Online Supplemental Material

Normalization, Convergence, and Speed in Demand and Gravity Inversion

Miguel Zerecero
University of California, Irvine

Not for Publication

This Supplemental Material documents the numerical exercises in the paper. It describes the controlled random-coefficients demand experiments and full estimation benchmarks. All the codes to replicate the results can be found here: https://github.com/zerecero/norm_conv_speed. This supplement also records the theorem and proof by Bifulco, Glück, Krebs, and Kukharskyy (2025) used in the main text.

1 Utility and market shares

There are $J = 25$ inside goods and an outside good indexed by $j = 0$. Consumer i obtains utility

$$u_{ij} = \delta_j + \mu_{ij} + \varepsilon_{ij}, \qquad j = 0, 1, \ldots, J, \tag{1}$$

where ε_{ij} is independently distributed type-I extreme value. The outside good is normalized to $\delta_0 = \mu_{i0} = 0$. For each inside good, mean utility is

$$\delta_j(a) = a + \mathbf{x}_j'\bar{\boldsymbol{\beta}} + \bar{\alpha} p_j + \xi_j, \tag{2}$$

and the consumer-specific deviation is

$$\mu_{ij}(\kappa) = \kappa \left(\sigma_0 \nu_{i0} + \sum_{\ell=1}^{3} \sigma_\ell \nu_{i\ell} x_{j\ell} + \sigma_p \nu_{ip} p_j \right). \tag{3}$$

The common intercept a shifts every inside good relative to the outside option. The scale $\kappa \geq 0$ changes preference heterogeneity while holding the underlying standard-normal taste draws fixed.

Conditional choice probabilities and predicted market shares are

$$\pi_{ij}(\boldsymbol{\delta};\kappa) = \frac{\exp\{\delta_j + \mu_{ij}(\kappa)\}}{\sum_{k=0}^{J} \exp\{\delta_k + \mu_{ik}(\kappa)\}}, \qquad \mathcal{S}_j(\boldsymbol{\delta};\kappa) = \frac{1}{I}\sum_{i=1}^{I} \pi_{ij}(\boldsymbol{\delta};\kappa). \tag{4}$$

The simulations use $I = 1{,}000$ equally weighted draws.

2 Data-generating process

The demand-side design follows Appendix B.1 of Dubé, Fox, and Su (2012). The three product characteristics satisfy

$$\mathbf{x}_j \sim N(\mathbf{0}, \boldsymbol{\Sigma}_x), \qquad \boldsymbol{\Sigma}_x = \begin{pmatrix} 1 & -0.8 & 0.3 \\ -0.8 & 1 & 0.3 \\ 0.3 & 0.3 & 1 \end{pmatrix}.$$

Demand shocks ξ_j and price innovations η_j are independent standard normal draws, and price is generated as

$$p_j = \left| 0.5\xi_j + \eta_j + 1.1 \sum_{\ell=1}^{3} x_{j\ell} \right|.$$

The mean coefficients and random-coefficient variances are

$$\bar{\boldsymbol{\beta}} = (1.5, 1.5, 0.5)', \qquad \bar{\alpha} = -3, \qquad (\sigma_0^2, \sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_p^2) = (0.5, 0.5, 0.5, 0.5, 0.2).$$

The taste draws are independent standard normals:

$$\boldsymbol{\nu}_i = (\nu_{i0}, \nu_{i1}, \nu_{i2}, \nu_{i3}, \nu_{ip})' \sim N(\mathbf{0}, \mathbf{I}_5).$$

The random seed is 20260814. Product characteristics, prices, demand shocks, and taste draws are generated once and then held fixed in both experiments.

3 Controlled experiments

3.1 Moving the outside-good share

The first experiment fixes $\kappa = 1$ and varies only the common inside-good intercept a. The grid contains 41 equally spaced values from -5 to 5. For each value, equations (2)–(4) generate a new target share vector $\hat{\mathbf{s}}$. All other market primitives remain unchanged. Raising a makes every inside good more attractive and therefore reduces the outside-good share: in the simulated market, $\hat{s}_0$ falls from 0.9743 at $a = -5$ to 0.0032 at $a = 5$. This experiment isolates the small-reference-share mechanism.

3.2 Changing heterogeneity at a fixed outside-good share

The second experiment varies

$$\kappa \in \{0, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3\}.$$

At each value of κ, the common intercept is re-centered to solve

$$\mathcal{S}_0\big(\boldsymbol{\delta}(a(\kappa)); \kappa\big) = 0.10. \tag{5}$$

Thus the same product realization and standard-normal taste draws are used throughout, heterogeneity changes through equation (3), and the intercept offsets its effect on the aggregate outside share. The resulting intercepts range from 1.394 to 1.486. Holding $\hat{s}_0 = 0.10$ isolates the role of heterogeneity in individual outside-good probabilities from the direct reference-share effect.

