Interfaces
Vol. 36, No. 5, September–October 2006, pp. 407–419
ISSN 0092-2102 | EISSN 1526-551X | 06 | 3605 | 0407

DOI 10.1287/inte.1060.0232

A Major Appliance Manufacturer Rethinks Its Inventory Policies for Service Vehicles

Michael F. Gorman
Department of MIS, Operations Management and Decision Sciences, School of Business Administration, University of Dayton, 300 College Park, Dayton, Ohio 45469-2130, michael.gorman@udayton.edu

Sanjay Ahire
Moore School of Business, University of South Carolina, Columbia, South Carolina 29208, ahire@moore.sc.edu

A major appliance manufacturer needed to effectively manage $7 million in inventories of repair and replacement parts housed in more than 1,300 service vehicles used for warranty repairs. The manufacturer wanted to optimize the inventories while ensuring that its technicians carried the parts they needed to repair customers' appliances on the first visit. Typically, service vehicles carry only 400 of the 70,000 parts stocked centrally. Thus, the appropriate composition of vehicle stock is essential to attain acceptable levels of service and costs. We developed and applied an intuitive, efficient, effective, and easy-to-implement heuristic solution to this vehicle-stocking problem and found that the firm could trim more than $3 million per year from its warranty repair costs and increase its first-visit repair rate from 86 percent to 90 percent. The results of a pilot study supported these findings, and we estimated a payback of less than two years on full-scale implementation. The manufacturer is implementing these improved vehicle-stocking policies across its national network.

Key words: inventory: production, policies; industries: appliance.

A Fortune-100 firm manufacturing appliances manages $7 million worth of parts inventories for repairing appliances under warranty in more than 1,300 service vehicles across the United States. (Other after-market service providers perform out-of-warranty repairs.) The firm, which we call Alpha to preserve its anonymity, wanted to provide excellent service through first-visit repairs while keeping its inventory and labor costs for warranty service low.

Alpha determined that having its technicians fix problems in appliances under warranty on the first visit is the most important contributor to customer satisfaction. Customers dislike the inconvenience of waiting for appliances to be repaired. However, technicians do not know which parts they need until they diagnose the problem; therefore, they do not have the part they need for repair 14 percent of the time.

The cost of such stockouts includes the important but largely unquantifiable cost of potential lost future appliance sales and such hard costs as additional technician driving time, increased materials-handling and shipping costs for expedited or drop-shipped parts, and the overhead costs of managing open repair orders and scheduling additional service calls. Alpha focused on the hard-cost savings of our project but realized that the soft benefits were sizable.

Although firms can use historical demand patterns to plan for aggregate replacement-part inventories, when and where they will need particular parts is unpredictable because any specific product breakdown within one technician's service area is rare. Technicians typically provide an array of services that require stock so varied that the need for any one part over a time interval is a low-probability random event.

Finally, service vehicles have limited capacity for parts, materials, and equipment. Typically they carry only 400 of the 70,000 parts that could be needed on a service call. The challenge in stocking the service vehicle is to determine how many of the 70,000 parts to stock.

Vehicle Stocking

Alpha has a central parts warehouse and several regional warehouses that supply technicians with parts (Figure 1). The technician's use of a part from the vehicle triggers the replenishment process by generating a replacement order (Figure 2). A regional

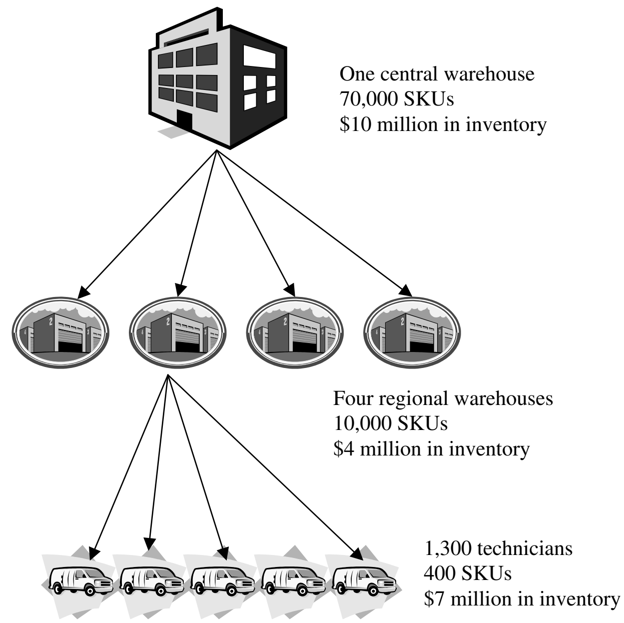

Figure 1: Alpha's parts-distribution network has three tiers, ending with inventory on the technicians' vehicles. Decisions on vehicle inventories strongly affect cost and service because the technicians are separated geographically from the regional warehouses.

part-courier network restocks technician vehicles. A four-day planned lead time from initiating the replacement order to delivery allows for lags in filling the order and delivery time from regional and central warehouses. When a technician needs a part that is not on the vehicle, Alpha ships it by air at increased expense to minimize the customer's dissatisfaction.

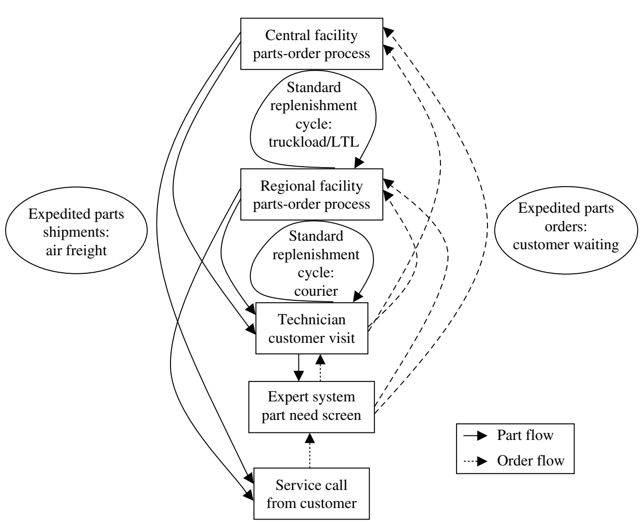

Figure 2: The standard replenishment cycle takes four days. If the technician does not have a replacement part needed for a repair on the vehicle, the repair is delayed until parts are expedited at high cost.

