Online Appendix: Abduction and the Demand Curve

Brian C. Albrecht James Traina

April 2026

This online appendix collects applications, worked examples, proofs, and connections to recent literature that supplement the results in the main text. Cross-references to main-text theorems, propositions, and equations use the numbering from the main text.

OA.A. Policy Applications as Rung 3 Problems

Pass-through, merger simulation, and consumer surplus all require the demand curve at a specific market's latent state, not an average across markets. These are equilibrium counterfactuals (Remark 1) that use the demand counterfactual as an input.

Pass-through. Consider an excise tax τ imposed on a single-good market with demand $Q = D(p, \xi)$ and inverse supply $P = S(Q) + \tau$. The pre-tax equilibrium satisfies $Q^* = D(P^*, \bar{\xi})$ and $P^* = S(Q^*)$. The post-tax equilibrium solves $Q = D(P, \bar{\xi})$ and $P = S(Q) + \tau$. By the implicit function theorem, equilibrium pass-through is

$$\rho(\bar{\xi}) \equiv \frac{dP^*}{d\tau} = \frac{1}{1 - S'(Q^*) \cdot \partial D(P^*, \bar{\xi})/\partial P},$$

evaluated at the realized latent state $\bar{\xi}$. The derivative $\partial D/\partial P$ is market-specific whenever demand is not additively separable. Substituting $\mathbb{E}[\partial D/\partial P \mid X]$ into the pass-through formula yields $\rho(\mathbb{E}[D'])$, which generally differs from both the market-specific pass-through $\rho(D'(\bar{\xi}))$ and the true population-average pass-through $\mathbb{E}[\rho(D'(\xi)) \mid X]$.[1]

Return to the crossing-curves example. Suppose inverse supply is $P = S(Q) = Q$ (slope $S' = 1$). The two markets share the same pre-tax equilibrium ($P^* = 3, Q^* = 3$) but have different demand slopes:

The Rung 2 policymaker, who knows $\mathbb{E}[\partial D/\partial P] = -2$ but not the realized type, computes pass-through of 1/3. If the market is actually type ξ_1, the true pass-

[1] As Berry and Haile (2021, p. 11) note, "This ratio is not a LATE. By definition, a LATE averages over the latent variables; this is not the same thing as holding them fixed."

TABLE OA.1. Pass-through and welfare in the crossing-curves example ($\tau = 1$)

	$\partial D/\partial P$	$\rho(\bar{\xi})$	Post-tax P	Post-tax Q	ΔCS
Type ξ_1 (steep)	−3	1/4 = 0.25	3.25	2.25	−0.66
Type ξ_2 (flat)	−1	1/2 = 0.50	3.50	2.50	−1.38
Rung 2 plug-in ($\pi = 0.5$)	−2	1/3 ≈ 0.33	3.33	2.33	—

Notes. The Rung 2 post-tax quantity (2.33) is the population average $\mathbb{E}[Q \mid \mathrm{do}(P = 3.33)] = 9 - 2(3.33) \approx 2.33$, which lies between the two market-specific values but equals neither. The ΔCS entry is blank because welfare requires integrating the demand level at a specific $\bar{\xi}$, not an average slope.

through is 1/4—the Rung 2 estimate overstates the price increase by 33%. If the market is type ξ_2, the true pass-through is 1/2—the Rung 2 estimate understates it by 33%.[2] The consumer surplus losses are −0.66 for type ξ_1 and −1.38 for type ξ_2—a ratio exceeding 2.[3] The error comes from averaging over the latent state before evaluating a nonlinear function of the demand derivative, not from sampling noise or estimation imprecision.

Merger simulation. Post-merger price predictions require the demand function at the merging firms' pre-merger demand conditions. The merged entity's first-order conditions depend on shares, own-price derivatives, and cross-price derivatives—all functions of δ^* and hence of ξ.[4] Different markets face different substitution patterns and hence different post-merger price increases, even at identical pre-merger prices. A merger analysis that uses the population-average Jacobian rather than the market-specific Jacobian will mispredict the post-merger equilibrium.

