1}^G$, with $g(j)$ denoting j 's nest; the outside option forms a singleton nest. Given $\mathbf{a}_t$ with $a_{jt} = 1$, define $s_{0t} \equiv 1 - \sum_{i \in \mathcal{J}_t^a} s_{it}$ and $s_{g,t}^a \equiv \sum_{i \in \mathcal{J}_g: a_{it}=1} s_{it}$. In the standard nested logit model (Berry, 1994), with nesting parameter $\sigma \in [0, 1)$, the share inversion is

$$\begin{aligned} d_{jt}^{-1}(s_t^a, \sigma) &= \ln\left(\frac{s_{jt}}{s_{0t}}\right) - \sigma \ln\left(\frac{s_{jt}}{s_{g(j),t}^a}\right) \\ &= \delta_{jt} = \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \zeta_{jt}. \end{aligned}$$

The shares above are positive, so the inverse is well defined whenever $a_{jt} = 1$. ■

Equation (5) for product j is available for estimation precisely when $a_{jt} = 1$. For values of $\mathbf{x}_t$ such that $\Pr(a_{jt} = 1 \mid \mathbf{x}_t) > 0$, the selection-bias function is $\mathbb{E}(\zeta_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$. Subtracting this conditional mean from ζ_{jt} yields a residual with conditional mean zero given $(\mathbf{x}_t, a_{jt} = 1)$.

Because firms make interdependent entry decisions in a game with common unobservables, rivals' decisions $\mathbf{a}_{-jt} \equiv (a_{it} : i \in \mathcal{J}, i \neq j)$ can be informative about ζ_{jt} . Hence, the configuration-specific selection-bias function $\mathbb{E}(\zeta_{jt} \mid \mathbf{x}_t, a_{jt} = 1, \mathbf{a}_{-jt})$ may vary with $\mathbf{a}_{-jt}$. Rival entry does not, however, change the structural form of product j 's inverted demand equation. For every configuration $\mathbf{a}_t$ with $a_{jt} = 1$, applying the inverse for that offered set yields equation (5) with the same parameters and demand shock ζ_{jt} as in equation (2), even though the inverse mapping, equilibrium prices, and market shares generally depend on the configuration. This differs from censored simultaneous-equation settings in which censoring changes the subsystem determining the observed outcome. By iterated expectations, the selection-bias function $\mathbb{E}(\zeta_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$ is the average of the configuration-specific functions over the distribution of $\mathbf{a}_{-jt}$ conditional on $(\mathbf{x}_t, a_{jt} = 1)$.

Controlling for $\mathbb{E}(\zeta_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$ is sufficient to remove selection bias, but one could instead control for the finer selection-bias function $\mathbb{E}(\zeta_{jt} \mid \mathbf{x}_t, a_{jt} = 1, \mathbf{a}_{-jt})$. The conditioning set used for the selection-bias function determines which instruments are orthogonal to the resulting demand residual by construction. Under the coarser correction, the residual is orthogonal to every function of $\mathbf{x}_t$ over the selected sample $a_{jt} = 1$, but need not be orthogonal to functions of $\mathbf{a}_{-jt}$. The maintained assumptions therefore justify instruments constructed from the characteristics $\mathbf{x}_t$ observed for all potential products, but do not generally justify instruments that

depend on realized rival entry. For example, $\sum_{i \neq j} x_{it}^D$ is admissible on exogeneity grounds, whereas $\sum_{i \neq j} x_{it}^D a_{it}$ is not generally admissible. Under the finer correction, the demand residual has conditional mean zero given $(x_t, a_{jt} = 1, a_{-jt})$ and is therefore orthogonal to functions of both x_t and a_{-jt} .

Although the finer selection-bias function broadens the class of admissible instruments, conditioning on a_{-jt} , which can take up to 2^{J-1} configurations, creates additional dimensionality and support challenges for identification and estimation. We therefore focus on the coarser selection-bias function $\mathbb{E}(\xi_{jt} \mid x_t, a_{jt} = 1)$.

As discussed in Online Appendix B, Proposition 1 continues to hold for multi-product firms because it depends only on the demand system, not on product ownership. Multi-product ownership changes the representation of the entry game, which must be formulated over firm-level product portfolios, but the product-level selection-bias function remains $\mathbb{E}(\xi_{jt} \mid x_t, a_{jt} = 1)$.

2.2 Price competition

Let x_{jt}^{mc} and x_{jt}^{fc} be product j 's variable- and fixed-cost subvectors of x_t , respectively, and let Π_{jt} denote its profit when active in market t :

$$\Pi_{jt} = p_{jt} q_{jt} - c(q_{jt}, x_{jt}^{\text{mc}}, \omega_{jt}) - fc_{jt}, \quad (6)$$

where $q_{jt} \equiv H_t s_{jt}$ is the quantity sold and H_t is market size, $c(q_{jt}, x_{jt}^{\text{mc}}, \omega_{jt})$ is the variable cost function, fc_{jt} is the fixed entry cost, and ω_{jt} is a variable-cost unobservable. Throughout the paper, we normalize $\mathbb{E}(\omega_{jt}) = 0$.

Conditional on the entry configuration a_t , and under the maintained assumption of single-product firms, every interior pure-strategy Bertrand–Nash equilibrium satisfies the first-order conditions

$$p_{jt} = mc_{jt} - s_{jt} \left[\frac{\partial s_{jt}}{\partial p_{jt}} \right]^{-1} \quad \text{for every } j \in \mathcal{J}_t^a, \quad (7)$$

where $mc_{jt} \equiv \partial c(q_{jt}, x_{jt}^{\text{mc}}, \omega_{jt}) / \partial q_{jt}$ is marginal cost. The own-price derivative is computed after substituting equation (2) into the market-share function (3); it therefore accounts for the effect of price through both mean utility and the random-coefficient term, while holding rivals' prices fixed. All terms are evaluated at the

equilibrium price vector.

Let $\xi_t \equiv (\xi_{jt} : j \in \mathcal{J})$ and $\omega_t \equiv (\omega_{jt} : j \in \mathcal{J})$. For any rivals' entry profile $\mathbf{a}_{-j} \in \{0, 1\}^{J-1}$, let

$$VP_j(\mathbf{a}_{-j}, \mathbf{x}_t, \xi_t, \omega_t)$$

denote firm j 's *indirect variable profit*: the variable profit $p_{jt}q_{jt} - c(q_{jt}, \mathbf{x}_{jt}^{\text{mc}}, \omega_{jt})$ evaluated at the Bertrand–Nash equilibrium conditional on $(a_{jt} = 1, \mathbf{a}_{-j}, \mathbf{x}_t, \xi_t, \omega_t)$.²

2.3 Market-entry game

We model product entry as the outcome of a static game among the J potential entrants. The model nests complete-information games such as those in [Ciliberto and Tamer \(2009\)](#) and [Ciliberto et al. \(2021\)](#), as well as incomplete-information games with common-knowledge unobservables, as in [Grieco \(2014\)](#) and [Aguirregabiria and Mira \(2019\)](#). It allows firms' information about demand and variable-cost unobservables at entry to range from knowing only their distributions to observing their realizations, with imperfect common signals as intermediate cases. This flexibility is central because firms' information determines the scope for endogenous selection on demand and variable-cost unobservables. Online Appendix C discusses how the framework can be adapted to a dynamic game of product entry and exit.

ASSUMPTION 1. *At the time of entry in market t , firm j 's information set is $(\mathbf{x}_t, \kappa_t, \eta_{jt})$.*

- a. The vector $\mathbf{x}_t$ is observable to the researcher and common knowledge among the firms.*
- b. The latent market state κ_t , unobserved by the researcher, collects all information about demand and variable-cost unobservables (ξ_t, ω_t) and fixed entry costs that is common knowledge among firms. It may include the entire vector (ξ_t, ω_t) , in which case firms know all demand and variable-cost unobservables when making their entry decisions.*
- c. The vector η_{jt} represents firm j 's private information about its fixed entry cost. These vectors are independently and identically distributed across firms with common distribution F_η and are independent of $(\xi_t, \omega_t, \kappa_t, \mathbf{x}_t)$. If F_η is degenerate, the entry game has complete information.*

²If the pricing game admits multiple equilibria, we take as given an equilibrium-selection rule, possibly indexed by an additional market-level unobservable. We otherwise leave this rule unrestricted. The indirect variable profit is evaluated at the equilibrium selected by this rule, and we suppress the selection index from the notation.

d. All variables unobserved by the researcher— $(\xi_t, \omega_t, \kappa_t, \{\eta_{jt} : j \in \mathcal{J}\})$ —are jointly independent of the exogenous observables in $\mathbf{x}_t$. ■

Assumption 1 separates firms' information at entry into the observed common state $\mathbf{x}_t$, the latent market state κ_t , and firm-specific private information η_{jt} . The content of κ_t determines what firms know about demand and variable-cost unobservables and hence the scope for entry to select products on those unobservables. Private information η_{jt} permits incomplete-information entry; when F_η is degenerate, the game has complete information. Any demand and variable-cost unobservables not observed at entry become common knowledge before active firms set prices in the Bertrand–Nash equilibrium described above. Finally, part [d] imposes the standard exogeneity of $\mathbf{x}_t$ with respect to all variables unobserved by the researcher. This exogeneity condition underpins identification in both demand and entry models.

We normalize the payoff from remaining inactive to zero. To simplify notation in the following expected-profit expressions, we use ξ_t locally as shorthand for the combined vector of demand and variable-cost unobservables (ξ_t, ω_t) ; when the distinction matters, we write the two components separately. For a hypothetical profile of rivals' entry decisions $\mathbf{a}_{-j} \in \{0, 1\}^{J-1}$, let $\pi_j(\mathbf{a}_{-j}, \mathbf{x}_t, \kappa_t, \eta_{jt})$ denote firm j 's expected profit from entering, given its information. Under Assumption 1,

$$\pi_j(\mathbf{a}_{-j}, \mathbf{x}_t, \kappa_t, \eta_{jt}) = \int VP_j(\mathbf{a}_{-j}, \mathbf{x}_t, \xi_t) dF_{j,\xi}(\xi_t | \kappa_t) - fc(\mathbf{x}_{jt}^{\text{fc}}, \kappa_t, \eta_{jt}). \quad (8)$$

Here $F_{j,\xi}(\cdot | \kappa_t)$ denotes firm j 's conditional beliefs about the combined vector of demand and variable-cost unobservables. Parts [c]–[d] of Assumption 1 allow $\mathbf{x}_t$ and η_{jt} to be omitted from the conditioning set. The fixed entry cost is known to firm j at entry and therefore lies outside the integral; write $fc_{jt} \equiv fc(\mathbf{x}_{jt}^{\text{fc}}, \kappa_t, \eta_{jt})$. If κ_t reveals all demand and variable-cost unobservables at entry, this conditional distribution is degenerate and the integral reduces to $VP_j(\mathbf{a}_{-j}, \mathbf{x}_t, \xi_t)$.

As noted in Assumption 1[c], the entry game is one of incomplete information unless each η_{jt} has a degenerate distribution. We describe its equilibria as Bayesian Nash Equilibria (BNE), a solution concept that reduces to a complete-information Nash equilibrium when each η_{jt} is degenerate. Given $(\mathbf{x}_t, \kappa_t)$, a BNE of this game can be represented as a J -tuple of entry probabilities, one for each firm, $(P_{jt} : j \in \mathcal{J})$. To describe this BNE, we first define a firm's expected profit function that accounts

for its uncertainty about other firms' entry decisions.

$$\pi_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt}) = \sum_{\mathbf{a}_{-j} \in \{0,1\}^{J-1}} \left(\prod_{i \neq j} [P_{it}]^{a_i} [1 - P_{it}]^{1-a_i} \right) \pi_j(\mathbf{a}_{-j}, \mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt}). \quad (9)$$

Firm j 's best response is to enter the market if and only if this expected profit is nonnegative. We can then define a BNE in this game as follows.³

DEFINITION 1. Bayesian Nash Equilibrium. Under Assumption 1 and given $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$, a BNE can be represented as a J -tuple of probabilities $\{P_{jt} \equiv P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t) : j \in \mathcal{J}\}$ that solves the following system of J best-response equations in the space of probabilities:

$$P_{jt} = \int \mathbb{1} \left\{ \pi_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt}) \geq 0 \right\} dF_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{jt}). \quad \blacksquare \quad (10)$$

2.4 Special cases in the literature

Several existing models of demand with endogenous product entry are nested within our framework. To highlight this generality, Table 1 summarizes selected contributions, focusing on the features relevant to this paper: the presence and nature of endogenous selection on unobservables, particularly whether it operates through demand-side unobservables. Most empirical applications of endogenous product entry assume that firms do not know demand or variable-cost unobservables at entry. These models therefore rule out—by assumption—selection on the demand and variable-cost unobservables entering variable profits. Two important exceptions are the models in [Ciliberto et al. \(2021\)](#) and [Li et al. \(2022\)](#).

3 Structure of the selection problem

3.1 Selection-bias function

Recall from equation (5) that, for any observation with $a_{jt} = 1$, the demand inversion yields $d_{jt}^{-1}(s_t^a, \sigma) = \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \xi_{jt}$. Whether product j is offered is determined

³If the entry game admits multiple BNE, we take an equilibrium-selection rule as given. We interpret $\boldsymbol{\kappa}_t$ as including its index, which we suppress from the notation.

Table 1: Models of endogenous product entry with differentiated-product demand

Paper	Industry – Product Entry	Selection on demand or variable-cost unobservables	What firms know about ξ_t at entry	Unobservables in entry cost	Entry game
Aguirregabiria and Ho (2012)	US airlines – Route (city-pair)	NO, once airline & route FEs are accounted for.	ξ_t is unknown. Firms know airline & route FEs but not demand shocks ξ_t .	Only η_{jt} . Private information shocks assumed independent of ξ_t .	Incomplete information dynamic game
Sweeting (2013)	US radio – Station genre	NO, once observable lagged variables are accounted for.	ξ_t is unknown. Follows AR(1). Firms know lagged but not current ξ_t .	Only η_{jt} . Private information shocks assumed independent of ξ_t .	Incomplete information dynamic game
Eizenberg (2014)	US home PC – PC models	NO. Selection on entry-cost unobservables, which are assumed independent of ξ_t .	ξ_t is unknown. Firms only observe the realizations of ξ_t after committing to product choices.	Includes common-knowledge κ_t , assumed independent of ξ_t .	Complete information static game
Fan and Yang (2020)	US smartphones – Phone models	NO. Selection on entry-cost unobservables, which are assumed independent of ξ_t .	ξ_t is unknown. Firms know brand FEs but not demand shocks ξ_t .	Includes common-knowledge κ_t , assumed independent of ξ_t .	Complete information static game
Ciliberto et al. (2021)	US airlines – Route (city-pair)	YES	ξ_t is known. Demand unobservables are known to firms when making product choices.	Includes common-knowledge κ_t that can be correlated with ξ_t .	Complete information static game
Li et al. (2022)	US airlines – Route (city-pair)	YES	ξ_t is known. Demand unobservables are known to firms when making product choices.	Includes common-knowledge κ_t that can be correlated with ξ_t .	Complete information static game
Bontemps et al. (2023)	US airlines – Airline’s network of non-stop routes	NO. Selection on entry-cost unobservables, which are assumed independent of ξ_t .	ξ_t is unknown. Firms don’t know demand unobservables when making network choices.	Includes common-knowledge κ_t , assumed independent of ξ_t .	Complete information static game

by firm j 's equilibrium entry decision:

$$a_{jt} = \mathbb{1} \left\{ \pi_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt}) \geq 0 \right\}. \quad (11)$$

Since the sample of observations for which the demand equation holds is selected, we decompose the unobservable ξ_{jt} into its conditional mean and a residual: $\xi_{jt} = \lambda_j(\mathbf{x}_t) + \tilde{\xi}_{jt}$, where $\lambda_j(\mathbf{x}_t) \equiv \mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$ is the selection-bias function and $\mathbb{E}(\tilde{\xi}_{jt} \mid \mathbf{x}_t, a_{jt} = 1) = 0$ by construction. This gives the following regression equation for any observation with $a_{jt} = 1$:

$$d_{jt}^{-1}(\mathbf{s}_t^a, \boldsymbol{\sigma}) = \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \lambda_j(\mathbf{x}_t) + \tilde{\xi}_{jt}. \quad (12)$$

Let F_κ denote the marginal distribution of $\boldsymbol{\kappa}_t$. The structure of the selection-bias function plays a key role in the identification of $(\alpha, \boldsymbol{\beta}, \boldsymbol{\sigma})$. To characterize this structure, define the *ordinary* propensity score—the probability of product j being offered, conditional on observables—as:

$$\bar{P}_j(\mathbf{x}_t) \equiv \Pr(a_{jt} = 1 \mid \mathbf{x}_t) = \int \mathbb{1} \left\{ \pi_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt}) \geq 0 \right\} dF_\kappa(\boldsymbol{\kappa}_t) dF_\eta(\boldsymbol{\eta}_{jt}). \quad (13)$$

For values of $\mathbf{x}_t$ such that $\bar{P}_j(\mathbf{x}_t) > 0$, the selection-bias function can be written as:

$$\lambda_j(\mathbf{x}_t) = \frac{\mathbb{E}(\xi_{jt} a_{jt} \mid \mathbf{x}_t)}{\bar{P}_j(\mathbf{x}_t)}. \quad (14)$$

It is well known that estimating equation (12) by instrumental variables—treating $\lambda_j(\mathbf{x}_t) + \tilde{\xi}_{jt}$ as the composite error—is generally invalid. If $\lambda_j(\mathbf{x}_t)$ is omitted, instruments constructed from $\mathbf{x}_t$ generally enter the composite error through the omitted selection-bias function and therefore violate the exclusion restriction (Wooldridge, 2010). A natural alternative is a selection-function approach that explicitly controls for $\lambda_j(\mathbf{x}_t)$. However, without additional structure, $\lambda_j(\mathbf{x}_t)$ remains an unrestricted function of the full observed state $\mathbf{x}_t$. In particular, the direct effect of $\mathbf{x}_{jt}^D$ on consumer demand ($\boldsymbol{\beta}$) cannot be separated from its indirect effect through selection; absent further restrictions, the demand parameters are not identified.

In this setting, the standard strategy in the literature is to impose conditions under which the selection-bias function depends only on the ordinary propensity score

$\bar{P}_j(\mathbf{x}_t)$, that is, $\lambda_j(\mathbf{x}_t) = g_j(\bar{P}_j(\mathbf{x}_t))$ for some function $g_j(\cdot)$. This restriction reduces the dimensionality of the nonparametric problem: g_j remains unrestricted, but its argument is the scalar $\bar{P}_j(\mathbf{x}_t)$ rather than the full vector $\mathbf{x}_t$. With sufficient variation in the demand regressors conditional on the ordinary propensity score, their direct effects can be distinguished from the selection effect captured by g_j . Estimation then follows a standard two-step procedure. In the first step, the ordinary propensity score $\bar{P}_j(\mathbf{x}_t)$ is estimated using data on $(a_{jt}, \mathbf{x}_t)$. In the second step, the structural parameters are estimated using, for example, the series estimators in [Das et al. \(2003\)](#) and [Newey \(2009\)](#), or the pairwise differencing approaches in [Powell \(2001\)](#) and [Aradillas-Lopez \(2012\)](#). After controlling for $g_j(\bar{P}_j(\mathbf{x}_t))$, price endogeneity due to simultaneity can be handled using familiar BLP-type instruments constructed from the exogenous characteristics of other potential products.

3.2 Failure of the ordinary propensity score

The restriction that the selection-bias function depend solely on the ordinary propensity score need not hold in models of endogenous product entry and oligopoly competition, even under simple specifications. Intuitively, entry decisions in these environments depend on equilibrium profitability, which is jointly determined by the demand and variable-cost unobservables of all competing products. As a result, the selection rule need not admit the single-index structure under which the ordinary propensity score controls for selection bias.

Proposition 2 summarizes benchmark conditions under which the ordinary propensity score is sufficient. Under the maintained exogeneity conditions, the monotonicity condition used in local average treatment effects (LATE) is equivalent to a scalar-threshold representation of the entry indicator, as in Theorem 1 of [Vytlacil \(2002\)](#). Either of these equivalent conditions is sufficient for the ordinary propensity score to control for selection bias, following Propositions 2 and 3 in [Angrist \(1997\)](#).

PROPOSITION 2. *Under the independence of $(\xi_t, \kappa_t, \eta_{jt})$ from $\mathbf{x}_t$ implied by Assumption 1[d], consider the following conditions:*

- a. *Monotonicity: For any two observable states $\mathbf{x}$ and $\mathbf{x}'$, either $\mathbb{1}\{\pi_j^P(\mathbf{x}, \kappa, \eta_j) \geq 0\} \geq \mathbb{1}\{\pi_j^P(\mathbf{x}', \kappa, \eta_j) \geq 0\}$ for almost every (κ, η_j) , or the reverse inequality holds for almost every (κ, η_j) .*

- b. *Single-index representation*: There are real-valued functions h_{1j} and h_{2j} such that, for every observable state $\mathbf{x}$, $\pi_j^P(\mathbf{x}, \boldsymbol{\kappa}, \boldsymbol{\eta}_j) \geq 0 \iff h_{1j}(\mathbf{x}) \geq h_{2j}(\boldsymbol{\kappa}, \boldsymbol{\eta}_j)$ for almost every $(\boldsymbol{\kappa}, \boldsymbol{\eta}_j)$.
- c. *Conditional independence*: the joint distribution of (ξ_{jt}, a_{jt}) conditional on $(\mathbf{x}_t, \bar{P}_j(\mathbf{x}_t))$ coincides with its distribution conditional on $\bar{P}_j(\mathbf{x}_t)$ alone.

Under the conditions of Theorem 1 in [Vytlacil \(2002\)](#), with the separability extension stated there for uncountable support, conditions (a) and (b) are equivalent. Furthermore, either condition (a) or (b) implies condition (c). Condition (c) is precisely the requirement that the ordinary propensity score $\bar{P}_j(\mathbf{x}_t)$ be sufficient for the joint distribution of (ξ_{jt}, a_{jt}) . For the selected demand equation and values with $\bar{P}_j(\mathbf{x}_t) > 0$, it implies the weaker conditional-mean restriction that $\lambda_j(\mathbf{x}_t)$ depends on $\mathbf{x}_t$ only through $\bar{P}_j(\mathbf{x}_t)$. ■

The following example provides a simple oligopoly model with endogenous entry satisfying Assumption 1 in which both condition (c) and the weaker conditional-mean restriction for the selected demand equation fail.

