Rocket Science Retailing and the Research Opportunities it Creates for Us

Marshall Fisher
The Wharton School

"The Role of Optimization in Supply Chain Management"
Institute for Mathematics and its Applications (IMA)
September 23-27, 2002
Harvard/Wharton Project: How can recent advances in Information Technology improve the way retailers forecast demand and plan supplies?

<table>
<thead>
<tr>
<th>Apparel, Footwear</th>
<th>Consumer Electronics and PCs</th>
<th>Books, CDs, Jewelry, Toys Theme Stores</th>
<th>Other Product Categories and Multiple Product Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>David’s Bridal</td>
<td>CompUSA</td>
<td>Borders Group</td>
<td>Ahold</td>
</tr>
<tr>
<td>Footstar</td>
<td>Office Depot</td>
<td>Bulgari</td>
<td>Christmas Tree Shops</td>
</tr>
<tr>
<td>Gap Inc.</td>
<td>Radio Shack</td>
<td>The Disney Store</td>
<td>CVS</td>
</tr>
<tr>
<td>GH Bass</td>
<td>Staples</td>
<td>Tiffany &amp; Co.</td>
<td>Federated</td>
</tr>
<tr>
<td>Maurices</td>
<td>the good guys!</td>
<td>TransWorld Enter</td>
<td>HE Butt</td>
</tr>
<tr>
<td>Nine West</td>
<td>Tweeter etc</td>
<td>Warner Bros.</td>
<td>Iceland Frozen Foods</td>
</tr>
<tr>
<td>The Limited World</td>
<td></td>
<td>Zany Brainy</td>
<td>JC Penney</td>
</tr>
<tr>
<td>Zara</td>
<td></td>
<td></td>
<td>Marks &amp; Spencer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>QVC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sears</td>
</tr>
</tbody>
</table>

Agenda

• Overview of retail operations – what do retailers do?

• State of retail supply chain management today – findings from the Harvard/Wharton Rocket Science Retailing Project

• ‘Rocket science’ retailing – injecting high grade analytics into retailing

• Some steps toward the vision

• Suggested research opportunities
<table>
<thead>
<tr>
<th>Research collaborators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gerard Cachon</td>
</tr>
<tr>
<td>Nicole DeHoratius</td>
</tr>
<tr>
<td>Karen Donohue</td>
</tr>
<tr>
<td>Vishal Gaur</td>
</tr>
<tr>
<td>Gurhan Kok</td>
</tr>
<tr>
<td>Anna McClelland</td>
</tr>
<tr>
<td>Rocco Mosconi</td>
</tr>
<tr>
<td>Kumar Rajaram</td>
</tr>
<tr>
<td>Walter Salmon</td>
</tr>
<tr>
<td>Zeynep Ton</td>
</tr>
<tr>
<td>Noel Watson</td>
</tr>
<tr>
<td>Giulio Zotteri</td>
</tr>
</tbody>
</table>
Retail operations

Goal = Max Profit = revenue – variable cost (cost of goods, inventory holding and obsolescence)

New Product Development/Selection

Raw Materials/Components Sourcing

Buyer/Planner
- Forecast Demand
- Set Inventory levels
- Pricing & Promotion
- Lead time management e.g. when to use air freight

Buyer & Designer
- Assortment planning

Store Manager
- Execution e.g. product display, customer service
- POS price changes
- Accuracy of inventory data

State of retail supply chain management today

Too many of the wrong products

Department store markdowns as a percentage of sales

Too few of the right products

“One third of customers entering a store leave without buying because they can’t find what they came for”
Two knit tops appeared well bought at the chain style color level ...

<table>
<thead>
<tr>
<th></th>
<th>Blue Hooded Top</th>
<th>Print Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bought</td>
<td>12,200</td>
<td>12,200</td>
</tr>
<tr>
<td>Sold</td>
<td>12,047</td>
<td>11,844</td>
</tr>
<tr>
<td>Average Price</td>
<td>$16.75</td>
<td>$16.75</td>
</tr>
<tr>
<td>Cost</td>
<td>$8.45</td>
<td>$8.45</td>
</tr>
<tr>
<td>Gross Margin</td>
<td>$99,990</td>
<td>$98,305</td>
</tr>
</tbody>
</table>

But at the style/color/size/store level, the blue hooded top was seriously under-bought, especially in the largest size
And SKU-door level analysis reveals huge lost margin

<table>
<thead>
<tr>
<th></th>
<th>Blue Hooded Top</th>
<th>Print Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bought</td>
<td>12,200</td>
<td>12,200</td>
</tr>
<tr>
<td>Sold</td>
<td>12,047</td>
<td>11,844</td>
</tr>
<tr>
<td>Average Price</td>
<td>$16.75</td>
<td>$16.75</td>
</tr>
<tr>
<td>Cost</td>
<td>$8.45</td>
<td>$8.45</td>
</tr>
<tr>
<td>Gross Margin</td>
<td>$99,990</td>
<td>$98,305</td>
</tr>
<tr>
<td>Estimated Lost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit Sales</td>
<td>171,919</td>
<td>5,111</td>
</tr>
<tr>
<td>Estimated Lost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lost Margin</td>
<td>$1,611,225</td>
<td>$42,859</td>
</tr>
</tbody>
</table>

What’s causing these problems?

