Abduction and the Demand Curve

Brian C. Albrecht James Traina

April 2026

Policy counterfactuals such as merger simulation require the demand curve at a specific market's realized conditions, not an average across markets. We characterize when the experimental average coincides with the market-specific curve and what structural estimation adds. The experimental average slope equals every market's demand slope if and only if demand is additively separable in price and the latent state. Linear demand satisfies this condition; standard discrete-choice models violate it. When separability fails, structural estimation identifies the market-specific demand curve if and only if the observed data identify the market's latent demand index; Berry inversion recovers this index from observed shares and prices—the abduction step in Pearl's causal hierarchy. We prove invertibility is necessary, not merely sufficient: without it, even price-only counterfactuals are set-identified. In a merger simulation, market-specific price predictions differ by a factor of two, driven entirely by unobserved demand conditions that experiments cannot distinguish.

Keywords: Demand estimation, abduction, causal hierarchy, counterfactual identification, additive separability

JEL Codes: C14, C36, C51, D12, L13

Albrecht: International Center for Law & Economics, mail@briancalbrecht.com. Traina: NYU Stern Abu Dhabi, james.traina@nyu.edu. We thank Alberto Bisin for helpful comments.

1. Introduction

A market's demand curve holds its unobserved conditions (quality, local tastes, recent advertising) fixed and varies price. Two markets that share observed characteristics but differ in these conditions face different demand curves. Policy counterfactuals such as merger simulation, tax pass-through, new-product introduction, and welfare analysis require this market-specific object, not an average across markets.

An experiment that randomizes prices reveals the average demand response at each price, but not which latent state generated any particular market's observed outcome. The experiment eliminates confounding between price and unobservables; it does not recover the unobservables themselves. Berry and Haile (2021, p. 11) put it directly: "A LATE averages over the latent variables; this is not the same thing as holding them fixed." This paper asks when the two objects coincide and what bridges the gap when they do not.

The slopes of the two objects coincide if and only if demand is additively separable in price and the unobserved state (Theorem 1). Separability means the unobserved state shifts demand levels but does not change price sensitivity; once the latent state interacts with price, the experimental average diverges from every market's slope. Even under separability, population-average elasticities differ from unit-level elasticities because the demand level itself varies with the latent state. Linear demand with additive shocks satisfies this condition. Standard discrete-choice models violate it because the implied demand shares are nonlinear in the latent demand index. Berry and Haile (2021) conclude that experiments "generally" do not identify demand; the theorem makes "generally" exact.

When separability fails—the empirically relevant case—structural estimation can recover the market-specific demand curve. It does so if and only if the observed data identify the market's latent demand state (Corollary 1). The Berry (1994) share inversion recovers this state from observed shares, prices, and characteristics (Theorem 2); counterfactuals that also change product characteristics require stronger recoverability conditions. Berry and Haile (2014) give sufficient conditions for identification of the demand system. We prove the converse. Invertibility is necessary for price-only counterfactuals, and additional recoverability is needed when counterfactuals change characteristics (Proposition 3).

Together the two characterizations give a practitioner a diagnostic. Check additive separability; if it fails (it will), check invertibility. The gap is not a theoretical curiosity; it is a quantitative fact about the workhorse models the profession uses. A merger calibration in Section 7 shows that two markets differing only in unobserved quality produce predicted merger price increases of 36% and 71%—a factor-of-two difference driven entirely by the latent state.

Existing work characterizes each boundary with sufficient conditions, not exact ones. Angrist, Graddy, and Imbens (2000) show that in a simultaneous equations model without additive residuals, linear IV identifies a weighted average of market-specific demand derivatives; with additive residuals, the slope of the average demand function equals every market's slope. Berry and Haile (2021) note that experiments "generally" do not identify demand. We prove the converse. If the experimental derivative matches every market's derivative, demand must be additively separable in price and the latent state. This completes the iff and generalizes beyond the additive-residual form $q_t^d(p) = q^d(p) + \varepsilon_t^d$ (no observables, scalar shock) to the full separable form $D(x, p, \xi) = f(x, p) + g(x, \xi)$ with observable characteristics and multidimensional latent states. Heckman and Pinto (2024) study related sufficient conditions for structural models more broadly. Berry and Haile (2014, 2018) establish identification conditions for demand systems and broader classes of nonseparable models; the question of which unit-level counterfactuals these conditions deliver has remained open. In treatment effects, Kline and Walters (2019) resolve apparent structural-vs-IV disagreements by showing both yield equivalent LATE estimates; any disagreement found is about the estimand, not the method.

Several recent papers approach the same boundary from complementary directions. Chen (2025) recasts the Berry–Haile framework in potential-outcomes language and shows that counterfactual homogeneity is necessary for point identification of unit-level counterfactuals; our results work within the maintained index/inversion framework and characterize which conditions are necessary for price-only versus characteristics-changing counterfactuals (Online Appendix E relates the two approaches). Borusyak et al. (2026) show that market-specific price counterfactuals require stronger identifying variation than average price effects, even though they require less than full demand identification. Andrews et al. (2025) addresses what comes after: once separability fails and the analyst must use structural estimation, the resulting counterfactuals remain trustworthy under misspecification only with sufficiently strong excluded-variable variation.

The formal language we use to make the gap precise comes from structural causal models (Pearl 2009).[1] Pearl (2009) poses three queries on the supply-demand model that illustrate what is now known as the *causal hierarchy* (Bareinboim et al. 2022): interventional questions (what experiments answer) versus counterfactual questions (what structural estimation answers). Moving from one to the other requires an *abduction* step; Berry inversion performs exactly this step. The hierarchy provides the objects that make the gap precise enough to characterize; the identification relies on functional-form conditions (additive separability, invertibility) that the graph alone cannot express. Section 2 introduces the demand model. Section 3 defines the two objects and constructs a simple example where they diverge. Section 4 characterizes exactly when the gap vanishes; Section 5 shows how structural estimation bridges it when it does not. These sections are self-contained. Section 6 translates the results into Pearl's causal hierarchy; Section 7 quantifies the gap with a merger calibration.

2. The Demand Model

This section introduces the primitives. The model is standard; we include it to fix notation.

2.1. Markets, units, and latent demand conditions

We study a market for differentiated products (or a single homogeneous good as a special case). A *market* is a product-market-time cell t.[2] Each market has observed characteristics X_t (product attributes, consumer demographics, income) and an unobserved demand shifter ξ_t that captures everything affecting demand that the econometrician does not observe—unobserved product quality, local tastes, recent advertising. The pair (X_t, ξ_t) constitutes the market's *demand conditions*. Price P_t is determined in equilibrium by the interaction of demand and supply conditions; we impose no restriction on the pricing mechanism. Quantity demanded (or market share) is $Q_t = D(X_t, P_t, \xi_t)$, where D is the structural demand function. In the discrete-choice framework, D is obtained by integrating individual consumers'

[1] Hünermund and Bareinboim (2025) survey how tools from Pearl's framework—do-calculus, data fusion, transportability—can be applied in econometric settings; Huntington-Klein (2022) reviews the reception of Pearl's causal program in empirical economics.

[2] In the BLP tradition, a "market" is a geographic area in a given time period, and a "product" is a particular model or brand available in that market. We use t as a generic market index.

choice probabilities over the distribution of consumer types, so that ξ is a market-level object even though consumers are heterogeneous.[3]

The demand counterfactual "what would this market demand at price p" holds the market's demand conditions (X_t, ξ_t) fixed and replaces the pricing equation, regardless of what supply-side conditions generated the observed price. The demand system is triangular: X and ξ are exogenous; P is set by demand and supply; $Q = D(X, P, \xi)$ is the outcome. The demand equation is *structural* in Haavelmo's (1943) sense: it remains invariant under interventions on P (see also Pearl 2015, for the connection to causal calculus). Section 6 develops the formal SCM interpretation; the identification results do not require it.

3. Two Distinct Objects

Berry and Haile (2021, p. 7) distinguish "average responses of demand to price changes" from "the levels and slopes of demand at specific points." The former is what experiments deliver. Set a price, average over markets. The latter is what merger simulation, pass-through, and welfare analysis require. The paper centers on two formal objects that correspond to this distinction. Throughout, we take the structural demand system as identified from cross-market variation; the identification question is whether unit-level counterfactuals are additionally point-identified for a specific market.

DEFINITION 1 (Market-specific demand curve). *For a market with observed characteristics $\bar{x} \in \mathcal{X}$ and unobserved demand conditions $\bar{\xi} \in \Xi$, the demand curve is*

$$q_{\bar{x},\bar{\xi}}(p) \equiv D(\bar{x}, p, \bar{\xi}), \quad p \in \mathcal{P}.$$

This holds the market's demand conditions $(\bar{x}, \bar{\xi})$ fixed and maps price to quantity demanded.

DEFINITION 2 (Experimental average). *The average quantity demanded when price*

[3] The treatment effects literature developed the prediction–planning distinction that motivates this analysis. Heckman and Vytlacil (2005) unified the two approaches through the marginal treatment effect (MTE), Vytlacil (2002) proved they impose identical restrictions on counterfactual data, and Kline and Walters (2019) showed that structural and reduced-form estimators agree on local average treatment effects but diverge when targeting different parameters—disagreements arise from the choice of estimand, not functional form alone.

is set to p across all markets with observed characteristics x is

$$m(p,x) \equiv \int D(x, p, \xi)\, dF_{\xi|X=x}(\xi).$$

A randomized price experiment that sets $P = p$ for a population of markets with characteristics x identifies $m(p,x)$. A valid instrument for price identifies m up to the instrument's first-stage variation.

The demand curve $q_{\bar{x},\bar{\xi}}(p)$ holds the market's unobserved conditions fixed. The experimental average $m(p,x)$ integrates over them. Merger simulation, pass-through, and welfare analysis require the demand curve: each operates at a specific market's realized conditions, not an average.[4] An experiment gives m. The question is when these two objects coincide.

Two markets that share the same observed characteristics x but differ in unobserved demand conditions ξ have the same experimental average $m(p,x)$ but different demand curves $q_{x,\xi}(p)$. The example below shows the gap is real. Even complete experimental knowledge cannot determine which demand curve generated the observed market. Section 4 gives the exact condition for the gap to vanish. Section 5 shows how structural estimation recovers the demand curve when it does not.

3.1. A simple example

We construct the simplest model exhibiting the gap. Two linear demand curves with different slopes cross at the observed equilibrium price. The proposition below formalizes the three claims.

PROPOSITION 1 (Constructive Rung 2–3 gap). *Suppress X (equivalently, condition on a fixed $X = x$ throughout). Assume the price support $\mathcal{P}$ contains more than one point. Consider the demand model with*

$$D(p, \xi) = \alpha(\xi) - \beta(\xi)\, p, \quad \beta(\xi) > 0,$$

[4] Pass-through, merger simulation, and welfare are equilibrium counterfactuals that take the demand curve as an input. They differ across markets because the relevant objects (incidence ratios, diversion ratios, compensating variation) are nonlinear functions of the demand slope, and the slope varies across markets whenever (AS) fails. Online Appendix A develops these three applications in detail.

and suppose $\xi \in \{\xi_1, \xi_2\}$ with $\beta(\xi_1) \neq \beta(\xi_2)$. Write $D(p', \xi_k)$ for the counterfactual demand of a market with $\xi = \xi_k$.

(a) ***Observational equivalence.*** *There exist parameter values such that two markets with different latent types generate the same observed data (Q^*, P^*).*

(b) ***Counterfactual divergence.*** *For every $p' \neq P^*$, $D(p', \xi_1) \neq D(p', \xi_2)$.*

(c) ***Experimental insufficiency.*** *For every $p' \neq P^*$, the identified set $\mathcal{I}(p'; e)$ (Definition 3) contains more than one point for a market observed at $e = (Q^*, P^*)$.*

PROOF. (a) The two demand curves cross at $P^* = (\alpha(\xi_1) - \alpha(\xi_2))/(\beta(\xi_1) - \beta(\xi_2))$, yielding common quantity $Q^* = \alpha(\xi_1) - \beta(\xi_1)P^*$. Concretely: $\alpha(\xi_1) = 12$, $\beta(\xi_1) = 3$, $\alpha(\xi_2) = 6$, $\beta(\xi_2) = 1$ gives $P^* = 3$, $Q^* = 3$.