EXAMPLE 2 (Failure of ordinary propensity-score sufficiency). Consider a two-product setting satisfying Assumption 1. Product k is always offered, while product j enters endogenously, and demand follows the standard logit model. We omit the market subscript and fix market size $H > 0$. Constant marginal costs $mc_i(\mathbf{x}_i^{\text{mc}})$, exogenous price-cost margins $PCM_i(\mathbf{x})$, and the fixed entry cost $fc_j(\mathbf{x}_j^{\text{fc}})$ depend only on observables. Define $v_i(\mathbf{x}) \equiv (\mathbf{x}_i^D)' \boldsymbol{\beta} + \alpha[mc_i(\mathbf{x}_i^{\text{mc}}) + PCM_i(\mathbf{x})]$ for $i \in \{j, k\}$. Demand shocks $\boldsymbol{\xi} = (\xi_j, \xi_k)$ are common knowledge at entry and there is no private information, so $\boldsymbol{\kappa} = \boldsymbol{\xi}$ and $\boldsymbol{\eta}_j$ is degenerate. Assume $0 < fc_j(\mathbf{x}_j^{\text{fc}}) < PCM_j(\mathbf{x})H$ on the support considered below. Firm j 's entry profit and the equivalent entry rule are

$$\begin{aligned} \pi_j^P &= PCM_j(\mathbf{x})H \frac{e^{v_j(\mathbf{x}) + \xi_j}}{1 + e^{v_j(\mathbf{x}) + \xi_j} + e^{v_k(\mathbf{x}) + \xi_k}} - fc_j(\mathbf{x}_j^{\text{fc}}), \\ \pi_j^P \geq 0 &\iff [PCM_j(\mathbf{x})H - fc_j(\mathbf{x}_j^{\text{fc}})]e^{v_j(\mathbf{x}) + \xi_j} \geq fc_j(\mathbf{x}_j^{\text{fc}})[1 + e^{v_k(\mathbf{x}) + \xi_k}], \\ a_j &= \mathbb{1}\left\{\xi_j \geq \ln[1 + e^{v_k(\mathbf{x}) + \xi_k}] - B_j(\mathbf{x})\right\}, \quad B_j(\mathbf{x}) \equiv v_j(\mathbf{x}) + \ln \frac{PCM_j(\mathbf{x})H - fc_j(\mathbf{x}_j^{\text{fc}})}{fc_j(\mathbf{x}_j^{\text{fc}})}. \end{aligned}$$

Suppose ξ_j and ξ_k are independent and jointly independent of $\mathbf{x}$. Let ξ_j have mean zero, $\mathbb{E}|\xi_j| < \infty$, and a continuously differentiable CDF F with density $f(z) > 0$ for every

$z \in \mathbb{R}$. Let $\xi_k \in \{u_L, u_H\}$ with equal probability, where $u_H = -u_L > 0$. Assume also that $(v_k(\mathbf{x}), B_j(\mathbf{x}))$ has a density that is strictly positive on a nonempty open set. For $m \in \{L, H\}$, define

$$t_m(v, b) \equiv \ln(1 + e^{v+u_m}) - b, \quad r_m(v) \equiv \frac{e^{v+u_m}}{1 + e^{v+u_m}},$$

$$P(v, b) \equiv 1 - \frac{1}{2}[F(t_L(v, b)) + F(t_H(v, b))], \quad N(v, b) \equiv \frac{1}{2} \sum_{m \in \{L, H\}} \int_{t_m(v, b)}^{\infty} z f(z) dz.$$

The threshold rule implies $\bar{P}_j(\mathbf{x}) = P(v_k(\mathbf{x}), B_j(\mathbf{x}))$ and $\mathbb{E}(\xi_j a_j \mid \mathbf{x}) = N(v_k(\mathbf{x}), B_j(\mathbf{x}))$. Since $\partial t_m / \partial v = r_m$ and $\partial t_m / \partial b = -1$, direct differentiation gives

$$\det\left(\frac{\partial(P, N)}{\partial(v, b)}\right) = \frac{1}{4}f(t_L)f(t_H)(r_H - r_L)(t_H - t_L) > 0,$$

because $u_H > u_L$ implies $r_H > r_L$ and $t_H > t_L$.

The inverse function theorem and the local-density assumption imply that (P, N) has a two-dimensional absolutely continuous component. It therefore cannot satisfy $N = n(P)$ almost surely for any measurable scalar function n , as condition (c) in Proposition 2 would require. Thus condition (c) fails. Since $0 < \bar{P}_j(\mathbf{x}) < 1$, the ratio $\lambda_j(\mathbf{x}) = \mathbb{E}(\xi_j a_j \mid \mathbf{x}) / \bar{P}_j(\mathbf{x})$ is not a function of the ordinary propensity score either. Finally, because either condition (a) or condition (b) would imply condition (c), both fail. ■

Example 2 shows that, even in simple environments, the ordinary propensity score need not be sufficient in the sense of condition (c): with $\kappa_t = \xi_t$, the rival's demand shock, ξ_{kt} , affects firm j 's entry threshold through oligopoly competition. As a result, selection depends on the full latent market state κ_t , whereas the ordinary propensity score averages over κ_t and loses relevant information.

3.3 Latent propensity scores and the selection-bias function

We now derive a representation of the selection-bias function $\lambda_j(\mathbf{x}_t)$ using the equilibrium entry probabilities from Definition 1, which satisfy

$$P_j(\mathbf{x}_t, \kappa_t) = \Pr(a_{jt} = 1 \mid \mathbf{x}_t, \kappa_t).$$

We refer to $P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t)$ as the *latent propensity score*, to highlight its dependence on the latent market state $\boldsymbol{\kappa}_t$. By Assumption 1[d], the ordinary propensity score is the average of the latent propensity scores over the marginal distribution of $\boldsymbol{\kappa}_t$:

$$\bar{P}_j(\mathbf{x}_t) = \int P_j(\mathbf{x}_t, \boldsymbol{\kappa}) dF_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}). \quad (15)$$

Proposition 3 shows that, under only the information and exogeneity conditions in Assumption 1, the selection-bias function admits a mixture representation in terms of these latent propensity scores. Unlike the ordinary propensity-score approach, this representation does not require any monotonicity or single-index assumption.

PROPOSITION 3. *Under Assumption 1, and for values of $\mathbf{x}_t$ such that $\bar{P}_j(\mathbf{x}_t) > 0$, the selection-bias function $\lambda_j(\mathbf{x}_t)$ admits the following mixture representation:*

$$\lambda_j(\mathbf{x}_t) = \int \left[\frac{P_j(\mathbf{x}_t, \boldsymbol{\kappa})}{\bar{P}_j(\mathbf{x}_t)} \right] \mu_j(\boldsymbol{\kappa}) dF_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}), \quad (16)$$

with $\mu_j(\boldsymbol{\kappa}) \equiv \mathbb{E}(\xi_{jt} \mid \boldsymbol{\kappa}_t = \boldsymbol{\kappa})$. ■

Proof: See Online Appendix D.2.

Proposition 3 provides the first building block of our identification strategy by expressing the selection-bias function as a mixture involving latent propensity scores. The next section develops two routes from this representation to identification of the selection-bias functions and demand parameters.

4 Identification and estimation

The researcher observes T local markets, each with the same J potential entrants. In every market t , the researcher observes the exogenous variables $\mathbf{x}_t$ and firms' entry decisions $\mathbf{a}_t \in \{0, 1\}^J$; for every active firm j , the researcher also observes p_{jt} and s_{jt} .

Let the vector of demand parameters be $\boldsymbol{\theta} \equiv (\alpha, \boldsymbol{\beta}', \boldsymbol{\sigma}')'$. We adopt a sequential identification strategy for $\boldsymbol{\theta}$. In the first step, we use the joint distribution of firms' entry decisions conditional on $\mathbf{x}_t$ to identify the probability objects required by the

two identification routes described below. In the second step, we use the selected-sample demand equations to identify the demand parameters θ , together with the remaining components of the selection-bias functions.

The econometric model consists of three sets of objects: (i) the selected-sample demand equations in (12); (ii) the selection-bias representation in (16); and (iii) the equilibrium-implied joint distribution of firms' entry decisions conditional on $\mathbf{x}_t$. Under Assumption 1, for every $\mathbf{a} \in \{0, 1\}^J$, its conditional probability mass function has the mixture representation

$$\Pr(\mathbf{a}_t = \mathbf{a} \mid \mathbf{x}_t) = \int \left(\prod_{j=1}^J P_j(\mathbf{x}_t, \boldsymbol{\kappa})^{a_j} [1 - P_j(\mathbf{x}_t, \boldsymbol{\kappa})]^{1-a_j} \right) dF_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}). \quad (17)$$

Conditional on $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$, firms' entry decisions depend only on their private shocks, which are assumed to be independent across firms. As a result, entry decisions are conditionally independent given $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$, and by Assumption 1[d] the conditional mixing distribution is the marginal distribution $F_{\boldsymbol{\kappa}}$ and does not depend on $\mathbf{x}_t$. All residual dependence across firms' entry decisions—beyond what is explained by observables—is therefore driven by the unobservables in $\boldsymbol{\kappa}_t$. This structure provides a key source of identifying variation. The joint distribution of $\mathbf{a}_t$ conditional on $\mathbf{x}_t$, and in particular cross-sectional dependence in entry decisions, therefore contains information about how the latent market state shifts firms' entry probabilities. The additional assumptions introduced below determine how this information is mapped into the selection-bias function.

We consider two complementary routes to an identified representation of the selection-bias function, each based on assumptions that may suit different economic applications. The first route—developed in Online Appendix E—assumes that the latent market state coincides with the vector of demand shocks, $\boldsymbol{\kappa}_t = \boldsymbol{\xi}_t$, and that $\boldsymbol{\xi}_t$ is jointly normal. Under additional index, exclusion, and smoothness restrictions, a multivariate version of Stein's lemma then delivers the closed form

$$\lambda_j(\mathbf{x}_t) = \sum_{k=1}^J \varrho_{jk} \frac{\partial \log \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}}, \quad (18)$$

where x_{k1t} is a continuous regressor entering product k 's latent entry index with coefficient τ_{k1} , $\Sigma_{jk} \equiv \text{Cov}(\xi_{jt}, \xi_{kt})$, and $\varrho_{jk} \equiv \Sigma_{jk} / \tau_{k1}$ (for details, see Proposition 5 in

Online Appendix E). The derivative regressors entering the selection-bias function can be constructed from a first-step estimate of $\bar{P}_j(\mathbf{x}_t)$, and their coefficients are estimated jointly with the demand parameters in the second step. The second route, developed in the remainder of this section, assumes instead an exact finite-mixture representation of the joint entry distribution and the selection-bias functions.

4.1 Two-step finite-mixture approach

The support $\mathcal{K}$ of the latent market state κ_t need not be finite or discrete. For example, in our differentiated-products setting, κ_t may collect the product-specific demand and variable-cost unobservables, $\kappa_t = (\xi_{jt}, \omega_{jt} : j \in \mathcal{J}) \in \mathbb{R}^{2J}$.

Equation (16) shows that the selection-bias function depends on the latent market state only through three objects: the latent propensity score $P_j(\mathbf{x}_t, \kappa)$, the marginal distribution F_κ , and the conditional mean function $\mu_j(\kappa)$. However, without additional structure, $\lambda_j(\mathbf{x}_t)$ remains an unrestricted nuisance function of the observed state $\mathbf{x}_t$. We therefore need further dimensionality-reducing restrictions under which the selection-bias functions can be represented as linear combinations of functions recoverable from the joint distribution of entry decisions, with the remaining coefficients identified jointly with the demand parameters.

For an integer $L \geq 1$, let $\{\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_L\}$ denote a partition of $\mathcal{K}$ into positive-probability cells. Each element of the partition can be interpreted as a *latent market type*. For each latent market type $\ell = 1, 2, \dots, L$, define its probability

$$\tilde{f}_\ell \equiv \Pr(\kappa_t \in \mathcal{K}_\ell) = F_\kappa(\mathcal{K}_\ell). \quad (19)$$

Define the type-specific propensity score

$$\tilde{P}_{j,\ell}(\mathbf{x}_t) \equiv \frac{1}{\tilde{f}_\ell} \int_{\mathcal{K}_\ell} P_j(\mathbf{x}_t, \kappa) dF_\kappa(\kappa), \quad (20)$$

as well as the type-specific average of the function $\mu_j(\kappa)$:

$$\tilde{\mu}_{j,\ell} \equiv \mathbb{E}(\xi_{jt} \mid \kappa_t \in \mathcal{K}_\ell) = \frac{1}{\tilde{f}_\ell} \int_{\mathcal{K}_\ell} \mu_j(\kappa) dF_\kappa(\kappa). \quad (21)$$

For every $\mathbf{a} = (a_j : j \in \mathcal{J}) \in \{0, 1\}^J$, these objects define the L -type finite-mixture

conditional probability mass function

$$\tilde{\text{Pr}}^L(\mathbf{a}_t = \mathbf{a} \mid \mathbf{x}_t) \equiv \sum_{\ell=1}^L \tilde{f}_\ell \left(\prod_{j=1}^J \tilde{P}_{j,\ell}(\mathbf{x}_t)^{a_j} [1 - \tilde{P}_{j,\ell}(\mathbf{x}_t)]^{1-a_j} \right). \quad (22)$$

For every such partition, the law of total probability gives

$$\bar{P}_j(\mathbf{x}_t) = \sum_{\ell=1}^L \tilde{P}_{j,\ell}(\mathbf{x}_t) \tilde{f}_\ell. \quad (23)$$

The same finite-mixture objects define the L -type selection-bias function

$$\tilde{\lambda}_j^L(\mathbf{x}_t) \equiv \sum_{\ell=1}^L \left[\frac{\tilde{P}_{j,\ell}(\mathbf{x}_t)}{\bar{P}_j(\mathbf{x}_t)} \right] \tilde{\mu}_{j,\ell} \tilde{f}_\ell. \quad (24)$$

Expression (24) is the finite-mixture analog of the mixture representation in Proposition 3. For a given partition, the finite-mixture selection-bias function is summarized by $\{\tilde{f}_\ell\}_{\ell=1}^L$, $\{\tilde{P}_{j,\ell}(\mathbf{x}_t)\}_{\ell=1}^L$, and $\{\tilde{\mu}_{j,\ell}\}_{\ell=1}^L$.

For a generic partition, the finite-mixture distribution in (22) and the finite-mixture selection-bias functions in (24) need not reproduce the true joint distribution of entry decisions and the selection-bias functions. The next assumption imposes exact reproduction of both objects together with an observable rank restriction.

ASSUMPTION 2 (Exact finite-mixture representation and rank). *There exist an integer $L^* < \infty$ and a partition $\{\mathcal{K}_1, \dots, \mathcal{K}_{L^*}\}$ of $\mathcal{K}$ into positive-probability cells whose implied objects, defined in (19)–(21), satisfy, almost surely in $\mathbf{x}_t$,*

$$\begin{aligned} \text{Pr}(\mathbf{a}_t = \mathbf{a} \mid \mathbf{x}_t) &= \tilde{\text{Pr}}^{L^*}(\mathbf{a}_t = \mathbf{a} \mid \mathbf{x}_t) && \text{for every } \mathbf{a} \in \{0, 1\}^J, \\ \lambda_j(\mathbf{x}_t) &= \tilde{\lambda}_j^{L^*}(\mathbf{x}_t) && \text{for every } j \text{ with } \bar{P}_j(\mathbf{x}_t) > 0. \end{aligned}$$

Moreover, for some partition of $\mathcal{J}$ into two nonempty groups, the matrix of their joint entry-profile probabilities conditional on $\mathbf{x}_t$, with rows and columns indexed by the two groups' entry profiles, has rank L^* for almost every $\mathbf{x}_t$ in a set with positive probability. Finally, the type probabilities are pairwise distinct. ■

The exact representation and the rank condition identify L^* from the joint distribution of entry decisions. For the grouping in Assumption 2, the L -type conditional

probability matrix in (22) is a sum of L rank-one matrices and therefore has rank at most L . The rank L^* imposed by Assumption 2 thus rules out every representation with fewer than L^* latent market types, while the exact representation in the assumption supplies one with exactly L^* types. Hence, L^* is the minimal number of latent market types that can reproduce the joint distribution of entry decisions and is identified from that distribution. Without the rank equality, the observable rank would provide only a lower bound. This argument follows Proposition 4 in Kasahara and Shimotsu (2014), which establishes the rank equality generically when L^* does not exceed the number of entry profiles in either group.

The latent market state need not have discrete support: Assumption 2 requires only that a finite partition reproduce the joint distribution of entry decisions and the selection-bias functions, not that the resulting representation recover F_κ itself or the latent propensity scores $P_j(\mathbf{x}_t, \kappa)$ pointwise. It also excludes latent heterogeneity driving selection that leaves no trace in cross-firm dependence in entry decisions. If the heterogeneity relevant to each firm's entry decision is firm-specific, independent across firms, and affects only that firm's latent propensity score, entry decisions are independent conditional on $\mathbf{x}_t$ and their joint distribution admits a one-type representation. The entry data then cannot reveal this heterogeneity even when it generates nonzero selection bias, so the exact-representation requirement fails.

4.2 First-step identification

The rank condition in Assumption 2 identifies L^* but not the type probabilities $\{\tilde{f}_\ell\}_{\ell=1}^{L^*}$ and type-specific propensity scores $\{\tilde{P}_{j,\ell}(\mathbf{x}_t)\}_{\ell=1}^{L^*}$. Conditional on L^* , Corollary 5 of Allman et al. (2009) establishes generic identification of these objects when $J \geq 2\lceil \log_2 L^* \rceil + 1$. To describe the underlying rank condition, for $L^* > 1$, partition the products into three nonempty groups and, for each group and almost every $\mathbf{x}_t$, form the matrix whose ℓ -th row contains the probabilities of that group's entry profiles for latent market type ℓ . Given the positive type probabilities in Assumption 2, Corollary 2 of Allman et al. (2009) identifies the type probabilities and the probabilities of the three groups' entry profiles, up to a simultaneous permutation of the latent market types, if the sum of the row Kruskal ranks of these matrices is at least $2L^* + 2$. Marginalizing these probabilities then identifies each $\tilde{P}_{j,\ell}(\mathbf{x}_t)$.

The preceding identification result is pointwise in $\mathbf{x}_t$. Assumption 1[d] makes

the type probabilities common across values of $\mathbf{x}_t$, and their pairwise distinctness in Assumption 2 permits the latent market types to be matched across those values. Without loss of generality, we label them so that $\tilde{f}_1 > \dots > \tilde{f}_{L^*} > 0$. This labeling fixes the remaining global permutation, which merely relabels the type-specific objects and leaves the finite-mixture distribution in (22) and the finite-mixture selection-bias functions in (24) unchanged.

4.3 Second-step identification

Under Assumption 2 and the preceding first-step identification conditions, the selection-bias function is linear in the unknown type-specific averages $\{\tilde{\mu}_{j,\ell}\}_{\ell=1}^{L^*}$. Their definition and the normalization $\mathbb{E}(\tilde{\zeta}_{jt}) = 0$ imply $\sum_{\ell=1}^{L^*} \tilde{\mu}_{j,\ell} \tilde{f}_\ell = 0$, so only $L^* - 1$ coefficients are free. Taking type L^* as the reference, so that $\tilde{\mu}_{j,L^*} = -\sum_{\ell=1}^{L^*-1} (\tilde{f}_\ell / \tilde{f}_{L^*}) \tilde{\mu}_{j,\ell}$, the selected-sample demand equation becomes

$$\begin{aligned} d_{jt}^{-1}(\mathbf{s}_t^a, \sigma) &= \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \sum_{\ell=1}^{L^*-1} \tilde{\mu}_{j,\ell} r_{j\ell t} + \tilde{\zeta}_{jt}, \\ r_{j\ell t} &\equiv \frac{\tilde{P}_{j,\ell}(\mathbf{x}_t) - \tilde{P}_{j,L^*}(\mathbf{x}_t)}{\tilde{P}_j(\mathbf{x}_t)} \tilde{f}_\ell, \quad \ell = 1, \dots, L^* - 1. \end{aligned} \tag{25}$$

Let $\tilde{\boldsymbol{\mu}}_j \equiv (\tilde{\mu}_{j,1}, \dots, \tilde{\mu}_{j,L^*-1})'$, $\mathbf{r}_{jt} \equiv (r_{j1t}, \dots, r_{j,L^*-1,t})'$, and stack $\tilde{\boldsymbol{\mu}} \equiv (\tilde{\boldsymbol{\mu}}_1', \dots, \tilde{\boldsymbol{\mu}}_J')$. Define $\boldsymbol{\vartheta} \equiv (\boldsymbol{\theta}', \tilde{\boldsymbol{\mu}}')'$, and let $\boldsymbol{\vartheta}_0$ denote its true value. Let $\mathbf{R}_{jt}$ be the $J(L^* - 1)$ -vector whose j -th $(L^* - 1)$ -dimensional block is $\mathbf{r}_{jt}$ and whose remaining blocks are zero, so that $\mathbf{R}_{jt}' \tilde{\boldsymbol{\mu}} = \mathbf{r}_{jt}' \tilde{\boldsymbol{\mu}}_j$. When $L^* = 1$, all these selection-function vectors are interpreted as empty.⁴ For each j , let $\mathbf{z}_{jt} \in \mathbb{R}^{K_z}$, with $K_z \geq \dim(\sigma) + 1$, contain excluded instruments constructed from $\mathbf{x}_t$, such as characteristics of potential entrants other than j , and define $\mathbf{w}_{jt} \equiv (\mathbf{z}_{jt}', \mathbf{x}_{jt}^{D'}, \mathbf{R}_{jt}')'$. Up to an irrelevant positive normalization, the pooled selected-sample moment function is

$$\mathbf{m}(\boldsymbol{\vartheta}) \equiv \sum_{j=1}^J \Pr(a_{jt} = 1) \mathbb{E} \left(\mathbf{w}_{jt} \left(d_{jt}^{-1}(\mathbf{s}_t^a, \sigma) - \alpha p_{jt} - \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} - \mathbf{R}_{jt}' \tilde{\boldsymbol{\mu}} \right) \mid a_{jt} = 1 \right).$$

By equation (25), the residual inside $\mathbf{m}(\boldsymbol{\vartheta})$, evaluated at $\boldsymbol{\vartheta}_0$, equals $\tilde{\zeta}_{jt}$. Because every component of $\mathbf{w}_{jt}$ is a function of $\mathbf{x}_t$, iterated expectations and $\mathbb{E}(\tilde{\zeta}_{jt} \mid \mathbf{x}_t, a_{jt} =$

⁴If $L^* = 1$, Assumption 2 and the normalization $\mathbb{E}(\tilde{\zeta}_{jt}) = 0$ imply $\lambda_j(\mathbf{x}_t) = 0$.

