Product Entry in the Global Automobile Industry

December 3, 2025

Alejandro Sabal

A Data Cleaning and Imputation of Product Characteristics

In this section, I explain the procedure to obtain the estimation sample from the raw data. I first describe preliminary data cleaning procedures. Then, I describe in detail how I aggregate and impute product characteristics to all potential products $\mathcal{A}$.

A.1 Preliminary Cleaning

The *IHS Markit* 2019 sample contains information on the universe of passenger vehicle registrations for 12 countries: Australia, Brazil, France, Germany, India, Italy, Japan, Mainland China, Mexico, Spain, the United Kingdom, and the United States. The data are at the quarterly-trim-country level.

I define 3 fuel type categories using information on fuel type in the dataset. First, I define "Electric" vehicles as any vehicle that fits in a plug-in category. For instance, I observe Plug-In Electric or Plug-in (Petrol or Diesel) vehicles. Any such fuel type will be considered part of the"Electric" category. Second, I define a "Hybrid" vehicle as any non-plug-in vehicle with a fuel type in the hybrid category. Finally, I define an internal combustion engine vehicle ("ICE") as any vehicle that is neither hybrid nor plug-in, be it Petrol or Diesel.

The dataset does not contain the vast majority of large pickups or vans, so I drop the small number of instances of such items from my sample. I use the SUV and Wagon categories to define the respective body types. I define the Convertible body type as including convertibles, retractable hardtops, and roadsters. Finally, I define the Car category as including either sedans, hatchbacks, or coupes. This gives 5 different possible body types.

In the sample, I observe some products in extreme price ranges. Since the goal of this paper is, in part, to estimate product development fixed costs, I drop the most luxurious products, which are likely designed and produced with very different technologies than the typical passenger vehicle. I drop all products with observed prices higher than $150,000 or

pertaining to the following brands: Lamborghini, McLaren, Bentley, Ferrari, Aston Martin, Maserati, Bugatti, and Rolls-Royce.

A.2 Demand Estimation Sample

I aggregate the data at the product-year level, where the product definition is a brand-body type-fuel type combination. I obtain total units sold by summing across all quarters and trims corresponding to my product definition. I aggregate the remaining characteristics (including horsepower, weight, length, width, and height) by taking a quantity-weighted average of such characteristics across all quarter-trim observations within my product definition. I merge this dataset with information from CEPII containing the distance between any two countries (population-weighted), which gives the distance between the headquarters country of the brand and the destination country. Market size is the product of the number of households and the average number of vehicles per household in market m, divided by the average tenure of car ownership, which I assume is 5 years, as in Grieco et al. (2023). The number of households in each country in 2019 and the average number of vehicles per household are obtained from the World Population Review and publicly available government sources (e.g., Department of Transportation).

I obtain market shares for each product in each market by dividing the total units sold by market size. As typical in the automobile literature that estimates mixed logit demand models, I drop very small market shares. More precisely, I drop all products with shares smaller than 0.00001 within a market. I only do this for demand and marginal cost estimation. In the estimation of fixed costs, I treat these products as part of the set of products in firms' global portfolios as well as in the set of product offerings in such markets.

A.3 Fixed Cost Estimation Sample and Product Characteristic Imputation

For the fixed cost estimation sample, I require a dataset at the (potential) product-market level that contains the observed characteristics of each product-market pair. This is what is required to use the demand and marginal cost model to predict variable profits under any arbitrary market structure for each possible entry opportunity. Given that I defined the set of potential products as the set of all brand-body type-fuel type combinations, I need to impute all observed characteristics that enter the demand and marginal cost specification for all potential products. These characteristics include horsepower, horsepower/weight, and size (length × width). To do this, I use the characteristics of the observed products.

I impute size for all products not observed in the *IHS Markit* sample according to the following sequential procedure: (1) using the mean size of observed products of the same body type sold by the same brand; (2) if there are no such products, I use the mean size of observed products of the same body type sold by the same parent company; (3) if there are no such products, I use the mean size across observed products of said body type.

I impute horsepower (and horsepower/weight) for all products not observed in the *IHS Markit* sample according to the following procedure: (1) using the mean horsepower (horsepower/weight) of observed products of the same fuel type and body type sold by the same parent company; (2) if there are no such products, I use the mean horsepower (horsepower/weight) of observed products with the same body type and fuel type offered in that country; (3) if there are no such products, I use the mean horsepower (horsepower/weight) of all products with the same fuel type and body type.

Finally, I treat a product as offered in a market ($O_{jm} = 1$) if and only if at least one unit is sold in that market. This is standard in previous papers studying product entry. I define a product as being in a firm's product portfolio ($G_j = 1$) if and only if it is observed to be sold in at least one market in my sample. This definition assumes that there are no products that are only offered in markets that are not included in my sample. This is a decent assumption, given that the markets in my sample account for most of the global demand. A caveat is that, in my model, a firm could add a product to its portfolio but face high market entry costs everywhere, leading it not to sell the product. Misclassifying such cases as $G_j = 0$ when $G_j = 1$ would understate the share of products introduced and likely overstate portfolio fixed costs. This bias should be minimal, since with large portfolio costs and independent entry-cost shocks, it is unlikely a firm would develop a product and then sell it nowhere.

B Proof of Theorem 1 (Moment Inequalities)

In this section, I derive the key results that I use for the estimation of $(\theta_e, \sigma_e, \theta_g, \sigma_g)$. In Section B.1, I derive moment inequalities that bound (θ_e, σ_e). In Section B.2, I derive moment inequalities that bound $(\theta_e, \sigma_e, \theta_g, \sigma_g)$. Recall from the main text that Γ_{jm} denotes the CDF of F_{jm}^e. I let Λ_j denote the CDF of F_j^g. I introduce notation and write, $\mathbb{1}_{\mathcal{C}^f} = \mathbb{1}\{\mathcal{C}^f \text{ is chosen by firm } f\}$, where $\mathcal{C}^f$ is a bundle of products chosen by firm f.

B.1 Stage 2: Market Entry Inequalities

I start by proving formally the upper bound inequality (9). At this stage, firm f has already chosen portfolio $\mathcal{G}^f$. By revealed preference and best response, for any $j \in \mathcal{G}^f$,

$$
\begin{aligned}
(\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}}) \, \mathbb{1}\{\mathbb{E}[\Delta_{jm}(\Omega_m^f, \Omega_m^{-f}) \mid \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} \\
= (\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}}) \, \mathbb{1}_{\Omega_m^f}.
\end{aligned}
\tag{15}
$$

Equation (15) says that conditional on firm f choosing Ω_m^f or $\Omega_m^f \setminus \{j\}$, the firm chooses Ω_m^f if and only if it is weakly preferred to $\Omega_m^f \setminus \{j\}$.

Under Assumption 4, the largest possible change in expected profits from introducing product j in market m can be obtained by offering no product other than j in market m. Thus, I obtain inequality,

$$
\begin{aligned}
(\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}}) \mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f}) | \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} \\
\geq (\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}}) \mathbb{1}_{\Omega_m^f}.
\end{aligned}
\tag{16}
$$

In inequality (16), inside the indicator function, the expected marginal value is evaluated at a set of entry decisions that is not optimal and is therefore independent of $\{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}$ conditional on $\mathcal{I}$. Also, due to Assumption 1, rivals' product offerings Ω_m^{-f} are independent of $\{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}$ conditional on $\mathcal{I}$. Thus, removing conditioning on the shocks,

$$
\begin{aligned}
(\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}}) \mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f}) | \mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} \\
\geq (\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}}) \mathbb{1}_{\Omega_m^f}.
\end{aligned}
\tag{17}
$$
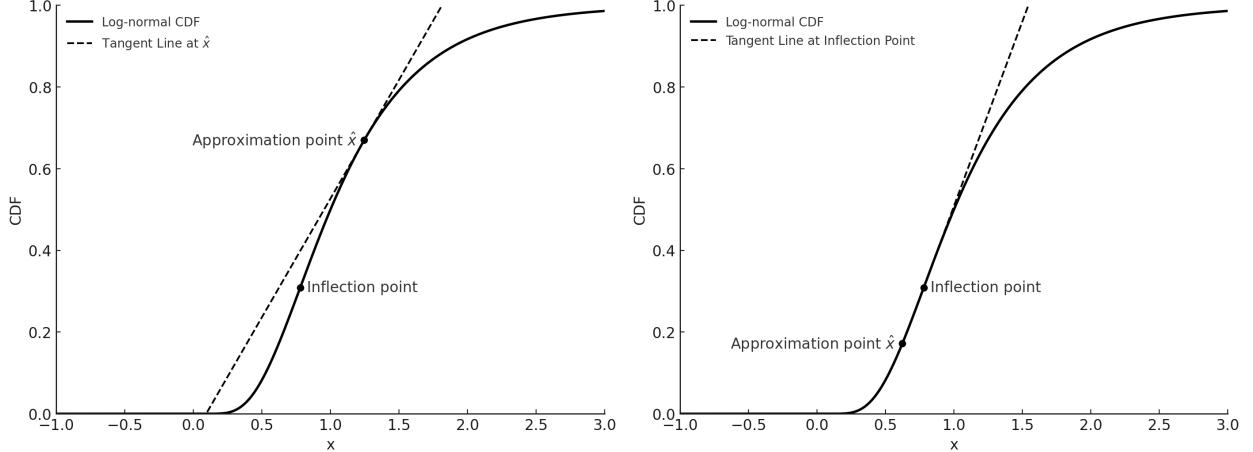
Note that inequality (17) holds for all bundles Ω_m^f containing product j, and that the term $\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f}) | \mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\}$ is now independent of both Ω_m^f and $\Omega_m^f \setminus \{j\}$. Therefore, realizing that $O_{jm} = \sum_{\Omega_m^f : j \in \Omega_m^f} \mathbb{1}_{\Omega_m^f}$ and $1 - O_{jm} = \sum_{\Omega_m^f : j \in \Omega_m^f} \mathbb{1}_{\Omega_m^f \setminus \{j\}}$, I sum inequality (17) across all such mutually exclusive bundles to obtain,

$$
\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f}) | \mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} \geq O_{jm}.
\tag{18}
$$

That is, if product j is offered in market m, the largest possible change in profits from offering product j in market m must be weakly positive. The term $\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f}) | \mathcal{I}, \mathcal{G}^f]$ is independent of ν_{jm}^e, so I take expectations conditional on $\mathcal{I}$ and $\mathcal{G}^f$ to derive,

$$
\Gamma_{jm}\big(\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f}) | \mathcal{I}, \mathcal{G}^f]; \theta_e, \sigma_e\big) \geq \mathbb{E}[O_{jm} | \mathcal{I}, \mathcal{G}^f],
\tag{19}
$$

Figure B.1: Convex Upper Bounds of a Log-Normal CDF

where recall that Γ_{jm} denotes the cumulative distribution function (CDF) of F^e_{jm} given θ_e and σ_e. Inequality (19) gives an upper bound on the probability that product j is offered in market m, conditional on $\mathcal{I}$ and $\mathcal{G}^f$. The key insight for estimation is to now use a convex upper bound of CDF Γ_{jm}. If F^e_{jm} is log-normal, it has inflection point $\check{x}_{jm}(\theta_e, \sigma_e) = \exp(Z'_{jm}\theta_e - \sigma_e^2)$, so one can define,

$$\overline{\Gamma}^1_{jm}(x; \theta_e, \sigma_e) := \Gamma_{jm}(x; \theta_e, \sigma_e)\mathbb{1}\{x < \check{x}_{jm}(\theta_e, \sigma_e)\}$$
$$+ [\Gamma_{jm}(\check{x}; \theta_e, \sigma_e) + \gamma_{jm}(\check{x}_{jm}; \theta_e, \sigma_e)(x - \check{x}_{jm}(\theta_e, \sigma_e))]\mathbb{1}\{x \geq \check{x}_{jm}(\theta_e, \sigma_e)\} \tag{20}$$

where γ_{jm} denotes the probability density function (PDF) of the corresponding log-normal distribution. In the empirical implementation, I use,

$$\overline{\Gamma}_{jm}(x, \hat{x}; \theta_e, \sigma_e) := \overline{\Gamma}^1_{jm}(x; \theta_e, \sigma_e)\mathbb{1}\{\hat{x} < \check{x}_{jm}(\theta_e, \sigma_e)\}+$$
$$\max\{\Gamma_{jm}(\hat{x}; \theta_e, \sigma_e) + \gamma_{jm}(\hat{x}; \theta_e, \sigma_e)(x - \hat{x}), \Gamma_{jm}(x; \theta_e, \sigma_e)\}\mathbb{1}\{\hat{x} \geq \check{x}_{jm}(\theta_e, \sigma_e)\}, \tag{21}$$

where $\hat{x}$ is a $\mathcal{I}$-measurable approximation point.

That is, if the $\mathcal{I}$-measurable approximation point lies below the inflection point, I use the convex upper bound given by equation (20). Otherwise, I use a linear approximation at the approximation point (which is on the concave part of the CDF), bounded below by the CDF, Γ_{jm}. Figure B.1 illustrates the convex upper bounds.[1] Given convex upper bound $\overline{\Gamma}_{jm}(x, \hat{x}_{jm}; \theta_e, \sigma_e)$, I derive,

$$\overline{\Gamma}_{jm}\big(\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}, \mathcal{G}^f], \hat{x}_{jm}; \theta_e, \sigma_e\big) \geq \mathbb{E}[O_{jm}|\mathcal{I}, \mathcal{G}^f].$$

---

[1]Note that these are global bounds and, unlike the CDF, take values larger than 1.

I now apply Jensen's inequality and obtain,

$$\mathbb{E}[\overline{\Gamma}_{jm}\big(\Delta_{jm}(\{j\}, \Omega_m^{-f}), \hat{x}_{jm}; \theta_e, \sigma_e\big) - O_{jm}|\mathcal{I}, \mathcal{G}^f] \geq 0 \tag{22}$$

Inequality (22) is a conditional moment inequality that can be used for estimation. Given this conditional moment inequality, unconditional moment inequalities can be derived using positive functions of $\mathcal{I}$. Moreover, provided the approximation points $\hat{x}_{jm}$ are $\mathcal{I}$-measurable, one can use observation-specific upper bound functions $\overline{\Gamma}_{jm}$.[2]

To derive the lower bound inequality, I start in a manner similar to equation (15), by writing, for any bundle Ω_m^f such that $j \notin \Omega_m^f$,

$$(\mathbb{1}_{\Omega_m^f \cup \{j\}} + \mathbb{1}_{\Omega_m^f})\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\Omega_m^f \cup \{j\}, \Omega_m^{-f}) - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \leq 0\}$$
$$= (\mathbb{1}_{\Omega_m^f \cup \{j\}} + \mathbb{1}_{\Omega_m^f})\mathbb{1}_{\Omega_m^f}. \tag{23}$$

Equation (23) says that conditional on firm f choosing either $\Omega_m^f \cup \{j\}$ or Ω_m^f, it will choose Ω_m^f if and only if it is preferred to $\Omega_m^f \cup \{j\}$.

