

CompeteGrok Analysis Report

Query: explain BLP demand estimation step by step

Timestamp: 2025-12-17 15:52:51.754131

Routes: []

Agent Output

BLP Demand Estimation: Step-by-Step Explanation

The Berry-Levinsohn-Pakes (BLP) model (from Berry, Levinsohn, and Pakes, 1995) is a cornerstone of empirical industrial organization (IO) economics for estimating demand in markets with differentiated products. It extends the basic logit model to allow for random coefficients on product characteristics, addressing issues like price elasticities varying by consumer tastes and endogeneity of prices/characteristics. The estimation combines a nonlinear inversion (to recover mean utilities) with Generalized Method of Moments (GMM) to handle endogeneity.

Here's a **step-by-step breakdown** of the estimation process:

Step 1: Model Setup (Random Coefficients Logit Demand)

- **Consumer utility:** For consumer i and product j in market t :

$$u_{ijt} = x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}$$

- x_{jt} : Observed product characteristics (e.g., price, size, features). **Price is endogenous** due to unobserved quality ξ_{jt} .
- $\beta_i = \bar{\beta} + \Sigma\nu_i$: **Random coefficients** on characteristics, capturing consumer heterogeneity ($\nu_i \sim N(0, I)$, integrated over).
- ξ_{jt} : Unobserved (product-market-specific) quality, correlated with x_{jt} .
- ϵ_{ijt} : i.i.d. Type I Extreme Value (Gumbel), enabling closed-form choice probabilities.

- **Choice probability** (share for product j):

$$s_{jt}(\delta_{jt}) = \int_{\nu_i} \frac{\exp(\delta_{jt} + \mu_{ijt}(\nu_i))}{1 + \sum_k \exp(\delta_{kt} + \mu_{ikt}(\nu_i))} dG(\nu_i)$$

- $\delta_{jt} = x_{jt}\bar{\beta} + \xi_{jt}$: **Mean utility**.
- $\mu_{ijt}(\nu_i) = z_{jt}\Sigma\nu_i$: Random utility shocks (e.g., on price or other chars).
- Outside good share $s_{0t} = 1 - \sum_j s_{jt}$.

- Observed market shares s_{jt} are averages over consumers.

Key innovation: Unlike plain logit (constant elasticities), random coeffs allow flexible substitution patterns (e.g., Coke/Pepsi substitutes more for cola lovers).

Step 2: Inversion to Recover Mean Utilities δ_{jt}

- No closed-form $\delta = \delta(s)$ due to nonlinearity from integration.
- Solve **BLP contraction mapping** (fixed-point iteration) for each candidate parameter vector:

$$\delta_{jt}^{r+1} = \ln s_{jt} - \ln s_{0t} + \frac{1}{N_s} \sum_{h=1}^{N_s} \ln \left(1 + \sum_k \exp(\delta_{kt}^r + \mu_{kth}(\theta)) \right) - \ln \left(1 + \sum_k \exp(\delta_{kt}^r + \mu_{kth}(\theta)) \right).$$

- Simplified: Simulate shares $\hat{s}_{jt}^h(\delta^r)$ over N_s draws of ν_i^h , then iterate δ^{r+1} until $\|\delta^{r+1} - \delta^r\| < \epsilon$.
- This inverts $s_{jt} = s_{jt}(\delta_{jt}; \theta)$, where $\theta = (\bar{\beta}, \Sigma)$.

Step 3: GMM Objective Function (Handle Endogeneity)

- $\delta_{jt} = x_{jt}\theta + \xi_{jt}$, so regress δ on x linearly, but ξ correlated with x .
- Use **instruments** Z_{jt} : Exogenous vars correlated with x_{jt} but not ξ_{jt} (e.g., product attributes of rivals, cost shifters like input prices, Hausman instruments).
- **Moment conditions**:

$$E[Z'_{jt}\xi_{jt}] = 0 \implies E[Z'_{jt}(\delta_{jt} - x_{jt}\theta)] = 0$$

- **Objective** (minimize over θ):

$$Q(\theta) = \hat{\delta}(\theta)' Z(Z'Z)^{-1} Z' \hat{\delta}(\theta)$$

- Or weighted: $\hat{\delta}' W \hat{\delta}$, where $\hat{\delta}_{jt}(\theta) = \delta_{jt} - x_{jt}\theta$.
- W : Optimal weighting matrix (2-step: identity first, inverse covariance second).

Step 4: Full Estimation Algorithm (Nested Loops)

1. **Initialize** θ^0 (e.g., plain logit estimates).
2. **Inner loop (Inversion)**: For fixed θ^k , solve contraction for $\delta^k = \Delta(s; \theta^k)$.
3. **Outer loop (GMM)**: Compute residuals $\hat{\delta}^k = \delta^k - x\theta^k$, form moments $g(\theta^k) = Z'\hat{\delta}^k/T$, update:

$$\theta^{k+1} = \arg \min_{\theta} g(\theta)' W g(\theta)$$

(Closed-form linear projection).

4. **Iterate** until convergence ($|\theta^{k+1} - \theta^k| < \epsilon$).
5. **2-Step GMM**: Refit with optimal W from Step 1 residuals.

- Computationally intensive: Inversion per θ (simulations + fixed-point), nested optimization.

Step 5: Post-Estimation (Elasticities, etc.)

- **Own-price elasticity:**

$$\epsilon_{jj} = -\frac{\partial s_j}{\partial p_j} \frac{p_j}{s_j} = -p_j \left(\int \frac{\partial \Pr_j}{\partial \delta_j} dG(\nu) \right) \alpha$$

(Varies by product/consumer segment).

- Compute full demand matrix, markups (with supply side), welfare.
- Standard errors: Bootstrap or analytic GMM variance.

Practical Tips & Caveats

- **Software:** Python (pyblp), R (BLPEstimatorR), MATLAB (Nevo's code).
- **Data:** Market-level shares, characteristics (e.g., Nielsen scanner data).
- **Instruments:** Critical—weak IVs cause bias (test with Hansen J-stat).
- **Extensions:** BLP handles dynamics, entry, etc.
- **Common pitfalls:** Poor convergence (tune sim draws $N_s \approx 500 - 1000$, damping in contraction), multiple equilibria (rare, check monotonicity).

For code/examples, see pyblp docs or original papers. If you need math derivations, empirics, or caveats for a specific application, let me know!

Agent Output

Synthesis: See above.

Privacy: Ephemeral RAG; zero retention.

Disclaimer: Not legal advice. Models have caveats (e.g. IIA assumption). Verify 2025 data.

LaTeX: Inline x , display

$$E = mc^2$$

.