1) = 0 establish $\mathbf{m}(\boldsymbol{\vartheta}_0) = \mathbf{0}$. The standard rank condition for local identification in [Rothenberg \(1971\)](#) then yields the following result.

PROPOSITION 4. *Suppose that $\Pr(a_{jt} = 1) > 0$ for every j , $\boldsymbol{\vartheta}_0$ lies in the interior of the parameter space, the expectations defining $\mathbf{m}$ are finite, $\mathbf{m}$ is continuously differentiable in a neighborhood of $\boldsymbol{\vartheta}_0$, and*

$$\mathbf{M}(\boldsymbol{\vartheta}_0) \equiv \left. \frac{\partial \mathbf{m}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}'} \right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0}$$

has full column rank. Then $\boldsymbol{\vartheta}_0$ is locally identified. ■

Proposition 4 completes the sequential identification argument. The first step identifies L^* , the type probabilities, and the type-specific propensity scores, and hence the selection-function regressors $\mathbf{r}_{jt}$. Assumption 2 guarantees that these regressors span the selection-bias functions. Conditional on these objects, the pooled demand moments locally identify the common demand parameters $\boldsymbol{\theta}$ jointly with the product-specific selection-function coefficients $\{\tilde{\boldsymbol{\mu}}_j\}_{j=1}^J$. Thus, identification of $\boldsymbol{\theta}$ requires neither a specification of entry payoffs nor a pricing equation.

Importantly, as discussed after Proposition 1 in Section 2.1, our selection correction conditions on $(\mathbf{x}_t, a_{jt} = 1)$, but not on the rivals' entry decisions $\mathbf{a}_{-jt}$. Admissible instruments $\mathbf{z}_{jt}$ can therefore be functions of the characteristics $\mathbf{x}_t$ of all potential entrants, but cannot generally depend on the rivals' actual entry decisions. Using such instruments would require either additional restrictions or the finer selection correction $\mathbb{E}(\zeta_{jt} \mid \mathbf{x}_t, a_{jt} = 1, \mathbf{a}_{-jt})$, complicating both identification and estimation.

4.3.1 Identification of marginal costs and fixed entry costs

Once demand parameters are identified and a form of competition is specified, price-cost margins can be recovered from firms' profit maximization conditions. Combining these with observed prices yields realized marginal costs mc_{jt} .

Suppose that marginal cost, possibly after a known transformation (the logarithm in our application), is modeled as a function of observable product characteristics $\mathbf{x}_{jt}^{\text{mc}}$ plus the additive variable-cost unobservable ω_{jt} . Estimating this marginal-cost equation raises the same selection-bias issues as in demand estimation. In particular, the equation involves the selection-bias function $\lambda_j^{\text{mc}}(\mathbf{x}_t) = \mathbb{E}(\omega_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$.

By Assumption 1[c]–[d], the private entry shock η_{jt} is independent of ω_t conditional on (x_t, κ_t) , and x_t is independent of the unobservables, so this selection-bias function has the same structure as in Proposition 3:

$$\lambda_j^{\text{mc}}(x_t) = \int \left[\frac{P_j(x_t, \kappa)}{\bar{P}_j(x_t)} \right] \mu_j^{\text{mc}}(\kappa) dF_\kappa(\kappa), \quad (26)$$

with $\mu_j^{\text{mc}}(\kappa) \equiv \mathbb{E}(\omega_{jt} \mid \kappa_t = \kappa)$ and $\tilde{\mu}_{j,\ell}^{\text{mc}} \equiv \mathbb{E}(\omega_{jt} \mid \kappa_t \in \mathcal{K}_\ell)$. The normalization $\mathbb{E}(\omega_{jt}) = 0$ implies $\sum_{\ell=1}^{L^*} \tilde{\mu}_{j,\ell}^{\text{mc}} \tilde{f}_\ell = 0$. Under an extension of Assumption 2 requiring the same partition also to reproduce the marginal-cost selection-bias functions, the same reference-type reparameterization and finite-mixture identification strategy can be applied to recover the marginal-cost parameters and the corresponding selection-bias functions.

To recover fixed costs, we further associate each latent market type with the type-specific conditional means $\tilde{\mu}_{j,\ell}$ and $\tilde{\mu}_{j,\ell}^{\text{mc}}$ of the demand and variable-cost unobservables. Under this additional type-level restriction and an identified variable-cost specification, the demand and marginal-cost parameters and these type-specific conditional means determine equilibrium variable profits for any given x_t , latent market type $\ell = 1, \dots, L^*$, and counterfactual entry configuration $\mathbf{a} \in \{0, 1\}^J$. In particular, for a given x_t and latent type ℓ , we obtain equilibrium price-cost margins, market shares, and variable profits by setting $\xi_{jt} = \tilde{\mu}_{j,\ell}$ and $\omega_{jt} = \tilde{\mu}_{j,\ell}^{\text{mc}}$ for all $j \in \mathcal{J}$.

Let $VP_{j\ell}(a_j = 1, \mathbf{a}_{-j}, x_t)$ denote the equilibrium variable profit of firm j evaluated at $\ell \in \{1, \dots, L^*\}$ when firm j enters and its rivals' entry configuration is $\mathbf{a}_{-j} \in \{0, 1\}^{J-1}$. Using the type-specific propensity scores and conditional independence of rival entry decisions within each latent type, expected variable profits at the time of entry decisions can be written as:

$$VP_{j\ell}^P(x_t) = \sum_{\mathbf{a}_{-j} \in \{0, 1\}^{J-1}} VP_{j\ell}(a_j = 1, \mathbf{a}_{-j}, x_t) \prod_{i \neq j} \tilde{P}_{i,\ell}(x_t)^{a_i} \left[1 - \tilde{P}_{i,\ell}(x_t) \right]^{1-a_i}. \quad (27)$$

So far, the private information entering fixed costs has been allowed to be the general vector η_{jt} . To recover fixed costs from the type-specific propensity scores, we now assume that this private information can be summarized by a scalar η_{jt} entering additively in fixed costs, with mean zero and known strictly increasing CDF F_η . Under the exclusion restriction that, conditional on ℓ , firm j 's fixed

cost depends on $\mathbf{x}_t$ only through $\mathbf{x}_{jt}^{\text{fc}}$, suppose fixed costs take the form $fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}}) + \sigma_{\eta_j}\eta_{jt}$, where $\sigma_{\eta_j} > 0$. Then, the type-specific propensity scores satisfy $\tilde{P}_{j,\ell}(\mathbf{x}_t) = F_{\eta} \left(\left[VP_{j\ell}^P(\mathbf{x}_t) - fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}}) \right] / \sigma_{\eta_j} \right)$. For $0 < \tilde{P}_{j,\ell}(\mathbf{x}_t) < 1$, this relationship can be inverted to obtain:

$$F_{\eta}^{-1} \left(\tilde{P}_{j,\ell}(\mathbf{x}_t) \right) = \frac{1}{\sigma_{\eta_j}} VP_{j\ell}^P(\mathbf{x}_t) - \frac{1}{\sigma_{\eta_j}} fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}}). \quad (28)$$

Provided that functions of the components of $\mathbf{x}_t$ excluded from $\mathbf{x}_{jt}^{\text{fc}}$ shift $VP_{j\ell}^P(\mathbf{x}_t)$ sufficiently while not entering $fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})$, this equation identifies the scale parameter σ_{η_j} and the fixed-cost function $fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})$.

4.4 Estimation

We implement the finite-mixture approach with a two-step estimator for each candidate number of latent market types L . The first step estimates the type probabilities and type-specific propensity scores by semiparametric sieve Maximum Likelihood Estimation (MLE); the second constructs the selection-function regressors and estimates the demand and selection-function parameters by Generalized Method of Moments (GMM). We then discuss selection among candidate values of L . The continuous- κ_t approach has the same two-step structure: its derivative regressors are constructed from a first-step estimate of the ordinary propensity score and enter the same second-step regression, as detailed in Online Appendix E.

4.4.1 First step: Type probabilities and type-specific propensity scores

For a candidate number of latent market types L , we approximate each type-specific propensity score using a logistic sieve (Hirano et al., 2003; Chen, 2007). Let $\mathbf{b}_t \equiv (b_1(\mathbf{x}_t), \dots, b_{N_X}(\mathbf{x}_t))'$ collect N_X basis functions. For $j = 1, \dots, J$ and $\ell = 1, \dots, L$, we specify

$$\tilde{P}_{j,\ell}(\mathbf{x}_t) = \Lambda \left(\mathbf{b}_t' \boldsymbol{\gamma}_{j,\ell} \right), \quad (29)$$

where Λ is the logistic CDF and $\boldsymbol{\gamma}_{j,\ell} \in \mathbb{R}^{N_X}$.

Let $\boldsymbol{\gamma}$ collect the JLN_X sieve coefficients and $\tilde{\mathbf{f}} \equiv (\tilde{f}_1, \dots, \tilde{f}_L)'$ the type probabilities, with $\tilde{f}_{\ell} > 0$ and $\sum_{\ell=1}^L \tilde{f}_{\ell} = 1$. The finite-mixture distribution in equation (22)

implies the sample log-likelihood

$$\ln \mathcal{L}_{1st}(\tilde{f}, \gamma) = \sum_{t=1}^T \ln \left[\sum_{\ell=1}^L \tilde{f}_{\ell} \prod_{j=1}^J \tilde{P}_{j,\ell}(\mathbf{x}_t)^{a_{jt}} \left\{ 1 - \tilde{P}_{j,\ell}(\mathbf{x}_t) \right\}^{1-a_{jt}} \right]. \quad (30)$$

We maximize (30) using the Expectation-Maximization (EM) algorithm (Pilla and Lindsay, 2001).⁵ For fixed L and N_X , the first-step model is finite-dimensional, and under correct specification, interiority, local identification, and nonsingularity, its MLE has the usual $\sqrt{T}$ asymptotics. With continuous covariates and $N_X \rightarrow \infty$, the sieve estimates generally converge more slowly and require the usual approximation, rate, and smoothness conditions.

4.4.2 Second step: Demand parameters and selection-bias functions

For a candidate L , interpret equation (25) and its associated selection-function vectors with L in place of L^* . Given the first-step estimates $(\hat{\tilde{f}}, \hat{\gamma})$, we construct the selection-function regressors

$$\hat{r}_{j\ell t} = \frac{\hat{\tilde{P}}_{j,\ell}(\mathbf{x}_t) - \hat{\tilde{P}}_{j,L}(\mathbf{x}_t) \hat{\tilde{f}}_{\ell}}{\hat{\tilde{P}}_j(\mathbf{x}_t)} \hat{\tilde{f}}_{\ell}, \quad \ell = 1, \dots, L-1, \quad (31)$$

where $\hat{\tilde{P}}_j(\mathbf{x}_t) \equiv \sum_{\ell=1}^L \hat{\tilde{P}}_{j,\ell}(\mathbf{x}_t) \hat{\tilde{f}}_{\ell}$ is the estimated ordinary propensity score. We then apply GMM to equation (25) to jointly estimate the demand parameters θ and the product-specific selection-function coefficients $\{\tilde{\mu}_j\}_{j=1}^J$. With L and N_X fixed as above, standard smoothness, moment, and rank conditions yield a $\sqrt{T}$ -asymptotically linear joint two-step estimator for fixed J , with markets as the independent sampling units (Newey, 2009). Inference must account for first-step estimation error and its covariance with the second-step moments. The linearized common-weight bootstrap in Online Appendix F does so without re-estimating the first-step EM algorithm in every replication.

⁵Recent applications of EM to nonparametric discrete-choice mixtures include Bunting (2022); Bunting et al. (2022); Hu and Xin (2022); Williams (2020).

4.4.3 Selecting the number of latent types

Although Assumption 2 identifies L^* in the population, implementation in finite samples requires choosing L from a set of candidate values. A first-step likelihood-based Bayesian information criterion (BIC) and the precision of the entry estimates indicate which candidates are adequately supported by the joint distribution of entry decisions. Because L determines the selection functions entering the selected-sample demand equation and our main objective is to estimate the structural demand parameters, we choose among the supported values of L using a second-step residual-based BIC-style criterion and Hausman-type comparisons of the demand parameters.

First-step BIC. For given L , the likelihood-based BIC for the entry model is

$$\text{BIC}_{1\text{st}}(L) = -2 \ln \mathcal{L}_{1\text{st}}(\hat{\mathbf{f}}, \hat{\boldsymbol{\gamma}}) + d_{1\text{st}}(L) \ln(T), \quad (32)$$

where $d_{1\text{st}}(L)$ is the number of estimated first-step parameters. If all sieve coefficients are type-specific, $d_{1\text{st}}(L) = JLN_X + L - 1$; coefficients restricted to be common across latent types are counted only once. Smaller values of $\text{BIC}_{1\text{st}}(L)$ favor a candidate L . We also examine the estimated type probabilities and the precision of the first-step estimates: a near-empty type or a nearly singular information matrix indicates that L is too large for the available entry variation.

Second-step BIC-style criterion. Among the values of L receiving adequate first-step support, our main information criterion is based on the fit of the selected-sample demand equation. Let N_{sel} denote the number of demand observations used in the second step. We define

$$\text{BIC}_{2\text{nd}}(L) = N_{\text{sel}} \ln \hat{\sigma}_{\xi}^2(L) + (L - 1)J \ln(N_{\text{sel}}), \quad (33)$$

where $\hat{\sigma}_{\xi}^2(L) \equiv N_{\text{sel}}^{-1} \sum_{(j,t): a_{jt}=1} \hat{\xi}_{jt}(L)^2$ is the selected-sample residual mean square from equation (25). The penalty counts the $J(L - 1)$ selection-function coefficients (the number of demand parameters is constant across L). Smaller values of $\text{BIC}_{2\text{nd}}(L)$ favor a candidate L .

Hausman-type tests on demand parameters. Information criteria do not answer the essential question of whether a larger L yields demand estimates that are different

from those at smaller L or from estimates that ignore selection. We therefore complement them with Hausman-type tests on the demand parameters.

Because the selection-function coefficients vary in dimension with L , comparisons are based on the demand-parameter vector θ , which is common to all specifications. Consider two second-step specifications A and B that differ in how selection is handled. Under the null that both recover the same θ , we use the Hausman-type Wald statistic

$$\mathcal{H}(A, B) = (\hat{\theta}^B - \hat{\theta}^A)' \hat{V}_{AB}^{-1} (\hat{\theta}^B - \hat{\theta}^A), \quad (34)$$

where $\hat{V}_{AB}$ estimates the variance-covariance matrix of $\hat{\theta}^B - \hat{\theta}^A$ and is computed using the linearized common-weight bootstrap of Online Appendix F. Under joint asymptotic normality and nonsingularity of this matrix, $\mathcal{H}(A, B) \xrightarrow{d} \chi^2(\dim(\theta))$. The same construction applies to any q -dimensional common subvector of θ , with limiting distribution $\chi^2(q)$.

5 Empirical application

5.1 Data and descriptive statistics

We apply our method to estimate demand in the US airline industry. The challenge of endogenous product entry in demand estimation in this industry has recently been explored by [Ciliberto et al. \(2021\)](#) and [Li et al. \(2022\)](#).

Data sources. Our sample consists of the publicly available DB1B and T100 datasets from the US Department of Transportation, complemented with population data from the Census Bureau’s American Community Survey (ACS). Specifically, we use quarterly data spanning 2012-Q1 to 2013-Q4 for routes between the airports at the 100 largest Metropolitan Statistical Areas (MSA) in the United States. These account for 142 airports, as several of these MSAs are served by more than one airport.

Airlines. The airlines included in our analysis are American (AA), Delta (DL), United (UA), US Airways (US), Southwest (WN), a combined group of Low-Cost Carriers (LCC), and a combined group of the remaining carriers (Others).⁶ Given

⁶Following [Ciliberto et al. \(2021\)](#), the list of airlines included in the group LCC is: Alaska, JetBlue, Frontier, Allegiant, Spirit, Sun Country, and Virgin. The group Others collects other smaller carriers.

the large number of carriers in Others, we do not consider this combined group as a player in the entry game. In the notation of our model, j always indexes an airline.

Markets in the demand model. In the demand model, a market t is defined as a *directional airport pair* in a given quarter. For example, New York La Guardia (LGA) $\rightarrow$ Chicago O'Hare (ORD) in 2012-Q1 and ORD $\rightarrow$ LGA in 2012-Q1 are two distinct demand markets. In each demand market t , consumers choose among the airlines offering non-stop flights on that directional route (up to seven airlines, plus the outside option). Thus, s_{jt} denotes the market share of airline j in demand market t .

Markets in the entry model. In the entry model, with some abuse of notation, a market t is defined as a *non-directional* airport pair in a given quarter, where, for example, ORD to LGA in 2012-Q1 is the same market as LGA to ORD in 2012-Q1. Each non-directional entry market thus corresponds to two directional demand markets. There are potentially 10,011 non-directional airport pairs among the 142 airports, i.e., $142 \times 141 / 2$. However, many of these airport pairs have not had an incumbent airline with non-stop flights for several decades. These are typically airport pairs that are geographically too close or in smaller metropolitan areas (MSAs). In our sample, we consider a non-directional airport pair if it was served with non-stop flights either in at least 50 quarters between 1994 and 2018, or in at least one quarter of our 2012–2013 sample period. This results in 2,652 non-directional airport pairs and 20,859 entry markets.⁷

Potential entrants. An airline is a *potential entrant* in a market if it operates non-stop flights out of both airports in that quarter, while it is an *entrant* if it operates non-stop flights between the two airports in that quarter. Of the 20,859 markets, we exclude 1,315 for which none of the modeled airlines is a potential entrant, leaving 19,544 for the estimation of the entry model.

Market size and distance. Following the empirical literature on the airline industry, we define market size as the geometric mean of the populations in the MSAs of the two airports and market distance as the geodesic distance between the two airports.

Observable variables x_t for entry model. The vector of exogenous observable variables at the airline-market level includes: market size, market distance, the

⁷Given 2,652 non-directional airport pairs and eight quarters, there are at most $2,652 \times 8 = 21,216$ potential entry markets. We discard from the analysis the 357 in which the airport pair had no non-stop flights in that quarter and was not served in at least 50 quarters between 1994 and 2018, leaving 20,859 markets.

combined hub size at the origin and destination airports of the airline and of each of its six competitors, airline-by-quarter indicators, and airport indicators. We define the hub size of an airline at an airport as the number of non-stop routes that the airline operates from that airport, and its combined hub size for a given airport pair as the sum of its hub sizes at the two endpoints.

Table 2 presents the distribution of the number of entrants and averages of the market size and distance. Notably, in more than 35% of these markets, there are no airlines providing non-stop flights. Among the markets served by non-stop flights, more than 90% are monopolies or duopolies. Furthermore, there is a strong positive correlation between the number of incumbents, market size, and distance.

Table 2: Distribution of Markets by Number of Entrants

Number of airlines	Frequency # markets (%)	Avg. market size in millions of people	Avg. market distance in miles
0 airlines	6,878 (35.19%)	2.66	638
1 airline	8,770 (44.87%)	3.24	914
2 airlines	2,708 (13.86%)	4.24	952
3 airlines	871 (4.46%)	5.25	1109
4 airlines	235 (1.20%)	5.24	1153
5 airlines	72 (0.37%)	8.16	1255
≥ 6 airlines	10 (0.05%)	6.84	320
Total	19,544 (100.00%)	3.31	835

Table 3 reports each airline’s entry frequency and the average size and distance of the markets it enters. Entry frequencies range from 24.9% for WN to 9.6% for AA; the correlations of entry with market size and distance also differ substantially across airlines. WN enters markets that do not differ significantly in size from those it does not enter (3.34 versus 3.30 million people), whereas AA enters much larger markets (5.12 versus 3.12 million people; averages for non-entered markets are not reported in the table). Entry also differs by distance: DL and US typically enter shorter-haul markets, averaging 874 and 882 miles, whereas markets served by LCC average 1,140 miles.

5.2 First step: Estimation of the entry model

For the entry decisions, we consider the semiparametric finite-mixture model in equation (30). We present results for b_t specified as a second-degree polynomial in

Table 3: Entry Frequency by Airline

Airline	Frequency # markets (%)	Avg. market size in millions of people	Avg. market distance in miles
WN	4,863 (24.9%)	3.34	987
DL	3,332 (17.0%)	3.93	874
UA	3,292 (16.8%)	4.36	962
LCC	2,744 (14.0%)	4.36	1140
US	2,059 (10.5%)	3.94	882
AA	1,869 (9.6%)	5.12	964

x_t .⁸ Table 4 presents the estimates of a probit model (with $L = 1$) and of three nested specifications of the mixture logit model (with $L = 2, 3, 4$).⁹ For each specification, Table 4 reports the estimated probabilities of the latent market types and, for each type, the average elasticity of an airline’s entry probability with respect to: market size, market distance, the airline’s own hub size, and its competitors’ hub size, where the latter is an average of the elasticities with respect to each of the six competitors’ combined hub sizes.

Most estimates in Table 4 have the expected signs: entry probabilities rise with market size and own hub size and fall with distance and competitors’ hub size. The estimated type probabilities are 0.55 and 0.45 for $L = 2$ and 0.38, 0.34, and 0.28 for $L = 3$; all are sizable and precisely estimated. Elasticities vary markedly across types—for $L = 3$, the distance elasticity ranges from -11.95 to -25.42 —indicating substantial heterogeneity in entry behavior across markets. Most notably, the competitors’ hub-size elasticity is small and statistically insignificant without latent types (-0.20 ; s.e. 0.15) but sizable and significant with $L = 2$ (-0.87 and -1.36). Unobserved market characteristics that raise all airlines’ profitability induce a positive correlation between competitors’ presence and entry, confounding competition’s deterrent effect if left uncontrolled.

Our choice of L uses the first- and second-step BICs (32) and (33), the precision of the entry-model estimates, and the robustness of the demand estimates. In

⁸The nine continuous regressors in x_t —market size, market distance, and the combined origin-and-destination hub sizes of the airline and of its six competitors—are each rescaled to $[-1, 1]$, and b_t is the full Chebyshev tensor product of total degree two: the nine linear terms, their nine squares, and all 36 pairwise interactions, i.e., 54 polynomial terms per airline.

⁹As will become clear below, for $L = 1$, we estimate a probit rather than a logit to construct selection-correction terms in the spirit of Heckman and for the case of continuous κ_t described in Online Appendix E.

