Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

OM Forum—OM Research: From Problem-Driven to Data-Driven Research

David Simchi-Levi

To cite this article:
David Simchi-Levi (2014) OM Forum—OM Research: From Problem-Driven to Data-Driven Research. Manufacturing & Service Operations Management 16(1):2-10. http://dx.doi.org/10.1287/msom.2013.0471

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 16, No. 1, Winter 2014, pp. 2–10
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)

informs
http://dx.doi.org/10.1287/msom.2013.0471

OM Forum

OM Research: From Problem-Driven to Data-Driven Research

David Simchi-Levi
Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, dslevi@mit.edu

This essay, based on my 2013 MSOM Distinguished Fellow lecture, presents my view on new opportunities and challenges for research in operations management.

Key words: operations management; research; impact on practice; problem-driven research; data-driven research; data-driven models; DDM
History: Received: October 1, 2013; accepted: November 18, 2013. Published online in *Articles in Advance* December 20, 2013.

Receiving the 2013 MSOM Distinguished Fellow Award is a unique opportunity to reflect on my research over the last two decades. It is also an opportunity to examine research in operations management (OM) in general and identify important challenges faced by our profession. Indeed, this is the time to reflect on changes in society and technology that should influence our research. This is exactly the objective of my essay.

For this purpose, I will tell a few stories describing the evolution of my research starting from my early years as a young assistant professor at Columbia University. As you will see, each story represents a shift in research direction driven by a combination of luck, intuition, and curiosity.

The first story is about the implementation of a school bus routing system in New York City. At the heart of this story is a shift in my research from theory to practice. The second story is about the implementation of process flexibility at Pepsi Bottling Group (PBG), a division of PepsiCo. It demonstrates how practice led to the development of a new theory that explains the observed effectiveness of process flexibility on companies from industries such as consumer packaged goods all the way to automotives. The third tale is about the implementation of a new approach for online retailing where theory and practice merge.

As you read these stories, it will become evident that there is a deep difference between the approach my colleagues and I took in the first two stories and the approach applied for the last tale. In some sense, the focus of the first two stories is on *problem-driven research*; that is, the starting point is a problem identified by an academic or an industry professional, and the challenge is to find answers by developing new theory or new insights about the problem at hand. The third story is different! Here the starting point is not a specific problem, but rather a large data set that allowed us to identify new opportunities for an online retailer. This is what I will refer to as *data-driven research*.

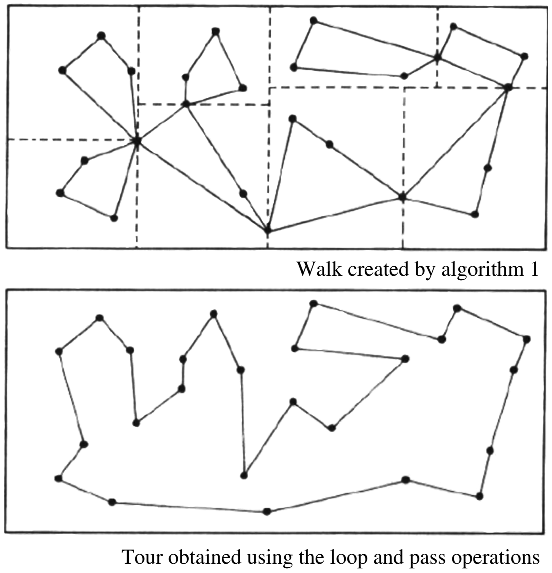
Thus, in the last section of this essay, I will spend time characterizing exactly what I mean by data-driven research, why it is relevant today more than ever before, and why it provides new opportunities for more creativity and a bigger and sometimes surprising impact on the organization. As you will see, this line of research can be quite different from what some in our profession refer to as empirical research.

1. From Theory to Practice

Theoretical analysis of algorithms for NP-hard problems has gone through remarkable developments since the early 1970s. My early research was influenced by Karp (1977), who observed that while the traveling salesman problem in the plane is NP-hard, there are simple, fast, local improvement heuristics that generate near-optimal solutions. To explain this behavior, he introduced region partitioning algorithms, algorithms that subdivide the region containing the cities into smaller regions, and construct an optimal tour connecting all cities in each subregion. These tours are connected to form a traveling salesman tour through all the cities in the entire region; see Figure 1.

In his paper, Karp (1977) showed that if cities' locations are randomly distributed, then the relative error between the length of the tour generated by his region partitioning algorithm and the length of the optimal

Figure 1 Karp's (1977) Illustration of a Region Partitioning Algorithm



Source. Reprinted by permission from Karp (1977, p. 212). © 1977, the Institute for Operations Research and the Management Sciences.

tour decreases to zero as problem size, i.e., the number of cites, increases, with probability one. These types of heuristics are referred to as *asymptotically optimal heuristics*.