Due to submodularity, the lowest possible expected change in profits from offering product j in market m is obtained whenever the firm is offering all products in its portfolio $\mathcal{G}^f$ in market m, which implies that,

$$(\mathbb{1}_{\Omega_m^f \cup \{j\}} + \mathbb{1}_{\Omega_m^f})\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f})|\{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \leq 0\}$$
$$\geq (\mathbb{1}_{\Omega_m^f \cup \{j\}} + \mathbb{1}_{\Omega_m^f})\mathbb{1}_{\Omega_m^f}. \tag{24}$$

As with the upper bound inequality, I remove conditioning on $\{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}$ in the expectation in inequality (24) due to Assumption 1. Then, realizing that $\sum_{\Omega_m^f: j \notin \Omega_m^f} \mathbb{1}_{\Omega_m^f} = 1 - O_{jm}$ and $\sum_{\Omega_m^f: j \notin \Omega_m^f} \mathbb{1}_{\Omega_m^f \cup \{j\}} = O_{jm}$, I sum inequality (24) over all bundles Ω_m^f with $j \notin \Omega_m^f$ to obtain,

$$\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f})|\mathcal{I}, \mathcal{G}^f] - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \leq 0\} \geq 1 - O_{jm}. \tag{25}$$

Taking expectations conditional on $\mathcal{I}$ and $\mathcal{G}^f$ on both sides of inequality (25) yields,
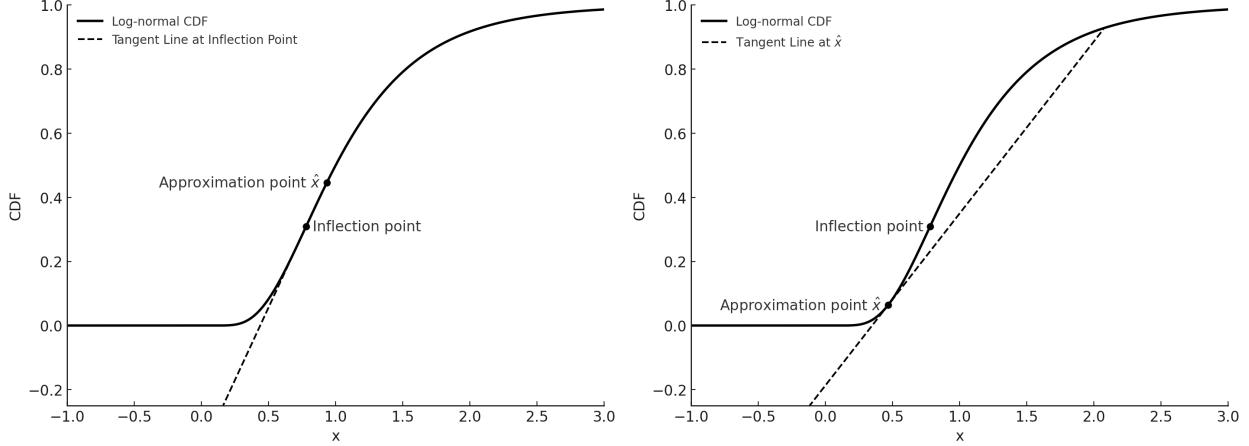
$$\Gamma_{jm}\big(\mathbb{E}[\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f})|\mathcal{I}, \mathcal{G}^f]; \theta_e, \sigma_e\big) \leq \mathbb{E}[O_{jm}|\mathcal{I}, \mathcal{G}^f]. \tag{26}$$

Inequality (26) provides a lower bound on the probability that product j is offered in market m, conditional on $\mathcal{I}$ and $\mathcal{G}^f$. I now follow a very similar logic as with the upper

---

[2]Porcher et al. (2024) also employ observation-specific linear approximations to derive moment inequalities, though in a single-agent setting and using odds-based inequalities rather than bounding choice probabilities using convex upper and concave lower bounds of the CDF of the unobserved shock.

Figure B.2: Concave Lower Bounds of a Log-Normal CDF

bound and use a *concave lower bound* for Γ_{jm} before applying Jensen's inequality. A family of such concave lower bounds is given by,

$$
\begin{aligned}
\underline{\Gamma}^1_{jm}(x; \theta_e, \sigma_e) := {} & \Gamma(x; \theta_e, \sigma_e) \mathbb{1}\{x \geq \check{x}_{jm}(\theta_e, \sigma_e)\} \\
& + [\Gamma_{jm}(\check{x}; \theta_e, \sigma_e) + \gamma_{jm}(\check{x}_{jm}; \theta_e, \sigma_e)(x - \check{x}_{jm}(\theta_e, \sigma_e))]\mathbb{1}\{x < \check{x}_{jm}(\theta_e, \sigma_e)\}
\end{aligned}
\tag{27}
$$

where recall that the inflection point is given by, $\check{x}_{jm}(\theta_e, \sigma_e) = \exp(Z'_{jm}\theta_e - \sigma_e^2)$. In the empirical implementation, I use,

$$
\begin{aligned}
\underline{\Gamma}_{jm}(x, \hat{x}; \theta_e, \sigma_e) := {} & \underline{\Gamma}^1_{jm}(x; \theta_e, \sigma_e) \mathbb{1}\{\hat{x} \geq \check{x}_{jm}(\theta_e, \sigma_e)\} \\
& + \min\{\Gamma_{jm}(\hat{x}; \theta_e, \sigma_e) + \gamma_{jm}(\hat{x}; \theta_e, \sigma_e)(x - \hat{x}), \Gamma_{jm}(x; \theta_e, \sigma_e)\}\mathbb{1}\{\hat{x} < \check{x}_{jm}(\theta_e, \sigma_e)\}.
\end{aligned}
\tag{28}
$$

That is, if the approximation point lies above the inflection point, I use the concave lower bound given in equation (27). Otherwise, I use a linear approximation at the approximation point (which is in the convex part of the CDF), bounded above by the CDF, Γ_{jm}. Figure B.2 illustrates the concave lower bound functions.

Given such a concave lower bound, together with Jensen's inequality, I obtain

$$
\mathbb{E}[\underline{\Gamma}_{jm}(\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f}), \hat{x}_{jm}; \theta_e, \sigma_e) - O_{jm}|\mathcal{I}, \mathcal{G}^f] \leq 0.
\tag{29}
$$

As with the upper bound, conditional moment inequality (29) can now be used for estimation given positive-valued functions of $\mathcal{I}$, which yield unconditional moment inequalities. Also, as before, provided $\hat{x}_{jm}$ is $\mathcal{I}$-measurable, one can use observation-dependent lower bound functions $\underline{\Gamma}_{jm}$.

7

Note that because by assumption, market entry fixed cost shocks $\{\nu_{jm}^e\}_{j\in\mathcal{G}^f,m}$ are unobserved at the time of choosing $\mathcal{G}^f$, I condition on the observed product portfolios when using conditional moment inequalities (22) and (29) for estimation.

This concludes the derivation of the inequalities used to estimate bounds on (θ_e, σ_e). In the next subsection, I derive inequalities that provide bounds on (θ_g, σ_g).

B.2 Stage 1: Global Portfolio Choice Inequalities

At this stage, firms choose their global product portfolios $\mathcal{G}^f$. The derivation of the inequalities at this stage follows a similar logic as the derivations in Stage 2. For the derivations that follow, I will use the following result:

Proposition 1 *The function $g: \mathbb{R} \to \mathbb{R}$ given by,*

$$g(y) = \mathbb{E}_X[\mathbb{1}(X \leq y)(y - X)] = \mathbb{E}_X[y - X | X \leq y]\mathbb{P}_X(X \leq y)$$

is convex in y for any continuous random variable X provided $F_X(y) := \mathbb{P}_X(X \leq y) > 0$ i.e., the conditional expectation is well defined.

Proof. Differentiating with respect to y, I obtain,

$$g'(y) = F_X(y)\left[1 - \frac{f_X(y)}{F_X(y)}(y - \mathbb{E}[X|X \leq y])\right] + \mathbb{E}[y - X|X \leq y]f_X(y)$$

$$= F_X(y)$$

where f_X denotes the density of random variable X. But then, clearly, $g''(y) = f(y) > 0$, which proves convexity. ∎

To derive the inequalities at this stage, I first define the value of a given portfolio for a firm f as,

$$\mathcal{V}_f(\mathcal{G}^f, \{\nu_j^g\}_{j\in\mathcal{A}^f}, \mathcal{I}) = \mathbb{E}\left[\sum_{m\in\mathcal{M}} \max_{\{\Omega_m^f\subseteq\mathcal{G}^f\}} \Pi_m^f(\Omega_m^f; \mathcal{G}^f, \{\nu_{jm}^e\}_{j\in\mathcal{G}^f,m}, \mathcal{I})\Big|\mathcal{I}\right]$$
$$- \sum_{j\in\mathcal{A}^f} G_j F_j^g(\nu_j^g; \theta_g, \sigma_g), \tag{30}$$

where

$$\Pi_m^f(\Omega_m^f; \mathcal{G}^f, \{\nu_{jm}^e\}_{j\in\mathcal{G}^f,m}, \mathcal{I})$$

$$= \mathbb{E} \left[\sum_{j \in \Omega_m^f} [\pi_{jm}^*(\Omega_m^f, \Omega_m^{-f}) - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e)] \middle| \mathcal{G}^f, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{I} \right].$$

Given any chosen product portfolio $\mathcal{G}^f$ and realized global portfolio fixed cost shocks, firm f computes its Stage 1 value as the expected maximal profits it will obtain once it realizes its market entry fixed cost shocks for each product and chooses offerings in each market optimally. The expectation in equation (30) is only conditional on $\mathcal{I}$, so the firm must also integrate over its own market entry fixed costs shocks. At Stage 2, firm f's optimal entry decisions depend on the portfolio it chooses in Stage 1 and the market entry fixed cost shocks it realizes in Stage 2. I write the Stage-2 decision rule in each market as $\Omega_m^f(\mathcal{G}^f, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f}, \mathcal{I})$.

Henceforth, I abuse notation and let $\Omega_m^{\mathcal{G}^f}$ denote the decision rule under portfolio $\mathcal{G}^f$ (at given market entry fixed cost shocks and at $\mathcal{I}$) and $O_{jm}^{\mathcal{G}^f}$ denote the optimal entry decision rule for product j in market m under $\mathcal{G}^f$ corresponding to $\Omega_m^{\mathcal{G}^f}$.

I start by deriving an upper bound inequality. First, note that,

$$(\mathbb{1}_{\mathcal{G}^f} + \mathbb{1}_{\mathcal{G}^f \setminus \{j\}}) \mathbb{1} \left\{ \mathcal{V}_f(\mathcal{G}^f, \{\nu_j^g\}_{j \in \mathcal{A}^f}, \mathcal{I}) - \mathcal{V}_f(\mathcal{G}^f \setminus \{j\}, \{\nu_j^g\}_{j \in \mathcal{A}^f}, \mathcal{I}) \geq 0 \right\}$$
$$= (\mathbb{1}_{\mathcal{G}^f} + \mathbb{1}_{\mathcal{G}^f \setminus \{j\}}) \mathbb{1}_{\mathcal{G}^f}. \tag{31}$$

That is, conditional on choosing either portfolio $\mathcal{G}^f$ or portfolio $\mathcal{G}^f \setminus \{j\}$, the firm chooses $\mathcal{G}^f$ if and only if it is preferred to $\mathcal{G}^f \setminus \{j\}$. A lower bound of $\mathcal{V}_f(\mathcal{G}^f \setminus \{j\}, \{\nu_j^g\}_{j \in \mathcal{A}^f}, \mathcal{I})$ can be obtained by using the fact that the entry decisions in all markets for products $j \in \mathcal{G}^f \setminus \{j\}$ must weakly dominate the optimal-under-$\mathcal{G}^f$ entry decision rules for such products. This means that the equality (31) implies the following inequality,

$$(\mathbb{1}_{\mathcal{G}^f} + \mathbb{1}_{\mathcal{G}^f \setminus \{j\}}) \mathbb{1} \left\{ \sum_{m \in \mathcal{M}} \left(\mathbb{E}[\Pi_m^f(\Omega_m^{\mathcal{G}^f}) - \Pi_m^f(\Omega_m^{\mathcal{G}^f} \setminus \{j\}) | \mathcal{I}] \right) - F_j^g(\nu_j^g; \theta_g, \sigma_g) \geq 0 \right\}$$
$$\geq (\mathbb{1}_{\mathcal{G}^f} + \mathbb{1}_{\mathcal{G}^f \setminus \{j\}}) \mathbb{1}_{\mathcal{G}^f}. \tag{32}$$

Equation (32) says that the change in value from including j into the portfolio starting from a suboptimal entry decision rule in the second stage of the game – which mandates choosing the *same* entry decisions in each market under $\mathcal{G}^f \setminus \{j\}$ as under $\mathcal{G}^f$ for all products $j' \neq j$ – must be higher than the actual change in expected value from introducing product j into the firm's portfolio at the best response. I re-write inequality (32) as,

$$(\mathbb{1}_{\mathcal{G}^f} + \mathbb{1}_{\mathcal{G}^f \setminus \{j\}}) \times$$

$$\mathbb{1}\left\{\sum_{m\in\mathcal{M}}\left(\mathbb{E}\big[O_{jm}^{\mathcal{G}^f}\big[\mathbb{E}\big[\Delta_{jm}(\Omega_m^{\mathcal{G}^f}\setminus\{j\}),\Omega_m^{-f})|\mathcal{I},\{\nu_{jm}^e\}_{j\in\mathcal{G}^f,m},\mathcal{G}^f\big]-F_{jm}^e(\nu_{jm}^e;\theta_e,\sigma_e)|\mathcal{I}\big]\right)\right.$$

$$\left.-F_j^g(\nu_j^g;\theta_g,\sigma_g)\geq 0\right\}\geq(\mathbb{1}_{\mathcal{G}^f}+\mathbb{1}_{\mathcal{G}^f\setminus\{j\}})\mathbb{1}_{\mathcal{G}^f}. \tag{33}$$

Inequality (33) shows that if product j is introduced in firm f's global portfolio, then the change in profits from offering it in each market in which firm f chooses to offer product j (evaluated at the optimal entry decisions under $\mathcal{G}^f$) is weakly positive net of the portfolio fixed cost. Under Assumption 4, I bound the expression inside the indicator function in inequality (33) by above and make it independent of the optimal portfolio choice. I obtain,

$$(\mathbb{1}_{\mathcal{G}^f}+\mathbb{1}_{\mathcal{G}^f\setminus\{j\}})\times$$

$$\mathbb{1}\left\{\sum_{m\in\mathcal{M}}\left(\mathbb{E}\big[O_{jm}^{\mathcal{G}^f}\big[\mathbb{E}\big[\Delta_{jm}(\{j\},\Omega_m^{-f})|\mathcal{I},\{\nu_{jm}^e\}_{j\in\mathcal{G}^f,m},\mathcal{G}^f\big]-F_{jm}^e(\nu_{jm}^e;\theta_e,\sigma_e)|\mathcal{I}\big]\right)\right.$$

$$\left.-F_j^g(\nu_j^g;\theta_g,\sigma_g)\geq 0\right\}\geq(\mathbb{1}_{\mathcal{G}^f}+\mathbb{1}_{\mathcal{G}^f\setminus\{j\}})\mathbb{1}_{\mathcal{G}^f}. \tag{34}$$

Therefore,

$$(\mathbb{1}_{\mathcal{G}^f}+\mathbb{1}_{\mathcal{G}^f\setminus\{j\}})\mathbb{1}\left\{\sum_{m\in\mathcal{M}}\left(\mathbb{E}\big[O_{jm}^{\{j\}}\big[\mathbb{E}\big[\Delta_{jm}(\{j\},\Omega_m^{-f})|\mathcal{I}\big]-F_{jm}^e(\nu_{jm}^e;\theta_e,\sigma_e)|\mathcal{I}\big]\right)\right.$$

$$\left.-F_j^g(\nu_j^g;\theta_g,\sigma_g)\geq 0\right\}\geq(\mathbb{1}_{\mathcal{G}^f}+\mathbb{1}_{\mathcal{G}^f\setminus\{j\}})\mathbb{1}_{\mathcal{G}^f}, \tag{35}$$

where

$$O_{jm}^{\{j\}}=\mathbb{1}\big\{\mathbb{E}[\Delta_{jm}(\{j\},\Omega_m^{-f})|\mathcal{I}]-F_{jm}^e(\nu_{jm}^e;\theta_e,\sigma_e)\geq 0\big\}.$$

It follows that the contribution of product j into firm f's Stage 1 value has to be smaller than the contribution if it were the only product the firm sold in each market, conditional on positive profits from product j in each market net of the fixed market entry cost.