Long lead times

Up to 11 months for apparel products sourced from Asia, but some – World and Zara – achieved 2 week lead times

Inaccurate forecasts

50% -100% errors for a season forecast at chain level is typical
Preseason forecasts are highly inaccurate

(Each Blue Dot Corresponds to a Particular Style/Color of Women’s Dress for a Major Cataloger)

Average Error is 55%
<table>
<thead>
<tr>
<th>What’s causing these problems?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long lead times</td>
</tr>
<tr>
<td>Inaccurate forecasts</td>
</tr>
<tr>
<td>Store level execution cannot be assumed</td>
</tr>
</tbody>
</table>
Albert Heijn example – store sales vs customer entrances reveals problems

1 unit left over - looks like 100% service, but possible shrink results in lower service

Incorrect Refill from backroom

Graph showing time on the x-axis and customers and sales on the y-axis.
Retailers have data that can help, but don’t use it

Scanners and data warehousing has created huge databases, but retailers lack ability to analyze this data

“How can history possibly be useful in a fashion business?”

“We are awash in data and starved for information”

History -> seasonality, price elasticity, store/sku patterns, size mix, reaction to promotion, etc.
Retailers have data that can help, but don’t use it

Inadequate tools - Planning paradigms are replenishment of staples and ‘one and done’ category planning for fashion products

Reality is many products like shoes, home fabrics, books & music, toys, etc have features of staples and fashion – limited life but can replenish & items matter

• Need to manage at the SKU store level

But can’t do this without analytic tools – too much data for a manual approach
Too many decisions to actively manage

Data for median retailer in our survey

- $1.1 billion annual sales
- 2.3 million store/sku stock points e.g. 230 stores x 10,000 sku’s
- 45% gross margin

--- $4.14 gross margin per week per stock point

Only way to actively manage this is by computer
The payoff from fixing this can be huge

Better analytics can easily double a retailer's profit

Data + analytic tools

= better in-stock + fewer markdowns

= profit increase of 5 – 10% of revenue
10% higher in stock leading to 10% more sales doubles profit

<table>
<thead>
<tr>
<th>Gross Margin</th>
<th>Profit before tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jewelry Stores</td>
<td>52.4%</td>
</tr>
<tr>
<td>Consumer Electronics and Computer Stores</td>
<td>33.1%</td>
</tr>
<tr>
<td>Apparel and Accessory Stores</td>
<td>36.1%</td>
</tr>
<tr>
<td>Department Stores</td>
<td>32.2%</td>
</tr>
</tbody>
</table>
Retail operations

Goal = Max Profit = revenue – variable cost (cost of goods, inventory holding and obsolescence)

New Product Development/Selection

Raw Materials/Components Sourcing

Plants

DC

Buyer & Designer
• Assortment planning

Buyer/Planner
• Forecast Demand
• Set Inventory levels
• Pricing & Promotion
• Lead time management e.g. when to use air freight

Store Manager
• Execution e.g. product display, customer service
• POS price changes
• Accuracy of inventory data

Stores
Consumers

Scientific retailing

- New Product Development/Selection
- Raw Materials/Components Sourcing
- Plants
- DC
- Stores
- Consumers
- Sales & Inventory by SKU/Store
- Loyalty Card Information
- Data Warehouse for model calibration - seasonality, price elasticity, size mix, store differences
- Data Accuracy
- Estimate Demand Density
- Track Data Accuracy
- Determine Optimal Inventory, Price and Promotion over Time for Season
Product Life Cycle Planning

**Pregnancy**
- Forecast initial demand
  - Survey experts
  - Mall intercepts
  - Prior products
  - Merchandise test
- Initial buy
- Preposition materials

**Youth**
- Read early sales, update forecast and reorder as needed
- Expedite resupply?
- Rollout to more stores or drop stores?