This idea, that simple partitioning algorithms can generate solutions whose relative error with respect to the optimal solution decreases to zero, was then applied by Haimovich and Rinnooy Kan (1985) to a simple distribution problem, which is a special case of the classical capacitated vehicle routing problem. In their paper, Haimovich and Rinnooy Kan (1985) analyzed a distribution system where a set of customers with equal demand has to be served by a fleet of identical vehicles of limited capacity. The vehicles are initially located at a given depot. The objective is to find a set of routes for the vehicles of minimal total length. Each route begins at the depot, visits a subset of customers, and returns to the depot without violating the capacity constraint. Haimovich and Rinnooy Kan (1985) showed that the capacitated vehicle routing problem with equal demand is asymptotically solvable via several different region partitioning schemes.

At the end of their paper, Haimovich and Rinnooy Kan (1985) posed the following challenge to the community: is it possible to identify asymptotically optimal algorithms for general vehicle routing problems that include both unequal customer demand and a customer-specific time window constraint? The time window constraint specifies a period of time, referred to as a time window, in which this delivery must occur. In this version of the problem, much like in practice, the demand of a customer cannot be split over several vehicles; each customer's load must be delivered by a single vehicle.

My Ph.D. student Julien Bramel and I took on this challenge in the early 1990s. In a series of papers (see, e.g., Bramel and Simchi-Levi 1995, 1996), we showed that in an asymptotically optimal strategy, the total distance traveled by all vehicles will have a very simple structure consisting of two parts. The first is a round-trip distance the vehicle travels from the depot to the subregion where customers are located, which is referred to as a simple tour. The second is the additional distance incurred by visiting each customer in the region, which is called an insertion cost. This insight led to modeling general routing problems as large-scale facility location problems; see Bramel and Simchi-Levi (1995, 1996).

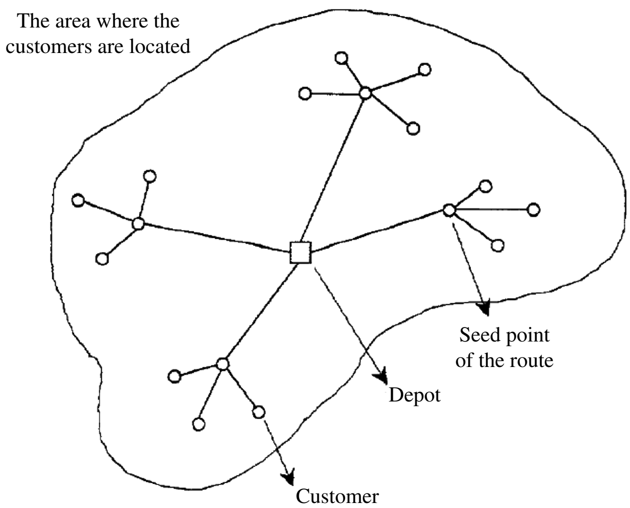
Facility location problems locate a set of facilities—possibly factories or warehouses—such as to minimize the fixed cost of opening the facilities and the variable transportation cost from the facilities to customer locations. Interestingly, when there are no time window constraints, the asymptotic results imply that the vehicle routing problem reduces to the celebrated capacitated facility location problem, where there is limited capacity on the amount of demand that each facility can serve. This asymptotically optimal heuristic, called the location-based heuristic, assigns customers to vehicles so as to minimize the sum of the length of all simple tours plus the total insertion costs of customers into each simple tour without violating the capacity or time window constraints.

To see the connection between the location problem and the routing problem, consider each customer in the routing problem as a potential site in the location problem. Solving the location problem by choosing a subset of the sites and assigning each customer to a single site corresponds to a solution to the routing problem where a simple tour is constructed between each site and the depot. Customers who are assigned to this site are inserted to the simple route in an efficient way; see Figure 2.

The reader may wonder about the effectiveness of such a heuristic, since location problems are also NP-hard. Although this is true, not all NP-hard problems are equally difficult to solve in practice. Indeed, as observed in Simchi-Levi, Chen, and Bramel (2013), large-scale location and network design problems can be solved efficiently by today's mixed integer programming software.