Previous Stocking-Level Policies

Until our project, Alpha relied on simplistic and often ineffective vehicle-stocking policies based on aggregate part-usage information. It used heuristics, such as "we should stock any part used more than five times in the last two months," which were coded in the central inventory distribution system. (These rules may have been invented by the programmer who developed the software.) Alpha's interest in minimizing global stocking levels caused excessive technician revisits for needed parts. Further, a single centralized policy for stocking service vehicles with parts failed to capture the differences in technicians' call patterns.

The technicians, whose chief productivity measure was completed repairs, recognized the problems with the centralized policies and sometimes refused to follow them. Thus, Alpha eventually dropped its central stocking policy and instead recommended the stocking level for each vehicle based on global average usage rates and stocking policies but allowed service technicians to determine their own vehicle's stock. Subsequently, the actual stock they carried deviated from Alpha's recommendations and depended more on the individual technician's intuition and opinion than on any rigorous or quantitative methodology. This subjective approach based on narrowly defined personal experience or preferences tended to overemphasize recently needed parts.

Previous Expert System

Alpha had tried to reduce the uncertainty in the parts required for each visit by screening customer calls with series of questions based on an expert system to identify the likely cause of the problem and the parts required. If Alpha could diagnose a problem in advance of the technician visit, it could check the service vehicle's stock to make sure it included the correct part before the technician made a first visit. Alpha's expert system allowed it to prestock the parts required for approximately one third of the visits that would otherwise have resulted in stockouts, reducing the number of visits technicians made without the appropriate part from approximately 20 percent of the visits to 14 percent.

The expert system was not as complete as Alpha would have liked because it needed to cover an enormous, ever-changing variety of appliances. Call-center employees could not always diagnose problems over the phone from the information customers provided, making it impossible to determine the parts needed. The system's credibility suffered when diagnoses were incorrect and unneeded parts were expedited or when technicians discovered that the repair required different parts than diagnosed.

Alpha realized that the expert system screening was helpful only when the part was not already in the vehicle's inventory. In 80 percent of the call screenings, the system determined needed parts that were routinely stocked on vehicles. When it identified a prestocking opportunity, the system could avoid technicians making second trips, but Alpha would still incur the costs of lost customer goodwill from repair delays, expedited shipments, and the overhead of technician scheduling if the part was not on the vehicle when the customer requested service. The probability that screening all calls would save technician trips was low, and the savings realized by screening were low. We therefore focused on ensuring that technicians would have the right parts on their vehicles when they made their calls, which would both improve service and reduce costs.

Six-Sigma Methodology

Alpha champions the six-sigma approach to improving operations and business processes. We began our project with a classic six-sigma framework, DMAIC (define, measure, analyze, improve, and control). Beyond the initial stages of defining and measuring the problem, the suite of six-sigma statistical tools (for example, Pareto analysis, regression analysis, and fishbone diagrams) proved insufficient. We persuaded Alpha to consider a broader set of operations research techniques. In the end, the business unit leaders acknowledged that the traditional six-sigma statistical tools would not have led them to the solution we developed.

Our Project

A team of three students, two faculty members, and four Alpha employees participated extensively in our project. To improve vehicle stocking, we needed to forecast part usage, devise a method of determining vehicle stock based on that forecast, and implement our solution (Figure 3).

Forecasting Part Usage

We wanted to predict part needs for a specific technician over the four-day replenishment lead time. We established the appropriate horizon for historical data, the format for forecast output, and statistical methodology to meet these needs.

We had to decide on the level of aggregation of different technicians' part-usage histories on which to base a technician forecast. On one hand, each technician had a unique order history based on the repairs he or she made. Such factors as customer income, demographics, and geographic region affect the types of calls technicians make and the models to be repaired. Even such subtle attributes as urban, suburban, or rural service areas affect the likely need for a part because of their correlation with customers' preferences for different product types. On the other hand, a 12-month history from a single technician may not provide enough information to predict part usage accurately. Anomalous events in a technician's history could skew recommendations for that technician too much toward those random events, leading to overstocking parts that happened to have a high replacement rate for the technician and understocking others that did not.

We therefore combined the part-usage histories of many technicians with similar attributes to form a larger sample on which to base a forecast. We used technician groups of four technicians from the same geographic region who served similar customer and product bases for our proof-of-concept analyses. We think that improved clustering can improve forecast accuracy and financial results and reduce the overhead for managing individual technicians' vehicle-stock levels. Alpha conducted cluster analyses of technicians' part-usage histories for its implementation to see whether it could create larger clusters to increase the historical basis for forecasts while still maintaining strong within-group homogeneity.

We recognized that accurately forecasting an appliance breakdown over a tight time horizon in a particular technician's territory is infeasible. The expected

