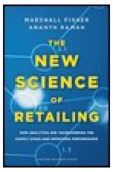
Chapters to Go



The New Science of Retailing: How Analytics Are Transforming the Supply Chain and Improving Performance

by Marshall Fisher and Ananth Raman
Harvard Business Press. (c) 2010. Copying Prohibited.

Chapter Three: Product Life Cycle Planning—How to Reinvent Forecasting, Inventory Optimization, and Markdown Pricing

Overview

If you're a retailer, you're being hurt by your inventory. All retailers, at one time or another, end up carrying too much or too little. Too much, of course, leads to end-of-season surpluses that must be cleared with margin-killing markdowns. Too little causes stockouts and missed sales. That frustrates customers and can prompt them to shop elsewhere, sometimes forever.

How do you walk the tightrope between margin lost and sales forgone? The key lies in managing three ingredients that, together, enable product availability: an accurate forecast, a flexible supply chain, and, of course, an appropriate supply of inventory.

Any of these three factors, handled right, will suffice, ensuring that your customers get what they want when they want it. But too many retailers err on the side of carrying excessive inventory because of their inaccurate forecasts or slow, rigid supply chains. If they improved their forecasts and the flexibility of their supplies, they could please their customers without having to sink so much of their cash into goods that might not sell.

In this chapter, we'll provide you with tools to help you improve forecast accuracy and decide how much inventory of each product to carry in each of your stores at any point. We'll also advocate an unorthodox, but much more profitable, way of managing markdowns.

Retail products exhibit all the variation of human life. Many endure for decades, but some die young, and a few never make it out of prelaunch planning. Their lives can stretch from a few weeks, for a hit CD, to years, for a brand of flour peddled in a grocery store or a soap sold in the pharmacy. Unilever's Lifebuoy has existed for more than a hundred years. Many goods, ranging from cars to refrigerators, last only a year in stores because the industry operates on an annual new-product-planning cycle.

Figure 3-1 shows the events in the life of a product. These may be grouped into three phases: beginning-of-life launch, midlife replenishment, and end-of-life exit.

Long before a product reaches stores, a retailer will predict how well it will sell and use that estimate to determine an initial buy quantity. Retailers use a variety of approaches for forecasting sales. One is to identify a similar product from a prior season and use its sales as an estimate of the new item's prospects. This works well if the new product is just a variation of the old one. Other approaches include asking buyers to estimate the product's prospects or conducting a test by placing a small quantity of the product in a few stores and extrapolating from the sales.

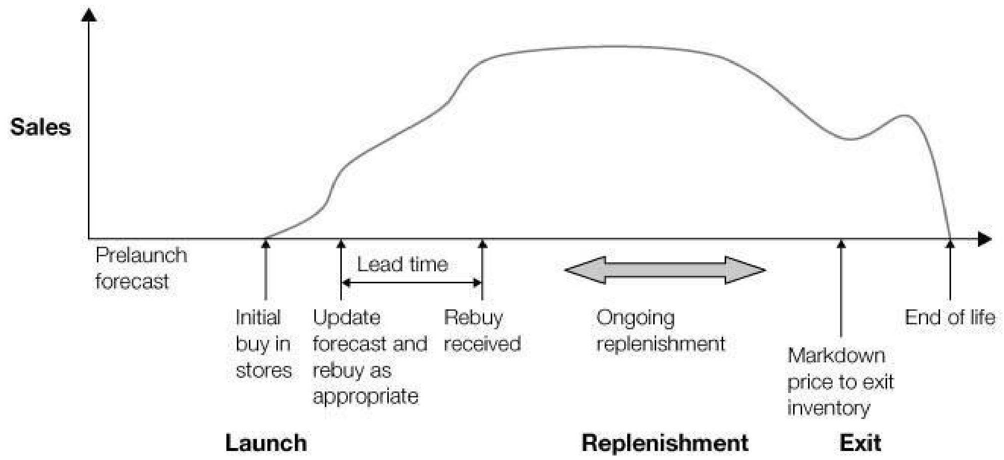


FIGURE 3-1: Product life cycle planning

Once sales of a new product stabilize, replenishment becomes the rule. Each week, the retailer ships units to stores based on what sold in the preceding week, bringing inventory up to a desired level. Eventually, either the end of the season arrives or sales slow, and the retailer decides to discontinue the product and sell off the remaining inventory through judicious markdowns.

You can split life cycle planning into three categories based on the ratio of length of life to replenishment lead time.

- If the lead time is greater than or equal to the life, then replenishment is impossible, and you're forced to make a single buy. That's known as the "one and done" mode of operation.
- At the other extreme, if product life is much longer than lead time—say, a life of a year and a lead time of four weeks—then the product goes on replenishment soon after launch.
- In the intermediate case, lead time is a little less than life cycle. Perhaps, the life is twenty-six weeks, and the lead time is twelve weeks. So you can order only a few replenishments. A variation is when you can make a single buy from a supplier but can allocate part of it to your distribution center and replenish stores from that supply.

Launch

We'll use a women's apparel catalog company to walk you through the key decisions faced in launching new products. We'll also describe a methodology for making those decisions that we developed while working with the cataloger. [1]

Our challenge was to decide how much to buy of each of an assortment of women's dresses and other clothing to be offered in a catalog that would last for twenty-six weeks. Because of long production lead times, the company needed to make orders several months before the catalog's mailing. If an item sold out during the twenty-six-week life of the catalog, it often could order more, giving customers the option of back-ordering. For simplicity in this explanation, we'll assume that customers couldn't back-order, so an order for a sold-out product became a lost sale. The firm sold any inventory remaining at the end of the catalog's life at a deep discount.

"We" Beats "Me": A New Way of Deliberating

A committee of four buyers decided how much of each item to purchase. [2] In making their decisions, the buyers studied drawings of each item, discussed target consumers, and considered the popularity of prior similar products.

In observing the committee's deliberations, we noticed that one of the buyers was more articulate and assertive than the others. Often, she swayed her colleagues, so the final decisions represented her preferences rather than the collective wisdom. The other buyers, though more reticent, didn't have less fashion sense, and we were concerned that the firm was losing 75 percent of the committee's wisdom. We therefore asked the buyers to vary their usual process. Each one would write down her forecast of each item's sales over the twenty-six-week life of the catalog, and we'd then compare them. Table 3-1 shows these individual forecasts for three of the items, together with the average and standard deviation of the forecasts.

Table 3-1: Forecast of four buyers for three products

	Anna	Laurie	Julie	Kim	Committee average	Committee standard deviation
Navy turtleneck	89	86	102	102	95	7
Red cardigan	51	100	152	39	86	45
Blue vest	30	91	183	76	95	56
The data in this table has been disguised.						

Note the disparity between Julie's and Kim's forecasts for the red cardigan and between Anna's and Julie's for the blue vest. The usual groupthink process would have hidden these differences of opinion.

Given these dispersed forecasts, how much of each item should the cataloger have ordered? For now, assume that production lead times are so long that the company can't replenish during the season and must buy each item at the start of the season. One possible answer is to buy exactly what was the average forecast for each item—that is, 95 navy turtlenecks, 95 blue vests, and 86 red cardigans. That's what this company typically did.

Initial Forecasts: Groping in the Dark

Twenty-six weeks after the company mailed the catalog, we knew the total demand for each product and could evaluate how well this approach had worked. Table 3-2 shows actual demand and forecast error for the three items. Notice that forecast errors are generally high, but the forecast for the navy turtleneck, where the four buyers tended to agree, was reasonably accurate.

A common measure of forecast error, called *mean absolute deviation* , or MAD, equals the total absolute error as a percentage of total actual demand. For this example, MAD equals (10 + 47 + 66) / (85 + 132 + 29), or 50 percent. Notice that we measure error as the *absolute deviation* of the forecast from actual demand, so the errors of −10 and −66 drop their negative signs in the calculation.

Table 3-2: Forecast errors

	Committee average	Committee standard deviation	Actual demand	Error	Absolute deviation
Navy turtleneck	95	7	85	−10	10
Red cardigan	86	45	132	47	47
Blue vest	95	56	29	−66	66

Figure 3-2 shows actual versus forecast demand for all items, and figure 3-3 shows forecast error versus standard deviation for all items. Note that forecast errors are generally quite high, with a *mean absolute percentage error* (MAPE) of 55 percent, but tend to be lower when the four buyers' forecasts are closer to each other and the standard deviation is smaller. In fact, the items in the right half of figure 3-3, with standard deviation of more than 200, had nearly six times the forecast error as the items with standard deviation of less than 200. In our experience, high forecast error for new items is typical; we generally see the presales forecast errors on new items range from 50 percent to 100 percent.

Forecast Errors Cost Money

What's the cost of these errors? Because demand is hard to predict, we make mistakes, buying too little of some products and too much of others. Table 3-3 shows the per-unit cost of under- and overbuying the three items. If you buy too little, you lose the gross margin on sales that you could have made. Every missed sale of the red cardigan costs the per-unit gross margin of $83 (that is, $160 − $77). This $83 lost margin is called an *opportunity cost* ; it doesn't appear on the income statement. Rather, it's the cost of *not* doing something. The per-unit underbuy cost for the three items is computed as the normal price minus cost.