Table 4: Estimation of Market Entry Model

	Probit $L = 1$	Mixture Logit $L = 2$	Mixture Logit $L = 3$	Mixture Logit $L = 4$
$(\tilde{f}_1, \dots, \tilde{f}_L)$		0.55, 0.45 (0.01, 0.01)	0.38, 0.34, 0.28 (0.01, 0.01, 0.01)	0.29, 0.27, 0.23, 0.21
Elasticities:				
Market size	1.63 (0.35)	1.95, 0.44 (0.38, 0.68)	10.91, 4.71, -5.73 (8.98, 1.36, 3.67)	1.56, 19.99, -7.67, 8.85
Distance	-2.10 (0.24)	-1.16, -17.25 (0.23, 1.50)	-11.95, -25.42, -15.30 (12.76, 2.03, 4.21)	-42.91, -38.83, -13.28, -1.61
Hub size	4.63 (0.18)	4.39, 10.39 (0.20, 0.80)	12.93, 15.29, 20.21 (7.17, 1.10, 3.63)	21.05, 50.71, 75.22, 52.31
Competitor hub size	-0.20 (0.15)	-0.87, -1.36 (0.15, 0.31)	-3.35, -3.00, -5.69 (2.74, 0.46, 1.38)	-2.14, -5.07, -4.55, -3.85
Airline $\times$ Quarter FE	Y	Y	Y	Y
Airport FE	Y	Y	Y	Y
Observations	19,544	19,544	19,544	19,544
Parameters	482	807	1,132	1,457
Log-likelihood	-13493.6	-11047.8	-10002.2	-9268.8
BIC _{1st}	31749.6	30069.1	31189.1	32933.4

Notes: The table reports estimates of a probit entry model without latent market types ($L = 1$) and of finite-mixture logit entry models with $L = 2$, $L = 3$, and $L = 4$ latent market types. In all specifications, the market observables x_t include market size, market distance, the combined hub size at the origin and destination airports of the airline and of each of its six competitors, airline $\times$ quarter indicators, and airport indicators. For each latent type ℓ , the table reports the estimated type probability $\tilde{f}_\ell$ and the average elasticity of an airline's entry probability with respect to each continuous regressor in x_t , where the elasticity denoted by "Competitor hub size" represent the average across the airline's six competitors' combined hub sizes (six distinct continuous regressors in x_t). Standard errors are in parentheses. All specifications are estimated by maximum likelihood using the EM algorithm. BIC_{1st} is the first-step criterion defined in equation (32).

Table 4, moving from $L = 1$ to $L = 2$ substantially raises the log-likelihood and lowers BIC_{1st} from 31,749.6 to 30,069.1,¹⁰ showing that latent market types capture strong correlation among airline entry decisions unexplained by x_t . Although $L = 3$ further raises the likelihood, its additional parameters increase BIC_{1st} to 31,189.1; hence $L = 2$ ranks first and $L = 3$ second, still below $L = 1$. At $L = 4$, BIC_{1st} rises to 32,933.4, the largest value, and the estimates become extremely imprecise; we therefore omit their standard errors.¹¹ Thus, the first step favors $L = 2$ or $L = 3$ and

¹⁰This improvement is not due to the probit at $L = 1$: comparing a logit at $L = 1$ with a finite-mixture logit at $L = 2$ gives the same result.

¹¹For $L = 2, 3$, and 4, our EM algorithm uses 32, 48, and 64 starts, respectively, combining random starts with starts constructed from the $L - 1$ solution, and retains the highest-likelihood solution. It then performs twelve basin-hopping rounds that perturb the best solution's posterior-type probabilities, re-run EM, and retain any improvement.

rules out $L = 4$; the next section chooses between $L = 2$ and $L = 3$ using $\text{BIC}_{2\text{nd}}$ and the robustness of the demand estimates.

5.3 Second step: Estimation of demand parameters

For the demand system, we follow [Ciliberto et al. \(2021\)](#) and estimate a nested logit model in which the seven airlines active in the directional market t form a single nest with nesting parameter σ , and the outside option is a separate alternative.

$$\ln \left(\frac{s_{jt}}{s_{0t}} \right) = \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \sigma \ln \left(\frac{s_{jt}}{1 - s_{0t}} \right) + \lambda_j(\mathbf{x}_t) + \tilde{\xi}_{jt}. \quad (35)$$

We compute airline j 's market share s_{jt} in demand market t as the number of passengers traveling that route on airline j 's non-stop flights (multiplied by 10, since the data are a 10% survey of total traffic), divided by market size—the geometric mean of the origin and destination MSA populations. The outside-option share is $s_{0t} = 1 - \sum_j s_{jt}$, where the sum runs over the airlines in the demand model. Price p_{jt} is the average fare on airline j 's non-stop flights in market t , in hundreds of dollars. The vector of exogenous demand characteristics $\mathbf{x}_{jt}^D$ includes market distance (in thousands of miles), its square, airline j 's airport presence, and airline $\times$ quarter and airport indicators. Airport presence is the share of airline j 's network operated from the route endpoints: the sum of its hub sizes at the origin and destination divided by its total hub size across all airports that quarter. For the selection-bias function $\lambda_j(\mathbf{x}_t)$, we consider the specifications below: two Heckman-type benchmarks and the two corrections proposed in Section 4—the continuous- κ_t correction of Online Appendix E and the finite-mixture correction with latent market types.

Heckman approach: single-index and normal unobservables. The unobservables of the entry model satisfy a single-threshold-crossing condition: $a_{jt} = \mathbb{1}\{\eta_{jt} \leq \kappa_t' \boldsymbol{\rho}_j + \mathbf{b}_t' \boldsymbol{\gamma}_j^P\}$, where $\boldsymbol{\eta}_t \equiv (\eta_{1t}, \dots, \eta_{Jt})'$ collects the scalar private shocks; $\boldsymbol{\xi}_t$, κ_t , and $\boldsymbol{\eta}_t$ are jointly normally distributed; $\boldsymbol{\rho}_j$ and $\boldsymbol{\gamma}_j^P$ are parameters; and $\mathbf{b}_t$ contains basis functions of $\mathbf{x}_t$. Let $v_{jt} \equiv \eta_{jt} - \kappa_t' \boldsymbol{\rho}_j$. The standard Heckman correction is

$$\lambda_j(\mathbf{x}_t) = \mathbb{E} \left(\tilde{\xi}_{jt} \mid v_{jt} \leq \mathbf{b}_t' \boldsymbol{\gamma}_j^P \right) = \varrho_j^H \frac{\phi \left(\Phi^{-1}[\bar{P}_j(\mathbf{x}_t)] \right)}{\bar{P}_j(\mathbf{x}_t)},$$

where ϕ and Φ are the standard normal density and CDF, and the airline-specific

composite coefficient $\varrho_j^H \equiv -\text{Cov}(\xi_{jt}, v_{jt}) / \sqrt{\text{Var}(v_{jt})}$. We construct the inverse Mills ratio from $\bar{P}_j(\mathbf{x}_t) \equiv \Phi(\mathbf{b}'_t \boldsymbol{\gamma}_j^P)$, the ordinary propensity score implied by the probit entry model with $L = 1$ from Table 4. With six airlines included in the entry model (all except Others), this correction introduces six additional control variables, one inverse Mills ratio for each modeled airline.

Semiparametric Heckman approach. We maintain the single-threshold-crossing structure of the entry model $a_{jt} = \mathbb{1}\{\eta_{jt} \leq \boldsymbol{\kappa}'_t \boldsymbol{\rho}_j + \mathbf{b}'_t \boldsymbol{\gamma}_j^P\}$ but allow $\boldsymbol{\xi}_t$, $\boldsymbol{\kappa}_t$, and $\boldsymbol{\eta}_t$ to have unrestricted distributions with continuous support. The selection-bias term is then a deterministic function of $\mathbf{b}'_t \boldsymbol{\gamma}_j^P$ or, equivalently, of the propensity score $\bar{P}_j(\mathbf{x}_t) \equiv F_{v,j}(\mathbf{b}'_t \boldsymbol{\gamma}_j^P)$, where $F_{v,j}$ is the (invertible) CDF of v_{jt} . Because the basis functions in $\mathbf{b}_t$ can approximate any smooth function, the probit propensity score with $L = 1$ from Table 4 entails no loss of flexibility. Following Newey (2009), we approximate the selection-bias function with a cubic polynomial in Heckman’s inverse Mills ratio, with airline-specific coefficients, so that the selection correction introduces 18 additional control variables.¹²

Continuous- $\boldsymbol{\kappa}_t$ approach: multivariate normal unobservables. The two approaches above restrict the unobservables driving selection to enter through the scalar index v_{jt} . As an alternative, we consider the approach in Online Appendix E, which allows for multi-dimensional common-knowledge unobservables, $\boldsymbol{\kappa}_t = \boldsymbol{\xi}_t$, under the assumption that these are jointly normally distributed (similar to Ciliberto et al., 2021). As shown in Online Appendix E, the selection-bias function then is:

$$\lambda_j(\mathbf{x}_t) = \sum_k \varrho_{jk} \frac{\partial \log \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}},$$

where x_{k1t} is a continuous regressor that shifts airline k ’s latent entry index, and the composite coefficients ϱ_{jk} encode the covariances between the demand unobservables. We implement this correction using the same probit entry model with $L = 1$ from Table 4 as in the two approaches above. For each carrier k , we use its own combined hub size at the route’s endpoints as the special regressor x_{k1t} . Thus, for every airline j , the correction includes seven derivatives of $\log \bar{P}_j(\mathbf{x}_t)$: one with respect to j ’s own hub size and six with respect to the competitors’ hub sizes. With six airlines included in the entry model (all except Others), this yields $6 \times 7 = 42$

¹²We obtain very similar estimates when using higher-order polynomial approximations.

additional control variables.¹³

Finite-mixture logit with latent types. This is the entry model estimated in Section 5.2. For each airline j and latent type $\ell = 1, \dots, L - 1$, the selection-function regressor is as in equation (31). The $L - 1$ regressors per airline are stacked into $\hat{\mathbf{r}}_{jt} = (\hat{r}_{j1t}, \hat{r}_{j2t}, \dots, \hat{r}_{j,L-1,t})'$, so with six airlines in the entry model the selection correction adds $6(L - 1)$ controls: 6 for $L = 2$, 12 for $L = 3$, and 18 for $L = 4$.

For all the two-stage least squares (2SLS) estimators of (35), we use as instruments, for each of an airline's competitors, airport presence if it is a potential entrant in the market (and zero otherwise), together with the number of competitors that are potential entrants. As detailed in Online Appendix F, standard errors are computed using a linearized common-weight bootstrap with 200 replications.

Table 5 reports demand-parameter estimates and Table 6 their implied median price elasticities and Lerner indexes. Correcting for the endogeneity of prices and within-nest market shares by moving from OLS to 2SLS without selection correction more than doubles the magnitude of the median own-price elasticity (-1.03 to -2.30) and lowers the median Lerner index from 97.1% to 43.4%. Accounting additionally for endogenous product entry yields median own-price elasticities ranging from -4.75 to -2.37 and median Lerner indexes from 21.1% to 42.3% across the six selection-corrected estimators. Median cross-price elasticities range from 0.81 to 1.68, versus 0.91 under 2SLS, with the largest increases under the continuous- κ_t and finite-mixture corrections. Despite variation across corrections, accounting for endogenous product entry generally yields more elastic demand estimates, with implications for any subsequent analyses using them as inputs.

The last four rows of Table 5 report Hausman-type specification tests comparing each selection-corrected estimator with the baseline 2SLS, based on the statistic $\mathcal{H}(2SLS, \cdot)$ defined in equation (34) of Section 4.4.3 and computed with the linearized common-weight bootstrap of Online Appendix F using 200 replications. We report the joint test on the demand parameters $(\alpha, \sigma, \beta_{\text{dist}}, \beta_{\text{dist}^2}, \beta_{\text{pres}})$ ($\chi^2(5)$) and the restricted test on the price and nesting parameters (α, σ) ($\chi^2(2)$), which govern the price elasticities in Table 6. The $\chi^2(5)$ test rejects equality with the baseline 2SLS

¹³In the notation of Online Appendix E, we interpret equation (41) so that, among the airline-specific hub measures, each latent index ζ_{kt} includes only airline k 's own combined hub size x_{k1t} ; rivals' hub sizes do not enter ζ_{kt} . Nevertheless, because $P_j(\mathbf{x}_t, \boldsymbol{\zeta}_t) = G_j(\zeta_{1t}, \dots, \zeta_{Jt})$, the ordinary propensity score $\bar{P}_j(\mathbf{x}_t)$ generally depends on every airline's x_{k1t} . Accordingly, the reduced form index of each airline in the entry model includes the combined hub sizes of all airlines.

at the 5% level or lower for every selection-corrected estimator except $L = 4$; the restricted $\chi^2(2)$ test rejects for the three $L = 1$ corrections and for $L = 3$, but not for $L = 2$. The non-rejections in the $L = 4$ column reflect the wide bootstrap standard errors induced by the imprecise estimation of the entry model.

Table 5: Estimation of Demand Parameters

	<i>Not control. for sel.</i>		<i>Controlling for endogenous selection</i>					
	OLS	2SLS	2SLS Heckman $L = 1$	2SLS Semipar. $L = 1$	2SLS Cont. κ $L = 1$	2SLS Fin.-Mix. $L = 2$	2SLS Fin.-Mix. $L = 3$	2SLS Fin.-Mix. $L = 4$
Price (100\$) (α)	-0.342 (0.015)	-0.837 (0.157)	-1.293 (0.222)	-1.384 (0.224)	-1.786 (0.311)	-0.851 (0.157)	-0.954 (0.158)	-0.993 (0.201)
Within share (σ)	0.675 (0.009)	0.592 (0.034)	0.473 (0.041)	0.437 (0.043)	0.556 (0.061)	0.602 (0.034)	0.602 (0.035)	0.621 (0.047)
Distance (1000mi)	-0.016 (0.055)	0.513 (0.169)	0.844 (0.228)	0.902 (0.231)	1.416 (0.318)	0.421 (0.168)	0.603 (0.166)	0.686 (0.210)
Distance ²	-0.112 (0.022)	-0.180 (0.031)	-0.168 (0.037)	-0.159 (0.038)	-0.261 (0.047)	-0.139 (0.032)	-0.185 (0.030)	-0.209 (0.037)
Airport presence	4.593 (0.221)	7.438 (0.550)	7.566 (0.655)	7.646 (0.666)	8.820 (0.723)	7.313 (0.544)	7.801 (0.598)	7.573 (0.892)
Airline $\times$ Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Airport FE	Y	Y	Y	Y	Y	Y	Y	Y
# control var. entry	0	0	6	18	42	6	12	18
Observations	37,184	37,184	37,184	37,184	37,184	37,184	37,184	37,184
BIC _{2nd}		-21916.8				-22057.7	-22433.4	-22677.7
Hausman $\chi^2(5)$ vs 2SLS			189.80	165.49	83.12	55.71	12.23	3.09
p-value			[0.000]	[0.000]	[0.000]	[0.000]	[0.032]	[0.686]
Hausman $\chi^2(2)$ vs 2SLS			101.06	97.29	64.81	3.98	6.74	1.94
p-value			[0.000]	[0.000]	[0.000]	[0.137]	[0.034]	[0.380]

Notes: The table reports estimates of the nested-logit demand parameters across the baseline OLS and 2SLS specifications and across the alternative selection-correction specifications described in the text. Standard errors are computed using the linearized common-weight bootstrap of Online Appendix F with 200 replications. BIC_{2nd} is the second-step BIC-style criterion defined in equation (33). The last four rows report the Hausman-type specification test $\mathcal{H}(2SLS, \cdot)$ from equation (34), comparing each selection-corrected estimator against the baseline 2SLS: the $\chi^2(5)$ statistic is joint over $(\alpha, \sigma, \beta_{\text{dist}}, \beta_{\text{dist}^2}, \beta_{\text{pres}})$, while the $\chi^2(2)$ statistic is joint over (α, σ) .

Following the combined selection strategy of Section 4.4.3, we select $L = 3$ using BIC_{1st} in Table 4, BIC_{2nd} in Table 5, and the pairwise Hausman-type tests in Online Appendix Table A.3. Although BIC_{2nd} alone favors $L = 4$, that specification ranks last by BIC_{1st}, has an imprecisely estimated entry model, and yields demand estimates statistically indistinguishable from those for $L = 3$. By contrast, the $L = 3$ estimates differ significantly from those for $L = 2$ ($\chi^2(5) = 27.32$, $p < 0.01$), and moving from $L = 2$ to $L = 3$ improves BIC_{2nd}.

Table 6: Median Price Elasticities and Lerner Indexes

	Not control. for sel.		Controlling for endogenous selection					
	OLS	2SLS	2SLS Heckman $L = 1$	2SLS Semipar. $L = 1$	2SLS Cont. κ $L = 1$	2SLS Fin.-Mix. $L = 2$	2SLS Fin.-Mix. $L = 3$	2SLS Fin.-Mix. $L = 4$
<i>Own-Price Elasticity</i>	-1.030	-2.304	-3.220	-3.369	-4.750	-2.366	-2.651	-2.812
AA	-1.269	-2.653	-3.620	-3.780	-5.430	-2.738	-3.068	-3.291
DL	-1.073	-2.486	-3.598	-3.791	-5.189	-2.543	-2.849	-3.002
UA	-1.273	-2.817	-3.973	-4.139	-5.841	-2.896	-3.246	-3.438
US	-1.078	-2.352	-3.306	-3.475	-4.817	-2.416	-2.706	-2.884
WN	-0.771	-1.842	-2.755	-2.913	-3.889	-1.879	-2.105	-2.205
LCC	-0.953	-2.017	-2.671	-2.763	-4.079	-2.083	-2.334	-2.504
Others	-1.174	-2.357	-2.943	-2.994	-4.673	-2.445	-2.740	-2.980
<i>Cross-Price Elasticity</i>	0.532	0.910	0.874	0.810	1.680	0.963	1.079	1.215
AA	0.629	1.074	1.029	0.952	1.983	1.138	1.274	1.436
DL	0.774	1.323	1.267	1.173	2.441	1.401	1.569	1.768
UA	0.555	0.949	0.908	0.842	1.751	1.005	1.125	1.267
US	0.593	1.014	0.971	0.899	1.871	1.074	1.202	1.355
WN	0.514	0.880	0.845	0.783	1.625	0.932	1.043	1.175
LCC	0.428	0.732	0.702	0.650	1.352	0.775	0.868	0.978
Others	0.292	0.499	0.478	0.443	0.921	0.529	0.592	0.667
<i>Lerner Index</i>	97.1%	43.4%	31.1%	29.7%	21.1%	42.3%	37.7%	35.6%
AA	78.8%	37.7%	27.6%	26.5%	18.4%	36.5%	32.6%	30.4%
DL	93.2%	40.2%	27.8%	26.4%	19.3%	39.3%	35.1%	33.3%
UA	78.6%	35.5%	25.2%	24.2%	17.1%	34.5%	30.8%	29.1%
US	92.7%	42.5%	30.3%	28.8%	20.8%	41.4%	37.0%	34.7%
WN	129.7%	54.3%	36.3%	34.3%	25.7%	53.2%	47.5%	45.4%
LCC	105.0%	49.6%	37.4%	36.2%	24.5%	48.0%	42.8%	39.9%
Others	85.2%	42.4%	34.0%	33.4%	21.4%	40.9%	36.5%	33.6%
Airline $\times$ Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Airport FE	Y	Y	Y	Y	Y	Y	Y	Y
# control var. entry	0	0	6	18	42	6	12	18
Observations	37,184	37,184	37,184	37,184	37,184	37,184	37,184	37,184

Notes: The table reports median own-price elasticities, cross-price elasticities, and Lerner indexes implied by the demand estimates in Table 5. For each specification, these quantities are first computed at the airline-market level using the corresponding column's demand parameter estimates; the table then reports their medians, overall and separately by airline. Cross-price elasticities are reported by price-setting airline.

Online Appendix Figures A.1 and A.2 report the empirical distributions of estimated own- and cross-price elasticities under 2SLS, the continuous- κ_t correction, and the finite-mixture correction with $L = 3$. Relative to 2SLS, both selection corrections shift the own-price distributions leftward and increase their dispersion for every airline, most markedly under the continuous- κ_t correction; they also imply stronger substitution across airlines. Thus, ignoring endogenous entry understates the magnitude of both own- and cross-price elasticities.

5.4 Estimation of costs and profits

Although our main focus is demand estimation with endogenous product entry, as discussed in Section 4.3.1, our framework also allows us to identify the remaining ingredients required for a wide range of counterfactual experiments: marginal costs, equilibrium variable profits, and fixed costs. In this section, we illustrate these ideas. We first estimate marginal costs and then use these estimates, together with the first-step type-specific propensity scores and the second-step demand parameters, to construct equilibrium variable profits and to recover fixed costs. Finally, we introduce an ANOVA–Sobol variance decomposition of the net entry-profit index that quantifies the contribution of unobserved market-type heterogeneity to endogenous selection, distinguishing the expected-variable-profit channel from the fixed-cost channel.

5.4.1 Marginal costs

Following Ciliberto et al. (2021), we assume that the airlines engage in Bertrand–Nash competition and that each airline has a marginal cost function that does not depend on output. Then, given demand equation (35), the marginal cost of airline j in demand market t can be estimated from the following pricing equation:

$$p_{jt} + \frac{1 - \sigma}{\alpha(1 - \sigma s_{jt|g} - (1 - \sigma)s_{jt})} = mc_{jt}, \quad (36)$$

where g denotes the nest that contains all the airlines, $s_{jt|g} \equiv s_{jt}/(1 - s_{0t})$, and mc_{jt} denotes the marginal cost. We then estimate the following marginal cost function:

$$\ln(mc_{jt}) = (\mathbf{x}_{jt}^{\text{mc}})' \boldsymbol{\varphi} + \lambda_j^{\text{mc}}(\mathbf{x}_t) + \tilde{\omega}_{jt}, \quad (37)$$

where $\mathbf{x}_{jt}^{\text{mc}}$ includes market distance and market distance squared, together with airline $\times$ quarter and airport indicators, $\lambda_j^{\text{mc}}(\mathbf{x}_t)$ is a selection-bias function specified as for the demand equation (35), and $\tilde{\omega}_{jt} \equiv \omega_{jt} - \lambda_j^{\text{mc}}(\mathbf{x}_t)$ is the residual variable-cost unobservable, with conditional mean zero in the selected sample.

Online Appendix Table A.4 reports the average marginal costs obtained from equation (36) using the demand estimates of Table 5 (see Online Appendix Figure A.3 for the corresponding empirical distributions), while Online Appendix

Table A.5 presents the estimates of φ from equation (37). The $\ln(mc_{jt})$ regressand in each column is constructed from the demand estimates of the corresponding column of Table 5, while equation (37) is always estimated by OLS: once selection is controlled for, its regressors are exogenous. Two facts stand out. First, price and within-nest-share endogeneity matters sizably: the demand estimates that do not correct for it (OLS column) imply average marginal costs close to zero, with roughly half of the observations dropped because the implied marginal cost is not positive. Second, after accounting for demand-side selection when constructing the dependent variable $\ln(mc_{jt})$, endogenous selection in the marginal-cost regression still matters, although less than in demand: only the $L = 1$ corrections yield estimates significantly different from the baseline 2SLS.

