Planogram Optimization in Support of Small Format Retail Inventory Management

by

Miles Kurtz

Submitted to the MIT Sloan School of Management and
Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degrees of
Master of Business Administration
and
Master of Science in Civil and Environmental Engineering
in conjunction with the Leaders for Global Operations program
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2023

Authored by: Miles Kurtz
MIT Sloan School of Management
and
Department of Civil and Environmental Engineering
May 12, 2023

Certified by: Retsef Levi
J. Spencer Standish (1945) Professor of Operations Management,
Thesis Supervisor

Certified by: Franz-Josef Ulm
Professor of Civil and Environmental Engineering,
Thesis Supervisor

Accepted by: Maura Herson
Assistant Dean, MBA Program, MIT Sloan School of Management

Accepted by: Colette L. Heald
Professor of Civil and Environmental Engineering
Graduate Officer

Planogram Optimization in Support of Small Format Retail Inventory Management

by

Miles Kurtz

Submitted to the MIT Sloan School of Management
and
Department of Civil and Environmental Engineering
on May 12, 2023, in partial fulfillment of the
requirements for the degrees of
Master of Business Administration
and
Master of Science in Civil and Environmental Engineering

Abstract

Target is in the midst of building its "stores-as-hubs" capabilities, relying on stores to support in-store shopping and serve as ecommerce fulfillment hubs. To execute this strategy, Target has further expanded its footprint into urban and dense suburban geographies. The stores in these areas, referred to as Small Format stores, have less than half of the square footage compared to a traditional Target location and carry an order of magnitude less SKUs. The dynamics of Target's urban retailing, which are characterized for the first time in this study, require specific inventory strategies to maintain service levels with a smaller product assortment and fewer customer choices.

One metric to measure inventory management is 'Fit', which considers an item's risk of generating backroom inventory in stores and the days of expected demand covered. Excess inventory decreases worker productivity, while insufficient inventory is associated with stockouts and lost sales. A mixed-integer linear program is developed to suggest the optimal shelf capacity for each product to maximize Fit. The decision model suggests sacrificing space allocated to high cube items to display more units of smaller items, and provides strong evidence for localizing Small Format assortments. A pilot of 10 test display units (planograms) was set and the effects measured via Synthetic Control Design (SCD). This research is part of a multi-year partnership between Target and MIT and is the first implementation of an in-store intervention.

Thesis Supervisor: Retsef Levi
Title: J. Spencer Standish (1945) Professor of Operations Management at MIT Sloan

Thesis Supervisor: Franz-Josef Ulm
Title: Professor of Civil and Environmental Engineering

Acknowledgments

My name may be at the top of this Thesis, but the work was made possible by a massive network of friends, family, mentors, and colleagues (these groups are not mutually exclusive).

Professor Retsef Levi, Professor Vivek Farias, and Professor Fraz-Josef Ulm contributed their ideas and expertise to this work. I am grateful for their investment in this research and for supporting me with resources to expand my analytical toolkit. As I developed the scope for this research, I also must thank Professor Sean Willems for helping me develop my own personal and ever-changing objective function for life.

I am coming out of this research a loyal Target customer and feel proud to have worked with a company so dedicated to developing its culture and supporting all team members. To the GSCL team in particular, I have greatly enjoyed working together. Sara Thomas, who ends every meeting by asking, "What do you need from me?" made it possible to take an idea and set it in stores within 6 months. Gretchen McCarthy, who stepped into the Chief Supply Chain and Logistics Officer role during my time at Target, demonstrated how to care for your team as a leader. Gretchen, I aspire to build a culture like you have and am excited for the continuation of your leadership. Tim Hotze, Heath Holtz, Sarah Deuth, and Arthur Valdez all invested time and provided their attention and feedback throughout this project. Daniel Smith (Daniel.Smith1) was a true partner in translating this research to store shelves.

Finally, I have to acknowledge all of the people in my corner. I could not have asked for a better group in the LGO Class of 2023 to go through this experience with. My parents, Sharon and Michael, whose own professional accomplishments are both admirable. My sister, Alyssa, and brother-in-law, Danny, for providing me with a workspace and my first nephew. To baby Asher, who always made sure I had something to buy on my Target runs - whether your first bassinet or a never-ending supply of diapers. Of course, Dr. Wendy Roldan, who will never let me forget that two Masters' degrees is still not a PhD. Thank you for all of your support and for the sacrifices you made so that I could study at MIT.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

At the time of this writing in Spring of 2023, there are 1,948 Target stores in the United States [2]. The proliferation of stores is part of Target's strategy to reach more customers through Target's "stores-as-hubs" strategy. As customers are offered increasing flexibility of how they want to shop, whether brick-and-mortar retail, ecommerce, or Buy Online, Pickup in Store (BOPIS), Target's retail locations fulfill these orders. Increasing accessibility to customers is also a strategic operations decision to allow for lower costs and increased speed of fulfillment. Behind these stores lies Target's distribution network which consists of a variety of facility types and nodes to receive, consolidate, and send out inventory.

As the number of retail stores increases, so too does the variety of customers and geographies served. Target has traditionally been associated with large-format stores in suburban geographies, but has increasingly moved into more urban and dense suburban areas; nothing demonstrates this more than the opening of the Times Square Target location in April 2022.

Target's stores in urban and dense suburban settings, which are referred to as Small Format stores, typically have different patterns of sales and require different inventory management policies compared to more traditional stores. This thesis seeks to characterize the unique aspects of Small Format retailing, particularly at Target, and then apply these insights to inform improved inventory management policies. The research focuses on planogram, or POG, modifications where a POG is a term to refer to the visual layout of products on a store's shelves.

The scope of this thesis was prompted by the special characteristics of Small Format retailing. Firstly, these stores have a smaller footprint for both selling and backroom area. As a result, the sales floor holds a smaller assortment of products and backrooms can quickly become overwhelmed by inventory, creating safety hazards and limiting team member productivity. Across the smaller assortment on the sales floor, demand becomes more concentrated leading to faster turnover of products and increased risk of products stocking out.

In response to these dynamics, this thesis addresses questions such as:

1. How do we strike the right balance between different inventory metric target levels by modifying shelf capacities and planogram design?

2. How can points of inflexibility (bottlenecks) in the planogram design process be mitigated to incorporate more data into decision making?

3. How can the impact of changes to inventory policies be evaluated in Small Format stores?

This research is a continuation of a multi-year relationship between Target and MIT to enhance productivity in stores through the optimization of Target's distribution network. Prior studies have sought to understand the impacts of unit of measure (UOM), lead time, and planogram (POG) design on key inventory metrics. The work is led by Target's Global Supply Chain and Logistics (GSCL) organization with the ultimate goal of minimizing backroom inventory to improve the efficiency of team members working in stores.

This thesis seeks to continue building upon on prior work, while for the first time implementing research recommendations in stores. The levers which will be used to drive change do not fall under the control of GSCL and will require the collaboration and buy-in of other key stakeholders, including the Merchandising and Space, Presentation & Transition (SPT) teams. This work has led to the development of a metric called 'Fit' to quantify how well products are represented on retail shelving, as well as to the development of a decision support tool to optimize Fit for a given basket of items. The results of this tool were translated to actual POG designs and set across two categories (Facial Tissue and Air Care) in five stores, for a total of ten redesigned and deployed POGs.

1.1 Research Overview

Below, the primary components of this research will be briefly outlined.

Small Format Characterization: Characterization of the dynamics of Small Format stores, particularly at Target, and how they are affected by network design, assortment selection, and the end-to-end POG design process.

POG Design Process Flow: In order to fully understand POG design and influence POGs, the specific process of analyzing and designing new POGs was approached from a value chain perspective. The stakeholders and key process steps are noted, with the goal of identifying bottlenecks in the process. The study of Target's POG design process is critical in informing the rest of the research so that practical solutions can be developed which match Target's capabilities and culture.

POG Optimization: Decision modeling of the recommended capacity for each product on a shelf in stores to optimize the in-store shopping experience and support inventory management objectives. A Mixed Integer Linear Program (MILP) was developed to incorporate the constraints identified during the POG Design Process Flow and maximize Fit on shelves.

Synthetic Control Design: The results from the above POG optimization were implemented in five stores in the New York City market, at two categories in each store (Facial Tissue and Air Care). Synthetic Control Design (SCD) is a method which compares the results of an intervention in a particular population to the expected results in the same population, had no intervention taken place. This is a powerful analytical tool which can both evaluate the real-world effects of POG optimization, and find a larger place within Target's strategic toolkit.

Results

The results of the optimization model suggest Fit can be maximized by reducing the number of large items on shelves in order to fit more total units (for example, replacing bulk packs of tissues with individual boxes). Additionally, the benefits of selecting an assortment local to each store became more visible. Quantitatively, the in-store implementation of these insights did not show a statistically significant change in unit sales; however, this outcome was expected due to the limited scope of the pilot compared to the natural unexplainable variance in sales.

1.2 Thesis Overview

Chapter 2 provides sufficient background on Target and their key metrics to facilitate understanding of this research. After setting a baseline understanding, Chapter 3 reviews relevant literature related to planogram optimization, assortment selection, and retailing, particularly in Small Format environments. Chapters 4 characterizes the Small Format retailing dynamics and idiosyncrasies at Target for the first time in the MIT research partnership. After characterizing the network, Chapter 5 will introduce an optimization model to maximize 'Fit' on planograms and present the results of the model, which were set in a collection of five stores. Chapter 6 will evaluate the results of the in store experiment (based on the optimization model) by applying a Synthetic Control Design methodology. The final chapter, Chapter 7, summarizes the research discussions as well as future research areas which would deepen the impact of this work at Target.

1.3 Data Confidentiality

For the purposes of protecting proprietary information, confidential information used as an input in this research has been disguised. Metrics and data presented have been normalized or presented as percentages unless pulled from public sources.

Chapter 2

Background, Approach, and Metrics

2.1 Background

Target's growth and history as an organization has been predicated upon large format stores in suburban geographies. Their supply chain and distribution network has been organized to support this strategy and retail footprint, but has transformed over the past years to respond to two meaningful changes in the retail environment: (1) Online and omni-channel fulfillment, and (2) Small Format stores in dense, urban environments.

Consistent across these changes are certain design principles of the distribution network, for example, electing to increase complexity upstream in distribution centers rather than in stores. This reflects a commitment to customer service; the time saved upstream in the supply chain improves team member productivity and enables store employees to focus their attention on guests.

2.2 Target Supply Chain Outline

A typical supply chain consists of multiple nodes to receive, sort, and ship products. The nodes in Target's supply chain are specialized for specific types of tasks and products. This is reflected by multiple facility types, including: Import Warehouses (IW), Upstream Distribution Centers (UDC), Fulfillment Centers (FC), Food Distribution

Centers (FDC), Regional Distribution Centers (RDC), and Flow Centers. RDCs and Flow Centers impact stores most directly, as these facilities are typically the origin points for store deliveries.

Products are displayed on planograms (POGs) in stores, which is the term used to refer to the layout of products on shelf displays. The amount of a certain product a POG can hold is referred to as the POG capacity of that product. For the example in Figure 2-1: if a product has 4 facings and a depth of 3, the total POG capacity for that item at a specific store is 12. The depth can also be referred to and considered as the capacity per facing (CpF).

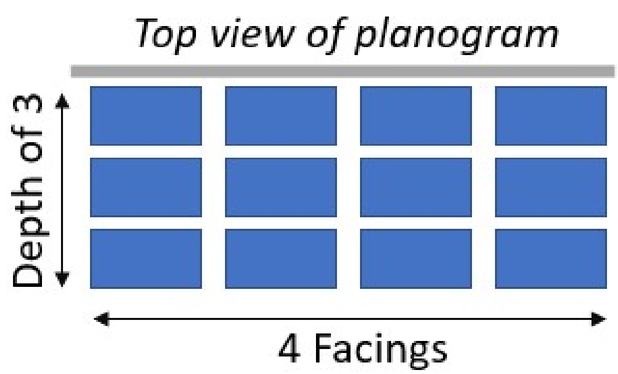

Figure 2-1: Item on POG with 4 facings and a depth of 3, corresponding to a POG capacity of 12.

For last-mile fulfillment to customers, Target is in the midst of executing its "stores-as-hubs" strategy. In this model, the store is viewed as the final node before customer delivery regardless of fulfillment method. This strategy places products closer to customers for easier, less expensive, and faster home delivery. It also has supply chain implications, because shelves are not only shopped by guests, but also by team members packing orders for ecommerce fulfillment.