Summing all inequalities of the form of inequality (35) across all bundles containing product j yields,

$$\mathbb{1}\left\{\sum_{m\in\mathcal{M}}\left(\mathbb{E}\big[O_{jm}^{\{j\}}\big[\mathbb{E}\big[\Delta_{jm}(\{j\},\Omega_m^{-f})|\mathcal{I}\big]-F_{jm}^e(\nu_{jm}^e;\theta_e,\sigma_e)|\mathcal{I}\big]\right)\right.$$

$$\left.-F_j^g(\nu_j^g;\theta_g,\sigma_g)\geq 0\right\}\geq G_j. \tag{36}$$

The expectation inside the indicator function in inequality (36) is an expectation such as that in Proposition 1.[3] Applying Jensen's inequality gives,

$$
\mathbb{1}\left\{\mathbb{E}\left[\sum_{m\in\mathcal{M}}\Gamma_{jm}\big(\Delta_{jm}(\{j\},\Omega_m^{-f})\big)\big[\Delta_{jm}(\{j\},\Omega_m^{-f})\right.\right.
$$
$$
\left.\left.-\mathbb{E}[F_{jm}^e(\nu_{jm}^e)|\mathcal{I},F_{jm}^e(\nu_{jm}^e)\leq\Delta_{jm}(\{j\},\Omega_m^{-f})]\big]\Big|\mathcal{I}\right]-F_j^g(\nu_j^g;\theta_g,\sigma_g)\geq0\right\}\geq G_j. \qquad (37)
$$

Henceforth, the analysis is identical to the derivation of the Stage 2 inequalities. As before, ν_j^g is now independent of the conditional expectation inside the indicator function in inequality (37). I take expectations on both sides conditional on $\mathcal{I}$ to obtain,

$$
\Lambda_j\left(\mathbb{E}\left[\sum_{m\in\mathcal{M}}\Gamma_{jm}\big(\Delta_{jm}(\{j\},\Omega_m^{-f})\big)\big[\Delta_{jm}(\{j\},\Omega_m^{-f})\right.\right.
$$
$$
\left.\left.-\mathbb{E}[F_{jm}^e(\nu_{jm}^e)|\mathcal{I},F_{jm}^e(\nu_{jm}^e)\leq\Delta_{jm}(\{j\},\Omega_m^{-f})]\big]\Big|\mathcal{I}\right];\theta_g,\sigma_g\right)\geq\mathbb{E}[G_j|\mathcal{I}]. \qquad (38)
$$

As with the Stage 2 inequalities, I use the convex upper bound of Λ_j of the form of that in equation (21), at an $\mathcal{I}$-measurable approximation point $\hat{x}_j$, such that, applying Jensen's inequality,

$$
\mathbb{E}\left[\overline{\Lambda}_j\left(\sum_{m\in\mathcal{M}}\Gamma_{jm}\big(\Delta_{jm}(\{j\},\Omega_m^{-f})\big)\big[\Delta_{jm}(\{j\},\Omega_m^{-f})\right.\right.
$$
$$
\left.\left.-\mathbb{E}[F_{jm}^e(\nu_{jm}^e)|\mathcal{I},F_{jm}^e(\nu_{jm}^e)\leq\Delta_{jm}(\{j\},\Omega_m^{-f})]\big],\hat{x}_j;\theta_g,\sigma_g\right)-G_j\Big|\mathcal{I}\right]\geq0. \qquad (39)
$$

Inequality (39) can now be used to estimate θ_g and σ_g. Note that $\mathbb{E}\big[F_{jm}^e(\nu_{jm}^e)|\mathcal{I},F_{jm}^e(\nu_{jm}^e)\leq\Delta_{jm}(\{j\},\Omega_m^{-f})\big]$ can be computed using numerical integration or Gaussian quadrature. I use the QuadGK Julia package to evaluate this conditional mean.

To derive a lower bound conditional moment inequality, I start from the fact that the following equality must hold at the best response for any $\mathcal{G}^f$ with $j\notin\mathcal{G}^f$,

$$
(\mathbb{1}_{\mathcal{G}^f\cup\{j\}}+\mathbb{1}_{\mathcal{G}^f})\mathbb{1}\big\{\mathcal{V}_f(\mathcal{G}^f\cup\{j\},\{\nu_j^g\}_{j\in\mathcal{A}^f},\mathcal{I})-\mathcal{V}_f(\mathcal{G}^f,\{\nu_j^g\}_{j\in\mathcal{A}^f},\mathcal{I})\leq0\big\}
$$
$$
=(\mathbb{1}_{\mathcal{G}^f\cup\{j\}}+\mathbb{1}_{\mathcal{G}^f})\mathbb{1}_{\mathcal{G}^f}. \qquad (40)
$$

Equation (40) says that conditional on firm f choosing either $\mathcal{G}^f\cup\{j\}$ or $\mathcal{G}^f$, it chooses $\mathcal{G}^f$ if and only if $\mathcal{G}^f$ is preferred to $\mathcal{G}^f\cup\{j\}$. Consider now the following sub-optimal second-stage decision rules under portfolio $\mathcal{G}^f\cup\{j\}$ in all markets m,

---

[3]I lighten notation by removing the (θ_e,σ_e) arguments from Γ_{jm} and F_{jm}^e.

$$\underline{\Omega}_m^{\mathcal{G}^f \cup \{j\}} = \begin{cases} O_{j'm}^{\mathcal{G}^f}, & j' \neq j \\ \underline{Q}_{j'm}, & j' = j. \end{cases}$$

i.e., firm f chooses the decision rule used for products $\mathcal{G}^f$ under portfolio $\mathcal{G}^f \cup \{j\}$ for all non-j products and uses the $\underline{Q}_{jm}$ decision rule for product j, where for now this decision rule is any arbitrary decision rule that must depend on $\mathcal{I}$ and any set of realized market entry fixed cost realizations. Combining these sub-optimal decision rules with equation (40),

$$(\mathbb{1}_{\mathcal{G}^f \cup \{j\}} + \mathbb{1}_{\mathcal{G}^f}) \times$$
$$\mathbb{1}\left\{ \sum_{m \in \mathcal{M}} \mathbb{E}\left[\Pi_m^f(\underline{\Omega}_m^{\mathcal{G}^f \cup \{j\}}) - \Pi_m^f(\Omega_m^{\mathcal{G}^f}) | \mathcal{I} \right] - F_j^g(\nu_j^g; \theta_g, \sigma_g) < 0 \right\} \geq (\mathbb{1}_{\mathcal{G}^f \cup \{j\}} + \mathbb{1}_{\mathcal{G}^f}) \mathbb{1}_{\mathcal{G}^f},$$

or

$$(\mathbb{1}_{\mathcal{G}^f \cup \{j\}} + \mathbb{1}_{\mathcal{G}^f}) \mathbb{1}\left\{ \sum_{m \in \mathcal{M}} \mathbb{E}\left[\underline{Q}_{jm} \left[\mathbb{E}\left[\Delta_{jm}(\Omega_m^{\mathcal{G}^f}, \Omega_m^{-f}) | \mathcal{I}, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{G}^f \right] - F_{jm}^e(\nu_{jm}^e) | \mathcal{I} \right] \right.$$
$$\left. - F_j^g(\nu_j^g; \theta_g, \sigma_g) < 0 \right\} \geq (\mathbb{1}_{\mathcal{G}^f \cup \{j\}} + \mathbb{1}_{\mathcal{G}^f}) \mathbb{1}_{\mathcal{G}^f}. \tag{41}$$

As with the upper bound, I use submodularity of variable profits to further bound the expression inside the indicator function in inequality (41). In particular,

$$(\mathbb{1}_{\mathcal{G}^f \cup \{j\}} + \mathbb{1}_{\mathcal{G}^f}) \mathbb{1}\left\{ \sum_{m \in \mathcal{M}} \mathbb{E}\left[\underline{Q}_{jm} \left[\mathbb{E}\left[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f}) | \mathcal{I} \right] - F_{jm}^e(\nu_{jm}^e) | \mathcal{I} \right] \right.\right.$$
$$\left.\left. - F_j^g(\nu_j^g; \theta_g, \sigma_g) < 0 \right\} \geq (\mathbb{1}_{\mathcal{G}^f \cup \{j\}} + \mathbb{1}_{\mathcal{G}^f}) \mathbb{1}_{\mathcal{G}^f}. \tag{42}$$

Inequality (42) holds for all bundles $\mathcal{G}^f$ that do not contain product j. Thus, summing all such inequalities across all bundles $\mathcal{G}^f$ that do not contain product j yields,

$$\mathbb{1}\left\{ \sum_{m \in \mathcal{M}} \mathbb{E}\left[\underline{Q}_{jm} \left[\mathbb{E}\left[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f}) | \mathcal{I} \right] - F_{jm}^e(\nu_{jm}^e) | \mathcal{I} \right] - F_j^g(\nu_j^g; \theta_g, \sigma_g) \geq 0 \right\} \leq G_j. \tag{43}$$

Next, let,

$$\underline{Q}_{jm} = \mathbb{1}\left\{ \mathbb{E}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f}) | \mathcal{I}] - F_{jm}^e(\nu_{jm}^e) \geq 0 \right\}.$$

Under convexity (Proposition 1), and letting $y := \Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})$, a lower bound of

$$g(\mathbb{E}(y|\mathcal{I})) := \mathbb{E}[\mathbb{1}\{\mathbb{E}(y|\mathcal{I}) - F_{jm}^e(\nu_{jm}^e) \geq 0\}[\mathbb{E}(y|\mathcal{I}) - F_{jm}^e(\nu_{jm}^e)]|\mathcal{I}]$$

is obtained by taking a first-order approximation around some $\mathcal{I}$-measurable point $\hat{x}_{jm}$. This yields,

$$
\begin{aligned}
g(\mathbb{E}(y|\mathcal{I})) &\geq \Gamma_{jm}(\hat{x}_{jm})[\mathbb{E}[y|\mathcal{I}] - \mathbb{E}[F^e_{jm}(\nu^e_{jm})|F^e_{jm} \leq \hat{x}_{jm}, \mathcal{I}]] \\
&= \Gamma_{jm}(\hat{x}_{jm})[\mathbb{E}[\Delta_{jm}(\mathcal{A}^f, \Omega^{-f}_m)|\mathcal{I}] - \mathbb{E}[F^e_{jm}(\nu^e_{jm})|F^e_{jm}(\nu^e_{jm}) \leq \hat{x}_{jm}, \mathcal{I}]].
\end{aligned} \tag{44}
$$

Plugging inequality (44) into inequality (43), I obtain,

$$
\mathbb{1}\Bigg\{ \mathbb{E}\Bigg[\sum_{m \in \mathcal{M}} \Gamma_{jm}(\hat{x}_{jm}) \big[\Delta_{jm}(\mathcal{A}^f, \Omega^{-f}_m) - \mathbb{E}[F^e_{jm}(\nu^e_{jm})|F^e_{jm}(\nu^e_{jm}) \leq \hat{x}_{jm}, \mathcal{I}]\big] \Big| \mathcal{I} \Bigg]
$$
$$
- F^g_j(\nu^g_j; \theta_g, \sigma_g) \geq 0 \Bigg\} \leq G_j \tag{45}
$$

where recall Γ_{jm} was the CDF of F^e_{jm} (given θ_e and σ_e). Given that the expectation term is now independent of ν^g_j conditional on $\mathcal{I}$, I take expectations conditional on $\mathcal{I}$ and obtain,

$$
\Lambda_j\Bigg(\mathbb{E}\Bigg[\sum_{m \in \mathcal{M}} \Gamma_{jm}(\hat{x}_{jm}) \big[\Delta_{jm}(\mathcal{A}^f, \Omega^{-f}_m)
$$
$$
- \mathbb{E}[F^e_{jm}(\nu^e_{jm})|F^e_{jm}(\nu^e_{jm}) \leq \hat{x}_{jm}, \mathcal{I}]\big] \Big| \mathcal{I} \Bigg]; \theta_g, \sigma_g \Bigg) \leq \mathbb{E}[G_j|\mathcal{I}]. \tag{46}
$$

Finally, I use the concave lower bound (as in equation (28)) $\underline{\Lambda}_j$ at an $\mathcal{I}$-measurable approximation point $\hat{x}_j$ and apply Jensen's inequality to obtain,

$$
\mathbb{E}\Bigg[\underline{\Lambda}_j\Bigg(\sum_{m \in \mathcal{M}} \Gamma_{jm}(\hat{x}_{jm}) \big[\Delta_{jm}(\mathcal{A}^f, \Omega^{-f}_m)
$$
$$
- \mathbb{E}[F^e_{jm}(\nu^e_{jm})|F^e_{jm}(\nu^e_{jm}) \leq \hat{x}_{jm}, \mathcal{I}]\big], \hat{x}_j; \theta_g, \sigma_g \Bigg) - G_j \Big| \mathcal{I} \Bigg] \leq 0. \tag{47}
$$

I have now shown how to derive the inequalities that can be used for estimation of parameters θ_g and σ_g. These are given by inequalities (39) and (47). In sum, I have proven Theorem 1.

C Solution Method Based on Inequalities

In this section, I provide a method to bound the equilibrium distribution of product offerings in each market. The method relies on first-order stochastic dominance for multivariate random vectors. Before defining first-order stochastic dominance, I define a partial order in $\mathbb{R}^n$. Throughout, I employ the standard partial order on $\mathbb{R}^n$: $\boldsymbol{x} \geq \boldsymbol{y}$ if and only if $x_i \geq y_i$ for

all $i \in \{1, ..., n\}$. An upper set in $\mathbb{R}^n$ is any set U of the form $U(\boldsymbol{y}) = \{\boldsymbol{x} : \boldsymbol{x} \geq \boldsymbol{y}\}$.

Definition 2 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two random vectors in $\mathbb{R}^n$ such that,*

$$\mathbb{P}(\boldsymbol{X} \in U) \geq \mathbb{P}(\boldsymbol{Y} \in U) \qquad \text{for all upper sets } U \subseteq \mathbb{R}^n.$$

Then $\boldsymbol{X}$ is said to first-order stochastically dominate $\boldsymbol{Y}$.