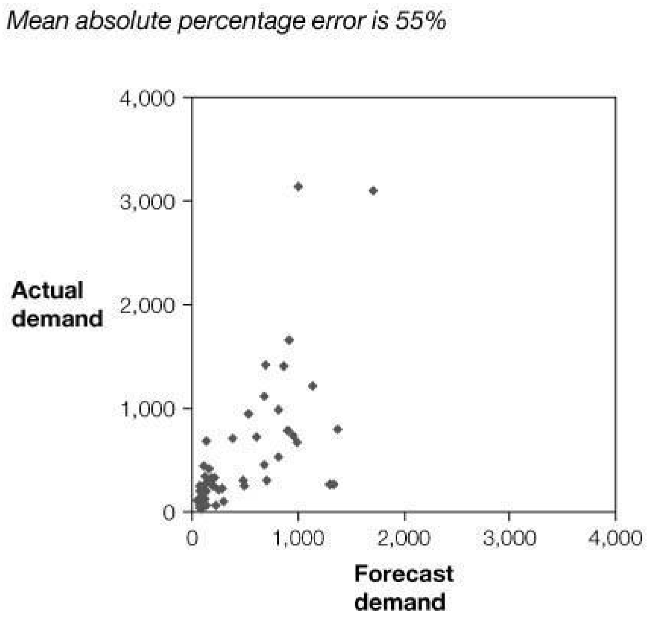


FIGURE 3-2: Actual versus forecast demand for all items

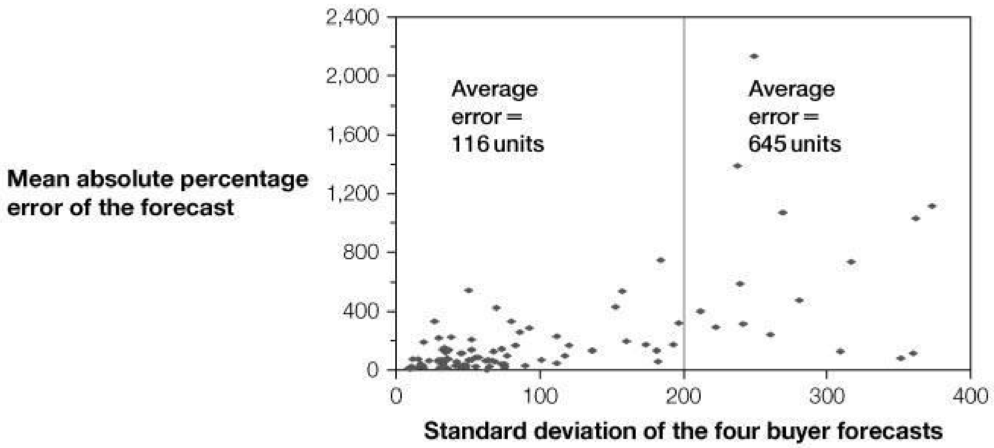


FIGURE 3-3: Forecast error versus committee standard deviation for all items

Table 3-3: The costs of under- and overbuying differ

	Cost	Normal price	Markdown price	Per-unit under buy cost	Per-unit over-buy cost
Navy turtleneck	$40	$60	$18	$20	$22
Red cardigan	77	160	48	83	29
Blue vest	65	110	33	45	32
The data in this table has been disguised.					

Buying too much incurs a markdown. Every blue vest left over after twenty-six weeks costs the cataloger $32 (that is, $65 – $33). The per-unit overbuy cost for the three items is computed as cost minus the markdown price.

Table 3-4 displays a financial evaluation of four increasingly sophisticated strategies for buying the three items.

Table 3-4: Profitability of four increasingly sophisticated buying strategies

Buy the forecast

	Buy qty	Actual demand	Sales	Gross margin on sales	Mark-down units	Loss on mark-downs	Net profit	Lost sales	Lost margin
Navy turtleneck	95	85	85	$1,700	10	$220	$1,480	0	$0
Red cardigan	86	132	86	7,138	0	0	7,138	46	3,818
Blue vest	95	29	29	1,305	66	2,112	–807	0	0
Totals			200	10,143	76	2,332	7,811	46	3,818

Hedge: Buy the forecast + 10%

	Buy qty	Actual demand	Sales	Gross margin on sales	Mark-down units	Loss on mark-downs	Net profit	Lost sales	Lost margin
Navy turtleneck	105	85	85	$1,700	20	$440	$1,260	0	$0
Red cardigan	95	132	95	7,885	0	0	7,885	37	3,071
Blue vest	105	29	29	1,305	76	2,432	–1,127	0	0
Totals			209	10,890	96	2,872	8,018	37	3,071

Use a probabilistic risk adjusted hedge

	Buy qty	Actual demand	Sales	Gross margin on sales	Mark-down units	Loss on mark-downs	Net profit	Lost sales	Lost margin
Navy turtleneck	89	85	85	$1,700	4	$88	$1,612	0	$0
Red cardigan	100	132	100	8,300	0	0	8,300	32	2,656
Blue vest	91	29	29	1,305	62	1,984	–679	0	0
Totals			214	11,305	66	2,072	9,233	32	2,656

Risk adjust using a continuous Gamma distribution

	Buy qty	Actual demand	Sales	Gross margin on sales	Mark-down units	Loss on mark-downs	Net profit	Lost sales	Lost margin
Navy turtleneck	94	85	85	$1,700	9	$198	$1,502	0	$0
Red cardigan	109	132	109	9,047	0	0	9,047	23	1,909
Blue vest	96	29	29	1,305	67	2,144	–839	0	0
Totals			223	12,052	76	2,342	9,710	23	1,909

We will be describing these four strategies in the next few pages. The top tableau computes net profit, assuming that the retailer bought the committee members' average forecast. In this table, the sales value is the minimum of the buy quantity ("Buy qty" column) and actual demand; gross margin on sales is the per-unit underbuy cost from table 3-3 times sales; the markdown units value is the buy quantity minus sales; loss on markdowns is the per-unit overbuy cost from table 3-3 times markdown units; net profit is gross margin minus loss on markdowns; lost sales is actual demand minus sales; and lost margin is the per-unit underbuy cost times lost sales. The buy of blue vests vastly exceeded the demand of 29 units, so the retailer had 66 units left over, which cost it a total of $2,112 to liquidate. This loss exceeded the margin earned on the 29 units sold, so, overall, the firm lost money on this item.

While the $2,112 markdown loss looms large, an even bigger number is the $3,818 of *lost margin* on the red cardigan because the 86 units didn't come close to meeting the actual demand of 132 units. This is consistent with our experience: lost margin dwarfs the cost of ditching overstocks. Yet when we talk to managers about inventory problems, they usually grouse about some item that they thought would sell briskly but didn't, leaving a year's supply moldering in the warehouse. Mistakes like that are obvious and embarrassing. The inventory sits there inviting recrimination. Eventually, the value of the goods has to be written down, which shows up on the income statement. That creates still more shame. If the mistake was big enough and the company is public, it can even lead to a decrease in stock price, when analysts and investors catch on. In contrast, an unsatisfied customer who didn't get what she wanted is invisible; the margin lost because of the missed sale doesn't appear on any financial record.

Cutting the Cost of Errors

You might guess that this retailer could have increased net profit by buying more than the forecast. The second tableau in table 3-4 shows that this supposition is correct. The tableau presents the same financial calculation described above, but this time the retailer simply buys 10 percent more than the committee members' average forecast.

Why did buying 10 percent more increase profit? The extra buy actually increases markdown units for the navy turtleneck and blue vest by a total of 20 units, and these extra markdowns cost an additional $540 in markdown losses. However, this loss is outweighed by the additional gross margin of $747 earned on the additional 9 sales of the red cardigan. Simply put, hedging high increases profit because the cost of underbuying is much higher than the cost of overbuying for the red cardigan.

Finding an Optimal Hedge

So hedging helps, but how much should you hedge? One way is to estimate the probability of different possible outcomes and then follow a strategy that works best on average. Assume, for example, that each of the four catalog buyers is equally likely to be right, and thus assign a probability of 0.25 to each of their forecasts. Following this approach for the navy turtleneck, you'd assume there was a 25 percent chance of orders of 86 units, 25 percent of 89, and 50 percent of 102, since both Julie and Kim predicted 102 for this item.

If you bought the committee average of 95 and demand were 86, you'd have overbought by 9 units, at a total cost of $198 (9 units × $22). If demand were 89, you'd have overbought by 6 units, at a cost of $132. And if demand were 102, you'd have underbought by 7 units, at a cost of $140 (7 units × $20). Your probability-weighted cost in this case therefore is (0.25 × $198) + (0.25 × $132) + (0.5 × $140), or $152.50. This probability-weighted cost is called the *expected cost* . The next step is to find the expected cost for all other buys and choose the one that does best on average.

Table 3-5 shows the expected cost for many different quantities. The purchase with the lowest expected cost is 89 and is highlighted in the table. The same analysis applied to the other two items shows that buying 100 red cardigans and 91 blue vests minimizes the expected cost of lost margin and markdowns.

Table 3-5: Finding an optimal buy for the blue turtleneck, assuming probabilistic demand

IF DEMAND= 86			IF DEMAND= 89		IF DEMAND= 102		
Buy quantity	**Lost margin**	**Mark-** down cost	**Lost margin**	**Mark-** down cost	**Lost margin**	**Mark-** down cost	**Probability weighted total cost**
80	120	0	180	0	440	0	295.0
81	100	0	160	0	420	0	275.0
82	80	0	140	0	400	0	255.0
83	60	0	120	0	380	0	235.0

84	40	0	100	0	360	0	215.0
85	20	0	80	0	340	0	195.0
86	0	0	60	0	320	0	175.0
87	0	22	40	0	300	0	165.5
88	0	44	20	0	280	0	156.0
89	0	66	0	0	260	0	146.5
90	0	88	0	22	240	0	147.5
91	0	110	0	44	220	0	148.5
92	0	132	0	66	200	0	149.5
93	0	154	0	88	180	0	150.5
94	0	176	0	110	160	0	151.5
95	0	198	0	132	140	0	152.5
96	0	220	0	154	120	0	153.5
97	0	242	0	176	100	0	154.5
98	0	264	0	198	80	0	155.5
99	0	286	0	220	60	0	156.5
100	0	308	0	242	40	0	157.5

The third tableau of table 3-4 displays financial results for these purchases. Notice that taking into account the high cost of lost margin relative to the cost of markdowns on the red cardigan has led to buying more of that item, while the reverse logic has led to buying fewer navy turtlenecks. Overall, net profit has shot up.

A Fancy (and Useful) Formula—The Gamma Distribution

One valid objection to this approach is that we've only allowed for four possible values for demand (corresponding to the four buyer forecasts), while actual demand can vary continuously from zero upward. We can fix this problem using various formulas statisticians have created to represent the probabilities of real-life phenomena. One well-known example is the famous bell-shaped normal distribution. However, the bell curve doesn't work well for uncertain demand because it assigns a positive probability to negative values, and negative demand is impossible.

The *Gamma distribution* is less well known but frequently used for modeling uncertain demand. It resembles the normal distribution but allows only positive values. Figure 3-4 is the Gamma distribution with a mean of 95 and a standard deviation of 7, which would be appropriate for the navy turtleneck, since these are the values of the average and the standard deviation in table 3-2. Figure 3-5 gives the Gamma distribution with a mean of 95 and a standard deviation of 56, which would be appropriate for the blue vest, given the values in table 3-2. Note that the shape of a Gamma distribution can vary considerably, depending on the value of the mean and the standard deviation. The Gamma for the navy turtleneck looks very much like a standard normal, whereas the Gamma for the blue vest is significantly skewed to the left.