Consumer surplus. Welfare analysis requires integrating the demand curve over a price range at the market's realized $\bar{\xi}$. In single-good settings, consumer surplus is $\mathrm{CS} = \int_{P^*}^{\bar{P}} q_{\bar{x},\bar{\xi}}(p)\,dp$, which depends on the entire demand curve, not just its

[2]The post-tax equilibrium for type ξ_k solves $Q = D(P, \xi_k)$ and $P = Q + \tau$ simultaneously. For type ξ_1: $Q = 12 - 3P$ and $P = Q + 1$ yield $P = 3.25$, $Q = 2.25$. For type ξ_2: $Q = 6 - P$ and $P = Q + 1$ yield $P = 3.5$, $Q = 2.5$.

[3]For linear demand $Q = a - bP$, the change in consumer surplus from a tax-induced price increase is $\Delta\mathrm{CS} = \mathrm{CS}(P + \rho\tau) - \mathrm{CS}(P^*)$ where $\mathrm{CS}(P) = (a - bP)^2/(2b)$. For type ξ_1: $\mathrm{CS}(3) = 1.5$, $\mathrm{CS}(3.25) = 0.84$, $\Delta = -0.66$. For type ξ_2: $\mathrm{CS}(3) = 4.5$, $\mathrm{CS}(3.5) = 3.13$, $\Delta = -1.38$. Pre-tax surplus levels already differ by a factor of 3 (1.5 vs. 4.5); the tax-induced losses differ by a factor of 2.1.

[4]This is the standard BLP-style merger simulation exercise; see Nevo (2000). The dependence of the predicted price increase on market-specific demand conditions is well understood by practitioners but is not usually framed in terms of the causal hierarchy.

slope at P^*. The pass-through table above quantifies the welfare consequences: the pre-tax surplus levels are 1.5 and 4.5 for types ξ_1 and ξ_2 respectively, differing by a factor of 3 despite identical observed outcomes.[5] In differentiated-products settings, the relevant object is the inclusive-value (log-sum) expression evaluated at the market's demand index.[6] The same logic applies: the latent state must be held fixed.

All three applications share a common structure: a policy-relevant quantity depends nonlinearly on the demand function evaluated at a specific market's $\bar{\xi}$, and the population average of that nonlinear function does not equal the function evaluated at the average $\bar{\xi}$. The gap is Jensen's inequality applied to the policy formula.

OA.B. Worked Example: Logit Market

To make the three-step procedure concrete, we walk through a simple logit example with $J = 3$ products, a homogeneous price coefficient $\alpha = 1$, and known parameters $\theta = (\alpha, \beta)$ with $\beta = 2$.

Setup. Market t has observed data: shares $s = (0.30, 0.20, 0.10)$, outside-good share $s_0 = 0.40$, prices $p = (3, 2, 4)$, and a single observed characteristic $x = (1, 0.5, 0.8)$. The structural model is $u_{ij} = \delta_j - \alpha p_j + \varepsilon_{ij}$ with ε_{ij} i.i.d. Type I extreme value, and $\delta_j = x_j\beta + \xi_j$. The share function is $\sigma_j(\tilde{\delta}) = \exp(\tilde{\delta}_j)/(1 + \sum_k \exp(\tilde{\delta}_k))$ where $\tilde{\delta}_j \equiv \delta_j - \alpha p_j$ is the "inclusive" mean utility.[7]

Step 1 (Abduction). The Berry inversion recovers inclusive mean utilities from observed shares:

$$\tilde{\delta}_j = \log s_j - \log s_0.$$

This is the closed-form expression for the logit; in random-coefficients models, the inversion is numerical but equally well-defined given C2.[8] Substituting the

[5]For linear demand $Q = a - bP$ with choke price a/b, consumer surplus at price P is $(a-bP)^2/(2b)$.

