Application of Flexible Labor and Standard Work in Fulfillment Center Produce Operations

By

Vrajesh Y. Modi

S.B. Mechanical Engineering, S.B. Management Science
Massachusetts Institute of Technology, 2011

SUBMITTED TO THE MIT SLOAN SCHOOL OF MANAGEMENT
AND THE DEPARTMENT OF MECHANICAL ENGINEERING
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREES OF

MASTER OF BUSINESS ADMINISTRATION
AND
MASTER OF SCIENCE IN MECHANICAL ENGINEERING

IN CONJUCTION WITH THE LEADERS FOR GLOBAL OPERATIONS PROGRAM
AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2015

Signature of Author: Signature redacted
MIT Sloan School of Management, Department of Mechanical Engineering
May 8, 2015

Certified by: Signature redacted
Y. Karen Zheng, Thesis Supervisor
Assistant Professor of Operations Management

Certified by: Signature redacted
Sanjay E. Sarma, Thesis Supervisor
Professor of Mechanical Engineering

Accepted by: Signature redacted
Maura Herson
Director, MBA Program
MIT Sloan School of Management

Accepted by: Signature redacted
David E. Hardt
Professor of Mechanical Engineering
Chair, Mechanical Engineering Committee on Graduate Students

Application of Flexible Labor and Standard Work in Fulfillment Center Produce Operations

By

Vrajesh Y. Modi

Submitted to the MIT Sloan School of Management and the Department of Mechanical Engineering on May 8, 2015 in Partial Fulfillment of the Requirements for the Degrees of Master of Business Administration and Master of Science in Mechanical Engineering

ABSTRACT

This thesis demonstrates the applicability of flexible labor and standard work in increasing labor productivity and improving quality in fulfillment center produce operations. Three hypotheses were tested: (1) labor productivity would be increased by implementing a flexible labor staffing model and shift-by-shift labor requirement projections; (2) additional labor productivity from redesigned products would be captured using the flexible labor staffing model with updated projections; and (3) product quality would be improved by defining and refining standard work for inspection processes and optimizing inspection frequency. Indeed, implementation of flexible labor generated an average 44% productivity improvement. Furthermore, introducing redesigned products and updating projections generated an additional average 16% productivity improvement. While stalled implementation of standard work made it difficult to show that standardizing inspection processes and optimizing inspection frequency improves product quality via improved inspection consistency, the project did identify several produce categories that did not require inspection and established a starting point to standardize inspection processes by documenting inspection best practices. During these changes, setting easily achievable targets that required an increase in performance improved team morale, while overly aggressive targets would have worsened morale and hindered implementation. AmazonFresh produce processes can be further improved by ensuring the Produce Receive function serves as a "first line of defense" against quality issues, redesigning more products to reduce downstream labor requirements, implementing standard work in inspections, experimenting with lower bin inspection frequencies, institutionalizing quality and productivity metrics, and deploying changes across all AmazonFresh sites. Concurrently, AmazonFresh leadership may consider installing an incentive system that supports productivity and quality improvements, create roles within the central organization to support process improvements, and incorporate volume and service level requirements when designing new sites. Finally, future research opportunities include observing the effect of inspection frequency on inspector performance and assessing whether recent advances in sensor and conveyance technology can improve or replace existing processes.

Thesis Supervisor: Y. Karen Zheng
Title: Assistant Professor of Operations Management

Thesis Supervisor: Sanjay E. Sarma
Title: Professor of Mechanical Engineering

Acknowledgements

I want to acknowledge the MIT Leaders for Global Operations Program and specifically the partnership between MIT and Amazon for making this work possible. I am incredibly fortunate to have the opportunity to work with such a terrific team while conducting my research. I want to thank my faculty advisors and Amazon supervisors and mentors – Professor Sanjay Sarma, Professor Karen Zheng, David Mabe, Brent Hill, Torsten Pilz, Michael Huffaker, and Mark Mastandrea – for their advice and help throughout the project. I also want to thank the members and colleagues of the Worldwide AmazonFresh and Prime Pantry Operations Team – Garrett Lewis, Aaron Thurgood, David McCalib, Bryan Maybee, Carrie Tachiyama, Aimee Constantine, Mark Rowe, Eric Nemetz, Brent Rudser, Yichun Li, Anthony Auffray, Dilshika Wijesekera, Susanne King, Chao-Lin Hsieh, Eda Leka, and Nobuhito Iguchi – for providing ideas and encouragement. Testing and implementation of these ideas would not be possible without the openness, positive attitude, and willingness to experiment from the entire fulfillment center team and especially from John McKenna, Jeff Heckel, Mark Sibon, Mike Pimentel, Ashlea Savona, Mike Trageser, the Seattle Produce Team, John McLaughlin, David Nielson, Matt Kelly, Kyle Smith, David Clow, and Anthony Epperley. Finally, I want to thank the AmazonFresh Produce Retail Team for supporting this research and assisting with brainstorming and data collection throughout the project.