With Bernoulli random vectors, $\boldsymbol{X}$ first order stochastically dominating $\boldsymbol{Y}$ amounts to,

$$\mathbb{P}\Big(\bigcap_{i \in \Omega} \{X_i = 1\} \Big) \geq \mathbb{P}\Big(\bigcap_{i \in \Omega} \{Y_i = 1\} \Big) \qquad \text{for any } \Omega \subseteq \{1, ..., n\}.$$

I will also employ the following result.

Theorem 3 (Shaked 2007) *$\boldsymbol{X}$ FOSD $\boldsymbol{Y}$ if and only if $\mathbb{E}[u(\boldsymbol{X})] \geq \mathbb{E}[u(\boldsymbol{Y})]$ for all non-decreasing functions u for which the expectations exist.*

In counterfactual exercises, I am interested in learning about changes in the equilibrium market structure and product offerings in response to policies subsumed in $\mathcal{I}$. Notice that a change in $\mathcal{I}$ could be a subsidy, a tax, a change in the ownership structure of firms, or any other change in market conditions that is known at the time of making portfolio and market offerings choices. To obtain counterfactual bounds to these entry probabilities, I use Bayes' rule to write,

$$\mathbb{P}(O_{jm} = 1 | \mathcal{I}) = \mathbb{P}(O_{jm} = 1 | \mathcal{I}, G_j = 1) \mathbb{P}(G_j = 1 | \mathcal{I}).$$

In deriving the moment inequalities, I showed that inequalities (18) and (25) hold. These inequalities state that conditional on product $j \in \mathcal{G}^f$, $\underline{O}^*_{jm} \leq O^*_{jm} \leq \overline{O}^*_{jm}$, where

$$\overline{O}^*_{jm} := \mathbb{1}\{\mathbb{E}_{\boldsymbol{\mu}^*_m}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}] - F^e_{jm}(\nu^e_{jm}) \geq 0\} \tag{48}$$

$$\underline{O}^*_{jm} := \mathbb{1}\{\mathbb{E}_{\boldsymbol{\mu}^*_m}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}] - F^e_{jm}(\nu^e_{jm}) \geq 0\}, \tag{49}$$

Moreover, recall from (36) and (43) that, $\underline{G}^*_j \leq G^*_j \leq \overline{G}^*_j$, where

$$\overline{G}^*_j := \mathbb{1}\Big\{ \sum_{m \in \mathcal{M}} \mathbb{E}\big[\overline{O}^*_{jm}[\mathbb{E}_{\boldsymbol{\mu}^*_m}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}] - F^e_{jm}(\nu^e_{jm})|\mathcal{I}] - F^g_j(\nu^g_j) \geq 0 \Big\}, \tag{50}$$

$$\underline{G}^*_j := \mathbb{1}\Big\{ \sum_{m \in \mathcal{M}} \mathbb{E}\big[\underline{O}^*_{jm}[\mathbb{E}_{\boldsymbol{\mu}^*_m}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}] - F^e_{jm}(\nu^e_{jm})|\mathcal{I}] - F^g_j(\nu^g_j) \geq 0 \Big\}. \tag{51}$$

In equations (48)-(51), $\boldsymbol{\mu}_m^*$ denotes the distribution of firms' offerings decisions in market m at whichever equilibrium emerges under $\mathcal{I}$.[4]

Lemma 1 *Under Assumptions 1-4, for each market $m \in \mathcal{M}$, the sequences $\{\underline{O}_{jm}^*\}_{j\in\mathcal{A}}$, $\{\overline{O}_{jm}^*\}_{j\in\mathcal{A}}$, $\{\underline{G}_j^*\}_{j\in\mathcal{A}}$, and $\{\overline{G}_j^*\}_{j\in\mathcal{A}}$ are sequences of independent random variables conditional on $\mathcal{I}$ bounding the equilibrium binary decisions $\{O_{jm}^*\}_{j\in\mathcal{A}}$ and $\{G_j^*\}_{j\in\mathcal{A}}$.*

Proof. The fact that $\underline{G}_j^* \leq G_j^* \leq \overline{G}_j^*$ and $\underline{O}_{jm}^* \leq O_{jm}^* \leq \overline{O}_{jm}^*$ follow from the derivations leading to inequalities (36) and (35), and (18) and (25), respectively. Assumptions 2-3 impose conditional independence, so the independence result is immediate. ∎

Importantly, the true equilibrium sequences are *not* independent sequences of random variables given that firms' product introduction choices are correlated due to interdependencies coming from cannibalization. The asterisks in the definitions of the bounding random variables in (48)-(51) denote that firms' expectations are with respect to the *true* joint distribution of entry decisions made by other firms. Firms care about which products are ultimately sold in each market and, therefore, about the joint distribution (across products and markets) of random variables $\{V_{jm}^*\}_{j\in\mathcal{A}}$, where $V_{jm}^* := O_{jm}^* G_j^* = O_{jm}^*$, that is, the event that product j is offered in market m, which requires it being developed and being offered. The latter equality holds because, in equilibrium, $O_{jm}^* = 1$ implies $G_j^* = 1$. Then, $\underline{V}_{jm}^* \leq V_{jm}^* \leq \overline{V}_{jm}^*$, where $\underline{V}_{jm}^* := \underline{O}_{jm}^* \underline{G}_j^*$ and $\overline{V}_{jm}^* := \overline{O}_{jm}^* \overline{G}_j^*$. The independence of first and second-stage shocks implies that $\overline{O}_{jm}^*$ and $\overline{G}_j^*$ as well as $\underline{O}_{jm}^*$ and $\underline{G}_j^*$, are independent conditional on $\mathcal{I}$. A corollary of Lemma 1 is that within each market m, $\{\underline{V}_{jm}^*\}_{j\in\mathcal{A}}$ and $\{\overline{V}_{jm}^*\}_{j\in\mathcal{A}}$ are conditionally (on $\mathcal{I}$) independent sequences of random variables. Therefore, conditional on $\mathcal{I}$, it follows that,

$$\mathbb{P}(\underline{O}_{jm}^* = 1|\mathcal{I})\mathbb{P}(\underline{G}_j^* = 1|\mathcal{I}) = \mathbb{P}(\underline{V}_{jm}^* = 1|\mathcal{I}) \leq \mathbb{P}(O_{jm}^* = 1|\mathcal{I})$$
$$\leq \mathbb{P}(\overline{V}_{jm}^* = 1|\mathcal{I}) = \mathbb{P}(\overline{O}_{jm}^* = 1|\mathcal{I})\mathbb{P}(\overline{G}_j^* = 1|\mathcal{I}). \qquad (52)$$

So, given any equilibrium distribution of product offerings, the inequalities (52) must hold. By Lemma 1, we also have that,

$$\mathbb{E}[\prod_{j\in\mathcal{W}} \underline{V}_{jm}^*|\mathcal{I}] \leq \mathbb{E}[\prod_{j\in\mathcal{W}} V_{jm}^*|\mathcal{I}] \leq \mathbb{E}[\prod_{j\in\mathcal{W}} \overline{V}_{jm}^*|\mathcal{I}], \qquad (53)$$

given any set of products $\mathcal{W}$. We will use this result to bound not only the marginal probabilities that products are offered, but also the joint distribution of product offerings in each market.

---

[4]Firm f only integrates over rival firms' offerings decisions. I simplify notation by using a $\boldsymbol{\mu}_m^*$ subscript, rather than a $\boldsymbol{\mu}_m^{*,-f}$ subscript.

C.1   Proof of Theorem 2

First, define:

$$\overline{O}_{jm}^k := \mathbb{1}\{\mathbb{E}_{\underline{\boldsymbol{\mu}}_m^{k-1}}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}] - F_{jm}^e(\nu_{jm}^e) \geq 0\}, \tag{54}$$

$$\underline{O}_{jm}^k := \mathbb{1}\{\mathbb{E}_{\overline{\boldsymbol{\mu}}_m^{k-1}}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}] - F_{jm}^e(\nu_{jm}^e) \geq 0\}, \tag{55}$$

$$\overline{G}_j^k := \mathbb{1}\Big\{\sum_m \mathbb{E}\big[\overline{O}_{jm}^k\big[\mathbb{E}_{\underline{\boldsymbol{\mu}}_m^{k-1}}\big[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}\big] - F_{jm}^e(\nu_{jm}^e)\big]\big|\mathcal{I}\big] - F_j^g(\nu_j^g) \geq 0\Big\}, \tag{56}$$

$$\underline{G}_j^k := \mathbb{1}\Big\{\sum_m \mathbb{E}\big[\underline{O}_{jm}^k\big[\mathbb{E}_{\overline{\boldsymbol{\mu}}_m^{k-1}}\big[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}\big] - F_{jm}^e(\nu_{jm}^e)\big]\big|\mathcal{I}\big] - F_j^g(\nu_j^g) \geq 0\Big\}, \tag{57}$$

To lighten notation in the proof, I define denote the probability that each of $\underline{O}_{jm}^k$, $\overline{O}_{jm}^k$, $\underline{G}_j^k$, $\overline{G}_j^k$, are equal to 1, respectively, as,

$$\underline{\mu}_{jm}^{cond}(\overline{\boldsymbol{\mu}}_m^{k-1}) := \Gamma_{jm}\big(\mathbb{E}_{\overline{\boldsymbol{\mu}}_m^{k-1}}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}]\big), \tag{58}$$

$$\overline{\mu}_{jm}^{cond}(\underline{\boldsymbol{\mu}}_m^{k-1}) := \Gamma_{jm}\big(\mathbb{E}_{\underline{\boldsymbol{\mu}}_m^{k-1}}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}]\big), \tag{59}$$

$$\underline{\mu}_j^{port}(\overline{\boldsymbol{\mu}}^{k-1}) := \Lambda_j\Big(\sum_{m \in \mathcal{M}} \mathbb{E}\big[\underline{O}_{jm}^k[\mathbb{E}_{\overline{\boldsymbol{\mu}}_m^{k-1}}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}] - F_{jm}^e(\nu_{jm}^e)]\big|\mathcal{I}\big]\Big), \tag{60}$$

$$\overline{\mu}_j^{port}(\underline{\boldsymbol{\mu}}^{k-1}) := \Lambda_j\Big(\sum_{m \in \mathcal{M}} \mathbb{E}\big[\overline{O}_{jm}^k[\mathbb{E}_{\underline{\boldsymbol{\mu}}_m^{k-1}}[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f})|\mathcal{I}] - F_{jm}^e(\nu_{jm}^e)]\big|\mathcal{I}\big]\Big), \tag{61}$$

where $\boldsymbol{\mu}^k := (\boldsymbol{\mu}_m^k)_{m \in \mathcal{M}}$. Note that $\{\underline{O}_{jm}^k\}_{j \in \mathcal{A}}$, $\{\overline{O}_{jm}^k\}_{j \in \mathcal{A}}$, $\{\underline{G}_j^k\}_{j \in \mathcal{A}}$, and $\{\overline{G}_j^k\}_{j \in \mathcal{A}}$ are sequences of independent random variables conditional on $\mathcal{I}$, and offering and portfolio choice bounds are conditionally independent, e.g., $\overline{O}_{jm}^k$ is independent of $\overline{G}_j^k$. **Proof.** I first prove that for any product j and any market m, $\{\overline{\mu}_{jm,k}\}_{k=0}^\infty$ is a decreasing sequence and $\{\underline{\mu}_{jm,k}\}_{k=0}^\infty$ is an increasing sequence. I prove this by induction. By construction $\underline{\mu}_{jm,1} > 0 = \underline{\mu}_{jm,0}$ and $\overline{\mu}_{jm,1} < 1 = \overline{\mu}_{jm,0}$. It follows that the hypothesis is true for $k = 1$. Assume that the hypothesis is true up to index $K > 1$. I show that (i) $\underline{\mu}_{jm,K+1} \geq \underline{\mu}_{jm,K}$ and (ii) $\overline{\mu}_{jm,K+1} \leq \overline{\mu}_{jm,K}$. To prove (i), note that by definition,

$$\underline{\mu}_{jm,K+1} = \underline{\mu}_{jm}^{cond}(\overline{\boldsymbol{\mu}}_m^K)\underline{\mu}_j^{port}(\overline{\boldsymbol{\mu}}^K).$$

By the inductive hypothesis, $\overline{\mu}_{jm,K} \leq \overline{\mu}_{jm,K-1}$ for each j and m. By definition of first-order stochastic dominance, and due to independence of the bounds across products within markets, this implies that $\overline{\boldsymbol{\mu}}_m^K \leq_{FOSD} \overline{\boldsymbol{\mu}}_m^{K-1}$ for all m. Moreover, due to Assumption 4 and

16

Theorem 3, both $\underline{\mu}_{jm}^{cond}$ and $\underline{\mu}_{j}^{port}$ are decreasing (in the FOSD sense). Therefore,

$$\underline{\mu}_{jm,K+1} = \underline{\mu}_{jm}^{cond}(\overline{\boldsymbol{\mu}}_m^K)\underline{\mu}_j^{port}(\overline{\boldsymbol{\mu}}^K) \geq \underline{\mu}_{jm}^{cond}(\overline{\boldsymbol{\mu}}_m^{K-1})\underline{\mu}_j^{port}(\overline{\boldsymbol{\mu}}^{K-1}) = \underline{\mu}_{jm,K}.$$

I have therefore proven that for each product j and each market m, the sequence of lower bound probabilities is increasing. Again, due to independence, this also implies that $\underline{\boldsymbol{\mu}}_m^{K+1} \geq_{FOSD} \underline{\boldsymbol{\mu}}_m^K$ for each K and each market m. I analogously prove (ii). First, write $\overline{\mu}_{jm,K} \geq \overline{\mu}_{jm,K-1}$ for each j and in each m. By definition of first-order stochastic dominance, and due to independence, this implies that $\underline{\boldsymbol{\mu}}_m^K \geq_{FOSD} \underline{\boldsymbol{\mu}}_m^{K-1}$. It follows from Assumption 4 and Theorem 3 that,

$$\overline{\mu}_{jm,K+1} = \overline{\mu}_{jm}^{cond}(\underline{\boldsymbol{\mu}}_m^K)\overline{\mu}_j^{port}(\underline{\boldsymbol{\mu}}^K) \leq \overline{\mu}_{jm}^{cond}(\underline{\boldsymbol{\mu}}_m^{K-1})\overline{\mu}_j^{port}(\underline{\boldsymbol{\mu}}^{K-1}) = \overline{\mu}_{jm,K}.$$