[6]In the logit, consumer surplus per capita is $\text{CS} = (1/\alpha)\log(1 + \sum_j \exp(\tilde{\delta}_j))$ where $\tilde{\delta}_j = \delta_j - \alpha p_j$. This depends on δ and hence on ξ. See Berry, Levinsohn, and Pakes (1995), Section 5.

[7]In the logit, σ depends on δ and p only through $\tilde{\delta} = \delta - \alpha p$. This is a property of the logit functional form, not a general feature of discrete choice models.

[8]Berry (1994) proved existence and uniqueness of the logit inversion. Berry, Gandhi, and Haile (2013) extended the result to a broad class of models satisfying "connected substitutes."

observed shares:

$$\tilde{\delta}_1 = \log(0.30) - \log(0.40) \approx -0.29,$$
$$\tilde{\delta}_2 = \log(0.20) - \log(0.40) \approx -0.69,$$
$$\tilde{\delta}_3 = \log(0.10) - \log(0.40) \approx -1.39.$$

Net-of-price mean utilities: $\delta_j = \tilde{\delta}_j + \alpha p_j$, giving $\delta \approx (2.71, 1.31, 2.61)$. These are the demand indices that the market's observed shares and prices imply, given the model. The unobserved quality of each product is $\xi_j = \delta_j - x_j\beta$, yielding $\xi \approx (0.71, 0.31, 1.01)$. This completes abduction: from observed (s^*, p^*, x^*) and the model (α, β), we have recovered the latent state (δ^*, ξ^*) specific to market t.

Step 2 (Action). Suppose we want to predict what happens if product 1 raises its price from $p_1 = 3$ to $p_1' = 4$, holding all other prices and characteristics fixed. The intervention replaces the pricing equation for product 1; the demand equation (4) and the recovered δ^* remain unchanged. Note that δ does not change—it is the mean utility *net of price*, so it depends on (x, ξ) but not on p. What changes is the inclusive utility: $\tilde{\delta}_1' = \delta_1 - \alpha p_1' = 2.71 - 4 = -1.29$.

Step 3 (Prediction). Counterfactual inclusive utilities:

$$\tilde{\delta}' = (-1.29,\ -0.69,\ -1.39).$$

Counterfactual shares:

$$s_j' = \frac{\exp(\tilde{\delta}_j')}{1 + \sum_k \exp(\tilde{\delta}_k')}.$$

Computing: $\sum_k \exp(\tilde{\delta}_k') = e^{-1.29} + e^{-0.69} + e^{-1.39} \approx 0.275 + 0.502 + 0.249 = 1.026$, so the denominator is 2.026. The counterfactual shares are approximately:

$$s' \approx (0.14,\ 0.25,\ 0.12), \quad s_0' \approx 0.49.$$

Product 1's share falls from 0.30 to 0.14—a large response. Product 2 gains (from 0.20 to 0.25) as consumers substitute. The outside good also gains.

The counterfactual is market-specific. This computation used market t's recovered $\delta^* = (2.71, 1.31, 2.61)$. A different market t' with different ξ would have different δ^*

and hence different counterfactual shares, even at the same prices. To illustrate, suppose market t' has lower unobserved quality for product 1: $\xi_1' = -0.29$ instead of 0.71, giving $\delta_1' = 1.71$, while products 2 and 3 are unchanged. Then $\tilde{\delta}_1' = 1.71 - 4 = -2.29$ under the same counterfactual price, and product 1's counterfactual share falls to $s_1' \approx 0.05$—less than half of what it was in market t. The demand curve is steeper in market t (product 1 has a larger customer base to lose) and flatter in market t'. A policymaker who used a population-average δ instead of market t's specific δ^* would get the wrong counterfactual shares.