5.4.2 Contribution of unobservables to endogenous selection

The preceding estimates show that unobserved market heterogeneity generates selection bias in demand estimation and, to a lesser extent, in marginal-cost estimation. We now assess its role in airlines' entry decisions by quantifying the fraction of cross-market variation in the net entry-profit index driven by the latent market state κ_t and, within this latent-type channel, the parts operating through expected variable profits and fixed entry costs. We do this for our preferred specification with $L = 3$ using an airline-by-airline ANOVA–Sobol variance decomposition. All derivations and implementation details are in Online Appendix G.

Following Section 4.3.1, let $VP_{j\ell t} \equiv VP_{j\ell}^P(\mathbf{x}_t)$ denote expected variable profit at entry, which averages equilibrium variable profits over rivals' possible entry configurations using their type-specific propensity scores. Let $fc_{j\ell t} \equiv fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})$ denote the component of fixed entry cost excluding the private shock $\sigma_{\eta_j}\eta_{jt}$, and let $\Pi_{j\ell t} \equiv VP_{j\ell t} - fc_{j\ell t}$ denote the net entry-profit index before that shock. As detailed in Online Appendix G, we estimate σ_{η_j} by 2SLS from the inverted entry equation and recover $fc_{j\ell t}$. The variance of each $Y \in \{\Pi, VP, fc\}$ across markets and latent types decomposes as $V_j^Y = V_{j,x}^Y + V_{j,\ell}^Y + V_{j,\text{res}}^Y$, where the three components capture variation attributable to observables alone, latent types alone, and their interaction, respectively. We define $V_{j,\text{unobs}}^Y \equiv V_{j,\ell}^Y + V_{j,\text{res}}^Y$ as the variation associated with latent types, either directly or through their interaction with observables. Since

$$\Pi = VP - fc,$$

$$\frac{V_{j,\text{unobs}}^{\Pi}}{V_j^{\Pi}} = \frac{V_{j,\text{unobs}}^{VP}}{V_j^{\Pi}} + \frac{V_{j,\text{unobs}}^{fc}}{V_j^{\Pi}} - \frac{2 \text{Cov}_{j,\text{unobs}}^{VP,fc}}{V_j^{\Pi}}.$$

Table 7 reports, for each airline, $\hat{\sigma}_{\eta_j}$, the corresponding first-stage F -statistic, the unobservable share of net-entry-profit variance and its pure-type and interaction components, and the three terms in this identity, each normalized by V_j^{Π} .

Table 7: Contribution of Unobservables to Endogenous Selection

Airline	$\hat{\sigma}_{\eta}$	$F\text{-stat}$	$\frac{V_{\text{unobs}}^{\Pi}}{V^{\Pi}}$	$\frac{V_{\ell}^{\Pi}}{V^{\Pi}}$	$\frac{V_{\text{res}}^{\Pi}}{V^{\Pi}}$	$\frac{V_{\text{unobs}}^{VP}}{V^{\Pi}}$	$\frac{V_{\text{unobs}}^{fc}}{V^{\Pi}}$	$\frac{-2 \text{Cov}_{\text{unobs}}^{VP,fc}}{V^{\Pi}}$
AA	125.65	492.75	0.60	0.27	0.34	0.27	0.89	-0.55
DL	47.45	550.33	0.53	0.06	0.47	0.37	1.01	-0.85
UA	88.27	1235.89	0.36	0.05	0.31	0.43	0.86	-0.92
US	152.97	387.13	0.62	0.16	0.46	0.24	0.95	-0.57
WN	348.12	1275.94	0.40	0.02	0.38	1.04	1.62	-2.26
LCC	364.51	218.36	0.42	0.05	0.37	0.32	0.80	-0.70

Notes: The table reports airline-specific measures from the ANOVA–Sobol decomposition of the net entry-profit index, computed, for each airline, over its potential-entry markets. The sample is restricted to route-quarters for which both directional demand markets are observed, since expected variable profits are constructed by aggregating the two directions of each route. All definitions and estimation details are in Online Appendix G. $\hat{\sigma}_{\eta_j}$ is the airline-specific scale parameter estimated from the 2SLS regression in equation (54); F -stat is the corresponding first-stage statistic, clustered by route-quarter.

The estimated private-shock scales differ substantially across airlines, and all six first stages are strong. Latent market types and their interactions with observables account for 36%–62% of the variation in the net entry-profit index. The interaction component exceeds the pure-type component for every airline, with the latter sizable mainly for AA and US. For every airline, the gross fixed-cost contribution exceeds the expected-variable-profit contribution. The unobservable components of expected variable profits and fixed costs also covary positively: latent market states associated with higher expected variable profits tend to entail higher fixed costs, so the two channels partially offset in the net entry-profit index. These results complement our earlier finding that endogenous selection matters more for demand estimation than for marginal-cost estimation: the ANOVA–Sobol decomposition additionally reveals that fixed-cost heterogeneity is the larger gross channel through which latent market states affect airlines’ entry decisions.

6 Conclusions

Firms offer products in markets where they expect to earn profits, using information about market conditions partly unobserved by researchers. When this information correlates with demand unobservables, entry creates endogenous selection bias in demand estimation. Existing approaches either restrict firms' information at entry or jointly estimate a full structural model of demand, pricing, and entry, making demand estimates heavily dependent on supply-side assumptions.

This paper addresses this form of endogenous selection with a two-step identification approach. When firms' market-entry decisions are correlated through latent market states, we show that the selection-bias function in demand can be represented using latent propensity scores: entry probabilities conditional on observables and latent market states. We provide conditions under which these scores are identified from the joint distribution of entry decisions and can be used to control for endogenous selection. These identification results lead to a computationally convenient two-step estimator of demand that accommodates price endogeneity without either ruling out endogenous selection through restrictions on firms' information or jointly estimating demand, pricing, and entry.

We illustrate the method using data from the airline industry. The results highlight the empirical importance of correcting demand estimates for endogenous product entry allowing for correlation across firms' market-entry decisions. Failing to correct for endogenous entry substantially understates the magnitude of own- and cross-price elasticities and overstates market power.

References

- Aguirregabiria, V., Collard-Wexler, A., and Ryan, S. (2021). Dynamic games in empirical industrial organization. In Ho, K., Hortaçsu, A., and Lizzeri, A., editors, *Handbook of Industrial Organization, Volume 4*, pages 225–343. Elsevier.
- Aguirregabiria, V. and Ho, C.-Y. (2012). A dynamic oligopoly game of the us airline industry: Estimation and policy experiments. *Journal of Econometrics*, 168(1):156–173.
- Aguirregabiria, V. and Mira, P. (2019). Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity. *Quantitative Economics*, 10(4):1659–1701.

- Ahn, H. and Powell, J. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58(1-2):3–29.
- Allman, E., Matias, C., and Rhodes, J. (2009). Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics*, 37(6A):3099–3132.
- Angrist, J. D. (1997). Conditional independence in sample selection models. *Economics Letters*, 54(2):103–112.
- Aradillas-Lopez, A. (2012). Pairwise-difference estimation of incomplete information games. *Journal of Econometrics*, 168(1):120–140.
- Aradillas-Lopez, A., Honoré, B., and Powell, J. (2007). Pairwise difference estimation with nonparametric control variables. *International Economic Review*, 48(4):1119–1158.
- Armstrong, T. B., Bertanha, M., and Hong, H. (2014). A fast resample method for parametric and semiparametric models. *Journal of Econometrics*, 179(2):128–133.
- Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.
- Berry, S., Carnall, M., and Spiller, P. T. (2006). Airline hubs: costs, markups and the implications of customer heterogeneity. *Competition Policy and Antitrust*, pages 183–213.
- Berry, S. and Jia, P. (2010). Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43.
- Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.
- Bontemps, C., Gualdani, C., and Remmy, K. (2023). Price competition and endogenous product choice in networks: Evidence from the us airline industry. Technical report, Working Paper.
- Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). Economics of the family. *Cambridge Books*.
- Bunting, J. (2022). Continuous permanent unobserved heterogeneity in dynamic discrete choice models. *arXiv preprint arXiv:2202.03960*.
- Bunting, J., Diegert, P., and Maurel, A. (2022). Heterogeneity, uncertainty and learning: Semiparametric identification and estimation. *Working Paper*.

- Caoui, E. H. and Steck, A. (2026). Diversification, market entry, and the global internet backbone. *Review of Economic Studies*.
- Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. *Handbook of Econometrics*, 6:5549–5632.
- Chen, X., Linton, O., and van Keilegom, I. (2003). Estimation of semiparametric models when the criterion function is not smooth. *Econometrica*, 71(5):1591–1608.
- Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–201.
- Ciliberto, F., Murry, C., and Tamer, E. (2021). Market structure and competition in airline markets. *Journal of Political Economy*, 129(11):2995–3038.
- Ciliberto, F. and Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828.
- Das, M., Newey, W., and Vella, F. (2003). Nonparametric estimation of sample selection models. *The Review of Economic Studies*, 70(1):33–58.
- De Paula, Á., Richards-Shubik, S., and Tamer, E. (2018). Identifying preferences in networks with bounded degree. *Econometrica*, 86(1):263–288.
- Draganska, M., Mazzeo, M., and Seim, K. (2009). Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *Quantitative Marketing and Economics*, 7:105–146.
- Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *The Review of Economic Studies*, 81(3):1003–1045.
- Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628.
- Fan, Y. and Yang, C. (2020). Competition, product proliferation, and welfare: A study of the us smartphone market. *American Economic Journal: Microeconomics*, 12(2):99–134.
- Galichon, A. and Salanié, B. (2022). Cupid’s invisible hand: Social surplus and identification in matching models. *The Review of Economic Studies*, 89(5):2600–2629.
- Gonçalves, S., Hounyo, U., Patton, A. J., and Sheppard, K. (2023). Bootstrapping two-stage quasi-maximum likelihood estimators of time series models. *Journal of Business & Economic Statistics*, 41(3):683–694.
- Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica*, 85(4):1033–1063.

- Grieco, P. (2014). Discrete games with flexible information structures: An application to local grocery markets. *The RAND Journal of Economics*, 45(2):303–340.
- Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189.
- Hu, Y. and Xin, Y. (2022). Identification and estimation of dynamic structural models with unobserved choices. *Available at SSRN 3634910*.
- Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316.
- Kasahara, H. and Shimotsu, K. (2014). Non-parametric identification and estimation of the number of components in multivariate mixtures. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):97–111.
- Li, S., Mazur, J., Park, Y., Roberts, J., Sweeting, A., and Zhang, J. (2022). Repositioning and market power after airline mergers. *The RAND Journal of Economics*, 53(1):166–199.
- Liu, H. and Luo, Y. (2025). Demand analysis under price rigidity and endogenous assortment: An application to china’s tobacco industry. *arXiv preprint arXiv:2501.17251*.
- Newey, W. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal*, 12:S217–S229.
- Pilla, R. S. and Lindsay, B. G. (2001). Alternative em methods for nonparametric finite mixture models. *Biometrika*, 88(2):535–550.
- Powell, J. (2001). Semiparametric estimation of censored selection models. *Nonlinear Statistical Modeling*, pages 165–96.
- Rothenberg, T. (1971). Identification in parametric models. *Econometrica*, 39(3):577–591.
- Schaumans, C. and Verboven, F. (2015). Entry and competition in differentiated products markets. *Review of Economics and Statistics*, 97(1):195–209.
- Smith, H. (2004). Supermarket choice and supermarket competition in market equilibrium. *The Review of Economic Studies*, 71(1):235–263.
- 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.

- Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70(1):331–341.
- Williams, B. (2020). Nonparametric identification of discrete choice models with lagged dependent variables. *Journal of Econometrics*, 215(1):286–304.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Online Appendix

A Notation guide

Table A.1: Notation Used in the Main Text

Symbol	Definition	Symbol	Definition
<i>Indices and observed variables</i>		<i>Entry and selection</i>	
$h;$	Consumer; potential product and	$\pi_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt});$	Entry profit averaged over rivals'
$j \in \mathcal{J} \equiv \{1, \dots, J\};$	firm (in the single-product model);	$P_{jt} \equiv P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t)$	BNE choices; BNE entry probability,
$t \in \{1, \dots, T\}$	market.		or latent propensity score.
$\mathbf{x}_t; \mathbf{x}_{jt}^D, \mathbf{x}_{jt}^{\text{mc}}, \mathbf{x}_{jt}^{\text{fc}}$	All exogenous variables observed by the researcher, including every potential product's characteristics; product- j demand, variable-cost, and fixed-cost subvectors.	$\bar{P}_j(\mathbf{x}_t)$	Ordinary propensity score, $\Pr(a_{jt} = 1 \mid \mathbf{x}_t) = \int P_j(\mathbf{x}_t, \boldsymbol{\kappa}) dF_{\boldsymbol{\kappa}}(\boldsymbol{\kappa})$.
$a_{jt}; \mathbf{a}_t; \mathcal{J}_t^a$	Offer indicator; vector of offer indicators; set of products offered in market t .	$\mu_j(\boldsymbol{\kappa})$	$\mathbb{E}(\xi_{jt} \mid \boldsymbol{\kappa}_t = \boldsymbol{\kappa})$, the conditional mean demand shock.
$p_{jt}; s_{jt}; \mathbf{s}_t^a$	Price; market share; vector of shares of the offered products.	$\lambda_j(\mathbf{x}_t); \tilde{\xi}_{jt}$	Selection-bias function $\mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$; residual $\xi_{jt} - \lambda_j(\mathbf{x}_t)$.
<i>Demand</i>		<i>Finite-mixture representation</i>	
$\alpha, \boldsymbol{\beta}, \sigma, \xi_{jt}$	Price coefficient, observed-characteristic coefficients, random-coefficient parameters, and mean-zero demand unobservable.	$L; \{\mathcal{K}_\ell\}_{\ell=1}^L$	Generic or candidate number of latent market types; positive-probability L -cell partition of $\mathcal{K}$.
δ_{jt}	Mean utility, $\alpha p_{jt} + (\mathbf{x}_{jt}^D)' \boldsymbol{\beta} + \xi_{jt}$.	$\tilde{f}_\ell$	Type probability $F_{\boldsymbol{\kappa}}(\mathcal{K}_\ell)$.
$\mathbf{v}_{ht}, F_v; \boldsymbol{\Omega}_\sigma; \varepsilon_{hjt}$	Consumer heterogeneity and its distribution; random-coefficient loading matrix; i.i.d. type-I extreme-value shock.	$\tilde{P}_{j,\ell}(\mathbf{x}_t); \tilde{\mu}_{j,\ell}$	Type-specific propensity score and mean demand shock: the respective cell averages of $P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t)$ and ξ_{jt} over $\mathcal{K}_\ell$.
$d_{jt}^{-1}(\mathbf{s}_t^a, \sigma)$	Product- j mean utility obtained by inverting the offered-product demand system.	L^*	Population number of latent market types satisfying Assumption 2; the minimal number reproducing the joint entry distribution.
$\boldsymbol{\theta} \equiv (\alpha, \boldsymbol{\beta}', \sigma')'$	Demand-parameter vector. In the nested-logit specialization, σ reduces to the scalar nesting parameter σ .	$r_{j\ell t}; \mathbf{r}_{jt}; \mathbf{R}_{jt}$	Selection-function regressor $\tilde{f}_\ell(\tilde{P}_{j,\ell} - \tilde{P}_{j,L^*})/\bar{P}_j$; its $L^* - 1$ stack; corresponding $J(L^* - 1)$ -vector with only product j 's block nonzero.
<i>Costs and information at entry</i>		<i>Identification and estimation</i>	
$\omega_{jt}; mc_{jt}; fc_{jt}; \xi_t, \omega_t$	Mean-zero variable-cost unobservable, marginal cost, and fixed entry cost; vectors of product-level demand and variable-cost unobservables.	$\tilde{\boldsymbol{\mu}}; \boldsymbol{\theta} \equiv (\boldsymbol{\theta}', \tilde{\boldsymbol{\mu}})'$	Stacked product-specific selection-function coefficients; full second-step parameter vector.
$\boldsymbol{\kappa}_t; \mathcal{K}; F_{\boldsymbol{\kappa}}$	Latent market state, unobserved by the researcher and common knowledge among firms; its support and marginal distribution.	$\mathbf{z}_{jt}; \mathbf{w}_{jt}; \mathbf{m}(\boldsymbol{\theta}); \mathbf{M}(\boldsymbol{\theta}_0)$	Excluded instruments; vector of excluded instruments and included regressors; pooled selected-sample moment function; its Jacobian at the true value.
$\boldsymbol{\eta}_{jt}; F_{\boldsymbol{\eta}}; \boldsymbol{\eta}_{jt}; F_{\boldsymbol{\eta}}; \sigma_{\boldsymbol{\eta}_j}$	General private fixed-cost information and distribution (vectors i.i.d. across firms); scalar recovery specialization, known CDF, and firm-specific scale.	$\mathbf{b}_t, N_X; \boldsymbol{\gamma}_{j,\ell}, \boldsymbol{\gamma}; \Lambda; \tilde{\mathbf{f}}$	Vector of N_X sieve basis functions; product- and type-specific coefficients and their stack; logistic CDF; type-probability vector.

Table A.2: Additional Notation Used in the Online Appendix

Symbol	Definition
<i>Multi-product firms</i>	
$f \in \mathcal{F}; \mathcal{J}_f$	Firm index and set of products owned by firm f ; in this extension, j indexes products only.
$\mathbf{a}_f, \mathbf{a}_{-f}; \mathbf{a}_{ft}, \mathbf{a}_{-ft}$	Generic firm- f portfolio and rival-product entry profile; their market- t counterparts.
$\mathbf{x}_{ft}^{\text{fc}};$	Portfolio fixed-cost observables and vector of portfolio-specific private cost shocks.
$\boldsymbol{\eta}_{ft} \equiv (\eta_{ft}(\mathbf{a}_f) : \mathbf{a}_f \in \{0, 1\}^{ \mathcal{J}_f })$	
<i>Dynamic game of product entry and exit</i>	
$t; \mathbf{x}_t^{ex}; \mathbf{x}_t \equiv (\mathbf{x}_t^{ex}, \mathbf{a}_{t-1})$	Time within a market (whose index is suppressed); exogenous state; dynamic observed state, including lagged availability.
$fc(a_{jt}, a_{j,t-1}, \mathbf{x}_{jt}^{\text{fc}}, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt})$	Dynamic fixed and transition cost.
$P = (P_i : i \in \mathcal{J}); v_j^P$	Candidate profile of conditional choice probabilities (CCPs), which are latent propensity scores at an MPE; difference between firm j 's choice-specific value functions for $a_{jt} = 1$ and $a_{jt} = 0$ under P .
$F_{\kappa x}(\cdot \mathbf{x}_t)$	Conditional distribution of the latent market state in the dynamic model.
<i>Identification under normal heterogeneity</i>	
$\boldsymbol{\tau}_j; \zeta_{jt} \equiv \mathbf{x}_t' \boldsymbol{\tau}_j + \zeta_{jt}; \boldsymbol{\zeta}_t; G_j, G_{jk}$	Latent-index coefficients, product- j index, and vector of indexes; $P_j(\mathbf{x}_t, \boldsymbol{\zeta}_t) = G_j(\boldsymbol{\zeta}_t)$ and $G_{jk} = \partial G_j / \partial \zeta_k$.
$x_{k1t}; \tau_{k1}; \Sigma_{jk} \equiv \text{Cov}(\zeta_{jt}, \zeta_{kt}); \varrho_{jk} \equiv \Sigma_{jk} / \tau_{k1}$	Product- k special regressor and its index coefficient; demand-shock covariance; composite coefficient on the corresponding derivative regressor.
<i>Linearized common-weight bootstrap</i>	
$c_\ell; \boldsymbol{\psi}$	Log ratio $\log(\tilde{f}_\ell / \tilde{f}_1)$, with $c_1 = 0$; unconstrained first-step parameter vector combining the log ratios and γ .
$\ell_t^{1st}, \dot{\ell}_t^{1st}, \hat{\mathbf{m}}_t; \hat{\mathbf{H}}$	Market first-step log-likelihood contribution and score; market second-step moment contribution; information-matrix estimator (distinct from H_t).
$w_t^{*(b)}; b = 1, \dots, B_{\text{boot}}; \hat{\boldsymbol{\psi}}^{*(b)}; \hat{\boldsymbol{g}}^{*(b)}; \hat{\mathbf{W}}^{*(b)}$	Common market weight and replication index; updated first-step parameters, weighted second-step moments, and GMM weighting matrix in replication b .
$\hat{\Delta}_{AB}^{*(b)}; \hat{\mathbf{V}}_{AB}$	Bootstrap difference $\hat{\boldsymbol{\theta}}_q^{B,*(b)} - \hat{\boldsymbol{\theta}}_q^{A,*(b)}$ for the common parameter subvector; its estimated variance-covariance matrix.
<i>ANOVA–Sobol decomposition</i>	
$d \in \{1, 2\}; T_j$	Direction index within non-directional entry market t (distinct from the demand map d_{jt}); number of retained potential-entry markets for airline j .
$VP_{j\ell t} \equiv VP_{j\ell}^P(\mathbf{x}_t); fc_{j\ell t}; \Pi_{j\ell t} \equiv VP_{j\ell t} - fc_{j\ell t}; Y_{j\ell t} \in \{\Pi_{j\ell t}, VP_{j\ell t}, fc_{j\ell t}\}$	Type-specific expected variable profit; fixed-cost component excluding the private shock; net entry-profit index before the private shock (distinct from Π_{jt}); generic object being decomposed.
$\bar{Y}_j; m_j^Y(t); n_j^Y(\ell); u_{j\ell t}^Y$	Overall, market-specific, and type-specific means; market-type interaction residual (the m_j^Y notation is unrelated to GMM moments).
$V_j^Y; V_{j,x}^Y; V_{j,\ell}^Y; V_{j,\text{res}}^Y; V_{j,\text{unobs}}^Y; \text{Cov}_{j,\ell}^{VP,fc}, \text{Cov}_{j,\text{res}}^{VP,fc}, \text{Cov}_{j,\text{unobs}}^{VP,fc}$	Total variance; observable, latent-type, interaction, and total unobservable components; corresponding covariance components of expected variable profit and fixed cost.

B Multi-product firms

This section sketches how the argument of Proposition 3 can be extended to multi-product firms. Products remain indexed by $j \in \mathcal{J}$, and firms are indexed by $f \in \mathcal{F}$. Let $\{\mathcal{J}_f : f \in \mathcal{F}\}$ be a partition of $\mathcal{J}$, where $\mathcal{J}_f$ is the set of products owned by firm f . Firm f 's portfolio choice in market t is represented by $\mathbf{a}_{ft} \equiv (a_{jt} : j \in \mathcal{J}_f) \in \{0, 1\}^{|\mathcal{J}_f|}$, while $\mathbf{a}_{-ft} \equiv (a_{jt} : j \in \mathcal{J} \setminus \mathcal{J}_f)$ collects the entry indicators for products owned by other firms.