(b) The crossing condition $D(P^*, \xi_1) = D(P^*, \xi_2)$ implies $\alpha(\xi_1) - \alpha(\xi_2) = (\beta(\xi_1) - \beta(\xi_2))P^*$, so $D(p', \xi_1) - D(p', \xi_2) = (\beta(\xi_1) - \beta(\xi_2))(P^* - p') \neq 0$ for all $p' \neq P^*$, since $\beta(\xi_1) \neq \beta(\xi_2)$ by assumption. Concretely: at $p' = 4$, $D(4, \xi_1) = 0$ and $D(4, \xi_2) = 2$.

(c) The evidence is $e = (Q^* = 3, P^* = 3)$. Two latent types, ξ_1 and ξ_2, are both consistent with e by part (a). The average demand $m(p, x) = \pi D(p, \xi_1) + (1 - \pi)D(p, \xi_2)$, where $\pi \equiv \mathbb{P}(\xi = \xi_1 \mid X = x)$, is fully determined by $(\alpha(\cdot), \beta(\cdot), \pi)$ and does not depend on which type generated e: knowledge of $m(p, x)$ for all p cannot distinguish ξ_1 from ξ_2 at the observed evidence. By part (b), $D(p', \xi_1) \neq D(p', \xi_2)$ for every $p' \neq P^*$; so both types are consistent with e yet produce distinct counterfactuals. Hence $\mathcal{I}(p'; e)$ contains more than one point for any $p' \neq P^*$. □

Figure 1 illustrates the geometry.

Experiments succeed at what they target: identifying $m(p, x)$ at every p. But even experimentally set prices do not resolve the gap. Both types generate $(Q^* = 3, P^* = 3)$ regardless of how $P^* = 3$ was assigned. The experiment eliminates confounding between P and ξ—the identification problem that instruments address—but does not determine which latent type generated the data. The gap is in price sensitivity; at the crossing point, the two types' slopes differ by a factor of three. Two markets that look identical from experimental data respond differently at every other price. A demand curve is a market-specific object.

In Online Appendix A, pass-through and pre-tax consumer surplus differ by factors of 2 and 3 respectively across the same two markets. A merger authority using the experimental average would predict the wrong price change in both markets, overstating it where ξ is low and understating it where ξ is high. The error is not noise; it is systematic and predictable from the theory.

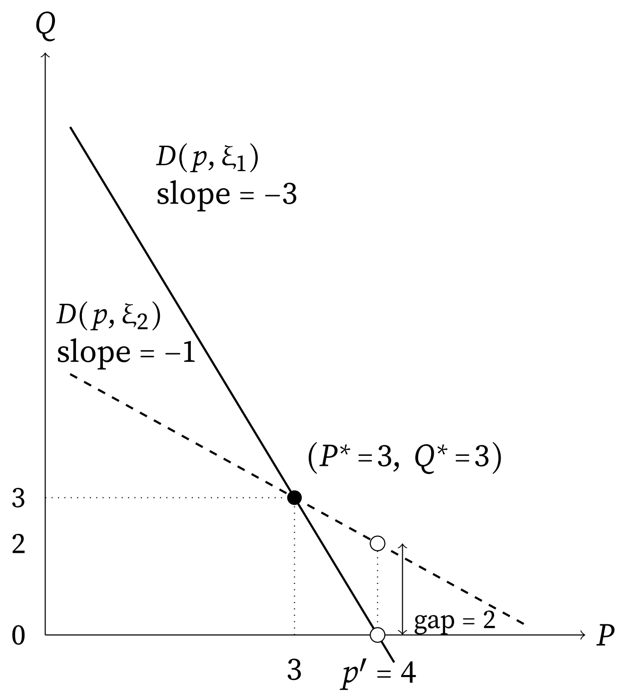

FIGURE 1. Two demand curves that cross at $(P^*, Q^*) = (3, 3)$. At the observed price, both types produce the same quantity. At the counterfactual price $p' = 4$, type ξ_1 (solid, steep) demands $Q = 0$ while type ξ_2 (dashed, flat) demands $Q = 2$. A randomized experiment identifies the experimental average $m(p, x)$ at every p, yet cannot determine which demand curve belongs to the observed market.

The constructive example uses a discrete two-type model; Proposition 2 (Section 4) establishes the same gap for any continuously differentiable demand function with slope heterogeneity. The question is not whether the gap exists (it does, generically, whenever price sensitivity varies with the latent state) but what functional form condition separates models where the gap vanishes from those where it does not.

4. When Does Experimental Identification Suffice?

Experimental information identifies $m(p, x)$, not the demand curve $q_{\bar{x},\bar{\xi}}(p)$ of any specific market. When exactly do they coincide?

4.1. The identification gap

In demand the gap takes a concrete form. We observe one equilibrium outcome per market, $e = (Q^*, P^*, X^*)$, and want the entire response function $p \mapsto q_{X^*,\bar{\xi}}(p)$ for that market's realized $\bar{\xi}$.

DEFINITION 3 (Identified set for counterfactual demand). *Fix the demand function D and observed evidence $e = (Q^*, P^*, X^*)$. The identified set for counterfactual demand at price p', given knowledge of $m(p, x)$ for all (p, x), is*

$$\mathcal{I}(p'; e) \;\equiv\; \left\{ D(X^*,\, p',\, \xi) \;:\; D(X^*, P^*, \xi) = Q^*, \quad \xi \in \operatorname{supp}(\xi \mid X^*) \right\}.$$

The counterfactual $D(X^*, p', \xi)$ is point-identified from experimental information and evidence e if $\mathcal{I}(p'; e)$ is a singleton.

The identified set collects all counterfactual demands at p' across latent states ξ consistent with the observed evidence at the equilibrium price.[5] Non-identification means $|\mathcal{I}| > 1$: distinct latent states consistent with the same evidence yield different counterfactual demands.

4.2. The gap is generic

Non-identification holds for any continuously differentiable demand function with slope heterogeneity, not just the two-type case.

PROPOSITION 2 (General Rung 2–3 non-identification). *Let $D : \mathcal{P} \times \Xi \to \mathbb{R}$ be continuously differentiable in p and suppress X. Suppose:*

- *(i)* ***Crossing with slope heterogeneity.*** *There exist ξ_a, ξ_b in the support of ξ and a price $P^* \in \mathcal{P}$ such that $D(P^*, \xi_a) = D(P^*, \xi_b)$ and $\frac{\partial D}{\partial p}(P^*, \xi_a) \neq \frac{\partial D}{\partial p}(P^*, \xi_b)$.*
- *(ii)* ***Support richness.*** *Both ξ_a and ξ_b have positive probability (or positive density) conditional on $X = x$.*

Then:

- *(a)* ***Local non-identification.*** *There exists a punctured neighborhood of P^* on which $D(p', \xi_a) \neq D(p', \xi_b)$. Consequently, for evidence $e = (Q^*, P^*)$ with $Q^* = D(P^*, \xi_a) = D(P^*, \xi_b)$, the identified set $\mathcal{I}(p'; e)$ contains more than one point for every p' in that punctured neighborhood.*

[5] The standard definition of "identification" in econometrics—uniqueness of the structural parameter across DGPs consistent with the reduced form—is a question about the demand function D. Definition 3 takes D as given and asks the sharper question: is the counterfactual at a specific market's conditions uniquely determined by the evidence?

(b) ***Divergence rate.*** *The width of the identified set satisfies*

$$w(p') \equiv |D(p', \xi_a) - D(p', \xi_b)| = \left|\frac{\partial D}{\partial p}(P^*, \xi_a) - \frac{\partial D}{\partial p}(P^*, \xi_b)\right| \cdot |p' - P^*| + o\big(|p' - P^*|\big).$$

PROOF. See Appendix B. □

Condition (i) combines two requirements, observational equivalence at P^* (the negation of the abduction condition in Section 5) and slope heterogeneity at P^* (the negation of (AS), as characterized by Theorem 1). Neither alone suffices—crossing without slope heterogeneity means counterfactuals coincide; slope heterogeneity without crossing means experimental information can distinguish the types at the observed price.[6] Part (b) gives the size of the gap. To first order, the identified set widens linearly with distance from P^*, at a rate equal to the absolute slope difference; higher-order terms may widen or narrow the true width. For a merger that raises price by 10%, the first-order width of the identified set for the counterfactual quantity is approximately 10% of the absolute slope difference. In the crossing-curves example ($\beta(\xi_1) = 3$, $\beta(\xi_2) = 1$), a 10% price change produces a counterfactual gap of 0.6 units, roughly 20% of the observed quantity. The question is not whether the gap exists, but how large it is.

4.3. When the gap vanishes

Proposition 1 and Proposition 2 show that the gap exists. When exactly does it vanish? The answer is a functional-form condition. The gap vanishes in derivatives if and only if D is additively separable in (p, ξ), the demand analogue of "no essential heterogeneity" in the treatment effects literature (Heckman and Vytlacil 2005). No prior result establishes that this condition is both necessary and sufficient.

THEOREM 1 (When does the Rung 2–3 gap vanish?). *Let $D : \mathcal{X} \times \mathcal{P} \times \Xi \to \mathbb{R}$ be the demand function, continuously differentiable in p. For part (b), assume Assumptions A4 and A2 in Appendix B.*

(a) ***Levels.*** *$m(p, x) = D(x, p, \bar{\xi})$ for all $\bar{\xi}$ in the support of $\xi \mid X = x$ and all $p \in \mathcal{P}$, if and only if $D(x, p, \xi)$ is constant on the support of $\xi \mid X = x$ for each p.*

[6] Different slopes alone do not guarantee a crossing on the price support; the crossing condition requires non-invertibility of $D(P^*, \cdot)$ at the observed quantity, which is a separate condition from slope heterogeneity. With multi-dimensional ξ ($J > 1$), non-injectivity is the generic case.

(b) ***Marginal price effects.*** $\dfrac{\partial m}{\partial p}(p,x) = \dfrac{\partial D}{\partial p}(x,p,\bar{\xi})$ *for all* $\bar{\xi}$ *in the support of* $\xi \mid X = x$ *and all* $p \in \mathcal{P}$*, if and only if* D *is additively separable:*

$$\text{(AS)} \qquad D(x,p,\xi) = f(x,p) + g(x,\xi)$$

for some functions f *and* g*, for all* ξ *in the support of* $\xi \mid X = x$*, simultaneously for all* $p \in \mathcal{P}$*.*

PROOF. (a) ($\Rightarrow$, equality implies separability): Fix p and x. Let $h(\xi) \equiv D(x,p,\xi)$. The hypothesis says $\mathbb{E}[h(\xi) \mid X = x] = h(\bar{\xi})$ for every $\bar{\xi}$ in the support. Since $\mathbb{E}[h(\xi) \mid X = x]$ is a constant c, we have $h(\bar{\xi}) = c$ for all $\bar{\xi}$, so $D(x,p,\cdot)$ is constant on the support. Since p was arbitrary, D does not vary with ξ. ($\Leftarrow$, separability implies equality): Immediate.

(b) ($\Rightarrow$, slope homogeneity implies separability): Fix (x,p) and let $h(\xi) \equiv \partial D(x,p,\xi)/\partial p$. The hypothesis says $\mathbb{E}[h(\xi) \mid X = x] = h(\bar{\xi})$ for every $\bar{\xi}$ in the support. By the argument of part (a), h is constant on the support of $\xi \mid X = x$; call this constant $\phi(x,p)$. Since p was arbitrary, $\partial D(x,p,\xi)/\partial p = \phi(x,p)$ for all ξ in the support and all $p \in \mathcal{P}$. Integrate:

$$D(x,p,\xi) = \int_{p_0}^{p} \phi(x,t)\,dt + D(x,p_0,\xi),$$

where $p_0 \in \mathcal{P}$ is any fixed reference price; the integral is well-defined because $\mathcal{P}$ is an interval (Assumption A4). Setting $f(x,p) \equiv \int_{p_0}^{p} \phi(x,t)\,dt$ and $g(x,\xi) \equiv D(x,p_0,\xi)$ establishes (AS). ($\Leftarrow$, separability implies slope homogeneity): If $D = f + g$, then $\partial D/\partial p = \partial f/\partial p$ does not depend on ξ, and the conclusion follows (Assumption A2 ensures the dominated convergence step in the integration). $\square$

The characterization has a natural decomposition into level shifts and slope shifts (Section 4.4). Section 6.5 shows why linear demand with additive shocks—Pearl's running example—satisfies (AS) and thus obscures the gap.