2.2.1 Backrooms

Within a store, the backroom is the hub of operations and plays a few key roles including receiving inventory, storing inventory for replenishment, and fulfilling ecommerce orders. Given the activities that are ongoing in the backroom, having an excessively full backroom can increase the need for labor in stores and draw attention away from in-store guests. Finally, as Target continues to invest in its stores as hubs strategy, backrooms are used to pack ecommerce orders which are sent to customers.

2.3 Metrics

For the purposes of analyzing planograms through the lens of inventory management, there are metrics utilized by this research and within Target that will be helpful for the reader to understand.

2.3.1 Product Hierarchy

The stock-keeping unit, or SKU, is universally recognized as a unique code that corresponds to an individual product. Retailers have the option of keeping a vendor's SKU or developing different coding. In addition to a SKU, Target developed a unique identifier for each product called a DPCI. The DPCI captures an item's hierarchy by listing its **d**epartment, **p**roduct, **c**lass, and **i**tem.

The most granular way to describe a DPCI is as an item-location combination. Each Target location (or store) has a unique identification number, and an item-location level analysis includes the metrics for a specific product at an individual store. Item-location is the level of granularity for the optimization model in this study.

2.3.2 Backroom Inventory Risk Score

The BRI risk score is a one through five categorization that indicates how likely an item is to be responsible for generating backroom inventory each time it is shipped to a specific store (i.e., the categorization occurs at an item-location level). The risk score is generated by comparing the POG capacity for an item to key inventory metrics, including:

- **Demand Coverage (DC)**: Demand Coverage refers to the expected daily demand at a 98% coverage level (or confidence level) over a time period (typically a sales day).
- **Ideal Beginning on Hand (BOH)**: This is the target inventory level at a location to begin each sales period (day).
- **Order to Level (OTL)**: When an order is triggered, it is placed up to the OTL. The OTL consists of expected demand over the review period and lead time, plus the greater of the presentation minimum or safety stock.
- **High Stock Level (HSL)**: This represents the worst-case high inventory level for a product. It occurs when an order is placed (to the OTL), but no units are sold over the review period and lead time. It can also be represented as the OTL + case pack quantity - 1. In the few stores where case pack quantity is one (due to breakpack processes upstream in the supply chain), the HSL is equivalent to the OTL.

With an understanding of these key inventory metrics, logic can be applied to generate risk scores. A description of each score is below:

1. **Fits (HSL $<$ POG)**: We defined HSL as worst-case high inventory level, so if the POG capacity is always greater than this, there will never be backroom inventory generated by a replenishment shipment.
2. **Low risk of BRI (HSL $>$ POG $>$ OTL)**: If the POG capacity sits above the OTL but below the HSL, there are circumstances where BRI could be generated although the risk is lower.
3. **High risk of BRI (OTL $>$ POG $>$ Ideal BOH)**: This category begins to increase the risk of BRI; if the POG capacity is below the OTL, each time a replenishment shipment arrives there is a significant chance of BRI. Here, BRI would be generated if the forecast for demand over the review plus lead time exceeds what actually materializes.

4. **Never Fits (Ideal BOH > POG > DC)**: If POG capacity is below the targeted BOH, inventory must always be positioned in the backroom.

5. **Never Fits (DC > POG)**: Similarly, if POG capacity does not meet expected daily demand, inventory must always be positioned in the backroom. This is broken out as a separate category from the above because it also impacts store labor; in this case, replenishment trips must be made from the backroom during the day, or lost sales will be incurred.

Based on the above scoring methodology, a smaller risk score is favorable for items; this metric is one of the key inputs into the optimization objective function which will be introduced in Chapter 5. Figure 2-2 visually summarizes the translation of the relationship between POG capacity and inventory metrics to a BRI risk score.

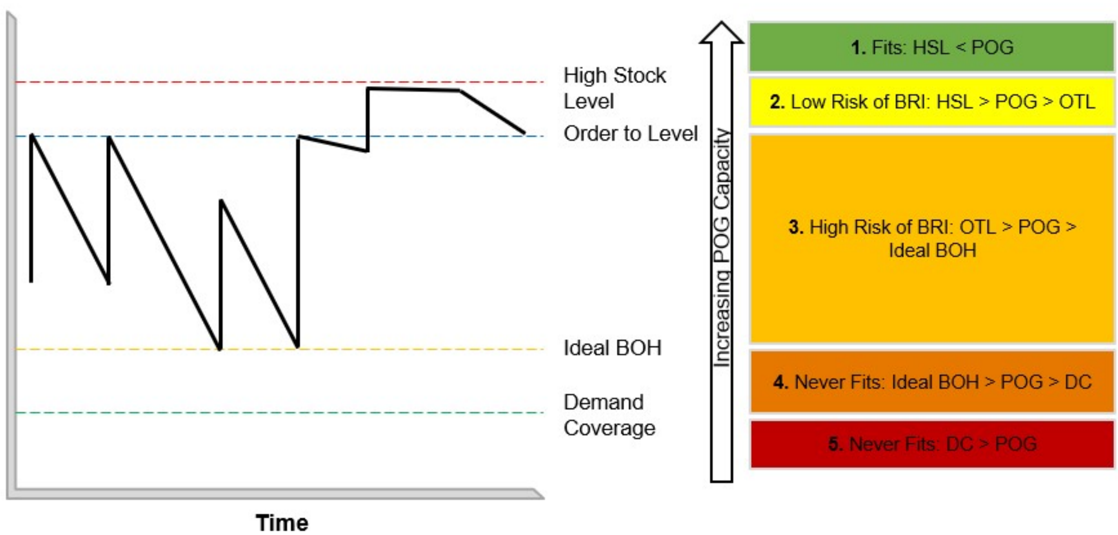

Figure 2-2: The BRI risk score is a rating from one through five that indicates an item's risk of generating BRI by comparing POG capacity to key inventory metrics.

2.3.3 Days of Supply

Days of Supply (DoS) is another metric used internally in Target, and builds upon Demand Coverage. Intuitively, Days of Supply is how many multiples of Demand

Table 2.1: Days of Supply scores

Days of Supply	**Score**
1	5+ Days
2	3-5 Days
3	2-3 Days
4	1-2 Days
5	<1 Day

Coverage fit on a shelf. If POG capacity is 30 units and DC is 10 units, the Days of Supply is consequently 3.

To parallel BRI risk scores, Days of Supply can be rated based on favorability from a score of one through five (Table 2.1). Depending on the category, the targeted Days of Supply can be modified from the above but continue to leverage the same framework. Additionally, sales are not constant by day of week, with lift observed on weekends. An additional target to incorporate into Days of Supply is coverage over an entire weekend (Saturday plus Sunday) of sales.

2.3.4 'Fit'

The last metric to introduce is called 'Fit' and is a term that has gained importance within Target, yet has a broad definition. At its core, this metric is intended to capture how well items fit on a shelf and was initially defined by casepack quantity. Casepack quantity, or how many units are shipped in a package, is often compared to POG capacity. If an item is shipped in casepack quantities of 5, but POG capacity is 12, shipments will be more likely to generate BRI (the least common multiple of those numbers is 60, so almost all shipments will have a remainder). However, the understanding of fit has broadened to consider concepts like assortment, "How well do these items complement one another on a shelf?"

For the purposes of this research, Fit is operationalized as a weighted average of an item's BRI risk score and Days of Supply score. The weighting gives 75% to the BRI risk score and 25% to a DoS score. The weighting was suggested as part of this research and the overarching objective of reducing BRI. Given that facings must be integer values, minor changes in the weighting do not impact results (i.e., even if the recommended facing goes from 1.9 to 2.4, 2 facings will still be recommended). To make the Fit score calculation more clear: an item with a BRI score of 4 and a DoS score of 2 would have a Fit score of 3.5. This metric will inform the objective function in the optimization model introduced in Chapter 5.

In conclusion, the Fit metric addresses the tension of having sufficient inventory in place to provide demand coverage, while not increasing inventory levels to the point where BRI would be generated and decrease store productivity.

2.4 Target Small Format Stores

The first Small Format, or Target Express, store was launched in 2014. These stores have continued to proliferate since then - reaching 203 planned stores in 2023 (Figure 2-3). In 2021, Target opened 32 new stores in total, 28 of which were Small Format, demonstrating the strategic focus on entering these markets.

Compared to a typical 'chain store' these stores have less than half of the average sales floor space ($<$60,000 ft^2 vs. 150,000 ft^2) and an order of magnitude less SKUs (Table 2.2). Given the smaller assortment, meeting local demand (referred to as localization by Target) is critical in these stores and they have become highly segmented. Small Format stores can be classified as either Central Business District, College Campus, Urban Neighborhood, and Dense Suburban. Even within these segments, further distinction is made; for example, College Campus stores could be segmented as 'Standard College' or 'College Food'. The high degree of segmentation is a delicate balance between the benefits of localizing assortment and the effort required to provide individualized attention to smaller subsets of stores; this tradeoff is further discussed in Chapter 4.

This analysis represents the first in-depth investigation into the dynamics of Target's Small Format stores as part of the relationship with MIT. Understanding these dynamics is necessary to build the optimization model and propose enhancements to

Table 2.2: Comparison of key metrics between Small Format and traditional 'Chain' Target locations

	Small Format	**'Chain' Store**
Store Size	<60,000 ft2	150,000 ft2
Number of SKUs	15,000	150,000
Vendor Fill Rate	75%	79%
Shopping Cart	Two-tiered or baskets only	Standard carts and baskets

the operation of these stores. Thus, sharing these learnings is viewed as a foundation of this research as well as future work at Target.

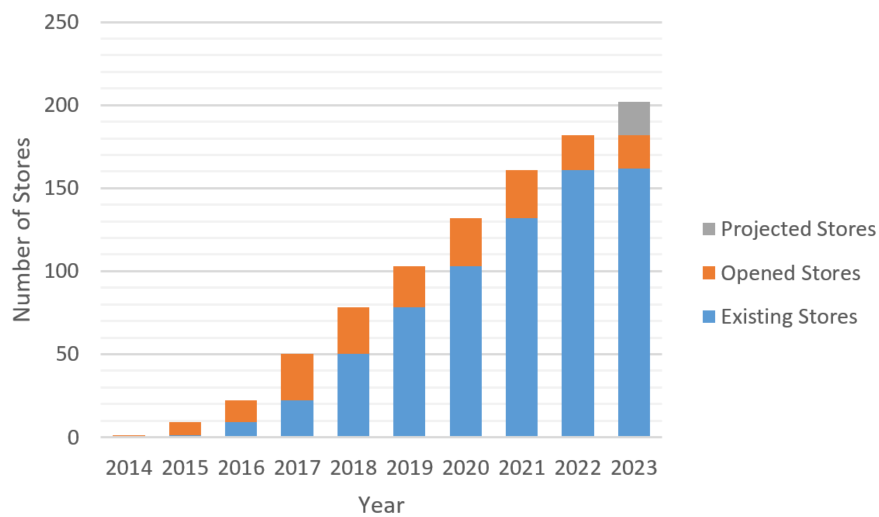

Figure 2-3: Proliferation of Target Small Format stores over time

2.4.1 Inventory Management Considerations

The descriptive statistics in Table 2.2 describe a high velocity sales environment which Target's GSCL team must respond to. With a dramatically reduced store footprint, the number of SKUs in a store is reduced even further to pool demand and focus on providing high service levels of the chosen product assortment. Fewer SKUs, particularly high performing SKUs, reduce opportunities for demand transference in stores; if the product you are seeking is out of stock, it is unlikely a similar product is carried by the store. Furthermore, given the small store footprint and smaller homes in these markets, the average basket size is smaller. This can be interpreted physically by guests using baskets rather than shopping carts, but also financially, with more frequent transactions of smaller dollar amounts. Increased sales velocity puts even more pressure on restocking shelves, maintaining higher on-shelf capacities, and further reducing the number of SKUs in a store. It is a reinforcing cycle which must be addressed through strategic inventory management decisions including assortment selection, planogram design, and by planning for guest dishonesty.