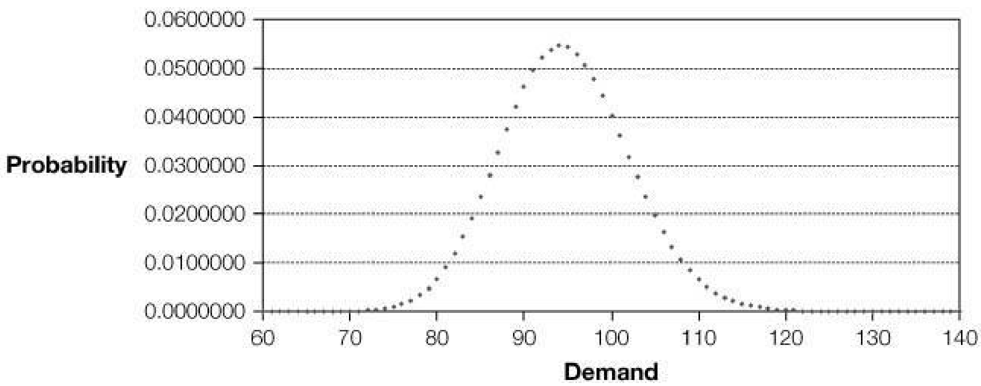


FIGURE 3-4: Gamma probability distribution for the navy turtleneck

You can find an optimal buy quantity with these probability distributions using exactly the same approach as before, except that now, instead of evaluating four demand scenarios, you evaluate hundreds. Conceptually, the approach is exactly the same. For each possible quantity, you compute the markdown and lost-margin cost for every possible demand value. You then multiply these costs by their probabilities and sum all of the possible demands to compute an expected cost. That way, you find the buy with the lowest expected cost.

Table 3-6 gives numeric probabilities for the Gamma distribution for the navy turtleneck. The table gives a demand value, the probability of that demand value, and the probability (called the *cumulative probability*) that demand will be less than or equal to that value.

Previously, you had three demand values—86, 89, and 102—with probabilities 0.25, 0.25, and 0.5, respectively. Now you have fifty-two demand values, ranging from 70 to 121, with the probabilities given in the table. If you want to evaluate buying 95 units, you compute a cost for each demand value. For example, if demand is 70, you'll have 25 units left over, at a cost of $550 (that is, $22 × 25). If demand is 121, you'll fail to satisfy 26 units of demand, at a cost of $520 ($20 × 26). A similar calculation can be done for each possible demand value, and then multiplied by the probability of that value and summed to find the expected cost of buying 95, which is $123.04. Section A-1 in the appendix describes the details of this approach. Following this approach produces expected cost–minimizing buys of 94 for the navy turtleneck, 109 for the red cardigan, and 96 for the blue vest. The fourth tableau of table 3-4 gives a financial evaluation of these buys for the three items.

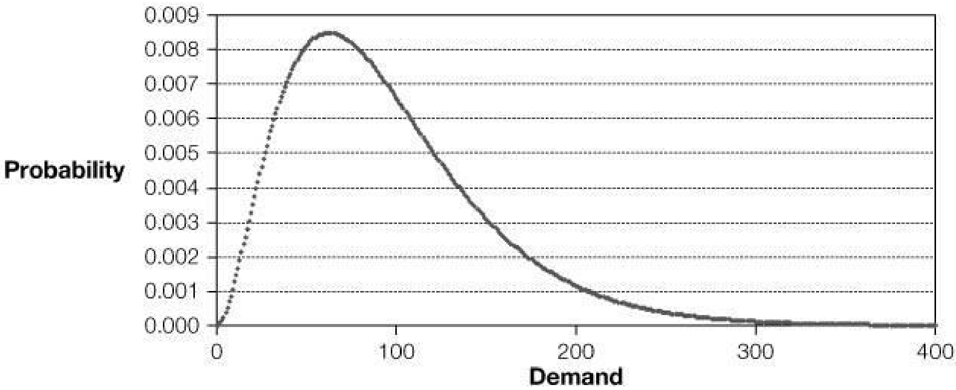


FIGURE 3-5: Gamma probability distribution for the blue vest

Table 3-6: Probability values for the navy turtleneck, using a Gamma distribution

Demand	Probability	Cumulative probability	Demand	Probability	Cumulative probability
70	0.000	0.000	96	0.053	0.604
71	0.000	0.000	97	0.051	0.654
72	0.000	0.001	98	0.048	0.702
73	0.000	0.001	99	0.044	0.746
74	0.001	0.002	100	0.040	0.787
75	0.001	0.003	101	0.036	0.823
76	0.002	0.004	102	0.032	0.854
77	0.002	0.006	103	0.028	0.882
78	0.003	0.010	104	0.023	0.905
79	0.005	0.015	105	0.020	0.925
80	0.007	0.021	106	0.016	0.941
81	0.009	0.030	107	0.013	0.955
82	0.012	0.042	108	0.011	0.965
83	0.015	0.058	109	0.008	0.974
84	0.019	0.077	110	0.007	0.981
85	0.023	0.100	111	0.005	0.986
86	0.028	0.128	112	0.004	0.990
87	0.033	0.161	113	0.003	0.993

88	0.038	0.199	114	0.002	0.995
89	0.042	0.241	115	0.002	0.996
90	0.046	0.287	116	0.001	0.997
91	0.049	0.336	117	0.001	0.998
92	0.052	0.388	118	0.001	0.999
93	0.054	0.442	119	0.000	0.999
94	0.055	0.497	120	0.000	0.999
95	0.054	0.551	121	0.000	1.000

Fortunately, there is a simpler way to find optimal buy quantities with this approach, using a concept called *marginal profitability analysis* . Suppose you have justified buying at least 93 units of the navy turtleneck, and you're wondering whether to buy a 94th. If it sells, you'll earn a $20 margin. If it doesn't, you'll lose $22 when you have to mark it down at the end of the season. The probability that it won't sell is the probability that demand is less than or equal to 93, which, as you see, equals 0.442, the cumulative probability shown next to a demand value of 93 in table 3-6. The probability that it will sell is thus 1 − 0.442 = 0.558. The expected net profit of the 94th unit is ($20 × 0.558) − ($22 × 0.442) = $1.44, and it is profitable to buy this unit. However, a 95th unit is not profitable because its expected net profit is negative. In this case, the equation is ($20 × 0.503) − ($22 × 0.497) = −0.87.

There is a formula that shortcuts this trial-and-error process for finding the point at which incremental profit shifts from positive to negative. It seems complicated, but it's really just a matter of plugging in values—as long as you've made good forecasts of demand.

Here's how it works. Let *p* denote the probability that the marginal unit sells, and then seek a buy quantity such that the net expected profit on the marginal unit is positive, but turns negative if we buy one more unit. The formula looks like this:

(normal price − cost)p − (cost − markdown price)(1 − p) = 0

Solving this equation for *p* gives you the following:

p = (cost − markdown price) / (normal price − markdown price)

Plugging in numbers for the navy turtleneck gives a target sales probability for the marginal unit of *p* = (40 − 18) / (60 − 18) = 0.524. We see in the row for demand of 93 in table 3-6 that the cumulative probability that demand is less than or equal to 93 is 0.442, and hence a 94th unit has a 0.558 (that is, 1 − 0.442) chance of selling, which is greater than our target of 0.524. The net expected profit on the 94th unit therefore would be positive. Similarly, the probability demand that is less than or equal to 94 is 0.497, so a 95th unit would have a 0.503 chance of selling and a negative net expected profit. Thus 94 is the optimal buy quantity.

Using Early Sales to Improve Accuracy

This kind of analysis enables you to make smart gambles based on inaccurate forecasts, but it also cries out for better forecasts if yours are shaky. In trying to help the catalog company, we searched for ways to improve its forecasts. To this end, we examined past catalogs and sales cycles and found that its apparel sold at predictable rates throughout the season. Sales chugged along in the first week, rose quickly in the second, and tailed off as the end of the season approached. Table 3-7 shows the average sales rates that we computed. We used table 3-7 to update forecasts during the season by extrapolating early sales based on what fraction of the total we estimate they constitute. For example, suppose an item had sold 11 units by the end of the second week. Then we'd forecast total sales of 100 for the twenty-six-week season, reasoning that these 11 sales represented 11 percent of total-season sales.

Table 3-7: Distribution of sales by week

Week	% of total sales by weekend	Week	% of total sales by weekend
1	2%	14	55%
2	11%	15	59%
3	15%	16	62%
4	18%	17	65%
5	21%	18	69%

6	25%	19	73%
7	29%	20	77%
8	33%	21	81%
9	37%	22	85%
10	40%	23	88%
11	43%	24	92%
12	47%	25	96%
13	51%	26	100%

These midseason corrections proved to be remarkably accurate. While we've used three items to illustrate our approach, the category we were planning actually had twenty-seven items, and figure 3-6 shows forecast results for all twenty-seven items. The left panel of figure 3-6 shows the original preseason forecasts, based on the assessment of the four buyers, while the right panel shows forecasts for the same items created by extrapolating the first two weeks of sales. Clearly, *obtaining even a small amount of initial sales data can have a dramatic impact on forecast accuracy* .

We have since applied this approach in a variety of companies and found the picture shown in figure 3-6 to be remarkably consistent. Before product launch, without current sales data, forecasts typically have an error rate of 50 percent to 100 percent, but just a few weeks after launch, updated forecasts based on initial sales have an error rate of only 10 percent to 20 percent.

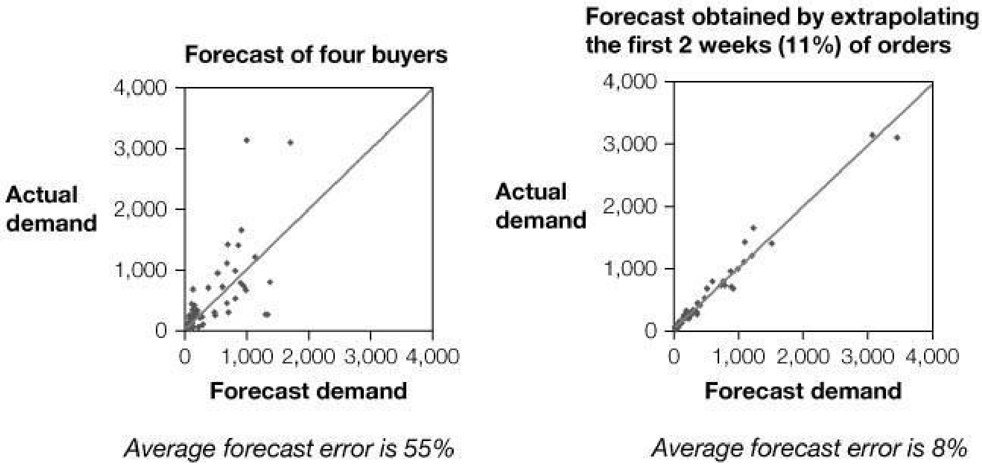


FIGURE 3-6: A little sales data dramatically improves forecast accuracy

Wouldn't it be ideal if you could buy using the forecasts in the right panel? You can, at least partially. For simplicity's sake, our examples have so far assumed that the lead time was so long that it allowed only a single buy at the start of the season. But a typical lead time for the replenishment of an existing item is twelve weeks. This means you could reorder at the end of the second week based on the accurate forecasts shown above, and the additional product would arrive at the end of the fourteenth week of the season, with twelve weeks left to sell it.