Theorem 1(b) sharpens an observation of Angrist, Graddy, and Imbens (2000). Without additive residuals, the linear IV estimand in a simultaneous equations model is a weighted average of market-specific demand derivatives, with weights that depend on the instrument. Their Corollary 3 shows that the additive-residual model $D(x,p,\xi) = f(x,p) + \varepsilon(x,\xi)$ collapses this weighted average to a single slope—the slope of the average demand function, which then equals every market's slope. Angrist, Graddy, and Imbens (2000) prove the sufficient direction.

Theorem 1(b) adds the converse. If the experimental derivative equals every market's derivative, (AS) must hold. The iff characterization makes Berry and Haile's (2021) "generally" exact. Heckman and Pinto's (2024) P3–P4 policy problems—forecasting modified or never-implemented policies—are exactly the problems that require market-specific identification; Theorem 1 gives the functional-form condition they leave unspecified.

The experimental average $m(p, x)$ gives the distribution of $D(x, p, \xi)$ at each p separately (marginals, not the joint). Recovering a specific market's demand curve requires matching that market's response *across* prices, and without (AS) this matching generally fails. The gap between any two markets varies with price, so knowing their difference at one price does not pin down their difference at another. Merger analysis cannot rely on average demand responses because the merger occurs in a specific market with a specific $\bar{\xi}$, and the predicted price change depends on $\partial D/\partial p$ at $\bar{\xi}$.

Additive separability is highly restrictive: any (p, ξ) interaction moves demand off the separable subspace. Random-coefficients logit has additively separable utility, but the market shares that result from aggregation are not additively separable in (p, ξ), because the logit transformation creates the interaction. Utility-level separability is standard; demand-level separability (AS) is the knife-edge. Allen and Rehbeck (2019) show that when *utility* is additively separable in observables and unobservables, utility indices and average welfare are nonparametrically identified from conditional means alone. Their separability is a restriction on the primitive (the utility function); ours is a restriction on the solution (the demand function).[7] Section 4.5 shows that every standard discrete-choice system violates (AS).

[7]Whether unobservables enter additively is a central boundary in nonparametric identification (Matzkin 2007). When the structural function is strictly monotone in a scalar unobservable, Torgovitsky (2015) and Chernozhukov and Hansen (2005) show point identification without separability—monotonicity provides a different path to market-specific identification. Imbens and Newey (2009) obtain a similar result via control functions. In the partial identification setting, Tebaldi, Torgovitsky, and Yang (2023) estimate demand nonparametrically without distributional assumptions on unobservables. Borusyak et al. (2026) show that price counterfactuals require less identifying variation than full demand identification but more than average price effects; Online Appendix E develops the connection.

4.4. Level shifts versus slope shifts

The gap binds when the latent state affects the *slope* of demand; a pure level shift preserves (AS) and the gap closes. Write $D(x, p, \xi) = \bar{D}(x, p) + \eta(x, p, \xi)$, where $\bar{D}(x, p) \equiv \mathbb{E}[D(x, p, \xi) \mid X = x]$ is the population-average demand and η captures the deviation of each market from the average. If η does not depend on p, meaning the latent state shifts demand up or down without altering the slope, then $D(x, p, \xi) = \bar{D}(x, p) + \eta(x, \xi)$, which is additive separability (AS), and the identification gap vanishes in derivatives. If $\partial\eta/\partial p \neq 0$ for some ξ, the latent state affects how responsive demand is to price, and the gap binds.

4.5. Why workhorse demand models fail separability

Every workhorse discrete-choice demand model violates additive separability, because choice probabilities are nonlinear functions of the latent demand index. In the random-coefficients logit, the own-price derivative $\partial\sigma_j/\partial p_j$ is a weighted average of α_i over individual choice probabilities π_{ij}, both of which depend on δ and hence on ξ. Higher mean utility shifts the composition of marginal consumers, changing the derivative even at the same price—slope heterogeneity in the sense of Section 4.4. The dependence survives in the homogeneous logit, the nested logit, and CES (Appendix A gives the formulas). By Theorem 1(b), separability fails in every discrete-choice model where the choice probabilities are nonlinear functions of mean utilities. The failure is not specific to discrete choice; for any structural function $Y = g(X, P, \xi)$, condition (AS) requires $\partial^2 g/\partial p\,\partial\xi = 0$, which fails whenever g is nonlinear in ξ in a way that interacts with p.

The gap extends to elasticities: even under (AS), population-average elasticities generically differ from unit-level elasticities because the denominator D depends on ξ (Appendix D). The Berry and Haile (2021, p. 11) critique—"a ratio of averages is not equal to the average ratio"—is a consequence. If the same market is observed at multiple exogenously assigned prices while ξ remains fixed, the demand curve is directly estimable; but this panel shortcut rarely applies (see the end of Appendix D).

The gap is strict and generic outside the knife-edge. If demand is additively separable, the gap does not arise: the experimental derivative already equals every market's derivative. If separability fails, the analyst must recover the market's latent demand index from observed data. Theorem 1 characterizes the first exit;

Theorem 2 provides the second. The two results are complements: one says when the gap is zero, the other says how to cross it when it is not. In practice, crossing the gap means identifying which demand curve, among all those consistent with the population data, generated the observed market. The structural model provides exactly this map.

5. Bridging the Gap: The Structural Model as Abduction

When additive separability fails, experimental information does not determine the demand curve of any specific market. Structural estimation bridges the gap by recovering the market's demand index from observed data. The procedure has three steps.

a. **Abduction.** Recover the market's demand index δ^* from observed shares, prices, and characteristics. If the counterfactual also changes characteristics, separately recover the underlying latent demand state $\bar{\xi}$.
b. **Action.** Replace the pricing equation with $P \leftarrow p'$, holding the demand function and the market's latent state fixed.
c. **Prediction.** Evaluate counterfactual demand at the recovered state—market-specific, not population-average.

Berry inversion performs step 1. Given abduction, steps 2 and 3 are mechanical. Standard BLP-style merger analyses already execute this procedure—estimate $\hat{\theta}$ from cross-market moments, invert each market's observed shares to recover δ^*, then evaluate counterfactual shares at the new price. We prove that the conditions for this procedure to deliver point identification are tight.

5.1. Berry inversion as abduction

THEOREM 2 (The structural model bridges the Rung 2–3 gap). *Consider the demand model with $Q = D(X, P, \xi)$. Suppose:*

(C1) ***Index structure.*** *There exists a function σ known up to a finite-dimensional parameter $\theta \in \Theta$ and a latent demand index $\delta : \mathcal{X} \times \Xi \to \mathbb{R}^J$ such that $D(x, p, \xi) = \sigma(\delta(x, \xi; \theta), p, x; \theta)$.*[8]

[8] "Known up to θ" means the functional form of σ is specified; only the parameter θ is unknown. The results extend to the semiparametric case where σ is nonparametrically specified, as in Berry and Haile (2014), since the identification argument uses the structure of σ at the true parameter value.

(C2) ***Invertibility.*** *For each (p, x), the map $\delta \mapsto \sigma(\delta, p, x; \theta)$ is invertible, so observed shares Q^* admit a unique $\delta^* = \sigma^{-1}(Q^*, p, x; \theta)$.*

(C3) ***Recoverability.*** *For each x, the mapping $\xi \mapsto \delta(x, \xi; \theta)$ is injective, so there exists a known (up to θ) function r with $\xi = r(\delta, x; \theta)$.*

Then for any market observed at (Q^*, P^*, X^*) (with $Q^* \equiv s^*$ in the share-vector case):

Step 1 (Abduction):

$$\delta^* = \sigma^{-1}(Q^*, P^*, X^*; \theta), \qquad \bar{\xi} = r(\delta^*, X^*; \theta).$$

Step 2 (Action): Replace the pricing equation with $P \leftarrow p'$.

Step 3 (Prediction):

$$D(\bar{x},\ p',\ \bar{\xi}) = \sigma\big(\delta^*,\ p',\ X^*;\ \theta\big).$$

If the counterfactual also changes characteristics to x', then $\delta' = \delta(x', \bar{\xi}; \theta)$ and $D(x',\ p',\ \bar{\xi}) = \sigma(\delta', p', x'; \theta)$.

PROOF. Observed quantities satisfy $Q^* = \sigma(\delta^*, P^*, X^*; \theta)$ with $\delta^* = \delta(X^*, \bar{\xi}; \theta)$ by C1. By C2, $\delta^* = \sigma^{-1}(Q^*, P^*, X^*; \theta)$. By C3, $\bar{\xi} = r(\delta^*, X^*; \theta)$. This completes abduction. Under the intervention $P \leftarrow p'$, the demand function remains $Q = \sigma(\delta, p, x; \theta)$ (structural invariance). For a price-only counterfactual, X and ξ (hence δ^*) are unchanged, so $D(\bar{x}, p', \bar{\xi}) = \sigma(\delta^*, p', X^*; \theta)$. For a counterfactual that also changes characteristics to x', the recovered $\bar{\xi}$ is used to form $\delta' = \delta(x', \bar{\xi}; \theta)$, and $D(x', p', \bar{\xi}) = \sigma(\delta', p', x'; \theta)$. □

An ideal experiment identifies the interventional demand response—the mapping from imposed prices to average quantities. Structural estimation adds the market's position on it. Proposition 3 shows the conditions above are required, not merely convenient. Without invertibility, even price-only counterfactuals are set-identified; without recoverability, characteristics-changing counterfactuals remain ambiguous even when δ^* is unique.

Conditions C1–C2 are jointly sufficient for the recovery map from (Q^*, P^*, X^*) to a latent index δ^* to be single-valued. The inversion step in demand estimation, familiar since Berry (1994), turns a single equilibrium observation into an inferred latent state.[9] The Berry–Haile identification conditions—which build on Matzkin's

[9] In the multinomial logit, the inversion is closed-form: $\delta_{jt} = \log s_{jt} - \log s_{0t}$. Berry and Haile (2021) call this step a "trick" yielding equations each with one demand shock. Theorem 2 makes precise what the trick achieves: recovery of this market's latent demand index.

(2008) nonparametric identification framework for nonseparable models and are extended to micro data in Berry and Haile (2024)—are the conditions for this inversion to be well-defined. Both characterize when inversion identifies the demand system; we show their conditions are also necessary for market-specific counterfactuals (Corollary 1), and Theorem 1 characterizes separately when inversion is needed at all.

Learning $\sigma(\cdot;\theta)$ from cross-market variation is the task of experimental identification—it uses population variation to pin down the demand function. Evaluating $D(\bar{x}, p', \bar{\xi})$ market by market is structural identification. The parameters θ come from cross-market moments; each market's δ^* comes from inverting that market's observed shares given θ. No averaging over ξ occurs at the prediction stage.

Online Appendix B traces the three-step procedure through a logit market with $J = 3$ products, computing counterfactual shares step by step. Appendix C discusses the identification content of each condition, and Online Appendix C establishes their necessity.