The population intervention $\mathbb{E}[s_1 \mid \mathrm{do}(P_1 = 4)]$ averages the counterfactual shares over the distribution of ξ. The unit-level counterfactual $s_1(p_1' = 4; u)$ evaluates at market t's specific δ^*. These are different objects, and no amount of Rung 2 information—no matter how precisely estimated—can substitute for the market-specific δ^* that abduction provides. This example makes concrete what Theorem 2 formalizes. Abduction inverts observed shares to recover δ^*, the counterfactual price replaces the observed price, and the share function evaluated at the recovered δ^* yields the counterfactual quantity. Every computation uses the market's own data, not population averages.

OA.C. Detailed Proofs for Proposition 3

The formal statement appears as Proposition 3 in the main text; a proof sketch for part (a) and the full proof of part (b) appear there as well. Here we provide the detailed computation for part (a), showing how the identified set behaves in the crossing-curves example.

By hypothesis, 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'. Both are consistent with the observed data, so abduction does not pin down δ^* uniquely. At any such p', the identified set $\mathcal{I}(p'; e, \mathcal{L}_2) \supseteq \{\sigma(\delta_1, p', x^*; \theta), \sigma(\delta_2, p', x^*; \theta)\}$ contains at least two distinct points.

To see this concretely, return to the crossing-curves example of Proposition 1. The observed data is $(Q^* = 3, P^* = 3)$. Without model structure that identifies the type (i.e., without C2), the pre-image is $\{\xi_1, \xi_2\}$. The identified set for the counterfactual at $p' = 4$ is:

$$\left\{D(4, \xi_1),\ D(4, \xi_2)\right\} = \left\{12 - 3(4),\ 6 - 1(4)\right\} = \{0,\ 2\}.$$

At $p' = 2$: $\{12 - 6, 6 - 2\} = \{6, 4\}$. At $p' = 5$: $\{12 - 15, 6 - 5\} = \{-3, 1\}$. The width of the identified set grows with the distance from the observed price, because the two demand curves diverge as they move away from the crossing point. For the pass-through formula, the identified set for the demand slope is $\{-3, -1\}$, yielding a pass-through identified set of $\{0.25, 0.50\}$—a range of uncertainty that is large enough to affect policy conclusions.

OA.C.1. Geometry of partial identification

The identified set has a geometric structure that generalizes beyond the two-type example. Return to the crossing-curves case: $D(p, \xi_1) = 12 - 3p$, $D(p, \xi_2) = 6 - p$, observed data $(Q^* = 3, P^* = 3)$. Without C2, the pre-image is $\{\xi_1, \xi_2\}$. The identified set for the counterfactual quantity at price p' is

$$\mathcal{Q}(p') \;\equiv\; \{D(p', \xi) : \xi \in \{\xi_1, \xi_2\}\} \;=\; \{12 - 3p',\; 6 - p'\}.$$

At $p' = P^* = 3$, both elements coincide at $Q^* = 3$: the identified set is a singleton. As p' moves away from P^*, the two demand curves diverge at rate $|\beta(\xi_1) - \beta(\xi_2)| = 2$ per unit of price. The width of the identified set is

$$w(p') \;\equiv\; |D(p', \xi_1) - D(p', \xi_2)| \;=\; 2\,|p' - P^*|.$$

At $p' = 4$: $w = 2$. At $p' = 5$: $w = 4$. At $p' = 2$: $w = 2$. The set widens linearly with the distance from the observation point. Figure OA.1 displays this.

The widening is not an artifact of linearity. Proposition 2(b) establishes the general result: in any setting with two candidate latent states and continuously differentiable demand, a first-order expansion around P^* gives

$$w(p') = \left|\frac{\partial D}{\partial p}(P^*, \xi_1) - \frac{\partial D}{\partial p}(P^*, \xi_2)\right| \cdot |p' - P^*| + o\big(|p' - P^*|\big).$$

The leading coefficient is the absolute difference in demand slopes at the two candidate types. In the linear example, $w(p') = |\beta(\xi_1) - \beta(\xi_2)| \cdot |p' - P^*|$ exactly. For the equilibrium counterfactuals of Online Appendix A, the identified set translates directly into policy uncertainty: pass-through is $\{0.25, 0.50\}$ (a ratio of 2) and the consumer surplus loss is $\{-0.66, -1.38\}$ (a ratio exceeding 2).[9] In differentiated-

[9] The pass-through identified set follows from substituting each candidate slope into $\rho = 1/(1 - S' \cdot D')$; the consumer surplus identified set follows from evaluating $\mathrm{CS}(P^* + \rho\tau) - \mathrm{CS}(P^*)$ for each type.