This suggests a strategy of "buy a little, sell a little, update your forecast, and buy more, if needed." How much should you buy initially? Enough that you probably won't sell out in the initial fourteen-week period, but not so much that you'll have excess inventory that you'll have to dump at a loss at the end of the season. To gauge the probability of selling out in the first fourteen weeks, you need a probability distribution of demand over this period. Fortunately, you can derive this from the distribution that you already have of total-season demand. Notice from table 3-7 that 55 percent of total predicted season sales should have occurred by the end of week 14. Thus the probability that demand in the first fourteen weeks will be less than or equal to X equals the probability that total-season demand is less than or equal to X / 0.55. Figure 3-7 shows the probability distributions for the navy turtleneck derived using this approach, and table 3-8 gives the exact probabilities for different demand values. [3]

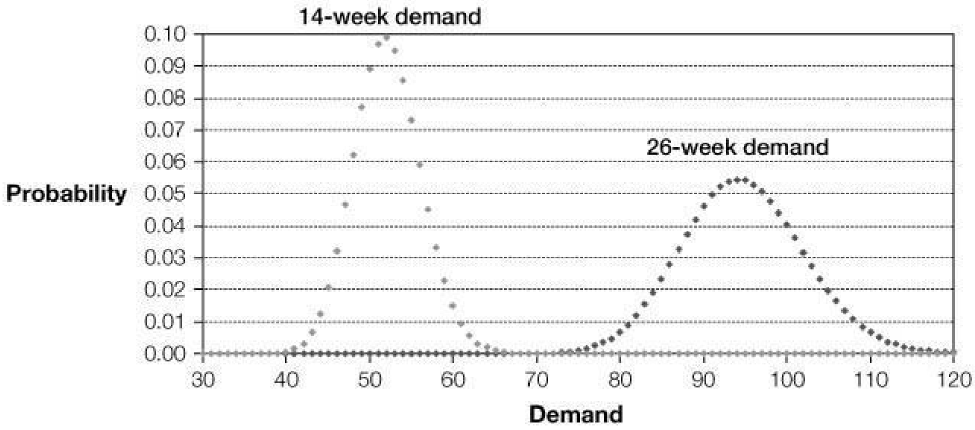


FIGURE 3-7: 14-week and 26-week probability distributions for the navy turtleneck: 14-week demand

Table 3-8: Probabilities for first 14 weeks and for 26-week season for the navy turtleneck

Demand	Season probability	14-week probability	Season cumulative probability	14-week cumulative probability
30	0.000000	0.000000	0.000000	0.000000
31	0.000000	0.000000	0.000000	0.000000
32	0.000000	0.000000	0.000000	0.000000
33	0.000000	0.000000	0.000000	0.000000
34	0.000000	0.000000	0.000000	0.000000
35	0.000000	0.000001	0.000000	0.000002
36	0.000000	0.000006	0.000000	0.000008
37	0.000000	0.000023	0.000000	0.000031
38	0.000000	0.000077	0.000000	0.000108
39	0.000000	0.000231	0.000000	0.000338
40	0.000000	0.000622	0.000000	0.000960
41	0.000000	0.001511	0.000000	0.002472
42	0.000000	0.003332	0.000000	0.005804
43	0.000000	0.006692	0.000000	0.012495
44	0.000000	0.012294	0.000000	0.024789
45	0.000000	0.020743	0.000000	0.045532
46	0.000000	0.032262	0.000000	0.077793
47	0.000000	0.046414	0.000000	0.124208
48	0.000000	0.061969	0.000000	0.186176
49	0.000000	0.077013	0.000000	0.263190
50	0.000000	0.089346	0.000000	0.352535
51	0.000000	0.097022	0.000000	0.449557
52	0.000000	0.098868	0.000000	0.548425
53	0.000000	0.094770	0.000000	0.643195
54	0.000000	0.085646	0.000000	0.728841
55	0.000000	0.073128	0.000000	0.801968
56	0.000000	0.059113	0.000000	0.861082
57	0.000000	0.045326	0.000000	0.906408
58	0.000000	0.033027	0.000000	0.939435
59	0.000000	0.022908	0.000000	0.962343
60	0.000000	0.015150	0.000000	0.977493

61	0.000000	0.009569	0.000000	0.987062
61	0.000000	0.009569	0.000000	0.987062
62	0.000000	0.005780	0.000000	0.992841
63	0.000000	0.003344	0.000001	0.996185
64	0.000001	0.001855	0.000002	0.998041
65	0.000002	0.000989	0.000004	0.999029
66	0.000005	0.000506	0.000009	0.999536
67	0.000010	0.000250	0.000019	0.999785
68	0.000020	0.000119	0.000039	0.999904
69	0.000039	0.000054	0.000077	0.999959
70	0.000072	0.000024	0.000149	0.999983
71	0.000130	0.000010	0.000279	0.999993
72	0.000227	0.000004	0.000506	0.999997
73	0.000385	0.000002	0.000891	0.999999
74	0.000631	0.000001	0.001522	1.000000
75	0.001005	0.000000	0.002528	1.000000

When we examine figure 3-7, an initial buy of about 70 makes sense. It's unlikely that demand in the first fourteen weeks will exceed 70, but it's also unlikely that demand over twenty-six weeks will be less than 70, and you'll be forced to markdown a portion of the purchase at the end of the season.

We can take a more accurate approach to the task of determining an initial buy by using a marginal expected profitability calculation similar to what we did previously when we were making only a single buy for the season. This analysis will lead to an initial buy of 69, very close to the 70 we guessed to be right by looking at the graph. And the beauty of this approach is that the calculations can be automated within Microsoft Excel, saving the manual effort of guesstimating from a graph.

The reasoning is essentially a cost-benefit analysis on the marginal, 69th unit we buy. The benefit of this unit is that if demand exceeds 68, we will sell it, and it reduces by 1 our lost sales. The probability of demand exceeding 68 is 1 minus the probability demand is less than or equal to 68, or, using the "14-week cumulative probability" column for row 68 in table 3-8, 1 − 0.999904 = 0.000096.

The probability that we need to markdown this 69th unit is the probability that twenty-six-week demand is less than or equal to 68, which can be read from the "Season cumulative probability" column for row 68 in table 3-8 to be 0.000039. Using the margin of $20 for this item and the per-unit markdown cost of $22, we can compute the expected net profitability of the 69th unit to be (20 × 0.000096) − (22 × 0.000039) = .00106. Since this is greater than 0, the 69th unit is profit justified. It's easy to confirm that a 70th unit has negative expected net profitability and hence is not profit justified. The same analysis for the other two products shows that 71 is an optimal initial buy for the red cardigan and 68 is an optimal initial buy for the blue vest.

Once sales begin, you update your forecast and make a second buy if needed. How big should the second buy be? Table 3-9 shows sales in the first two weeks for the three items, an updated twenty-six-week forecast obtained by dividing the first two weeks' sales by 0.11 (that is, the 11 percent discussed above), and a forecast for the first fourteen weeks obtained by multiplying the twenty-six-week forecast by 0.55 (that is, the 55 percent discussed above). These calculations produce fractional values, which we have rounded to the nearest integer in table 3-9. The updated twenty-six-week forecast still has a margin of error, which you would take into account in determining a second buy, just as you would have with the initial buy. But this margin of error is much smaller, so for simplicity, assume that the updated forecast is perfectly accurate.

Table 3-9: Determining the second buy

	Initial buy quantity	**Sales in first 2 weeks**	**Updated 26-week forecast**	**Forecast sales for weeks 1–14**	**Forecast lost sales for weeks 1–14**	**Second buy quantity**
Navy turtleneck	69	9	82	45	0	13
Red	71	15	136	75	4	61

cardigan						
Blue vest	68	2	18	10	0	0

You clearly don't want to buy more of the blue vest. You're forecasting total-season sales of roughly 18, and you've already bought 68. The second buy for the navy turtleneck is set to the total-season forecast minus what you've already bought, or 13 (that is, 82 − 69), which avoids lost sales.

The second purchase for the red cardigan is trickier. You could follow the same approach that you used for the navy turtleneck and set the second buy to 65 (that is, 136 − 71). But this ignores the fact that forecast sales of 75 units for the first fourteen weeks exceeds the initial buy of 71, so you would have sold out during the first fourteen weeks and lost 4 units of demand. Thus a better second-period buy is 136 minus the 71 units already bought and sold and also minus the 4 units of lost demand, for 61 units of expected demand in weeks 15 through 26.

The first tableau of table 3-10 provides a financial summary of results for this sequence of two buys. Notice that going from one buy to two has dramatically improved net profit, from $9,710 in the best one-buy scenario shown in table 3-4 to $12,653 in the two-buy scenario.

Profits Take Flight

The twelve-week replenishment lead time in this example assumes that a Chinese factory made the product and shipped it by boat to the United States, which takes about a month. An alternative would be to ship the replenishment order by air, raising shipping cost by $1 per unit, but reducing lead time to eight weeks. The second tableau of table 3-10 shows revised initial and second buys as well as financial results for this approach. The second buy quantity was determined using the approach described above. Net profit has increased by $384, to $13,037. This increase is offset by the $102 cost of airfreight for 102 replenishment units at a cost of $1 per unit. Still, the arrangement yields a profit.