PROPOSITION 3 (Necessity of abduction conditions). *Consider the demand model and notation of Theorem 2.*

(a) ***Failure of invertibility.*** *If $\sigma(\cdot, p, x; \theta)$ is not injective at (p^*, x^*) and the pre-image elements are not observationally equivalent—i.e., there exist $\delta_1 \neq \delta_2$ in $\sigma^{-1}(\{Q^*\}, p^*, x^*; \theta)$ with $\sigma(\delta_1, p', x^*; \theta) \neq \sigma(\delta_2, p', x^*; \theta)$ for some p'—and both δ_1, δ_2 lie in the range of $\xi \mapsto \delta(X^*, \xi; \theta)$ for $\xi \in \operatorname{supp}(\xi \mid X^*)$, then the identified set $\mathcal{I}(p'; e)$ contains more than one point.*

(b) ***Failure of recoverability.*** *If $\xi \mapsto \delta(x, \xi; \theta)$ is not injective and the pre-image elements produce distinct demand indices at some counterfactual characteristics—i.e., there exist $\xi_1 \neq \xi_2$ in $\operatorname{supp}(\xi \mid X^*)$ with $\delta(x^*, \xi_k; \theta) = \delta^*$ and $\delta(x', \xi_1; \theta) \neq \delta(x', \xi_2; \theta)$ for some x'—then even when δ^* is point-identified, counterfactuals $D(x', p', \bar{\xi})$ are set-identified.*

Proof of part (a). Under C1, $D(X^*, P^*, \xi) = Q^*$ iff $\sigma(\delta(X^*, \xi; \theta), P^*, X^*; \theta) = Q^*$, so any ξ with $\delta(X^*, \xi; \theta) \in \{\delta_1, \delta_2\}$ is consistent with the evidence. By the range hypothesis, such $\xi_1, \xi_2 \in \operatorname{supp}(\xi \mid X^*)$ exist with $\delta(X^*, \xi_k; \theta) = \delta_k$; both types produce Q^* at P^* but differ at p'. Hence the identified set satisfies

$$\mathcal{I}(p'; e) \supseteq \left\{\sigma(\delta_1, p', x^*; \theta),\ \sigma(\delta_2, p', x^*; \theta)\right\}.$$

At any p' where $\sigma(\delta_1, p', x^*; \theta) \neq \sigma(\delta_2, p', x^*; \theta)$, this set contains at least two distinct points. $\square$

Proof of part (b). Suppose C2 holds, so δ^* is unique, but there exist $\xi_1 \neq \xi_2$ with $\delta(x^*, \xi_k; \theta) = \delta^*$ for $k = 1, 2$. For a price-only counterfactual (holding x^* fixed), δ^* is sufficient, and $D(\bar{x}, p', \bar{\xi}) = \sigma(\delta^*, p', x^*; \theta)$ is point-identified regardless of which ξ generated δ^*. But for a counterfactual that changes characteristics to $x' \neq x^*$, the new demand index is $\delta' = \delta(x', \xi; \theta)$, which depends on ξ directly. By the non-degeneracy hypothesis, there exist such ξ_1, ξ_2 with $\delta(x', \xi_1; \theta) \neq \delta(x', \xi_2; \theta)$ for some x'. Different unobserved types that happened to produce the same δ^* at x^* produce different demand indices at x'. Conversely, any ξ consistent with the observed evidence satisfies $\delta(x^*, \xi; \theta) = \delta^*$ (since C2 pins down δ^* uniquely), so the identified set is exactly $\{\sigma(\delta(x', \xi; \theta), p', x'; \theta) : \delta(x^*, \xi; \theta) = \delta^*\}$. This set is sharp: every element is attainable by some $\xi \in \operatorname{supp}(\xi \mid X^*)$ consistent with the evidence, because the constraint $\delta(x^*, \xi; \theta) = \delta^*$ with $\xi \in \operatorname{supp}(\xi \mid X^*)$ is both necessary and sufficient for consistency under C2.[10] $\square$

COROLLARY 1 (Share inversion and point identification are equivalent). *Under the demand model with* C1 *maintained, and subject to the non-degeneracy qualifications of Proposition 3:*

(a) *Price-only counterfactuals $D(\bar{x}, p', \bar{\xi})$ are point-identified from experimental information and evidence e if and only if* C2 *holds.*

(b) *Characteristics-changing counterfactuals $D(x', p', \bar{\xi})$ for a fixed x' are point-identified if and only if* C2 *holds and the image of the set of latent states consistent with the observed market under $\xi \mapsto \delta(x', \xi; \theta)$ is a singleton—i.e., all $\xi \in \operatorname{supp}(\xi \mid X^*)$ satisfying $\delta(x^*, \xi; \theta) = \delta^*$ yield the same $\delta(x', \xi; \theta)$. Global* C3 *is sufficient for all x' simultaneously (since it makes the consistent set itself a singleton) but stronger than necessary for a given x'.*

PROOF. The "if" direction of (a): if C2 holds, share inversion recovers δ^* uniquely, so $\sigma(\delta^*, p', x^*; \theta)$ is point-identified. The "if" direction of (b): if C2 holds and all ξ consistent with the observed market yield the same $\delta(x', \xi; \theta)$, then $\delta^* = \delta(x^*, \bar{\xi}; \theta)$ uniquely determines $\delta(x', \bar{\xi}; \theta)$, so $D(x', p', \bar{\xi})$ is point-identified. For the "only if" direction of (a), Proposition 3(a) applies directly. For (b), if C2 fails then price-only counterfactuals are already set-identified; if C2 holds but the image is not

[10] This is why the "new product" counterfactual—predicting demand for a product that does not yet exist—requires both C2 and C3. The new product has characteristics $x' \neq x^*$ for any existing product, so forming $\delta(x', \bar{\xi}; \theta)$ requires knowing $\bar{\xi}$, not just δ^*. This is the additional identification content of C3 over C2.

a singleton, there exist $\xi_1 \neq \xi_2$ consistent with the evidence with $\delta(x', \xi_1; \theta) \neq \delta(x', \xi_2; \theta)$, so $D(x', p', \bar{\xi})$ is set-identified. □

Three distinct objects are in play. The demand system (σ, θ) is identified from cross-market variation. The market-specific index δ^* is recovered by inversion (C2). The underlying latent state $\bar{\xi}$ is separated from δ^* by recoverability (C3), and matters only when the counterfactual changes product characteristics. For a price-only counterfactual, δ^* suffices. Changing price does not change δ, so $\sigma(\delta^*, p', x^*; \theta)$ is point-identified. For a counterfactual that also changes characteristics, the new index $\delta' = \delta(x', \bar{\xi}; \theta)$ depends on $\bar{\xi}$ directly, and two markets with the same δ^* but different $\bar{\xi}$ will diverge.

The crossing-curves example and Berry inversion are the same problem seen from opposite sides. Without abduction, two demand curves pass through the observed (Q^*, P^*) and the counterfactual is indeterminate. With abduction, inversion selects one curve from the crossing pair. Corollary 1 closes an open loop. Berry and Haile (2014) and Berry, Gandhi, and Haile (2013) establish sufficient conditions for demand identification, but neither establishes what is lost without invertibility. The necessity direction answers this. Without C2, even price-only counterfactuals are set-identified; for characteristics-changing counterfactuals, point identification requires all latent states consistent with the observed market to yield the same demand index at the counterfactual characteristics. The identified set widens as invertibility weakens. This is an identification result, not a restatement of Berry inversion in different language.

In standard discrete-choice models with connected substitutes, $\sigma(\cdot, p, x; \theta)$ is globally injective (Berry 1994; Berry, Gandhi, and Haile 2013), so C2 holds and price-only counterfactuals are point-identified. The non-degeneracy qualifications in Proposition 3 hold universally in standard parametric families: for logit, nested logit, and CES, $\sigma(\cdot, p, x; \theta)$ is analytic in δ, so distinct pre-images $\delta_1 \neq \delta_2$ produce distinct shares at all but isolated prices—the non-degeneracy condition $\sigma(\delta_1, p', x^*; \theta) \neq \sigma(\delta_2, p', x^*; \theta)$ for some p' is guaranteed. When invertibility does fail, the identified set widens linearly with distance from the observed price (Proposition 2(b); Online Appendix C develops the geometry), paralleling the sharp bounds of Mogstad, Santos, and Torgovitsky (2018) and Chesher and Rosen (2017).

The three-step procedure applies to any structural model $Y = h(\delta(\varepsilon; \theta), X; \theta)$ where the index is invertible; Section 8 develops examples beyond demand.

5.2. Averaging at the input stage versus the output stage

In practice, θ is estimated by pooling data across markets. GMM moment conditions $\mathbb{E}[Z'\xi] = 0$ average over markets (Berry, Levinsohn, and Pakes 1995), and estimation consistency relies on a law of large numbers across markets. This cross-market averaging uses population variation to learn the demand function $\sigma(\cdot;\theta)$. But once $\hat{\theta}$ is in hand, each market's δ^* is recovered by inversion (C2), and counterfactuals are evaluated market-by-market, $D(\bar{x}, p', \bar{\xi}) = \sigma(\delta^*, p', x^*; \hat{\theta})$. No averaging over ξ occurs at the counterfactual stage.

The experimental average $m(p', x)$ averages at the *output* stage, since the target itself is a population average. Structural estimation averages at the *input* stage. Learning θ requires cross-market variation, but evaluation is unit by unit at the output stage. This distinction clarifies a common confusion. Critics of structural estimation sometimes object that the model's predictions "depend on functional form." They are right about the demand function. Its shape depends on the specification. But the market-specific content of structural estimation is not the shape of the demand function; it is the location of a particular market on it. That location comes from data, not from functional form.

Estimation uncertainty in $\hat{\theta}$ propagates through the inversion at the standard $\sqrt{T}$ rate but does not change the identification analysis (Remark A1 in Appendix C). When C1–C3 are relaxed, unit-level counterfactuals are set-identified, with width governed by slope heterogeneity as in Proposition 2(b) (Remark A2 in Appendix C).

Before turning to the merger calibration, we translate the identification analysis into the language of Pearl's causal hierarchy.

6. Connection to Pearl's Causal Hierarchy

The results in Sections 3–5 are self-contained IO identification results. This section develops their precise interpretation within Pearl's (2009) structural causal model framework. The translation delivers three things: a formal answer to Pearl's (2023) open question about whether ceteris paribus demand can be defined in counterfactual language (Proposition 4); an exact correspondence between Berry inversion and Pearl's abduction step (Section 6.4); and a diagnosis of why Pearl's own running example obscures the gap (Section 6.5). The identification theorems do not require this framework.

6.1. The demand model as a structural causal model

The demand system consists of four equations. Observed characteristics X and demand shock ξ are determined by separate exogenous processes, independent of each other and of supply conditions. Price P is set by the interaction of demand and supply; no restriction is placed on the pricing mechanism. Quantity demanded Q is the outcome of the structural demand function $D(X, P, \xi)$, deterministic conditional on ξ, with sampling variability absorbed into ξ. This is a structural causal model (Pearl 2009, Ch. 7).[11] The exogenous variables are $U = (U_X, U_\xi, U_P)$, the endogenous variables are $V = (X, \xi, P, Q)$, and the structural functions are:

$$X = g_X(U_X), \tag{1}$$

$$\xi = g_\xi(U_\xi), \tag{2}$$

$$P = g_P(X, \xi, U_P), \tag{3}$$

$$Q = D(X, P, \xi). \tag{4}$$

We write Q generically; in single-good examples $Q \in \mathbb{R}$, while in differentiated-products applications Q is the vector of market shares $s \in \Delta^J$. The functions g_X and g_ξ are unrestricted and play no role in demand estimation.

Equations (1)–(4) form a triangular system; the exogenous variables (U_X, U_ξ, U_P) are mutually independent in the standard BLP specification (Berry, Levinsohn, and Pakes 1995), and each endogenous variable is determined by its causes in the model. The demand equation (4) is *structural* in Haavelmo's (1943) sense. It remains invariant under interventions on P. Changing how prices are set does not alter how consumers respond to prices.

6.2. Three rungs of causal knowledge

The causal hierarchy for demand has three strictly nested layers, each requiring information unavailable at the rung below.

Rung 1 (Observation). The joint distribution of observed variables:

$$\mathcal{L}_1 \equiv \mathbb{P}(Q, P, X).$$

[11] Each structural function is *autonomous*: it can be modified independently of the others, which gives the do$(\cdot)$ operator its meaning.

This includes correlations, conditional expectations, and regressions—but not the causal effect of price on demand, since P and ξ are correlated through the pricing equation (Pearl 2009, Ch. 1).

Rung 2 (Intervention). The complete family of interventional distributions.

DEFINITION 4 (Rung 2 information set). *The Rung 2 information set is the collection of conditional distributions*

$$\mathcal{L}_2 \equiv \left\{F_{Q|\mathrm{do}(P=p),X=x} : p \in \mathcal{P},\ x \in \mathcal{X}\right\},$$

where $\mathcal{P} \subseteq \mathbb{R}$ is the price support and $\mathcal{X}$ is the support of observed characteristics.[12]

The $\mathrm{do}(\cdot)$ operator replaces equation (3) with $P = p$, holding all other equations fixed. The population mean under intervention is

$$m(p,x) \equiv \mathbb{E}[Q \mid \mathrm{do}(P = p), X = x] = \int D(x, p, \xi)\, dF_{\xi|X=x}(\xi),$$

which is precisely the experimental average m defined in Section 3 (Definition 2).