Of course, this type of analysis that identifies asymptotically optimal heuristics has important drawbacks. First, it requires large size instances. Second, for this analysis to be tractable, it is sometimes necessary to assume independent customer behavior. Finally, the general vehicle routing problem analyzed by Bramel and Simchi-Levi (1996) is still a simplified version

Figure 2 An Example of the Location-Based Heuristic



Source. Reprinted by permission from Bramel and Simchi-Levi (1995, p. 651). © 1995, the Institute for Operations Research and the Management Sciences.

of the usually more complex problems that appear in practice. Typically, a vehicle routing problem will have many constraints on the type of routes that can be constructed including multiple vehicle types, distance constraints, and complex restrictions on items that can be delivered on the same vehicle. But, this type of analysis has one important advantage. It provides insight into the structure of efficient algorithms, both from cost and computational time points of view.

This body of research could have stayed pretty remote from practice had I not received a phone call in 1992 from the New York City Board of Education asking for help in developing a decision support system for school bus routing. At that time, the New York City school system comprised 1,069 schools and approximately one million students. About 100,000 students rode around 1,150 buses that were leased by the Board of Education.

The size of the school bus routing problem in New York City provided a golden opportunity to test the effectiveness of the location-based heuristic—an asymptotically optimal heuristic—in practice. We started with the "Manhattan Project," an implementation and validation of the heuristic with data only from Manhattan. The algorithm showed great promise—it used substantially fewer buses than what the city used; see Braca et al. (1997). These results motivated the city, with our help, to integrate the algorithm as part of a decision support system that included a large database, a module for visualization of pick-up points, routes displayed on a geographic information system, and our optimization engine. In 1994, this system won the first-place prize in the Government/Public Sector category of the Windows World Open, a competition sponsored by, among others, Microsoft, Borland, AT&T, the Windows World Conference, and the magazine *Computerworld*.

This is just one example of how my research began with a problem posed in the academic literature, the development of new theory around this problem, and the implementation of this theory to a specific public or private organization. This insight, that applying theory to practice can make a big impact, was the foundation that led to LogicTools (now part of IBM), a software company that my wife, Edith, and I founded in 1996. When LogicTools was sold in 2007, we had more than 300 clients using our technology to improve supply chain performance.

2. From Practice to Theory

In early 2005, Pepsi Bottling Group approached me at the Massachusetts Institute of Technology (MIT) with a daunting challenge. Consumer preference was shifting from carbonated drinks to noncarbonated drinks and from cans to bottles. At that time, PBG produced these newly preferred products in a limited number of plants, resulting in half of their plants operating at capacity and leading to service problems during periods of peak demand; see Simchi-Levi (2010).

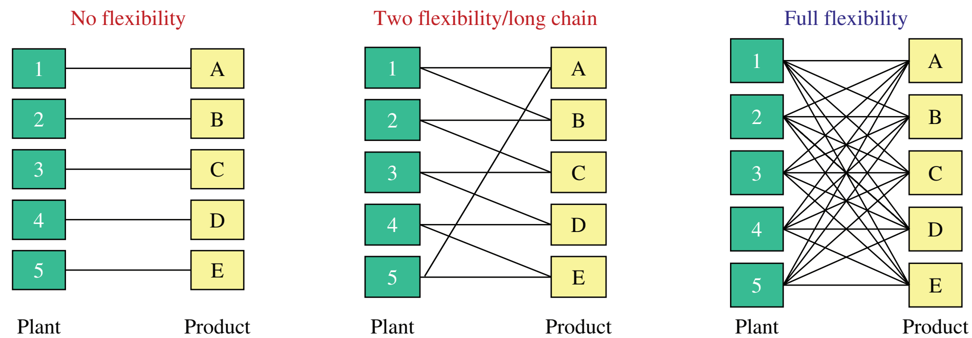
The MIT-PBG approach to the challenge was surprisingly simple. It focused on process flexibility defined as the ability to produce different products in the same manufacturing plant at the same time. In this strategy, each plant is capable of producing just a few of the products, and every quarter, PBG would assign products to plants to match with consumer preference. This strategy improved supply chain performance by significantly reducing out-of-stock levels, effectively adding one and a half production lines' worth of capacity to PBG's supply chain without any capital expenditure; see Simchi-Levi (2010).

The amazing observation we made about the implementation of process flexibility is that its effectiveness at PBG was independent of the time of year or the business unit where PBG implemented the new strategy. For the purpose of serving its customers in North America, the region was divided into seven business units, each served with a different number of plants. But, independent of the number of plants, process flexibility always significantly increased service levels at PBG. The natural question to ask was, why?

This is the challenge that my Ph.D. student Yehua Wei and I took in 2008: to develop a theory that explains the effectiveness of process flexibility and, hopefully, to generate new insights that will help introduce more robust strategies that help match supply with demand even under disruption.