Table 3-10: Profit under the two-buy scenario

PROFIT UNDER THE TWO-BUY SCENARIO WITH A 12-WEEK LEAD TIME

	Actual demand weeks 1–14	Actual demand weeks 15–26	Initial buy qty	2nd buy qty	Lost sales in weeks 1–14	Lost sales in weeks 15–26	Sales	Gross margin on sales	Mark-down units	Loss on mark-downs	Net profit	Lost sales	Lost margin
Navy turtleneck	41	44	69	13	0	3	82	1,640	0	0	1,640	3	60
Red cardigan	70	62	71	61	0	0	132	10,956	0	0	10,956	0	0
Blue vest	17	12	68	0	0	0	29	1,305	39	1,248	57	0	0
Totals								13,901	39	1,248	12,653	3	60

PROFIT UNDER THE TWO-BUY SCENARIO WITH AN 8-WEEK LEAD TIME

	Actual demand weeks 1–10	Actual demand weeks 11–26	Initial buy qty	2nd buy qty	Lost sales in weeks 1–10	Lost sales in weeks 11–26	Sales	Gross margin on sales	Mark-down units	Loss on mark-downs	Net profit	Lost sales	Lost margin
Navy turtleneck	31	54	58	24	0	3	82	1,640	0	0	1,640	3	60
Red cardigan	50	82	58	74	0	0	132	10,956	0	0	10,956	0	0
Blue vest	10	19	56	0	0	0	29	1,305	27	864	441	0	0
Totals			172	102				13,901	31	980	13,037	3	60

Putting It All Together

We have used the catalog buying example to describe many different strategies for buying, including just buying what we forecast, hedging the buy by a somewhat arbitrary 10 percent, optimally hedging using a discrete and a continuous Gamma approach, and reacting to early sales with twelve- and eight-week lead times. Tables 3-4 and 3-10 show the steady increase in profit as we increased the sophistication of the buying strategy. Note that the big increase in profit comes from reacting to early sales, something we have found to be true in general.

Figure 3-8 shows the results of the methodology described here when applied to all twenty-seven items in the category. We estimated that taking a sharper pencil to the buying decisions that factored in risk increased margin by an amount that equaled 3.5 percent of what the base case revenues would have been under the current policy of simply buying what was forecast and then chasing the winners. We were also able to simulate what would happen with a shorter lead time; and, remarkably, cutting lead time from the current twelve weeks to nine weeks, as might be done with airfreight, adds 4 percent of revenue to margin dollars. This is because with a nine-week lead time, we can buy less initially, which reduces markdowns, and then jump on winners sooner, which avoids lost sales due to stockouts. Chapter 4 provides a recipe for reducing lead time and achieving overall supply chain flexibility.

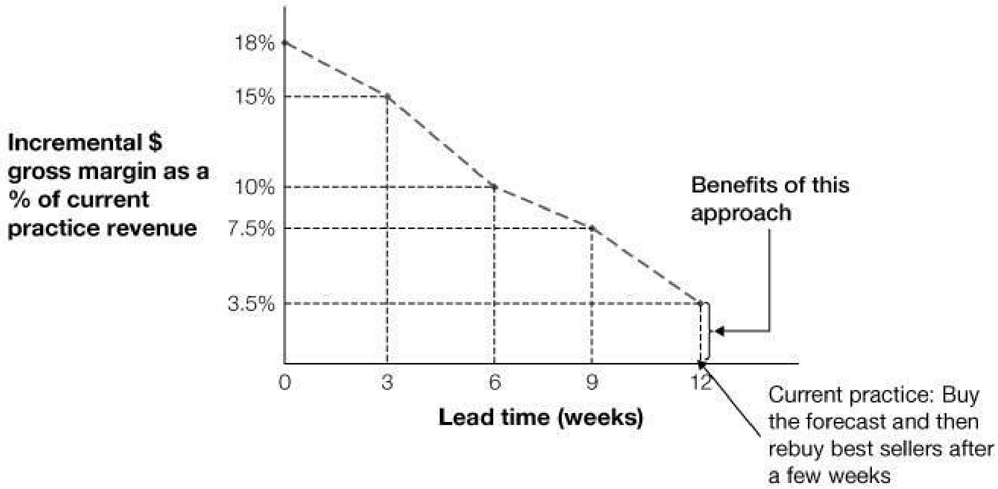


FIGURE 3-8: Profit increase when this approach is applied to all items

[1]The methodology described here was developed in collaboration with one of our former doctoral students, Kumar Rajaram, now a professor at UCLA's Anderson School. Additional details are reported in Marshall Fisher, Kumar Rajaram, and Ananth Raman, "Optimizing Inventory Replenishment of Retail Fashion Products," *Manufacturing & Service Operations Management* 3, no. 3 (2001): 230–241.

[2]The committee forecasting process we'll describe here, and its use in estimating the standard deviation of forecasts, was first developed in a project we conducted, together with Janice Hammond of Harvard Business School, at Sport Obermeyer, a fashion skiwear firm. The idea was suggested by Wally Obermeyer when he was president of Sport Obermeyer and working with us on this project. For more details see Marshall Fisher, Janice Hammond, Walter Obermeyer, and Ananth Raman, "Making Supply Meet Demand in an Uncertain World," *Harvard Business Review* , May–June 1994.

[3]In using this approach, we are ignoring for simplicity the fact that the 55 percent value also varies from season to season. This approach makes sense if this variation is small, which it is. We used the relationship that the probability that demand in the first fourteen weeks is less than or equal to X equals the probability that total season demand is less than or equal to X / 0.55 [1] to first find a continuous distribution, and then converted this to a discrete distribution using the approach described in table A-1.

Other Examples for Improving the Accuracy of Launch Forecasts

Reference Product from a Prior Season

The most common approach to initial demand forecasting is to identify a similar product, sold previously, and extrapolate from that product's prior sales. In using past sales, be sure to correct for sales lost due to stockouts and for the impact of other sales drivers that are not going to be repeated, such as promotions, or holidays like Easter and Thanksgiving, whose timing varies year to year.

The strength of this kind of forecasting is that it is easy to understand and do. The weakness is we humans. We're fallible, and we have to execute it. Choosing a reference product and adjusting its sales to factor out one-time events requires judgment. Thus this approach works best when the new product closely resembles one you sold before.

Using Test Sales

Test sales cost more, in time and money, but also yield more reliable results. With this method, you place small batches of the new product in a few stores and measure how well they sell. If the tests disappoint, you might even cancel the product. A variation of this approach is to introduce a new product in a few stores and then gradually roll it out to more stores as it proves itself. Here, you might use test sales to gauge which types of stores a product will sell best in.

As an example, consider a fall product like a wool sweater, sold September through December. Its lead time from order to placement in stores is four months. April would thus be a reasonable month in which to test it. On the basis of these sales, you'd place an order for delivery into stores at the start of September.

A challenge would be tempting customers to buy a product in April designed to be worn in the fall. You might sidestep the problem by testing during August. Assuming the manufacturer is in Asia, you'd ask that it ship the test goods via air, rather than boat, so they'd arrive in time for the test. If the product sold well, you'd order more at the end of August for arrival at the start of December to support Christmas sales.

Key questions in designing tests are picking stores to test in, including the right number, and predicting chain sales from the tests. We had a chance to ponder these issues in depth while working with a value-priced apparel seller that extensively tested products but doubted the accuracy of its forecasts. [4]

This outfit believed in the old quip "Will it play in Peoria?" For years, the conventional wisdom in consumer packaged-goods testing was that Peoria worked well as a test market because of its averageness. The city, in central Illinois, is about a three-hour drive from Chicago and is located near the country's geographic center. Retailers regarded it as average in terms of age, income, and family size. This particular retailer's philosophy was, "If one Peoria is good, then twenty-five are perfect." It would test products in twenty-five stores chosen to be as close as possible to chainwide averages on sales and other metrics. To predict chain-season sales, it simply multiplied the test sales, accounting for the total number of stores in the chain, and weeks in the season. If it sold 80 units in three weeks in twenty-five stores, its forecast for a twenty-four-week season in two hundred fifty stores would be 80 × 8 × 10 = 6,400. The factor of 8 scaled from three weeks to twenty-four, and the factor of 10 scaled the twenty-five stores to two hundred fifty stores. Trouble is, this approach ignores the effects of seasonality. A twenty-four-week season might have a higher (or lower) weekly sales rate than a three-week test.

To evaluate the retailer's approach, we matched actual sales in a recent fall season for knit tops with our reconstruction of the retailer's forecasts. (The original forecasts were no longer available.) The left panel of figure 3-9 shows our results, with each dot corresponding to a style/color combination. The horizontal axis shows the forecast, and the vertical axis shows actual demand. As you can see, the retailer's doubts about the accuracy of its tests were well founded. That didn't surprise us. It had more than one thousand stores and much diversity in its customers. Just within the Philadelphia region, it had urban, suburban, and rural locations, and the demographics of shoppers at each kind of location differed. We suggested trying to capture that diversity in the test sample and proposed clustering stores into similar groups. To conduct a test in ten stores, for example, we recommended identifying ten clusters of similar stores and choosing a test store from each.

We considered clustering on geography and demographics but concluded that what we should cluster on was taste, which was reflected in the mix of products that customers buy. We formed clusters based on the product mix that stores had sold in a prior season, following the process described in section A-2 in the appendix.

The right panel of figure 3-9 shows forecast versus actual results using this technique. Notice that the forecasts are much more accurate. Table 3-11 summarizes extensive testing of this approach at the apparel retailer and at three other retailers and shows that it cuts forecast error approximately in half compared with traditional testing.

A common alternative is clustering based on attributes such as store size, location, or climate. This approach is rife with problems. Location data is typically broken out by zip code, but many people don't shop where they live. Likewise, it's not clear which store attributes are the right ones to use, and choosing the wrong ones can do more harm than good. Many retailers, for example, cluster based on region or store size, but we found that differences in sales mix were essentially uncorrelated with location and store sales volume and instead correlated with ethnicity and average temperature at the store location.

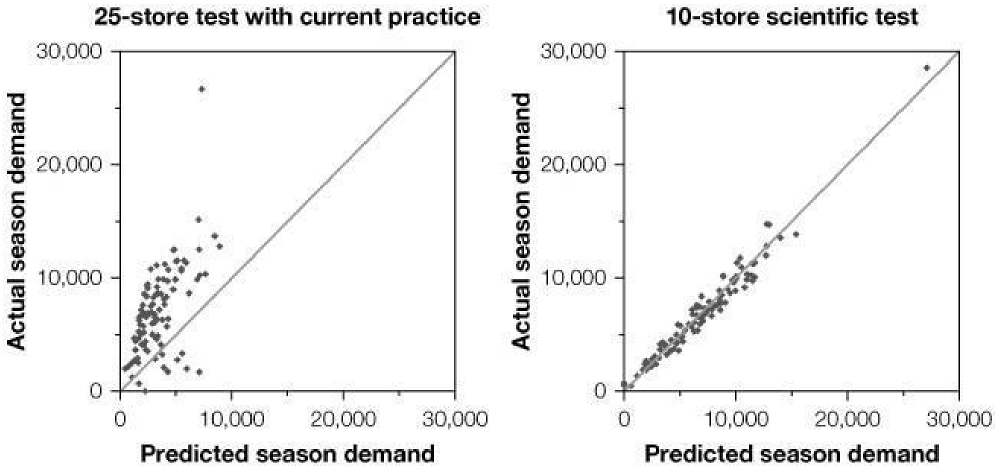


FIGURE 3-9: Scientific testing produces more accurate forecasts with fewer test stores

Store clusters have applications beyond testing. Retailers frequently use them to guide in planning, buying, and allocating merchandise. Cluster analysis enables a retailer to cater to differences in consumer taste stemming from factors such as gender, age, ethnicity, wealth, and climate. In this respect, clustering resembles segmentation in marketing. In the previous chapter, we described another approach to sales-based clustering, motivated by retailers' efforts to localize assortments. We have found that both approaches produce very similar clusters, and, for convenience, we would recommend using one set of clusters for both purposes.