Proposition 1 continues to hold because its inversion argument depends only on the demand system, not on product ownership. Consequently, for any product j and values of $\mathbf{x}_t$ such that $\bar{P}_j(\mathbf{x}_t) > 0$, the product-level selection-bias function remains $\mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, a_{jt} = 1)$.

At the entry stage, a firm's private information must be indexed by the portfolios it may choose. Let $\mathbf{a}_f \equiv (a_j : j \in \mathcal{J}_f) \in \{0, 1\}^{|\mathcal{J}_f|}$ denote a generic portfolio and define $\boldsymbol{\eta}_{ft} \equiv (\eta_{ft}(\mathbf{a}_f) : \mathbf{a}_f \in \{0, 1\}^{|\mathcal{J}_f|})$. Let $\mathbf{x}_{ft}^{\text{fc}}$ collect the components of $\mathbf{x}_t$ that enter firm f 's portfolio fixed cost, including the relevant market-level variables and characteristics of the products in $\mathcal{J}_f$. We maintain the analogue of Assumption 1[c]: the vectors $\boldsymbol{\eta}_{ft}$ are independent across firms and independent of $(\boldsymbol{\xi}_t, \boldsymbol{\omega}_t, \boldsymbol{\kappa}_t, \mathbf{x}_t)$, while their components may be dependent within a firm. The analogue of Assumption 1[d] is also maintained, with $\{\boldsymbol{\eta}_{ft} : f \in \mathcal{F}\}$ replacing $\{\boldsymbol{\eta}_{jt} : j \in \mathcal{J}\}$. For example, for a two-product firm, $\boldsymbol{\eta}_{ft} \equiv (\eta_{ft}(0, 0), \eta_{ft}(0, 1), \eta_{ft}(1, 0), \eta_{ft}(1, 1))$, where $\eta_{ft}(1, 0)$ is the private component of fixed cost when the firm offers the first product but not the second. We normalize the entire fixed cost of the empty portfolio to zero, including $\eta_{ft}(0, 0) = 0$ in this example.

Under this information structure, and using $\boldsymbol{\xi}_t$ in the following expected-profit expression as shorthand for both demand and variable-cost unobservables, the analogue of equation (8) for a profile of rival product choices $\mathbf{a}_{-f} \in \{0, 1\}^{J-|\mathcal{J}_f|}$ is

$$\begin{aligned} \pi_f(\mathbf{a}_f, \mathbf{a}_{-f}, \mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{ft}) &= \int VP_f(\mathbf{a}_f, \mathbf{a}_{-f}, \mathbf{x}_t, \boldsymbol{\xi}_t) dF_{f, \boldsymbol{\xi}}(\boldsymbol{\xi}_t \mid \boldsymbol{\kappa}_t) \\ &\quad - fc(\mathbf{a}_f, \mathbf{x}_{ft}^{\text{fc}}, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{ft}). \end{aligned}$$

Here $F_{f, \boldsymbol{\xi}}(\cdot \mid \boldsymbol{\kappa}_t)$ denotes firm f 's conditional beliefs about $\boldsymbol{\xi}_t$. Maintaining product-separable variable costs, $VP_f(\mathbf{a}_f, \mathbf{a}_{-f}, \mathbf{x}_t, \boldsymbol{\xi}_t)$ is firm f 's indirect variable profit: the sum of the equilibrium variable profits of the products in $\mathcal{J}_f$ offered under $\mathbf{a}_f$,

evaluated at the Bertrand–Nash equilibrium selected under the convention in the main text, in which every multi-product firm jointly sets the prices of its offered products. The fixed-cost function depends on η_{ft} through the component $\eta_{ft}(a_f)$ associated with the chosen portfolio.

Given these payoffs, a BNE is characterized, conditional on (x_t, κ_t) , by one probability distribution over mutually exclusive product portfolios for each firm, induced by each firm’s maximization of expected profit after averaging over rival firms’ portfolio choices. For any product $j \in \mathcal{J}_f$, firm f ’s conditional portfolio-choice probabilities induce the latent propensity score for product j :

$$P_j(x_t, \kappa_t) \equiv \Pr(a_{jt} = 1 \mid x_t, \kappa_t) = \sum_{a_f: a_j=1} \Pr(a_{ft} = a_f \mid x_t, \kappa_t).$$

Because firm f ’s equilibrium portfolio choice depends on the unobservables only through η_{ft} conditional on (x_t, κ_t) , the maintained independence of η_{ft} from $(\xi_t, \omega_t, \kappa_t, x_t)$ implies that a_{jt} and ξ_{jt} are conditionally independent given (x_t, κ_t) . Together with the maintained analogue of Assumption 1[d], Proposition 3 therefore continues to hold: for $\bar{P}_j(x_t) > 0$, the selection-bias function $\lambda_j(x_t)$ admits the mixture representation in equation (16), with the latent propensity score defined above. This argument requires no conditional independence among the entry indicators of products owned by the same firm. Multi-product ownership therefore changes how the latent propensity score for each product is obtained from firm-level portfolio choices, but not the general selection-bias representation.

C Dynamic game of product entry and exit

This section sketches how our framework extends when forward-looking firms’ product-availability decisions arise from a Markov Perfect Equilibrium (MPE) of a dynamic game of product entry and exit. We write firm j ’s fixed and transition cost as $fc(a_{jt}, a_{j,t-1}, x_{jt}^{\text{fc}}, \kappa_t, \eta_{jt})$. The configurations $(a_{jt}, a_{j,t-1}) = (1, 0), (0, 1), (1, 1)$, and $(0, 0)$ correspond, respectively, to entry, exit, continued operation, and continued inactivity; we normalize $fc(0, 0, x_{jt}^{\text{fc}}, \kappa_t, \eta_{jt}) = 0$.

ASSUMPTION 1-Dyn. *Suppose that t indexes time within a market; we suppress the market index. Conditions [a]–[c] of Assumption 1 continue to hold, with η_{jt} representing*

private information about current fixed and transition costs. We do not impose condition [d], because the observed state now contains lagged availability decisions and may therefore be correlated with the latent market state. The following additional conditions hold.

- e. Let $\mathbf{x}_t \equiv (\mathbf{x}_t^{ex}, \mathbf{a}_{t-1})$ denote the observed state at period t , where $\mathbf{x}_t^{ex}$ collects the exogenous market- and product-level observables and $\mathbf{a}_{t-1} \equiv (a_{j,t-1} : j \in \mathcal{J})$ collects lagged availability decisions.
- f. The joint process $\{(\mathbf{x}_t^{ex}, \boldsymbol{\kappa}_t)\}_t$ is first-order Markov, with time-invariant components allowed.
- g. The private-information vectors $\boldsymbol{\eta}_{jt}$ are i.i.d. across firms and over time.
- h. Conditional on the latent market state $\boldsymbol{\kappa}_t$, the current demand and variable-cost unobservables $(\boldsymbol{\xi}_t, \boldsymbol{\omega}_t)$ are independent of $\mathbf{x}_t^{ex}$ and the past market history. Consequently, $\mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, \boldsymbol{\kappa}_t) = \mu_j(\boldsymbol{\kappa}_t)$. ■

The conditions in Assumption 1-Dyn are standard in the literature on empirical dynamic games of oligopoly competition (see [Aguirregabiria et al., 2021](#)). Given a candidate profile P , let $P_{it} \equiv P_i(\mathbf{x}_t, \boldsymbol{\kappa}_t)$ denote firm i 's conditional choice probability (CCP) of choosing $a_{it} = 1$ at state $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$, and let $v_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt})$ denote the difference between firm j 's choice-specific value functions associated with $a_{jt} = 1$ and $a_{jt} = 0$. This difference equals the difference in current profits plus the discounted difference in expected continuation values.

DEFINITION 2. Markov Perfect Equilibrium. Under Assumption 1-Dyn, an MPE can be represented as a J -tuple of CCPs $\{P_{jt} \equiv P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t) : j \in \mathcal{J}\}$ that solves the following system of J fixed-point equations at every state:

$$P_{jt} = \int \mathbb{1} \left\{ v_j^P(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt}) \geq 0 \right\} dF_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{jt}). \quad \blacksquare$$

At an MPE, P_{jt} is also the latent propensity score for firm j 's product-availability decision. Conditional on $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$, this decision is a function of $\boldsymbol{\eta}_{jt}$; hence, Assumption 1[c] implies that a_{jt} and ξ_{jt} are conditionally independent given $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$. Together with

Assumption 1-Dyn[h], the conditioning argument underlying Proposition 3 yields

$$\begin{aligned}\bar{P}_j(\mathbf{x}_t) &\equiv \Pr(a_{jt} = 1 \mid \mathbf{x}_t) = \int P_j(\mathbf{x}_t, \boldsymbol{\kappa}) dF_{\boldsymbol{\kappa}|\mathbf{x}}(\boldsymbol{\kappa} \mid \mathbf{x}_t), \\ \lambda_j(\mathbf{x}_t) &= \int \left[\frac{P_j(\mathbf{x}_t, \boldsymbol{\kappa})}{\bar{P}_j(\mathbf{x}_t)} \right] \mu_j(\boldsymbol{\kappa}) dF_{\boldsymbol{\kappa}|\mathbf{x}}(\boldsymbol{\kappa} \mid \mathbf{x}_t), \quad \bar{P}_j(\mathbf{x}_t) > 0,\end{aligned}$$

where $F_{\boldsymbol{\kappa}|\mathbf{x}}(\cdot \mid \mathbf{x}_t)$ is the conditional distribution of the latent market state given the observed state. The conditional rather than marginal distribution is required because $\mathbf{x}_t$ contains lagged availability decisions and may therefore be correlated with $\boldsymbol{\kappa}_t$.

D Proofs of Propositions

D.1 Proof of Proposition 1

Fix a market t , its prices and observed demand characteristics, a configuration $\mathbf{a}_t$, and a value of σ . Let $n \equiv |\mathcal{J}_t^a|$, and write $\boldsymbol{\delta} \equiv \boldsymbol{\delta}_t^a$ and $\mathbf{d}(\boldsymbol{\delta}) \equiv \mathbf{d}_t^a(\boldsymbol{\delta}, \sigma)$, suppressing the fixed arguments. If $n = 0$, both $\mathbb{R}^0$ and $\mathcal{S}^a$ contain only the empty vector, so the result is immediate. Hence, suppose that $n \geq 1$.

For $j \in \{0\} \cup \mathcal{J}_t^a$, let $\tilde{d}_j(\boldsymbol{\delta}, \mathbf{v})$ denote product j 's conditional demand share implied by equation (3), that is, its logit choice probability given $\mathbf{v}$. Define

$$d_j(\boldsymbol{\delta}) \equiv \int \tilde{d}_j(\boldsymbol{\delta}, \mathbf{v}) dF_v(\mathbf{v}), \quad j \in \{0\} \cup \mathcal{J}_t^a.$$

For $j \in \mathcal{J}_t^a$, this is the j -th component of $\mathbf{d}(\boldsymbol{\delta})$; $d_0(\boldsymbol{\delta})$ is the outside share. Pointwise, every conditional share is strictly positive and

$$\tilde{d}_0(\boldsymbol{\delta}, \mathbf{v}) + \sum_{j \in \mathcal{J}_t^a} \tilde{d}_j(\boldsymbol{\delta}, \mathbf{v}) = 1.$$

The derivatives of the conditional logit shares with respect to $\boldsymbol{\delta}$ are continuous and uniformly bounded, so differentiation under the integral is valid and $\mathbf{d}$ is continuously differentiable. Its $n \times n$ Jacobian $\mathbf{D}(\boldsymbol{\delta})$ has entries

$$D_{ji}(\boldsymbol{\delta}) = \int \tilde{d}_j(\boldsymbol{\delta}, \mathbf{v}) \left[\mathbb{1}\{i = j\} - \tilde{d}_i(\boldsymbol{\delta}, \mathbf{v}) \right] dF_v(\mathbf{v}), \quad i, j \in \mathcal{J}_t^a.$$

For every nonzero $\mathbf{q} = (q_j : j \in \mathcal{J}_t^a) \in \mathbb{R}^n$,

$$\mathbf{q}' \mathbf{D}(\delta) \mathbf{q} = \int \left\{ \sum_{j \in \mathcal{J}_t^a} \tilde{d}_j(\delta, \mathbf{v}) q_j^2 - \left[\sum_{j \in \mathcal{J}_t^a} \tilde{d}_j(\delta, \mathbf{v}) q_j \right]^2 \right\} dF_v(\mathbf{v}) > 0.$$

Indeed, the Cauchy–Schwarz inequality and the fact that the conditional inside shares sum to $1 - \tilde{d}_0$ imply that the expression in braces is at least $\tilde{d}_0 \sum_{j \in \mathcal{J}_t^a} \tilde{d}_j q_j^2 > 0$, where the arguments of the conditional shares are suppressed. Thus, $\mathbf{D}(\delta)$ is symmetric positive definite and therefore nonsingular everywhere.

For any $\delta^{(0)} \neq \delta^{(1)}$, let $\mathbf{q} \equiv \delta^{(1)} - \delta^{(0)}$. The fundamental theorem of calculus gives

$$\mathbf{q}' [\mathbf{d}(\delta^{(1)}) - \mathbf{d}(\delta^{(0)})] = \int_0^1 \mathbf{q}' \mathbf{D}(\delta^{(0)} + \tau \mathbf{q}) \mathbf{q} d\tau > 0.$$

Hence, distinct mean-utility vectors generate distinct market-share vectors, so $\mathbf{d}$ is injective. Moreover, $\mathbf{d}(\mathbb{R}^n) \subseteq \mathcal{S}^a$, because every inside share is positive and their sum equals $1 - d_0(\delta) < 1$. Let $\mathcal{I} \equiv \mathbf{d}(\mathbb{R}^n)$. Since the Jacobian is nonsingular everywhere, the inverse function theorem implies that $\mathcal{I}$ is open in $\mathbb{R}^n$ and therefore relatively open in $\mathcal{S}^a$.

To show that $\mathcal{I}$ is also relatively closed, consider any sequence

$$\mathbf{d}(\delta^{(m)}) \longrightarrow \mathbf{s} \in \mathcal{S}^a.$$

If $\{\delta^{(m)}\}$ were unbounded, then there would be a fixed $j \in \mathcal{J}_t^a$ and a subsequence, not relabeled, along which either $\delta_j^{(m)} \rightarrow -\infty$ or $\delta_j^{(m)} \rightarrow +\infty$. In the first case, $\tilde{d}_j(\delta^{(m)}, \mathbf{v}) \rightarrow 0$ pointwise, so dominated convergence implies $d_j(\delta^{(m)}) \rightarrow 0$, contradicting $s_j > 0$. In the second case, $\tilde{d}_0(\delta^{(m)}, \mathbf{v}) \rightarrow 0$ pointwise, so dominated convergence implies $d_0(\delta^{(m)}) \rightarrow 0$. But convergence of the inside shares also implies

$$d_0(\delta^{(m)}) \longrightarrow 1 - \sum_{j \in \mathcal{J}_t^a} s_j > 0,$$

a contradiction. Therefore, $\{\delta^{(m)}\}$ is bounded. It has a convergent subsequence, and continuity of $\mathbf{d}$ implies that $\mathbf{s} \in \mathcal{I}$. Thus, $\mathcal{I}$ is relatively closed in $\mathcal{S}^a$. The set $\mathcal{S}^a$ is convex and hence connected. Since $\mathcal{I}$ is nonempty and both relatively open and relatively closed in $\mathcal{S}^a$, it follows that $\mathcal{I} = \mathcal{S}^a$. Therefore, $\mathbf{d}_t^a(\cdot, \sigma)$ is a bijection from

$\mathbb{R}^n$ onto $\mathcal{S}^a$, and $(d_t^a)^{-1}(\cdot, \sigma)$ is well defined on $\mathcal{S}^a$. ■

D.2 Proof of Proposition 3

Fix $\mathbf{x}_t$ such that $\bar{P}_j(\mathbf{x}_t) > 0$, as in the proposition. By definition,

$$\lambda_j(\mathbf{x}_t) = \mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, a_{jt} = 1) = \frac{\mathbb{E}(\xi_{jt} a_{jt} \mid \mathbf{x}_t)}{\Pr(a_{jt} = 1 \mid \mathbf{x}_t)}. \quad (38)$$

Under the equilibrium entry rule, a_{jt} is a function of $(\mathbf{x}_t, \boldsymbol{\kappa}_t, \boldsymbol{\eta}_{jt})$. Assumption 1[c] makes $\boldsymbol{\eta}_{jt}$ independent of ξ_{jt} conditional on $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$; hence, a_{jt} and ξ_{jt} are conditionally independent given $(\mathbf{x}_t, \boldsymbol{\kappa}_t)$. It follows that

$$\begin{aligned} \mathbb{E}(\xi_{jt} a_{jt} \mid \mathbf{x}_t) &= \mathbb{E}(\mathbb{E}(\xi_{jt} a_{jt} \mid \mathbf{x}_t, \boldsymbol{\kappa}_t) \mid \mathbf{x}_t) \\ &= \mathbb{E}(\mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, \boldsymbol{\kappa}_t) \mathbb{E}(a_{jt} \mid \mathbf{x}_t, \boldsymbol{\kappa}_t) \mid \mathbf{x}_t) \\ &= \mathbb{E}(\mu_j(\boldsymbol{\kappa}_t) P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t) \mid \mathbf{x}_t). \end{aligned} \quad (39)$$

The first equality is the law of iterated expectations, and the second uses the conditional independence just established. The third uses the definition of the latent propensity score,

$$\mathbb{E}(a_{jt} \mid \mathbf{x}_t, \boldsymbol{\kappa}_t) = P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t),$$

and Assumption 1[d], which implies

$$\mathbb{E}(\xi_{jt} \mid \mathbf{x}_t, \boldsymbol{\kappa}_t) = \mathbb{E}(\xi_{jt} \mid \boldsymbol{\kappa}_t) = \mu_j(\boldsymbol{\kappa}_t).$$

Substituting (39) into (38) and using $\Pr(a_{jt} = 1 \mid \mathbf{x}_t) = \bar{P}_j(\mathbf{x}_t)$ gives

$$\lambda_j(\mathbf{x}_t) = \frac{\mathbb{E}(\mu_j(\boldsymbol{\kappa}_t) P_j(\mathbf{x}_t, \boldsymbol{\kappa}_t) \mid \mathbf{x}_t)}{\bar{P}_j(\mathbf{x}_t)}.$$

Assumption 1[d] also implies that the conditional distribution of κ_t given $\mathbf{x}_t$ equals its marginal distribution F_κ . Hence,

$$\lambda_j(\mathbf{x}_t) = \int \left[\frac{P_j(\mathbf{x}_t, \boldsymbol{\kappa})}{\bar{P}_j(\mathbf{x}_t)} \right] \mu_j(\boldsymbol{\kappa}) dF_\kappa(\boldsymbol{\kappa}).$$

This proves Proposition 3. $\blacksquare$

E Identification under normal heterogeneity

This section develops the first of the two identification routes introduced in Section 4. Here, the latent market state coincides with the jointly normal vector of demand unobservables, $\kappa_t = \boldsymbol{\xi}_t$. Under the index, exclusion, and smoothness restrictions stated below, a multivariate version of Stein's lemma expresses the selection-bias function $\lambda_j(\mathbf{x}_t)$ as a linear combination of derivatives of the logarithm of the ordinary propensity score.

E.1 Setup

We specialize the general framework of Section 4 by setting $\kappa_t = \boldsymbol{\xi}_t \equiv (\xi_{1t}, \dots, \xi_{Jt})'$, so that the demand unobservables exhaust the latent market state and are common knowledge at entry. The conditional-mean function in Proposition 3 then satisfies $\mu_j(\boldsymbol{\kappa}) = \mathbb{E}(\xi_{jt} \mid \kappa_t = \boldsymbol{\kappa}) = \kappa_j$, where κ_j is the j -th component of $\boldsymbol{\kappa}$. Hence, equation (16) becomes

$$\lambda_j(\mathbf{x}_t) = \frac{1}{\bar{P}_j(\mathbf{x}_t)} \mathbb{E}[\xi_{jt} P_j(\mathbf{x}_t, \boldsymbol{\xi}_t)]. \quad (40)$$

We assume that the latent propensity score $P_j(\mathbf{x}_t, \boldsymbol{\xi}_t)$ depends on $(\mathbf{x}_t, \boldsymbol{\xi}_t)$ only through the vector of product-level latent indexes $\boldsymbol{\zeta}_t = (\zeta_{1t}, \dots, \zeta_{Jt})'$, where

$$\zeta_{jt} \equiv \mathbf{x}_t' \boldsymbol{\tau}_j + \xi_{jt}. \quad (41)$$

That is, $P_j(\mathbf{x}_t, \boldsymbol{\xi}_t) = G_j(\zeta_{1t}, \dots, \zeta_{Jt})$ for some function $G_j : \mathbb{R}^J \rightarrow [0, 1]$. Write $G_{jk}(\boldsymbol{\zeta}) \equiv \partial G_j(\boldsymbol{\zeta}) / \partial \zeta_k$ for the partial derivative with respect to the k -th argument. The additional assumptions required for this approach are collected below.

- ASSUMPTION 3.** (i) *Normality.* The demand unobservables are jointly normal: $\xi_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$, where Σ is a positive definite $J \times J$ covariance matrix with (j, k) entry Σ_{jk} .
- (ii) *Smoothness.* For every $j \in \mathcal{J}$, the function $G_j(\xi)$ is continuously differentiable in ξ . The regularity conditions required to apply Stein's lemma and differentiate under the integral sign hold, and $\mathbb{E} |G_{jk}(\xi_t)| < \infty$ for all $j, k \in \mathcal{J}$.
- (iii) *Index relevance and exclusion.* For every product $k \in \mathcal{J}$, let x_{k1t} denote a continuously distributed component of $\mathbf{x}_t$ associated with product k . It enters ζ_{kt} with coefficient $\tau_{k1} \neq 0$, is excluded from ζ_{it} for every $i \neq k$, and has support, conditional on the remaining regressors, containing an open interval. Thus, x_{k1t} affects every latent propensity score only through ζ_{kt} . ■

E.2 Multivariate Stein's lemma and selection-bias representation

We use the following standard result. If $\xi_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$ and $m : \mathbb{R}^J \rightarrow \mathbb{R}$ is continuously differentiable with the required integrability conditions, then

$$\mathbb{E}(\xi_t m(\xi_t)) = \Sigma \mathbb{E}(\nabla_{\xi} m(\xi_t)). \quad (42)$$

Componentwise, for every $j \in \mathcal{J}$,

$$\mathbb{E}(\xi_{jt} m(\xi_t)) = \sum_{k=1}^J \Sigma_{jk} \mathbb{E}\left(\frac{\partial m(\xi_t)}{\partial \xi_{kt}}\right). \quad (43)$$

Applying this lemma to the latent propensity score yields the following representation of the selection-bias function.