Of course, this is not the first attempt to explain the effectiveness of process flexibility. Previous research

Figure 3 **Three Different Process Flexibility Designs**



had been fruitful in explaining the effectiveness of process flexibility when the system size is large, i.e., by applying an asymptotic analysis. But our experience with PBG was that process flexibility was effectively matching supply with demand independent of system size.

To present the theory developed in Simchi-Levi and Wei (2012), consider Figure 3, which depicts three different system designs of a supply chain with five manufacturing facilities and five product families. Under full flexibility, each plant is capable of producing all products, and hence when the demand for one product is higher than expected while the demand for a different product is lower than expected, a flexible manufacturing system can quickly make adjustments by shifting production capacities appropriately. By contrast, with a "dedicated" strategy (sometimes called "no flexibility"), each plant is responsible for a single product and hence does not have the same ability to match supply with demand. Of course, the dedicated strategy reduces equipment requirements, raw material inventory, labor training, and manufacturing cost relative to full flexibility.

Between these two extremes there are various designs that try to balance the low cost of no flexibility with the ability to match supply and demand in full flexibility. For instance, in a two-flexibility design, each plant produces exactly two product families, and each product is produced by exactly two plants.

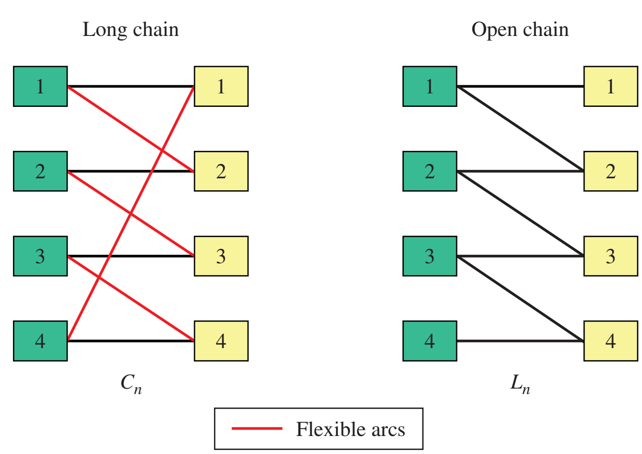
Of course, there are many ways to implement sparse designs, and the challenge is to identify an effective one. An important concept analyzed in the literature and applied in practice by various companies is the concept of the long chain. The long chain is a two-flexibility design that creates one cycle connecting all the plants and all the products; see Figure 3 for an example.

The first to observe the power of the long chain were Jordan and Graves (1995), who, through empirical analysis, showed that the long chain design can provide almost as much benefit as full flexibility. In particular, they found that for randomly generated demand, the expected amount of demand that can be satisfied by a long chain design is very close to that of a full flexibility design. Unfortunately, with a few exceptions (see Chou et al. 2010), there is very little theory to explain why the long chain design works so well. This challenge is addressed by Simchi-Levi and Wei (2012).

In our paper, we showed that many observations about the behavior of the long chain in practice can be explained by two deterministic properties, *supermodularity* and *decomposition*. To introduce these properties, we need a few notations and assumptions.

Consider a manufacturing network with n plants, each with unit capacity, and n products facing independent and identically distributed (i.i.d.) demand whose expected value is one. We index plants from 1 to n and products from 1 to n. An arc connecting plant i to product j implies that plant i is capable of producing product j. We refer to arc (i, i) connecting plant i to product i as a dedicated arc, and arc (i, j) connecting plant i to product j, $i \neq j$, as a flexible arc. We use C_n and L_n to denote the long chain design and the open chain design, respectively, containing n plants and n products. An open chain is defined as a long chain without one of its flexible arcs; see Figure 4.

Figure 4 **Long Chain, Flexible Arcs, and Open Chain**



The first main property we derived, supermodularity of the long chain, states that, given realized demand, the flexible arcs in the long chain are supermodular with respect to each other; that is, the existence of each flexible arc makes all other flexible arcs more valuable. A byproduct of this property is that under i.i.d. demand, the increase in expected sales obtained when one starts with a dedicated design and adds flexible arcs one at a time to construct the long chain is always increasing. This explains the numerical observations made by Graves (2008) and Hopp et al. (2004). Indeed, it highlights the importance of "closing the chain," that is, creating a long chain by adding the last flexible arc to an open chain, as a way to increase expected sales.

The second deterministic property, decomposition of the long chain, states that given realized demand, the sales of the long chain can be decomposed into a sum of n quantities, where each quantity is equal to the difference between sales of two open chains, one with n products and n plants and one with only $n-1$ products and $n-1$ plants. In particular, under i.i.d. demand and for any positive integer n, this property implies that the expected sales of C_n, the long chain, equal n times the difference between the expected sales of L_n, an open chain of size n, and L_{n-1}, an open chain of size $n-1$. Thus, instead of the need to characterize sales of the long chain, it is sufficient to characterize sales of open chains.

