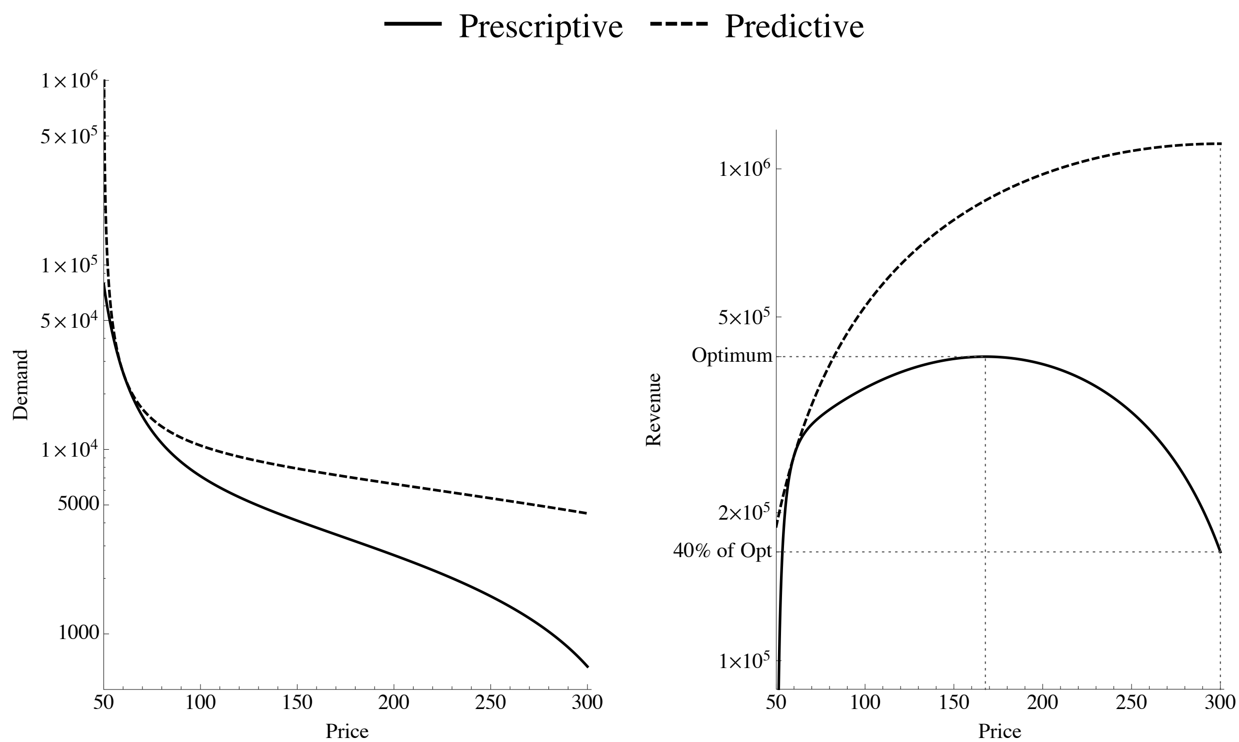
Prediction vs Prescription in Data-Driven Pricing

Dimitris Bertsimas, Nathan Kallus

Prediction vs Prescription in Data-Driven Pricing

Price optimization is a fundamental problem in revenue management that addresses the tradeoff between consumer demand for a product when sold at a particular price and the per-unit revenues netted at this sale price (see Phillips (2005)). In data-driven pricing, an optimal pricing strategy is to be prescribed based on historical data (see Besbes et al. (2010)). The standard approach to data-driven pricing in practice and in the literature involves fitting a **predictive** model for demand given price, such as a linear, logistic, or nonparametric regression, and maximizing the predicted demand times revenue given price. However, a fundamental misunderstanding about the subtle but consequential difference between **prediction** and **prescription** in data-driven pricing, present in such an approach, leaves managers with suboptimal pricing strategies and significant potential revenues on the table.

In this paper, we explore this problem by focusing on the fundamental building block of data-driven pricing: the choice of a price for a single product offering based on historical observations of the outcomes of past sale offerings, also called *customized price optimization* in Phillips (2005). We distinguish between the **prediction** problem, in which demand and revenue are to be predicted for a sale offering where the price set is given and demand is hidden (see Spirtes (2010)), and the **prescription** problem, in which the sale price to set is to be prescribed so that revenues are maximized. We show that under the standard framework considered in the literature, the optimal price is in general a *statistically non-identifiable quantity*; specifically, the price response function (PRF), which specifies the expected demand for the product when sold at a given price, is not identifiable. Thus, under the standard framework, data-driven pricing has no hope. We redeem the problem by establishing conditions under which both the optimal price and PRF are in fact



(a) Demand

(b) Revenue

Figure 1: The true demand and revenue curves (solid lines) and the spurious ones that would arise from an invalid predictive analysis (dashed lines), which would lead to a 60% loss in revenues compared to the true, prescriptive optimum.

identifiable. We provide concrete examples connected with practice and collected from a literature survey that show that incorrect analyses based on prediction leads to significant loss of revenue (see e.g. Figure 1).

To correctly address the **prescription** problem in data-driven pricing, we develop a particular, non-parametric solution and prove that it asymptotically converges to the true, optimal price as more data is gathered under mild conditions. Motivated by work that suggests that parametric approaches are often sufficient for price optimization (Besbes and Zeevi (2015)), we also develop a novel parametric approach to the newly formulated **prescription** problem in data-driven pricing. To assess the success of solutions to the prescription problem – whether generated by our new parametric approach or by a parametric or non-parametric predictive analysis – we develop a new statistical hypothesis test for the hypothesis that a particular price prescription provides optimal revenue in the **prescription** problem.

Finally, we provide an empirical study of the problem of prescribing interest rates for automobile loans based on historical data. The problem and dataset have been previously studied in Besbes et al. (2010) but our theoretical results suggest that the analysis therein is invalid. Indeed, the corresponding pricing strategies are rejected by our new test in almost every case as suboptimal, supporting this conclusion empirically. On the other hand, our new parametric approach to pricing passes the test of prescriptive revenue optimality in almost every case.

In conclusion, our results and literature survey suggest that a misunderstanding of the difference between **prescription** and **prediction** in data-driven pricing is commonplace and that it has detrimental consequences on real revenue-optimizing decisions. The new tools we develop allow practitioners to correctly address real-world data-driven pricing problems and to assess revenue optimality.

References

Besbes O, Phillips R, Zeevi A (2010) Testing the validity of a demand model: An operations perspective. *Manufacturing & Service Operations Management* 12(1):162–183.

Besbes O, Zeevi A (2015) On the (Surprising) Sufficiency of Linear Models for Dynamic Pricing with Demand Learning. *Management Science* forthcoming.

Phillips R (2005) *Pricing and Revenue Optimization* (Stanford University Press, Stanford, CA).

Spirtes P (2010) Introduction to causal inference. *The Journal of Machine Learning Research* 11: 1643–1662.