Thus, $\{\overline{\mu}_{jm,k}\}_{k=0}^{\infty}$ is decreasing and $\overline{\boldsymbol{\mu}}_m^K \leq_{FOSD} \overline{\boldsymbol{\mu}}_m^{K-1}$ due to independence. Next, I prove that for all k, necessarily,

$$\underline{\boldsymbol{\mu}}_m^k \leq_{FOSD} \boldsymbol{\mu}_m^* \leq_{FOSD} \overline{\boldsymbol{\mu}}_m^k \tag{62}$$

where asterisks denote any equilibrium distribution of product offering decisions under any arbitrary information set $\mathcal{I}$, for each $m \in \mathcal{M}$. I also prove this by induction. Inequalities (62) clearly hold for $k = 1$. Assume they hold for arbitrary $K \in \mathbb{N}$. I now prove that it must hold for $K + 1$. To do so, take any arbitrary subset $\mathcal{W} \subseteq \mathcal{A}$. Note that $\mathcal{W} = \mathcal{W}^1 \cup \mathcal{W}^2 \cup ... \cup \mathcal{W}^F$, a partition across firms. Then,

$$\mathbb{P}\left(\bigcap_{j \in \mathcal{W}} \{\overline{V}_{jm}^{K+1} = 1\} | \mathcal{I} \right) = \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{P}(\overline{V}_{jm}^{K+1} = 1 | \mathcal{I}) \tag{63}$$

$$= \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \overline{\mu}_{jm}^{cond}(\underline{\boldsymbol{\mu}}_m^K)\overline{\mu}_j^{port}(\underline{\boldsymbol{\mu}}^K) \geq \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \overline{\mu}_{jm}^{cond}(\underline{\boldsymbol{\mu}}_m^*)\overline{\mu}_j^{port}(\underline{\boldsymbol{\mu}}^*) \tag{64}$$

$$= \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{P}(\overline{O}_{jm}^* = 1 | \mathcal{I})\mathbb{P}(\overline{G}_j^* = 1 | \mathcal{I}) = \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{P}(\overline{V}_{jm}^* = 1 | \mathcal{I}) = \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{E}(\overline{V}_{jm}^* | \mathcal{I})$$

$$= \mathbb{E}\left[\prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \overline{V}_{jm}^* | \mathcal{I} \right] \geq \mathbb{E}\left[\prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} V_{jm}^* | \mathcal{I} \right] = \mathbb{P}\left(\bigcap_{j \in \mathcal{W}} \{V_{jm}^* = 1\} | \mathcal{I} \right), \tag{65}$$

where the inequalities in lines (63) and (65) follow from the (conditional) independence of the upper bounds and (64) follows from the inductive hypothesis and Theorem 3. This

17

proves $\overline{\boldsymbol{\mu}}^{K+1} \geq_{FOSD} \boldsymbol{\mu}^*$. I analogously prove the other direction. Indeed,

$$\mathbb{P}\left(\bigcap_{j \in \mathcal{W}} \{\underline{V}_{jm}^{K+1} = 1\} | \mathcal{I} \right) = \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{P}(\underline{V}_{jm}^{K+1} = 1 | \mathcal{I}) \tag{66}$$

$$= \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \underline{\mu}_{jm}^{cond}(\overline{\boldsymbol{\mu}}_m^K) \underline{\mu}_j^{port}(\overline{\boldsymbol{\mu}}^K) \leq \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \underline{\mu}_{jm}^{cond}(\overline{\boldsymbol{\mu}}_m^*) \underline{\mu}_j^{port}(\overline{\boldsymbol{\mu}}^*) \tag{67}$$

$$= \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{P}(\underline{O}_{jm}^* = 1 | \mathcal{I}) \mathbb{P}(\underline{G}_j^* = 1 | \mathcal{I}) = \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{P}(\underline{V}_{jm}^* = 1 | \mathcal{I}) = \prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \mathbb{E}(\underline{V}_{jm}^* | \mathcal{I})$$

$$= \mathbb{E}\left[\prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} \underline{V}_{jm}^* | \mathcal{I} \right] \leq \mathbb{E}\left[\prod_{f=1}^{F} \prod_{j \in \mathcal{W}^f} V_{jm}^* | \mathcal{I} \right] = \mathbb{P}\left(\bigcap_{j \in \mathcal{W}} \{V_{jm}^* = 1\} | \mathcal{I} \right), \tag{68}$$

where (66) and (68) follow from (conditional) independence of the lower bounds and (67) follows from the inductive hypothesis and Theorem 3.

This concludes the proof. ∎

Notice that this proof only relies on Assumption 1, Assumption 4, and conditional independence of fixed cost shocks. It does not rely on the log-normality assumption.

C.2 Implementation

To implement Algorithm 1 in practice, I need to compute expectations $\mathbb{E}_{\boldsymbol{\mu}_m}$, which are with respect to the distribution of rival firms' offerings decisions given $\boldsymbol{\mu}_m$ and over demand and marginal cost shocks. Thus, I first take 100 draws from the joint distribution of $(\boldsymbol{\xi}, \boldsymbol{\omega})$ across 9766 product markets, yielding a 9766×100 matrix of demand shocks and a similar matrix of marginal cost shocks. Second, I draw a 9766×100 matrix U^{cond} and a 1530×100 matrix U^{port} of uniformly distributed draws on $(0, 1)$. I hold all draws fixed throughout all iterations.

The U^{port} matrix yields thresholds for the portfolio decisions given by probabilities $\overline{\mu}_{j,k}^{port}$ and $\underline{\mu}_{j,k}^{port}$. For instance, if product j has a draw smaller than $\overline{\mu}_{j,k}^{port}$, then that product is introduced in the firm's portfolio under that draw at iteration k at the upper bound. Otherwise, at iteration k and under such a draw, product j is not introduced in the firm's global product portfolio at the upper bound. The U^{cond} matrix yields thresholds for the offering choices given by probabilities $\overline{\mu}_{jm,k}^{cond}$ and $\underline{\mu}_{jm,k}^{cond}$. For instance, if for product j in market m the U^{cond} draw is less than the $\overline{\mu}_{jm,k}^{cond}$ probability at iteration k, then that draw corresponds to such a product being offered in such a market at iteration k (conditional on being in the firm's portfolio). Otherwise, that product is not offered in market m under iteration k and that draw. For a product to be offered in a market at a given iteration k, it

must satisfy both the offering and portfolio thresholds.

It follows that under each iteration k, I need to compute 9766×100 marginal values across all product-market pairs. Then, I average them across the 100 draws to obtain simulated values of the expected marginal values at the product-market level. This averaging yields the expectations that enter the fixed cost CDFs in equations (58)-(61). To compute the marginal values under any draw, I use the Morrow and Skerlos (2011) contraction mapping for the pricing equilibria with 500 income and normal draws for each market.

I run Algorithm 1 for 6 iterations under each parameter vector $(\theta_e, \sigma_e, \theta_g, \sigma_g)$ and each policy counterfactual. I find that after $k > 6$, the additional gains in informativeness are small in my application.

D   Estimation Implementation

D.1   Demand and Marginal Cost Estimation

Demand estimation follows Petrin (2002). The assumption that $(\boldsymbol{\xi}, \boldsymbol{\omega})$ are realized after firms make offerings choices implies that $\mathbb{E}[(\xi_{jm}, \omega_{jm})|\mathcal{I}] = 0$. Under this assumption, standard Berry et al. (1995) or Gandhi and Houde (2019) instruments are valid. I use size, horsepower, and horsepower/weight to build Gandhi and Houde (2019) differentiation instruments. These instruments are constructed by using the characteristics of products that are "close" in the characteristics space. I use the PyBLP Python package to construct the differentiation instruments (Conlon and Gortmaker 2020). For each of the three characteristics, I build two instruments. Let $x_{jm\ell}$ denote the value of any such characteristic, indexed by ℓ. I construct, for each $\ell \in \{$size, horsepower, horsepower/weight$\}$:

$$z_{jm\ell}^{other} = \sum_{k \in \Omega_m^f \setminus \{j\}} \mathbb{1}\{|d_{jkm\ell}| < SD_\ell\}, \qquad z_{jm\ell}^{rival} = \sum_{k \in \Omega_m^{-f}} \mathbb{1}\{|d_{jkm\ell}| < SD_\ell\}$$

where $d_{jkm\ell} = x_{jm\ell} - x_{km\ell}$ and SD_ℓ denotes the standard deviation of all such pairwise differences computed across all markets. This yields 6 differentiation instruments.

To improve the precision of my estimates, I include 4 additional moments. First, I include (i) the log of distance to the brand's HQ country, which enters my marginal cost specification, and (ii) the average price of the same product in other markets / the average price of products of the same parent company in other markets (when a given product is only sold in one market). The latter set of instruments resembles those in Hausman (1996), which are valid provided demand and marginal cost shocks $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$ are uncorrelated across markets.

Second, I use micro-moments, similarly to Petrin (2002). I use micro-data from the 2019 MRI-Simmons Crosstab report to help pin down the heterogeneity in preferences for prices within countries. I match the probabilities that a consumer in the United States is within a given income group conditional on purchasing a vehicle in a given price range. More specifically, I match the following two moments:

(a) $\quad\quad\quad\quad$ $\mathbb{E}[income_i > \$100,000 | price_{jm} > \$50,000, m = \text{United States}]$

(b) $\quad\quad\quad\quad$ $\mathbb{E}[income_i \in [\$60,000, \$100,000] | price_{jm} > \$50,000, m = \text{United States}]$.

I jointly estimate demand and marginal costs using Python's PyBLP package. To integrate over the distribution of income in each market/country, I take 20,000 simulation draws in each market m from a log-normal distribution with location and scale,

$$\mu_m = \log(GDP\_per\_capita\_PPP_m) - \sigma_m^2/2, \quad\quad \sigma_m = \sqrt{2}\Phi^{-1}\left(\frac{Gini_m + 1}{2}\right),$$

where Φ is the standard normal CDF. This parametrization ensures that income is drawn from a log-normal distribution with mean and Gini coefficients equal to the observed values. I obtain the Gini coefficient and PPP GDP per capita in each market m from the World Bank. The main results from the estimation are in the main text, Table 1. Below, I report the matched micro-moments:

Table D.1: Micro-Moments

Moment	Observed	Estimated	Difference
(a)	0.631	0.612	0.0188
(b)	0.212	0.245	-0.0329

D.1.1 Distribution of $(\boldsymbol{\xi}, \boldsymbol{\omega})$

I fit a bivariate normal distribution for the distribution of demand and marginal cost shocks, which I assume firms know at the time of making product portfolio and product offerings decisions. I estimate a variance of 2.36 for ξ, a variance of 0.024 for ω, and a covariance of 0.081.

D.2 Fixed Cost Estimation - Instruments

To estimate fixed costs, I implement the moment inequalities derived in Sections B.1 and B.2. As described in the main text, I first construct instruments using an approach

similar to two-stage least squares. I then use these instruments to build unconditional moment inequalities. I use the empirical counterparts of these moment inequalities to estimate $(\theta_e, \sigma_e, \theta_g, \sigma_g)$. The steps are described in the main text, Section 4.

To construct the instruments, I use the Stage 3 model for variable profits, together with the fitted distribution of $(\boldsymbol{\xi}, \boldsymbol{\omega})$ to compute the ingredients necessary to implement the inequalities in Theorem 1. I simulate $S = 200$ draws $\{\xi_{jm}^s, \omega_{jm}^s\}_{s=1, j \in \mathcal{A}, m \in \mathcal{M}}^{200}$ from the fitted bivariate normal distribution and construct, for each product j and each market m,

$$\widehat{\Delta}_{jm}(\{j\}, \Omega_m^{-f}) = \frac{1}{S} \sum_{s=1}^{S} \Delta_{jm}(\{j\}, \Omega_m^{-f}; \xi_{jm}^s, \omega_{jm}^s) \tag{69}$$

$$\widehat{\Delta}_{jm}(\mathcal{G}^f, \Omega_m^{-f}) = \frac{1}{S} \sum_{s=1}^{S} \Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f}; \xi_{jm}^s, \omega_{jm}^s) \tag{70}$$

$$\widehat{\Delta}_{jm}(\mathcal{A}^f, \Omega_m^{-f}) = \frac{1}{S} \sum_{s=1}^{S} \Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f}; \xi_{jm}^s, \omega_{jm}^s). \tag{71}$$

To obtain an exogenous ($\mathcal{I}$- measurable) predictor of the above marginal values, I project the log of the predicted values in (69)-(71) on the following objects: (i) the log of $\tilde{\delta}_{jm} = \beta_m + \beta_{b(j)} + \beta^x \boldsymbol{X}_{jm}$, the non-price mean utility of product j in market m net of unobserved heterogeneity ξ_{jm}, (ii) $\tilde{c}_{jm} = \log(c_{jm}) - \omega_{jm}$, the (log) marginal cost of supplying market m with product j net of the unobserved heterogeneity ω_{jm}, and (iii) $\log(\tilde{\delta}_{jm}) \times \tilde{c}_{jm}$. To carry out the projection, I estimate 4 PPML specifications. I first project (69)-(70), conditional on $j \in \mathcal{G}^f$, using, $\hat{y}_{jm} = \exp(\kappa^1 M_m + \kappa_m^2 \tilde{\delta}_{jm} + \kappa_m^3 \tilde{c}_{jm} + \kappa_m^4 \tilde{c}_{jm} \times \tilde{\delta}_{jm}) + \varepsilon_{jm}$. I then use the same specification to project (69) and (71), using the full sample of potential products. As described in the main text, I then use predicted values from these regressions to construct the unconditional moments that can be used for estimation.

To implement the Stage 2 inequalities that partially identify (θ_e, σ_e), I use tercile bin indicators for the predicted values in the PPML regressions that project (69)-(70), which I denote by $\hat{\bar{x}}_{jm}^g$ and $\hat{\underline{x}}_{jm}^h$, respectively. I also use the squares of such (positive) predicted values interacted with an indicator function denoting whether the log of the predicted value (in billions of USD) is greater than -2, which selects more profitable products. This helps to provide bounds on the scale parameter σ_e, as argued in Section 4.

To implement the Stage 1 inequalities that partially identify (θ_g, σ_g), I use quintile bin indicators (5 bins) for the *sum* across markets of the predicted values arising from PPML regressions that project (69) and (71) using the full sample of potential products – $\hat{\bar{x}}_{jm}^a$ and $\hat{\underline{x}}_{jm}^l$, respectively –, as well as squares of such sums of predicted values, interacted with an indicator function for whether the log of the sum (in billions of USD) is greater than 0.

D.2.1 Approximation Points Used for Convex and Convex CDF Bounds

To implement the moment inequality procedure, I use observation-specific upper and lower bounds of the fixed-cost CDFs. I use the convex and concave families of functions in equations (21) and (28), respectively. The approximation points I use for the Stage 2 upper and lower bound inequalities are $\hat{\bar{x}}^g_{jm}$ and $\hat{\underline{x}}^h_{jm}$, respectively. I use $\sum_m \hat{\bar{x}}^a_{jm}$ and $\sum_m \hat{\underline{x}}^l_{jm}$, respectively, for the Stage 1 upper and lower bound inequalities.