It turns out that it is easier to characterize sales of open chains than sales of long chains. Thus, the decomposition property allows us to provide a strong theoretical justification to several key observations on the performance of the long chain and other sparse flexibility designs. For example, it allows us to show that the long chain design maximizes expected sales among all two-flexibility designs.

To present additional insights, define the fill rate of a specific flexibility design and i.i.d. demand to be the ratio between expected sales associated with this design and total expected demand across all products. Interestingly, supermodularity and decomposition lead to a risk pooling result implying that the fill rate of the long chain increases with the number of products, but this increase converges to zero exponentially fast. This has interesting implications for system design. It implies that in large systems, a collection of several disjoint large chains can work as well as a single long chain!

Equally important, these properties allow us to bound the difference between the fill rate of full flexibility and that of the long chain under systems of any finite size by the asymptotic difference between the fill rate of full flexibility and that of the long chain. This result connects nicely with the asymptotic result of Chou et al. (2010). For example, when product demand is drawn from the Normal distribution with a coefficient of variation equal to $1/3$, the asymptotic difference between the fill rate of full flexibility and that of the long chain is shown to be 4%. Hence, our results imply that the difference between the fill rate of full flexibility and that of the long chain for any system size is bounded by 4%, quite a remarkable performance for such a sparse flexibility design.

Motivated by the insights from the analysis, Simchi-Levi, Wang, and Wei (2013) developed a new approach for supply chain resiliency that combines process flexibility and strategic inventory. We measured the level of supply chain robustness by proposing the concept of time to survive, defined as the maximum time that no customer demand is lost, regardless of which plant is disrupted. Recently, the Ford Motor Company implemented a related approach to mitigate supply chain disruption; see Simchi-Levi, Schmidt, and Wei (2014).

3. Merging Theory and Practice

Increased computing power and the explosion of data are changing the way organizations capture data, analyze information, and make decisions. These changes provide opportunities for the OM community to *analyze extensive data to identify new models that drive decisions and actions.* But connecting the realms of data, models, and decisions will require our community to move out of our comfort zone and address intellectual challenges that demand bringing together statistics, computational science, and operations research (OR) techniques.

At its heart, this research direction is about letting the data tell a story that helps identify new opportunities for research, including, in particular, new models that we have not analyzed before. An interesting question is how much data-driven research already exists in the OR and management science literature in general or in the OM literature in particular. Surprisingly, this research is hard to find. Perhaps the reason for this is that some practice-oriented research actually starts with data and without any specific problem or model in mind, yet this research is not presented in such a way in papers; however, in my experience, practice-oriented research, even when initiated by industry professionals, typically begins with a specific problem in mind.

A rare exception is the paper by my colleague Richard Larson (1990), "The Queue Inference Engine: Deducing Queue Statistics from Transactional Data." In his paper, Larson (1990) asks the following question: Given customer transactional data generated by ATMs, including recorded times of service commencement and service completion for each customer, is it possible to infer the number of customers waiting to use the machines? His answer involved developing a model that allows bank managers to derive queue statistics from customer transactional data and hence decide if and when they need to add or remove ATMs.

In a recent discussion with Professor Larson, he observed that initially he focused on a specific question proposed to him by BayBanks, one of the largest banks in the late 1980s, but as soon as he received the data, he realized the new opportunities for insightful research. As Professor Larson told me in a recent message, "I asked the question: Is there any additional information I can get out of all this transactional data beyond the usual steady state queue results? And so the QIE [queue inference engine] was born" (Larson 2013). This is just a beautiful example of how data drive new research, but sadly we have very few examples of this type.

In what follows, I report a second example of data-driven research from recent work with an online retailer; see Johnson et al. (2013) and Simchi-Levi, Wang, and Weinstein (2013). This retailer offers new products via online 48-hour sales events. Since every event involves products from different designers and with different styles, the online retailer has very little knowledge of customer demand as a function of price. As a result, the retailer applies a single price during a typical 48-hour event by surveying prices for similar products offered by other retailers.

The data provide an insight into the effectiveness of such a policy. For this purpose, consider Figure 5, where we focus on three different product categories: European luxury products, home décor products, and men's products. The horizontal coordinate records, for each category, the fraction of initial inventory sold for a particular SKU–event pair, referred to a "sell through." The vertical coordinate provides information on the frequency that this level of sell-through occurs for each product category.