Glossary

5S: Indirect waste reduction process; stands for Sort, Straighten, Shine, Standardize, and Sustain
AM: Area Manager
Associate: Warehouse employee; typically trained to receive, stow, and pick products
FC: Fulfillment Center
FC SME: "Subject Matter Expert;" a support role that reports directly to the FC Site Leader
IB: Inbound
NYR: "Not Yet Received;" products that have arrived at the FC but have not formally been received
OB: Outbound
PO: Purchase Order
Produce: Fruits and vegetables; this document will not use the verb form of "produce" (i.e., "to make or create")
Produce AM: Produce Area Manager; the manager responsible for all Produce Specialists and Produce Associates at a particular FC
Produce Associate: Employee on the FC Produce Team who may or may not have some past experience working with produce
Produce Specialist: Employee on the Produce Team who typically has several years of experience working with produce and may serve as a Produce Team shift leader
SKU: Stock Keeping Unit; a unique identifier for a particular product; all instances of the product have the same SKU (i.e. an SKU is different from a serial number)
VTO: "Voluntary Time Off;" offered by FC managers to Associates if there is not enough work

Table of Contents

Chapter 1: Introduction

This thesis demonstrates the applicability of flexible labor and standard work in increasing labor productivity and improving quality in FC produce operations. The data for this thesis was generated at AmazonFresh during a 28-week project, which this chapter summarizes. Data are normalized throughout this thesis to protect Amazon Confidential information; however, percentages, ratios, and other relative relationships are accurate.

A. Project context

AmazonFresh, a subsidiary of Amazon that provides online grocery fulfillment, seeks to simultaneously expand into new geographies, improve product quality, reduce operational costs, and promise short cycle time from order to delivery. Concurrently achieving these seemingly conflicting objectives requires a fundamental redesign of the company's existing processes (described in Chapter 4). Produce operations within the FC were identified as an area of opportunity for the project due to unique product requirements and high product variability relative to traditional Amazon operations.

B. Project goals

The project goals were to increase labor productivity and improve product quality. In particular, the objective was to reduce long-term end-to-end unit fulfillment costs for produce (without increasing costs elsewhere) and reduce frequency of produce quality-related complaints. It was presumed that customer demand would not materially change, so reducing labor hours was seen as the primary lever for driving unit cost reduction and labor productivity increases.

C. *Hypotheses*

Throughout the project, my ingoing assumption was that practices that increased labor productivity and improved quality in traditional manufacturing were applicable to FC produce operations. Specifically, the first hypothesis was that labor productivity would be increased by implementing a flexible labor staffing model together with shift-by-shift labor requirement projections. The second hypothesis was that the flexible labor staffing model, together with updated labor projections, would help to realize the labor productivity improvement from a

redesigned product. The third hypothesis was that product quality would be improved by defining standard work for inspection processes, analyzing the root causes of common defects, and iteratively refining the standard work to eliminate or mitigate these root causes.

D. Approach

This project examined all produce-specific operations (receiving, replenishing, stowing, banana processing,[1] trimming, packing, inspecting) at the AmazonFresh Seattle FC to identify and reduce sources of waste and variation. In particular, the project:

- Established a baseline productivity level by making qualitative and quantitative observations of the Seattle FC Produce Team for eight weeks (Weeks 6-13).
- Reduced the amount of "slack" by implementing a flexible labor staffing model, providing shift-by-shift labor requirement projections, and holding shift leaders accountable for meeting projections (Weeks 14-19).
- Captured labor savings from a banana product that was redesigned[2] to significantly reduce the labor needed to process it. Savings were realized by updating the labor projections and holding shift leaders accountable for meeting new projections (Weeks 20-23).
- Defined and started to implement standard work for produce quality inspections to enable continuous improvement. Inspection frequency was a particular area of interest.

Findings from the changes in Seattle were used to set productivity and quality targets across the network and at new sites.

E. Thesis overview

This thesis will focus on expected and actual impact of these changes to demonstrate applicability of flexible labor and standard work in FC produce operations. Chapter 2 will provide background information on the grocery industry, Amazon, and AmazonFresh. Chapter 3 will review literature on flexible labor, standard work, and control charts. Chapter 4 will describe AmazonFresh FC operations and AmazonFresh produce operations in Seattle, as of February 2014. Chapter 5 will discuss the initial implementation of flexible labor. Chapter 6

---

[1] See Chapter 4, Section B and Chapter 6, Section A for descriptions of the existing banana process and improvement opportunities.

[2] The banana product was redesigned from a bunch of five bananas to a prepackaged two-pound bag of bananas.

will discuss the subsequent application of flexible labor to capture the savings from a redesigned product. Chapter 7 will discuss the implementation of standard work and the experimentation with inspection frequency. Finally, Chapter 8 will draw conclusions, present recommendations, and suggest areas for future research.