**Middle age**
- Traditional long life cycle replenishment
  - Set target stock levels for each SKU-store

**Retirement**
- Manage down replenishment flow
- Drop product in selected stores
- Markdown price or transfer to outlet stores

**Diagram Notes**
- Merchandise Test
- Read & React Period
- Replenishment / Model Stock Setting
- Phase Down / End of Life / Markdown
## Merchandise Depth Test

### Typical calendar for Fall season

<table>
<thead>
<tr>
<th>Test</th>
<th>Early Read</th>
<th>Primary sales season</th>
<th>Closeouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>May</td>
<td>Aug</td>
<td>Sept</td>
<td>Dec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Jan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Feb</td>
</tr>
</tbody>
</table>

Goal is to predict total chain primary season sales from test sales

Test design issues

How many stores to test in?

Which stores to test in?

How to forecast chain sales for season from test results?
Testing practice

• **Written survey**
  - 25 out of 27 conduct tests
  - Median effectiveness = 6 on 10 point scale -> most retailers test but think they don’t test well

• **Interviews**
  - Retailer Y says “Retailer X has found 2 stores that perfectly predict next season.”
  - Retailer X says “Our tests are inaccurate because of wrong test stores, poor execution, weather …. “
Typical testing process

- Choose 25 average stores e.g. average sales rate
- Place product in those stores for 4 weeks
- Forecast chain season sales as

  \[(\text{stores in chain}/25) \times (\text{weeks in season}/4) \times \text{Total test sales}\]
Predicted vs actual using typical process with 25 test stores at apparel retailer.
New testing process uses store clustering to find representative test stores

- **Cluster stores** based on similarity of sales mix. Use K-median clustering for this.

- Choose most representative **store from each cluster** to be a test store.

- **Fit formula** that most accurately predicts chain season sales from test store sales. Test stores aren’t weighted equally.

- Can also **predict sales of each cluster**, and hence store sales
Test store selection problem

Season sales of store $i$

Difference in sales mix between store $i$ & store $j$

\[
\min \sum_{i \in I} \sum_{j \in I} w_i d_{ij} x_{ij}
\]

Subject to

\[
\sum_{i \in I} x_{ij} = 1, \quad j \in I
\]

\[
\sum_{j \in I} y_j = k
\]

\[0 \leq x_{ij} \leq y_j \leq 1, \quad i, j \in I\]

\[x_{ij} \text{ and } y_j \text{ integral}, \quad i, j \in I\]
Predicted vs actual using new method with 10 test stores at same retailer
## Performance of the new method for four retailers

<table>
<thead>
<tr>
<th>Retailer</th>
<th>New Test Process Forecast Error 10 Test stores</th>
<th>Traditional Test Forecast Error 10 Test stores</th>
<th>Profit increase due to reduced lost sales and markdowns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women’s Apparel</td>
<td>12.9%</td>
<td>41.9%</td>
<td>18% of revenue</td>
</tr>
<tr>
<td>Nine West</td>
<td>17.6%</td>
<td>27.9%</td>
<td>11% of revenue</td>
</tr>
<tr>
<td>Meldisco</td>
<td>12.7%</td>
<td>22.7%</td>
<td></td>
</tr>
<tr>
<td>Home Fabrics</td>
<td>17%</td>
<td>43%</td>
<td></td>
</tr>
</tbody>
</table>
What Drives Mix? Correlation Between Sales Mix and Store Descriptors

<table>
<thead>
<tr>
<th>Store Descriptor</th>
<th>Fashion Apparel Retailer</th>
<th>Nine West</th>
<th>Meldisco</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Temperature</td>
<td>0.7</td>
<td>0.6</td>
<td>0.65</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0.4</td>
<td>0.7</td>
<td>0.75</td>
</tr>
<tr>
<td>Store Type</td>
<td>0.35</td>
<td>0.5</td>
<td>0.52</td>
</tr>
<tr>
<td>Location</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Store Size</td>
<td>0.2</td>
<td>0.3</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Product Life Cycle Planning

**Merchandise Test**
- Read & React Period

**Replenishment / Model Stock Setting**
- Phase Down / End of Life / Markdown

1. **Pregnancy**
   - Forecast initial demand
   - Survey experts
   - Mall intercepts
   - Prior products
   - Merchandise test
   - Initial buy
   - Preposition materials

2. **Youth**
   - Read early sales, update forecast and reorder as needed
   - Expedite resupply?
   - Rollout to more stores or drop stores?

3. **Middle age**
   - Traditional long life cycle replenishment
     - set target stock levels for each SKU/store

4. **Retirement**
   - Manage down replenishment flow
   - Drop product in selected stores
   - Markdown price or transfer to outlet stores
Data from a fashion cataloger – early sales is highly predictive

Expert Forecast by a Committee of Four Merchandisers

Forecast Obtained by Extrapolating the First 2 Weeks (11%) of Orders

Average Forecast Error is 55%

Average Forecast Error is 8%
Reading and reacting to early sales

This portion of demand can be supplied based on an accurate forecast.