Assortment

Given that a Small Format store has an order of magnitude less SKUs than a chain store (15,000 vs. 150,000), it is expected that the best-performing SKUs are selected for a store based on overall performance and localized demand. Determining localized demand is accomplished by careful merchandising in addition to store leadership pulling high-performing products into their locations. Filling a store with only high performers has the expected effect of a higher sales velocity, and slightly lower vendor fill rates.

Planogram Design

Planogram design for Small Format stores, which will be analyzed in more detail in Chapter 4, follows a process similar to other Target locations. The sales behavior of planograms, however, differs greatly for Small Format stores. Given the constraint on space in the store, the same basket of products will have less shelf capacity in a Small Format store. For example, if one considers a facial tissue POG at a Small Format store, there are 55 facings in total. That same basket of products has an average of 95 facings across regular chain stores (almost double the capacity), plus a larger selection of items which are not present in Small Format stores.

Within a Small Format store, planograms are more likely to be 'flexed' by store leaders. Flexing refers to the practice where a planogram is modified from what was sent to the store - for example, giving more capacity to a certain product. This behavior is often compensating for items which are out of stock, or promoting items which sell well. On a smaller shelf, the impact of a single item being out of stock affects both sales and presentation quality more so than on a larger shelf with more selection. The phenomenon of flexing is a rational response to Small Format dynamics and emphasizes the need to maximize the use of every available linear foot of sale space in a store.

Guest Dishonesty

Small Format stores experience higher rates of guest dishonesty, specifically theft and self-checkout errors. Combined with a goal of minimizing backroom inventory, the effects of theft can be quite pronounced if a store has no product with which to replenish lost items. As a result, in certain stores, team members intentionally maintain a backroom supply of high-theft products by stocking shelves with a lower depth per facing. This process is not codified in replenishment systems, but is a logical response from the store. If a shelf is entirely emptied and there is no product remaining in the store, no sales will be made until the next replenishment delivery.

Secondly, locking POGs have been introduced at a number of stores and a range of categories. These can range from physically locked cabinets to sliders which make a loud noise when moved to select a product from behind them. This affects presentation as building the locking mechanism may take up additional capacity and further reduce the capacity of the POG. Furthermore, sales may be impacted if some proportion of customers do not ask a team member for assistance purchasing an item. At the same time, the OOS for these specific items should decrease, improving product availability. With these opposing forces, the intention is a positive net effect on product sales.

2.4.2 Supporting Supply Chain Dynamics

In response to these dynamics, the supply chain supporting Small Format stores has been carefully designed and operated. A schematic of fulfillment is included in Figure 2-4 and discussed further in the sections below.

Network Design

The establishment of Flow centers has been a key response to high-velocity, Small Format stores. The New York City stores included in this study are fulfilled by the Perth-Amboy Flow Center in New Jersey, which has capabilities to break casepacks

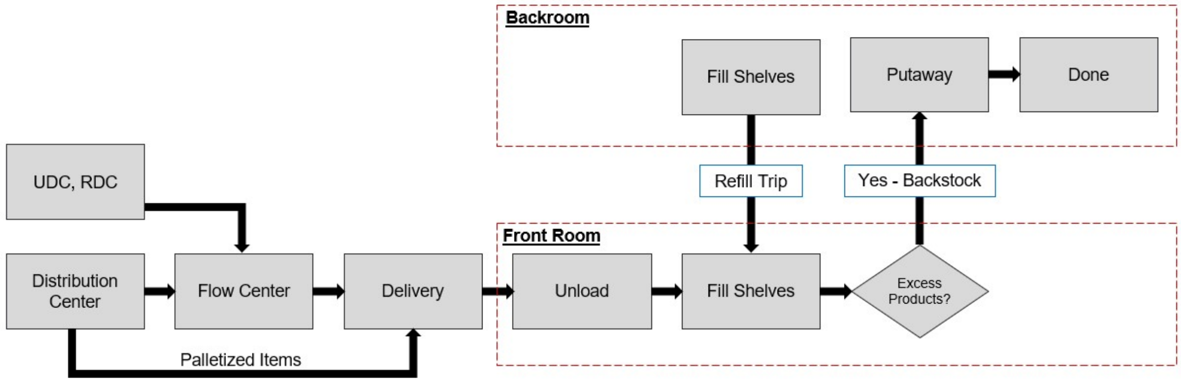

Figure 2-4: Process flow of goods from UDCs and Distribution Centers to Target shelves, with a focus on Small Format stores fulfilled via Flow Centers.

and replenish stores in eaches (i.e., UOM is equal to one). This follows the principle observed to move complexity upstream in the supply chain, where the investment in breaking casepacks reduces complexity in store and concerns of casepack fit. Certain products are not included in the casepack strategy due to safety (e.g., pallets of bottled water, which are heavy) or due to size (e.g., pallets of paper towels). These items arrive at the Flow center from upstream facilities, are cross-docked, and loaded onto trailers alongside items from the Flow Center.

With regards to Target's "stores-as-hubs" strategy, Small Format stores are not included in Ship from Store initiatives. However, Buy Online, Pickup in Stores (BOPIS) is offered at these stores although many do not have parking lots or curbside pickup.

Shelf Replenishment

Trailers arrive at stores for replenishment, though for Small Format stores there is not a consistent loading path to a store's backroom. Some stores have their own loading dock leading to the backroom, while others receive deliveries through the same front doors that customer use.

The time of each delivery varies by store as well; while many Target locations are restocked in the evening and early morning when guests are not in the stores, this is often not the case in the New York market. Depending on the store, restocking can occur during business hours when customers are shopping in the store. Due to higher labor costs and security, team members do not remain in stores overnight in certain store locations.

When trailers do arrive for replenishment, items are loaded on carts which are intended to flow directly to the floor. Each cart corresponds to a specific section of the store per guidelines that the store sets. Ideally, items arrive on the carts and are set on the sales floor 1:1 for items that were sold in the prior sales period. This is measured as a metric called Perfect Flow. For items that do not fit on the floor, they are brought to the backroom and stored. Once an item has inventory in the backroom, the priority is to restock the floor from the backroom before items that are received each day (a FIFO policy). This standard is not always followed in practice and is primarily a concern for perishable items or seasonal products.

Organizational Design

As Target's network design has evolved to meet the needs of Small Format stores, so too has their organizational design. Within larger groups of the Global Supply Chain & Logistics team, there are separate teams focused on operation of Small Format stores. Inventory analysts and category merchants can be assigned solely to Small Format store, learning the unique characteristics and management techniques for these markets. They have support in the field from Small Format operations coordinators who oversee and provide consultative support for a number of stores clustered geographically.

Chapter 3

Literature Review

3.1 Prior Work

The research relationship between Target and MIT has existed since 2019, and this work continues to build on the collaborative partnership. Initial research in 2019 by Das Durgesh considered Replenishment Spillover (or BRI generation) through the lens of POG size, unit of measure (UOM), and replenishment frequency. Durgesh then calculated an optimal UOM, and through simulation concluded that optimizing replenishment frequency alongside UOM could achieve beyond a 40% decrease in BRI in some cases. The results of the BRI root cause analysis (Figure 3-1) justify further pursuit of POG capacity as a lever to reduce BRI levels [3]. In the context of this study of Small Format stores replenished through the Perth Flow Center, the unit of measure is less important as it is equal to one for all stores in scope.

The simulation approach was continued by Lydia Thurman in 2020, who continued studying the effects of modifying UOM and replenishment frequency. Meaningfully, Thurman discovered that configured leadtimes were frequently overstated and that reducing them (to be more accurate) could drive down BRI and at the same time maintain service levels. Secondly, Thurman proved how adjusting UOM could improve both BRI and service metrics [4]; again, due to functionality built into Target's Perth Flow Center, UOM was not a factor considered in this research.

Ben Sidell's work pivoted to classifying items by their risk of generating backroom

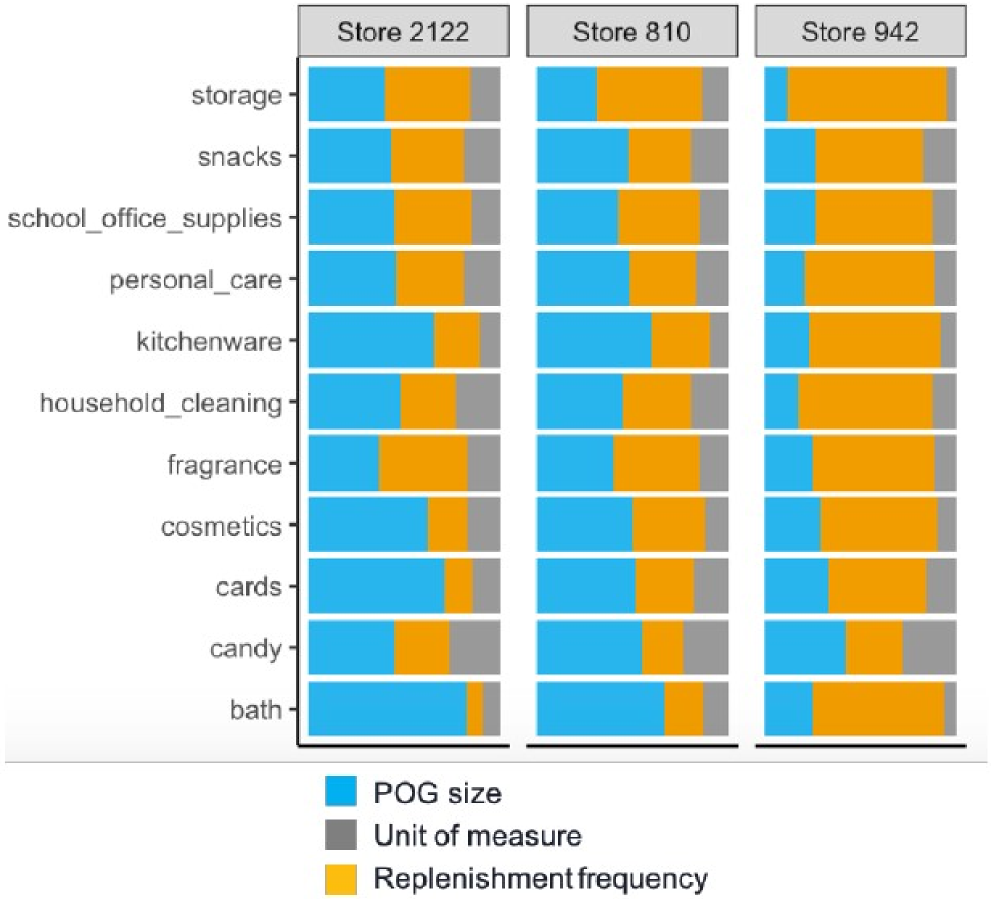

Figure 3-1: Analyzing the root cause of Replenishment Spillover justifies investigating POG capacity as a primary lever.

inventory (which became the BRI risk score introduced in Chapter 2), and suggesting a broad optimization model to modify facings of items to lower their BRI risk score. The results, depicted in Figure 3-2, show the results of the optimization engine in a 'No Worse' scenario, where the score for an individual product cannot become any worse, as well as a 'Fully Optimal' scenario with less constraints. Sidell provides a thorough discussion that notes the limitations of translating this optimization to practice; this optimization applies to an entire class at an individual store, but POGs are built at a more granular level than the class. As Sidell notes, "Realistically the tool would be given a subset of this class, which would only contain a shelf or an aisle. With the reduced space for optimization, the end results could be less optimal. Second, the tool is still working with systematic risk of BRI generation as the primary and only goal." [5]These limitations inspire the research included here, which focuses on optimization at the individual shelf level, and also expands the goal or objective function to include factors beyond risk of BRI generation.

While the circumstances of Target have been well studied through their partnership with MIT, they are not unique to this retailer. In the remainder of this chapter, related studies of inventory management, particularly in urban retailing, will be discussed in the context of their applicability to the research performed in this thesis.