Rung 2 answers population questions: “If we set price to p across all markets with characteristics x, what is the distribution of quantity demanded?” Randomized experiments, valid instruments, and do-calculus all target Rung 2 objects.[13] In the potential outcomes framework (Rubin 1974), Rung 2 knowledge is equivalent to knowing the marginal distributions of potential outcomes $Y(p)$ under all treatment values.

The Rung 2–3 distinction is orthogonal to the selection-on-observables versus selection-on-unobservables distinction; both are Rung 2 assumptions. Whether identification comes from random assignment, instrumental variables, or a selection correction, the identified object is still a population average over latent types. Potential outcomes alone do not explain how to reach Rung 3 from Rung 2. That requires the structural function D and the abduction step, which is the content of the SCM that potential outcomes abstract away.

[12] “Knowing $\mathcal{L}_2$” means knowing the full conditional distribution of Q given $\mathrm{do}(P = p)$ and $X = x$, for every (p, x) in the support—not merely the conditional mean or a finite collection of moments. This is the strongest possible Rung 2 assumption; our impossibility results hold a fortiori when only partial Rung 2 information is available.

[13] Do-calculus (Pearl 2009) derives interventional distributions from observational data and a causal graph; its identification algorithms are complete for Rung 2 queries (Shpitser and Pearl 2006), not Rung 3.

Rung 3 (Counterfactual). The unit-level counterfactual response function:

$$\mathcal{L}_3 \equiv \{Q_p(u) : p \in \mathcal{P},\ u \in \mathcal{U}\}, \qquad Q_p(u) = D\big(g_X(u_X),\ p,\ g_\xi(u_\xi)\big).$$

For a specific market u, $Q_p(u)$ is the quantity demanded at price p when all exogenous variables are held at their realized values. It is deterministic given u—a consequence of the SCM structure, not an additional assumption. The demand curve is $p \mapsto Q_p(u)$ for a fixed u, which is precisely $q_{\bar{x},\bar{\xi}}(p) = D(\bar{x}, p, \bar{\xi})$ from Definition 1.

PROPOSITION 4 (The demand curve is a Rung 3 object). *The ceteris paribus demand curve $q_{\bar{x},\bar{\xi}}(p)$ is a Rung 3 object: it is the unit-level counterfactual $Q_p(u)$ for a market with realized conditions $(\bar{x}, \bar{\xi})$. The experimental average $m(p, x)$ is a Rung 2 object: it integrates $Q_p(u)$ over the population distribution of market conditions.*

Pearl (2023) notes that the econometric literature lacks a formal definition of counterfactuals in terms of structural equations. Proposition 4 supplies one for demand: the ceteris paribus demand curve is the unit-level counterfactual $Q_p(u)$ for a fixed u, which is precisely Rung 3 in Pearl's hierarchy.

REMARK 1 (Demand curve versus policy counterfactuals). *In the SCM framework, the demand curve $Q_p(u)$ holds the pricing equation fixed at $P = p$ but keeps all other equations intact. A policy counterfactual, such as "what is the welfare effect of a price floor," may require modifying equation (3) in a way that changes the distribution of ξ—for example, if the floor causes some products to exit. The demand curve and the policy counterfactual are distinct objects; Rung 3 covers only the former.*

6.3. The hierarchy is strict

The first separation ($\mathcal{L}_1 \not\Rightarrow \mathcal{L}_2$) is the identification problem that occupies most of the causal inference literature; observational distributions do not determine interventional distributions without additional assumptions. In the demand setting, P is correlated with ξ through the pricing equation (3): high-quality products command both high prices and high demand, biasing the naive price coefficient toward zero. Instrumental variables, natural experiments, and randomized pricing address this—Berry and Haile's (2021, p. 4) "first fundamental challenge"—by providing exogenous variation in P that is independent of ξ, delivering Rung 2 objects like $m(p, x) = \mathbb{E}[Q \mid \mathrm{do}(P = p), X = x]$ from Rung 1 data.

The second separation ($\mathcal{L}_2 \not\Rightarrow \mathcal{L}_3$): complete knowledge of all interventional distributions does not, in general, determine unit-level counterfactuals. This is the identification gap of Sections 3–4 translated into Pearl's language. The obstacle is that a single equilibrium observation (Q^*, P^*, X^*) does not reveal which latent state $\bar{\xi}$ generated it—multiple values of ξ are consistent with both the evidence and $\mathcal{L}_2$ (the posterior over U given the evidence is non-degenerate, in Pearl's terminology). Proposition 1 exhibits two latent types within a single model that share identical Rung 2 objects (m) but have different demand curves; it demonstrates that the gap is structural, not an artifact of incomplete data.

The general Causal Hierarchy Theorem (Bareinboim et al. 2022) establishes these separations for generic SCMs; our proofs are self-contained and specific to the demand setting, yielding exact conditions for when the gap vanishes (Theorem 1) and when it can be bridged (Theorem 2). The underlying structure is a marginals-versus-joint distinction: $\mathcal{L}_2$ specifies the distribution of Q under each $\mathrm{do}(P = p)$ separately, but the demand curve concerns a single market's response across all prices simultaneously. Two latent types that share the same price-by-price marginals can have different demand curves—Proposition 1 exhibits exactly this.

6.4. Share inversion as abduction

In Pearl's framework, moving from Rung 2 to Rung 3 requires three steps: *abduction* (update beliefs about U given evidence), *action* (set the counterfactual intervention), and *prediction* (evaluate the outcome). Theorem 2 is exactly this three-step procedure:

Step 1 (Abduction): From observed (Q^*, P^*, X^*), recover $\bar{\xi}$ via C1–C3. The posterior over U collapses to a point: $\bar{\xi} = r(\sigma^{-1}(Q^*, P^*, X^*; \theta), X^*; \theta)$.

Step 2 (Action): Replace equation (3) with $P \leftarrow p'$, holding all other equations fixed.

Step 3 (Prediction): Evaluate $Q_{p'}(u) = \sigma(\delta^*, p', X^*; \theta)$.

The conditions C1–C3 are not computational conveniences; they are the conditions under which abduction has a unique solution—the demand-model counterpart of the identification conditions in Berry and Haile (2014). Although the Berry inversion is often introduced as an estimation step, the demand-identification literature treats the associated index and invertibility restrictions as substantive conditions for recovering the latent demand state from market-level data. C1–C2

TABLE 1. Correspondence between Pearl's SCM framework and structural demand estimation.

Pearl's SCM	Structural econometrics	This paper
Unit u	Entity with realized ε	Market with (X_t, ξ_t)
$Y_x(u)$	$g(\bar{x}, \bar{\varepsilon})$	$D(\bar{x}, p, \bar{\xi})$
Rung 2 ($\mathcal{L}_2$)	ATE, LATE	IV/experimental demand responses
Rung 3 ($\mathcal{L}_3$)	MTE, unit counterfactual	Market-specific demand curve
$p \mapsto Q_p(u)$	Unit response function	Ceteris paribus demand $q_{\bar{x},\bar{\xi}}(p)$
Abduction	Recover ε from data	Berry inversion / recover latent demand index
Action	Set $P \leftarrow p'$	Set $P \leftarrow p'$
Prediction	Evaluate $D(\bar{x}, p', \bar{\xi})$	Evaluate $\sigma(\delta^*, p')$

The table maps SCM objects into the languages of structural econometrics and demand estimation. The key rows for this paper are Rung 2, Rung 3, and abduction.

make the abduction step single-valued; C3 is needed to recover ξ when characteristics change. The language is different; the objects are the same.

6.5. Pearl's linear example as the knife-edge

Pearl (2009) poses three queries on the supply-demand model. His running example uses linear demand with additive errors: $Q = aP + b\xi$. This model satisfies (AS): the derivative $\partial Q/\partial p = a$ does not depend on ξ. The Rung 2 slope already equals every market's slope, so the derivative gap vanishes. Pearl's own framework predicts this: the gap is zero precisely when (AS) holds, and linear demand with additive errors is the leading case of (AS). The gap that motivates the present paper does not appear in Pearl's running example because the running example satisfies exactly the knife-edge condition that eliminates it.

Table 1 summarizes the correspondence.

7. Quantitative Illustration

The characterization is exact; the gap is quantitative. The experimental approach identifies the average response to a price intervention across markets; the policy problem requires the response of this market at its realized demand conditions. How large is the difference in a workhorse demand model? We compare two markets that share the same prices but differ in unobserved quality ξ_1, first at fixed prices (isolating the derivative mismatch) and then in Bertrand equilibrium (where the mismatch translates into different predicted merger price effects).

Consider a market with $J = 2$ products and an outside good ($j = 0$). Let the random coefficient on price be $\alpha_i = \exp(\mu_\alpha + \tau_\alpha \nu_i)$ with $\nu_i \sim N(0,1)$, $\mu_\alpha = 0$, $\tau_\alpha = 0.5$, and let ε_{ij} be i.i.d. Type I extreme value. We compare two markets that differ only in ξ_1. Two regimes are distinguished below: fixed-price comparisons (holding $p = (2.0, 1.5)$ across markets) and Bertrand equilibrium outcomes (where prices are endogenous).

Market A. $\delta = (1.0,\ 0.5)$, $p = (2.0,\ 1.5)$. Product 1 has relatively low unobserved quality (ξ_1 is low) and commands a market share of $s_1^A \approx 0.21$. Product 2 has $s_2^A \approx 0.20$. The outside good share is 0.60.

Market B. $\delta = (2.0,\ 0.5)$, $p = (2.0,\ 1.5)$. Same prices, same product 2, but product 1 has higher unobserved quality ($\delta_1 = 2.0$ instead of 1.0). Product 1's share rises to $s_1^B \approx 0.39$; product 2's falls to $s_2^B \approx 0.14$.

The own-price derivatives and elasticities differ across markets despite identical prices:

	Market A ($\delta_1 = 1.0$)		Market B ($\delta_1 = 2.0$)	
	s_j	$\partial s_j/\partial p_j$	s_j	$\partial s_j/\partial p_j$
Product 1	0.21	−0.14	0.39	−0.20
Product 2	0.20	−0.14	0.14	−0.12
Outside	0.60	—	0.47	—

Product 1's own-price derivative is about 51% larger in magnitude in Market B than Market A at identical prices. Higher ξ_1 shifts product 1's market share, changing the composition of marginal consumers and the weighted average of α_i in the derivative—$\partial^2 \sigma_1 / \partial p_1 \, \partial \xi_1 \neq 0$, the separability failure of Theorem 1.

The fixed-price comparison isolates the derivative mismatch. What begins as a difference in price derivatives propagates through the Bertrand equilibrium and becomes a difference in predicted merger effects; Table 2 reports equilibrium outcomes. Consider Markets A and B, plus a "midpoint" market at $\delta_1 = 1.5$. Suppose products 1 and 2 merge. Standard Bertrand merger simulation (Nevo 2000) solves the merged firm's first-order conditions, which depend on diversion ratios and own-price derivatives—objects that vary with ξ.

TABLE 2. Merger price effects across markets that differ only in ξ_1.

	Market A ($\delta_1 = 1.0$)	Market B ($\delta_1 = 2.0$)	Pop. avg. ($\delta_1 = 1.5$)	Ratio B/A
Share, product 1 (s_1)	0.23	0.35	0.29	1.5
Own-price elasticity (ε_{11})	−1.20	−1.16	−1.18	1.0
Diversion ratio (2 → 1)	0.29	0.39	0.35	1.4
Merger Δp_2	+36%	+71%	+51%	2.0
Consumer surplus change	−21%	−21%	−21%	1.0

Notes. Random-coefficients logit with $J = 2$ and an outside good. $\alpha_i \sim \text{Lognormal}(0, 0.25)$, 100,000 simulation draws. Pre-merger: two single-product Bertrand firms. Post-merger: common ownership of both products. Markets differ only in the unobserved quality of product 1 (ξ_1, absorbed into δ_1); marginal costs are identical. Marginal costs calibrated so that the population-average market ($\delta_1 = 1.5$) has equilibrium prices $p = (2.0, 1.5)$.