Updating the Forecast Based on Early Sales

In 1998, a leading shoe seller contacted us, saying that it wanted to improve its forecasts within a season based on early sales. Shoes have a spring and a fall season, each six months long. The spring season starts in early January and extends through June. This retailer didn't sell a lot of shoes in January and February, but it did sell some, so using those sales to forecast the rest of the season made sense to us, and we proposed doing that. The retailer had reservations, pointing out that early buyers are more fashion-forward than later ones and thus not good predictors. We pressed on, arguing that a simple forecast would at least be a start. We tabulated total sandal sales by week for the 1997 season and noticed that 10.7 percent of total-season sales occurred in the first eight weeks of the season, so we created a 1998 total-season forecast by dividing sales in the first eight weeks by 0.107. We applied this model to history and found that it had an average forecast error of 34 percent.

Table 3-11: Results for four retailers

Retailer	New test process forecast error, 10 test stores	Traditional test forecast error, 10 test stores
Women's apparel	12.9%	41.9%
Shoe retailer 1	17.6%	27.9%
Shoe retailer 2	12.7%	22.7%
Home fabrics	17%	43%

Pondering the results, we thought, "Ouch, they were right. Simple extrapolation of early sales didn't work. A 34 percent forecast error stinks." We then talked with managers and discovered something. Often, the retailer had run out of inventory for style/color combinations that sold worse than forecast. In contrast, for those that sold better than forecast, the shoe seller had taken significant price markdowns relatively early in the season. We therefore tweaked our forecast to incorporate the impact of price and inventory on sales. This significantly improved forecast accuracy, as shown in figures 3-10 and 3-11. For a step-by-step explanation of how you can make the same sort of calculation, see section A-3 in the appendix.

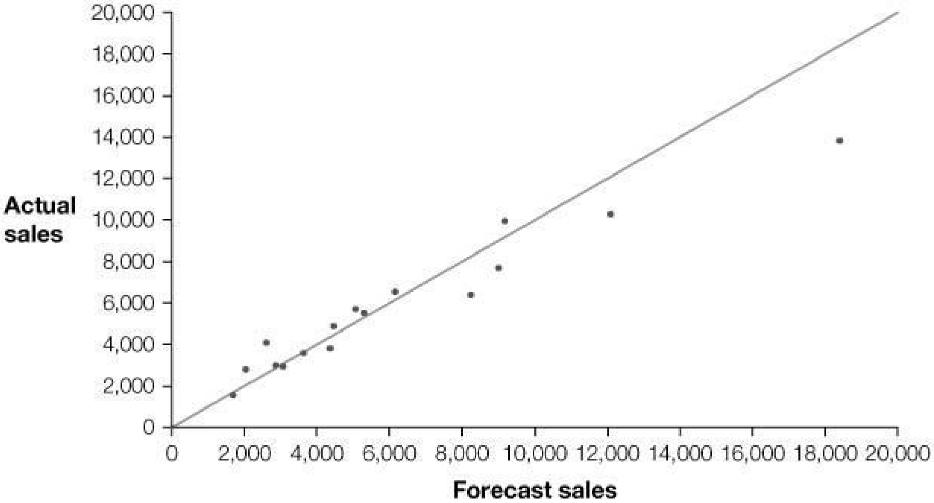


FIGURE 3-10: Taking price and inventory into account cut forecast error to 16%

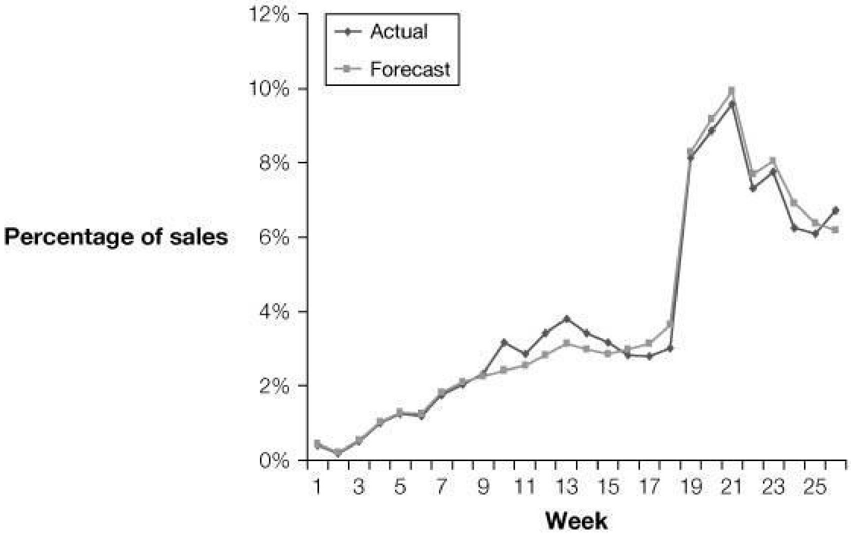


FIGURE 3-11: Weekly forecast versus actual for one sandal

At the end of February 1999, we used this formula to obtain a forecast of sandal sales for the 1999 spring season. Our results predicted that one style would be a runaway hit; our forecast based on early sales was much higher than the retailer's expectations. It had placed orders for twelve thousand pairs and wasn't planning to order more. But our analysis indicated that the retailer should order another eight thousand pairs and that this would increase its margin by nearly $200,000. So, as an experiment, it placed the order. The results are shown in table 3-12. Despite the fact that not all of the additional shoes arrived in time—supply can be just as uncertain as demand— our forecast proved to be accurate. The shoe seller increased its gross margin from $359,627 to $522,309.

[4]The approach to testing described here was developed in collaboration with one of our former doctoral students, Kumar Rajaram, now a professor at UCLA's Anderson School. Additional details are reported in Marshall Fisher and Kumar Rajaram, "Accurate Testing of Retail Merchandise: Methodology and Application," *Marketing Science* 19, no. 8 (2000): 266–278.

Replenishment

Many retail products have a life cycle of a year or more. Sometime during the launch phase, the retailer decides whether the product is a "keeper." If it is, it becomes part of the regular lineup and goes on replenishment. Most retailers have millions of store/SKU combinations on replenishment. For these, they collect sales data continuously and ship additional quantities to stores periodically to restock. They establish a target maximum inventory level, called an *order point* , for a given product and regularly ship enough to each store to replace what has sold. It's important to adjust the target inventory week to week as the sales rate changes due to seasonality and shifting popularity of the product.

Table 3-12: Application of this approach to one sandal

	Original plan	Optimized plan based on our analysis	Actual results
Unit receipts	12,469	20,410	18,575
Unit sales	9,882	14,875	14,579
Gross margin $	$385,497	$580,274	$562,269
End-of-season inventory units	2,587	5,535	3,996
Inventory purge cost	$25,870	$55,350	$39,960
Maintained margin $	$359,627	$524,924	$522,309

Sometimes, a store sells out of a product before new stock arrives. The higher the order point, the lower the chance of selling out, but a higher order point also implies more inventory and greater carrying costs. A key question is, "What's the right trade-off between carrying costs and the cost of lost sales due to stockouts?"

Here's a test of your retailing intuition. Suppose you tracked a group of items over time, and their in-stock rate was 95 percent. (The in-stock rate for a SKU is the percentage of stores in a week that *didn't* run out of that product before replenishment arrived.) Is that good or bad? Most people believe, correctly, that they can't answer this question without more information. They want to know such facts as the products' gross margins, the cost of carrying inventory, the likelihood a customer will buy a substitute product if they encounter a stockout, and how well competitors do at keeping merchandise in stock.

It's also important to realize that a 5 percent stockout rate can result in much greater lost sales. Table 3-13 shows an example of six products where the in-stock rate averages to 95 percent, and therefore the stockout rate is 5 percent. But because most of the stockouts are concentrated in the fastest-selling, highest-priced item, the lost-sales rate is 24 percent.

Many retailers report higher stockout rates on fast movers than on slow ones, and they often fixate on in-stock rate as a performance metric rather than *sales-capture rate* . Because it takes less inventory to maintain in-stock for a slow mover than a hot seller, they often end up increasing their in-stock rate by providing worse in-stock performance on their most popular products, to the detriment of their sales-capture rate.

Kevin Freeland, former senior vice president for inventory management of Best Buy, saw this problem firsthand when he visited one of that chain's stores. [5] He discovered that the store had sold out of three best-selling personal computers. At the time, PCs and peripherals contributed to more than a third of Best Buy's annual sales.

Freeland remarked to the staffer showing him around, "This is terrible. We're losing all these sales." The assistant zeroed in on the slow sellers, pointing out that because the store had plenty of inventory for those, its numbers looked good. Freeland promptly changed Best Buy's service metric from in-stock rate to sales-capture rate.

Table 3-13: Lost revenue can be much greater than the stockout rate indicates

Item	In-stock rate	Demand units per week	Price	Potential revenue	Average lost revenue per week
A	0.99	0.25	$10	$2.50	$0.03
B	0.99	0.5	9	4.50	0.05
C	0.99	1	5	5.00	0.05
D	0.99	0.4	15	6.00	0.06
E	0.99	0.9	10	9.00	0.09
F	0.75	10	100	1,000.00	250.00
Totals				1,027.00	250.27
Average in-stock rate = 95% Lost revenue = $250.27 / $1,027.00 = 24%					

Customers can respond in a variety of ways when they encounter a stockout, and how they respond determines the impact on your sales. In the best case, the stockout has no impact because the customer either returns after you've restocked or purchases a substitute. Another possibility, of course, is that you lose the sale because the customer either decides she can do without the item or buys it elsewhere. A third possibility is that you lose the sales on an entire basket of goods the customer planned to buy. And the fourth—and worst—possibility is that a customer, usually after encountering repeated stockouts, abandons your company and shops elsewhere.

Now that you've considered the behavior of this imaginary customer, ponder this question: what happens to your sales if your in-stock rate rises by 1 percent? Do sales rise by 1 percent, too? Or do they rise less than that or more? The answer, it turns out, depends on the relative percentage of times that you lose none, one, or more than one sale on that stockout. Our studies, in which we regressed sales against the in-stock rate for a large number of stores over several years, have found cases where the sales lift from a 1 percent in-stock improvement exceeded 1 percent because a stockout frequently caused the loss of a basket of potential purchases.