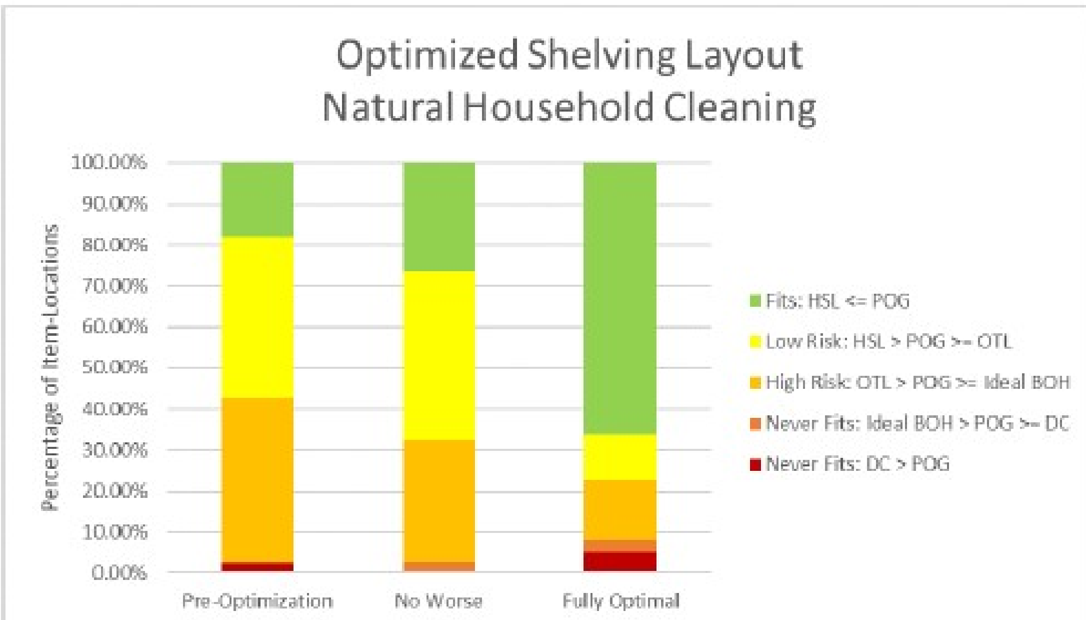

Figure 3-2: Example results from Sidell's research depicting opportunity of an optimized shelving layout for a specific class and store. No worse and fully optimal variant results are depicted to show differences in opportunity.

3.2 Small Format Retailing

Small Format stores can be defined as ones in which a larger retailer offers only a portion of their products in a more compact physical location. Retailers opening such formats have increased as a strategic response to retailers like Toys "R" Us, Sears, and Macy's being forced to close over 100 stores each. Their closings are a direct result of rising ecommerce and a resulting excess of retail space in the United States.

Tarjome reports that Small Format stores are a successful strategy to capture lost sales from nearby stores that close, appeal to highly targeted populations, and be used as pickup and return facilities for online sales.[6] As we will see, this is exactly aligned with the strategy Target has adopted for their Small Format stores.

Broadly, the goal of these compact stores is density, combining high product variety with a small footprint; still, compared to larger stores, these smaller stores sacrifice product variety in order to provide convenience and speed to customers. While written over 20 years ago, Ketzenberg *et al.* demonstrate how excessive inventory levels in dense stores create a barrier to profitability by crowding out other goods. [6] This contradicts common thinking of the asymmetric penalties between erring on the low versus high side of inventory, which has historically kept inventory levels higher to provide better service.

Target is not alone in their pursuit of entering more urban markets. In 2011, WalMart began a pilot of Small Format stores at 12,000-15,000 square feet, but closed 102 of these stores in 2016 after they underperformed [7]. Best Buy has more recently been experimenting with store formats like an Outlet, but also a 5,000 square-foot Small Format store [8]. Similarly, Kohl's announced in 2022 that they will be rolling out 100 Small Format stores over the next four years to enter new markets; however, their new stores will be around 35,000 square feet compared to their typical store at 80,000 square feet [9]. Compared to the other specialty retailers, the Kohl's Small Format stores are more similar in size and scope to Target's stores. Retailers' experimentation around the format demonstrates the challenges associated with Small Format retailing, as well as the potential opportunity. Target's focus on improving operations surrounding their Small Format stores are timely and set them up as a competitive player in these markets.

3.3 Assortment Selection

The most fundamental decision a retailer can make is, "What is it that I want to sell?" Selecting assortment is a careful balance; studies have demonstrated that increasing assortment stimulates sales, while assortment variety is constrained by both physical space and cost. Additionally, increasing assortment can impact the depth of a category (for example, offering more size configurations of a single item leaves less space for other items). These effects are magnified in Small Format stores, where the physical space constraint is even more pronounced.

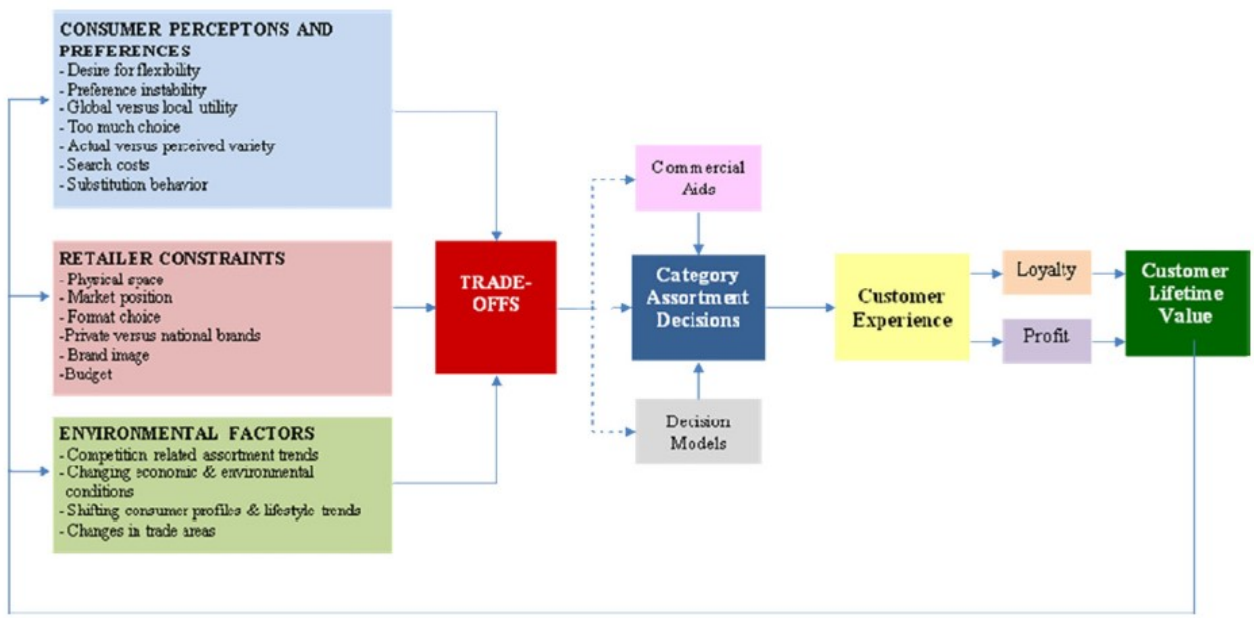

Figure 3-3: Product assortment planning model proposed by Mantrala *et al.* [1]

Other considerations within assortment selection include the relationships items have with one another, and whether they are substitute or complementary products. [10] When a customer shops for a particular product and does not find it, they may either forgo the purchase or find a substitute product. If customers have a high propensity to substitute in a certain category, then providing more depth or high service rates becomes less important. For complementary products, similar effects can be observed. There may be two products which provide higher utility to a consumer together, and in the absence of one product, the other may not be purchased.

Mantrala summarizes the considerations introduced above as tradeoffs in category assortment, balancing consumer preferences with retailer constraints and other market factors. A graphic overview of these factors is included in Figure 3-3.

3.4 Planograms

Given the almost infinite possible combination of products on a shelf, planogram design has traditionally (similar to what we will see in Chapter 4) been done manually with some or little guidance from software. Researchers have advanced heuristics and algorithms to support planogram design, while technology companies such as Apollo and Spaceman have commercialized similar methods and sold them to companies.[11]

Most relevant to this research are heuristics similar to the Knapsack problem to design planograms to maximize profit. The Knapsack problem is a common framework in combinatorial optimization that involves packing a backpack with items of the most value without exceeding a weight constraint. This translates well to shelf optimization because one seeks to maximize the selling potential (or other objective) of items while constrained by space. While this research will focus more on operational objectives like BRI generation and DoS rather than financial metrics, the framework of the model remains constant.

More recent papers, such as one by Hubner and Schaal in 2017, enhance the real-world applicability of POG modeling by including both demand substitution and space elasticity effects. [12] Space elasticity is a studied correlation between shelf capacity and demand stimulation - typically, as space is added for an item on the shelf, sales of that item also increase. In their work, Hubner and Schaal conclude that following their algorithm and accounting for these two effects can increase profits by up to 60%.

Chapter 4

POG Design Process Flow

To use planogram design as the lever to reduce BRI, it is first necessary to understand the end-to-end design process for POGs. The process is described using a process flow or value stream mapping (VSM) analysis, commonly associated with manufacturing and lean processes. A VSM displays all steps in a process, both value-added and non-value added to visualize inefficiencies, such as hidden work, redundant work, or sequential work that could be de-coupled [13]. A coarse VSM for Target's POG design process flow, which is consistent across many retailers, is analyzed below (Figure 4-1). Internally to Target, and at other organizations, a more detailed VSM can be employed to streamline processes. Particularly for repetitive and ongoing processes, unlocking efficiency can provide great rewards over time.

Target's POG design process is a process that takes place over the course of a year (about 50 weeks), which means that it begins well in advance of the actual POGs being set in stores; this is particularly true for long lead time business lines, such as imports of apparel. The process begins with a decision point of whether a POG will be refreshed; typically, this occurs annually or on a set calendar cycle. For example, after the Consumer Electronics Show (CES) each year, the electronics POGs are reset to incorporate new products.

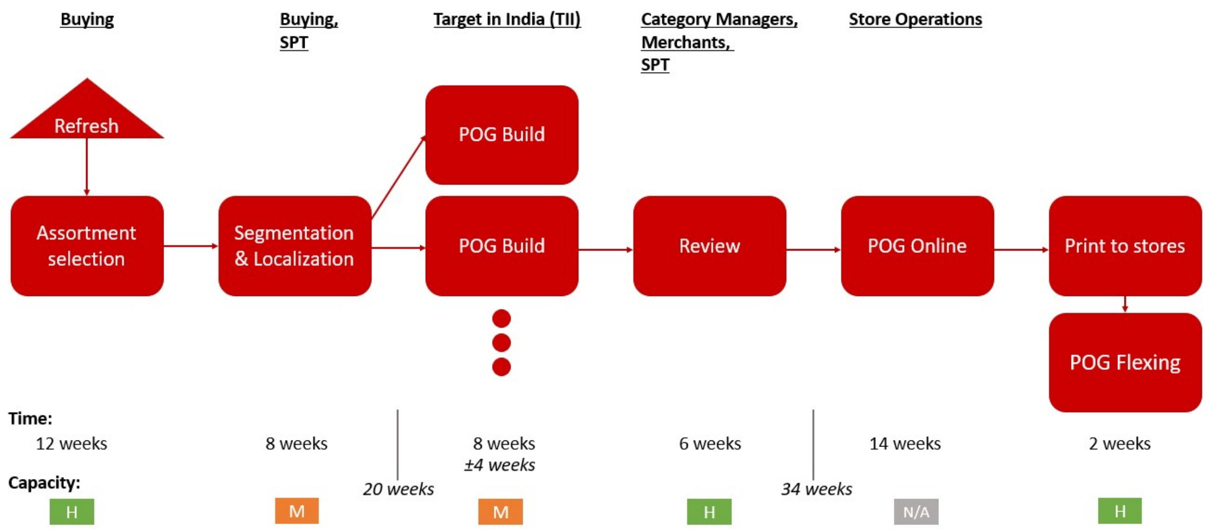

Figure 4-1: Process flow map for POG design, highlighting key stakeholders and bottlenecks in the process. This was used to develop constraints for the POG optimization model.

4.0.1 Assortment Selection

Selecting which products to sell and how to purchase them is the responsibility of buyers. There are two main patterns of purchasing among buyers. One type makes decisions based on 'trend' and the other type makes decisions based on historical performance. Both approaches are likely valuable to continue to stock shelves with items successful with guests, maintain relationships with key suppliers, while at the same time avoiding the opportunity cost of missing out on the next retail trend.

From listening to product pitches and conducting product line reviews, buyers hold much of the power in assortment selection. Buyers are also making decisions on packaging, product dimensions (which impacts on-shelf capacities), and casepack quantity.

Assortment selection occurs over a 12-week period and is not a capacity-constrained step. Each department has their own buyers, and because buying occurs on a calendar cycle, when each step of the process kicks off is predictable.