The key driver is the diversion ratio from product 2 to product 1 (Farrell and Shapiro 2010). As Conlon and Mortimer (2021) emphasize, diversion is not a single number but depends on the counterfactual scenario, making it inherently a Rung 3 object. In Market B, product 1's high unobserved quality ($\delta_1 = 2.0$) attracts a larger customer base ($s_1 = 0.35$), so a price increase on product 2 diverts more consumers to product 1. The diversion ratio is 0.39, versus 0.29 in Market A. The merged firm internalizes this diversion, raising product 2's price by about 71% in Market B versus 36% in Market A—a factor-of-two difference driven entirely by unobserved demand conditions.

An analyst limited to experimental estimates knows $\mathbb{E}[\sigma(\delta, p)]$ and $\mathbb{E}[\partial\sigma/\partial p(\delta, p)]$ as functions of p, but not any market's δ^*. Plugging these population-average objects into the merged firm's first-order conditions yields a single prediction of 47% for every market—overstating the true effect in Market A by 11 percentage points and understating it in Market B by 24 percentage points. The error is not estimation noise but Jensen's inequality: σ is nonlinear in δ, so averaging over δ before evaluating the merger formula produces a prediction that corresponds to no actual market.

Online Appendix D reports sensitivity analysis varying δ_1 from 0.5 to 2.5; the structural merger prediction ranges from 24% to 95%, a factor of 3.9. The experimental prediction remains 47% throughout.

The fixed-price comparison is Theorem 1 made visible: two markets at the same price have different derivatives because ξ enters the share function nonlinearly, violating (AS). The Bertrand comparison is Corollary 1 made consequential: recovering δ^* by inversion pins down the derivative at the realized market. The merger formula is the same in both cases. What differs is which demand curve supplies its inputs. The risk is not the wrong formula, but the wrong demand curve—and the distance between them is a factor of two.

8. Discussion

Taken together, Theorems 1 and 2 answer the question posed in the introduction. When is the interventional question enough, and when must the counterfactual one be asked? Pearl's challenge and economists' objections turn out to be compatible. The third question is genuinely different, and answering it requires functional-form restrictions that graphs cannot encode (Imbens 2020, 2022; Heckman and Pinto 2024).

Market-specific structural estimation and valid instruments are complements, not substitutes. Instruments deliver the demand function; inversion delivers the market's position on it. Share inversion is an identification requirement, not a computational trick. The conditions under which it succeeds are exactly the conditions for market-specific counterfactuals to be point-identified. The relevant conditions (additive separability, invertibility, recoverability) are properties of the structural functions, not of the causal graph. The graph for the demand model is the same whether $\partial^2 D/\partial p\,\partial\xi = 0$ or not; the conditions that determine whether experiments suffice for counterfactuals live below the resolution of any directed acyclic graph.

The multi-product extension is in Appendix A. The separability condition becomes Jacobian invariance in (p, δ), and the invertibility condition is ensured by the Berry, Gandhi, and Haile (2013) connected-substitutes conditions, which require every product to connect to the outside good through a chain of strict substitution. The gap is not specific to demand.[14] In production function estimation, firm-specific productivity is the latent state and cross-firm regressions deliver

[14] Other parallels include dynamic discrete choice (Rust 1987), where the value function at a specific agent's unobserved state is Rung 3; the sufficient statistics approach (Chetty 2009), which succeeds when the relevant elasticities are Rung 2 objects; transportability (Bareinboim and Pearl 2014), where applying estimates to a new market requires abducting the target's latent state; and Wu and Wang (2026), who independently classify treatment effect estimands within the hierarchy.

the average input–output relationship (Rung 2), but the production function at a specific firm's productivity requires inverting the production function. This is the same abduction step. In macroeconomic policy, Caravello, McKay, and Wolf (2024) reach the same conclusion from the other direction. Identified monetary shocks pin down near-term counterfactuals, but structure is needed "for one sole purpose: to extrapolate" to persistent policy changes, the macro analogue of our abduction step. In treatment effects, the LATE (Imbens and Angrist 1994) is Rung 2 and the MTE is Rung 3; Mogstad and Torgovitsky (2018, 2024) show that extrapolation requires structural restrictions on how effects vary with the unobserved margin, the treatment-effects analogue of our separability conditions.

The practical upshot is a diagnostic. A researcher who inverts for δ^* is performing abduction whether or not she uses the term. If demand satisfies (AS), the inversion is unnecessary. If it does not, there is no shortcut. The conditions under which inversion succeeds are exactly the conditions for unit-level counterfactuals to be point-identified.

Structural econometrics has always been, at bottom, a defense of functional form as a tool for counterfactual reasoning.[15] This paper shows the defense is necessary. Without it, the demand curve is not point-identified from experimental information alone.

References

Allen, Roy and John Rehbeck. 2019. "Identification With Additively Separable Heterogeneity." *Econometrica* 87 (3): 1021–1054.

Andrews, Isaiah, Nano Barahona, Matthew Gentzkow, Ashesh Rambachan, and Jesse M. Shapiro. 2025. "Structural Estimation Under Misspecification: Theory and Implications for Practice." *Quarterly Journal of Economics* 140 (3): 1801–1855.

Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *Review of Economic Studies* 67 (3): 499–527.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics." *Journal of Economic Perspectives* 24 (2): 3–30.

[15] The exchange between Angrist and Pischke (2010) and Nevo and Whinston (2010) makes the point. Each merger's demand conditions differ enough that average causal estimates cannot stand in for the market-specific counterfactual.

Bareinboim, Elias, Juan D. Correa, Duligur Ibeling, and Thomas Icard. 2022. "On Pearl's Hierarchy and the Foundations of Causal Inference." In *Probabilistic and Causal Inference: The Works of Judea Pearl*, pp. 507–556. ACM Books.

Bareinboim, Elias and Judea Pearl. 2014. "External Validity: From Do-Calculus to Transportability Across Populations." *Statistical Science* 29 (4): 579–595.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25 (2): 242–262.

Berry, Steven T., Amit Gandhi, and Philip A. Haile. 2013. "Connected Substitutes and Invertibility of Demand." *Econometrica* 81 (5): 2111–2144.

Berry, Steven T. and Philip A. Haile. 2014. "Identification in Differentiated Products Markets Using Market Level Data." *Econometrica* 82 (5): 1749–1797.

Berry, Steven T. and Philip A. Haile. 2018. "Identification of Nonparametric Simultaneous Equations Models with a Residual Index Structure." *Econometrica* 86 (1): 289–315.

Berry, Steven T. and Philip A. Haile. 2021. "Foundations of Demand Estimation." *Handbook of Industrial Organization* 4 (1): 1–62.

Berry, Steven T. and Philip A. Haile. 2024. "Nonparametric Identification of Differentiated Products Demand Using Micro Data." *Econometrica* 92 (4): 1135–1162.

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

Borusyak, Kirill, Jiafeng Chen, Peter Hull, and Lint Lei. 2026. "Nonparametric Identification of Demand without Exogenous Product Characteristics." Working Paper 34842, National Bureau of Economic Research.

Caravello, Tomás E., Alisdair McKay, and Christian K. Wolf. 2024. "Evaluating Monetary Policy Counterfactuals: (When) Do We Need Structural Models?" Working Paper 32988, National Bureau of Economic Research.

Chen, Jiafeng. 2025. "Reinterpreting Demand Estimation." Working Paper, Stanford University.

Chernozhukov, Victor and Christian Hansen. 2005. "An IV Model of Quantile Treatment Effects." *Econometrica* 73 (1): 245–261.

Chesher, Andrew and Adam M. Rosen. 2017. "Generalized Instrumental Variable Models." *Econometrica* 85 (3): 959–989.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1: 451–488.

Christensen, Timothy and Benjamin Connault. 2023. "Counterfactual Sensitivity and Robustness." *Econometrica* 91 (1): 263–298.

Conlon, Christopher and Julie Holland Mortimer. 2021. "Empirical Properties of Diversion Ratios." *RAND Journal of Economics* 52 (4): 693–726.

Farrell, Joseph and Carl Shapiro. 2010. “Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition.” *The B.E. Journal of Theoretical Economics* 10 (1): 1–39.

Haavelmo, Trygve. 1943. “The Statistical Implications of a System of Simultaneous Equations.” *Econometrica* 11 (1): 1–12.

Heckman, James J. and Rodrigo Pinto. 2024. “Econometric Causality: The Central Role of Thought Experiments.” *Journal of Econometrics* 243 (1–2): 105719.

Heckman, James J. and Edward Vytlacil. 2005. “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.” *Econometrica* 73 (3): 669–738.

Hünermund, Paul and Elias Bareinboim. 2025. “Causal Inference and Data Fusion in Econometrics.” *The Econometrics Journal* 28 (1): 41–82.

Huntington-Klein, Nick. 2022. “Pearl before Economists: The Book of Why and Empirical Economics.” *Journal of Economic Methodology* 29 (4): 326–334.

Imbens, Guido W. 2020. “Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics.” *Journal of Economic Literature* 58 (4): 1129–1179.

Imbens, Guido W. 2022. “Causality in Econometrics: Choice vs. Chance.” *Econometrica* 90 (6): 2541–2566.

Imbens, Guido W. and Joshua D. Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62 (2): 467–475.

Imbens, Guido W. and Whitney K. Newey. 2009. “Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity.” *Econometrica* 77 (5): 1481–1512.

Kline, Patrick and Christopher R. Walters. 2019. “On Heckits, LATE, and Numerical Equivalence.” *Econometrica* 87 (2): 677–696.

Matzkin, Rosa L. 2007. “Nonparametric Identification.” In *Handbook of Econometrics*, vol. 6B, edited by James J. Heckman and Edward E. Leamer, chap. 73. Elsevier.

Matzkin, Rosa L. 2008. “Identification in Nonparametric Simultaneous Equations Models.” *Econometrica* 76 (5): 945–978.

Mogstad, Magne, Andres Santos, and Alexander Torgovitsky. 2018. “Using Instrumental Variables for Inference on Policy Relevant Treatment Parameters.” *Econometrica* 86 (5): 1589–1619.

Mogstad, Magne and Alexander Torgovitsky. 2018. “Identification and Extrapolation of Causal Effects with Instrumental Variables.” *Annual Review of Economics* 10: 577–613.

Mogstad, Magne and Alexander Torgovitsky. 2024. “Instrumental Variables with Unobserved Heterogeneity in Treatment Effects.” Working Paper 32927, National Bureau of Economic Research. Prepared for the Handbook of Labor Economics.

Molinari, Francesca. 2020. “Microeconometrics with Partial Identification.” *Handbook of Econometrics* 7A: 355–486.

Nevo, Aviv. 2000. “Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry.” *RAND Journal of Economics* 31 (3): 395–421.

Nevo, Aviv and Michael D. Whinston. 2010. “Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference.” *Journal of Economic Perspectives* 24 (2): 69–82.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii. 2015. “Moment Inequalities and Their Application.” *Econometrica* 83 (1): 315–334.

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2nd ed.

Pearl, Judea. 2015. “Trygve Haavelmo and the Emergence of Causal Calculus.” *Econometric Theory* 31 (1): 152–179.

Pearl, Judea. 2023. “Comments on Nick Huntington-Klein’s Review “Pearl before Economists”.” *Journal of Economic Methodology* 30 (1): 63–67.

Rubin, Donald B. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66 (5): 688–701.

Rust, John. 1987. “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher.” *Econometrica* 55 (5): 999–1033.

Shpitser, Ilya and Judea Pearl. 2006. “Identification of Joint Interventional Distributions in Recursive Semi-Markovian Causal Models.” In *Proceedings of the 21st National Conference on Artificial Intelligence*, pp. 1219–1226.

Tebaldi, Pietro, Alexander Torgovitsky, and Hanbin Yang. 2023. “Nonparametric Estimates of Demand in the California Health Insurance Exchange.” *Econometrica* 91 (1): 107–146.

Torgovitsky, Alexander. 2015. “Identification of Nonseparable Models Using Instruments with Small Support.” *Econometrica* 83 (3): 1185–1197.