Chapter 2:
Background

A. Grocery industry overview

Upon hearing about the AmazonFresh service, people immediately recall Webvan, considered by many to be the first major venture in online grocery fulfillment. It was started in 1999 by Louis Borders and went bankrupt in July 2001.[3] While operating, it expanded rapidly to San Francisco, Dallas, San Diego, Los Angeles, Chicago, Seattle, Portland, Atlanta, Sacramento, and Orange County, CA.[4,5] Generally, Webvan's failure is attributed to the company's rapid expansion. Due to technical issues and lack of sufficient customers, its operations were never profitable, but the company continued to invest in facilities and technology and ultimately ran out of cash.[6]

Today, there are several new companies with various operating models in the United States offering online grocery shopping service with home delivery, such as FreshDirect, Instacart, and Google Express. FreshDirect offers next day delivery and operates in various counties in New York, New Jersey, Connecticut, Pennsylvania, and Delaware.[7] FreshDirect claims that it is able to cut costs and deliver fresh food four to seven days more quickly by buying directly from the source, cutting out middlemen, and operating its own distribution center.[8] Instacart offers same day delivery and operates in Atlanta, Austin, Boston, Boulder, CO, Chicago, Denver, Houston, Los Angeles, New York City, Philadelphia, Portland, San Francisco, San Jose, CA, Seattle, and Washington, DC.[9,10] However, rather than operating its own distribution centers and owning its inventory, Instacart sends

---

[3] Lunce, Stephen E., et al. "Success and failure of pure-play organizations: Webvan versus Peapod, a comparative analysis." *Industrial Management & Data Systems*, Vol. 106, No. 9, pp. 1344-1358, 2006.
[4] Glasner, Joanna. "Why Webvan Drove Off a Cliff." July 10, 2001. Accessed December 16, 2014 at http://archive.wired.com/techbiz/media/news/2001/07/45098.
[5] Sandoval, Greg. "Webvan to close Sacramento operations." April 24, 2001. Accessed December 16, 2014 at http://news.cnet.com/Webvan-to-close-Sacramento-operations/2100-1017_3-256370.html.
[6] Glasner, Joanna. "Why Webvan Drove Off a Cliff." July 10, 2001. Accessed December 16, 2014 at http://archive.wired.com/techbiz/media/news/2001/07/45098.
[7] "FreshDirect - Help – FAQs." Accessed December 16, 2014 at https://www.freshdirect.com/help/faq_home.jsp?page=home_delivery.
[8] "FreshDirect - Help – FAQs." Accessed December 16, 2014 at https://www.freshdirect.com/help/faq_home.jsp?page=what_we_do.
[9] "Instacart - Same-day Grocery Delivery in Atlanta, Austin, Boston, Chicago, Denver, Los Angeles, New York City, Philadelphia, San Francisco Bay Area, Seattle, and Washington, D.C." Accessed December 16, 2014 at https://www.instacart.com/faq.
[10] Rigby, Darrell, et al. "Digital Darwinism: Winning with the best of digital and physical." *Bain Retail Holiday Newsletter*, Issue 4, December 19, 2013.

its employees to stores such as Whole Foods and Costco to do groceries on behalf of the customer. Similarly, Google Express (formerly Google Shopping Express) offers same day delivery in San Francisco, San Jose, West Los Angeles, Manhattan, Chicago, Boston, and Washington, DC.[11] Like Instacart, Google Express does not hold its own inventory but rather does shopping in physical stores on behalf of customers. However, Google Express goes beyond groceries and offers to shop at a range of retailers, such as Babies "R" Us, Kohls, and Staples.[12] Finally, an interesting online grocery retail player is Ocado, a UK-based company that made a profit in 2012. Ocado operates its own "Customer Fulfillment Centers," a "Non-food distribution center," and local delivery hubs ("Spokes").[13] While Ocado subsequently lost money in 2013, its positive 2012 profitability does suggest that it is possible for companies operating very efficiently to break even or profit in online grocery fulfillment.

Despite these new ventures, the grocery market is still dominated by traditional retailers. The top four grocery retailers by sales in 2014 were Wal-Mart, Kroger, Target, and Safeway.[14] Interestingly, a recent trade journal article commented, "We don't expect Amazon to show up on the Super 50 any time soon. But its entry into the grocery game is a very real concern for traditional players fighting to retain the market share that many of them took for granted for generations."[15] These major grocery retailers operate extensive supply chains that terminate in their physical retail stores.[16]

A final traditional retailer worth mentioning is Whole Foods, which AmazonFresh often compares itself to. Whole Foods is "the largest retailer of natural and organic foods in the U.S. and the 7th

---

[11] "Delivery areas - Express Help." Accessed December 16, 2014 at https://support.google.com/shoppingexpress/answer/4559799?hl=en.
[12] "The stores on our site - Express Help." Accessed December 16, 2014 at https://support.google.com/shoppingexpress/answer/4562192?hl=en.
[13] "Annual Report for the 52 weeks ended 1 December 2013." Accessed December 16, 2014 at http://www.ocadogroup.com/investors/reports-and-presentations/2013.aspx.
[14] Chanil, Debra. "Seizing the Day | Progressive Grocer." May 1, 2014. Accessed December 16, 2014 at http://www.progressivegrocer.com/node/68197.
[15] Ibid.
[16] Senauer, Ben and John Seltzer. "The Changing Face of Food Retailing." *Choices: The Magazine of Food, Farm, and Resource Issues*, a publication of the Agricultural & Applied Economics Association. 4th Quarter 2010 | 25(4). Accessed December 16, 2014 at http://core.kmi.open.ac.uk/download/pdf/6220630.pdf.

largest public food retailer overall."[17] In particular, Whole Foods believes that its high quality standards differentiate it from other grocers.[18]

B. Amazon overview and history[19]

Founded in 1994 by current CEO Jeff Bezos, Amazon.com described itself in its 1997 Annual Report as "the leading online retailer of books." Over the years, the company expanded its product selection significantly and entered various new businesses including the manufacture and sale of electronic devices (such as the Kindle and Fire Phone), the hosting of technology infrastructure (i.e., Amazon Web Services), and the sale of digital products such as e-books and streaming video. Over the past decade, Amazon has grown its net sales at a Compound Annual Growth Rate of 26.8%; in 2013, Amazon's net sales were $74 billion (Figure 1).