4.0.2 Segmentation & Localization

With the product assortment determined for a category, the Space Planning & Transitions (SPT) and Buying teams are collaboratively responsible for segmentation of stores and allocating products to each segment (referred to as localization). Key decisions in this step are segmenting products and locations as high-volume or low-volume, deciding on the total number of POGs that are needed, and adjacency planning. Adjacency planning allocates space in a store to a category, such as deciding that a certain category will have 8 feet of shelf space. To add space to one category requires removing space from another, making this a delicate balance. Based on the results of segmentation, a category's strategy can be revised and more or less inventory could be required of certain products. If the strategy is revised during this analytical step, buyers then revisit quantities with suppliers.

Data used to segment stores falls into either customer/geography demographics or store characteristics. For example, understanding income distribution in a specific geography will influence the entry-level or high-end product lines carried in a specific department.

This step occurs over an 8-week period, and similar to buying is not capacity constrained. While there is little variability in the timelines for Assortment Selection and Segmentation, POG Build must begin after 20 weeks to ensure sufficient time for design and to place purchase orders with suppliers.

4.0.3 POG Build

The majority of POG designs are built manually by the Target in India (TII) team. This team receives a product assortment as an input, as well as display guidelines. These guidelines include blocking rules; for example, whether to group products by sub-type or by brand. Additionally, there are inputs on how many facings each brand should have and prominent placement (the height of an item on a POG). It is well documented that placing products at an average adult's eye level will stimulate demand, while lower shelves may be for products targeted towards children.

The POG build step brings together multiple data sources, which may present competing recommendations, to set the optimal POG for a given objective. Typically this relates to sales dollars, inventory turnover, or operational objectives such as BRI.

POG Build is an 8-week process, but does differ by category. This step introduces the first capacity constraint in the POG design process, and it is possible that completed POGs can be sent late (by up to 4 weeks), but are rarely completely early. All POG designs for a category are submitted at once as a batch, rather than following a policy where POGs are visible once each is completed.

4.0.4 POG Review

Before they can be finalized, all POGs are reviewed manually in a two step process. First, Merchants (within Target) or Category Managers (external employees of suppliers) review the POGs based on the selected category strategy. This includes product mix, number of facings, and brand blocking, for example. Once merchants approve the designs, SPT evaluates the designs for feasibility and practicality, like ensuring products will fit based on shelf height.

SPT's review process entails proofing roughly 10% of POGs due to their capacity constraints. The POGs to review are selected based on dollar value importance, so there is predictability in which POGs will be reviewed. This is intended to catch quality issues, but when time is constrained, the likelihood of a quality issue increases. Here, quality issues can refer to item grouping issues such as an item separated from the rest of a brand because the specific item was too tall. Similarly, a quality issue could be a product that was left on the POG that should have been removed during the transition.

The POG review step has 6-weeks allocated to it, but with high variability. The review phase, particularly feasibility review by SPT, acts as a buffer to absorb any delays earlier in the process. Quality is put at risk when the work must occur on a shorter timeline, but the POG Online phase must begin 16 weeks prior to set date.

4.0.5 POG Online

16 weeks before POGs are set, they must enter a holding period referred to internally as ‘POG Online’. During this phase, the POGs have already been finalized and new POG data is rolled to downstream partners like inventory analysts. At this stage, the data becomes digestible for other teams to help calculate inventory metrics like presentation minimums and stage inventory across nodes prior to the POG transition.

The steps that occur during POG online explain why POG review must be completed by week 34 even if there are delays earlier in the process.

4.0.6 Print to stores

The last step in the process of transitioning to a new POG is sending that POG to stores, which occurs 16 days before the POG’s official set date. When a new POG is sent to stores, they may receive new shelving or need to rearrange the existing shelving. Additionally, labels for the shelf are sent to the store. Resetting a new POG is time intensive both to configure the POG and for team members to read the new guidelines and convert it to a display. For this reason, POGs are sent 16 days in advance and stores may transition the POG 1-2 weeks before the official set date to better balance their team member capacity.

For many reasons, there is variability between POGs sent to stores and POGs that are actually set in the field. This practice, referred to as Flexing (and introduced in Chapter 2), is particularly prevalent in Small Format stores when items are out of stock. Team leaders in stores have valid reasons for flexing; however, decisions made in stores are not captured in upstream data systems which creates an issue of data integrity. To make this more clear, imagine a store leader who brings tissue boxes out of the backroom and onto the sales floor to fill up shelf space for paper towels, which were out of stock. Sales of the tissue boxes will increase due to the added facings and inventory, while the data systems only see the initial or expected amount of tissues on the sales floor. Thus, it will appear as if the demand increase is stimulated by a lower level of inventory.

4.1 Bottlenecks in POG Design Process

In many systems, there is a singular bottleneck that is well-known or even designed into the system. In the POG Design Process, there are three potential bottlenecks described below. While in manufacturing, for example, the bottleneck can move in a system as it is product mix dependent, here there are multiple potential bottlenecks due to variability in the system. If POG Design by the TII team is delayed due to many segmentations, a bottleneck is introduced that may impact quality of the final design. Additionally, because there are hard deadlines in the POG process, such as segmentation completed after 20 weeks or POG online 16 weeks prior to set date, each discrete phase of the process can experience a bottleneck.

- **Analysis**: Early analysis, which here includes both assortment selection and the segmentation & localization steps can be a bottleneck in the design process. The additional time required here could be tied to data collection, data availability, or time spent on manual analysis.

- **POG Build**: Given that each POG is built manually by the TII team, this step can serve as a bottleneck. The root cause of this is the information that is incorporated into POGs from disparate systems. For example, the optimization model discussed in Chapter 5 would suggest a number of facings for each item, while a separate model would cover prominent placement and display guidelines. Additionally, based on store layouts, designers may need to make changes due to structural columns in shelves, etc.

- **POG Review**: The final review from category managers, who have their own intentions and knowledge, can serve as the bottleneck. Particularly as the segmentation of POGs increases, the greater number of reviews that needs to take place.

4.2 Recommendations

Capacity analysis: There are ways to unlock capacity in the design process by de-coupling steps and removing points of inflexibility. To continue the manufacturing analogy that underlies this value stream map, the TII team currently produces POG designs in a batch process; the designs are not passed on for review until they are all completed. Moving to a continuous process where a design is sent for review as it is completed can de-couple steps in the review process by allowing build and review to occur in parallel. This would help to create a buffer around the firm deadline at 34 weeks in the end-to-end design process.

Quality: Given that POG transition dates are set on firm timelines, when bottlenecks create delays, the impact may be to quality. Currently, this is mitigated by reviewing 10% of POGs based on revenue. Considering lean methodologies and the application of statistical methods, random sampling may be more effective in catching quality issues. This method is used in many manufacturing facilities to confirm specifications of raw materials, such as the hardness of steel.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

POG Optimization for Fit

This chapter provides a description of the decision model developed to optimize the number of facings for each product on a shelf in a Small Format store. The linear program builds upon prior research projects at Target, and is well suited to address the tradeoffs inherent to inventory management and POG design. Prior research has leveraged optimization models to understand the benefits of modifying facings on BRI levels. This model bounds the problem by optimizing individual POGs, rather than whole categories of products. Narrowing the scope in this way allows recommendations to be translated to in-store operational enhancements.

As introduced in Chapter 3, the Knapsack problem is an appropriate optimization framework to apply to POG design. In a 2001 paper, Yang discusses the translation of the Knapsack problem to a shelf space allocation problem (SSAP), noting that shelf space allocation has additional policy constraints beyond capacity.[11] This SSAP can be summarized in equations 5.1 to 5.4 below, where P is the total profit, N denotes the set of natural numbers, and a store has n products in the product mix across m shelves. The length of shelf k is T_{k} and each facing of product i is a_{i} long. U_{i} and L_{i} are the upper and lower bounds for the number of facings of product i. Lastly, $p_{\mathrm{i,k}}$ is the profit per facing of product i on shelf k, which is assumed to be linear with $x_{\mathrm{i,k}}$.

$$Maximize : P = \sum_{i=1}^{n} \sum_{k=1}^{m} P_{i,k} x_{i,k} \qquad (5.1)$$

subject to

$$\sum_{i=1}^{n} a_i x_{i,k} \leq T_k, \quad k = 1, ..., m, \tag{5.2}$$

$$L_i \leq \sum_{k=1}^{m} x_{i,k} \leq U_i, \quad i = 1, ..., n, \tag{5.3}$$

$$x_{i,k} \in N \cup 0, \quad i = 1, ..., n, k = 1, ..., m. \tag{5.4}$$

This algorithm consists of three phases: first, the preparatory phase assesses feasibility. In the second phase (the allocation phase), space is allocated to products in priority order where priority is determined by profitability. Lastly, in the termination phase, the total profit of the resulting solutions is computed.

Before describing the optimization methodology in full detail below, it is worthwhile noting that while Yang's model focuses on profit, the methods are transferable to optimizing Inventory Management metrics, specifically BRI risk. The model described below optimizes for a 'Fit' objective rather than profit. Inventory management is one piece of a complex POG building puzzle that this research will advocate eventually expanding to include other merchandising considerations such as product financials, consumer psychology, and shopper trends.

5.1 Variable Definition

A description of all variables included in the optimization model can be found in Table 5.1 for consistency, beginning with input variables and ending with decision variables.

Each unique item-location under consideration (variable i) was sourced from the existing planogram for that category. The item's characteristics (number of facings, f_i, and width per facing, W_i) are included in the POG level data. Calculated input variables, POG Fit Category (C_{if}) and Demand Coverage score(D_{if}), were determined for each combination of an item and number of facings from one through ten (reference Chapter 2 for a reminder on how these scores are calculated). The demand-level data and inventory metrics underlying these calculations are available in internal Target

Table 5.1: Variable Definition for POG Optimization Linear Program

Variable	Definition	Type
i	Unique item-location under consideration	Input
f_i	Number of Facings for Item-Location	Input
W_i	Width of a Facing for item-location	Input
C_{if}	POG Fit Category for item, i, based on number of facings, f	Input
D_{if}	Demand Coverage score for item, i, based on number of facings, f	Input
J_{if}	Binary decision variable to allocate item i, f facings on a shelf	Decision

data sets.

5.2 Objective Function

The objective of the optimization is to minimize the risk of generating BRI for a given planogram, while also minimizing the risk of stockouts due to fluctuations in demand. Each of these metrics (BRI risk and DoS coverage) has been operationalized as a score from one through five, where lower scores are associated with more favorable scenarios (hence the minimization objective). A detailed overview of these scores can be found in Section 2.3.

Another rationale for considering DoS, which was not an input in prior research, is to gain buy-in from Category Managers, who review POG designs before they are set in stores (See Chapter 4). Days of Supply is often top of mind for category managers and accounting for this in the model earns trust with these partners. Combined, these two enhancements (optimizing existing POGs versus entire categories, considering Demand Coverage) minimize barriers to in-store implementation and drive towards action in stores.

Based on prioritization, the BRI risk score is given a weighting of 75% while the Demand Coverage score is given a weighting of 25%. Given that the number of facings must be an integer value, adjusting the weighting is relatively insensitive in providing different outcomes; however, given additional resources, this would be worthwhile to explore.

Based on the above, the objective function, which is also referred to as a 'Fit' score is defined as follows:

$$Minimize : \sum_{J=1}^{i} \sum_{f=1}^{10} J_{if}(0.75 \times C_{if} + 0.25 \times D_{if}) \tag{5.5}$$

The objective function is designed to be modular in the sense that new data components can be readily incorporated, so long as they are operationalized on a 1-5 scale and provided a weighting by the business.

5.3 Constraints

Each DPCI can have one, and only one, facing value: Every item is assigned a single number of facings out of 10 possible values. The selection of facings from 1-10 is intended to allow any number of facings, had this limit been raised the solver still would not suggest greater than 10 facings.

$$\forall \text{i in POG}, \ \sum_{f=1}^{10} C_{if} = 1 \tag{5.6}$$

The new shelf design should not exceed the space available: No space is added to an existing POG, and all products must fit in the allotted space. There is no corresponding guidance to use all the space available, as the premise of this optimization is that the store is already space constrained.

$$\sum_{J=1}^{i} \sum_{f=1}^{10} J_{if} \times f \times W_i \leq \text{Shelf Width} \tag{5.7}$$

Demand Coverage cannot be less than one day (or any specified number of days): The DoS score cannot be 5 or 4, depending on Category Manager guidance; in some cases, this constraint was infeasible with a solution and needed to be relaxed.

$$\forall \text{i in POG}, \ \sum_{f=1}^{10} J_{if} \times D_{if} \geq 2 \tag{5.8}$$

5.4 Scope

As introduced, this optimization model focuses on Small Format stores. Five stores were selected in the Manhattan District, and within each store, the Facial Tissue and Air Care planograms were optimized; thus, ten planograms were optimized in total.

