The New Frontier of Price Optimization

Identifying the optimal prices for products was once a time-consuming process. That's changing as businesses start to take advantage of advances in machine learning, increases in computing speed, and greater availability of data.

David Simchi-Levi

[PRICING]

The New Frontier of Price Optimization

Identifying the optimal prices for products was once a time-consuming process. That's changing as businesses start to take advantage of advances in machine learning, increases in computing speed, and greater availability of data.

BY DAVID SIMCHI-LEVI

Until fairly recently, price optimization has been restricted primarily to certain industries that have limited inventory, such as airlines and hotels. It's complex work, demanding the analysis of vast quantities of data and a deep understanding of competitors' behavior. Few organizations could optimize the price of more than a handful of products at one time.

But this is changing. Thanks to the growing availability of internal and external data, advances in machine learning, and increases in computing speed, price optimization can be applied more broadly. We have developed a way to set optimal prices for hundreds of stock units in near real time and on an ongoing basis.

In trials of our pricing technology with three online retailers, we found that we were able to increase each retailer's revenue, market share, and profit for selected products by double digits. What's more, although the examples described in this article involve online retailers, the price optimization method we developed is also appropriate for brick-and-mortar retailers; we recently implemented a similar method at a brewing company, where we optimized the company's promotion and pricing in various retail channels with similar results.

A Three-Step Process

Our system for generating better price predictions includes three steps:

1. Forecast. We match a cluster of products with similar sales characteristics to those of the product being optimized. Then we use a machine learning technique called a *regression tree*,[1] which consists of a set of if-then statements that yield a prediction. Using

a company's historical sales data, our algorithm generates as many as 20 if-then statements that can be used to predict the relationship between demand and price. That information, in turn, can be used to generate a price.

2. Learn. Next, we test our price against actual sales, redrawing our pricing curve to match actual results. At the end of the learning period, we know how well the product sold and can use that information to refine our demand-price curve for it.

3. Optimize. Once the learning period is over, we apply the new curve and optimize pricing across hundreds of products and time periods.

Price Optimization in Practice

In reality, not every step is used in every situation. For example, Boston, Massachusetts-based Rue La La Inc. did not want us to change prices in the course of its 48-hour sales, so we skipped the learning step. And when we worked with Chicago, Illinois-based Groupon Inc., we realized that the nature of its business made a demand forecast difficult to generate, so we focused instead on learning from current sales. Here's how my colleagues and I applied our price optimization techniques at three online retailers: Rue La La, Groupon, and B2W Digital.

Optimizing Pricing for Limited-Time Offers at Rue La La This online fashion retailer offers limited-time discounts ("flash sales") on designer apparel and accessories. Flash sale businesses like Rue La La aim to create a feeling of urgency and scarcity of products by offering great deals but for only a limited time (often just a few days) and with limited inventory. On Rue La La's website, the customer sees a number of "events," each representing a collection of similar products. Each event shows a countdown timer informing the customer of the time remaining until it will no longer be available.

PLEASE NOTE THAT GRAY AREAS REFLECT ARTWORK THAT HAS BEEN INTENTIONALLY REMOVED.
THE SUBSTANTIVE CONTENT OF THE ARTICLE APPEARS AS ORIGINALLY PUBLISHED.

FALL 2017 **MIT SLOAN MANAGEMENT REVIEW 23**

The New Frontier of Price Optimization (Continued from page 23)

One of Rue La La's main challenges was pricing items that it had never sold before. The company refers to them as "first-exposure" items, and they account for the majority of its sales. For example, in one department, about half of the first-exposure items sold out before the end of the event — suggesting that Rue La La could have raised prices on those items while still achieving high sell-through. On the other hand, many first-exposure items sell less than half of their inventory by the end of the sale period, indicating that their prices may have been too high.

In practice, Rue La La marketers set prices following the traditional cost-plus pricing method — simply adding a markup percentage to the product cost. However, the number of stock outs for some products and amount of leftover inventory for others suggested that the

combinations that did not sell out before the sale ended, while Group 2 included those item-event combinations that did result in stock outs.