Vytlacil, Edward. 2002. “Independence, Monotonicity, and Latent Index Models: An Equivalence Result.” *Econometrica* 70 (1): 331–341.

Wu, Peng and Linbo Wang. 2026. “Position: A Potential Outcomes Perspective on Pearl’s Causal Hierarchy.” ArXiv:2601.20405.

Appendix A. Multi-Product Formalization

The main text develops all results for a scalar demand function $Q = D(X, P, \xi)$ with a single price P and a single unobservable ξ. The workhorse demand models in industrial organization involve J differentiated products, vector-valued shares, vector-valued prices, and a vector of unobserved demand shocks. This appendix extends each main result to the multi-product setting. The scalar characterization (Theorem 1(b)) generalizes to require invariance of the entire $J \times J$ price Jacobian, a joint restriction on J^2 derivatives rather than a collection of separate scalar conditions.

Setup

J differentiated products compete for a population of consumers. Market shares are $s = (s_1, \ldots, s_J) \in \Delta^J$, the interior of the J-simplex (with an outside good share $s_0 = 1 - \sum_j s_j > 0$). Prices are $p = (p_1, \ldots, p_J) \in \mathbb{R}_+^J$, observed characteristics are $x = (x_1, \ldots, x_J) \in \mathbb{R}^{J \times K}$, and unobserved demand shocks are $\xi = (\xi_1, \ldots, \xi_J) \in \mathbb{R}^J$. The demand function is the vector-valued share function

$$s = \sigma(\delta, p, x; \theta), \qquad \sigma : \mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^{J \times K} \to \Delta^J,$$

where $\delta = (\delta_1, \ldots, \delta_J) \in \mathbb{R}^J$ is the vector of mean utilities, with $\delta_j = \delta_j(x_j, \xi_j; \theta)$. In the BLP framework, $\delta_j = x_j'\beta + \xi_j$ (the additive index), but the results below require only injectivity of $\xi \mapsto \delta(x, \xi; \theta)$, not the specific functional form.

Proposition 4 (vector version)

For each product j, the unit-level counterfactual share at price vector p' is

$$s_j(p'; u) = \sigma_j(\delta^*, p', x; \theta), \qquad \delta^* = \delta(x^*, \bar{\xi}; \theta),$$

which holds (x, ξ) fixed at the unit's realized values and varies only the price vector. This is a market-specific counterfactual, a deterministic function of the unit u (through δ^*) and the counterfactual price p'. The experimental average

$$m_j(p', x) = \int \sigma_j(\delta, p', x; \theta)\, dF_{\delta|X=x}(\delta)$$

averages over the distribution of δ (equivalently, of ξ) in markets with characteristics x. The proof is identical to the scalar case, applied coordinate-by-coordinate.

Theorem 1 (vector version)

The scalar characterization extends as follows. Define the $J \times J$ Jacobian matrix of shares with respect to prices:

$$J_p(\delta, p, x; \theta) \equiv \frac{\partial \sigma(\delta, p, x; \theta)}{\partial p'} \in \mathbb{R}^{J \times J}, \qquad [J_p]_{jk} = \frac{\partial \sigma_j}{\partial p_k}.$$

The population-average Jacobian is $\bar{J}_p(p, x; \theta) \equiv \mathbb{E}[J_p(\delta, p, x; \theta) \mid X = x]$, integrating over δ.

PROPOSITION A1 (Vector analogue of Theorem 1(b)). *Fix x and a connected component $\mathcal{P}_0$ of the open price support. Assume each share component $\sigma_j(\delta, p, x; \theta)$ is continuously differentiable in each price component, and for each (j, k) the derivative $\partial \sigma_j / \partial p_k$ is dominated on compact price sets by an integrable envelope (entrywise analogue of Assumptions A1–A2). Then the identification gap vanishes in all price derivatives on $\mathcal{P}_0$—that is,*

$$\bar{J}_p(p, x; \theta) = J_p(\delta, p, x; \theta) \quad \text{for all } p \in \mathcal{P}_0 \text{ and all } \delta \text{ in the conditional support}$$

—if and only if the share system is additively separable in (p, δ) on $\mathcal{P}_0$: there exist vector-valued functions f and g such that

$$\text{(VAS)} \qquad \sigma(\delta, p, x; \theta) = f(p, x; \theta) + g(\delta, x; \theta)$$

for all δ in the conditional support of $\delta \mid X = x$, simultaneously for all $p \in \mathcal{P}_0$. Equivalently, this occurs if and only if there exists a matrix-valued function $\Phi(p, x; \theta)$ such that

$$\text{(JI)} \quad J_p(\delta, p, x; \theta) = \Phi(p, x; \theta) \quad \text{for all } \delta \text{ in the conditional support and all } p \in \mathcal{P}_0.$$

PROOF. ($\Rightarrow$): Fix (j, k, p, x) and let $h(\delta) \equiv \partial \sigma_j(\delta, p, x; \theta) / \partial p_k$. The hypothesis and the argument of Theorem 1(a) imply that h is constant on the conditional support of δ; denote the matrix of these constants by $\Phi(p, x; \theta)$. Since this holds for each (j, k, p), we have $J_p(\delta, p, x; \theta) = \Phi(p, x; \theta)$ for all δ in the conditional support and all $p \in \mathcal{P}_0$.

Fix a reference price $p_0 \in \mathcal{P}_0$ and any δ_0 in the conditional support. Define

$$f(p,x;\theta) \equiv \sigma(\delta_0,p,x;\theta) - \sigma(\delta_0,p_0,x;\theta), \qquad g(\delta,x;\theta) \equiv \sigma(\delta,p_0,x;\theta).$$

For any δ in the conditional support, let

$$H_\delta(p) \equiv \sigma(\delta,p,x;\theta) - f(p,x;\theta) - g(\delta,x;\theta).$$

Then $\partial H_\delta(p)/\partial p' = 0$ on $\mathcal{P}_0$ because $\sigma(\delta,p,x;\theta)$ and $f(p,x;\theta)$ have the same Jacobian $\Phi(p,x;\theta)$, while $g(\delta,x;\theta)$ is independent of p. Also $H_\delta(p_0) = 0$. Since $\mathcal{P}_0$ is connected, H_δ is constant on $\mathcal{P}_0$, hence identically zero. This yields (VAS).

($\Leftarrow$): If $\sigma(\delta,p,x;\theta) = f(p,x;\theta) + g(\delta,x;\theta)$, then $J_p(\delta,p,x;\theta) = \partial f(p,x;\theta)/\partial p'$ does not depend on δ, so $\bar{J}_p = J_p$. □

Condition (VAS) is the multi-product analogue of (AS): once the entire Jacobian with respect to prices is δ-invariant, the share system can differ across markets only through a price-independent vector shift. Equivalently, condition (JI) says that the full matrix of own- and cross-price derivatives is invariant to δ. This is far more restrictive than the scalar condition, because it requires invariance of J^2 derivatives simultaneously. In the random-coefficients logit, the own-price derivative is

$$\frac{\partial\sigma_j}{\partial p_j} = -\int \alpha_i\,\pi_{ij}(\delta,p)(1-\pi_{ij}(\delta,p))\,dG(\alpha_i),$$

and the cross-price derivative is

$$\frac{\partial\sigma_j}{\partial p_k} = \int \alpha_i\,\pi_{ij}(\delta,p)\,\pi_{ik}(\delta,p)\,dG(\alpha_i), \qquad j \neq k.$$

Both depend on δ through the individual choice probabilities π_{ij}. Changing ξ_j for a single product changes the entire column and row of J_p associated with product j. Jacobian invariance fails whenever G is non-degenerate; it also fails for the homogeneous logit, the nested logit, and CES.

Separability failure in scalar models

The δ-dependence of the price derivative is visible in every standard special case. In the *homogeneous logit* ($\alpha_i = \bar{\alpha}$ for all i), the own-price derivative and cross-price derivative are

$$\frac{\partial \sigma_j}{\partial p_j} = -\bar{\alpha}\, \sigma_j(1-\sigma_j), \qquad \frac{\partial \sigma_j}{\partial p_k} = \bar{\alpha}\, \sigma_j \sigma_k, \quad j \neq k,$$

both of which depend on shares (σ_j, σ_k) and hence on δ. Even without taste heterogeneity, the nonlinearity of the logit choice probability creates the interaction between δ and the derivative that Theorem 1(b) characterizes.

In the *CES model*, with $\sigma_j \propto (\delta_j/p_j)^\eta$, the own-price elasticity is $-\eta(1-\sigma_j)$, which depends on σ_j and hence on δ.

In the *nested logit*, products are partitioned into nests $g = 1, \ldots, G$, with within-nest correlation parameter $\lambda_g \in (0, 1]$. Shares take the form

$$\sigma_j = \frac{\exp(\tilde{\delta}_j/\lambda_g)}{D_g^{1-\lambda_g} \cdot (1 + \sum_{g'} D_{g'}^{1-\lambda_{g'}})}, \qquad D_g \equiv \sum_{k \in g} \exp(\tilde{\delta}_k/\lambda_g),$$

where $\tilde{\delta}_j = \delta_j - \alpha p_j$. The own-price derivative is

$$\frac{\partial \sigma_j}{\partial p_j} = -\frac{\alpha}{\lambda_g}\, \sigma_j \Big(1 - \lambda_g \sigma_{j|g} - (1-\lambda_g)\sigma_j\Big),$$

where $\sigma_{j|g} = \exp(\tilde{\delta}_j/\lambda_g)/D_g$ is the within-nest share. This derivative depends on both $\sigma_{j|g}$ and σ_j, which shift with δ. Even at fixed prices, two markets with different ξ have different within-nest compositions and hence different own-price sensitivities. The cross-price derivative similarly depends on whether products j and k are in the same nest, creating a δ-dependent substitution matrix.

For any structural function $Y = g(X, P, \xi)$, condition (AS) requires $\partial^2 g/\partial p\, \partial \xi = 0$. In every model above, the demand function is nonlinear in ξ (through δ) in a way that interacts with p, so the cross-partial is generically nonzero.

Theorem 2 (vector version)

Recovering the market-specific δ^* requires inverting the share system. In the multi-product case, this inversion solves a system of J nonlinear equations, mapping observed share vectors to mean-utility vectors. The three conditions generalize to:

(C1′) **Index structure.** $s = \sigma(\delta(x, \xi; \theta), p, x; \theta)$ with $\delta : \mathbb{R}^{J \times K} \times \mathbb{R}^J \to \mathbb{R}^J$.

(C2′) **Global invertibility.** For each (p, x), the map $\delta \mapsto \sigma(\delta, p, x; \theta)$ is a bijection from $\mathbb{R}^J$ to the range of σ.

(C3′) **Recoverability.** $\xi \mapsto \delta(x, \xi; \theta)$ is injective for each x.

In the scalar case ($J = 1$), invertibility of $\sigma(\cdot, p, x; \theta)$ in δ follows from monotonicity whenever $\partial\sigma/\partial\delta > 0$. In the multi-product case, local invertibility (nonsingularity of the Jacobian $\partial\sigma/\partial\delta'$) does not imply global invertibility: a locally invertible map from $\mathbb{R}^J$ to $\mathbb{R}^J$ need not be globally injective when $J > 1$.

Berry, Gandhi, and Haile (2013) establish global invertibility under "connected substitutes," which requires two conditions: (i) the Jacobian $\partial\sigma/\partial\delta'$ satisfies $\partial\sigma_j/\partial\delta_j > 0$ (own effect) and $\partial\sigma_j/\partial\delta_k \le 0$ for $j \ne k$ (substitutes), and (ii) the substitution pattern is connected in the sense that for any partition of products into two groups, at least one product in each group is a substitute for some product in the other group.[16] Under connected substitutes, the map $\delta \mapsto \sigma(\delta, p, x; \theta)$ is globally invertible for each (p, x), so that abduction $\delta^* = \sigma^{-1}(s^*, p^*, x^*; \theta)$ has a unique solution.