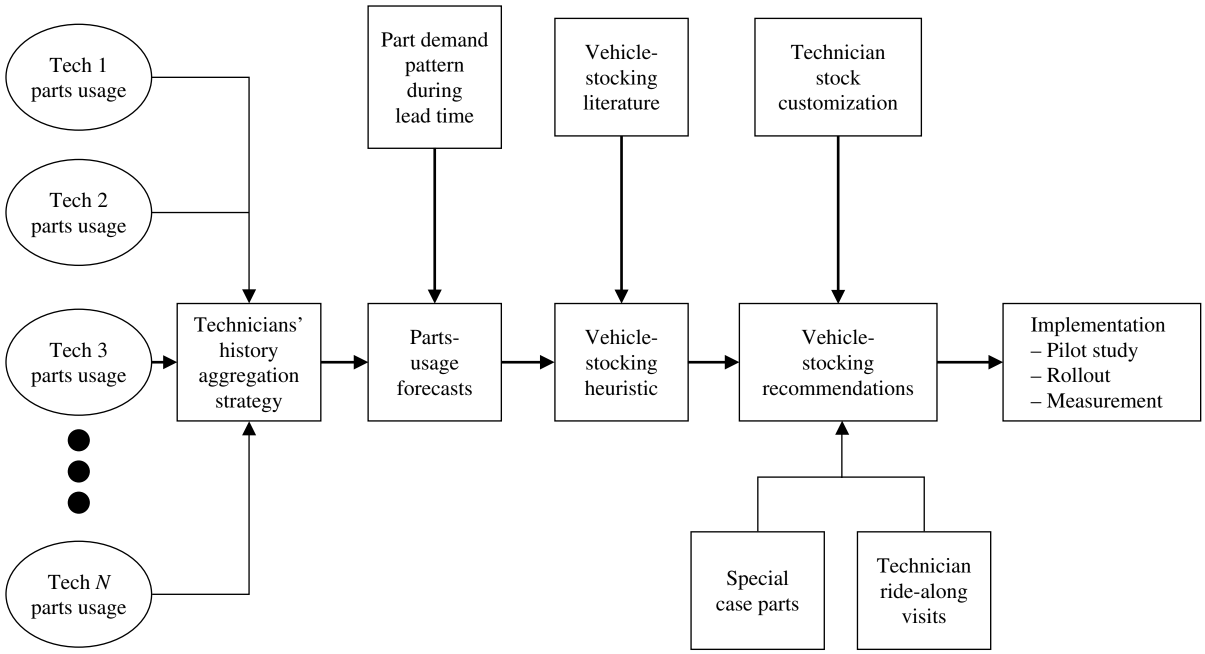

Figure 3: The steps we took in developing our recommendations for stocking technicians' vehicles included aggregating technicians' histories for developing part-usage forecasts and creating a vehicle-stocking heuristic. We rode along with technicians and customized vehicle stocks to suit technicians' preferences, which improved our credibility and technicians' acceptance of our recommendations.

number of uses of any one part may be very low during the replenishment lead time (zero, one, or two uses), so relative forecast errors can be high. We forecasted the total part usage for our technician group for the year to evaluate forecast accuracy and subsequently to estimate a mean arrival rate of orders per technician over the four-day lead time based on the aggregate estimate.

With a short forecast period with very low average demand, we faced the question, How many units of inventory should be held if the expected demand during the lead time is less than one? We decided that the goal of our forecast is not necessarily to determine a point estimate of demand for a part over the four days because the average demand during the lead time does not describe the possible range of demand that a technician might experience. We sought to predict the probability of different numbers of orders arriving over that lead time so that we could predict the probability of a stockout for a given inventory level.

We described the arrival pattern of orders over the lead time as a Poisson distribution (a discrete distribution that is a function of only its mean orders over a time interval). We had already observed that the mean expected arrival rate of any part from our forecast showed little seasonality and was independent of other part breakdowns. We calculated the probability of 1, 2, 3, up to M arrivals over the replenishment lead time under the assumption that order arrivals followed a Poisson distribution. We explored other, nonparametric order arrival rates but found that the Poisson performed well for the parts of interest and was easy to implement for the large number of parts to forecast.

Once we established our strategy for aggregating technicians and probabilistically forecasting, we considered many levels of sophistication to use in forecasting the expected number of part orders during the lead time. We found that product life-cycle effects were negligible. Although pronounced patterns emerge in part needs throughout the life cycle of an appliance, because our forecast is short-term, recent part-usage history (for example, the last 12 months) has the most predictive power. Shorter part-usage history (for example, the last three months) overemphasized recent orders, and the small sample allowed anomalous part orders to be too easily included in the vehicle stock. Longer forecast horizons tended to understate the need for parts for new appliances because past history contains no demand for these parts.

Alpha faces little seasonality in its part demand; its products are used year-round. Ultimately, we found that a simple average of the past 12 months' usage performed sufficiently well. We experimented with weighted-moving-average forecasts, but they contributed little to the accuracy of forecasts and, more important, had little impact on the estimated probabilities of order arrivals, a key determinant in our model.

Special circumstances, such as discontinuation of service for an appliance or special maintenance programs based on part recalls, posed an exceptional problem. For these few affected parts, weighted-moving-average forecasts adjusted more quickly than simple averages but still did not recognize known changes in business conditions quickly enough to be useful for forecasting. We therefore manually adjust forecasts for parts with special circumstances that make historical data inadequate for forecasting.

Vehicle-Stocking Literature

In developing our vehicle-stocking methodology, we drew on the established literature on the subject. Teunter (1998) and Teunter and Haneveld (2002a, b) provided comprehensive reviews of the assumptions employed in the literature on what they call the repair kit problem. Motivated by the copier-servicing problem at Xerox, Smith et al. (1980) presented one of the first published studies for strategies to minimize the costs of holding repair-parts inventory and of service failures. Subsequently, Graves (1982) solved a similar problem as an inventory-cost-minimization problem given a minimum service level. In a series of papers, Mamer and Smith (1982, 1985), March and Scudder (1984), and Brumelle and Granot (1993) relaxed the assumptions on independent part-failure rates and multiple parts per job that Smith et al. and Graves had employed. Because most of our repairs were based on independent, single-part failures, we did not use the methods employed in this line of research (appendix).

In early efforts, the researchers assumed that vehicles were restocked after each job; in other words, the technicians stocked their vehicles for a single customer visit (or single demand period). Thus, multiple jobs requiring the same part had no cumulative effect, and the decision becomes a binary (stock it or not) decision. Unfortunately, this is not the case in our problem. Because of the great distances from the central and regional warehouses and the geographically dispersed technicians, Alpha's typical (nonexpedited) restocking time for a part is four days. In that time, a technician may make 35 to 50 customer calls, any of which could require a recently used part that is currently on order.