PROPOSITION 5. Under the index structure above and Assumptions 1 and 3, for every $j \in \mathcal{J}$ and every $\mathbf{x}_t$ such that $\bar{P}_j(\mathbf{x}_t) > 0$,

$$\lambda_j(\mathbf{x}_t) = \mathbb{E}(\xi_{jt} | \mathbf{x}_t, a_{jt} = 1) = \sum_{k=1}^J \frac{\Sigma_{jk}}{\tau_{k1}} \frac{\partial \log \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}}. \quad \blacksquare \quad (44)$$

Proof. Step 1: Selection representation. By definition, $\lambda_j(\mathbf{x}_t) = \mathbb{E}(\xi_{jt} a_{jt} | \mathbf{x}_t) / \bar{P}_j(\mathbf{x}_t)$. The law of iterated expectations and the definition of the latent propensity score

give $\mathbb{E}(\xi_{jt} a_{jt} \mid \mathbf{x}_t) = \mathbb{E}[\xi_{jt} P_j(\mathbf{x}_t, \xi_t) \mid \mathbf{x}_t]$. Assumption 1[d] makes the distribution of ξ_t independent of $\mathbf{x}_t$. Therefore,

$$\lambda_j(\mathbf{x}_t) = \frac{\mathbb{E}(\xi_{jt} P_j(\mathbf{x}_t, \xi_t))}{\bar{P}_j(\mathbf{x}_t)}. \quad (45)$$

Here and below, expectations integrate over the marginal distribution of ξ_t , with $\mathbf{x}_t$ held fixed.

Step 2: Stein's lemma. Apply equation (43) with $m(\xi_t) = P_j(\mathbf{x}_t, \xi_t) = G_j(\zeta_{1t}, \dots, \zeta_{Jt})$. Because ξ_{kt} enters only ζ_{kt} , with unit coefficient, the chain rule gives

$$\frac{\partial P_j(\mathbf{x}_t, \xi_t)}{\partial \xi_{kt}} = G_{jk}(\zeta_{1t}, \dots, \zeta_{Jt}). \quad (46)$$

Stein's lemma then gives

$$\mathbb{E}(\xi_{jt} P_j(\mathbf{x}_t, \xi_t)) = \sum_{k=1}^J \Sigma_{jk} \mathbb{E}(G_{jk}(\zeta_{1t}, \dots, \zeta_{Jt})). \quad (47)$$

Step 3: Connecting ξ_t -derivatives to $\mathbf{x}_t$ -derivatives. By Assumption 3(iii), x_{k1t} affects $P_j(\mathbf{x}_t, \xi_t)$ only through ζ_{kt} , and $\partial \zeta_{kt} / \partial x_{k1t} = \tau_{k1}$. The chain rule therefore gives

$$\frac{\partial P_j(\mathbf{x}_t, \xi_t)}{\partial x_{k1t}} = \tau_{k1} G_{jk}(\zeta_{1t}, \dots, \zeta_{Jt}). \quad (48)$$

Moreover, Assumption 1[d] implies $\bar{P}_j(\mathbf{x}_t) = \mathbb{E}[P_j(\mathbf{x}_t, \xi_t)]$, where the distribution of ξ_t does not depend on $\mathbf{x}_t$. Assumption 3(ii) permits differentiation under this expectation. Taking expectations of the preceding identity therefore gives

$$\mathbb{E}(G_{jk}(\zeta_{1t}, \dots, \zeta_{Jt})) = \frac{1}{\tau_{k1}} \frac{\partial \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}}. \quad (49)$$

Step 4: Final expression. Substituting (47) and (49) into (45) yields

$$\lambda_j(\mathbf{x}_t) = \frac{1}{\bar{P}_j(\mathbf{x}_t)} \sum_{k=1}^J \frac{\Sigma_{jk}}{\tau_{k1}} \frac{\partial \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}} = \sum_{k=1}^J \frac{\Sigma_{jk}}{\tau_{k1}} \frac{\partial \log \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}}, \quad (50)$$

where the second equality uses $\partial \log \bar{P}_j / \partial x = (1/\bar{P}_j) \partial \bar{P}_j / \partial x$, which is valid because

$$\bar{P}_j(\mathbf{x}_t) > 0. \quad \square$$

E.3 Regression representation

Define the composite parameters $\varrho_{jk} \equiv \Sigma_{jk}/\tau_{k1}$. Proposition 5 implies that, for observations with $a_{jt} = 1$, the demand equation takes the form

$$d_{jt}^{-1}(\mathbf{s}_t^a, \sigma) = \alpha p_{jt} + \mathbf{x}_{jt}^{D'} \boldsymbol{\beta} + \sum_{k=1}^J \varrho_{jk} \frac{\partial \log \bar{P}_j(\mathbf{x}_t)}{\partial x_{k1t}} + \tilde{\zeta}_{jt}, \quad (51)$$

where $\mathbb{E}(\tilde{\zeta}_{jt} \mid \mathbf{x}_t, a_{jt} = 1) = 0$. The correction therefore uses derivatives of the logarithm of the ordinary propensity score $\bar{P}_j(\mathbf{x}_t)$ with respect to the continuous regressors x_{k1t} that shift the corresponding latent indexes ζ_{kt} . The coefficients $\varrho_{jk} = \Sigma_{jk}/\tau_{k1}$ combine the covariance between $\tilde{\zeta}_{jt}$ and ζ_{kt} with the inverse index loading $1/\tau_{k1}$.

F Linearized common-weight bootstrap

The estimator in Section 4.4 has two steps. The first estimates the type probabilities and type-specific propensity scores by sieve MLE. The second uses these estimates to construct the selection-function regressors and applies GMM to estimate the demand parameters $\boldsymbol{\theta}$ jointly with the stacked product-specific selection-function coefficient vector $\tilde{\boldsymbol{\mu}}$ defined in Section 4.1. Under standard regularity conditions for semiparametric two-step estimators (Das et al., 2003; Newey, 2009), the resulting estimator is asymptotically linear. We describe a linearized common-weight bootstrap that accounts for first-step estimation error and its covariance with the second-step moments without re-estimating the first-step MLE in every replication. General bootstrap results for semiparametric estimators with preliminary nonparametric components are provided by Chen et al. (2003); Armstrong et al. (2014) and Gonçalves et al. (2023) study related computationally efficient bootstrap procedures for multi-step estimators.

Fix L and N_X , and treat the resulting finite-dimensional first-step specification as correctly specified. We present the likelihood-based update for the finite-mixture estimator. For another correctly specified finite-dimensional likelihood-based first

step, its log-likelihood and smooth mapping into the selection-function regressors can be substituted below; a non-likelihood first step instead requires its corresponding first-order influence-function update. Let $\boldsymbol{\psi}$ combine γ with an unconstrained reparameterization of the type probabilities. In particular, set $c_1 = 0$ and define $c_\ell \equiv \log(\tilde{f}_\ell/\tilde{f}_1)$ for $\ell = 2, \dots, L$. The corresponding softmax map is

$$\tilde{f}_1 = \frac{1}{1 + \sum_{k=2}^L \exp(c_k)}, \quad \tilde{f}_\ell = \frac{\exp(c_\ell)}{1 + \sum_{k=2}^L \exp(c_k)}, \quad \ell = 2, \dots, L.$$

Thus, $\boldsymbol{\psi} \equiv (c_2, \dots, c_L, \gamma')'$ can be updated in Euclidean coordinates, while the softmax map preserves the simplex constraint on $\tilde{\mathbf{f}}$. Because the type probabilities are pairwise distinct, their full-sample labeling is locally stable. We retain those labels in every bootstrap replication; the probability that a first-order update crosses a labeling boundary converges to zero as $T \rightarrow \infty$.

Let $\ell_t^{1\text{st}}(\boldsymbol{\psi})$ denote market t 's contribution to the first-step log-likelihood in equation (30), and let $\dot{\ell}_t^{1\text{st}}(\boldsymbol{\psi}) \equiv \partial \ell_t^{1\text{st}}(\boldsymbol{\psi}) / \partial \boldsymbol{\psi}$ be its score. Also let $\hat{\mathbf{m}}_t(\boldsymbol{\theta}, \tilde{\boldsymbol{\mu}}; \boldsymbol{\psi})$ denote the stacked second-step GMM moment contribution of market t , aggregating over the selected products in that market. Define the outer-product-of-gradients estimator of the information matrix by

$$\hat{\mathbf{H}} \equiv \frac{1}{T} \sum_{t=1}^T \dot{\ell}_t^{1\text{st}}(\hat{\boldsymbol{\psi}}) \dot{\ell}_t^{1\text{st}}(\hat{\boldsymbol{\psi}})'$$

Under correct specification and a nonsingular Fisher information matrix, $\hat{\mathbf{H}}$ consistently estimates both the per-market Fisher information and the negative expected Hessian of a market-level log-likelihood contribution, or equivalently of the average first-step log-likelihood.

1. **Step 1: Draw common bootstrap weights.** For replication b , draw multinomial weights $\{w_t^{*(b)} : t = 1, \dots, T\}$ satisfying $\sum_{t=1}^T w_t^{*(b)} = T$, equivalently by sampling markets with replacement. Use the same weights in both steps.
2. **Step 2: Update the first step.** Construct the one-step Fisher-scoring update

$$\hat{\boldsymbol{\psi}}^{*(b)} = \hat{\boldsymbol{\psi}} + \hat{\mathbf{H}}^{-1} \left[\frac{1}{T} \sum_{t=1}^T (w_t^{*(b)} - 1) \dot{\ell}_t^{1\text{st}}(\hat{\boldsymbol{\psi}}) \right].$$

This is the first-order approximation to re-estimating the first-step likelihood on the weighted bootstrap sample. Recover $\hat{\mathbf{f}}^{*(b)}$ from the updated log ratios using the softmax map; the remaining components give $\hat{\boldsymbol{\gamma}}^{*(b)}$.

3. **Step 3: Reconstruct the selection-function regressors.** First compute the bootstrap type-specific and ordinary propensity scores:

$$\begin{aligned}\hat{P}_{j,\ell}^{*(b)}(\mathbf{x}_t) &= \Lambda\left(\mathbf{b}'_t \hat{\boldsymbol{\gamma}}_{j,\ell}^{*(b)}\right), \\ \hat{P}_j^{*(b)}(\mathbf{x}_t) &\equiv \sum_{\ell=1}^L \hat{P}_{j,\ell}^{*(b)}(\mathbf{x}_t) \hat{f}_\ell^{*(b)}.\end{aligned}$$

Then construct

$$\hat{r}_{j\ell t}^{*(b)} = \frac{\hat{P}_{j,\ell}^{*(b)}(\mathbf{x}_t) - \hat{P}_{j,L}^{*(b)}(\mathbf{x}_t) \hat{f}_L^{*(b)}}{\hat{P}_j^{*(b)}(\mathbf{x}_t)} \hat{f}_\ell^{*(b)}, \quad \ell = 1, \dots, L-1.$$

4. **Step 4: Re-estimate the second step.** Define the weighted bootstrap moment vector

$$\hat{\mathbf{g}}^{*(b)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\mu}}) = \frac{1}{T} \sum_{t=1}^T w_t^{*(b)} \hat{\mathbf{m}}_t(\boldsymbol{\theta}, \tilde{\boldsymbol{\mu}}; \hat{\boldsymbol{\psi}}^{*(b)}).$$

Obtain $(\hat{\boldsymbol{\theta}}^{*(b)}, \hat{\tilde{\boldsymbol{\mu}}}^{*(b)})$ by minimizing

$$\hat{\mathbf{g}}^{*(b)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\mu}})' \hat{\mathbf{W}}^{*(b)} \hat{\mathbf{g}}^{*(b)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\mu}}),$$

where $\hat{\mathbf{W}}^{*(b)}$ is constructed by the same rule as the weighting matrix in the original estimator. For a specification without an estimated first step, Steps 2–3 are omitted and Step 4 uses the common bootstrap weights directly.

Repeating these steps for $b = 1, \dots, B_{\text{boot}}$ yields draws $\{\hat{\boldsymbol{\theta}}^{*(b)}, \hat{\tilde{\boldsymbol{\mu}}}^{*(b)}\}_{b=1}^{B_{\text{boot}}}$. We use their sample variance-covariance matrix to estimate that of the second-step estimator. For the Hausman-type comparison of specifications A and B in Section 4.4.3, let $\hat{\boldsymbol{\theta}}_q^A$ and $\hat{\boldsymbol{\theta}}_q^B$ denote estimates of the same q -dimensional subvector of $\boldsymbol{\theta}$, with $q = \dim(\boldsymbol{\theta})$ for the full-vector test. Apply the respective procedures to the same sample of markets, using the same market-level weights in every replication, and

define

$$\begin{aligned}\hat{\Delta}_{AB}^{*(b)} &\equiv \hat{\theta}_q^{B,*(b)} - \hat{\theta}_q^{A,*(b)}, & \bar{\Delta}_{AB}^* &\equiv \frac{1}{B_{\text{boot}}} \sum_{b=1}^{B_{\text{boot}}} \hat{\Delta}_{AB}^{*(b)}, \\ \hat{V}_{AB} &\equiv \frac{1}{B_{\text{boot}} - 1} \sum_{b=1}^{B_{\text{boot}}} (\hat{\Delta}_{AB}^{*(b)} - \bar{\Delta}_{AB}^*) (\hat{\Delta}_{AB}^{*(b)} - \bar{\Delta}_{AB}^*)' .\end{aligned}$$

The common weights preserve the covariance between the two estimators, so $\hat{V}_{AB}$ consistently estimates the variance-covariance matrix of $\hat{\theta}_q^B - \hat{\theta}_q^A$ and is used in equation (34).

First-order validity requires the usual smoothness, moment, and rank conditions for the two-step estimator. In particular, conditional on fixed L and N_X , the finite-dimensional first-step specification must be correctly specified, its score must be sufficiently smooth, and its Fisher information matrix must be nonsingular. The type probabilities must be pairwise distinct, and the mapping from the first-step parameters into the selection-function regressors must be smooth. For the finite-mixture estimator, the ordinary propensity scores $\bar{P}_j(\mathbf{x}_t)$ must also be bounded away from zero uniformly over the support of $\mathbf{x}_t$. Under these conditions, the one-step update is first-order equivalent to re-estimating the first step on each weighted sample. Writing $\hat{\boldsymbol{\theta}} \equiv (\hat{\boldsymbol{\theta}}', \hat{\boldsymbol{\mu}}')'$, consistency of the simulated variance additionally requires conditional uniform integrability of $T\|\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}\|^2$ and $B_{\text{boot}} \rightarrow \infty$. For a Hausman-type comparison, these conditions must hold jointly for the fixed pair of specifications. The procedure conditions on each specification's L and N_X and therefore does not incorporate uncertainty from selecting either.

G ANOVA–Sobol decomposition of entry payoffs

We construct airline-specific ANOVA–Sobol decompositions of the net entry-profit index before the private shock and of its expected-variable-profit and fixed-cost components. Each decomposition separates variation attributable to observable market characteristics, latent market types, and their interaction. Throughout this section, “unobservables” refers to latent-type heterogeneity and its interaction with observables, not to the private entry shock, which is excluded from the index. We do not pool across airlines.

G.1 Setup and entry-payoff objects

Under the scalar private-entry-shock specification introduced in Section 4.3.1, the type-specific propensity scores satisfy $\tilde{P}_{j,\ell}(\mathbf{x}_t) = F_\eta\left(\left[VP_{j\ell}^P(\mathbf{x}_t) - fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})\right] / \sigma_{\eta_j}\right)$, where $VP_{j\ell}^P(\mathbf{x}_t)$ is airline j 's expected variable profit at latent market type ℓ and market characteristics $\mathbf{x}_t$, $fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})$ is the component of fixed entry cost excluding $\sigma_{\eta_j}\eta_{jt}$, and σ_{η_j} is the airline-specific scale of the private entry shock. Consistent with the logistic first-step specification in Section 4.4, we take $F_\eta = \Lambda$ in the empirical fixed-cost recovery and work with fitted propensity scores bounded away from zero and one. We suppress hats on plug-in objects except where they clarify the fixed-cost recovery. Inverting gives

$$VP_{j\ell}^P(\mathbf{x}_t) - fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}}) = \sigma_{\eta_j} \Lambda^{-1}\left(\tilde{P}_{j,\ell}(\mathbf{x}_t)\right). \quad (52)$$

For each airline j , latent type $\ell = 1, \dots, L$, and non-directional entry market t , define

$$VP_{j\ell t} \equiv VP_{j\ell}^P(\mathbf{x}_t), \quad fc_{j\ell t} \equiv fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}}), \quad \Pi_{j\ell t} \equiv VP_{j\ell t} - fc_{j\ell t}.$$

Thus, $\Pi_{j\ell t}$ is the type-specific net entry-profit index before the private shock, rather than the realized active-firm profit Π_{jt} defined in Section 2. Each object is a deterministic function of $\mathbf{x}_t$ and ℓ through the type-specific propensity scores, demand and marginal-cost parameters, and type-specific conditional means, and is defined for every market in which airline j is a potential entrant—not only for markets in which it is observed to enter.

G.2 Computing expected variable profits

We compute $VP_{j\ell}^P(\mathbf{x}_t)$ as in Section 4.3.1, averaging airline j 's equilibrium variable profit over its rivals' entry configurations $\mathbf{a}_{-j} \in \{0, 1\}^{J-1}$:

$$VP_{j\ell}^P(\mathbf{x}_t) = \sum_{\mathbf{a}_{-j} \in \{0,1\}^{J-1}} VP_{j\ell}(\mathbf{a}_{j=1}, \mathbf{a}_{-j}, \mathbf{x}_t) \prod_{i \neq j} \tilde{P}_{i,\ell}(\mathbf{x}_t)^{a_i} \left[1 - \tilde{P}_{i,\ell}(\mathbf{x}_t)\right]^{1-a_i}. \quad (53)$$

Within this subsection, let $d \in \{1, 2\}$ index the two directional demand markets associated with the non-directional entry market t . Under the type-level restriction used for fixed-cost recovery, for each configuration $\mathbf{a}$ and latent type ℓ we set

$\xi_{itd} = \tilde{\mu}_{i,\ell}$ and $\omega_{itd} = \tilde{\mu}_{i,\ell}^{\text{mc}}$ for every airline i . For each direction d , we then solve the Bertrand–Nash equilibrium of the nested logit demand system in equation (35). Let $\mathbf{x}_{itd}^D$ and $\mathbf{x}_{itd}^{\text{mc}}$ denote the demand and marginal-cost covariate vectors specified in equations (35) and (37), respectively. Airline i 's mean utility is $\delta_{i\ell td} = \alpha p_{i\ell td} + (\mathbf{x}_{itd}^D)' \boldsymbol{\beta} + \tilde{\mu}_{i,\ell}$, and its marginal cost is $mc_{i\ell td} = \exp((\mathbf{x}_{itd}^{\text{mc}})' \boldsymbol{\varphi} + \tilde{\mu}_{i,\ell}^{\text{mc}})$. Suppressing the dependence of equilibrium prices and shares on $\mathbf{a}$, equilibrium prices solve

$$p_{i\ell td} + \frac{1 - \sigma}{\alpha(1 - \sigma s_{i\ell td|g} - (1 - \sigma)s_{i\ell td})} = mc_{i\ell td}, \quad \forall i \text{ active in } \mathbf{a}, \quad d \in \{1, 2\},$$

which is equation (36) evaluated at the type- ℓ primitives. For each active airline j , configuration-level variable profit is

$$VP_{j\ell}(\mathbf{a}, \mathbf{x}_t) \equiv \sum_{d=1}^2 H_{td} s_{j\ell td} (p_{j\ell td} - mc_{j\ell td}),$$

where H_{td} is the size of directional demand market d . As an application-specific approximation, the “Others” group enters each direction’s nest inclusive value with its demand index evaluated at a zero demand unobservable and at its observed price, which is held fixed rather than re-optimized. Because entry is non-directional, we restrict the decomposition to route-quarters for which both directional demand markets are observed.

With $J = 6$ modeled airlines, equation (53) sums over $2^5 = 32$ rival configurations per focal airline. Across all six focal airlines, these are generated by the 63 nonempty configurations for each type and route direction. For a given market, we restrict the sum to profiles of its potential entrants; equivalently, every profile in which a nonpotential airline enters receives zero weight.

G.3 Estimation of σ_{η_j} and recovery of fixed costs

Given $VP_{j\ell}^P(\mathbf{x}_t)$ from the previous step, we estimate the airline-specific scale parameter σ_{η_j} by exploiting equation (52). Since $\Lambda^{-1}(\tilde{P}_{j,\ell}(\mathbf{x}_t)) = \mathbf{b}'_t \boldsymbol{\gamma}_{j,\ell}$ is the fitted first-step entry index, equation (52) can be written as

$$VP_{j\ell}^P(\mathbf{x}_t) = \sigma_{\eta_j} (\mathbf{b}'_t \boldsymbol{\gamma}_{j,\ell}) + fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}}). \quad (54)$$

We treat equation (54) as a regression of $VP_{j\ell}^P(\mathbf{x}_t)$ on the first-step entry index $\mathbf{b}_t' \gamma_{j,\ell}$, with the fixed-cost component $fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})$ as the unobserved regression component. The entry index is generally endogenous in this regression because it is correlated with $fc_{j\ell}(\mathbf{x}_{jt}^{\text{fc}})$. However, under the exclusion restriction that fixed costs depend only on j 's own characteristics $\mathbf{x}_{jt}^{\text{fc}}$, functions of the components of $\mathbf{x}_t$ that are excluded from $\mathbf{x}_{jt}^{\text{fc}}$ serve as instruments for $\mathbf{b}_t' \gamma_{j,\ell}$, provided they shift expected variable profits while remaining excluded from fixed costs.

In the application, we use as instruments, for each of airline j 's six competitors m (the five other modeled airlines and the Others group), competitor m 's share of non-stop activity at the market's endpoints: the sum, over the origin and destination airports, of m 's hub size at the airport divided by the combined hub size of all carriers at that airport, set to zero unless m is a potential entrant in the market. Of these six competitors, the five other modeled airlines enter expected variable profits through their type-specific propensity scores, whereas the Others group is not assigned a type-specific propensity score and enters through the inclusive-value approximation described above. The associated rival-presence variables remain excluded from airline j 's fixed costs. For each airline j , we then estimate σ_{η_j} by a separate 2SLS, pooling across latent types $\ell = 1, \dots, L$ over the route-quarters in which airline j is a potential entrant and both directional markets are observed, with latent-type-by-quarter fixed effects.