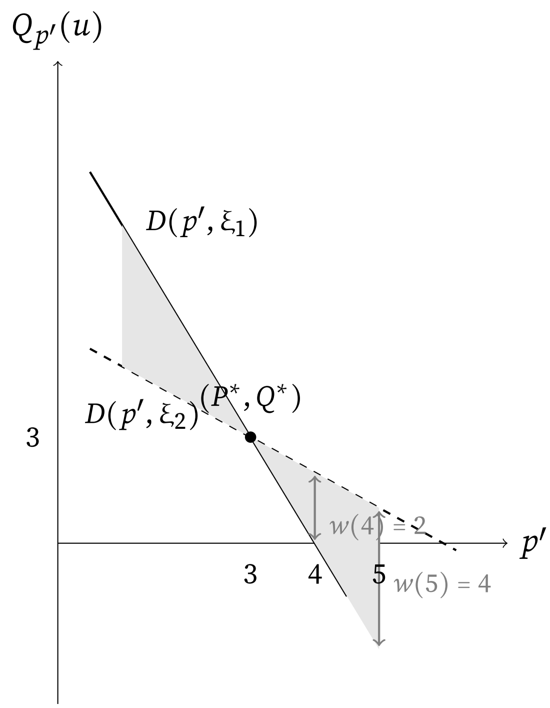

FIGURE OA.1. Identified set for the counterfactual $Q_{p'}(u)$ when C2 fails. The shaded region shows all demand responses consistent with the observed data ($P^* = 3, Q^* = 3$) and the two candidate latent types. The identified set is a singleton at $p' = P^*$ (where both demand curves agree) and widens linearly with the distance from the observed price. The rate of widening is $|\beta(\xi_1) - \beta(\xi_2)|$: the more heterogeneous the price sensitivity, the faster the identified set grows.

products settings, the pre-image $\sigma^{-1}(\{s^*\})$ may contain a continuum of candidate δ vectors, and the identified set depends on the curvature of σ and the local geometry of the pre-image.[10]

OA.D. Quantitative Illustration: Sensitivity Analysis

The model setup and merger results appear in Section 7 of the main text (Table 2). This appendix extends the analysis by varying δ_1 across a wider range.

Table OA.2 confirms the monotonicity. Varying δ_1 from 0.5 to 2.5 (a range of ±1 around the mean), the predicted merger price increase for product 2 ranges from 24% to 95%—a factor of 3.9. The diversion ratio from product 2 to product 1 rises

[10]When $\sigma(\cdot, p, x; \theta)$ is smooth but not globally injective, the pre-image is generically a $(J - k)$-dimensional manifold in $\mathbb{R}^J$, where k is the rank of $\partial\sigma/\partial\delta'$. See Molinari (2020) for a general treatment.

TABLE OA.2. Sensitivity of merger price effects to the unobserved demand state.

δ_1	s_1	Diversion ($2 \to 1$)	Merger Δp_2	Gap vs. pop. avg.
0.5	0.18	0.23	+24%	−27 pp
1.0	0.23	0.29	+36%	−15 pp
1.5	0.29	0.35	+51%	0 pp
2.0	0.35	0.39	+71%	+20 pp
2.5	0.41	0.43	+95%	+44 pp

Notes. Same model and marginal costs as Table 2. Only δ_1 (and hence ξ_1) varies; $\delta_2 = 0.5$ and marginal costs are identical across rows. "Pop. avg." is the row at $\delta_1 = 1.5$.

monotonically with δ_1, because higher unobserved quality draws a larger customer base to product 1 and makes it the primary outside option for product 2's marginal consumers. This is slope heterogeneity operating through the substitution matrix, exactly as Theorem 1(b) predicts.