For example, in almost 70% of the European luxury SKU–event combinations, the entire initial inventory was sold. This suggests that the price applied in these events may have been too low. In contrast, in about 55% of the home décor SKU–event combinations, none of the initial inventory was sold. This suggests that the price applied in these events may have been just too high.

The question, of course, is whether the retailer has any data that may help improve the retailer pricing policy. We worked directly with the retailer to obtain and understand point of sales data, product information, and customer behavior, and we used all of these data along with machine learning techniques to improve demand forecast accuracy. To translate a demand forecast into a pricing policy, we developed new theory around assortment price optimization and created and implemented a pricing recommendation tool that has shown significant margin improvement.

The access to detailed data allowed us to further explore and identify additional business improvement opportunities and research ideas. For example, we identified trends in the data that suggest that there exists an optimal range of the number of items for the retailer to include in an assortment; *displaying too few items or too many items seems to have a negative impact on sales per item*. We plan on further exploring this observation and will hopefully make a contribution to the assortment planning field.

Another observation suggested by the data is illustrated in Figure 6, where we show sales of one SKU

Figure 5 **Challenges with a Single Price Strategy**

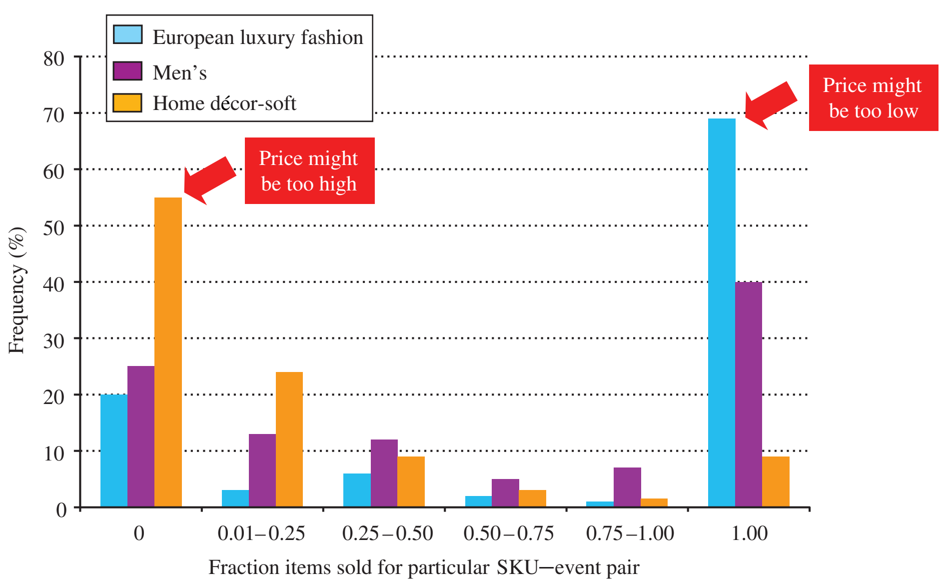
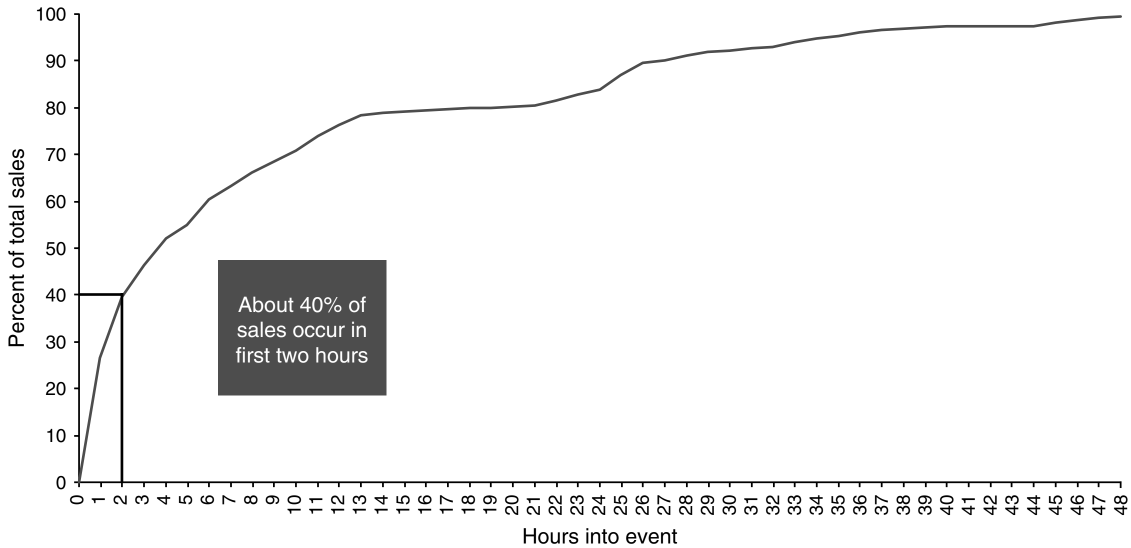