For each event-item combination in either group, we calculated the percentage of sales that occurred in every hour of the event. This gave us a demand curve for the proportion of sales that occur in the first, say, x hours of an event. We then aggregated all these demand curves in the group that did not sell out before the sale ended into a few distinct and interpretable curves. We did this by applying a clustering technique that looked for all demand curves with a similar structure and then averaged them into a smaller number of curves.

The clustering technique suggested that the various event-item combinations of those items that did not experience a

demand curves from Group 1. For instance, suppose an item that sold out is associated with a sales event that started at 3 p.m. and stocked out after 10 hours. The curve associated with sales events that start at 3 p.m. indicates that 60% of total sales occur by the 10th hour. Thus, the total demand for this item was estimated as the initial inventory divided by 0.6.

Once we applied the clustering technique and generated estimated demand for items that stock out, we were ready to predict future demand. The regression tree, a machine learning technique developed in the 1960s, has proven to be the best tool for predicting demand. There are two reasons for this. First, it can successfully partition all items sold in the past and use only the relevant ones to predict demand for the current item being analyzed. Second, it works in special pricing situations, such as the inverted price relationship common in luxury goods, where demand actually goes up when the price rises. Traditional linear regression techniques can't handle this special situation, but it's critical for pricing fashion and high-end products, where price is often interpreted as a signal of quality.

We then developed a novel reformulation of the price-optimization equation, creating an algorithm that allowed Rue La La to optimize prices on a daily basis in time for the next day's sales. To execute our price-optimization algorithm, we developed a fully automated pricing-decision support tool that ran every day, providing price recommendations to merchants for events starting the next day.

We quantified the financial and market impacts of our tool for styles in various price ranges using a field experiment with Rue La La that lasted six months and that included 6,000 products. In the end, the decision support software led to a 10% increase in revenue for the company. This increase in revenue translated into a direct impact on profit and margin.

The two biggest challenges we faced when building our demand prediction model were estimating lost sales due to stock outs and predicting demand for styles that had no historical sales data.

company was leaving money on the table. To increase revenue and market share, Rue La La needed a pricing algorithm that could set higher prices for some first-time items and lower prices for others.

Our approach was twofold and began with developing a demand prediction model for first-exposure styles. We then used this demand prediction data as input into a price optimization model to maximize revenue.[2] The two biggest challenges we faced when building our demand prediction model were estimating lost sales due to stock outs and predicting demand for styles that had no historical sales data.

To address the first challenge — estimating sales lost due to stock outs — we split historical sales data into two groups: Group 1 included all item-event

stock out (Group 1) could be described by a total of four curves, each of which was associated with the time of day and week when the sale started. (For example, sales progress at different hourly rates during sales events that start in the afternoon than they do for events that start on weekday evenings.)

We used the sales data from the items that did not sell out to estimate lost sales for those products that did sell out. We looked at all of the event-item combinations that did sell out in the course of the sale and identified which of the four curves was appropriate, based on the time the sale started. Since we knew the time each item stocked out, we could immediately estimate total demand for the product based on the corresponding

Learning About Demand for Groupon Offers Another company we worked with was Groupon, a large e-commerce marketplace for daily deals that offers subscribers discounts from local merchants. For example, a discount dinner might be purchased from Groupon at $17 and redeemed at a local restaurant for a meal that would cost $30 if ordered without the coupon.

Groupon launches thousands of new deals every day and offers them only for a short period. The combination of the huge product portfolio and short sales period makes demand prediction quite challenging. Groupon needed an effective demand prediction model to optimize its prices but found it impossible to develop such a forecast.

To address this challenge, we generated multiple forecasts at the time a product launched on the company website.[3] The idea behind our approach was to generate multiple demand functions, such that the true demand-price relationship is approximated by one of these demand functions.