OA.E. Connections and Extensions

Identification cost of abduction. Borusyak et al. (2026) study nonparametric identification of demand when product characteristics need not be exogenous. They show that price counterfactuals are identified using *recentered instruments*—functions that combine exogenous price instruments with possibly endogenous characteristics, recentered so validity comes from the exogenous price shocks. The key condition is *faithfulness*: a strength-of-variation condition under which inverse-demand candidates consistent with the recentered instrument test all imply the same price counterfactuals, even though demand is not fully identified.

The hierarchy provides a useful decomposition of what their conditions achieve. Standard price instruments can identify average price effects or interventional responses—the Rung 1-to-2 transition. Faithfulness serves the Rung 2-to-3 transition: it requires price-instrument variation strong enough in P, together with proxy variation in X rich enough about δ, that candidates consistent with those average effects cannot disagree about price counterfactuals holding δ fixed.[11]

[11] In our language, faithfulness is analogous to the extra condition needed to move from interventional average responses to price counterfactuals for a market with fixed latent demand conditions. Unlike C2, BCHL do not point-identify the full inverse demand function or the realized δ^*; the remaining transformation cancels in price-counterfactual calculations.

Without faithfulness, instruments may be adequate for average price effects yet still be too weak for market-specific price counterfactuals.

Their framework also makes explicit that the variation required for price counterfactuals is *more* than what average price effects require, but *less* than what full nonparametric demand identification requires—exactly the margin isolated by the hierarchy.

Counterfactual homogeneity. Chen (2025) reformulates the Berry–Haile market-level setup in potential-outcomes language, showing that the maintained structure appears as counterfactual homogeneity plus an additional functional-form restriction.[12]

In the SCM, a price-counterfactual version of this restriction follows from C1 and C2. If σ is common across markets (C1) and invertible (C2), then observing Q^* at conditions (P^*, X^*) uniquely determines δ^* via inversion, and δ^* determines counterfactual outcomes $\sigma(\delta^*, p', X^*; \theta)$ at any counterfactual price p'. The map from observed to counterfactual outcomes is therefore a function of (Q^*, P^*, X^*, p') alone, identical across markets. This is the price-counterfactual analogue of Chen's counterfactual homogeneity.

The SCM counterfactual $Q_p(u)$ and the Neyman–Rubin potential outcome $Q_i(p)$ are the same object, indexed differently.[13] Under C1–C2, the potential outcome can be written as

$$Q_i(p) = \sigma\big(\sigma^{-1}(Q_i(p^*), p^*, x_i; \theta),\ p,\ x_i;\ \theta\big),$$

which is a deterministic function of the observable $Q_i(p^*)$ and the known (p^*, x_i, p, θ). This means $\mathrm{Var}(Q_i(p) \mid Q_i(p^*), X_i) = 0$: knowing a market's outcome at one price determines its outcome at any other price. This is a conditional, price-only version of Chen's $\mathrm{Var}(Y_i(a) \mid Y_i(a')) = 0$ restriction, derived here as a consequence of the SCM structure, confirming that Chen's restriction captures the price-only implication of the abduction step.

When counterfactual homogeneity is only approximately true (as it must be in practice), Rung 3 outputs are best interpreted as model-based extrapolations

[12]This is our encoding of Chen's concept. His formulation is that "the relationship between counterfactual outcomes is assumed to be identical across markets."

[13]In the SCM, u denotes the unit's exogenous variables; in the potential outcomes framework, i indexes the unit directly. The equivalence $Q_p(u) = Q_i(p)$ holds because both denote "what market i would demand at price p, holding everything else fixed."

disciplined by Rung 2 variation: the structural model uses Rung 2 variation to learn σ and then extrapolates to Rung 3 via abduction.

References

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. 2021. "Foundations of Demand Estimation." *Handbook of Industrial Organization* 4 (1): 1–62.

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.

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

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.