These categories were selected based on:

- **Set date**: Given that this project intended to reset POGs in stores, departments were filtered by planned set dates to minimize disruption to the business cycle. Given the high workload of designing and setting POGs, this respected the time of the TII team and team members in stores.

- **Stability**: Categories were then selected based on stability by avoiding categories with high promotional activity and sales variability, which adds complexity to the optimization. Additionally, the selected categories exhibit high-tradeability (so stockouts are less impactful), allowing demand to be treated as exogenous.

- **Collaboration**: Lastly, category merchants and buyers were presented with the opportunity to participate in the research project. In order to make changes on a short timescale, the commitment of category teams was required to review and provide feedback on the commercial feasibility of the model's output.

Based on the above category selection considerations, one can see that the ability to translate research findings into operational changes was prioritized. Performance considerations, like the magnitude of opportunity or category trends, were not inputs into category selection.

In summary, the scope of the pilot (and data underlying the results) is from 30 days of data across five locations, and two POGs in each location. The two POGs in each location were for Air Care and Facial Tissue categories, given that these naturally had a transition around the time of the experiment, exhibited stability, and were the only POGs of the products in the store. Equally important, the category owners of these categories were invested in dedicating additional time to participate in this pilot.

5.5 Results

One example of a 'before' and 'after' from the optimization model can be seen in Figure 5-1, which will be referenced in the upcoming discussion. This individual example well captures trends which were observed across the 10 POGs optimized in this study. Based on analysis of existing versus optimized POG designs, there were two key trends systematically observed:

1. **Unit maximization**: The optimization model measures both BRI and DC with respect to units (rather than dollars); thus, the model attempts to pull units out of the backroom and onto the sales floor. This is intuitive - more units on the sales floor both improves demand coverage and reduces BRI. This is accomplished through the sacrifice of larger items in favor of smaller items. In Figure 5-1 which displays a facial tissue POG, facings of bulk tissue packs are reduced to include more smaller units, increasing the total number of units on the sales floor collectively. Mathematically, if a larger item sees a reduction in Fit score by 1 point, but three smaller cube items improve their Fit scores by 1 point each, the net gain is favorable by 2 points.

2. **Localization**: Optimizing and designing one POG in one location differs from the current practice where the same POG may be set at multiple stores. Focusing on demand data for a single location affects the POG design, and is reflected by items that do not reflect the local market. In NYC, this can be seen by the inclusion of certain large cube items on shelves (in this example, bulk facial tissue packs). In NYC and other urban markets, shoppers have limited storage space and must commute with their purchases without vehicles, creating a preference for smaller items. This compounds with the unit maximization effect to further support reducing bulk packs of tissues.

It is worthwhile noting that this effect of localization can be observed separately from the effects of unit maximization. When analyzing optimized Air Care POGs, car air fresheners are vastly reduced given the low levels of vehicle owernship and demand for these products in NYC. In return, there is an increase in oil diffusers and other scented products in stores like Hell's Kitchen, where apartments are often above odor-producing restaurants.

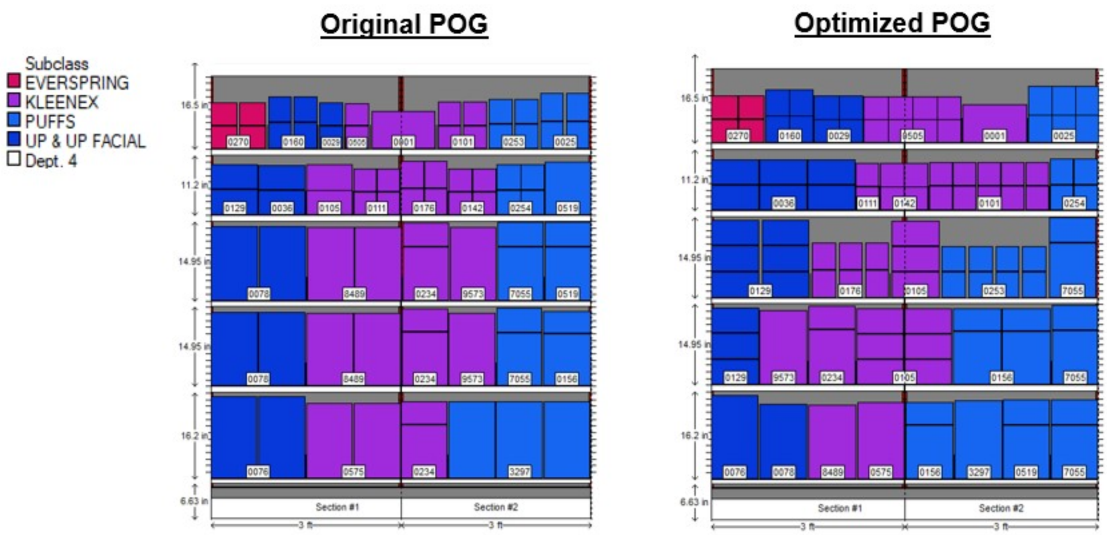

Figure 5-1: Visual display of optimized POGs highlights key impacts of the model, including a localized assortment and reduction in large cube items.

5.5.1 Fit Improvement

Quantitatively, the Fit scores of original and optimized POGs characterize the degree of impact from optimization. One could consider a planogram with 20 items, each with a 'Fit' score of 3. In this case, the aggregate Fit score for the shelf would be 60 (20 items times a score of 3 for each item). Comparing the before and after scores indicates the degree of improvement and opportunity on each shelf, which is depicted in Figure 5-2.

The optimized POGs yield an improvement in Fit scores across departments and locations, and though the magnitude of the increase is not constant, there is no systemic rationale for the magnitude of improvement at each location. Impressively, in some instances, such as Air Care at Location A, the Fit score decreases from 98 to 79, almost a 20% improvement.

As a tool for change management, this score is quite powerful based on the constraints incorporated in the optimization model. While keeping product assortment constant and neither adding nor removing space from a planogram, Fit scores can improve by 20% or an absolute score improvement of 19 in some cases. As a reminder, each integer increase in Fit represents extra Days of Supply or a reduced risk of BRI for a product.

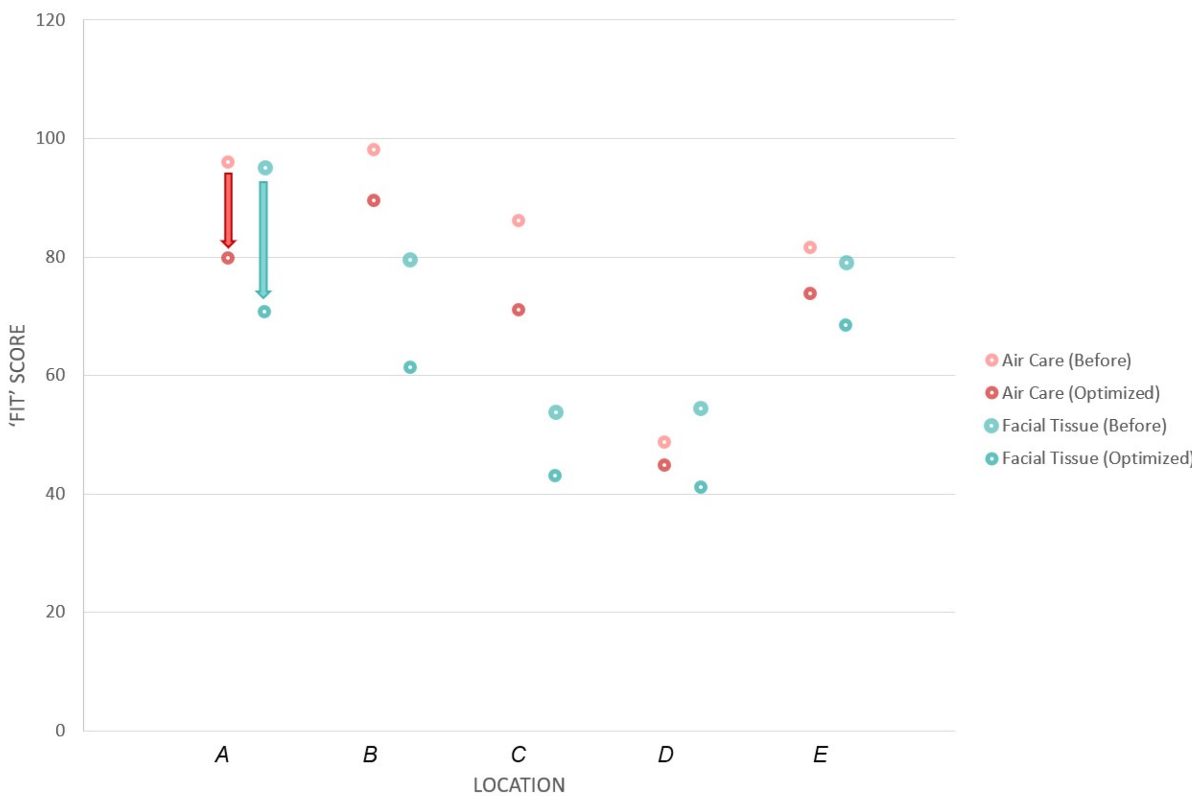

Figure 5-2: Results of 'Fit' score before and after POG optimization for 10 individual displays demonstrate the opportunity for improvement without assortment or space changes.

5.5.2 Simulated Results

Thus far we have observed that the optimization model pulls units from the backroom to the sales floor, and this change can be seen in the improvement in Fit score. The next logical question to ask before implementing this in stores is, "What does this actually mean for BRI levels?" To answer this, a simple simulation was developed that uses inventory accounting to compare inventory levels under the original and optimized POG designs for the prior 12 months.

The inventory accounting formula used is:

$$BRI = BOH + received - sales - POG\,capacity \tag{5.9}$$

Each day, the Beginning on Hand (BOH) is measured, as well as the units received by stores and unit sales. By subtracting the POG capacity from this, the remainder (if greater than 0), represents the number of units in the backroom. The results of this simulation show a marginally lower average BRI over the course of the year, driven particularly by smoothing during times of peak inventory levels. On average, BRI levels fall from 67 units to 60 units as a result of the optimization, which when applied to multiple POGs and categories has a meaningful impact. In addition to averages, it is useful to evaluate these policies as a pressure valve during high-demand periods. During peak times, such as in Summer 2021, these policies can clear almost 100 units from the backroom (Figure 5-3).

Beyond immediately practical purposes of redesigning existing POGs, this simulation can be used to run investigative "What if?" analyses. While for purposes of this optimization existing space was treated as a constraint, the model can easily be extended to consider how adding (or removing) space would impact BRI levels. Adding 96 linear inches to a POG reduced average BRI in the same example to 52 units for the POG, compared to 60 units, with even larger decreases observed during peak times (Figure 5-4).

In future analysis, such scenario planning could build evidence to increase space for a specific category and likewise identify categories to reduce space allocation.

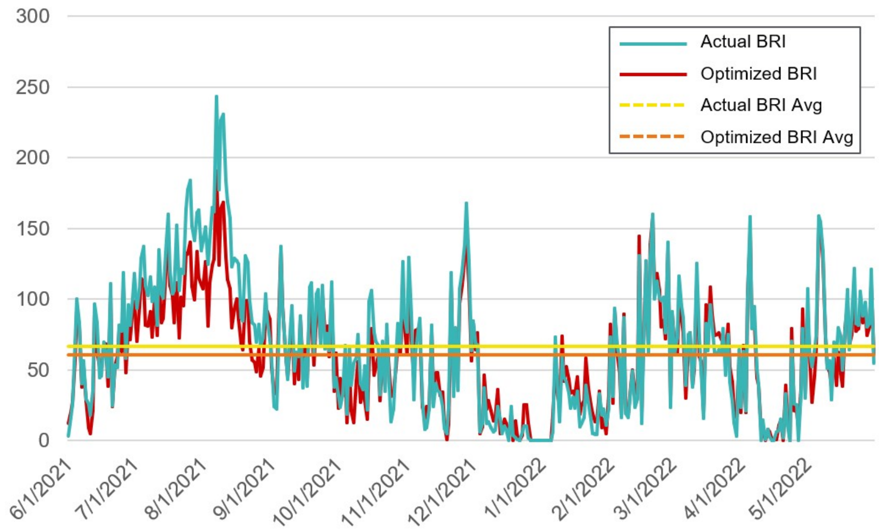

Figure 5-3: Results of BRI simulation for Air Care category at single location demonstrate a decrease in average BRI over the course of the year.