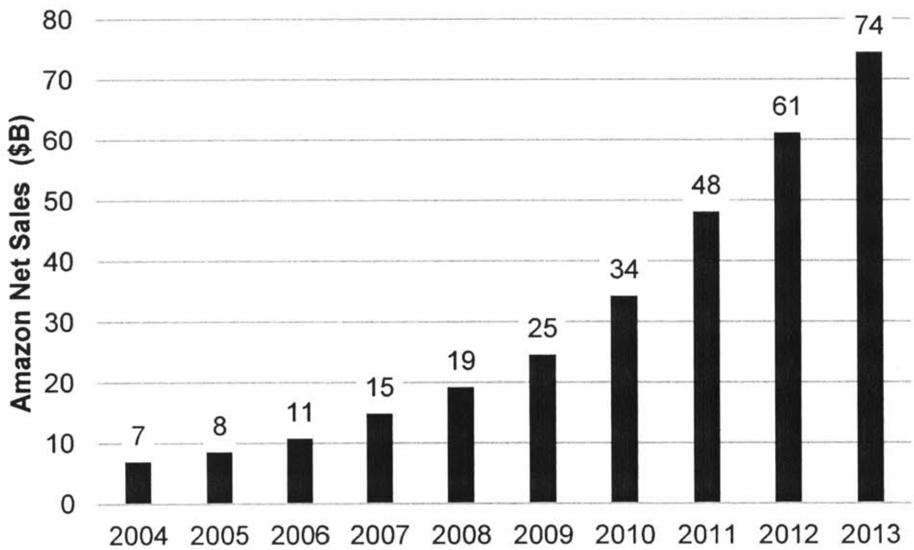

Figure 1: Amazon Net Sales ($B, 2004-2013)

[17] Whole Foods Market, Inc. Form 10-K, 2014. Accessed December 16, 2014 at http://assets.wholefoodsmarket.com/www/company-info/investor-relations/annual-reports/2014-WFM-10K.pdf.
[18] Ibid.
[19] Amazon Annual Reports (1997, 2007, 2010, 2013). Accessed December 16, 2014 at http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-reportsAnnual.

C. AmazonFresh history

AmazonFresh started in August 2007 by delivering groceries to Mercer Island, a suburb of Seattle.[20] For six years, various operating models were tested and refined, and the business expanded to encompass most of Seattle and its suburbs. The customer response was mixed. Some enthusiastically exclaimed, "I don't like spending the time to shop, so this is perfect!"[21] Meanwhile, others were less enthusiastic, remarking, "I'm an avid buyer of strawberries, and I never liked their strawberries. I'd only buy bagged lettuce and maybe a cucumber – things that wouldn't go bad right away."[22]

Then, in June 2013, AmazonFresh expanded to Los Angeles,[23] and in December 2013, the company added San Francisco.[24] However, the company continued to experiment in these new cities with different subscription, fulfillment, and delivery models. In particular, while the free loyalty program in Seattle (called "Big Radish") offered free or discounted delivery based on a customer's total spending within a certain time period and based on the order size, the subscription model in Los Angeles and San Francisco (called "Prime Fresh") was marketed as an upgraded version of Amazon Prime[25] that granted access to online grocery through AmazonFresh and unlocked free same-day delivery across a wide range of Amazon products on orders over $35. Unseen by customers, the building designs and layouts were also different at each of the three sites, creating a great opportunity during the project to analyze the pros and cons of various options. Finally, while delivery trucks transported completed orders directly from the Seattle FC to customers, in Los Angeles and San Francisco, orders were transported together from a distant FC to local delivery stations and then separated onto different delivery routes.

---

[20] Harris, Craig and John Cook. "Amazon starts grocery delivery service." August 1, 2007. Accessed December 16, 2014 at http://www.seattlepi.com/business/article/Amazon-starts-grocery-delivery-service-1245445.php.
[21] Martinez, Amy. "AmazonFresh set to expand?" January 26, 2013. Accessed December 16, 2014 at http://seattletimes.com/html/businesstechnology/2020208438_amazonfreshxml.html.
[22] Ibid.
[23] Horsey, Julian. "AmazonFresh Grocery Delivery Service Now Available In Los Angeles." June 10, 2013. Accessed December 16, 2014 at http://www.geeky-gadgets.com/amazonfresh-grocery-delivery-service-now-available-in-los-angeles-10-06-2013/
[24] Soper, Taylor. "Amazon Fresh launches in San Francisco with $299/year 'Prime Fresh' membership." December 11, 2013. Accessed December 16, 2014 at http://www.geekwire.com/2013/amazon-fresh-san-francisco/.
[25] An Amazon Prime membership costs $99/year. A "Prime Fresh" membership costs an additional $200/year.

