Post-Estimation Adjustments in Data-Driven Decision-Making with Applications in Pricing

Michael Albert
Darden School of Business
University of Virginia
albertm@darden.virginia.edu

Max Biggs
Darden School of Business
University of Virginia
biggsm@darden.virginia.edu

Ningyuan Chen
University of Toronto
ningyuan.chen@utoronto.ca

Guan Wang
University of Toronto
wguan.wang@rotman.utoronto.ca

Abstract

The predict-then-optimize (PTO) framework is a standard paradigm for data-driven decision-making. However, by separating estimation from optimization, it can yield suboptimal decisions, especially when the objective function is asymmetric with respect to estimation errors. This asymmetry, where underestimation and overestimation incur different penalties, systematically degrades the expected outcome. We propose a simple, data-driven post-estimation adjustment that corrects the initial parameter estimate to account for the downstream decision task. We show that for a broad class of problems satisfying a specific curvature condition—a constant ratio between the third and second derivatives of the surrogate reward—our adjustment takes a simple closed-form and is guaranteed to asymptotically improve performance over standard PTO. This condition holds for many canonical pricing models. We develop two data-driven methods, a closed-form plug-in rule and a more general bootstrap procedure, and show that they can surprisingly outperform an oracle benchmark. Numerical experiments on a data-driven pricing problem validate our theory, demonstrating consistent revenue improvements, particularly in the small-sample regimes common in practice.

1 Introduction

In data-driven decision-making, a common workflow is to first use historical data to estimate unknown parameters of a model and then plug these estimates into an optimization problem to find the best decision. This two-stage process, known as predict-then-optimize (PTO), is popular for its simplicity and modularity. However, its core weakness is that the estimation step is agnostic to the downstream optimization task. This can lead to suboptimal decisions, as statistical accuracy does not always translate to decision quality [10].

We focus on a class of problems where this suboptimality arises from an asymmetry in the objective function with respect to estimation errors. Consider a canonical data-driven pricing problem: a firm wants to set a price p to maximize revenue $p \cdot d(p)$, where demand is linear: $d(p) = a - \theta p$. The price sensitivity θ is unknown. The optimal price is $p(\theta) = a/(2\theta)$. In the PTO framework, one first obtains an estimate $\hat{\theta}$ from data and then sets the price to $p(\hat{\theta}) = a/(2\hat{\theta})$. The resulting revenue,

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: ML×OR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

viewed as a function of the estimate $\hat{\theta}$, is the *surrogate reward*:

$$R_\theta(\hat{\theta}) = p(\hat{\theta}) \cdot (a - \theta p(\hat{\theta})) = \frac{a^2}{2\hat{\theta}} - \frac{a^2\theta}{4\hat{\theta}^2}. \tag{1}$$

As shown in Figure 1, this function is highly asymmetric. While the reward is maximized at the true value $\hat{\theta} = \theta$, the penalty for underestimating θ is far more severe than for an equivalent overestimation. Consequently, even for an unbiased estimator ($\mathbb{E}[\hat{\theta}] = \theta$), the expected reward $\mathbb{E}[R_\theta(\hat{\theta})]$ is degraded by estimation noise. This suggests that a corrective adjustment that intentionally biases $\hat{\theta}$ could improve expected performance.

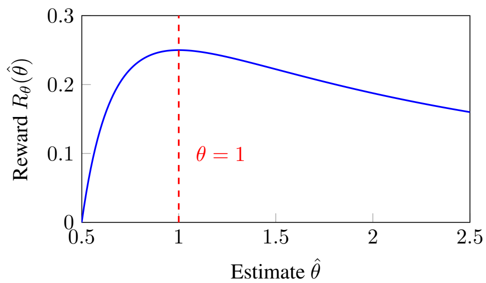

Figure 1: The surrogate reward $R_\theta(\hat{\theta})$ for the linear pricing problem with $a = 1, \theta = 1$. The function is asymmetric around its maximum at $\hat{\theta} = \theta$. Underestimation ($\hat{\theta} < \theta$) incurs a larger loss than overestimation.

In this paper, we develop a general framework for such a post-estimation adjustment. Our method preserves the modularity of PTO by applying a simple, data-driven correction after the initial prediction. Our contributions are: 1) We derive an *oracle adjustment* that provides a theoretical benchmark for improvement. 2) We identify a key structural condition on the surrogate reward—that the ratio of its third to second derivative is constant—which simplifies the adjustment and holds for many common pricing models. 3) We develop two practical data-driven methods: a simple *Plug-in* rule and a more general *Bootstrap* procedure. 4) We prove these methods uniformly improve upon PTO and can surprisingly outperform the oracle. 5) We validate our findings with numerical experiments, showing our method improves revenue, especially in small-sample settings.