4 Inversion and reported statistics

At every design point, target shares are computed at the true mean utilities. Both algorithms start from $\boldsymbol{\delta}^0 = \mathbf{0}$. Define

$$r_j(\boldsymbol{\delta}) = \log \hat{s}_j - \log \mathcal{S}_j(\boldsymbol{\delta}).$$

The reference-fixed algorithm keeps $\delta_0 = 0$ and updates $\delta_j^{t+1} = \delta_j^t + r_j(\boldsymbol{\delta}^t)$ for $j \geq 1$. The symmetric algorithm updates every equation and then restores the same normalization, so for

$j \geq 1$ it uses

$$\delta_j^{t+1} = \delta_j^t + r_j(\boldsymbol{\delta}^t) - r_0(\boldsymbol{\delta}^t).$$

Both stop when the common distance

$$d\big(\hat{\mathbf{s}}, \boldsymbol{\mathcal{S}}(\boldsymbol{\delta})\big) = \max_{j \neq 0} \left|\log \hat{s}_j - \log \mathcal{S}_j(\boldsymbol{\delta})\right|$$

is at most 10^{-12}. The outside-share equation is omitted because one share equation is redundant; the same equation is omitted for both methods. One market-share calculation is one evaluation of the entire vector $\boldsymbol{\mathcal{S}}(\boldsymbol{\delta})$. The local convergence factors are computed at the true solution. The figures also report $1 - \hat{s}_0$ and the reference-fixed lower bound

$$1 - \hat{s}_0 + \frac{\mathrm{Var}(\pi_{i0})}{1 - \hat{s}_0}.$$

5 Full-estimation Monte Carlo exercise

This exercise uses the comparatively difficult demand-estimation environment developed in Section 7 of Dubé et al. (2012), following the implementation of Pál and Sándor (2023) and Monardo (2025). There are $T = 50$ markets and $J = 25$ inside goods. The coefficients and random-coefficient variances are those reported in Section 2 above, while the common inside-good intercept takes the values

$$a \in \{-4, -2, 0, 2, 4\}.$$

Prices are generated according to

$$p_{jt} = 3 + \sum_{\ell=1}^{3} x_{j\ell} + w_{jt} + 1.5\xi_{jt} + e_{jt}, \qquad w_{jt}, \xi_{jt} \sim N(0,1), \quad e_{jt} \sim U(0,5).$$

The instruments include the cost shifter and the quadratic and cross-product differentiation instruments used by Pál and Sándor (2023). Market shares are computed with their fixed sample of 1,331 five-dimensional quasi-random standard-normal draws. The same draw matrix is used in every market and every Monte Carlo replication.

For each intercept, I generate 20 independently seeded datasets and estimate each from five common nonlinear-parameter starting values, giving 100 executions per algorithm. The first start is one-half of the absolute value of the homogeneous-logit two-stage least-squares estimates. The other four starts multiply that vector component by component by independent $U(0, 0.5)$ draws, following Pál and Sándor (2023) and Monardo (2025).

The estimations are performed in PyBLP (Conlon and Gortmaker, 2020). Both algorithms use the same simulated data, integration draws, instruments, starting values, one-step GMM problem, L-BFGS-B outer optimizer, analytic derivatives, and 10-process parallelization across markets. Only the inversion map changes. Both maps carry the most recently computed mean utilities to the next evaluation of the GMM objective and use the stopping distance defined above. Each market inversion has a common safeguard of 70,000 iterations.

Convergence is evaluated over all 100 executions for each intercept and algorithm. Runtime and mean inner iterations use only replications in which all five starts converge under both algorithms. Runtime is the total wall-clock time for the five sequential starts, averaged across these paired-complete replications. Mean inner iterations are the total iterations divided by the number of market inversions over the same sample. Executions that reach the inversion safeguard count as nonconvergent but are not assigned the cap when computing conditional speed. The main paper reports the resulting convergence, runtime, and iteration comparisons.

6 Original BLP full-estimation benchmark

The original BLP exercise follows the initial automobile supply-and-demand specification in Figure 6 of Conlon and Gortmaker (2020), as implemented in PyBLP. It uses the packaged original BLP data and importance-sampling integration rule, with 20 markets, 2,217 product observations, and 200 agents per market. The exercise does not implement the subsequent optimal-instrument and Halton- draw extensions or the post-estimation elasticity and markup calculations. Both runs use the same data, starting parameters, bounds, two-step GMM procedure, clustered inference, and analytic derivatives; only the inner inversion map changes. Both methods converge in 66 objective evaluations to an objective value of 497.3357. The symmetric map reduces market-share calculations from 401,818 to 356,930, or 11.2 percent, and runtime falls from 140.2 to 127.3 seconds.