Chapter 3:
Literature review

A. Flexible labor

A great deal has been written about flexible labor in various contexts. What differentiates this thesis is that it quantitatively demonstrates the applicability of this concept (primarily to reduce labor use) in the online grocery fulfillment industry. The term "flexible labor" can be used to refer to a variety of concepts, including temporary workers (i.e., short-term employees), part-time workers, personnel transfers between facilities, paid overtime, flexible working hours, and cross-training. In this thesis, the term is used to refer collectively to personnel transfers between facilities, voluntary time off (VTO), and cross-training.

The consensus in industrial engineering and supply chain literature is that process flexibility helps address variability in labor supply and market demand. Francas discusses the importance of flexibility especially where there is a risk of low capacity utilization, and his model shows that there can be complex relationships between different kinds of flexibility.[26] Cachon and Terweisch discuss flexibility in the context of "[attempting] to assign some workers to processes creating other products and to have the remaining workers handle multiple machines simultaneously for the process with the low-demand product."[27] In this thesis, we extend this concept from manufacturing ("machines") to warehousing operations ("processes" and "functions"). Cachon and Terweisch also write about a "second pool of (temporary) workers" for high-demand scenarios.[28]

Slomp agrees, "It has been shown repeatedly that increases in labor flexibility positively affect system performance," but adds, "Most of the positive effects are achieved without going to the extreme of total flexibility. Several authors have noted in recent years that limited amounts of

---

[26] Francas, David, et al. "Machine and Labor Flexibility in Manufacturing Networks." February 11, 2010. Accessed December 20, 2014 at https://plis.univie.ac.at/fileadmin/user_upload/abt_logistic/Working_Papers/Francas__Loehndorf__Minner__2011__Labor_Flexibility.pdf.
[27] Cachon, Gerard and Christian Terwiesch. Matching Supply with Demand: An Introduction to Operations Management (Second Edition). Section 10.7, pp. 214-215. *McGraw-Hill*, 2009.
[28] Ibid.

cross-training are sufficient to gain near optimal performance results." Consequently, he develops a model to select the optimal workers to cross-train.[29] He also discusses the formation of "chains" of workers based on their qualifications to balance workload by reallocating work.[30]

Hopp and Van Oyen suggest a framework for assessing the suitability of cross-trained workers in a given operation.[31] They discuss several potential benefits including near-term efficiency improvement, quality improvement, learning curve acceleration, and improved organizational culture; they also provide several potential tactical models including scheduled rotation, floating workers, zoned work-sharing, worker-prioritized work-sharing, and a "craft" approach.[32] Importantly, they also point out that to actually realize the cost savings that are created by cross-training, headcount needs to be reduced or volume increased. Finally, they point to several companies, including IBM, John Deere, and GE, that have implemented flexible labor to achieve various strategic objectives.[33] However, Easton points out that there are also limits to the benefits of cross-training due to the structure of the workday and due to absenteeism.[34]

B. Standard work

Standard work is a well-established concept in manufacturing. In this project, documenting and refining produce quality inspection standard work was used to make product quality improvement possible through more consistent inspections. Standard work is also a necessary precursor to applying flexible labor techniques optimally.

Taiichi Ohno, one of the creators of Toyota's Just-In-Time production system, discussed the written standard work sheet in the context of visual control, and identified three key elements of

---

[29] Slomp, Jannes, et al. "Cross-training in a cellular manufacturing environment." *Computers & Industrial Engineering* 48 (2005) 609–624.
[30] Ibid.
[31] Hopp, Wallace J. and Mark P. Van Oyen. "Agile Workforce Evaluation: A Framework for Cross-training and Coordination." *IIE Transactions* (2004) **36**, 919-940. Accessed on December 20, 2014 at http://webuser.bus.umich.edu/whopp/reprints/Agile%20Workforce%20Evaluation.pdf.
[32] Ibid.
[33] Ibid.
[34] Easton, Fred F. "Cross-training performance in flexible labor scheduling environments." *IIE Transactions* (2011) 43, 589–603. Accessed on December 20, 2014 at http://www.tandfonline.com/doi/pdf/10.1080/0740817X.2010.550906.

standard work procedures: cycle time, work sequence, and standard inventory.[35] Womack and Jones later characterized standard work more broadly as "the best way to get the job done right the first time, every time."[36] About standard work, the Lean Enterprise Institute writes, "Standardized work is one of the most powerful but least used lean tools. By documenting the current best practice, standardized work forms the baseline for kaizen or continuous improvement. As the standard is improved, the new standard becomes the baseline for further improvements, and so on. Improving standardized work is a never-ending process."[37]

C. Control charts

While flexible labor and standard work were the concepts that were applied to directly improve the productivity and produce quality, control charts were used to quantitatively measure the impact. A great deal of information is available about "statistical process control," a set of methods that use statistics to monitor and control the quality of a process. In this project, a specific type of control chart, called an attribute control chart, was used to track the quality of the produce.

Traditional control charts track a specific parameter (such as length or diameter) of the output of a manufacturing process. Statistical methods are used to set an "upper control limit" (UCL) and a "lower control limit" (LCL); when the parameter exceeds the UCL or falls below the LCL, the process is said to be "out of control," and an investigation is warranted to determine the cause of the variation.