Alpha asked us to recommend a stocking level for each part stocked given a known restocking period. In previous research, only Heeremans and Gelders (1995) had considered the multivisit model, in which the vehicle is not restocked after each repair. In this case, the decision changes from which parts should be carried to how many units of each part should be carried. However, Heeremans and Gelders assumed that the number of technician visits between vehicle restockings is known; in Alpha's case, we know the interval between restockings, but the exact number of the technician's customer visits varies over that period. The number of customer visits over the restocking period varies because of such factors as travel time to customers and the duration of visits, which depends on part availability and repair complexity.

Alpha emphasized inventory minimization going into this project, but given the high cost of customer-service failure, total cost minimization was a more accurate modeling goal. We differed from Heeremans and Gelders (1995) because we sought to minimize cost (like Smith et al. 1980), whereas they sought to minimize stocking cost subject to a service level constraint (like Graves 1982). Accordingly, they omitted the vehicle-space constraint in their solution approach because their approach naturally minimizes vehicle inventories.

Our Vehicle-Stocking Method

The vehicle-stocking problem is a complex combinatorial optimization problem that is time consuming to solve to optimality (appendix). To obtain good solutions quickly, we developed a single-pass greedy-heuristic approach. Our method considers three key part attributes when deciding which parts to hold in the vehicle stock: the probability of demand for the part, part cost, and part volume (in cubic feet). We must establish the mix and quantity of parts on the vehicle to best meet the service needs of the technician for the lead time between vehicle replenishments. In short, we recommend that the vehicle be stocked with parts that have the greatest expected net benefit per unit of volume consumed (NBV) in the service vehicle, subject to the constraint imposed by the size of the service vehicle (appendix).

We define the explicit cost during the lead time of holding each incremental unit, j, of inventory of part i (EC_i) as

$$\mathrm{EC}_i = C_i * h * L,$$

where h is the annual holding cost for inventory expressed in percentage terms, L is the lead time in years, and C_i is the price per unit of part i in dollar terms. Holding costs, h, may include any number of inventory-handling cost components, including interest costs, handling costs, shrinkage costs, and obsolescence costs. We define the cost of holding inventory over the lead time because this is the critical interval over which orders (demand for the part) may arrive without an opportunity to restock the service vehicle.

The expected cost of rescheduling, technician revisiting, and service failure determines the benefit of carrying a part in the service vehicle. We define the expected benefit of inventory part i, unit j (EB_{ij}), as

$$\mathrm{EB}_{ij} = \mathrm{RC} * (1 - P_{ij}),$$

where RC is the revisit cost associated with not carrying enough units of the required part (including loss of customer goodwill, scheduling overhead, and loss of technician productivity), and P_{ij} is the cumulative probability for all demand levels less than j units for part i during the restocking lead time. This benefit corresponds to the probability of demand greater than or equal to j for part i, and the revisit that would certainly result from not carrying the jth unit of part i in stock in the case when demand is greater than or equal to j. In summary, $(1 - P_{ij})$ represents the probability of a revisit corresponding to j units of stock for part i. As a cumulative probability, P_{ij} is nondecreasing in j; thus, EB_{ij} is nonincreasing in j for any demand distribution.

We recognized that our modeling of the problem would charge EB_{ij} twice if multiple parts were not available on a single job when, in actuality, this revisit cost would be incurred once (Mamer and Smith 1982, 1985; March and Scudder 1984; Brumelle and Granot 1993). We observed that for Alpha, the occurrences of both multiple part failure and multiple part inventory outages were rare and would affect only parts of marginal stocking value, so this simplification does not affect the accuracy of our recommendations.

The difference of the expected costs and benefits is the net benefit of carrying a unit j of part i (NB_{ij}):

$$\mathrm{NB}_{ij} = \mathrm{EB}_{ij} - \mathrm{EC}_i.$$

For some parts with very high value C_i (and thus high EC_i) and for others with highly unlikely demand, NB_{ij} is negative for the first unit of inventory ($j = 1$) and can be summarily excluded from consideration in the truck stock.

The opportunity cost of holding a part is based on the cubic volume consumed by unit j of part i in the service vehicle (v_i). We define the net benefit per unit of volume consumed (NBV) as

$$\mathrm{NBV}_{ij} = \mathrm{NB}_{ij} / v_i.$$

Each unit of inventory held implies less remaining space for other parts. The net benefit per volume (NBV_{ij}) is calculated from the cubic volume of each part. We can identify the parts that deliver high value for little sacrificed vehicle space using this metric.

We choose the parts that deliver the highest marginal benefit per cubic foot for stocking in service vehicles. We add candidate units, j, for all parts, i, in order of decreasing NBV_{ij}, until the vehicle reaches its cubic capacity.

Although not a likely situation for Alpha, which has high revisit costs, if the NBV_{ij} becomes negative before the capacity constraint is reached, then all parts with a positive NBV are stocked. In this case, Alpha could opt to carry additional units with negative NBV to attain some minimal target service level.

We considered alternative measures of space consumption (such as the square area of the part footprint), but we found cubic volume to be the best indicator of space consumption. We evaluated the importance of the weight of each part or the total

Cost of revisit (RC)	\$50
Days per year (D)	260 (52 weeks)
Lead time (T)	5 days
Lead time (years) ($L = T/D$)	0.0192 years
Replenishments per year (n)	52
Inventory carrying costs (h)	25% per year

Part number (i)	Unit cost (\$/unit) C_i	Unit volume (cubic feet/unit) v_i	Annual holding costs per unit (\$) $HC_i = h * C_i$	Explicit part cost during lead time (\$) $EC_i = HC_i * L$	Annual demand D_i	Average demand during lead time d_i
1	\$50	100	12.50	0.24	20	0.385
2	\$200	20	50.00	0.96	10	0.192
3	\$10	10	2.50	0.05	5	0.096
4	\$150	20	37.50	0.72	1	0.019
5	\$250	100	62.50	1.20	1	0.019

Table 1: In this example, we show the input data needed to determine the recommended vehicle stock. It depends not only on the cost of revisiting the customer (RC), the restocking lead time (L), the inventory carrying costs (h), and the product demand (d) and unit costs (C) but also on the cubic volume of the part (v).

allowable number of parts in the vehicle. As it turns out, because part packaging is lightweight and bulky, the vehicle's cubic capacity is exhausted before its weight capacity; thus, part weight is always a loose constraint and can be safely ignored.