D.3 Fixed Cost Estimation - Empirical Analogues of Moments in Theorem 1

I construct positive-valued instruments and interact them with the conditional moments in Theorem 1 to construct empirical unconditional moment inequalities,

$$\frac{1}{J_g M} \sum_{j \in \mathcal{G}, m \in \mathcal{M}} [\overline{\Gamma}_{jm}(\Delta_{jm}(\{j\}, \Omega_m^{-f}); \theta_e, \sigma_e) - O_{jm}] \mathbb{1}\{\hat{\bar{x}}^g_{jm} \in Q_\tau(\hat{\bar{x}}^g_{jm})\} \geq 0, \tag{72}$$
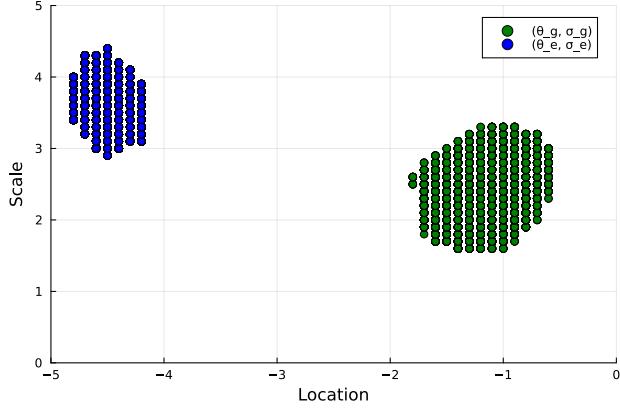
$$\frac{1}{J_g M} \sum_{j \in \mathcal{G}, m \in \mathcal{M}} [\underline{\Gamma}_{jm}(\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f}); \theta_e, \sigma_e) - O_{jm}] \mathbb{1}\{\hat{\underline{x}}^h_{jm} \in Q_\tau(\hat{\underline{x}}^h_{jm})\} \leq 0, \tag{73}$$

$$\frac{1}{J} \sum_{j \in \mathcal{A}} \left[\overline{\Lambda}_j \left(\sum_{m \in \mathcal{M}} \Gamma_{jm}(\Delta_{jm}(\{j\}, \Omega_m^{-f})) \left[\Delta_{jm}(\{j\}, \Omega_m^{-f}) \right. \right. \right.$$

$$\left. \left. - \mathbb{E}\left[F^e_{jm}(\nu^e_{jm})|\mathcal{I}, F^e_{jm}(\nu^e_{jm}) \leq \Delta_{jm}(\{j\}, \Omega_m^{-f})\right]\right]; \theta_g, \sigma_g \right) - G_j \right] \tag{74}$$

$$\times \mathbb{1}\left\{ \sum_{m \in \mathcal{M}} \hat{\bar{x}}^a_{jm} \in Q_\tau \left(\sum_{m \in \mathcal{M}} \hat{\bar{x}}^a_{jm} \right) \right\} \geq 0,$$

$$\frac{1}{J} \sum_{j \in \mathcal{A}} \left[\underline{\Lambda}_j \left(\sum_{m \in \mathcal{M}} \Gamma_{jm}(\hat{\underline{x}}^l_{jm}) \left[\Delta_{jm}(\mathcal{A}^f, \Omega_m^{-f}) \right. \right. \right.$$

$$\left. \left. - \mathbb{E}\left[F^e_{jm}(\nu^e_{jm})|F^e_{jm}(\nu^e_{jm}) \leq \hat{\underline{x}}^l_{jm}, \mathcal{I}\right]\right]; \theta_g, \sigma_g \right) - G_j \right]$$

$$\times \mathbb{1}\left\{ \sum_{m \in \mathcal{M}} \hat{\underline{x}}^l_{jm} \in Q_\tau \left(\sum_{m \in \mathcal{M}} \hat{\underline{x}}^l_{jm} \right) \right\} \leq 0, \tag{75}$$

where the expectation of the truncated fixed costs in inequalities (74)-(75) are computed using the QuadGK Julia package (Gaussian quadrature). J_g denotes the number of products in the sample that are offered in at least one market.

Figure E.1: Baseline Fixed Cost Estimates - Grid Search

E   Fixed Cost Estimation - Grid Search and Robustness

Table E.1: Stages 1 and 2 Parameter Confidence Set Limits

	95% Confidence Set Limits	
	More IV Bins	**Fewer IV Bins**
Stage 2: Market Entry Fixed Cost		
θ_e (Location)	[-4.9, -4.1]	[-4.8, -4.2]
σ_e (Scale)	[2.9, 4.4]	[2.8, 4.3]
Stage 1: Product Fixed Cost		
θ_g (Location)	[-1.9, -0.5]	[-1.9, -0.6]
σ_g (Scale)	[1.5, 3.6]	[1.5, 3.7]
Observations – Stage 2	3240	3240
Observations – Stage 1	739	739

Note: Confidence sets computed using Andrews and Soares (2010). First, I implement a grid search to compute a 97.5% confidence set for parameters (θ_e, σ_e) using the Stage 2 moment inequalities. Then, I use the Stage 1 moment inequalities to compute a 97.5% confidence set for (θ_g, σ_g), evaluating the moments at the accepted values of (θ_e, σ_e). The Bonferroni correction yields a 95% confidence set for all four parameters $(\theta_e, \sigma_e, \theta_g, \sigma_g)$. Marginal values are in billions of U.S. dollars.

Figure E.1 plots the grid search for the baseline fixed cost estimates.

Next, I report the estimation results using more instrument bins for both Stage 2 and Stage 1 inequalities. I construct 10 bins for the Stage 2 inequalities and 8 bins for the Stage 1 inequalities using the regressions from Section D.2. Recall that the baseline specification used 3 bins for Stage 2 and 5 bins for Stage 1. I keep the remaining polynomial-based instruments as they are in the main specification. I do not find that including more instrument bins leads to model rejection. Table E.1 shows that the confidence sets do not change very significantly

under more instrument bins. And, if anything, the limits of the confidence sets become wider. This is due to increased covariance across more instruments.

I also report the confidence sets under fewer instrument bins: 2 bins for both Stage 2 and Stage 1. Again, all remaining instruments are left unchanged. Table E.1 shows that the confidence sets do not change much relative to the main specification. This increases my confidence in the results and shows that they are robust to the number of instrument bins.

F   Moment Inequalities Extension - Economies of Scope

The model described in the main text assumes that portfolio fixed costs and market entry fixed costs are independent of the firms' chosen portfolio and offerings bundles. In this subsection, I show how this assumption can be relaxed. I allow market entry fixed costs to exhibit a version of economies of scope. Similar arguments can also be used to allow for economies of scope in global portfolio fixed costs.

Suppose that the total market entry fixed costs paid by firm f if it offers products Ω_m^f are,

$$F_m^{e,f} = \theta_0 \mathbb{1}\{|\Omega_m^f| \geq 1\} + \sum_{j \in \Omega_m^f} F_{jm}^e.$$

That is, firms pay a constant amount θ_0 to enter the market, and then market entry fixed costs with a specification as in the main text, given by Assumption 2. Throughout, I assume that a natural lower bound on θ_0 is 0 to be consistent with the fact that fixed costs are positive under the log-normal assumption. Then, I apply similar arguments to those in the main text to derive a lower and an upper bound inequality. As before, I start with revealed-preference equality,

$$(\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}})[\mathbb{1}\mathbb{E}[\Delta_{jm}(\Omega_m^f, \Omega_m^{-f})|\mathcal{I}, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{G}^f]$$
$$- \theta_0 \mathbb{1}\{\Omega_m^f = \{j\}\} - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0 - \mathbb{1}_{\Omega_m^f}] = 0,$$

where I have now allowed the fixed cost to be higher if j is the first product to be introduced into market m. Under submodularity and $\theta_0 \geq 0$, one can derive upper bound inequality,

$$(\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}})[\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{G}^f]$$
$$- F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} - \mathbb{1}_{\Omega_m^f}] \geq 0,$$

which results from maximizing the marginal value and minimizing the marginal fixed cost.

For the upper bound, I then follow the same arguments as in the main text to derive upper bound inequality,

$$\mathbb{E}[\overline{\Gamma}_{jm}(\Delta_{jm}(\{j\}, \Omega_m^{-f}); \theta_e, \sigma_e) - O_{jm}|\mathcal{I}, \mathcal{G}^f] \geq 0. \tag{76}$$

Inequality (76) does not provide any bound on θ_0. For the lower bound, I proceed in a similar fashion. Minimizing Δ_{jm} and maximizing marginal fixed costs, I derive inequality,

$$(\mathbb{1}_{\Omega_m^f} + \mathbb{1}_{\Omega_m^f \setminus \{j\}})[\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f})|\mathcal{I}, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{G}^f] - \theta_0$$
$$- F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} - \mathbb{1}_{\Omega_m^f}] \leq 0.$$

Following similar steps as in section B.1, I obtain conditional moment inequality,

$$\mathbb{E}[\underline{\Gamma}_{jm}(\Delta_{jm}(\mathcal{G}^f, \Omega_m^{-f}) - \theta_0; \theta_e, \sigma_e) - O_{jm}|\mathcal{I}] \leq 0. \tag{77}$$

Inequality (77) provides a lower bound on θ_0. Therefore, the inequalities discussed so far only identify a lower bound and not an upper bound of the new parameter of interest θ_0. In the following subsection, I discuss additional inequalities that partially identify θ_0.

F.1 Additional Inequalities

To derive additional inequalities, I condition on either a single product or no product being offered in market m. That is, I now start with,

$$(\mathbb{1}_{\{j\}_m} + \mathbb{1}_{\emptyset_m}) \times$$
$$[\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{G}^f] - \theta_0 - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} - \mathbb{1}_{\{j\}_m}] = 0.$$

To deal with the issue of selection, I first re-write the above inequality as,

$$(\mathbb{1}_{\{j\}_m} + \mathbb{1}_{\emptyset_m}) \times \tag{78}$$
$$\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}, \{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}, \mathcal{G}^f] - \theta_0 - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} = \mathbb{1}_{\{j\}_m}.$$

Since $(\mathbb{1}_{\{j\}_m} + \mathbb{1}_{\emptyset_m}) \leq 1$ and due to Assumption 1, I obtain,

$$\mathbb{1}\{\mathbb{E}[\Delta_{jm}(\{j\}, \Omega_m^{-f})|\mathcal{I}, \mathcal{G}^f] - \theta_0 - F_{jm}^e(\nu_{jm}^e; \theta_e, \sigma_e) \geq 0\} - \mathbb{1}_{\{j\}_m} \geq 0. \tag{79}$$

Inequality (79) says that if only a single product j is offered in market m, then the change in profits from doing so must necessarily be weakly positive. Letting O_{jm}^1 denote the event

that only product j is sold in market m by firm f, I follow steps similar to those in the main text to derive inequality,

$$\mathbb{E}[\overline{\Gamma}_{jm}(\Delta_{jm}(\{j\}, \Omega_m^{-f}) - \theta_0; \theta_e, \sigma_e) - O_{jm}^1 | \mathcal{I}] \geq 0.$$

To implement this moment inequality in practice, the empirical counterpart of $\mathbb{E}[O_{jm}^1 | \mathcal{I}]$ averages across firm-markets a dummy equal to 1 if a given firm introduces a singleton bundle in a given market. An upper bound on θ_0 is identified because a value of θ_0 that is too large would imply too few observations of singleton product entry across markets.

In practice, I do not observe sufficient instances of singleton product entry to obtain informative bounds on θ_0.

G   Simulating the Method

In this section, I use simulation to:

1. Test the behavior of the moment inequalities proposed in Section B under varying data-generating processes (DGPs),

2. Test the performance of the inference methods used in this paper, based on Andrews and Soares (2010), when a single realization of the product entry game is observed.

To simulate data, I need to solve the model fully. Thus, I simulate a solvable version of my product entry model, which I describe in Section G.1.

G.1   Fully Solvable Version of Global Product Entry Game

The solvable model features N symmetric J-product firms competing in M markets. I set $J = 3$ so that firms have 3 potential products that they can introduce in their global product portfolios and across markets. Both firms and products are symmetric in their profit shifters, but markets are allowed to be heterogeneous.

Profits for firm f from selling N_m^f products in market m take the form, $\Pi_m^f(N_m^f, N_m^{-f}) = A_m \frac{N_m^f}{1 + (N_m^f)^{\kappa_o}(N_m^{-f})^{\kappa_r}}$, where $\kappa_o \in (0, 1)$ regulates substitutability across the firm's own products, and $\kappa_r \in (0, 1)$ regulates substitution across rival firms' products. A_m is an exogenous market-level profit shifter. Firms draw log-normal fixed costs, as in the main text. I abuse notation and index all firms' potential products as 1, 2, or 3 (even though they are different products).

Timing: The timing is exactly as in the model in the main text. In the first stage, each firm f draws private fixed portfolio cost shocks $\{\nu_j^g\}$ for each of their 3 potential products. Upon observing this private information, firms choose which products to introduce in their portfolio. Next, they draw private fixed market entry shocks $\{\nu_{jm}^e\}_{j \in \mathcal{G}^f, m}$ for each product and choose how many of the products in their portfolio to offer in each market.

Best response: Firm f's best response in Stage 2 depends on portfolio size N_p^f. Let $FE_{mn}^{(k)}$ denote the kth order statistic among the n potential entry cost draws in market m, and let $FD^{(k)}$ denote the kth order statistic among 3 product development cost draws.

If $N_p^f = 1$, the firm sells in market m iff $\mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m1}^{(1)} \geq 0$.

If $N_p^f = 2$,

$$
\begin{cases}
2 \text{ products iff } \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m2}^{(2)} \geq 0; \\
1 \text{ product iff } \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m2}^{(1)} \geq 0, \\
\quad \text{and } \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m2}^{(2)} < 0.
\end{cases}
$$

If $N_p^f = 3$,

$$
\begin{cases}
3 \text{ products iff } \mathbb{E}[\Pi_m^f(3, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - FE_{m3}^{(3)} \geq 0; \\
2 \text{ products iff } \mathbb{E}[\Pi_m^f(3, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - FE_{m3}^{(3)} < 0, \\
\quad \text{and } \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m3}^{(2)} \geq 0; \\
1 \text{ product iff } \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m3}^{(2)} < 0, \\
\quad \text{and } \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}] - FE_{m3}^{(1)} \geq 0.
\end{cases}
$$

In Stage 1, define

$$
\mathcal{V}_f(N_p^f) = \mathbb{E}\Big[\sum_m \max_{N_m^f \leq N_p^f} \mathbb{E}[\overline{\Pi}_m^f(N_m^f, N_m^{-f})|\mathcal{I}]\Big|\mathcal{I}\Big].
$$

The firm develops:

$$
\begin{cases}
3 \text{ products iff } \mathcal{V}_f(3) - \mathcal{V}_f(2) - FD^{(3)} \geq 0; \\[4pt]
2 \text{ products iff } \mathcal{V}_f(3) - \mathcal{V}_f(2) - FD^{(3)} < 0, \\
\quad \text{and } \mathcal{V}_f(2) - \mathcal{V}_f(1) - FD^{(2)} \geq 0; \\[4pt]
1 \text{ product iff } \mathcal{V}_f(2) - \mathcal{V}_f(1) - FD^{(2)} < 0, \\
\quad \text{and } \mathcal{V}_f(1) - FD^{(1)} \geq 0.
\end{cases}
$$

The equilibrium solution is summarized by three market thresholds:

$$
t_{1m}^e = \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}], \quad t_{2m}^e = \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(1, N_m^{-f})|\mathcal{I}],
$$
$$
t_{3m}^e = \mathbb{E}[\Pi_m^f(3, N_m^{-f})|\mathcal{I}] - \mathbb{E}[\Pi_m^f(2, N_m^{-f})|\mathcal{I}] \tag{80}
$$

and three product development thresholds,

$$
t_1^g = \mathcal{V}_f(1), \quad t_2^g = \mathcal{V}_f(2) - \mathcal{V}_f(1), \quad t_3^g = \mathcal{V}_f(3) - \mathcal{V}_f(2), \tag{81}
$$

where $\mathcal{V}_f$ integrates over rival and own product choices, determined by $(t_{1m}^e, t_{2m}^e, t_{3m}^e)$ in equilibrium. Since firms are symmetric, note that the thresholds are not firm-specific.