1.1 Related Work

Our work contributes to decision-focused learning, also known as smart predict-then-optimize [10], end-to-end optimization [9], or integrated learning and optimization [16]. Prominent methods include decision rule learning [4, 5] and differentiable optimization layers [1]. Our work addresses settings like pricing where key parameters (e.g., price sensitivity) are latent and not directly observed. Another related paradigm is Operational Data Analytics (ODA) [11], which leverages homogeneity conditions in decision problems. While ODA provides uniformly optimal data-driven decisions for problems satisfying its specific scaling properties [7, 17], our linear pricing example does not fit this framework. Our work is particularly relevant to data-driven pricing. Prior research has explored regularizing contextual parameters for optimal regret rates [3, 15] or using doubly robust estimation for contextual pricing [6, 20]. Other works establish worst-case revenue bounds with limited observations [12, 2, 8]. In contrast, we focus on average-case performance and provide a closed-form adjustment. Our approach directly addresses the "optimizer's curse" [18, 14], where optimized objectives based on uncertain parameters systematically overestimate true performance. Unlike general debiasing methods [19], our analysis leverages specific demand function forms to precisely characterize the adjustment and its improvement rate. This connects to the broader statistics literature on regularization [21, 13], but our method is decision-aware, sometimes leading to expansion rather than typical shrinkage.

2 A Framework for Post-Estimation Adjustment

We consider a sequence of estimators $\hat{\theta}_n$ from datasets of size n. Our goal is to find an adjustment function $\pi(\cdot)$ that improves the expected reward, $\mathbb{E}[R_\theta(\pi(\hat{\theta}_n))] > \mathbb{E}[R_\theta(\hat{\theta}_n)]$, uniformly over θ.

2.1 Setup and Assumptions

We impose mild regularity conditions on the estimator and the surrogate reward function.

Assumption 1 *As $n \to \infty$, the estimator $\hat{\theta}_n$ for θ is root-n consistent with diminishing $o(n^{-1})$ bias and well-behaved higher moments.*

Let $R_\theta^{(k)}(x)$ be the k-th derivative of $R_\theta(x)$ with respect to x.

Assumption 2 *For any true parameter θ, we have $R_\theta^{(1)}(\theta) = 0$ and $R_\theta^{(2)}(\theta) < 0$; There exists a constant C, independent of θ, such that $\frac{R_\theta^{(3)}(\theta)}{R_\theta^{(2)}(\theta)} = \frac{C}{\theta}$; The fifth derivative $R_\theta^{(5)}(\cdot)$ is locally bounded.*

Part (2) is our key structural assumption. It holds for many canonical economic models, including linear ($C = -6$), log-linear ($C = -4$), and power-law demand models.

2.2 The Oracle Adjustment

We first consider an idealized oracle setting where the true θ and its asymptotic variance parameter σ_θ^2 are known. We analyze a multiplicative adjustment of the form $\pi(\hat{\theta}_n) = \hat{\theta}_n(1 + \lambda/n)$.

Proposition 1 (Oracle Adjustment) *Suppose Assumptions 1 and 2 hold. The optimal oracle adjustment coefficient is $\lambda^* = -\frac{(C+2)\sigma_\theta^2}{2\theta^2}$, and the asymptotic improvement in expected reward over PTO is*

$$\lim_{n\to\infty} n^2 \left(\mathbb{E}\left[R_\theta\left(\hat{\theta}_n(1 + \tfrac{\lambda^*}{n})\right)\right] - \mathbb{E}[R_\theta(\hat{\theta}_n)]\right) = -\frac{R_\theta^{(2)}(\theta)(C+2)^2\sigma_\theta^4}{8\theta^2} \geq 0. \tag{2}$$

The proof relies on a Taylor expansion of the expected reward. The constant C dictates the nature of the adjustment: if $C > -2$, the adjustment shrinks the estimate ($\lambda^* < 0$), but if $C < -2$, as in our pricing example ($C = -6$), it expands the estimate ($\lambda^* > 0$).

2.3 Data-Driven Adjustments

The oracle is impractical. We now present two data-driven methods that learn the adjustment from data.

Plug-in Adjustment. This method replaces the unknown quantities in the oracle formula with their sample estimates. Let $\hat{\sigma}_\theta^2$ be a consistent estimator for σ_θ^2. We propose an adjustment coefficient of the form $\lambda_n = k\frac{\hat{\sigma}_\theta^2}{\hat{\theta}_n^2}$, where k is a constant to be optimized.