As a result, our recommended vehicle-stock inventory carries a mix and quantity of parts that balances inventory-carrying costs, revisit costs, and the opportunity cost of each part carried in a vehicle.

Vehicle-Stocking Illustration

To illustrate our approach, we demonstrate the heuristic with an example of five parts of different dollar values, expected demand levels, and cubic volumes that drive their individual model-recommended stocking levels (Table 1). In this fictitious example, the estimated cost of a technician revisit is \$50 in lost productivity, and the cost of inventory holding is 25 percent of the part value per year. Lead time is five days out of a 260-day working year. There is no seasonality in order arrivals, so demand during lead time is a constant proportion of annual demand.

The revisit probabilities $(1 - P_{ij})$ corresponding to various values of j for the five parts with reference to the average lead-time demand for each part (Table 2) are based on a Poisson distribution. The expected benefit corresponding to each unit of inventory is derived from a revisit cost corresponding to not carrying the jth unit and the probability of the need for the jth unit. From the same example, the benefit associated with the first unit of part 1 is \$15.96 (\$50 * 31.93%); and for the second unit, is \$2.87 (\$50 * 5.75%). The expected cost of carrying a unit of part 1 over the five days (Table 1) is \$0.24 (\$50 * 25% * 5/260) and is constant for each unit of inventory of part 1. Thus, the net expected benefit of part 1 falls from \$15.96 to \$2.87 as the inventory level increases from one unit to two units. Given a volume of 100 cubic feet, the net benefit per cubic foot falls from \$0.16 to \$0.03.

We apply this method to all parts for all reasonable levels of inventory for each part i. We need consider only those units of each part that have a positive NBV for vehicle stocking. M_i equals the highest level of inventory j for part i with a positive NBV. The net benefit per cubic foot is positive for seven of the 20 (five parts times four inventory levels) candidate part inventory levels (Table 2). In this illustration, regardless of vehicle capacity, part 5 is completely excluded from vehicle-stock consideration; for the other parts, no more than three units are stocked because, for these cases, the cost of carrying the part exceeds the expected benefit of reduced revisits ($M_1 = 3$; $M_2 = 1$; $M_3 = 2$; $M_4 = 1$; $M_5 = 0$).

We determined the priority order of each part i and inventory unit j that enter the desired vehicle stock (Table 3). For example, the first unit of part 3 has the highest value for stocking and enters the vehicle's stock first, primarily because of its small size

Part number (i)	Probability $(1-P_{ij})$ of lead time demand level $\geq j$ given average lead time demand $= d_i$ (%)				Benefit of carrying inventory unit, j ($) $EB_{ij} = (1-P_{ij}) * RC$				Net benefit of unit of inventory, j ($) $NB_{ij} = EB_{ij} - EC_i$				Net benefit per cubic foot of each unit of inventory, j ($) $NBV_{ij} = Nb_{ij}/v_i$			
	$j=1$	$j=2$	$j=3$	$j=4$	$j=1$	$j=2$	$j=3$	$j=4$	$j=1$	$j=2$	$j=3$	$j=4$	$j=1$	$j=2$	$j=3$	$j=4$
1	31.93	5.75	0.71	0.07	15.96	2.87	0.36	0.03	15.72	2.63	0.12	(0.21)	**0.16**	**0.03**	0.001	(0.00)
2	17.49	1.63	0.10	0.00	8.75	0.81	0.05	0.00	7.79	(0.15)	(0.91)	(0.96)	**0.39**	(0.01)	(0.05)	(0.05)
3	9.17	0.43	0.01	0.00	4.58	0.22	0.01	0.00	4.54	0.17	(0.04)	(0.05)	**0.45**	**0.02**	(0.00)	(0.00)
4	1.90	0.02	0.00	0.00	0.95	0.01	0.00	0.00	0.23	(0.71)	(0.72)	(0.72)	**0.01**	(0.04)	(0.04)	(0.04)
5	1.90	0.02	0.00	0.00	0.95	0.01	0.00	0.00	(0.25)	(1.19)	(1.20)	(1.20)	(0.00)	(0.01)	(0.01)	(0.01)

Table 2: We base our stocking decisions on each part's net benefit per unit volume (NBV). The expected benefits (EB) of holding inventory depend on the probability of demand for each unit of each candidate part. (All probabilities are Poisson-based.) The expected cost (EC) of each part is the inventory holding cost during lead time. NBV is the difference between EB and EC divided by the unit volume of the part to account for the opportunity cost of the part's consumption of limited vehicle capacity.

and low cost. The first unit of part 2 enters second. The first and second units of part 1 enter the recommended stock of the vehicle as both the third and fourth highest value units (carry two units). Part 1 does not have the highest stocking priority because of its size; the opportunity cost of stocking numerous other parts in the space diminished its desirability. However, because of its high demand, the probability of a second arrival during lead time is so much greater than it is for the other parts with lower demand that it is also the fourth part to enter the vehicle stock. Thereafter, the second unit of part 3 enters the vehicle stock, and only then does the first unit of part 4 enter the vehicle stock. Part 4 has a high cost and low demand, but it is small, so it has a small positive net benefit per cubic foot. Part 5, because of its high cost and low demand, is not carried at all on the service vehicle. We continue to add units to the inventory until we reach M_i for each part, and we carry seven units of inventory in total. If this vehicle had a 300 cubic foot capacity, we would have reached the vehicle capacity before all parts could be stocked and stopped at the sixth unit of inventory (part 4, unit 1) because we would have hit our capacity constraint and would have had no room for the third unit of part 1. Correspondingly, there would have been a reduction in service and inventory driven by the limited capacity of the vehicle.