Profit Optimizing Replenishment Inventory

With this background, let's consider how to choose an optimal order point for a single item in a single store based on two costs: the cost of carrying inventory and the lost margin due to stockouts when we don't have inventory. Assume that the customer doesn't ever buy only that one item, but her other purchases are unaffected. Sure, that's unrealistic, but it will help illustrate the core concepts of inventory optimization.

We can find an optimal order point by using the marginal expected profitability approach we described for the catalog buying example. In that example, we imagined that we increased the buy quantity unit by unit and did an expected profitability calculation on each incremental unit. In the catalog example, the expected profit of an incremental unit was the probability it sold times the gross margin minus the probability it didn't sell times the markdown cost if it didn't sell.

In the replenishment case, an incremental unit of inventory injected by raising the reorder point by one only helps us in the weeks in which we would have stocked out without this unit. For example, if our in-stock rate at the current reorder point is 90 percent, then an incremental unit only helps us 10 percent of the time. For simplicity, assume we are making weekly replenishment shipments. Then an incremental unit provides value in only five weeks of the year, and in these weeks it increases sales by one unit. To obtain these five additional sales and the resulting gross margin, we need to carry the extra unit of inventory for one year. So, in general, the net expected profit from raising the order point by one is item gross margin times fifty-two weeks times (one minus the current in-stock rate) minus the cost of carrying an item for a year. We want to raise the order point to a level where the last unit we add is just marginally profitable, and we can find this point by solving the equation

item gross margin × 52 weeks × (1 − current in-stock rate) − the cost of carrying an item for a year = 0

to find that the optimal in-stock rate satisfies

optimal in-stock rate = (1 − cost of carrying an item for one year / 52 × item gross margin)

(Of course, if replenishment deliveries are made at a frequency other than weekly, we simply replace "52" in the formula above with the number of delivery cycles in a year.) Then the order point is set to the level where the probability of demand not exceeding that level in a week equals the optimal in-stock rate. This value can be found using a probability table of demand, such as the Gamma table we showed earlier in this chapter.

The formula provides a way to gauge whether you are carrying an economical level of inventory. If your in-stock rate is much higher than the formula would suggest, you're carrying too much, and vice versa. Of course, this formula is only approximate because of many real-world complications, such as the willingness of customers to substitute, but we have found it to be a good guide to see whether your inventory levels are in the right ballpark. Notice that a higher per-unit margin results in a higher in-stock rate, and a higher carrying cost results in a lower in-stock rate, which is intuitive.

An Implementation of These Ideas Had Major Impact

The two of us worked with 4R Systems, a software and consulting company, to develop a system for setting order points based on the concepts described here. That forced us to address many of the operating nuances of a real supply chain, such as seasonality, customer substitution, and varying lead times. Moreover, manufacturers ship most products in cases ranging from one unit to several dozen. This injects inventory into the system that serves as safety stock but also needs to be taken into account in setting order points. Likewise, you'll have to consider the minimum amount of inventory needed in your stores to create attractive displays.

Last but not least, we invested considerable effort in estimating the likely accuracy of forecasts. Forecast accuracy varied considerably by product, depending, for example, on sales velocity of the product and where the product was in its life cycle. The importance of estimating forecast accuracy was continually emphasized by 4R CEO Jiri Nechleba, who suggested that "admitting that one's forecasts have error, really uncertainty, is the first step towards making good decisions.”

4R Systems deployed its system in December 2002 at a major retailer of products for the home, including bedding and small appliances. Every week the retailer electronically transferred sales and inventory data on more than 5 million store/SKU combinations. 4R processed this data to compute order points for all store/SKU combinations. It then sent its results to the retailer, which shipped enough product to stores to bring on-hand plus on-order inventory to the order point.

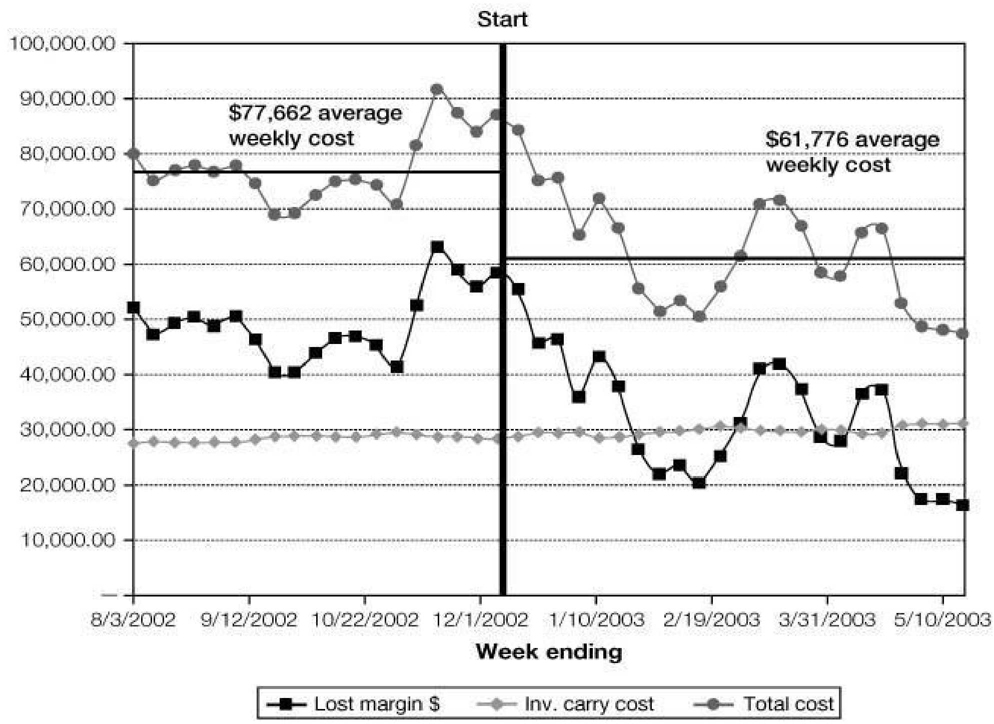


FIGURE 3-12: Results of an implementation

Figure 3-12 shows results for the first department in which the retailer used the 4R system. The line marked with diamonds shows inventory carrying cost, which equaled the inventory carrying cost rate multiplied by the dollar value of inventory on hand as shown on the department’s weekly inventory report. The inventory report also showed which items were sold out in each store. To estimate lost sales and lost margin, we used the historical sales rate of each store/SKU combination that sold out, paired with an assessment of the fraction of customers who would not substitute another item when they encountered a stockout. The line marked with squares shows this weekly lost-margin cost. The total cost line marked with circles is the sum of these two costs.

We determined the impact by comparing average total cost over a number of weeks before and after implementation. Average weekly cost decreased from $77,662 to $61,776, implying an annual benefit of $826,072—that is, 52 weeks × ($77,662 − $61,776). This was one of thirty-four departments, and the total annual benefit calculated this way for all departments was about $20 million.

[5]Freeland is currently chief operating officer, Advance Auto Parts.

Exit

Eventually, a product reaches the end of its life. Maybe the season changes, a manufacturer phases out a model, or sales slow because of new competitors. Regardless, as a product’s usefulness wanes, your goal changes from maintaining stock to selling off inventory as efficiently as possible.

The primary (and sometimes painful) tool for accomplishing this is a markdown. A typical product may be marked down several times in the store. If that doesn’t clear the inventory, it may eventually be sold to a discounter, usually for pennies on the dollar. In addition to markdowns, consolidating remaining inventory in a smaller number of stores can help dispose of goods like apparel, where the sales rate tends to slow toward the end of a season because the assortment is broken—that is, many sizes and colors have sold out. Five hundred stores of broken inventory might constitute one hundred stores’ worth of fully assorted inventory, so consolidation can accelerate sales. Transfers can also move product to stores where the climate allows for a longer sales season. You might, say, ship golf shirts from New England to Florida in the early fall, when New Englanders have begun to dig out their coats but Floridians are still seeing sunny, warm days.

Optimizing Markdowns

A crucial question in designing an exit is determining the depth and timing of markdowns to extract the greatest possible revenue. (You already own the goods, so product cost is a *sunk cost* , and maximizing revenue is equivalent to maximizing profit.) Doing this systematically involves three steps.

1. *Forecast* : estimate the sales lift that would result from a given markdown.
2. *Optimize* : Determine the depth and timing of markdowns to maximize revenue earned from the remaining inventory.
3. *Test* : Improve the accuracy of the estimates in step 1 through price tests in which you vary the price to determine the sales response to particular markdowns.

Figure 3-13 shows the sales lift experienced by several items that were marked down 20 percent, 50 percent, and 75 percent successively in a prior season. You can see that when these items were marked down by 50 percent, their sales rate increased by a factor of about 3. The curve shows the formula $e^{2.4\text{Markdown\%}}$, where e = 2.7182 is the base of the natural logarithm. This so-called exponential formula has been found to fit well how sales lift varies with depth of markdown, and in this case, 2.4 is the value in the exponent that best fits this data. [6]

The advantage of fitting a formula to the data is that next season you can use this formula as an estimate of the lift that would result from different markdown levels, assuming that you think customers will respond similarly. This enables optimization of markdowns.

We'll illustrate markdown optimization with an item for which you made a single buy of 3,300 units to cover demand for a sixteen-week season. The item is selling at 100 units per week, and for simplicity's sake, the example ignores seasonality. Without a markdown, you'd have 1,700 units left at the end of the season. Suppose you want to take a single markdown with either three weeks left in the season or ten weeks left. Which is the best time to markdown, and in each case, what level of markdown would maximize the revenue yield? To answer this question, we have compiled in table 3-14 the revenue that would result from different markdowns and for the two points in time at which a markdown could be taken, assuming that the revenue lift is given by the formula $e^{2.5\text{Markdown\%}}$, which you determined based on a prior season's markdown. We assume that the lift factor is the same whether you take a markdown with three weeks or ten weeks left in the season, although, usually, earlier markdowns produce a bigger lift. Unit sales in the table is computed as the minimum of the remaining inventory and the base sales rate of 100 times the sales lift factor times the number of weeks left in the season.