Computing equilibrium profits given threshold strategies: The probability that any firm offers n products, for $n \in \{1, 2, 3\}$, given the portfolio and market entry strategies, are given by,

$$
\begin{aligned}
p_1(\boldsymbol{t}^e, \boldsymbol{t}^g) =\ & \mathbb{P}(FD^{(3)} \leq t_3^g)[\mathbb{P}(FE_3^{(1)} \leq t_1^e) - \mathbb{P}(FE_3^{(2)} \leq t_2^e)] \\
& + [\mathbb{P}(FD^{(2)} \leq t_2^g) - \mathbb{P}(FD^{(3)} \leq t_3^g)][\mathbb{P}(FE_2^{(2)} \leq t_2^e) - \mathbb{P}(FE_2^{(1)} \leq t_1^e)] \\
& + [\mathbb{P}(FD^{(1)} \leq t_1^g) - \mathbb{P}(FD^{(2)} \leq t_2^g)][\mathbb{P}(FE_1^{(1)} \leq t_1^e)], \tag{82} \\
p_2(\boldsymbol{t}^e, \boldsymbol{t}^g) =\ & \mathbb{P}(FD^{(3)} \leq t_3^g)[\mathbb{P}(FE_3^{(2)} \leq t_2^e) - \mathbb{P}(FE_3^{(3)} \leq t_3^e)] \\
& + [\mathbb{P}(FD^{(2)} \leq t_2^g) - \mathbb{P}(FD^{(3)} \leq t_3^g)]\mathbb{P}[FE_2^{(2)} \leq t_2^e], \tag{83} \\
p_3(\boldsymbol{t}^e, \boldsymbol{t}^g) =\ & \mathbb{P}(FD^{(3)} \leq t_3^g)\mathbb{P}(FE_3^{(3)} \leq t_3^e). \tag{84}
\end{aligned}
$$

To compute expected profits in the first stage given any N_p^f, first realize that $\mathcal{V}_f$ takes the

following form,

$$\mathcal{V}_f(N_p^f) = \sum_m \sum_{i=1}^{N_p^f} \mathbb{P}(N_m^f = i | N_p^f)[\mathbb{E}[\Pi_m^f(i, N_m^{-f})|\mathcal{I}] - i\mathbb{E}[F_{jm}^e | F_{jm}^e \le \mathbb{E}[\Pi_m^f(i, N_m^{-f})|\mathcal{I}]]].$$

Equations (82)-(84) characterize $\mathbb{P}(N_m^f = i | N_p^f)$ as a function of $(\boldsymbol{t}^e, \boldsymbol{t}^p)$. Given expected variable profits, $\mathbb{E}[\Pi_m^f(i, N_m^{-f})|\mathcal{I}]$, the term $\mathbb{E}[F_{jm}^e | F_{jm}^e \le \mathbb{E}[\Pi_m^f(i, N_m^{-f})|\mathcal{I}]]$ is simple to compute numerically using Gaussian quadrature. I use the QuadGK package in Julia to compute this expectation.

To compute $\mathbb{E}[\Pi_m^f(i, N_m^{-f})|\mathcal{I}]$ for each i given $(\boldsymbol{t}^e, \boldsymbol{t}^g)$, I use the DSP package in Julia to perform a convolution which gives the probability distribution of the number of rival product offerings given the number of firms N and $p_n(\boldsymbol{t}^e, \boldsymbol{t}^g)$ for $n \in \{1, 2, 3\}$.

Solving for $(\boldsymbol{t}^e, \boldsymbol{t}^g)$: Given $(\boldsymbol{t}^e, \boldsymbol{t}^g)$, I showed how to evaluate $\mathcal{V}_f$ and $\mathbb{E}[\Pi_m^f(i, N_m^{-f})|\mathcal{I}]$ for any $i \in \{1, 2, 3\}$. Thus, I solve for $(\boldsymbol{t}^e, \boldsymbol{t}^g)$ by solving the non-linear system of equations (80)-(81). To do so, I use the NLsolve Julia package.

G.2 Behavior of Moment Inequalities Across DGPs

In this section, I study how the informativeness of the moment inequalities varies with parameters κ_o and κ_r and with the number of firms N. The informativeness of the inequalities depends on the loss from bounding the marginal value of introducing a product by evaluating it at extreme bundles and on the loss from applying Jensen's inequality to average out firms' expectational errors.

Implementation: I simulate $S = 500$ different realizations for each of $T = 12$ different "types" of global product and market entry game described in the previous section. In each game (s, t), there are N firms competing in 12 markets. I hold fixed the market-level profit shifters and set them at $A_m^{(s,t)} = 0.2mt$ for $m \in \{1, 2, ..., 12\}$, $t \in \{1, 2, ..., 12\}$, and $s \in \{1, 2, ..., 500\}$. Different values of t generate variation in profitability across different game types. In this subsection, I perform valid asymptotics as $ST \to \infty$. With $ST = 6000$, simulation noise is small, so I report the identified sets.

True parameters: I set the true parameters to be $(\theta_g, \sigma_g) = (3, 1)$ and $(\theta_e, \sigma_e) = (1, 1)$.

Instruments: For the Stage 2 inequalities, I condition on the realized set of portfolio decisions and construct instruments following the PPML procedure described in the main text. More precisely, I project the minimal (maximal) marginal values at the product-market level – evaluated at all (no) products in the portfolio being introduced in the market and at the realized set of rival offerings decisions – using specification $\gamma_0 \exp(\gamma_m^t A_m^{(s,t)})$. I then compute the predicted values of both regressions, sort them, and define 4 percentile

categories and bins associated with these categories. This gives 8 instruments in total – 4 for the upper bound inequality and 4 for the lower bound inequality.

For the Stage 1 inequalities, I implement an equivalent procedure as with the Stage 2 inequalities, but without conditioning on the realized product portfolios. I project the realized maximal and minimal (where now the minimal marginal value does not condition on the observed portfolio, so is evaluated at all 3 products being introduced in each market) marginal values on $\gamma_0 \exp(\gamma_m^t A_m^{(s,t)})$. I then compute the predicted values of both regressions and sum them across markets. Then, I sort the sums and define 4 percentile categories for both the lower and the upper bound and define 4 bins associated with these categories, yielding 8 instrument bins - 4 for the lower bound and 4 for the upper bound.

G.2.1   Baseline: $N = 10$, $\kappa_o = 0.1$, $\kappa_r = 0.1$

In this baseline case, I set the number of firms to $N = 10$ and set $\kappa_o = 0.1$ and $\kappa_r = 0.1$.

Figure G.1: Identified Sets
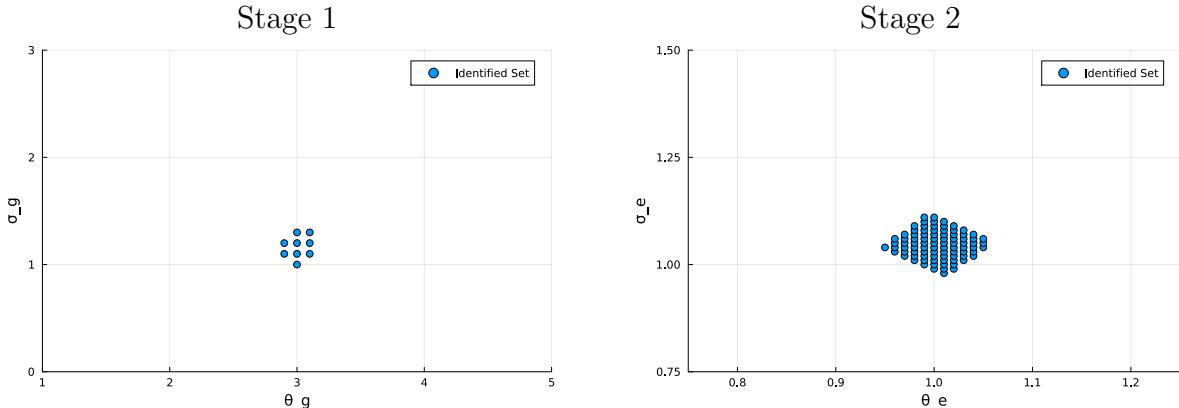
Figure G.1 shows the identified sets for (θ_g, σ_g) and (θ_e, σ_e), which I obtained via grid search. As seen in this figure, the identified set is quite informative both for the Stage 1 and the Stage 2 fixed cost parameters (note: the scale is not the same for the Stage 1 and the Stage 2 parameters).[5]

G.2.2   High Substitutability Within the Firm: $N = 10$, $\kappa_o = 0.25$, $\kappa_r = 0.1$
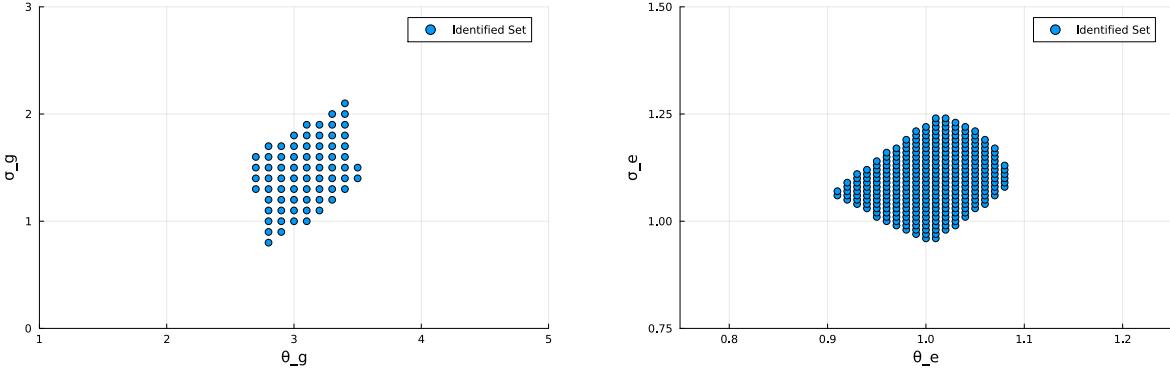
I now study a deviation from the baseline case in which the parameter determining substitution within the firm is larger.

Figure G.2 shows the results. Compared to the baseline cases, high substitution within the firm reduces the informativeness of the fixed cost parameter bounds, both for the Stage

---

[5]In Additional Materials C I also report the lower contours for the moment inequalities.

Figure G.2: Identified Sets

Stage 1 — Stage 2
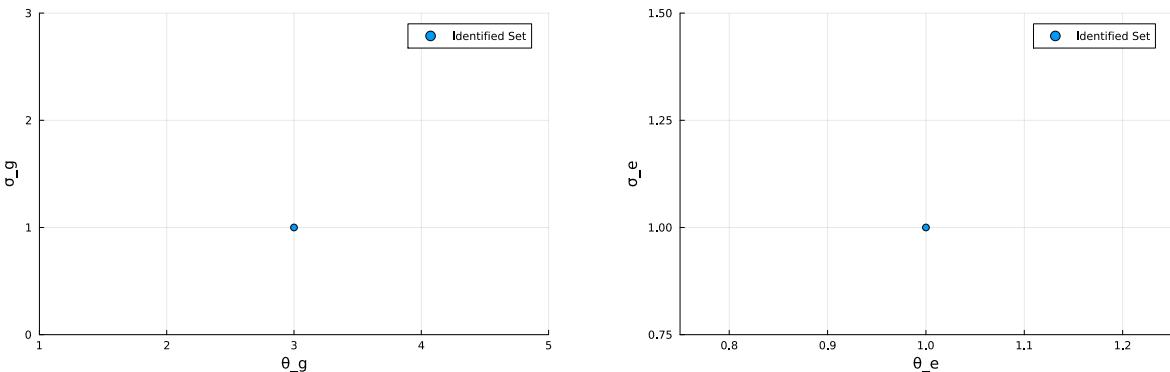
1 and for the Stage 2 fixed cost parameters. This is expected given that the loss from
bounding the marginal value of introducing a product at extreme bundles (all other products
vs. no other products offered) is greater when such products are highly substitutable. That
is, greater cannibalization within the firm makes the moment inequalities less informative
about the true parameters.

G.2.3 Low Substitutability within the Firm: $N = 10$, $\kappa_o = 0.01$, $\kappa_r = 0.1$

Figure G.3 show what happens when there is very low substitutability across the firms'
own products. Interestingly, the tightness of the inequalities increases greatly so that only
the true fixed cost parameters both in Stage 1 and Stage 2 are accepted (given my grid).
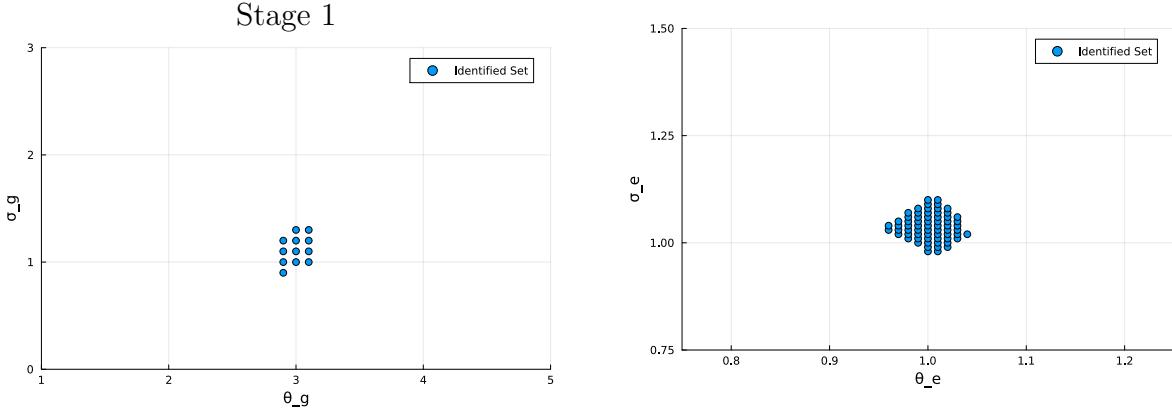
This result is of special significance for anyone wanting to use this estimation approach
in a setting where firms are single-product (rather than multi-product). Indeed, Figure G.3
shows that in the absence of cannibalization within the firm, my estimation procedure is
highly informative.

Figure G.3: Identified Sets

Stage 1 — Stage 2

G.2.4 High Substitutability Across Firms: $N = 10$, $\kappa_o = 0.1$, $\kappa_r = 0.25$

Figure G.4: Identified Sets

Interestingly, high substitution across firms does not seem to change the tightness of the inequalities much relative to the baseline case. While it seems to slightly reduce the informativeness of the Stage 1 inequalities, it does not seem to lead to reduced informativeness in Stage 2. If anything, in this case, the Stage 2 inequalities yield tighter bounds on θ_e. The impact of higher κ_r on the informativeness of the inequalities is therefore not as quantitatively important as the impact of κ_o. This makes sense because, by virtue of the unobservability of rivals' fixed cost shocks, I showed that one can use the realized set of rival offerings decisions to construct the moment inequalities. Greater substitution across firms reduces the rate of product introduction in equilibrium, but firms expect this, and the informativeness of the moment inequalities is not significantly affected.

G.2.5 Low Substitutability Across Firms $N = 10$, $\kappa_o = 0.1$, $\kappa_r = 0.01$

For completeness, Figure G.5 reports the identified sets for the case in which there is low substitution across rival firms' products. As expected, there are no substantial differences in the informativeness of the moment inequalities relative to the baseline case for the same reasons as in Section G.2.4.