However, produce quality cannot easily be reduced to a single parameter. As Cachon and Terweisch explain, "Rather than collecting data concerning a specific variable and then comparing this variable with specification limits to determine if the associated flow unit is defective or not, it is frequently desirable to track the percentage of defective items in a given sample. This is especially the case if it is difficult to come up with a single variable, such as

---

[35] Ohno, Taiichi. Toyota Production System. Chapter 2, pp. 21-22. *Productivity Press*, 1988. English translation of Toyota seisan hoshiki, published by *Diamond, Inc.*, 1978.
[36] Womack, James P. and Daniel T. Jones. Lean Thinking. Chapter 6, p. 113. *Simon & Schuster*, 1996.
[37] "Standardized Work: The Foundation for Kaizen." *Lean Enterprise Institute*, 2009. Accessed December 20, 2014 at http://www.lean.org/Workshops/WorkshopDescription.cfm?WorkshopId=20.

length or duration, that captures the degree of specification conformance."[38] Indeed, produce has a range of characteristics, including color, smell, shape, firmness, weight, bruising, taste, etc., that determine whether it is "good." Consequently, attribute control charts are the best way to track produce quality. Specifically, this project tracked the **p**roportion of defective units using a p-chart,[39] (as opposed to tracking the total or average **c**ount of defects using a c-chart).

---

[38] Cachon, Gerard and Christian Terwiesch. Matching Supply with Demand: An Introduction to Operations Management (Second Edition). Section 9.6, pp. 190-192. *McGraw-Hill*, 2009.
[39] "6.3.3. What are Attributes Control Charts?" NIST/SEMATECH e-Handbook of Statistical Methods. Created June 1, 2003; updated October 30, 2013. Accessed December 21, 2014 at http://www.itl.nist.gov/div898/handbook/pmc/section3/pmc33.htm.

Chapter 4:
Prior State: AmazonFresh and Seattle Produce Operations

A. AmazonFresh operations overview

There are four major grocery FC process paths – Receive, Stow, Pick/Pack, and Roll/Sort – that apply to nearly all products. This section describes them (in a slightly simplified way) and contrasts them with analogous processes in traditional, non-grocery Amazon (.com) FCs to shed light on potential challenges of grocery fulfillment.

Receive

Most products are delivered to the FC packaged in boxes on pallets. When a truck fulfilling a particular PO arrives at the FC, an Associate at the loading dock uses a thermometer to ensure that the product is arriving at an appropriate temperature and then "dock-receives" the PO. This makes it possible for Associates performing the "Receive" function to receive individual units against that PO. The pallet may sit on the loading dock (as "NYR") for some period of time depending on its priority and staffing conditions, but eventually it is brought from the loading dock to the "Receive" area inside the FC, where an Associate takes individual boxes off the pallet, opens them, and puts them on a conveyor belt. Associates performing the "Receive" function take a box off the belt, take the items out of the box, quality check the items, and then place them on a stow cart.

There are several differences between this process and the analogous one in a traditional Amazon FC. First, grocery FCs need to begin the process by taking the temperature of the product, while non-grocery FCs do not need to worry about this at all. This step can lead to confusion or disagreements with the delivery driver when a product arrives outside of specifications. Second, non-grocery FCs do not necessarily need to define priority levels for receiving, since there are no concerns about relative product perishabilities. Consequently, in non-grocery FCs, product can simply be received in the same order that it arrives at the FC, reducing complexity. Third, in grocery FCs, during the quality check, Receivers need to observe many more features of grocery products, such as expiration date, smell, color, texture, etc., depending on the product. It is also more important in grocery settings to read the product packaging more closely, because, for example, different flavors of yogurt or different types of milk may all look very similar.

Stow

An Associate performing the "Stow" function travels around the FC with the stow cart and puts the items away on shelves. In grocery FCs, each product is stowed in an appropriate area based on its durability and temperature requirements. For example, durable products like 12-packs of soda are stowed in areas where they will be picked first during a pick/pack sequence, while fragile products such as chips and eggs are stowed in areas such that they will be picked last (to minimize crushing). Some products are stowed in a frozen environment, while others are stowed in a refrigerated area, and yet others are stowed at room temperature

In contrast, in non-grocery FCs, products can essentially be stowed anywhere, eliminating a great deal of complexity.

Pick/Pack

Once several customers place orders, the orders are communicated to an Associate performing the "Pick" function. This Associate travels around the FC with a pick cart containing several totes (one or more for each order) and gathers items into the appropriate totes to fulfill the customer orders. This process takes into account variations in expiration date for products that have multiple instances within the FC. By the time the Associate has finished one trip through the FC, several customer orders have been picked and simultaneously "packed" into totes. Associates are trained to follow a certain packing order to ensure that fragile items are not damaged by more durable ones. Additionally, Associates are trained to recognize if an item has become spoiled while in the FC. Also, Associates are trained on items that should not be combined with others in a single tote. (For example, onions and garlic should not be combined with flowers, organic fresh produce should not be combined with conventional fresh produce, and fresh meat needs to be separated from other foods to prevent contamination risk.) Additional complication occurs when an order contains items from multiple temperature zones.