Figure 6 Demand Pattern with a Single Price



during a 48-hour event. Surprisingly, the data show that about 40% of sales occurred in the first two hours of the event; we found that this was typical for most events. Thus, the first few minutes of an event offer an insight into the customer demand function by observing these initial sales. This suggests that once an event starts, the retailer faces an exploration–exploitation trade-off between learning to gather more information about the demand function versus optimizing price to maximize revenue.

To help in the learning process, we use historical data to estimate a set of benchmark demand functions; we do not know which one correctly represents demand for the specific style sold in the current event, but historical data suggest that one of those benchmark demand functions is likely to be the right one.

We developed a learning and optimization method under an important business constraint that the retailer identified: apply only a few different prices during the learning period. The question is, can such a method be effective in practice? More importantly, can a learning and optimization method be effective with a single learning price? In this case, the retailer applies a single price during the learning period and then observes customers' buy/no buy decisions. Once the learning period is over, the retailer switches to an optimized price.

Our analysis shows that when n customers log on to an event, expected loss under a single price strategy, relative to an oracle who knows the true demand function, is $\mathbf{O(n)}$. On the other hand, with a single learning price, expected loss drops dramatically to $\mathbf{O(\log n)}$. Finally, we prove that if the retailer changes the price k times during the learning period, expected loss is proportional to $\mathbf{O(\log\log \cdots \log n)}$, iterated k times. Numerical experiments suggest that the single price learning strategy can increase expected revenue by 6%–7%.

So, this story is about merging theory and practice. It started with data from our retailer and their interest in using data to improve their pricing process. We brought together statistics, machine learning, and OR techniques to not only deliver an effective pricing optimization tool, but also identify opportunities and develop a new mechanism for learning and optimization that requires very few price switches.

4. Let the Data Drive the Model

Throughout my career, I have focused on research that is tied to both theory and practice; the three examples that I shared here are no different. Although both problem-driven research and data-driven research are practice oriented, the main difference lies within how the research is initiated. In problem-driven research, an academic or industry professional identifies a problem and uses models and data to develop insights and possibly make improvements in practice. In data-driven research, data are gathered from an organization before any specific model is developed; it is the academic's careful analysis of the data that sheds light on possible opportunities to make improvements.

In general, there are two types of data-driven research. In the first, the focus is on a specific goal—increase revenue, decrease cost, reduce the spread of an epidemic—and the challenge is to let the data identify the specific issues, opportunities, and models that the organization should focus on. In what follows, I will refer to this type as data-driven models (DDMs). The second is an open-ended search for correlations and relationships without any clear goal in mind. This is typically the objective of data mining: uncover economic or other relationships by analyzing a huge mass of data.

The work for the online retailer covered in §3 belongs to the DDM category of data-driven research, where the data drives models. The goal was to improve revenue, but it was not clear how to achieve that objective. It was the data that suggested that the retailer had an opportunity to learn about consumer behavior in the first few minutes of an event and then optimize pricing decisions. This observation led to the development of a learning and optimization model. Equally important, it was the data that suggested the need to carefully choose the number of items in an assortment and hence called for a model for assortment optimization.

The sort of data-driven research I envision for our community—whether it is the OR community in general or just the OM community—is of the DDM type because it is clearly linked with decision making. It allows our community to distinguish itself from those in economics and statistics who apply data mining and focus mostly on the second type of data-driven research. Indeed, the DDM type of data-driven research described above is a very good fit with our community because of our unique set of skills and tools that enable the development of models to improve decision making.

It should be apparent by now that the so-called empirical research done by some in our community can be quite different from my vision for data-driven research. Indeed, when the empirical research just identifies correlations and relationships, then it is no different from what economists and statisticians do when applying data mining techniques. However, if the empirical research drives a new model that assists decision making, then it fits the DDM type of data-driven research.

My sense is that today, being data driven in our community is often considered suspect. At the same time, there is an exaggerated adherence to the notion that theory and hypotheses should be formulated first before even looking at any data. Indeed, for classical statistics and hypothesis testing to work, one has to take this approach. But, the early history of OR was highly data driven, where repeated observations about a specific system led to models and decisions. The reader is referred to the book *Methods of Operations Research* by Philip M. Morse and George E. Kimball, published in 1951 (Morse and Kimball 1951), where the authors describe the emergence of the field of OR during World War II.

