

# Stage-4 Report — Distilling Personalized Pricing into Closed-Form Rules via Symbolic Regression

Pipeline state: Stage 4 ADVANCE (Gate-4 score 72 / 100, top-3 marketing tier). Pipeline halted at user request before Stage 5 (paper writing).

---

## Abstract

Hierarchical Bayesian demand models generate personalized pricing policies that substantially outperform uniform pricing, but the resulting price schedules depend on thousands of individual- or market-specific parameters that field managers cannot interpret and regulators cannot audit. We ask how much of the profit premium from personalized pricing survives when a high-dimensional Bayesian pricing policy is distilled into a single closed-form equation using symbolic regression. Across NielsenIQ scanner data for cereal, yogurt, and fluid milk, a five-parameter equation at the three-digit ZIP-code level captures 30 to 62 percent of the profit gain from full Bayesian personalization in cereal and yogurt—matching regression and decision-tree benchmarks while fitting on a single line. We also identify two clear boundaries where symbolic compression fails: finely disaggregated retail markets with noisy estimates, and categories with extreme, tail-driven price sensitivity.

---

## Contribution statement

1. A formal compression framework for HB pricing. We give the first treatment of symbolic-regression distillation of structural demand-derived pricing policies, with (i) a compression-loss decomposition lemma linking rule MSE to profit-loss via second-order expansion, (ii) an existence proposition for a deployable rule with non-trivial recovery, and (iii) a small-sample failure proposition that characterizes when symbolic compression is unsafe (Theory §7–§9).
2. An empirical map of the compression frontier. Across three CPG categories at three aggregation grains, we trace SR's recovery of HB-optimal lift over uniform pricing, with bootstrap variance benchmarked against OLS, ridge, CART ( $d=2,3,5$ ), and random forest. SR-5 occupies a unique point on the frontier: closed-form, single equation, accuracy competitive with much larger models.
3. A characterizing counterexample. Milk/ZIP3 has high in-sample  $R^2$  (0.79) and negative recovery (−33 %) because elasticity heterogeneity is tail-concentrated. The paper turns this into a *diagnostic* — posterior-kurtosis is a pre-screen for whether SR distillation will deliver — rather than presenting it as a robustness failure.
4. A model-implied / IPW evaluator-disagreement diagnostic. We show that under chain-uniform pricing, Smith-Seiler-Aggarwal IPW disagrees in sign with the model-implied profit ranking. We attribute the disagreement to near-degenerate propensities — extending the DellaVigna-Gentzkow (2019) chain-pricing observation into the personalization-evaluation literature — and recommend reporting both with a propensity-overlap diagnostic.
5. A deployable artifact. Three one-line equations (see Table 2) are usable today by a category manager pricing cereal or yogurt baseline SKUs at the 3-digit-ZIP level.

---

## Key results

Table 1 — Headline lift (model-implied; HB-optimal and SR-distilled vs uniform-best)

| Category   | Grain | n  | HB vs uniform | SR-5 vs uniform | SR-5 recovery of HB |
|------------|-------|----|---------------|-----------------|---------------------|
| cereal RTE | UPC   | 10 | +13.70 %      | n/a             | —                   |

| Category   | Grain    | n   | HB vs uniform | SR-5 vs uniform | SR-5 recovery of HB |
|------------|----------|-----|---------------|-----------------|---------------------|
| cereal RTE | ZIP3     | 785 | +6.25 %       | +1.90 %         | +30 %               |
| cereal RTE | retailer | 188 | +1.18 %       | −0.57 %         | −48 %               |
| yogurt     | UPC      | 10  | +0.19 %       | —               | —                   |
| yogurt     | ZIP3     | 693 | +9.33 %       | +3.79 %         | +41 %               |
| yogurt     | retailer | 160 | +1.61 %       | −12.25 %        | −759 %              |
| milk       | UPC      | 10  | +13.33 %      | —               | —                   |
| milk       | ZIP3     | 753 | +3.66 %       | −1.21 %         | −33 %               |
| milk       | retailer | 58  | +0.09 %       | −9.37 %         | −10 844 %           |

Bold rows are the deployable cases. UPC is motivating; retailer / milk-ZIP3 are the characterizing failure regimes.

Table 2 — SR-5 distilled equations at ZIP3

| Category   | Equation                                                                 | R <sup>2</sup> | Recovery |
|------------|--------------------------------------------------------------------------|----------------|----------|
| cereal RTE | $\hat{p} = 6.36 + 0.29 \cdot \text{meanP} + 0.13 \cdot n_{\text{hh}}$    | 0.55           | +30 %    |
| yogurt     | $\hat{p} = 1.14 + 0.155 \cdot \text{meanP} - 0.024 \cdot \text{meanP}^2$ | 0.58           | +41 %    |
| milk       | $\hat{p} = 3.32 + 0.46 \cdot \text{meanP} + 0.062 \cdot n_{\text{hh}}$   | 0.79           | −33 %    |

meanP = mean observed price; n\_hh = number of HMS panelist households. Mean observed price dominates demographics in every category (Implication I2).