In non-grocery FCs, none of these product-related complications apply. Furthermore, it is not even necessary for an Associate to pick entire orders for particular customers. Instead, Associates can simply pick an optimized aggregate collection of items to significantly reduce

time walking around, and another Associate can later sort the picked items into individual orders and pack the orders into cardboard boxes.

Roll/Sort

Finally, grocery totes are "rolled" closed and sorted onto the appropriate delivery truck route. This is similar to non-grocery FCs, where completed boxes are sorted by shipper.

B. Overview of produce operations in Seattle

The previous section described many ways in which grocery FC processes are more complex than non-grocery FC processes. This section further explains how produce-specific processes introduce additional challenges. As mentioned previously, there are several differences among the various AmazonFresh sites; this thesis primarily focuses on the processes in Seattle, and this section describes the produce processes as of the start of the project (February 2014).

In broad terms, produce is handled differently from most other products. It is received separately within the FC, inspected by an expert upon receipt, stored in separate rooms, and inspected regularly for quality and longevity. Additionally, several produce-related processes (e.g., cleaning/sanitizing, trimming) that take place in climate-controlled produce or trim rooms do not apply to other product categories. Furthermore, produce operations consist entirely of "specialized" processes: produce receive, replen, produce stow, produce pack, trim, inspection, banana process, and produce indirect. By "specialized," we mean processes that cannot, in theory, be accomplished without substantial incremental produce-specific training. Finally, produce is highly susceptible to damage (i.e., being squished, crushed, or bruised) and expiration (i.e., becoming spoiled, smelly, discolored, rotten, moldy, overripe, or infested).

These process differences are formalized by the organizational design. The Produce Team operates as a distinct FC department and has historically had a dedicated team of Produce Specialists and Produce Associates, who carry food handler permits (unlike FC Associates). In fact, until recently, the department was part of the Retail Team, which is responsible for relationships with vendors and customers. Now, the Produce Team is part of the FC Team but

reports up to the FC SME, not the FC Operations Manager. The Produce Team is comprised of Senior Produce Specialists, Produce Specialists, and Produce Associates.

Produce Receive

The Produce Receive function extends the FC Receive function and broadly consists of four steps: inspection, Advanced FC Receive, SKU transformation, and prep. Inspection will be described shortly. The Advanced FC Receive process is different from the standard FC Receive process because the product is received in a "batch" format rather than in a continuous format. Consequently, it is done by Specialists on their laptops using the Advanced FC Receive functionality (rather than using a handheld scanner). The SKU transformation converts a single case of a product into the actual quantity of the actual product in the computer system; this step involves physically counting the number of units of produce (e.g., limes) inside a case. The prep step varies by produce SKU; for example, pears are prepped by placing them individually inside foam sleeves. All cases are "prepped" by printing and sticking an SKU label on the case, writing the quantity on the case, and writing the date received on the case.

Replen

The current Replen process consists of picking the product from a "reserve" (i.e., "non-pickable") back-stock location, inspecting it, and then stowing it in a "pickable" location (i.e., one from which an FC Associate performing the "Pick" function is permitted to pick the product to fulfill a customer order). From a customer's perspective, this process is completely non-value added. However, it is necessary to maintain a separate back-stock if there is insufficient pickable bin space. There is no comparable FC process because there is generally sufficient bin space for other types of products in the FC.

Produce Stow

The Produce Stow process is a more complex version of FC Stow that requires some additional knowledge about durability and optimal temperature zones for different produce products.

Produce Pack

The exact Produce Pack process varies by SKU, but in broad terms it consists of inspecting the product and then putting it into a paper bag, plastic bag, plastic mesh bag, or clamshell. For example, a Produce Associate could create several 4lb mesh bags filled with oranges out of a single case of oranges. There is no comparable FC process.

Produce Trim

The exact Produce Trim process varies by SKU, but in broad terms it consists of inspecting the product, cutting off a part of the product (usually near the bottom) with a knife, and then washing/soaking the product in a sink so it "drinks up" the water and becomes crisper. Due to the relatively long cycle time of this process, trimming is generally done in large batches based on expected demand. There is no comparable FC process.

Inspection

Also called "rating," this process currently involves a Specialist using all five senses (including tasting samples) to make a qualitative judgment regarding the quality of the product and the product's longevity. These ratings are then communicated to customers on the AmazonFresh website. Inspections come in two forms: Receive Inspections and Bin Inspections. Receive Inspections take place when produce is initially received, and Bin Inspections take place at a Specialist-specified frequency once the product is stowed. More specifically, there are actually two types of produce SKUs: "shelf-life" and "inspection." Shelf-life SKUs (such as apples) are inspected during the Produce Receive process but not afterwards; instead, they are simply removed from pickable bins and discarded after a pre-specified number of days. Inspection SKUs (such as strawberries) are regularly examined, and the Specialist decides when to discard the product. There is no comparable FC process.

Banana process

The AmazonFresh website sells bananas by the bunch. Each bunch is expected to contain exactly five bananas. Consequently, the banana process consists of manually picking up a bunch of four to eight bananas from a case, tearing off and discarding bananas until the bunch consists

of five bananas, and then placing the bunch on a special rack for the picker. There is no comparable FC process.