G.2.6 Large Number of Firms: $N = 20$, $\kappa_o = 0.1$, $\kappa_r = 0.1$

Figure G.6 illustrates the effect of having more firms participating in the global entry game. The identified set becomes slightly smaller when there are more firms. This is due to two effects. First, when there are more firms, each firm makes smaller expectational errors. Second, when there are more firms, more products may be offered on average, which reduces the loss from bounding marginal values with extreme bundles due to submodularity.

G.2.7   Small Number of Firms: $N = 2$, $\kappa_o = 0.1$, $\kappa_r = 0.1$

Figure G.7 illustrate the effect of having only 2 firms competing globally. In this case, the inequalities become less informative for the same reasons as argued in Section G.2.6.

G.2.8   Main Takeaways

While the simulations abstract away from heterogeneity across products and across firms, they are still useful for understanding some of the key properties of the moment inequalities. I have established that much of the informativeness of the moment inequalities relies on the extent to which products *within* the firm are substitutable. High degrees of substitutability within the firm render the moment inequalities less informative, while lower degrees of substitutability make them tighter. Moreover, the number of firms (relative to the number of products per firm) matters. With smaller firms relative to the overall market, expectational errors are smaller. If the set of potential products of a firm is small relative to the set of products that are offered in the market, the loss from bounding a product's marginal value with extreme bundles within the firm is smaller. Finally, I showed that all else equal,

Figure G.5: Identified Sets
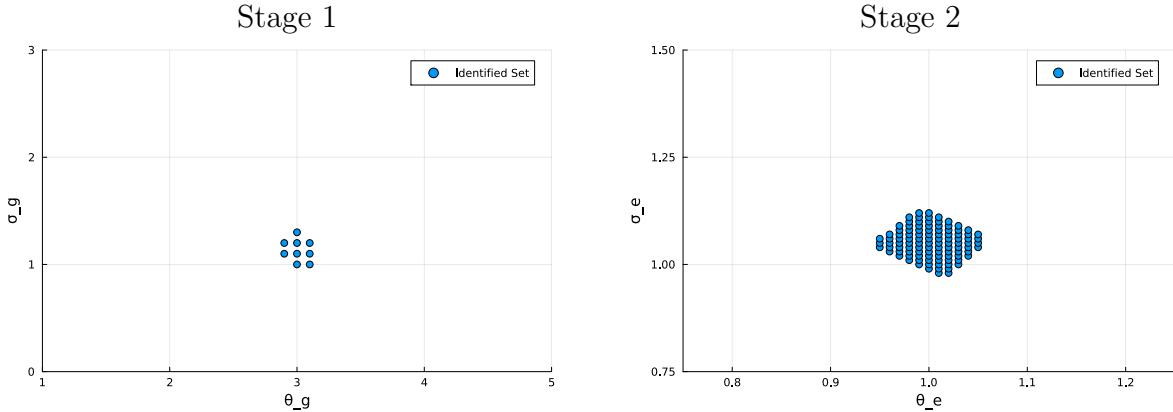
Figure G.6: Identified Sets
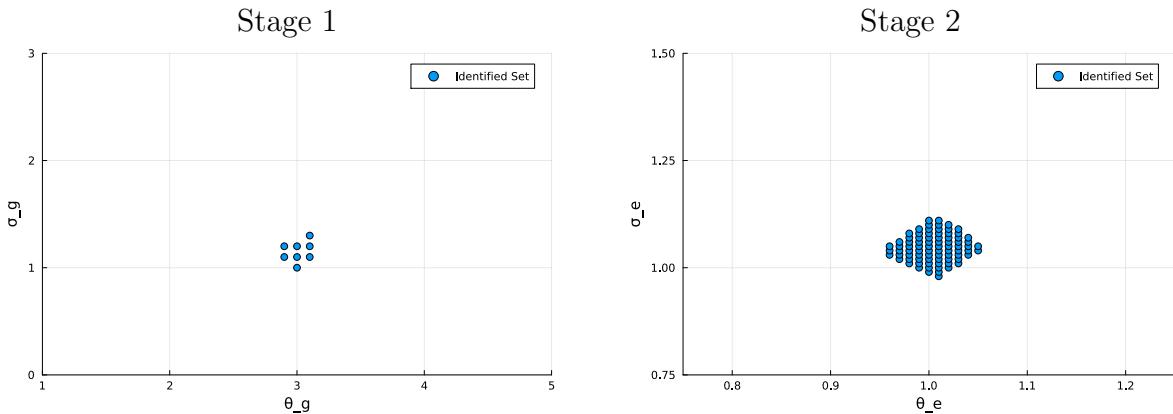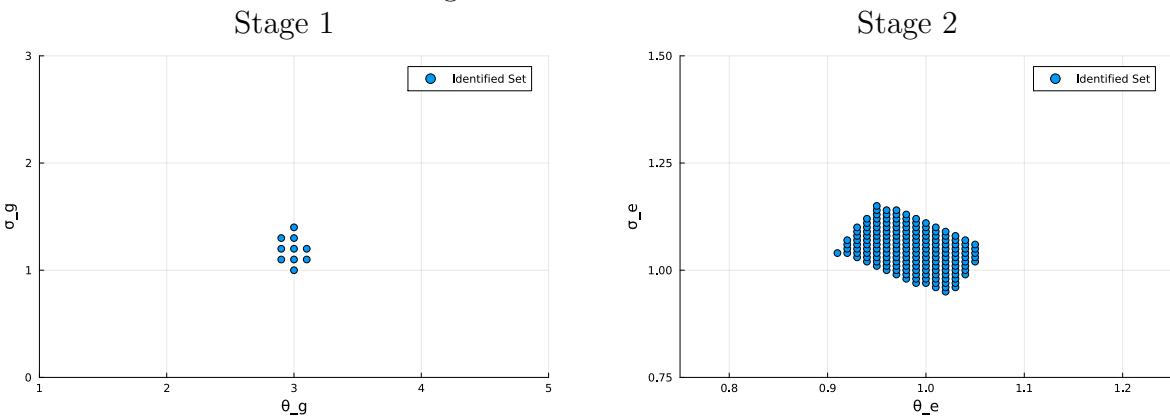
Figure G.7: Identified Sets

substitution across firms does not have much of an effect on the tightness of the moment inequalities.

G.3 Inference Under a Single Realization of Global Product Entry Game

In this section, I use the fully solvable version of the model to assess the properties of Andrews and Soares (2010) confidence sets in my setting. I simulate $S = 100$ realizations for each of $T = 12$ different "types" of global product and market entry games, just as in Section G.2. In each game (s, t), there are N firms competing in 12 markets. I hold fixed the market-level profit shifters and set them at $A_m^{(s,t)} = 0.2mt$ for $m \in \{1, 2, ..., 12\}$, $t \in \{1, 2, ..., 12\}$ and $s \in \{1, 2, ..., 100\}$. Different values of t generate variation in profitability across different game types.

True parameters: I set the true parameters to be $(\theta_g, \sigma_g) = (3, 1)$ and $(\theta_e, \sigma_e) = (1, 1)$.

Instruments: I construct the instruments as described in Section G.2. However, I run the PPML regressions at the (s, t)-level to mimic the actual implementation in the main text, where I only observe a single realization of the cross-section. Thus, the variation used to construct confidence sets in this section is across product-market pairs.

For each (s, t) pair, I construct confidence sets for parameters (θ_e, σ_e) using the procedure in Andrews and Soares (2010). For each of the $S \times T$ 95% confidence sets, I record: (i) whether the true values $(1, 1)$ are included in the confidence set (coverage), (ii) the length of the confidence set along the θ_e dimension (holding $\sigma_e = 1$ at the truth), (iii) the length of the confidence set along the σ_e dimension (holding $\theta_e = 1$ at the truth).

Panel A in Table G.1 reports the average coverage of the confidence set across all $S \times T$ realizations of the global product entry game. I do this both under a robust variance-covariance matrix and a clustered (market-level) variance-covariance matrix. Undercoverage

34

can occur, particularly when the number of firms is relatively small. As the number of firms increases, the coverage of the Andrews and Soares (2010) confidence sets tends to increase. Note that clustering does not necessarily yield higher coverage, though this is the case when the number of firms is large ($N = 75$).

Table G.1: Confidence Set Properties for (θ_e, σ_e)

SE Type	$N = 5$	$N = 25$	$N = 50$	$N = 75$
Panel A: Coverage of True Parameters (%)				
Robust	92.9	93.8	94.1	93.7
Clustered	92.4	94.1	94.1	95.0
Panel B: Median Length of Confidence Set				
Robust	(0.8, 2.8)	(0.4, 0.9)	(0.2, 0.6)	(0.2, 0.5)
Clustered	(0.9, 4.4)	(0.4, 1.6)	(0.3, 1.1)	(0.3, 0.9)

Note: Panel A reports the average coverage across all $S \times T$ simulations of the confidence sets for (θ_e, σ_e). Panel B reports the median length of these confidence sets along each dimension. The first coordinate corresponds to the θ_e dimension conditional on σ_e being at its true value, and the second to the σ_e dimension conditional on θ_e being at its true value. Confidence sets are computed at the (s, t) level, as in the empirical application.
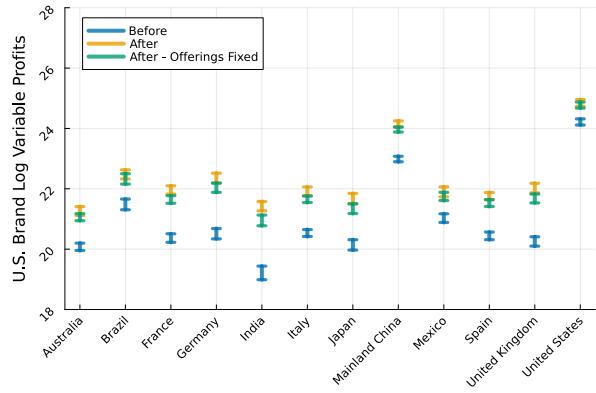
Panel B in Table G.1 reports the median length of the confidence set along each of the dimensions of (θ_e, σ_e), conditional on the confidence set not being empty. The confidence sets are smaller whenever there are more firms and without clustering. In the empirical implementation, I report confidence sets using a robust variance-covariance matrix because (i) Table G.1 (Panel A) shows that clustering does not necessarily improve coverage and undercoverage is not severe, and (ii) Table G.1 (Panel B) shows that clustering leads to a larger confidence set.

A caveat in this exercise is that the simulations require symmetry across firms and products within a (s, t) pair. While at Stage 2, even in the simulations, firms are not fully symmetric due to variation in how many products they have in their global product portfolio, greater symmetry relative to reality should worsen correlation across expectational errors since the realization of the market structure affects firms in (almost) the exact same way. However, greater symmetry also means that there are no "large" firms in the sample. Large firms with a lot of market power can deteriorate the asymptotic properties of the estimator.
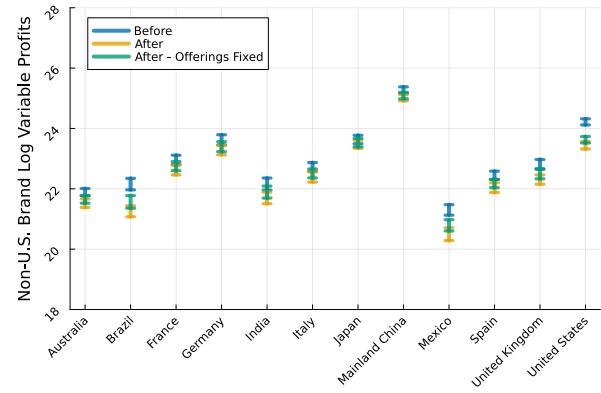
H Counterfactual Exercises - Variable Profit Plots

Figure H.1: Variable Profits under U.S. Policy Interventions
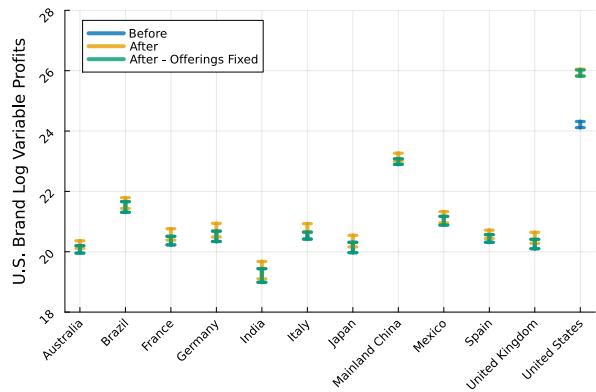
Panel A: U.S. Products – Marginal Cost Subsidy
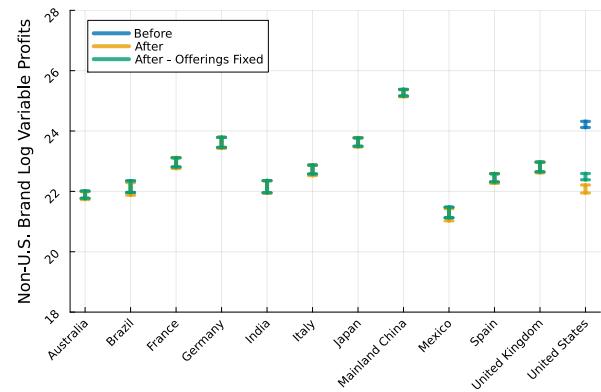
Panel B: Non-U.S. Products – Marginal Cost Subsidy

Panel C: U.S. Products – Consumer Subsidy

Panel D: Non-U.S. Products – Consumer Subsidy

Note: Panels A–B display bounds on the expected (log) total variable profits of U.S. and non-U.S. brands across countries before (blue) and after (orange) a 20% reduction in U.S. brands' marginal costs. Panels C–D show analogous bounds for a 50% consumer subsidy on U.S.-branded products in the United States. Expectations are taken with respect to bounds on the probability distribution of firms' offerings and demand and marginal cost shocks. The green intervals use the bounds on the distribution of product offerings in each market before the policy is implemented.

References

Donald W. K. Andrews and Gustavo Soares. Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection. *Econometrica*, 78(1):119–157, 2010.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–890, 1995.

Christopher Conlon and Jeff Gortmaker. Best Practices for Differentiated Products Demand Estimation with PyBLP. *The RAND Journal of Economics*, 51(4):1108–1161, 2020.

Amit Gandhi and Jean-François Houde. Measuring Substitution Patterns in Differentiated-Products Industries. *Mimeo*, 2019.

Paul L E Grieco, Charles Murry, and Ali Yurukoglu. The Evolution of Market Power in the U.S. Automobile Industry. *The Quarterly Journal of Economics*, page qjad047, 09 2023.

Jerry Hausman. Valuation of New Goods under Perfect and Imperfect Competition. In *The Economics of New Goods*, pages 207–248. National Bureau of Economic Research, Inc, 1996.

W. Ross Morrow and Steven J. Skerlos. Fixed-Point Approaches to Computing Bertrand-Nash Equilibrium Prices Under Mixed-Logit Demand. *Operations Research*, 59(2):328–345, April 2011.

Amil Petrin. Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy*, 110(4):705–729, 2002.

Charly Porcher, Eduardo Morales, and Thomas Fujiwara. Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach. *Mimeo*, May 2024.

Moshe Shaked. *Stochastic Orders*. Springer Science & Business Media, 2007.