Once the scale parameters are estimated, we recover fixed costs as the residual from equation (52):

$$\widehat{fc}_{j\ell t} = \widehat{VP}_{j\ell}^P(\mathbf{x}_t) - \widehat{\sigma}_{\eta_j} \Lambda^{-1} \left(\widehat{\mathbf{P}}_{j,\ell}(\mathbf{x}_t) \right).$$

G.4 Variance decomposition

For airline j , reindex its T_j retained potential-entry markets as $t = 1, \dots, T_j$ and assign probability $1/T_j$ to each market. Independently, latent type ℓ has probabilities $\tilde{f}_1, \dots, \tilde{f}_L$. The resulting product measure is consistent with the independence of $\mathbf{x}_t$ and the latent market state in Assumption 1[d]. Each object $Y_{j\ell t} \in \{\Pi_{j\ell t}, VP_{j\ell t}, fc_{j\ell t}\}$ is a deterministic function of these two inputs.

Define the unconditional, market-specific, and type-specific means by

$$\bar{Y}_j = T_j^{-1} \sum_{t=1}^{T_j} \sum_{\ell=1}^L \tilde{f}_\ell Y_{j\ell t}, \quad m_j^Y(t) = \sum_{\ell=1}^L \tilde{f}_\ell Y_{j\ell t}, \quad n_j^Y(\ell) = T_j^{-1} \sum_{t=1}^{T_j} Y_{j\ell t}.$$

The interaction residual is

$$u_{j\ell t}^Y \equiv Y_{j\ell t} - m_j^Y(t) - n_j^Y(\ell) + \bar{Y}_j.$$

The variance components are

$$\begin{aligned} V_j^Y &= T_j^{-1} \sum_{t=1}^{T_j} \sum_{\ell=1}^L \tilde{f}_\ell (Y_{j\ell t} - \bar{Y}_j)^2, & V_{j,x}^Y &= T_j^{-1} \sum_{t=1}^{T_j} (m_j^Y(t) - \bar{Y}_j)^2, \\ V_{j,\ell}^Y &= \sum_{\ell=1}^L \tilde{f}_\ell (n_j^Y(\ell) - \bar{Y}_j)^2, & V_{j,\text{res}}^Y &= T_j^{-1} \sum_{t=1}^{T_j} \sum_{\ell=1}^L \tilde{f}_\ell (u_{j\ell t}^Y)^2. \end{aligned}$$

Orthogonality of the three ANOVA components under the product measure gives

$$V_j^Y = V_{j,x}^Y + V_{j,\ell}^Y + V_{j,\text{res}}^Y.$$

The components capture variation due to observables alone, latent types alone, and their interaction, respectively.

G.5 Measures of the contribution of unobservables to endogenous selection

For $Y \in \{\Pi, VP, fc\}$ with $V_j^Y > 0$, define

$$V_{j,\text{unobs}}^Y \equiv V_{j,\ell}^Y + V_{j,\text{res}}^Y.$$

The ratio $V_{j,\text{unobs}}^Y / V_j^Y$ is the total-effect share of latent market types: the share of variance associated with them either directly or through their interaction with observable market characteristics. Its complement is $V_{j,x}^Y / V_j^Y$.

To distinguish the channels through which latent types affect selection, we decompose $V_{j,\text{unobs}}^\Pi$ into the corresponding expected-variable-profit and fixed-cost components. Write $\bar{VP}_j$ and $\bar{fc}_j$ for $\bar{Y}_j$ evaluated at $Y = VP$ and $Y = fc$. The

pure-type and interaction covariance components are

$$\begin{aligned}\text{Cov}_{j,\ell}^{VP,fc} &\equiv \sum_{\ell=1}^L \tilde{f}_\ell (n_j^{VP}(\ell) - \overline{VP}_j) (n_j^{fc}(\ell) - \overline{fc}_j), \\ \text{Cov}_{j,\text{res}}^{VP,fc} &\equiv \frac{1}{T_j} \sum_{t=1}^{T_j} \sum_{\ell=1}^L \tilde{f}_\ell u_{j\ell t}^{VP} u_{j\ell t}^{fc}, \\ \text{Cov}_{j,\text{unobs}}^{VP,fc} &\equiv \text{Cov}_{j,\ell}^{VP,fc} + \text{Cov}_{j,\text{res}}^{VP,fc}.\end{aligned}$$

Because $\Pi_{j\ell t} = VP_{j\ell t} - fc_{j\ell t}$,

$$V_{j,\text{unobs}}^\Pi = V_{j,\text{unobs}}^{VP} + V_{j,\text{unobs}}^{fc} - 2 \text{Cov}_{j,\text{unobs}}^{VP,fc}. \quad (55)$$

We report the three normalized terms

$$\frac{V_{j,\text{unobs}}^{VP}}{V_j^\Pi}, \quad \frac{V_{j,\text{unobs}}^{fc}}{V_j^\Pi}, \quad \frac{-2 \text{Cov}_{j,\text{unobs}}^{VP,fc}}{V_j^\Pi}.$$

They are, respectively, the gross expected-variable-profit contribution, the gross fixed-cost contribution, and the covariance adjustment. The first two are gross contributions and therefore need not lie between zero and one. By equation (55), the three terms sum to $V_{j,\text{unobs}}^\Pi / V_j^\Pi$.

H Additional results

Table A.3: Pairwise Hausman Specification Tests

	2SLS	Heckman	Semipar.	Cont. κ	F.M., $L = 2$	F.M., $L = 3$
<i>Panel A: All demand parameters, $\chi^2(5)$</i>						
Heckman	189.80***					
Semipar.	165.49***	48.57***				
Cont. κ	83.12***	17.64***	33.41***			
F.M., $L = 2$	55.71***	166.48***	149.24***	86.01***		
F.M., $L = 3$	12.23**	194.87***	170.68***	81.73***	27.32***	
F.M., $L = 4$	3.09	139.95***	143.29***	82.92***	16.17***	2.73
<i>Panel B: Price and nesting parameters, $\chi^2(2)$</i>						
Heckman	101.06***					
Semipar.	97.29***	23.26***				
Cont. κ	64.81***	6.22**	12.90***			
F.M., $L = 2$	3.98	98.98***	97.13***	66.36***		
F.M., $L = 3$	6.74**	86.93***	88.50***	54.66***	4.80*	
F.M., $L = 4$	1.94	44.47***	59.21***	44.13***	1.32	0.51
<i>Panel C: Marginal cost parameters, $\chi^2(2)$</i>						
Heckman	7.56**					
Semipar.	9.24***	1.81				
Cont. κ	13.72***	7.62**	10.64***			
F.M., $L = 2$	1.17	5.74*	6.99**	11.25***		
F.M., $L = 3$	5.36*	2.72	3.67	5.36*	3.00	
F.M., $L = 4$	2.86	19.59***	22.34***	10.96***	1.82	0.07

Notes: The table reports the pairwise Hausman-type test $\mathcal{H}(A, B)$ from equation (34) applied, within each panel, to every pair among the baseline 2SLS and the six selection-corrected estimators of Tables 5 and A.5; the OLS estimator is not included in these tests. Each cell gives the χ^2 statistic for the null that the row and column estimators recover the same parameter vector, with $\widehat{V}_{AB}$ computed using the linearized common-weight bootstrap of Online Appendix F with 200 replications. Panel A reports the joint tests on the demand parameters $(\alpha, \sigma, \beta_{\text{dist}}, \beta_{\text{dist}^2}, \beta_{\text{pres}})$, Panel B restricts the tests to the price and nesting parameters (α, σ) , and Panel C reports the joint tests on the marginal-cost parameters $(\varphi_{\text{dist}}, \varphi_{\text{dist}^2})$. Stars denote rejection of the null of equality: * at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A.4: Average Marginal Costs

	<i>Not control. for sel.</i>		<i>Controlling for endogenous selection</i>					
	OLS	2SLS	2SLS Heckman $L = 1$	2SLS Semipar. $L = 1$	2SLS Cont. κ $L = 1$	2SLS Fin.-Mix. $L = 2$	2SLS Fin.-Mix. $L = 3$	2SLS Fin.-Mix. $L = 4$
<i>Marginal Cost (100\$)</i>	−0.002	1.180	1.455	1.485	1.653	1.201	1.297	1.338
AA	0.265	1.348	1.596	1.622	1.791	1.369	1.459	1.501
DL	0.207	1.477	1.776	1.809	1.978	1.498	1.599	1.641
UA	0.448	1.571	1.830	1.857	2.025	1.592	1.684	1.725
US	0.081	1.254	1.526	1.556	1.723	1.275	1.370	1.411
WN	−0.459	0.877	1.193	1.228	1.395	0.897	1.002	1.043
LCC	−0.258	0.801	1.043	1.068	1.237	0.823	0.911	0.952
<i>Others</i>	−0.004	0.840	1.024	1.040	1.209	0.861	0.937	0.977
Airline $\times$ Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Airport FE	Y	Y	Y	Y	Y	Y	Y	Y
# control var. entry	0	0	6	18	42	6	12	18
Observations	37,184	37,184	37,184	37,184	37,184	37,184	37,184	37,184

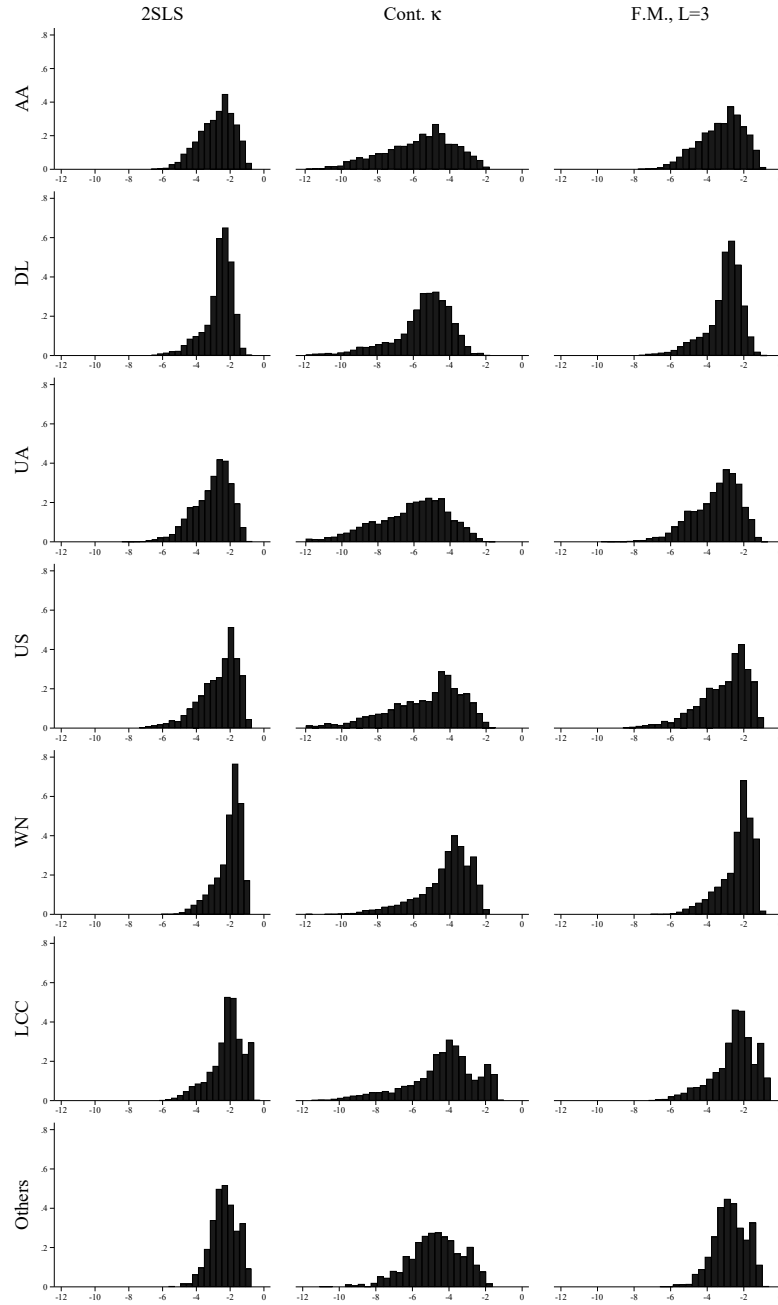
Notes: The table reports average marginal costs implied by the pricing equation (36), evaluated at the observed prices and market shares using the corresponding column's demand parameter estimates from Table 5. Marginal costs are first computed at the airline-market level and are then averaged overall and separately by airline.

Table A.5: Estimation of Marginal Cost Parameters

	<i>Not control. for sel.</i>		<i>Controlling for endogenous selection</i>					
	OLS	2SLS	2SLS Heckman $L = 1$	2SLS Semipar. $L = 1$	2SLS Cont. κ $L = 1$	2SLS Fin.-Mix. $L = 2$	2SLS Fin.-Mix. $L = 3$	2SLS Fin.-Mix. $L = 4$
Distance (1000mi)	1.368 (0.093)	1.390 (0.202)	0.904 (0.161)	0.864 (0.130)	0.764 (0.091)	1.325 (0.203)	1.143 (0.205)	1.103 (0.149)
Distance ²	−0.281 (0.030)	−0.318 (0.058)	−0.164 (0.042)	−0.150 (0.033)	−0.134 (0.023)	−0.296 (0.059)	−0.242 (0.058)	−0.231 (0.039)
Airline $\times$ Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Airport FE	Y	Y	Y	Y	Y	Y	Y	Y
# control var. entry	0	0	6	18	42	6	12	18
Observations	19,441	36,390	37,176	37,177	37,184	36,438	36,731	36,809
BIC _{2nd}		−61819.4				−63535.7	−75253.9	−78389.5
Hausman $\chi^2(2)$ vs 2SLS			7.56	9.24	13.72	1.17	5.36	2.86
p-value			[0.023]	[0.010]	[0.001]	[0.558]	[0.069]	[0.240]

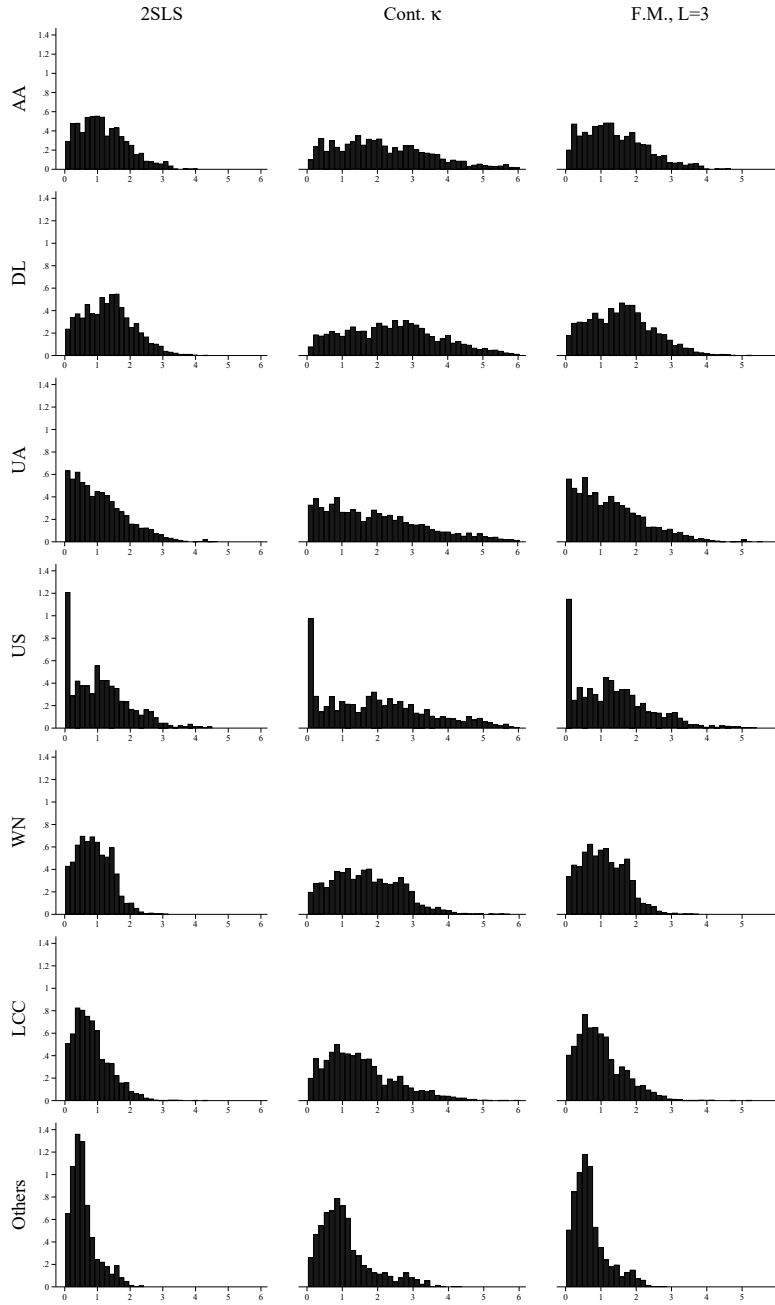
Notes: The table reports estimates of the marginal-cost parameters in equation (37). The dependent variable $\ln(mc_{jt})$ is constructed from the implied marginal costs obtained from equation (36) using the corresponding column's demand parameter estimates from Table 5. All columns are estimated by OLS. Standard errors are computed using the linearized common-weight bootstrap of Online Appendix F with 200 replications. The BIC_{2nd} row reports the analogous second-step BIC-style criterion for the log-marginal-cost regression, in the same spirit as equation (33). The last two rows report the Hausman-type specification test $\mathcal{H}(2SLS, \cdot)$ from equation (34), applied here to the two marginal-cost parameters $(\varphi_{\text{dist}}, \varphi_{\text{dist}^2})$ ($\chi^2(2)$). The number of observations varies across columns because observations with a non-positive implied marginal cost from equation (36) are excluded.

Figure A.1: Distribution of Estimated Own-Price Elasticities



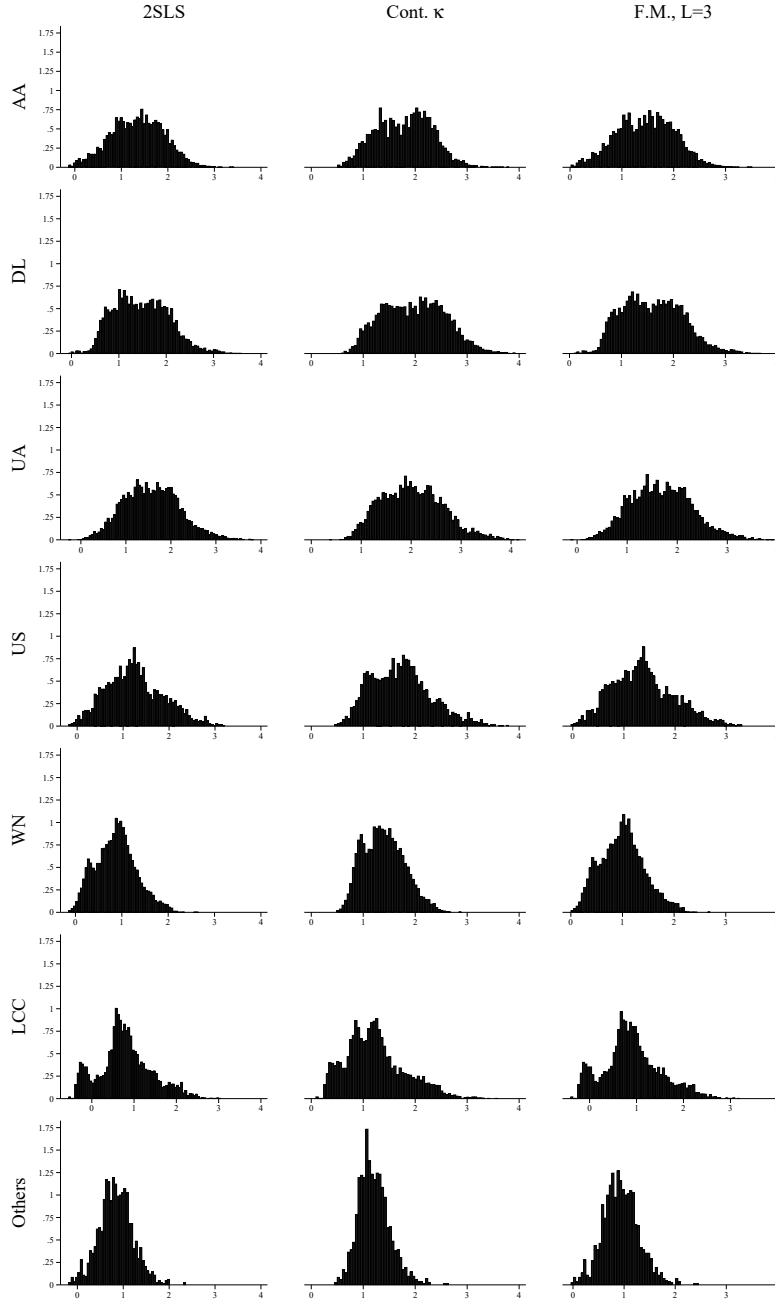
Notes: The figure reports the empirical distributions of estimated own-price elasticities at the airline-market level. Each row corresponds to the price-setting airline, while the three columns correspond respectively to the 2SLS estimator without selection correction, the continuous- κ_t correction with $L = 1$, and the finite-mixture correction with $L = 3$. Within each column, elasticities are evaluated using the corresponding demand parameter estimates $(\hat{\alpha}, \hat{\sigma})$ at the observed prices and market shares.

Figure A.2: Distribution of Estimated Cross-Price Elasticities



Notes: The figure reports the empirical distributions of estimated cross-price elasticities at the airline-market level. Each row corresponds to the price-setting airline, while the three columns correspond respectively to the 2SLS estimator without selection correction, the continuous- κ_t correction with $L = 1$, and the finite-mixture correction with $L = 3$. Within each column, elasticities are evaluated using the corresponding demand parameter estimates $(\hat{\alpha}, \hat{\sigma})$ at the observed prices and market shares.

Figure A.3: Distribution of Estimated Marginal Costs



Notes: The figure reports the empirical distributions of estimated marginal costs at the airline-market level. Each row corresponds to an airline, while the three columns correspond respectively to the 2SLS estimator without selection correction, the continuous- κ_t correction with $L = 1$, and the finite-mixture correction with $L = 3$. Within each column, marginal costs are computed from equation (36) using the corresponding demand parameter estimates and the observed prices and market shares.