Proposition 2 (Data-driven Plug-in Adjustment) *Suppose Assumptions 1 and 2 hold, along with standard regularity conditions on $\hat{\sigma}_\theta^2$. The optimal data-driven adjustment is achieved with $k^* = (2 - C)/2$, yielding*

$$\lim_{n\to\infty} n^2 \left(\mathbb{E}\left[R_\theta\left(\hat{\theta}_n(1 + \tfrac{\lambda_n^*}{n})\right)\right] - \mathbb{E}[R_\theta(\hat{\theta}_n)]\right) = -\frac{R_\theta^{(2)}(\theta)(2-C)^2\sigma_\theta^4}{8\theta^2} \geq 0. \tag{3}$$

Comparing the improvement terms in (2) and (3) reveals a surprising result: the oracle's improvement factor $(C+2)^2$ is replaced by $(2-C)^2$. For many problems where $C < 0$ (e.g., $C = -6$), we have $(2-C)^2 > (C+2)^2$. This means the data-driven adjustment achieves a *larger* asymptotic improvement than the oracle.

Bootstrap Adjustment. As a more general alternative, we can use a bootstrap procedure to find the adjustment. The intuition is to create a "bootstrap world" where the PTO estimate $\hat{\theta}$ is treated as the ground truth, and then find the adjustment λ that performs best in this simulated world. This method is particularly useful when the plug-in formula is complex or assumptions are hard to verify, as in multi-parameter problems.

Algorithm 1 Bootstrap-based Adjustment

Require: Observed data $\mathcal{D}_n$, PTO estimator $\hat{\theta}$, number of bootstrap draws B.

for $b = 1$ **to** B **do**

  Generate bootstrap sample $\mathcal{D}_n^{*b}$ from a parametric model using $\hat{\boldsymbol{\theta}}$.

  Compute bootstrap PTO estimator $\hat{\boldsymbol{\theta}}^{*b}$ from $\mathcal{D}_n^{*b}$.

end for

Define bootstrap objective: $\hat{R}_B(\lambda) \leftarrow \frac{1}{B}\sum_{b=1}^{B} R_{\hat{\boldsymbol{\theta}}}\big(\hat{\theta}^{*b}(1+\lambda/n)\big)$.

Compute $\hat{\lambda} \leftarrow \arg\max_\lambda \hat{R}_B(\lambda)$ via grid search.

return Adjusted estimator $\hat{\theta}(1+\hat{\lambda}/n)$.

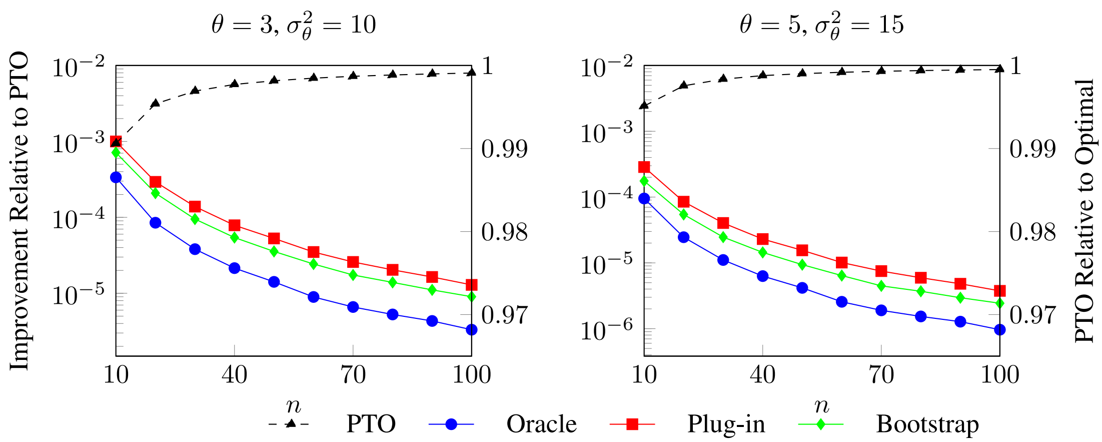

Figure 2: Performance in the linear pricing model. The left y-axis (log scale) shows the relative revenue improvement of adjusted policies over PTO. The right y-axis shows the performance of PTO relative to the optimum. The data-driven adjustments (Plug-in, Bootstrap) consistently and significantly outperform both PTO and the Oracle, as predicted by theory.