Of course, when the selling process for a specific deal starts, we have no idea which demand function among the multiple forecasts is the best to capture consumer behavior. To solve this problem, we split the time the deal is sold on the website into two parts: learning and price optimization. During the learning stage, we applied a test price and observed customers' decisions. At the end of the learning period, we knew how much was sold and therefore could identify the demand function closest to the level of sales at the learning price. This was the final demand-price function we used, and we used it as the basis for optimizing price during the optimization period.

The trade-offs in our algorithm are clear. If we run the test price for a long time, then we will develop a good understanding of the true demand function but with little time remaining for optimization. By contrast, if we study the sales response for just a short period of time, we won't have deep insight into customer demand but we will have a longer price optimization period.

When implementing our approach for the first time, we initially generated 10 demand functions for every deal. Very quickly, we realized that this was not nearly enough. In the final implementation, we generated around 100 demand functions for every new deal. These demand functions took into account the city or region where the deal was offered, the price range, and the size of the discount.

In the final implementation, however, Groupon imposed a few constraints on our approach. First, the learning price had been negotiated between Groupon and the local merchant and could not be set by our algorithm. Second, during the optimization phase, Groupon allowed us to lower the price only by a percentage ranging from

5% to 30%. Finally, the local merchant was to receive a fixed price. For example, in the $17 restaurant deal mentioned earlier, the local merchant received $10. If at the end of the learning period, our algorithm recommended decreasing the price to $15, the local merchant would still receive $10 and Groupon would absorb the difference by taking $5 instead of $7. This meant the local merchant always benefited from the price decrease, since they saw the same payment for every deal sold, while a lower price would presumably win them far more volume.

In a field experiment that included 1,295 deals for which the algorithm recommended a price decrease, we tracked two pieces of data: the total amount of money paid by customers to Groupon (called booking) and the portion of money that Groupon kept after paying local merchants (referred to as revenue). The results varied by the deal category (for example, whether the offer was for food and drink or for beauty-related offers), but overall, bookings increased by 116% and revenue by 21.7%.

Further analysis of the results from the field experiment showed that reducing price had a much bigger impact on low-volume deals. For deals with fewer bookings per day than the median, the average increase in revenue was 116%, while revenue increased only 14% for deals with more bookings per day than the median.

Forecasting, Learning, and Price Optimization at B2W Digital A third company we worked with was B2W Digital, a large Latin American online retailer that is based in Rio de Janeiro, Brazil, and

> If we run the test price for a long time, then we will develop a good understanding of the true demand function but with little time remaining for optimization.

competes with companies such as Amazon and Walmart. The rich portfolio of products offered by B2W, the accessible historical sales data for these products, and the ability to change prices multiple times a day provided us with a unique opportunity to combine forecasting, learning, and optimization. In the first step, we generated a forecast for each product in our study by combining internal and external data. Internal data included traffic to the site, pricing, discounts, advertisement spending, and competing products offered on the B2W website and their prices. External data included competitors' pricing on their websites, their advertisements, and weather conditions. As was the case for Rue La La, using a version of regression

trees[4] turned out to be the best way to predict demand-price relationships.

The second step involved learning. Every few hours, we observed the traffic to the website, demand for B2W products, and competitor behavior. We noted these changes and updated our input to the regression tree to better represent current

decreased relative to a control group by 264%, profit decreased by 416%, and the number of units sold fell by 216%. However, when the optimization engine was in place, revenue in the treatment group increased by 471% relative to the control group, profit increased by 366%, and the number of units sold increased by 391%.

understand and can implement this type of technology.

Most of all, executives need to overcome internal resistance by recognizing that price optimization technology is not going to replace merchandisers or sales executives. Although price optimization technology is an extremely useful and flexible decision-support tool, like most tools, it is most useful as a device to augment professional capacities, not replace them.

Indeed, in most organizations, merchants and pricing experts can focus on the top 10% or 20% of the company's products but cannot cover all products in every category. The ability to automate pricing extends this capacity by allowing companies to optimize pricing for more products than most organizations currently find possible.