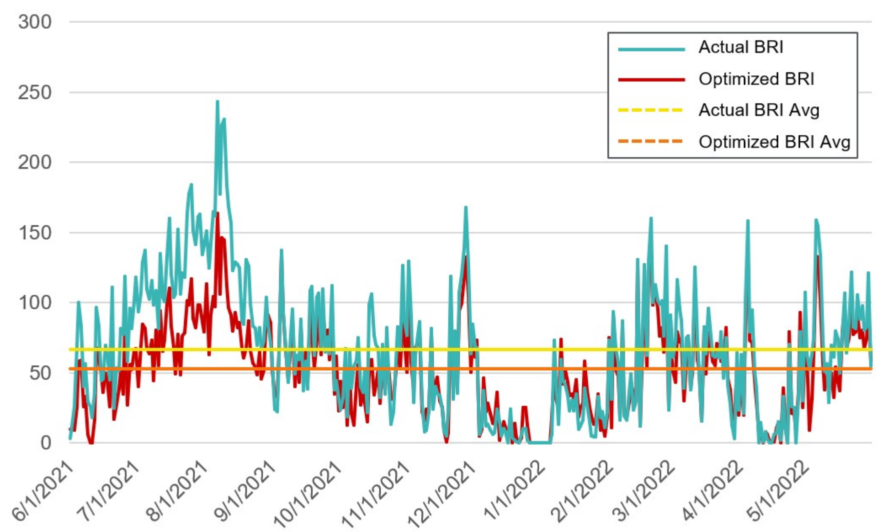

Figure 5-4: Adding 96 linear inches to POG demonstrates greater impact on BRI reduction over the course of a year for Air Care category at single location.

5.6 Extensions of the Model

For future consideration, there are additional features that when added to the model are expected to increase the impact in operational settings.

- **Display Constraints**: This includes two primary improvements. First, the existing model treats a POG as one linear shelf, rather than separate shelves.

When translating facings into a design, this can create challenges for the TII team because of spillover on shelves. One enhancement would be to optimize across each shelf individually. Secondly, in practice, there are certain display guidelines such as grouping by brand or by product sub-class. Incorporating these considerations in the model will not only improve the ability to translate the insight into practice, but continue to reduce the manual workload of designing POGs.

- **Demand Elasticity**: Adding or removing facings to products is well known to stimulate demand upwards or downwards. As we add facings to a product, the model does not account for a lift in sales. In this case, the impact is expected to be minimal because facings typically changed by 1 or 2 units and the categories selected did not experience high demand elasticity.

  However, this becomes more important when extending the model in future scenarios. Some categories, like Food and Beverage, experience higher demand elasticity (adding facings of chips or sports drinks increases demand). Additionally, if SKUs are consolidated or eliminated, the change in facings may be more drastic.

- **Vendor Performance**: Lastly, it does no good to optimize a product on a shelf if the vendor has poor performance. A vendor's fill rate or overall rating could be operationalized as a score from one through five and incorporated into the model, similar to a BRI risk score or DC score. This would prioritize items that can be stocked confidently and reduce empty shelves. In fact, when selecting categories for this study, many were excluded because of bare shelves driven by low vendor fill rates.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 6

Pilot Evaluation

The POG optimization pilot took the output of the model discussed in Chapter 5, and built actual POGs to send to stores by following the process described in Chapter 4. Specifically, the scope of the pilot was to redesign the Facial Tissue and Air Care POGs in five stores in New York, New York. After 30 days, data would be collected to evaluate the success of the pilot.

In order to evaluate whether the pilots set in the five stores were meaningful, the data post-intervention needed to be compared to a control. One option for the control is the same store before the intervention; however, seasonality, effects of marketing campaigns, and other unknown factors could influence our result and would need to be identified and measured. A different option would be to compare the store post-intervention to other similar stores. This still creates challenges, particularly in New York City where there is a smaller number of potential control stores to choose from and they are highly localized based on neighborhood. The best approach to address these challenges is to use a method called Synthetic Control Design (SCD), a relatively novel methodology, to evaluate the effects of the pilot.[14, 15]

In the evaluation of the pilot, the dependent variable measured was Unit Sales. Given that the objective was to reduce BRI and stockouts, the net impact would be reflected in increased unit sales. Placing the right amount of a product in the right place, at the right time (Target's supply chain strategy) will drive sales overall.

6.1 Synthetic Control Design

The premise of SCD is exactly what it sounds like - the control store that the intervention store will be compared to does not actually exist but is generated synthetically. Practically, this means that we do not need to find any single store or set of stores in the untreated group that represents the treatment stores. The synthetic control is a linear combination of an existing set of stores, which is intended to represent the behavior that would have been observed had no intervention taken place (Figure 6-1).

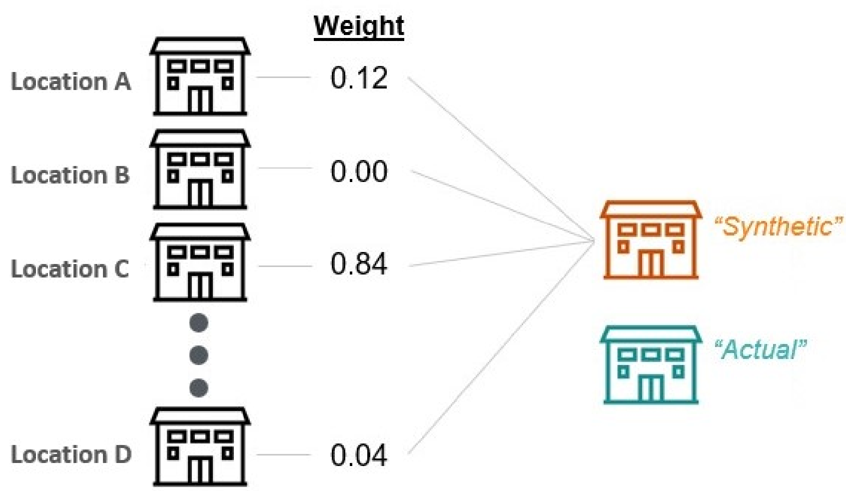

Figure 6-1: Illustrative example of how weighting of individual stores creates a synthetic control of the store where an intervention was made.

6.2 Methodology

The methodology utilized for this SCD was adopted from the open-source textbook, Causal Inference for the Brave and True [16]. We begin by assuming there are $J+1$ stores, where Store 1 is the store that is treated with the intervention. In the case of this experiment, five stores were affected with an intervention. The test store is consequently an average across these five stores to capture the effects of the trend.

What remains, stores J+2 through J+1, can be considered the donor pool of stores to build the synthetic control with. For this study, the donor pool consists of Small Format stores in the East Coast region. The time span for the study goes across T time periods where each period is one day of sales data. Here, we consider the past year of historic data. Therefore, for each store, j and time, t, the outcome $\mathrm{Y_{jt}}$ is observed for each store at a particular time. Then, the following can be defined:

Y^I_{jt}: The potential outcome with intervention

Y^N_{jt}; The potential outcome without intervention,

where the effect size of the intervention is represented as: $\tau_{jt} = Y^I_{jt} - Y^N_{jt}$

In the above formula, Y^I_{jt} is known, whereas Y^N_{jt} is not. The next step is to estimate the potential outcome without intervention, which has already been introduced as a weighted average of the stores in the donor pool (where the weights of the stores $\mathrm{W}=\omega_2,...<\omega_{\mathrm{j+1}}$), and is represented formulaically as:

$$Y^N_{jt} = \sum_{j=2}^{J+1} \omega_j Y_{jt} \qquad (6.1)$$

To assign weights to stores from the donor pool, a LASSO regression model is used. The regression will return the weights for each store in the donor pool to minimize the ordinary least squares (OLS) between the test store and stores in the donor pool.

6.3 Results

6.3.1 Significance

Before sharing results, one must first understand the magnitude of result which would be deemed significant. This threshold was developed using an A|A test, in which pre-intervention data was compared with pre-intervention data.

This test was run over a 12 month period, and divded the timeframe into equal 6-month periods for each segment of the A|A test. Based on the results, an effect magnitude of 1.5 unit lift in daily sales is observed (which corresponds to a roughly 3% unit sale lift). These results indicate that in order to deem the results of a test significant, they must be higher than this 3% lift threshold. In 2021, Target's revenue per square foot was $437, so a 3% lift would correspond to over a $10 increase in revenue per square foot (and Target has 243 million retail square feet). [17] All of this is to set reasonable expectations both for stakeholders and the reader that the level of idiosyncratic noise in the data is such that it would obscure the effects of the POG optimization intervention. A small change is associated with a small signal, and the magnitude of the optimization was not one that was hypothesized to have a single-digit percentage increase in sales. Still, at a billion dollar scale, a 0.1% increase is one million dollars.

This finding should not be interpreted as discouraging; the Next Steps section discusses opportunities to increase the power of the test in order to increase the sensitivity of measuring results.

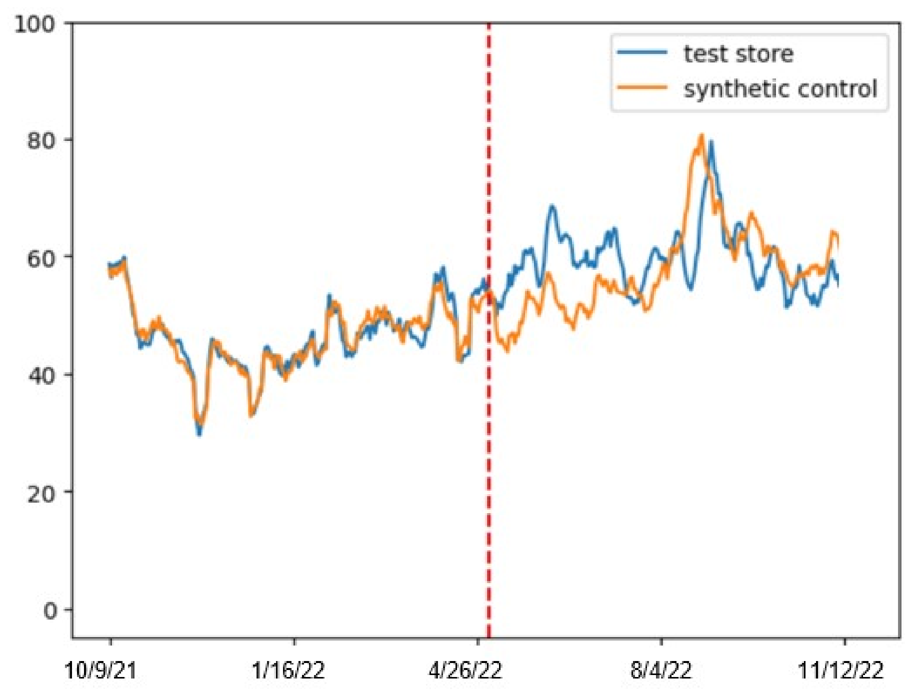

Figure 6-2: Results of A|A test on stores included in the POG optimization test.

6.3.2 Outcome

In brief, the outcome of the results is as expected based on the A|A testing performed. The data does not show a statistically significant change in unit sales over the month following the intervention in stores, which is displayed graphically in Figure 6-3. To emphasize once more, statistically significant results would have corresponded to a 3% lift in daily unit sales; based on the magnitude of intervention made in stores, which accommodated constraints to maintain assortment and avoid major display shifts, the results observed are as expected.

Beyond idiosyncratic noise, there are aspects of the experimental design which could also have lowered the effect of the signal. Most obviously based on earlier discussion of Flexing in stores is the potential that the POG set in store varied from the planned POG, or the same products were previously and continued to be presented in secondary locations in stores.