3 Numerical Experiment

We validate our framework on the linear demand pricing problem ($d(p) = a - \theta p$).

Setup. We set $a = 60$ and test two scenarios: ($\theta = 3, \sigma_\epsilon^2 = 10$) and ($\theta = 5, \sigma_\epsilon^2 = 15$), where σ_ϵ^2 is the variance of the demand noise. For sample sizes $n \in \{10, \ldots, 100\}$, we generate 10^5 datasets, estimate $\hat{\theta}$ via OLS, and compare four policies: **PTO**, **Oracle**, **Plug-in**, and **Bootstrap**. We measure the performance of PTO relative to the clairvoyant optimum and the percentage improvement of the adjusted policies over PTO.

Results. Figure 2 shows the results. The findings strongly support our theory:

1. **All adjustments improve performance.** The Oracle, Plug-in, and Bootstrap methods consistently yield positive revenue improvements over PTO across all settings.
2. **Data-driven methods outperform the Oracle.** As predicted for $C = -6$, where $(2-C)^2 = 64$ is much larger than $(C+2)^2 = 16$, the Plug-in and Bootstrap adjustments provide a substantially larger improvement than the Oracle.
3. **Gains are largest for small samples.** The improvements are most pronounced for small n, where estimation uncertainty is highest. This highlights the practical value of our method in data-scarce environments.

References

[1] Brandon Amos and J Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In *International conference on machine learning*, pages 136–145. PMLR, 2017.

[2] Moshe Babaioff, Yannai A Gonczarowski, Yishay Mansour, and Shay Moran. Are two (samples) really better than one? In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 175–175, 2018.

[3] Gah-Yi Ban and N. Bora Keskin. Personalized Dynamic Pricing with Machine Learning: High-Dimensional Features and Heterogeneous Elasticity. *Management Science*, 67(9):5549–5568, September 2021. ISSN 0025-1909. doi: 10.1287/mnsc.2020.3680.

[4] Gah-Yi Ban and Cynthia Rudin. The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1):90–108, 2019.

[5] Dimitris Bertsimas and Nihal Koduri. Data-driven optimization: A reproducing kernel hilbert space approach. *Operations Research*, 70(1):454–471, 2022.

[6] Victor Chernozhukov, Mert Demirer, Greg Lewis, and Vasilis Syrgkanis. Semi-parametric efficient policy learning with continuous actions. *Advances in Neural Information Processing Systems*, 32, 2019.

[7] Leon Yang Chu, Qi Feng, J George Shanthikumar, Zuo-Jun Max Shen, and Jian Wu. Solving the price-setting newsvendor problem with parametric operational data analytics (oda). *Management Science*, 2023.

[8] Constantinos Daskalakis and Manolis Zampetakis. More revenue from two samples via factor revealing sdps. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 257–272, 2020.

[9] Priya Donti, Brandon Amos, and J Zico Kolter. Task-based end-to-end model learning in stochastic optimization. *Advances in neural information processing systems*, 30, 2017.

[10] Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1):9–26, 2022.

[11] Qi Feng and J George Shanthikumar. The framework of parametric and nonparametric operational data analytics. *Production and Operations Management*, 32(9):2685–2703, 2023.

[12] Hu Fu, Nicole Immorlica, Brendan Lucier, and Philipp Strack. Randomization beats second price as a prior-independent auction. In *Proceedings of the sixteenth ACM conference on economics and computation*, pages 323–323, 2015.

[13] Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

[14] Shinji Ito, Akihiro Yabe, and Ryohei Fujimaki. Unbiased objective estimation in predictive optimization. In *International Conference on Machine Learning*, pages 2176–2185. PMLR, 2018.

[15] Adel Javanmard and Hamid Nazerzadeh. Dynamic pricing in high-dimensions. *Journal of Machine Learning Research*, 20(9):1–49, 2019.

[16] Utsav Sadana, Abhilash Chenreddy, Erick Delage, Alexandre Forel, Emma Frejinger, and Thibaut Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 2024.

[17] Andrew F Siegel and Michael R Wagner. Profit estimation error in the newsvendor model under a parametric demand distribution. *Management Science*, 67(8):4863–4879, 2021.

[18] James E Smith and Robert L Winkler. The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006.

[19] Charles M Stein. Estimation of the mean of a multivariate normal distribution. *The annals of Statistics*, pages 1135–1151, 1981.

[20] Jingwen Tang, Zhengling Qi, Ethan Fang, and Cong Shi. Offline feature-based pricing under censored demand: A causal inference approach. *Manufacturing & Service Operations Management*, 27(2):535–553, 2025.

[21] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.