Initial Shipment to Cover
Read/React Period +
Safety Stock Based on Error
Margin in the Forecast

Read Market
& Order More of
Hot Sellers

Replenishment
Supply

% of Total
Season
Sales
Achieved

0.0% 10.0 20.0 30.0 40.0 50.0 60.0 70.0 80.0 90.0 100.0

Week of season
Lead Time

0 1000 2000 3000 4000

0 1000 2000 3000 4000

0 1000 2000 3000 4000

### Forecast Committee Results

<table>
<thead>
<tr>
<th>Product</th>
<th>Anna</th>
<th>Laurie</th>
<th>Julie</th>
<th>Kim</th>
<th>Committee Average</th>
<th>Committee Standard Deviation</th>
<th>Actual Demand</th>
<th>Forecast Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navy Turtleneck</td>
<td>89</td>
<td>86</td>
<td>102</td>
<td>102</td>
<td>95</td>
<td>8</td>
<td>85</td>
<td>-10</td>
</tr>
<tr>
<td>Red Cardigan</td>
<td>63</td>
<td>76</td>
<td>152</td>
<td>51</td>
<td>86</td>
<td>46</td>
<td>199</td>
<td>113</td>
</tr>
<tr>
<td>Blue Vest</td>
<td>30</td>
<td>91</td>
<td>183</td>
<td>76</td>
<td>95</td>
<td>64</td>
<td>29</td>
<td>-66</td>
</tr>
</tbody>
</table>

When the Committee Agrees, They Tend to be Accurate
The Committee Process Is A Powerful Way To Determine What You Can And What You Cannot Predict

Standard Deviation of the Individual Forecasts of the Four Person Committee

Average Error = 116 units
Average Error = 645 units

High Agreement
Low Agreement
The Initial Buy Minimizes the Cost of Stockouts and Closeouts

- Initial Buy
- Replenishment Order Placed
- Replenishment Order Arrives
- End of Life

Demand Phases:
- Initial Demand
- Life cycle Demand

Risk Phases:
- Stockout Risk
- Closeout Risk

Probability Distribution:
- Initial Demand
- Life cycle demand

Initial Buy Quantity
Read/react can be used to evaluate the impact of Lead Time reduction.

The graph shows the incremental dollar gross margin as a percentage of the base case revenue over different lead times. The benefits of read/react are evaluated at each lead time point, ranging from 3 weeks to 12 weeks, with a decrease in gross margin from 18% to 3.5%.
Nine West Example

**Objective** To improve demand forecast using early sales data and decide mid-season replenishment order quantities

![Diagram showing supply and order timing]

- **Initial Supply in Stores**
- **Replenishment Orders**
- **Order Arrivals**

Data from a fashion cataloger – early sales is highly predictive

Expert Forecast by a Committee of Four Merchandisers

Forecast Obtained by Extrapolating the First 2 Weeks (11%) of Orders

Average Forecast Error is 55%

Average Forecast Error is 8%
Simple Extrapolation of Early Sales Didn’t Work for

Forecast of season sales = first 8 weeks sales / .107

34% average error
Because price markdowns and stockouts were affecting sales
Needed forecast model that considers impact of price and supply

Forecast of sales of SKU j in week t

\[ S_j(t) = k_j \cdot s(t) \cdot \min \left( \left( \frac{p(t)}{p_j(t)} \right)^\beta, 1 \right) \cdot \max \left( \left( \frac{p(t)}{p_j(t)} \right)^\mu, 1 \right) \cdot \min \left( \frac{I_j(t)}{I_0}, 1 \right) \cdot \left( 1 - \gamma \cdot \max \left( 0, 1 - \frac{I_j(t)}{I_0} \right) \right) \]
And then we got accurate forecasts.
Weekly forecast vs actual for one sandal
Application of this approach to one style-color

<table>
<thead>
<tr>
<th></th>
<th>Original plan</th>
<th>Optimized plan based on our work</th>
<th>Actual results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Receipts</td>
<td>6,210</td>
<td>14,160</td>
<td>12,325</td>
</tr>
<tr>
<td>Unit sales</td>
<td>9,882</td>
<td>14,875</td>
<td>14,579</td>
</tr>
<tr>
<td>Gross Margin $</td>
<td>$385,497</td>
<td>$580,274</td>
<td>$562,269</td>
</tr>
</tbody>
</table>
Product Lifecycle Planning