Additionally, there are constant enhancements to Target's supply chain systems taking place, one of which was deployed just prior to our study. The magnitude of the concurrent system change was quite large and could have muted a signal generated by this study.

6.4 Next Steps

There are a two main areas of next steps that will be meaningful for this research and continue to provide value to Target, and those researching interventions in retailing more generally. The first is the design of the specific experiment, and the second is the general use of synthetic control as a statistical tool to measure outcomes.

On experimental design, to increase the power of the test it could be implemented at more stores and run for longer time. The constraint here was the length of the engagement for Target with this specific project, though results will continue to be monitored and reported internally. Additionally, the idiosyncratic noise within the synthetic control could be reduced by accounting for more variables. Some examples could be weather, store foot traffic, promotions, or other changes to internal systems

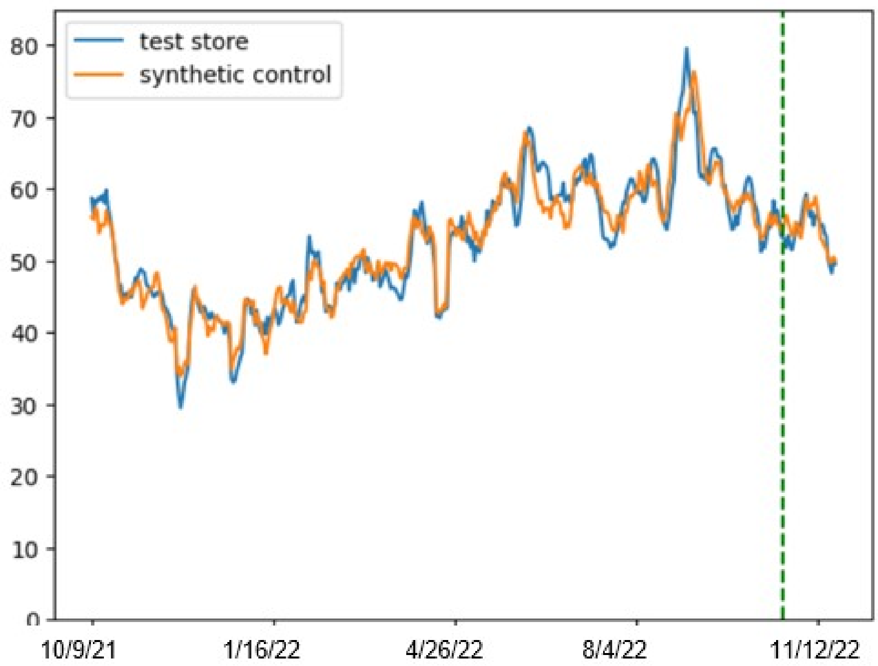

Figure 6-3: Synthetic control results for the Air Care department, displaying unit sales over time. No significant change in unit sales is observed in the 30 days following the intervention.

which could all have confounding effects on unit sales.

Finally, synthetic control should be pursed more broadly as a tool to measure ongoing initiatives at Target. Their store remodels, for example, have been a major driver of growth over the past year and synthetic control could be used to provide insights about such strategic initiatives. This is a valuable tool to add to the toolkit of statisticians at Target.

Chapter 7

Conclusion

This final chapter will summarize the methodology and findings discussed in the prior chapters, as well as draw connections explicitly for the reader who now understands the discrete components of this research. Based upon the learnings, future research that would be worthwhile for the Target and MIT collaboration will be presented as well.

7.1 Key Research Findings

Chapter 2 documented the unique attributes of Small Format stores, particularly the limited assortment in store that drives a much higher sales velocity compared to typical 'Chain' target stores. This introduces a greater emphasis on keeping products in stock to avoid lost sales, while at the same time minimizing backroom inventories to maximize store productivity.

In Chapter 4, the POG design process is introduced through the perspective of a process flow diagram or value stream map. By taking this analytical method commonly used in manufacturing settings and applying it to 'manufacturing' a POG, the bottlenecks in the process and points of inflexibility become visible. This can inspire change in stakeholder management, but is also used to inform constraints of the optimization model discussed in Chapter 5.

The POG optimization model quantitatively defines a term that is increasingly important to Target, 'Fit', as an objective function which weights BRI risk and Days of Supply. The decision model recommends the number of facings for each item on a POG, while keeping assortment constant. The results of the model recommend reducing the number of large items on shelves to fit additional quantities of smaller DPCIs, which both helps empty the backroom, reduces OOS risk, and matches customer demands for smaller items in an urban environment. Despite constraints around assortment and days of supply, in some cases, facing changes were recommended for 40% of DPCIs on the shelf - a nontrivial modification.

The effects of the proposed changes were simulated analytically, and also measured after 30 days of being implemented in stores. While no significant effects were measured in this time scale, as discussed in Chapter 6, this does not mean the changes were ineffective as the 3% hurdle rate of idiosyncratic noise is one that is quite large to overcome. More power can be given to the experiment by increasing its scale and length. Still, SCD remains a powerful tool at Target's disposal to continue monitoring this test and other related tests across their network.

7.2 Future Research

In many chapters, ways to continue the work or build on it with further research were introduced. Below, overarching opportunities to extend this research project are discussed. Continuing this work includes deepening an understanding of assortment and POG design, in addition to tangential research areas which will complement POG research and enable more precise implementation of the ideas uncovered in this study.

Segmentation & Localization: An advantage of the optimization model developed for this research is that it can be applied to localize facings at individual store levels. As discussed in Chapter 4, based on current manual POG design processes, it is impractical to optimize and design POGs at the individual store level; thus, it is critical to segment stores in such a way that the same POG design can be set across them with minimal impact from the uniformity. At the same time, within each segment, the localization of product mix should be considered and may differ. This could potentially be an ideal opportunity to apply algorithms such as K-Means clustering to group stores.

Assortment Reduction: An ongoing dilemma in retailing is the tradeoff between assortment reduction and its benefits to the bottom line and supply chain complexity, against increasing assortment and the stimulation of sales from the merchandising perspective. Trader Joe's has famously done this well, while WalMart lost sales as a result of their assortment reduction strategy. Particularly in Target's Small Format stores, this would be a worthwhile pursuit as it relates to POG optimization. Specifically, this research should entail quantifying demand transference for products removed so that remaining product capacities can be increased accordingly.

Omni-channel fulfillment: As Target leans in to the "stores-as-hubs", shelves and backrooms can be optimized by sales channel. During this research, demand signals for products in stores were not differentiated by traditional in-store shopping, BOPIS, or ship from store (this last channel is not applicable in NYC Small Format stores). Given that the overall goal of this research is to maximize store productivity by reducing BRI, understanding demand split by channel can help Target strategically position inventory on shelves or in backroom for the most efficient fulfillment.

Activity based costing: Driving store productivity improvements is a key rationale of this research and is connected to 'Fit'. While this research was the first to connect BRI risk score to inventory units, one additional linkage can be made between BRI and actual financial costs. The objective function of future optimizations should increasingly focus on dollars, rather than units. Quantifying the impact of reduced BRI on labor efficiency or store sales will ensure that Target's bottom line is supported by optimization results.

Information System Architecture: For research to scale from pilots to the network, there must be a robust enterprise-wide digital strategy behind the implementation. In this line of research, a business case could be investigated that compares costs to benefits of scaling tools to automate POG design or better capture specific DPCI dimensions, for example. A robust study could guide the investment thesis that ties together information systems, merchandising, and supply chain.

7.3 Closing Remarks

This study is the first in the relationship between Target and MIT that included an in-store intervention and measurement of the results. The implementation of research influenced the execution of algorithms, for example, by adding constraints around days of supply to the optimization model developed. Additionally, it provided process learnings on the design of POGs and opportunities to continue streamlining their design.

Starting off with a concept and making changes in some of Target's busiest stores in the heart of New York City is a testament to the commitment and talent of team members at Target. From corporate headquarters, across supply chain facilities, and in retail stores, team members approach Target's business with curiosity, a desire to listen and test new ideas, and above all, a focus on serving guests. This culture is core to Target and will allow Target to maintain their competitive edge regardless of changes in retailing over time.

Appendix A

Terminology

- **Assortment**: The set of products carried in each store at each point in time.
- **Backroom**: Location in stores where excess inventory is stored to replenish the sales floor. In stores with ship from store capability, acts as fulfillment center for customer orders.
- **Backroom Inventory Risk Score**: 1-5 score that corresponds to an item's risk of generating inventory in the backroom. Determined by comparing POG capacity (for a unique product, at a specific location) to key inventory metrics.
- **Buy Online, Pickup in Store (BOPIS)**: Method of fulfillment where guests place an order online, the order is shopped by team members, and guests arrive at the store to pickup a pre-packed bag of the items they requested.
- **Days of Supply (DOS)**: Metric to characterize POG capacity for a unique item at a specific location based on how many days of expected demand can be satisfied.
- **DPCI**: Captures a unique product's hierarchy by listing its **d**epartment, **p**roduct, **c**lass, and **i**tem.
- **Fit**: Metric developed which is a weighted average of BRI risk score and DOS to balance an item's likelihood of generating backroom inventory with risk of

stockouts.

- **Flexing**: A practice where local team member modify POG designs set to the stores. Often used in response to out of stocks, excess inventory, or when POG plans do not match a shelf layout.
- **Global Supply Chain & Logistics (GSCL)**: Organization in Target responsible for purchasing of goods and placement of inventory across network for customer fulfillment.
- **Guest**: Term used to refer to Target customers.
- **Merchandising**: Organization responsible for selecting assortment of products for each category.
- **Planogram (POG)**: Term used to refer to shelf displays at Target locations. POGs are built by the TII team.
- **Space Planning & Transitions (SPT)**: Organization responsible for segmenting POGs, designing POGs, and setting timelines to refresh each POG.
- **Team Member**: Term used to refer to Target employees, particular in field and store locations.
- **Target in India (TII)**: Team members located in Target's Bangalore, India location. This team designs POGs for Target, among a number of other strategic and analytical responsibilities.

Bibliography

[1] Murali K. Mantrala, Michael Levy, Barbara E. Kahn, Edward J. Fox, Peter Gaidarev, Bill Dankworth, and Denish Shah. Why is assortment planning so difficult for retailers? a framework and research agenda. *Journal of Retailing*, 85(1):71–83, 2009.

[2] Target Corporate Website. https://corporate.target.com/.

[3] Das Durgesh. Assessing sales floor capacity and replenishment strategy. Master's thesis, Massachusetts Institute of Technology, 2020.

[4] Lydia Thurman. Assessing inventory replenishment strategy. Master's thesis, Massachusetts Institute of Technology, 2021.

[5] Ben Sidell. Advancing replenishment efficiency utilizing unit of measure and planogram settings. Master's thesis, Massachusetts Institute of Technology, 2022.

[6] Michael Ketzenberg, Richard Mettersand, and Vicente Vargas. Inventory policy for dense retail outlets. *Journal of Operations Management*, 18(3):303–316, September 2000.

[7] Market Watch. *Wal-Mart to close Walmart Express stores as part of reorganization.* https://www.marketwatch.com/story/wal-mart-to-close-walmart-express-stores-as-part-of-reorganization-2016-01-15.

[8] Best Buy Corporate Website. https://corporate.bestbuy.com/.

[9] Kohl's Corporate Website. https://investors.kohls.com/.

[10] Narendra Agarwal and Stephen A. Smith. *Retail Supply Chain Management: Quantitative Models and Empirical Studies*. Springer, 2015.

[11] Ming-Hsien Yang. An efficient algorithm to allocate shelf space. *European Journal of Operations Research*, 131:107–118, 2001.

[12] Alexander Hubner and Kai Schaal. An integrated assortment and shelf-space optimization model with demand substitution and space-elasticity effects. *European Journal of Operations Research*, 261:302–316, 2017.

[13] A.R. Rahani and Muhammad al Ashraf. Production flow analysis through value stream mapping: A lean manufacturing process case study. *Engineering Procedia*, 41:1727–1734, 2012.

[14] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505, June 2010.

[15] Muhammad Amjad, Devavrat Shah, and Dennis Shen. Robust synthetic control. *Journal of Machine Learning Research*, 19(22):1–51, August 2018.

[16] Matheus Facure Alves. *Causal Inference for the Brave and True*. https://matheusfacure.github.io/python-causality-handbook/landing-page.html.

[17] Target Corp. Annual report 2021. https://corporate.target.com/investors.