The proof is identical to Theorem 2. Observed shares satisfy $s^* = \sigma(\delta^*, p^*, x^*; \theta)$ with $\delta^* = \delta(x^*, \bar{\xi}; \theta)$ by C1′. By C2′, $\delta^* = \sigma^{-1}(s^*, p^*, x^*; \theta)$. By C3′, $\bar{\xi} = r(\delta^*, x^*; \theta)$. Under the intervention $P \leftarrow p'$, the structural share system remains $s = \sigma(\delta, p, x; \theta)$. For a price-only counterfactual, x^* and $\bar{\xi}$ (hence δ^*) are unchanged, so $s(p'; u) = \sigma(\delta^*, p', x^*; \theta)$. For a counterfactual that also changes characteristics to x', the recovered $\bar{\xi}$ is used to form $\delta' = \delta(x', \bar{\xi}; \theta)$, and the counterfactual shares are $\sigma(\delta', p', x'; \theta)$. The inversion σ^{-1} is the abduction step, and the conditions for it to be well-defined are the Berry–Haile identification conditions.

Appendix B. Regularity Conditions

We collect the regularity conditions referenced in the main text.

ASSUMPTION A1 (Differentiability). *The demand function $D(x, p, \xi)$ is continuously differentiable in p for each (x, ξ).*

ASSUMPTION A2 (Dominated convergence). *For each $x \in \mathcal{X}$ and each compact $K \subset \mathcal{P}$, there exists an integrable function $M(\xi)$ with $\mathbb{E}[M(\xi) \mid X = x] < \infty$ such that $|\partial D(x, p, \xi)/\partial p| \le M(\xi)$ for all $p \in K$ and ξ in the conditional support of $\xi \mid X = x$. This ensures differentiation under the expectation is justified.*

[16] Formally, the matrix $-(\partial\sigma/\partial\delta')^{-1}$ has all non-negative entries (an M-matrix property). See Berry, Gandhi, and Haile (2013), Theorem 1 and Corollary 1, for the precise statement and proof.

ASSUMPTION A3 (Non-degenerate heterogeneity). $\mathrm{Var}(\xi \mid X = x) > 0$ *for all* $x \in \mathcal{X}$.[17]

ASSUMPTION A4 (Support richness). *The price support* $\mathcal{P}$ *is an interval in* $\mathbb{R}$. *The support of* ξ *conditional on* $X = x$ *is connected and contains at least two distinct points. In the multi-product setting, the price support is an open connected subset of* $\mathbb{R}^J$.

ASSUMPTION A5 (Invertibility regularity). *In the multi-product setting,* $\sigma(\cdot, p, x; \theta)$ *is continuously differentiable in* δ*, and the Jacobian* $\partial\sigma/\partial\delta$ *has full rank on the relevant domain.*

Proof of Proposition 2

(a) Define $\Delta(p) \equiv D(p, \xi_a) - D(p, \xi_b)$. By (i), $\Delta(P^*) = 0$ and $\Delta'(P^*) \neq 0$. By continuity of Δ', Δ is strictly monotone on a neighborhood of P^*, so $\Delta(p') \neq 0$ for p' in a punctured neighborhood of P^*.

The average $m(P^*, x) = \int D(P^*, \xi)\, dF_\xi(\xi)$ is consistent with any mixing distribution over the support that generates the observed average. Both ξ_a and ξ_b receive positive weight by (ii) and both produce Q^* at P^* by (i), so complete knowledge of $m(p, x)$ for all p cannot distinguish them at the observed evidence. The identified set $\mathcal{I}(p'; e)$ contains $\{D(p', \xi_a), D(p', \xi_b)\}$, which are distinct.

(b) Taylor-expand $D(p', \xi_k)$ around P^* for $k \in \{a, b\}$ and subtract. Using $D(P^*, \xi_a) = D(P^*, \xi_b)$ from (i), the zeroth-order terms cancel and the leading term of $\Delta(p')$ is $[D'_p(P^*, \xi_a) - D'_p(P^*, \xi_b)](p' - P^*)$. Condition (ii) is used only in part (a). □

Appendix C. Identification Content of Each Condition

Conditions C1–C3 correspond to the identification conditions of Berry and Haile (2014). Each serves a specific role in the Rung 2–3 transition. C1 reduces the dimensionality of abduction. C2 ensures the solution is unique. C3 decomposes the recovered index when the counterfactual changes characteristics. Without any one, abduction fails or delivers an ambiguous answer.

C1 (Index structure). All demand-relevant unobservables enter through a finite-dimensional index δ. Without C1, ξ could affect demand through an arbitrary,

[17] This rules out the degenerate case $\xi = \bar{\xi}$ a.s., under which the abduction step is trivial (δ^* uniquely determines $\bar{\xi}$) and Proposition 3(b) has no bite. Non-degeneracy ensures C3 is a genuinely binding restriction.

high-dimensional pathway, making recovery of $\bar{\xi}$ ill-posed.[18] The index restriction says that there is a finite-dimensional sufficient statistic δ for the unobservables. In BLP, $\delta_j = x_j'\beta + \xi_j$ (mean utility net of price); the vector of unobserved product qualities is summarized in δ. Note that σ may depend on (p, x) separately from δ; the index restriction is about how ξ enters, not about whether σ has other arguments.

C2 (Invertibility) ensures abduction has a unique solution. Given observed shares s^*, prices p^*, and characteristics x^*, the equation $s^* = \sigma(\delta, p^*, x^*; \theta)$ pins down δ uniquely. Without invertibility, multiple latent states are consistent with the data, and the unit-level counterfactual is not point-identified. Berry (1994) established invertibility for the multinomial logit; Berry, Gandhi, and Haile (2013) provided general conditions ("connected substitutes") under which the share function is globally invertible.

The crossing-curves example of Proposition 1 illustrates: two demand functions pass through $(Q^* = 3, P^* = 3)$, and without C2 the identified set for counterfactual demand at $p' = 4$ is $\{0, 2\}$—an interval of width 2.

In the simple logit with homogeneous price coefficient α, the *inclusive* mean utilities $\tilde{\delta}_{jt} \equiv \delta_{jt} - \alpha p_{jt}$ satisfy $\tilde{\delta}_{jt} = \log s_{jt} - \log s_{0t}$. Under our convention $\delta_{jt} = x_{jt}'\beta + \xi_{jt}$ (net of price), abduction recovers $\delta_{jt} = \log s_{jt} - \log s_{0t} + \alpha p_{jt}$ given α. The logit share function is strictly monotone in each δ_j (holding other δ_k fixed), so the inversion is globally unique—C2 holds by construction.[19] In random-coefficients models, the inversion remains well-defined but is not closed-form.

C3 (Recoverability) separates the recovered index δ into its observed and unobserved components. Berry and Haile's (2014) baseline uses the additive ("linear index") form $\delta(x, \xi; \theta) = x'\beta + \xi$, but they emphasize that the key requirement is injectivity of $\xi \mapsto \delta(x, \xi; \theta)$, which allows nonseparable indices provided monotonicity in ξ holds. Once δ^* is obtained by inversion (C2), recoverability yields $\bar{\xi} = r(\delta^*, x^*; \theta)$. Without C3, we recover the composite δ but cannot isolate ξ, which

[18] In statistical language, C1 says that δ is a sufficient statistic for ξ in the demand function, so $Q = \sigma(\delta, p, x; \theta)$ does not depend on ξ except through δ. This is the "index sufficiency" condition of Berry and Haile (2014), restated. Without it, the mapping from observables to latent states has infinite-dimensional fibers.

[19] For the logit, $\partial\sigma_j/\partial\delta_j = \sigma_j(1 - \sigma_j) > 0$ on the interior of the simplex. This ensures local invertibility. The closed-form inversion $\tilde{\delta}_j = \log s_j - \log s_0$ establishes global invertibility directly. In the BLP random-coefficients model, global invertibility is guaranteed by the connected-substitutes conditions of Berry, Gandhi, and Haile (2013); the inversion is numerical (typically via the contraction mapping of Berry, Levinsohn, and Pakes 1995).

is needed to evaluate counterfactuals that change x while holding ξ fixed. For price-only counterfactuals (holding x fixed), C3 is not needed: δ^* itself suffices, since $D(\bar{x}, p', \bar{\xi}) = \sigma(\delta^*, p', x^*; \theta)$. C3 becomes essential when counterfactuals also change product characteristics.

REMARK A1 (Estimation uncertainty). *If $\hat{\theta}_T \to \theta_0$ at rate $\sqrt{T}$ (standard GMM asymptotics), then by continuity of σ^{-1} the recovered $\hat{\delta}_t^* = \sigma^{-1}(s_t^*, P_t^*, X_t^*; \hat{\theta}_T)$ is consistent for δ_t^*. If, in addition, σ is continuously differentiable in (δ, θ) and the Jacobian $\partial\sigma/\partial\delta$ is nonsingular on the relevant domain (Assumption A5 in the multi-product case, with the scalar analogue when $J = 1$), then σ^{-1} is locally differentiable by the inverse function theorem, and the counterfactual $\hat{Q}_{p'}(u) = \sigma(\hat{\delta}_t^*, p', X_t^*; \hat{\theta}_T)$ inherits the $\sqrt{T}$ rate by the delta method. Estimation uncertainty in $\hat{\theta}$ propagates through the inversion, but it does not change the identification analysis: the hierarchy classification depends on the population objects (σ, θ_0), not on their estimators.*

REMARK A2 (Partial identification without C1–C3). *When C1–C3 are relaxed, unit-level counterfactuals are generally not point-identified but can be bounded. Without C2 (invertibility), multiple δ values rationalize observed shares, generating an identified set for counterfactual shares whose width grows linearly with the distance from the observed price at a rate proportional to slope heterogeneity (Proposition 2(b); Online Appendix C). Without C1 (index structure), the dimensionality of ξ exceeds what the data can pin down. These bounds connect to the partial identification literature (Molinari 2020; Pakes et al. 2015; Christensen and Connault 2023) and to Chen's (2025) observation that relaxing counterfactual homogeneity forecloses point-identification.*

Appendix D. Elasticity Gap and Reduced-Form Estimates

In discrete-choice models, the relevant objects are own-price elasticities:

COROLLARY A1 (Elasticity version of the gap). *Suppose $D(x, p, \xi) > 0$ on the relevant domain and define the own-price elasticity $\eta(x, p, \xi) \equiv \frac{\partial D}{\partial p}(x, p, \xi) \cdot \frac{p}{D(x,p,\xi)}$. Then the population-average elasticity equals every market's unit-level elasticity,*

$$\mathbb{E}[\eta(x, p, \xi) \mid X = x] = \eta(x, p, \bar{\xi}) \quad \textit{for all } \bar{\xi}, p,$$

if and only if $\eta(x, p, \xi)$ does not depend on ξ: there exists $\tilde{\eta}(x, p)$ such that, for each (x, p), $\eta(x, p, \xi) = \tilde{\eta}(x, p)$ for all ξ in the support of $\xi \mid X = x$.

PROOF. ($\Leftarrow$) If $\eta(x, p, \xi) = \tilde{\eta}(x, p)$ for all ξ in the support, then $\mathbb{E}[\eta \mid X = x] = \tilde{\eta}(x, p) = \eta(x, p, \bar{\xi})$ for all $\bar{\xi}$. ($\Rightarrow$) The argument is identical to Theorem 1(a) with η replacing D: if $\mathbb{E}[\eta \mid X = x] = \eta(\bar{\xi})$ for all $\bar{\xi}$, then η is constant in ξ on the support. $\square$

The corollary reveals a subtlety: even when the derivative gap closes, the elasticity gap remains. Under (AS), $\partial D/\partial p$ is ξ-free, but $\eta = (\partial D/\partial p) \cdot p/D$ depends on ξ through the denominator D. Two markets with the same slope but different demand levels have different elasticities. Population-average elasticities therefore generically differ from unit-level elasticities even under (AS).

COROLLARY A2 (When reduced-form estimates are informative). *The average derivative* $\frac{\partial m}{\partial p}(p, x)$*—identifiable from experiments or valid instruments—equals every market's demand slope if and only if* (AS) *holds.*

When (AS) fails, a reduced-form researcher who estimates the average price effect via an experiment or instrument is estimating $\partial m/\partial p$, which does not coincide with the slope any specific market faces. If the same market is observed at multiple exogenously assigned prices while ξ remains fixed, the demand curve is directly estimable and the recovery problem is moot. In the typical BLP setting, where each market-product cell is observed once, no such panel shortcut applies.