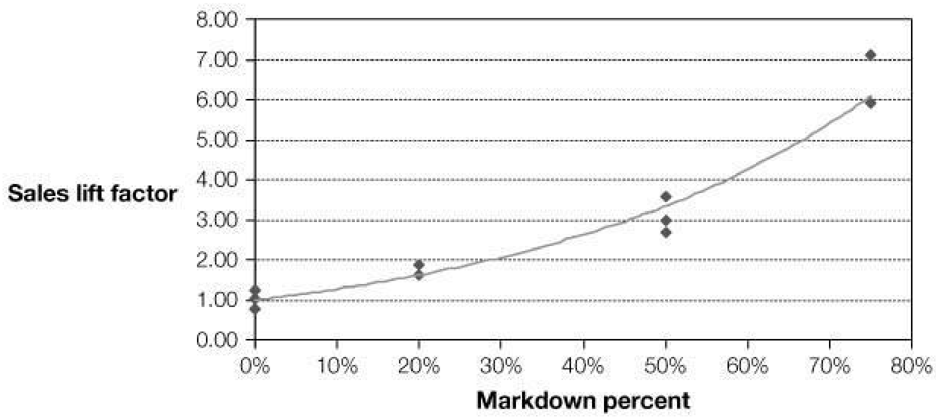


FIGURE 3-13: Estimating sales lift factor versus markdown percentage

Table 3-14: Markdown optimization example

Base sales rate			100 units per week			
Regular price			$60.00			
			3 weeks left in season		**10 weeks left in season**	
Markdown	**Sales lift factor**	**Sale price**	**Rest of season sales**	**Rest of season revenue**	**Rest of season sales**	**Rest of season revenue**

5%	1.13	57.00	340	19,377	1,133	64,589
10%	1.28	54.00	385	20,801	1,284	69,337
15%	1.45	51.00	436	22,261	1,455	74,205
20%	1.65	48.00	495	23,742	1,649	79,139
25%	1.87	45.00	560	25,221	1,868	84,071
30%	2.12	42.00	635	26,674	2,117	88,914
35%	2.40	39.00	720	28,067	2,399	93,556
40%	2.72	36.00	815	29,357	2,700	97,200
45%	3.08	33.00	924	30,494	2,700	89,100
50%	3.49	30.00	1,047	31,413	2,700	81,000
55%	3.96	27.00	1,187	32,036	2,700	72,900
60%	4.48	24.00	1,345	32,268	2,700	64,800
65%	5.08	21.00	1,524	31,994	2,700	56,700
70%	5.75	18.00	1,726	31,075	2,700	48,600
75%	6.52	15.00	1,956	29,344	2,700	40,500
80%	7.39	12.00	2,000	24,000	2,700	32,400
85%	8.37	9.00	2,000	18,000	2,700	24,300
90%	9.49	6.00	2,000	12,000	2,700	16,200
95%	10.75	3.00	2,000	6,000	2,700	8,100

You can see that for a markdown with three weeks left you'd achieve the maximum revenue of $32,268 by taking a markdown of 60 percent, and for a markdown with ten weeks left you would have achieved the maximum revenue of $97,200 with a markdown of 40 percent. Notice that these markdown levels do not sell all of the inventory. Retailers often consider markdowns a way to jettison *all* of their leftovers. But maximizing revenue makes more sense, assuming you can dispose of the leftovers at no cost by, say, donating them to charity.

Smart timing of a markdown matters as much as the magnitude. So which is better, a 60 percent markdown with three weeks remaining or a 40 percent markdown with ten weeks left? In the latter case, you'd receive $97,200 of revenue over ten weeks. In the former, you'd receive $32,268 during the three-week markdown period and $42,000 in the prior seven weeks (7 weeks × $60 × 100 units per week), for total revenue of $74,268. That's substantially less than you would've earned by taking a smaller markdown earlier.

Your buyers will be tempted to delay markdowns in the hopes that sales will increase, but usually they stunt your sales by doing this. A better rule is to take a markdown big enough to clear up the problem as soon as you realize that you bought too much. The consumer response to a small markdown in October will be much greater than to a larger markdown on December 26. In October, fall products are still in season and thus worth more.

Improving the Process with Price Testing

The shortcoming of the process described above is your ability to estimate consumer response to a markdown. Once again, a test can improve the accuracy of your estimates. We designed such a test for Zany Brainy, the toy retailer founded by David Schlesinger, to test alternative prices for three different toys. [7] This test was to set regular prices, but the same approach applies to markdowns.

We formed three matched panels of six stores chosen so that each store in a panel had a "twin" store in each of the other two panels that was as much like it as possible. We then identified low, medium, and high prices for each of the toys, as listed in table 3-15, and charged one of these prices in each of the three panels. We assigned the prices so that each panel had one toy each at the low, medium, and high prices. Table 3-16 shows sales at each of these prices over the six-week test period.

Product C surprised us. Its sales were much higher at the medium price than at the low or high ones. Here, we concluded that customers were using price as an indicator of quality. The product was an unbranded handheld two-way radio, so customers could not objectively evaluate the quality of its electronics. In contrast, product A was a branded product with known quality, while product B was an easy-to-assess wooden block set.

Table 3-15: Test prices and purchase costs

	PRICES($)			
	Low	**Medium**	**High**	**Purchase cost ($)**
Product A	19.99	24.99	29.99	11
Product B	24.99	29.99	34.99	18
Product C	14.99	19.99	24.99	11

Table 3-16: Total unit sales for the three test products at each of the three price levels

	TOTAL UNIT SALES		
	At low price	**At medium price**	**At high price**
Product A	7	4	3
Product B	33	26	14
Product C	47	74	36

For products A and B, we fit a demand curve and found optimal prices. Figure 3-14 shows the demand curve for product B. Figure 3-15 shows the optimal prices. We considered the 3 percent profit increase for product A to be significant.

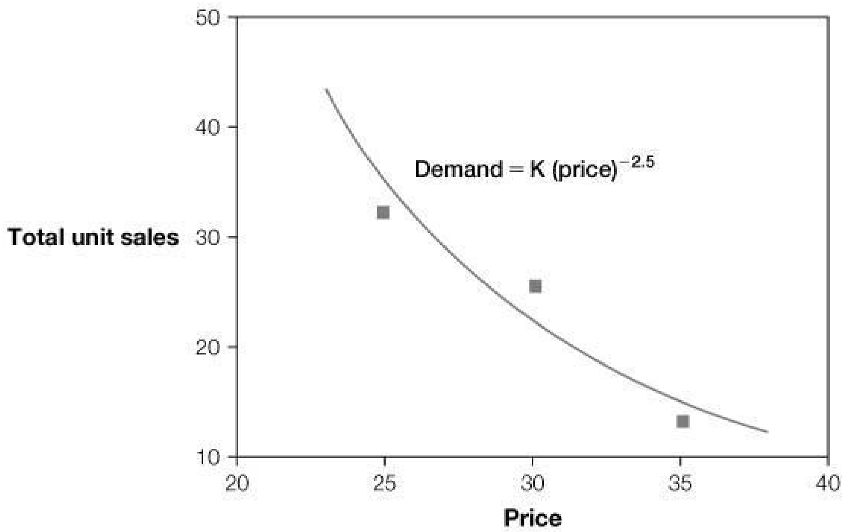


FIGURE 3-14: Demand curve estimation for product B

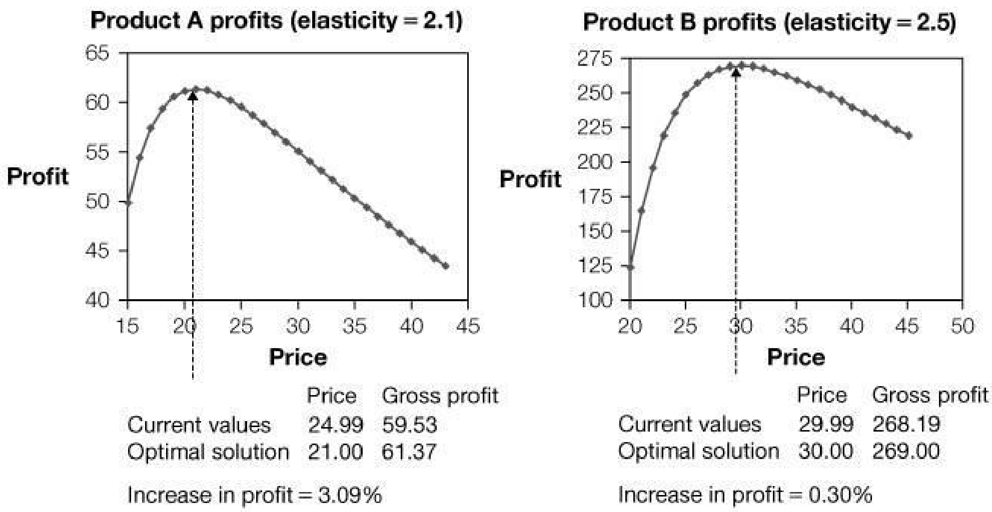


FIGURE 3-15: Finding optimal prices

An easy way to improve estimation accuracy while avoiding the effort of a formal test is to take a relatively small markdown at the start of the markdown season and use this to estimate consumer response. This estimate will be highly accurate because it is made based on current information about the particular products. Then take a second optimal markdown based on your estimate.

[6]To determine that 2.5 was the price elasticity that best fit the data, we made price elasticity a parameter in a cell of an Excel model containing the historical data, and then computed the forecast error measure mean absolute deviation (MAD) by, for each historical data point, finding the absolute difference between predicted and actual lift and summing across all points. Then we used the Excel function Solver to find the elasticity value that minimized MAD, which turned out to be 2.5. See Stephen Smith and Dale Achabal, “Clearance Pricing and Inventory Policies for Retail Chains,” *Management Science,* 1998, for additional details on estimating the sales impact of mark-downs.

[7]Facing severe competition, Zany Brainy discontinued operation in 2001. Schlesinger is now cofounder and president of Five Below, a leading extreme value retailer to the teen market and beyond.

Concluding Thoughts

Much of our focus in this chapter has been the launch of new products, because this is where forecasts are least accurate and inventory decisions most challenging. Here are the key principles to take away from our discussion.

Forecasting

- All forecasts are wrong, so you’ll have too much of some products and too little of others. But even a little sales data dramatically improves the accuracy of new-product forecasts.
- The margin of error on a forecast matters as much as the forecast itself.

Inventory

- Buying what you forecast in the preseason does not produce the best results. Buy more than you forecast if the cost of too little exceeds the cost of too much, and vice versa. The size of your hedge depends on the margin of error on the forecast.

Supply Flexibility

- For short-lived products, lead times often exceed product life, so you’ll be able to make only a single buy during the product’s life. But going from one buy to two—a preseason purchase and an early-season one—can have a huge impact on reducing stock-outs and end-of-season markdowns.
- Shortening lead time still further, perhaps through replacing traditional shipping with airfreight, can improve performance.