Unfortunately, over the years, our community has drifted away from being data driven, where models and tools are developed after, not before, data are analyzed; see also Corbett and Van Wassenhove (1993). The challenge to the OM community today is to reorient itself and emphasize DDMs in research and teaching. This is important to the health of our field not only because organizations are changing the way they capture data, analyze information, and make decisions, but also because we analyze systems that involve people. Such systems are difficult to understand unless we have access to behavioral data. Sadly, there is very little reliance on data in identifying new research opportunities, developing creative models, or teaching innovative concepts.

Data-driven research is our opportunity to be more creative and play an important role in the development of what some call "business analytics." It will foster the development of new engineering and scientific methods that explain, predict, and hopefully change behavior. To do that effectively, we need an open source data repository that allows scholars to let the data tell a story, be creative, and develop and test models and solution methods. This is of course a tall order, but without such an open source repository, it will be difficult for young scholars to be involved.

Acknowledgments

The author expresses special thanks to his Ph.D. students, some of whom collaborated with him on various projects described in this essay. This paper benefited from discussions with Kris Johnson (MIT), Edward Kaplan (Yale University), Richard Larson (MIT), Chung Piaw Teo (National University of Singapore), He Wang (MIT), Yehua Wei (Duke University), and Alexander Weinstein (MIT).

References

Braca J, Bramel J, Posner B, Simchi-Levi D (1997) A computerized approach to the New York City school bus routing problem. *IIE Trans.* 29(8):693–702.

Bramel J, Simchi-Levi D (1995) A location based heuristic for general routing problems. *Oper. Res.* 43(4):649–660.

Bramel J, Simchi-Levi D (1996) Probabilistic analyses and practical algorithms for the vehicle routing problem with time windows. *Oper. Res.* 44(3):501–509.

Chou MC, Chua GA, Teo C-P, Zheng H (2010) Design for process flexibility: Efficiency of the long chain and sparse structure. *Oper. Res.* 58(1):43–58.

Corbett CJ, Van Wassenhove LN (1993) The natural drift: What happened to operations research? *Oper. Res.* 41(4):625–640.

Graves SC (2008) Flexibility principles. Chhajed D, Lowe TJ, eds. *Building Intuition: Insights from Basic Operations Management Models and Principles*, Chap. 3 (Springer, New York), 33–51.

Haimovich M, Rinnooy Kan AHG (1985) Bounds and heuristics for capacitated routing problems. *Math. Oper. Res.* 10(4):527–542.

Hopp WJ, Tekin E, Van Oyen MP (2004) Benefits of skill chaining in serial production lines with cross-trained workers. *Management Sci.* 50(1):83–98.

Johnson K, Lee A, Simchi-Levi D (2013). Analytics in online retailing: Demand forecasting and price optimization. Working paper, MIT, Cambridge, MA.

Jordan WC, Graves SC (1995) Principles on the benefits of manufacturing process flexibility. *Management Sci.* 41(4):577–594.

Karp RM (1977) Probabilistic analysis of partitioning algorithms for the traveling salesman problem in the plane. *Math. Oper. Res.* 2(3):209–224.

Larson RC (1990) The queue inference engine: Deducing queue statistics from transactional data. *Management Sci.* 36(5): 586–601.

Larson RC (2013) Private email to the author on September 22.

Morse PM, Kimball GE (1951) *Methods of Operations Research* (Technology Press of MIT, Cambridge, MA). [Reprinted in 2003 by Dover Publications, Mineola, NY.]

Simchi-Levi D (2010) *Operations Rules: Delivering Customer Value Through Flexible Operations* (MIT Press, Cambridge, MA).

Simchi-Levi D, Wei Y (2012) Understanding the performance of the long chain and sparse designs in process flexibility. *Oper. Res.* 60(5):1125–1141.

Simchi-Levi D, Chen X, Bramel J (2013) *The Logic of Logistics: Theory, Algorithms and Applications for Logistics Management*, 3rd ed. (Springer-Verlag, New York).

Simchi-Levi D, Schmidt W, Wei Y (2014) From superstorms to factory fires: Managing unpredictable supply chain disruptions. *Harvard Bus. Rev.* (January–February).

Simchi-Levi D, Wang H, Wei Y (2013) Increasing supply chain robustness through process flexibility and strategic inventory. Working paper, MIT, Cambridge, MA.

Simchi-Levi D, Wang H, Weinstein AM (2013) Dynamic pricing and demand learning with limited price experimentation. Working paper, MIT, Cambridge, MA.