Stocking order	Part (i)	Unit (j)	NBV_{ij} ($)	Cumulative volume (cubic feet)	Recommended stock levels Part (i)	Unit (j)
1	3	1	0.45	10	1	3
2	2	1	0.39	30	2	1
3	1	1	0.16	130	3	2
4	1	2	0.03	230	4	1
5	3	2	0.02	240	5	0
6	4	1	0.01	260	Inventory units	7
7	1	3	0.00	360	Inventory value	$520

Table 3: After assigning NBV values to each unit of each part, we assign the highest-value parts to the vehicle stock and continue in decreasing order of NBV until we reach the vehicle capacity.

Results

We compared the attributes of our recommended vehicle-stock, service-performance, and financial results to the previous recommended stock for two different groups of technicians with different part histories and different recommended vehicle stocks (Table 4). Broadly speaking, we observed that our recommended unit count and dollar value of inventory are greater than the baseline policy before our analysis. Generally, we recommend the highest demand and smallest parts for stocking; however, some parts were small enough that, despite their infrequent use, we easily justified their inclusion in the vehicle stock because they can be carried in little space and have small opportunity costs when measured in space consumed.

A number of commonly ordered parts were obvious to include in the vehicle stock. However, because of the generally diminishing probability of each incremental order during the lead time, very few parts required multiunit stocking. As shown by the reduced average units per part, we recommended carrying fewer multiunit inventories of parts because the NBV for the second unit of a high-demand part was often significantly lower than the NBV for the first unit of the other parts.

To estimate the service performance of our recommended vehicle stock compared with the previously recommended stock, we tested two months of repair-parts orders (15 restocking cycles) for two clusters of four technicians (groups A and B). For the tests, we used two months that were not from our historical forecasting data (that is, an outside sample) to get a prediction of model performance. We evaluated each technician's stock performance in terms of recommending the appropriate parts and the correct inventory levels of each part. We show a two to four percent improvement in identifying the correct parts and a two to six percent improvement in identifying the correct stocking level of each part.

	Technician group A		Technician group B	
	Baseline	Recommended	Baseline	Recommended
Inventory				
Average unique parts	248	721	271	827
Average total parts in stock	371	878	337	979
Average units per part	1.5	1.22	1.24	1.18
Dollar value of parts in stock ($)	2,725	5,492	3,032	5,997
Average part cost ($)	7.34	6.26	8.99	6.13
Average part size (cubic feet)	0.16	0.073	0.19	0.072
Service performance				
Correct part on vehicle (%)	88	90	88	92
Enough units of each part on vehicle (%)	86	88	86	92
Financial performance				
Annual savings estimate per technician		$672 $2,491		$4,309
Overall savings opportunity (1,300 technicians)			$3,237,780	

Table 4: Our financial results in test cases for Alpha indicated that we should increase the recommended vehicle stock to improve service and to reduce total costs. Increasing the quantity and value of parts increased the total inventory holding cost, but those costs were more than justified by the increased technician productivity from reduced driving time for revisiting customers. We estimated the overall savings opportunity by multiplying the savings per technician by 1,300 technicians.

The resulting financial opportunity findings for two representative technician groupings (Table 4) conflicted with the corporate desire to reduce inventories; we recommended increasing them in terms of both the number and dollar value of parts on each vehicle. It may seem out of line to double inventory to achieve only a two to six percent increase in service performance. However, that improvement allows Alpha to reduce the number of revisits per technician from approximately 500 to 300 or 400 revisits per year (a 25 to 40 percent reduction), which constitutes a substantial savings in technician time alone, not including the value of customer goodwill.

The savings from increased technician productivity easily justified the increase in inventories. Based on the average savings per technician in our representative group, we estimate a companywide opportunity of $3.2 million from improved service savings net of increased inventory holding costs (Table 4). This savings estimate is based on a systemwide reduction in cost per customer visit derived from fewer second visits to customers. Technicians spend less driving time per customer order, resulting in more productive (new) customer visits per day, higher technician productivity, and eventually fewer technicians needed to maintain the customer base. Some additional cost savings that we did not include in our savings estimate are fewer expedited shipments for parts not in stock when needed and the nonquantified but undeniably substantial dollar value of improved customer service.

Implementing Our Vehicle-Stock Recommendations

We conducted an implementation project for five technicians from the same service region (different from those in our proof-of-concept study). We created technician clusters of approximately 40 technicians with similar usage patterns to create expected demand levels in this implementation. We discerned that the five pilot study technicians belonged to three different clusters based on their usage patterns, so we developed three different recommended vehicle stocks.

At the outset of the implementation, we found out that these technicians were already carrying nearly

	Service performance (First visit availability rate)		Efficiency (New calls per eight-hour day)	
Technician	Baseline (%)	Pilot (%)	Baseline	Pilot
Technician 1	84.5	87.5	6.6	7
Technician 2	84.9	83.7	7.3	7.5
Technician 3	80.2	83.4	6.2	6.4
Technician 4	74.5	77.7	6.3	6.6
Technician 5	83.1	82.9	7.3	6.9
Overall	81.4	83.0	6.74	6.88
Percentage improvement		1.6%		2.1%
Estimated dollar savings				$2,970,326

Table 5: We ran a test implementation of five technicians to validate the projected model results in practice before fully implementing the recommendations. We confirmed that improved part inventories on the vehicles improved technicians' efficiency. As a result, they served more customers satisfactorily in fewer technician hours, saving Alpha almost $3 million based on our implementation. We estimated overall dollar savings based on a 2.1 percent reduction in technician hours for 1,300 technicians. With these savings came improved service performance that will increase customer satisfaction and loyalty.

	Optimistic estimate	Conservative estimate
Ongoing benefits		
Improvement in first-visit repair	4%	2%
Technician time savings from improved first-visit repairs	$4,240,000	$2,120,000
Ongoing costs		
Inventory carrying costs	$378,620	$378,620
Implementation costs		
System programming	$8,500	$8,500
Warehouse processing costs for initial stocking	$2,061,037	$2,061,037
Return costs for unneeded parts on vehicle	$220,000	$220,000
Total implementation costs	$2,289,537	$2,289,537
Nondiscounted payback period (months)	7.1	15.8

Table 6: Alpha estimated that the improved truck stock has under a two-year nondiscounted payback period. It is developing a phased-in approach to limit the initial logistical costs of restocking vehicles (the bulk of the implementation costs). Although delaying the anticipated benefits, this approach reduces technicians' resistance to changing their vehicle inventories, thereby improving the probability of successful implementation.