4R Supply Planning Product Suite

Different Challenges require Different Algorithms

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Product Intro</td>
<td>Choosing representative stores for entire chain to test product and extrapolate a forecast</td>
</tr>
<tr>
<td>Initial Buy</td>
<td>Determining the optimal buy quantity, risk adjusted for stockouts and closeouts</td>
</tr>
<tr>
<td>Read React</td>
<td>Correctly interpret early sales to make an intelligent replenishment purchase</td>
</tr>
<tr>
<td>Replenishment</td>
<td>Accurately set stock levels at both the stores and the DC’s</td>
</tr>
<tr>
<td>End of Life</td>
<td>Find the optimal price and conditions to liquidate inventory</td>
</tr>
</tbody>
</table>
4R System Clients – Partial List

KENNETH COLE new york

LINENS-N-THINGS

BMG ENTERTAINMENT

Alitalia

BEST BUY

AMERICAN PACIFIC

NORDSTROM

BVLGARI

Retailer/Supplier Collaboration via Common 4R Planning Platform

4R Supply Planner

Sales & inventory data, supply chain parameters

Consumer forecast, DC and store inventory levels

Discussions

POS data & Orders

Consumer forecast, DC inventory levels, Asia buys

Supply chain parameters

Linens N' Things

DC

Buyers

Account planners

American Pacific

DC

Asian Suppliers

Stores

Stores

Stores

Product Flow

Information Flow

4R Systems

Retailer/Supplier Collaboration via Common 4R Planning Platform

Sales & inventory data, supply chain parameters

Consumer forecast, DC and store inventory levels

Discussions

POS data & Orders

Consumer forecast, DC inventory levels, Asia buys

Supply chain parameters

Linens N' Things

DC

Buyers

Account planners

American Pacific

DC

Asian Suppliers

Stores

Stores

Stores

Product Flow

Information Flow

4R Systems
On Hand inventory prior to 4R

OH-WOS

Lost sales prior to 4R

Lost Sales

-2.0%
-1.0%
0.0%
1.0%
2.0%
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%


4R
Cont
Dif.
Lost sales for 4R SKUs decreased from 6.6% to 2.2%
Weeks of supply relative to Control SKUs decreased 7%
Summary – the profit increase from retail analytics is large

- Merchandise Test
- Read & React Period
- Replenishment / Model Stock Setting
- Phase Down / End of Life / Markdown

**Pregnancy**
- Survey experts
- Initial buy
- Merchandise test

**Youth**
- Read early sales, update forecast and reorder as needed

**Middle age**
- Traditional long life cycle replenishment
- **3% reduction in lost sales with 7% less inventory from next generation replenishment algorithm**

**Retirement**
- Manage down replenishment flow

Profit increase = 11% - 18% of revenue

Profit increase = 3.5% of revenue

46% gross margin increase
Scientific retailing

New Product Development/Selection → Raw Materials/Components Sourcing → Plants → DC → Stores

Determine Optimal Inventory, Price and Promotion over Time for Season

Data Warehouse for model calibration - seasonality, price elasticity, size mix, store differences

Estimate Demand Density

Track Data Accuracy

• Sales & Inventory by SKU/Store
• Loyalty Card Information

Consumers
Research Opportunities in Retailing

• Enhancing traditional forecasting/inventory management for a real context
  – Short life cycle products
  – Correcting for lost sales due to stockouts
  – Measuring micro consumer response to stockout – substitute, lose sale or lose customer
  – Assortment planning
  – Forecasting product A based on sales of B, C, ...
  – Full optimization of all levers – inventory, price, advertising, store placement, etc.
  – Probability density estimation – forecasting extremal events with sparse data

• Retail experimentation – merchandise testing

• Store as factory – adapting TQM and TPS to store operations
  – Improving data accuracy
  – Avoiding phantom stockouts
  – Empirical studies to discover best practice like the MIT auto study
Real demand differs from textbook image

Textbook Image of Finished Good Demand

Real Demand Patterns

Research Opportunities in Retailing

- Enhancing traditional forecasting/inventory management for a real context
  - Short life cycle products
  - Assortment planning
  - Correcting for lost sales due to stockouts
  - Measuring micro consumer response to stockout – substitute, lose sale or lose customer. Combine marketing’s consumer view with operations supply chain view
  - Full optimization of all levers – inventory, price, advertising, store placement, etc.
  - Probability density estimation – forecasting extremal events with sparse data
  - Data aggregation/disaggregation

- Retail experimentation – merchandise testing

- Store as factory – adapting TQM and TPS to store operations
  - Improving data accuracy
  - Avoiding phantom stockouts
  - Empirical studies to discover best practice like the the MIT auto study