> ## Executives need to overcome internal resistance by recognizing that price optimization technology is not going to replace merchandisers or sales executives.

market conditions. Finally, we applied price optimization across all products offered by B2W that compete against each other in the same market.

In August 2015, B2W started using our system. The new forecast, generated through this learning process, was then used by the optimization model. This required no manual intervention; all prices were pushed directly to the website.

We compared the performance of two groups of equivalent products: a control group where B2W merchants applied their traditional pricing strategies and a treatment group where the new technology pushed prices directly to the B2W website. For low-priced products, revenue increased by 66%, profit by 44%, and number of units sold — which is a proxy for market share — increased by 141%. The results for fast-selling products were not as impressive but still quite strong: Revenue grew by 17%, profit by 30%, and the number of units sold by 30%.

One test we conducted illustrated the importance of the price optimization step. We experimented with using only the forecasting and learning steps, without price optimization, for premium products. In this test, when we didn't apply price optimization, revenue for the treatment group of premium products

Our entire field study ran for several months. When it concluded, B2W reported that the new technology had not only improved revenue, profit, and market share but also expanded the range of products sold: Optimization had enabled B2W to sell not only more products but more unique products than in its control group.

Prerequisites for Price Optimization

Innovation isn't only a matter of inventing a new tool; it's also about people using the tool. From that point of view, every new technology is an exercise in change management. Price optimization technology is no exception. These types of systems, and indeed, any big data analytics project, require more than just new technology to be successful. They also require executives to create the conditions that are needed to take advantage of the opportunity. To succeed with price optimization programs like this, executives first need to:
• Be open to using data from external sources, such as competitors' prices, to complement internal data.
• Break up silos between different functional areas within the organization so that data can be shared across functions.
• Hire data analytics specialists who

David Simchi-Levi is a professor of engineering systems at the Massachusetts Institute of Technology in Cambridge, Massachusetts, as well as a faculty member of the MIT Operations Research Center and the MIT Institute for Data, Systems, and Society. Comment on this article at http://sloanreview.mit.edu/x/59102, or contact the author at smrfeedback@mit.edu.

REFERENCES

1. A regression tree is a kind of decision tree where each node represents an "if-then" rule; when these rules are followed, the tree yields a numerical prediction. For our purposes, these if-then rules are created by an algorithm sifting through a company's historical sales.

2. For more details, see K.J. Ferreira, B.H.A. Lee, and D. Simchi-Levi, "Analytics for an Online Retailer: Demand Forecasting and Price Optimization," Manufacturing & Service Operations Management 18, no. 1 (winter 2016): 69-88.

3. For more details, see W.C. Cheung, D. Simchi-Levi, and H. Wang, "Technical Note — Dynamic Pricing and Demand Learning With Limited Price Experimentation," Operations Research, Articles in Advance, https://doi.org/10.1287/opre.2017.1629.

4. In this case, we used a more sophisticated version of regression trees called "random forest."

MIT**Sloan**
Management Review

PDFs ■ Reprints ■ Permission to Copy ■ Back Issues

Articles published in MIT Sloan Management Review are copyrighted by the Massachusetts Institute of Technology unless otherwise specified at the end of an article.

MIT Sloan Management Review articles, permissions, and back issues can be purchased on our Web site: sloanreview.mit.edu or you may order through our Business Service Center (9 a.m.-5 p.m. ET) at the phone numbers listed below. Paper reprints are available in quantities of 250 or more.

To reproduce or transmit one or more MIT Sloan Management Review articles by electronic or mechanical means (including photocopying or archiving in any information storage or retrieval system) **requires written permission.**

To request permission, use our Web site: sloanreview.mit.edu
or
E-mail: smr-help@mit.edu
Call (US and International):617-253-7170 Fax: 617-258-9739

Posting of full-text SMR articles on publicly accessible Internet sites is prohibited. To obtain permission to post articles on secure and/or password-protected intranet sites, e-mail your request to smr-help@mit.edu.