50 percent more than the dollar value of the previously recommended stock, directionally consistent with our recommendation to increase stock. Accordingly, we recommended on average only a 33 percent increase over their actual baseline inventory (as compared with doubling Alpha's recommended stocking levels). In the rest of this discussion, we compare our recommendations with the actual technicians' stock and performance (rather than with Alpha's recommended levels).

In practice, three of the five technicians experienced an improvement in part-availability performance; two fared worse (Table 5). Overall, we found a 1.6 percentage point improvement in the first-time part availability rate. We also observed a 2.1 percent increase in technician productivity when measured in jobs per eight-hour day, which translates to increased service calls per technician or to fewer technicians required to meet current service needs. Based on our experience in this implementation, we estimate that Alpha can save almost $3 million dollars (a 2.1 percent reduction in technician hours across 1,300 technicians).

Alpha estimated a payback of less than two years for the return on the initial investment of the project (Table 6). Although the implementation revealed some practical problems in rolling in the new truck-stock policies, these benefits outweigh these problems significantly with respect to full-scale implementation of the recommended strategies.

Implementation Insights

The implementation helped us to identify some interesting but mostly short-term challenges in implementing the revised truck-stock strategies. To start with, technicians thought that the new policy had psychological costs (inertia and change resistance being the primary reasons). The organization has an entrenched unionized work culture among its technicians, and we knew that obtaining technician buy-in would be a major challenge. Because of the ineffectiveness of previous recommended stocking strategies, the technicians had ignored them and developed their own informal strategies. The consistency of the new recommendations with the technicians' own strategies helped with the transition, but there were still differences. To improve our understanding of the problem and of technicians' needs, we rode along with technicians. This hands-on approach to the improvement strategies helped us to gain the technicians' acceptance better than forcing them to adopt the revised stocking policies would have. Still, convincing 1,300 technicians is challenging. Our logical analysis coupled with our hands-on approach with sample technicians helped us to meet this challenge.

Second, our increased variety of parts caused more stocking locations within the vehicle, requiring more technician effort to track parts. Misplaced parts were partly responsible for the observed lower-than-predicted actual technician performance in both the baseline and the implementation. We found that the existing internal design of the technicians' vehicles was not ideal for carrying the recommended mix. Alpha has plans to redesign the vehicle's storage space to suit the revised part mix.

Third, in addition to accommodating the increased variety of parts, accommodating odd-shaped and small parts can be challenging. For small and odd-shaped parts, we added a penalty factor, effectively letting the model know that their dimensions did not fully capture the opportunity cost for these parts. The penalty factor had the side benefit of reducing the number of small parts to be stocked in the vehicle, one of the problems we had in implementation. We also ran the model for smaller cubic footage per van than was actually available, allowing technicians to personalize our recommended vehicle stock to gain their acceptance and further improve part availability. We thus also allowed storage space for special-needs parts (such as recalls) that the model could not identify.

While Alpha pursues technician buy-in and approval for the initial investment to institute this new policy, we continue to seek ways to reduce the initial cost and streamline the implementation of our recommendations. For example, we are fine-tuning our recommendations to ease the financial and psychological costs of transition by implementing a more gradual stocking recommendation. Because the model identifies parts with highest benefits first, we have recommended a top-10 report on parts with the highest net benefit that are not currently stocked on each technician's vehicle. Conversely, we have also recommended a hit list of the worst parts with the lowest NBV that the technician should consider discontinuing on the vehicle. This approach allows for more technician input and reduces the initial implementation costs by more gradually phasing in recommended stock items. An ongoing technician-specific report showing the actual technician-expedited part orders and those based on the recommended vehicle stock will help us to close the gap between actual and recommended stock. Alpha realizes that this difficulty is primarily about adaptation to change, and it should realize the benefits of the recommended parts mix to their full extent in due course.

With our approach, we can help technicians understand that, given the chance, the new method can work to improve their performance. These revised inventories would benefit them in terms of reducing the need for individual judgment, providing broader perspectives on part demands, and improving parts availability for service calls. All of these benefits lead to improvements in technician productivity (measured in service calls completed per day), in customer service-performance, and in customer loyalty.

Concluding Remarks

Our analyses identified potential savings of more than $3 million per year in Alpha's warranty repair costs and improvements in the first-visit repair rate from 86 percent to 90 percent. In a systematic implementation, we verified these results with an expected payback of less than two years. Alpha is currently implementing the recommended stocking strategy throughout its US network.

Appendix

We summarize the differences in assumptions in our study by adapting a table from Teunter and Haneveld (2002b). Previous researchers used different assumptions surrounding independence and coincidence of part failures, jobs between restockings, and planning horizon. The objectives used have been cost minimization (C) or inventory minimization with a minimum service constraint (S). Fill rate has been measured based on job fill rate – percentage of jobs completed (all required parts are present) (J) or part fill rate – percentage of parts available for all jobs

Assumptions employed	Smith et al. (1980)	Graves (1982)	Mamer and Smith (1982)	March and Scudder (1984)	Mamer and Smith (1985)	Brumelle and Granot (1993)	Heeremans and Gelders (1995)	Teunter and Haneveld (2002a)	Gorman and Ahire (2006, this article)
Independent failures of different part types	X	X					X		X
One part of any type needed per job	X	X					X	X	
Jobs between restockings	1	1	1	1	1	1	Fixed	1	Variable
Infinite planning horizon	X	X	X	X	X	X	X		X
Objective	C	S	C	C	C	C	S	C	C
Fill rate	J, P	J	J	J	J	J	J	P	P

Table A: Differences in assumptions used in various studies, adapted from Teunter and Haneveld (2002b).

undertaken (P). Our study is the first to allow variable jobs between restockings.

