The Impact of SKU and Network Complexity on Inventory Levels

By Joseph McCord and David Novoa Garnica
Thesis Advisor: Dr. Bruce Arntzen

Topic Area: Inventory, Simulation

Summary: This research investigates whether observed inventory levels at a global consumer packaged goods firm are driven by two specific forms of supply chain complexity: the numbers of stock-keeping units (SKUs) in a given brand and the number of stocking locations for a given SKU. To determine the strength of this relationship we applied ordinary least squares regression and a simulation exercise. The research found that these two forms of complexity have little to no correlation with inventory levels, potentially because inventory levels appear to be set by non-forecast error-based heuristics.

**Introduction**

Firms that buy, produce, or sell physical goods hold inventory for several reasons. Holding stock allows these firms to meet general demands of customers and to hedge against uncertainties in supply and demand. Ideally, firms would prefer to hold the minimal amount of inventory possible while meeting their service objectives profitably.

However, the general desire to offer more distinct products to consumers means that certain forms of complexity are increasing. If firms better understood the relationship between complexity and required inventory levels they might be in a better position to weigh tradeoffs between benefits of certain forms of complexity – such as diverse product portfolios or short customer order fulfillment lead times – against an attributable increase in inventory holding cost. Any firm which aims to manage its inventory levels carefully should find value in understanding this relationship.

**Data and Methodology**

This research effort focused on the operations of a global consumer packaged goods firm across several regional market clusters. This firm manufactures and sells products across major product categories.

**KEY INSIGHTS**

1. Although it may be counter-intuitive, greater complexity does not translate into higher inventory levels.

2. Inventory quantities mirror patterns associated with simulated inventory management heuristics rather than traditional optimal inventory models.

3. Two potential measurements of complexity in the consumer packaged goods context are the number of SKUs per brand and the number of stocking locations per SKU.
including personal care, hair care, deodorant, and packaged foods.

Given the nature of the consumer packaged goods industry, the research sponsor was particularly motivated to understand the role of two specific forms of complexity: product or SKU complexity – measured here as the number of SKUs in a brand – and network complexity – measured as the number of stocking locations in a market cluster for an SKU.

Based on intuition and existing literature, the working hypothesis for this research was that either form of complexity would increase observed inventory levels by the traditional square-root law. Brands that included more SKUs or SKUs that used more stocking locations would be reducing pooled variance in demand, meaning that optimal safety stocks would rise according to a power function to increases in forecast error. This problem statement is depicted in Figure 1 below.

![Figure 1. Graphic Summary of Problem Statement](image)

Therefore, in order to measure the strength of this relationship within the operations of the research partner, the researchers applied ordinary least-squares regression between complexity and observed inventory levels against a power curve. The regression analysis was achieved by using historical data from several information systems of the research partner:

- An inventory optimization system which records average demand levels by SKU
- Records of actual inventory levels at SKU locations
- Sales history records (to support effective removal of obsolete SKUs from the analysis).

For each iteration of the analysis, the researchers calculated two measurements of complexity and inventory levels of brands and SKUs expressed as days of stock. Each measurement of complexity was then compared to the associated inventory levels within a scatterplot. The researchers used basic spreadsheet software to measure the correlation between complexity and inventory levels against a power curve line of best fit.

This approach was repeated for multiple product categories across two market clusters (geographic regional markets) across multiple points in time to determine the applicability of the results within the firm’s operating environment. Where possible, inventory levels were averaged across several close points in time to reduce the likelihood of results being driven by arbitrary sampling within the products’ review periods.

To complement this regression methodology and help explain results, the research also included a simulation exercise. Using an artificial database of products constructed to have similar properties to the actual datasets, anticipated inventory levels were calculated using several variants of common inventory control rules.

For example, one simulation applied a base stock policy in which safety stocks were calculated assuming that forecast error was solely driven by SKU complexity within brands. Another simulation assumed that inventory levels were not set by demand variance at all, but according to the common “ABC” heuristic. This approach assumes faster moving (higher sales) products can be managed with lower inventory levels while the long tail of slow moving “C” products have a higher target inventory level. The researchers conducted this exercise to compare hypothetical results with actual observed scatterplots from the regression analyses.

**Results**

Initial summary analytics of the categories analyzed showed a common “long tail” of slow moving SKUs when demand per SKU was observed. For one cluster, approximately 10% of the 4,000+ SKUs had daily demand of 500 units or more, while more than half of the SKUs had an average demand of 25 units or less per day. This result illustrates one aspect of
the significant complexity within the research firm’s product portfolios.

The distribution of network sizes per SKU varied between market clusters, with some showing a fairly even balance across a range of one to eight stocking locations per SKU and others with an average network size of two and a range of one to eleven locations.

Based on four iterations of the analysis, SKU complexity did not prove to have a correlation with inventory levels. Instead, the scatterplots showed a consistent pattern: a wide variation in inventory levels for brands with fewer SKUs and fairly stable and lower inventory levels for brands with more SKUs. In general this pattern, seen in Figure 2, had the opposite trend to the hypothesized relationship. All scatterplots had an R-squared value of between 0.04 and 0.06 for the linear regression.

Based on four iterations of the analysis, SKU complexity did not prove to have a correlation with inventory levels. Instead, the scatterplots showed a consistent pattern: a wide variation in inventory levels for brands with fewer SKUs and fairly stable and lower inventory levels for brands with more SKUs. In general this pattern, seen in Figure 2, had the opposite trend to the hypothesized relationship. All scatterplots had an R-squared value of between 0.04 and 0.06 for the linear regression.

As expected, when the simulated inventory levels were obtained from a safety stock equation which derived all of its forecast error from variance driven by the number of SKUs in a brand, inventory levels followed a power curve almost exactly. When additional sources of forecast error were introduced, the inventory levels began to drift from the line of best fit, but still displayed some correlation. However, when inventory levels were driven by relative demand and not forecast error, as in the “ABC” method, inventory levels displayed a similar pattern to the observed scatterplots of the analysis on historical data. This observation held true for both SKU and network complexity. Figure 3 shows the scatterplots produced through the simulation exercise.

Conclusions

The overall aim of this research effort was to answer the question of whether SKU or network complexity drives observable inventory levels. Based on ordinary least squares regression analysis using
calculated measures of complexity, the research found no observable changes in days of stock from changes in SKU or network complexity.

The simulation exercise offers some insight into why. While the research partner firm maintains an inventory optimization software to recommend optimal inventory levels, planning staff do not always adhere to its recommendations. Instead they might be applying simpler methods such as "ABC." The visual similarity between the simulated "ABC" inventory levels and the scatterplots based on historical data indicate that this could be the case. Instead of adhering to inventory targets calculated by an inventory optimization module, staff may prefer to respond to the complexity of managing hundreds of distinct products by applying non-safety stock-based approaches such as the "ABC" method.

Other firms in this industry or any other which manage large numbers of SKUs should consider the actual decision-making of planning staff when looking at the impact of complexity on inventory levels. As this research only looked at the experience of one firm, however, other settings may see different results if they operate under more centralized inventory control processes. Additionally, further research could look into the role of other sources of complexity such as sourcing lead time variance, frequency of product mix changes, and overall variance of demand within brands.