7 A theorem and proof from Bifulco et al. (2025)

Following Bifulco et al. (2025), let (Z, d) be a metric space, let $G : Z \to Z$, and let z^* be a fixed point of G. The space (Z, d) is *connected* if it cannot be written as the union of two nonempty, open, disjoint subsets. The map G is *globally non-expansive* if

$$d(G(z), G(z')) \leq d(z, z')$$

for all $z, z' \in Z$; equivalently, G is Lipschitz continuous with Lipschitz constant at most one. The fixed point z^* is *Lyapunov stable* if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d(z, z^*) < \delta$ implies $d(G^n(z), G^n(z^*)) < \varepsilon$ for all integers $n \geq 0$. It is *locally attractive* if there exists $\delta > 0$ such that $d(z, z^*) < \delta$ implies $G^n(z) \to z^*$ as $n \to \infty$. Finally, z^* is *locally asymptotically stable* if it is both Lyapunov stable and locally attractive, and *globally attractive* if $G^n(z) \to z^*$ for every $z \in Z$.

Theorem 1 (Theorem C.1 from Bifulco et al. (2025))**.** *Let (Z, d) be a metric space which is connected. Let $G : Z \to Z$ be a Lipschitz continuous function with Lipschitz constant at most 1, and let $z^* \in Z$ be a fixed point of G which is locally asymptotically stable.*

Then z^* is globally attractive, i.e., for each $z \in Z$ the sequence $G^n(z)$ converges to z^* as $n \to \infty$. In particular, z^* is the only fixed point of G.

Proof. Let B denote the basin of attraction of the fixed point z^*, i.e., the set of all $z \in Z$ for which $G^n(z) \to z^*$ as $n \to \infty$. Note that B is non-empty since $z^* \in B$.

Step 1: B is open. Since z^* is locally asymptotically stable, there exists an open neighborhood $\mathcal{U}$ of z^* such that $\mathcal{U} \subseteq B$. Let $z_0 \in B$. Since $G^n(z_0) \to z^*$, there exists N such that $G^N(z_0) \in \mathcal{U}$. By continuity of G^N, there exists an open neighborhood $\mathcal{V}$ of z_0 such that $G^N(\mathcal{V}) \subseteq \mathcal{U}$. Every point in $\mathcal{V}$ therefore belongs to B. Hence B is open.

Step 2: B is closed. Let $(z_k)_{k \in \mathbb{N}}$ be a sequence in B that converges to some $z \in Z$. We need to show $z \in B$, i.e., $G^n(z) \to z^*$ as $n \to \infty$.

Let $\varepsilon > 0$ be arbitrary. Since $z_k \to z$, choose k_0 such that $d(z_{k_0}, z) < \varepsilon/2$.

Since $z_{k_0} \in B$, we have $G^n(z_{k_0}) \to z^*$ as $n \to \infty$. Therefore, there exists n_0 such that for all $n \geq n_0$,

$$d(G^n(z_{k_0}), z^*) < \varepsilon/2.$$

Using non-expansiveness, for $n \geq n_0$,

$$\begin{aligned} d(G^n(z), z^*) &\leq d(G^n(z), G^n(z_{k_0})) + d(G^n(z_{k_0}), z^*) \\ &\leq d(z, z_{k_0}) + \varepsilon/2 \\ &< \varepsilon. \end{aligned}$$

Since this holds for arbitrary $\varepsilon > 0$, we have $G^n(z) \to z^*$, so $z \in B$. Hence B is closed.

Step 3: Connected space. Since Z is connected and B is both open and closed (and non-empty), we must have $B = Z$. Therefore, z^* is globally attractive.

Step 4: Uniqueness. Suppose $\widetilde{z}^*$ is another fixed point. Since $B = Z$, global attractiveness of z^* implies

$$G^n(\widetilde{z}^*) \to z^*.$$

But $G^n(\widetilde{z}^*) = \widetilde{z}^*$ for every n, so $\widetilde{z}^* = z^*$. Therefore the fixed point is unique. □

References

BIFULCO, P., J. GLÜCK, O. KREBS, AND B. KUKHARSKYY (2025): "Single and Attractive: Uniqueness and Stability of Economic Equilibria Under Monotonicity Assumptions," Tech. rep., CESifo.

CONLON, C. AND J. GORTMAKER (2020): "Best practices for differentiated products demand estimation with pyblp," *The RAND Journal of Economics*, 51, 1108–1161.

DUBÉ, J.-P., J. T. FOX, AND C.-L. SU (2012): "Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation," *Econometrica*, 80, 2231–2267.

MONARDO, J. (2025): "Comparing methods for estimating random coefficient logit demand models," Working paper, University of Bristol.

PÁL, L. AND Z. SÁNDOR (2023): "Comparing procedures for estimating random coefficient logit demand models with a special focus on obtaining global optima," *International Journal of Industrial Organization*, 88, 102950.