The vehicle-stocking model formulation is as follows:

$$\text{Maximize} \quad \sum_{i=1}^{N}\sum_{j=1}^{M}(\text{EB}_{ij} - \text{EC}_i)\cdot x_{ij} \qquad (1)$$

$$\text{subject to} \quad \sum_{i=1}^{N}\sum_{j=1}^{M} v_i \cdot x_{ij} \le V, \qquad (2)$$

$$\sum_{i=1}^{N}\sum_{j=1}^{M} p_{ij} \cdot x_{ij} \ge N\cdot\lambda, \qquad (3)$$

$$x_{ij} \in \{0,1\} \quad \text{for all } x_{ij}. \qquad (4)$$

Variable Definitions

$i =$ index for parts, $i = 1, \ldots, N$.
$j =$ index for a unit of a part, $j = 1, \ldots, M$.
$N =$ number of candidate parts.
$M_i =$ maximum units in vehicle inventory for each part, i, for which $\text{NBV}_{ij} > 0$.
$M =$ maximum units in vehicle inventory for any part $= \text{Max}(M_i)$ for all i, $i = 1, \ldots, N$.
$x_{ij} =$ binary decision variable, part i, unit j; 1 implies stock unit j of part i, 0 indicates do not carry unit j of part i.
$p_{ij} =$ probability of demand exactly equal to j units for part i during lead time, L.
$P_{ij} = \sum_{k=0}^{j-1} p_{ij}$; cumulative probability of demand less than j units for part i during lead time, L (P_{ij} is nondecreasing in j).
$1 - P_{ij} =$ probability of lead time demand greater than or equal to j units for part i.
$\text{RC} =$ technician revisit cost.
$\text{EB}_{ij} = (1 - P_{ij}) * \text{RC}$; expected benefit of carrying unit j of part i (nonincreasing in j).
$C_i =$ cost of part i.
$L =$ lead time (years).
$h =$]annual inventory-holding costs in percentage terms.
$V =$ total allowable vehicle-stocking volume (cubic area).
$\text{EC}_i = C_i * h * L$; explicit holding cost of part i during lead time in dollars.
$v_i =$ volume of a unit of part i in cubic feet.
$\lambda =$ minimum required part fill rate (0% to 100%).

Our vehicle-stocking heuristic can be summarized as follows:

(1) Find maximum units, M_i, and net benefit per unit of volume NBV_{ij} for each part i:

For $i = 1, \ldots, N$,
  $\text{EC}_i = C_i \cdot h \cdot L$,
  $M_i = V/v_i$,
  $j = 1$.
  Do
    $\text{EB}_{ij} = \text{RC}\cdot(1 - P_{ij})$.

If $\mathrm{EB}_{ij} < \mathrm{EC}_i$, then $M_i = j - 1$.
Else $\mathrm{NBV}_{ij} = (\mathrm{EB}_{ij} - \mathrm{EC}_i)/v_i$,
$j = j + 1$.
Until $(1 - P_{ij}) = 0$ or $\mathrm{EB}_{ij} < \mathrm{EC}_i$.
Next i.

(2) Establish stocking levels:

Sort stock units in order of decreasing NBV_{ij}.
Do while $\mathrm{NBV}_{ij} \geq 0$ and $\sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij} \cdot v_i \leq V$,
$x_{ij} = 1$.
End Do.

Acknowledgments
Our special thanks to Mary Jacquez, Maria Kruse, and Lee Lust for their help in executing this project.

References

Brumelle, S., D. Granot. 1993. The repair kit problem revisited. *Oper. Res.* **41**(5) 994–1006.

Graves, S. C. 1982. A multiple-item inventory model with a job completion criterion (note). *Management Sci.* **28**(11) 1334–1337.

Heeremans, D., L. Gelders. 1995. Multiple period repair kit problem with a job completion criteria: A case study. *Eur. J. Oper. Res.* **81**(2) 239–248.

Mamer, J. W., S. A. Smith. 1982. Optimizing field repair kits based on a job completion rate. *Management Sci.* **28**(11) 1328–1333.

Mamer, J. W., S. A. Smith. 1985. Job completion based inventory systems: Optimal policies for repair kits and spare machines. *Management Sci.* **31**(6) 703–718.

March, S. T., G. D. Scudder. 1984. On optimizing field repair kits based on a job completion rate. *Management Sci.* **30**(8) 1025–1028.

Smith, S. A., J. C. Chambers, E. Shlifer. 1980. Optimal inventories based on job completion rate for repairs requiring multiple items. *Management Sci.* **26**(8) 849–852.

Teunter, R. H. 1998. *Inventory Control of Service Parts in the Final Phase*. Labyrinth Publications, Groningen, The Netherlands.

Teunter, R. H., W. K. Haneveld. 2002a. Inventory control of service parts in the final phase. *Eur. J. Oper. Res.* **137**(3) 497–511.

Teunter, R. H., W. K. Haneveld. 2002b. Inventory control of service parts in the final phase: A central depot and repair kits. *Eur. J. Oper. Res.* **138**(1) 76–86.

The editor in chief has received a letter from an officer of the appliance company that includes the following statements:

"We were looking for ways to improve inventory levels in our service vehicles which are responsible for warranty and other repairs on appliances. Our predisposition was towards reducing inventory on each truck. However, after careful modeling and quantification of the costs and benefits of inventory, we found that smaller parts had a lower opportunity cost in terms of space taken in these vehicles, and that holding more of them and fewer large parts reduced our total cost of service.

"We learned that an increase in inventory value and parts will actually decrease our total cost of service from reduced 'revisits' of our technicians, and improved customer service by getting it fixed the first time. Based solely on reduced revisit costs, we estimate a save potential of \$3.2 million across the network of service vehicles. Additionally, we project a soft benefit from a 25% improvement in 'fix it the first time' percentage, improving our customer satisfaction. Based on this study, we have piloted the recommendations and hope to employ them networkwide.

"Dr. Ahire and Dr. Gorman's insights and contribution to the project were instrumental towards its success, and I endorse their work."