Table 3 — Frontier at cereal/ZIP3, n=785, 15-bootstrap SDs

| Distiller     | Complexity (params) | Refit R <sup>2</sup> | Profit  | Lift    | Recovery | Bootstrap SD |
|---------------|---------------------|----------------------|---------|---------|----------|--------------|
| uniform-best  | 1                   | —                    | \$4 408 | 0.00 %  | 0 %      | —            |
| SR-5          | 5                   | 0.467                | \$4 502 | +2.15 % | +34 %    | \$7.81       |
| OLS           | 12                  | 0.578                | \$4 494 | +1.96 % | +31 %    | \$7.81       |
| CART          | ~7                  | 0.524                | \$4 492 | +1.93 % | +31 %    | \$5.89       |
| d=3           |                     |                      |         |         |          |              |
| CART          | ~20                 | 0.613                | \$4 513 | +2.40 % | +38 %    | \$7.52       |
| d=5           |                     |                      |         |         |          |              |
| random forest | ~6 995              | 0.700                | \$4 541 | +3.03 % | +48 %    | \$5.64       |
| HB-optimal    | 785                 | —                    | \$4 683 | +6.25 % | 100 %    | —            |

SR-5 is the only closed-form continuous-policy point on the frontier; CART produces if-then trees, RF is non-interpretable.

Table 4 — Evaluator disagreement (IPW vs model-implied at ZIP3)

| Category | IPW(uniform) | IPW(HB) | Model(uniform) | Model(HB) | Model lift | IPW says     |
|----------|--------------|---------|----------------|-----------|------------|--------------|
| cereal   | \$42.23      | \$26.90 | \$4 408        | \$4 683   | +6.25 %    | uniform wins |
| yogurt   | \$24.24      | \$17.01 | \$2 089        | \$2 284   | +9.33 %    | uniform wins |
| milk     | \$62.39      | \$33.15 | \$5 957        | \$6 532   | +9.65 %    | uniform wins |

Sign reversal driven by near-degenerate propensities under chain-uniform pricing. Diagnostic, not refutation; reported alongside model-implied.

Figure 1 (sketched) — accuracy / complexity frontier at cereal/ZIP3

(deferred to Stage 5; conceptually:  $x$ -axis  $\log(\text{parameter count})$ ,  $y$ -axis recovery %; SR-5 is on the Pareto frontier between constant and CART- $d=3$ , dominating OLS strictly on parameter count.)

Figure 2 (sketched) — recovery vs aggregation grain

(retailer  $n \approx 200$ , recovery negative for SR-10/15; ZIP3  $n \sim 700$ , recovery 30–62 % for SR-5; UPC  $n=10$ , distillation undefined.)

## Posterior diagnostics for theory propositions

| Theory claim                                                     | Evidence                                                         | Status     |
|------------------------------------------------------------------|------------------------------------------------------------------|------------|
| Prop. 1 (existence of deployable rule with non-trivial recovery) | Table 1 ZIP3 cereal/yogurt rows                                  | Confirmed  |
| Prop. 2 (small- $n$ failure when $K \approx 10$ )                | Table 1 retailer rows; SR-10/15 recoveries below uniform         | Confirmed  |
| §7 Lemma (quadratic compression-loss)                            | Recovery scales with refit $R^2$ across distillers in Table 3    | Consistent |
| I2 (mean-price dominance over demographics)                      | Table 2 equations                                                | Confirmed  |
| I5 (tail-concentrated heterogeneity breaks SR)                   | Milk ZIP3 $R^2=0.79$ , recovery $-33\%$ , posterior kurtosis 7.4 | Confirmed  |

## What the pipeline did not do (transparency)

- Did not refit demand with PyBLP (queued; seed has working pipeline).
- Did not run a week-split holdout (queued; bootstrap SD reported instead).
- Did not pool 2008–2012 HMS (queued; could tighten retailer-grain posteriors).
- Did not produce the LaTeX paper — user halted before Stage 5.

## Files of record

| Path                                                                                                                                              | Purpose                                  |
|---------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|
| output/stage0/problem_statement.md,<br>literature_map.md                                                                                          | Stage 0                                  |
| output/stage1/selected_idea.md,<br>novelty_check_idea.md, idea_prototype.md                                                                       | Stage 1                                  |
| output/stage2/theory_v1.md, math_audit_v1.md,<br>novelty_check_v1.md                                                                              | Stage 2                                  |
| output/stage2b/exploration_report.md                                                                                                              | Stage 2b                                 |
| output/stage3/implications.md                                                                                                                     | Stage 3                                  |
| output/stage3a/empirical_results.md                                                                                                               | Stage 3a                                 |
| output/stage4/self_attack_v1.md,<br>scorer_v1.md, REPORT.md                                                                                       | Stage 4                                  |
| output/seed/Notes/aggregate_pricing_results.md,<br>distillation_frontier_results.md,<br>ipw_vs_model_implied.md,<br>hb_sr_distillation_results.md | Underlying empirical evidence            |
| process_log/pipeline_state.json                                                                                                                   | Pipeline state, scored 72 / 100, ADVANCE |