Produce Indirect

This function is currently used by Specialists and Produce Associates to capture hours for several indirect produce processes, such as misting trim, addressing trouble tickets, writing / responding to rejection/credit e-mails, writing / responding to other e-mails (e.g., questions, requests for rejection photos), creating shift "pass-down" documents, creating other reports, removing empty boxes from the shelves, and cleaning workstations. There is no comparable FC process.

Summary

As this section illustrates, produce SKUs are handled quite differently in the FC compared to other grocery SKUs. Almost always, the complexity is higher, so training requirements must be higher when produce is involved.

Chapter 5:
Aligning labor use with workload with flexible labor

This chapter describes the initial implementation of flexible labor within the Seattle FC Produce Team. These changes led to an average 44% increase in productivity.

A. Observation: significant "slack" in Produce Team

During the first few weeks of the project, it became clear that there was a significant amount of "slack" built into the Produce Team. During visits to the produce room at the FC, it was uncommon *not* to see one or more idle Produce Associates. This observation was substantiated by both qualitative and quantitative data.

In terms of qualitative data, the Produce AM shared a recent anecdote of a time when a Produce Specialist did not come to work, but all the work got done, anyway. Another team member recalled a day the whole team went home early because there was no work; earlier the same day, another team member had already gone home. Across the team, it was common knowledge that certain days were "busy," while others were "for 5S."

On the quantitative side, a linear regression was run using data from each shift for the 8-week baseline period in 2014 (Weeks 6-13). The two independent variables were OB units shipped and IB cases received, and the dependent variable was hours of labor staffed on the shift. The regression found that the two key workload drivers (OB units shipped, IB cases received) only explained 33% of the variation in the labor hours staffed.

Next, an optimization was formulated to find a lower bound on the amount of excess labor (rather than to provide the best fit model as the regression did). The optimization minimized excess labor by adjusting two variables – "IB rate" and "OB rate" – subject to the constraint that all work was accomplished on each shift. (It was known that there were no shifts during the baseline period during which the Produce Team failed to finish their IB and OB work.) Thus, the output from the optimization was a minimum threshold of excess labor and the corresponding OB and IB rates. Comparing actual labor usage to the calculated labor requirement using these

excess labor minimizing OB and IB rates revealed that the amount of labor used during the 8 week baseline period was at least 22% more than necessary.

Finally, the model was further validated by projecting expected labor needs for upcoming shifts and then observing actual staffing levels and the amount of idle time. At the start of each shift, the contents of expected deliveries for the shift were known, so it was possible to apply the IB rate generated by the optimization to project the labor needed for IB work. Similarly, at the start of each shift, there was a forecast of OB units available; this was used in conjunction with the OB rate from the optimization to project the labor needed for OB work. These two projections together comprised the overall labor projection for the shift.

B. Countermeasure: labor projection for shift

Based on this finding, a hypothesis was generated that labor would be used more efficiently if labor needs were predicted at the start of the shift and then the shift leader's labor use was compared to the projection. A simple tool was constructed in Excel to provide a labor projection for each shift. The tool had two built-in constants – the IB rate and OB rate – from the optimization described previously. At the start of the shift, the user would need to input expected IB cases and forecasted OB units for the shift, and the tool recommended a staffing level ([expected IB cases ÷ IB rate] + [forecasted OB units ÷ OB rate]). The Produce AM strongly encouraged Specialists to adhere to projections and empowered Produce Specialists to "flex" Produce Associates to the FC (assuming the FC needed the help and the Produce Associate was properly trained) and offer VTO. While Specialists were still permitted to deviate from projections in extenuating circumstances, the Produce AM held shift leaders accountable for meeting labor projections or explaining deviations. It helped that several Produce Associates were already trained in many FC processes.

During implementation of the change, for two weeks (Weeks 16-17), the shift leaders received a great deal of support. Start-of-shift conversations were used to retrain shift leaders on using the Excel tool and to help them think through their "game plan" for the shift. Mid-shift check-ins by the FC Operations Manager were used to determine how the shift was progressing relative to the projections, identify barriers, and reinforce commitment to the change from the FC leadership.

End-of-shift conversations were used to gather feedback from the shift leader and from the FC Operations Manager on how the shift went. The lessons learned from each shift were communicated to the entire team by e-mail and at start-of-shift meetings. These frequent check-ins were also used to assess sustainability of the change and team morale. Over the subsequent months, a weekly call was set up with the FC Produce AM and his counterpart in Corporate to discuss the implementation of the change and the past week's productivity data.

Several incentives were also provided to support the change. The FC Site Leader allocated funds for the Produce AM to buy lunch for the team to show support for their help. Also, the Produce AM had conversations to remind the team about how performance was linked to compensation. Additionally, managers were provided with reward credits to recognize team members that were performing exceptionally well; the credits could be redeemed for company apparel. Concurrently, a Senior Produce Specialist position was made available, so the best shift leader had the potential to be promoted. Finally, a great deal of encouragement was provided, and mistakes during the implementation were not penalized but were viewed as learning opportunities.

C. Result: 44% productivity improvement

In the six weeks following the implementation of the labor projections, productivity (measured as OB units per paid labor hour, excluding meetings, breaks, and lunch) improved significantly. The average productivity in the fifth and sixth weeks after the change was introduced (Weeks 18-19) was 44% higher than the average productivity during the baseline period (Figure 2).

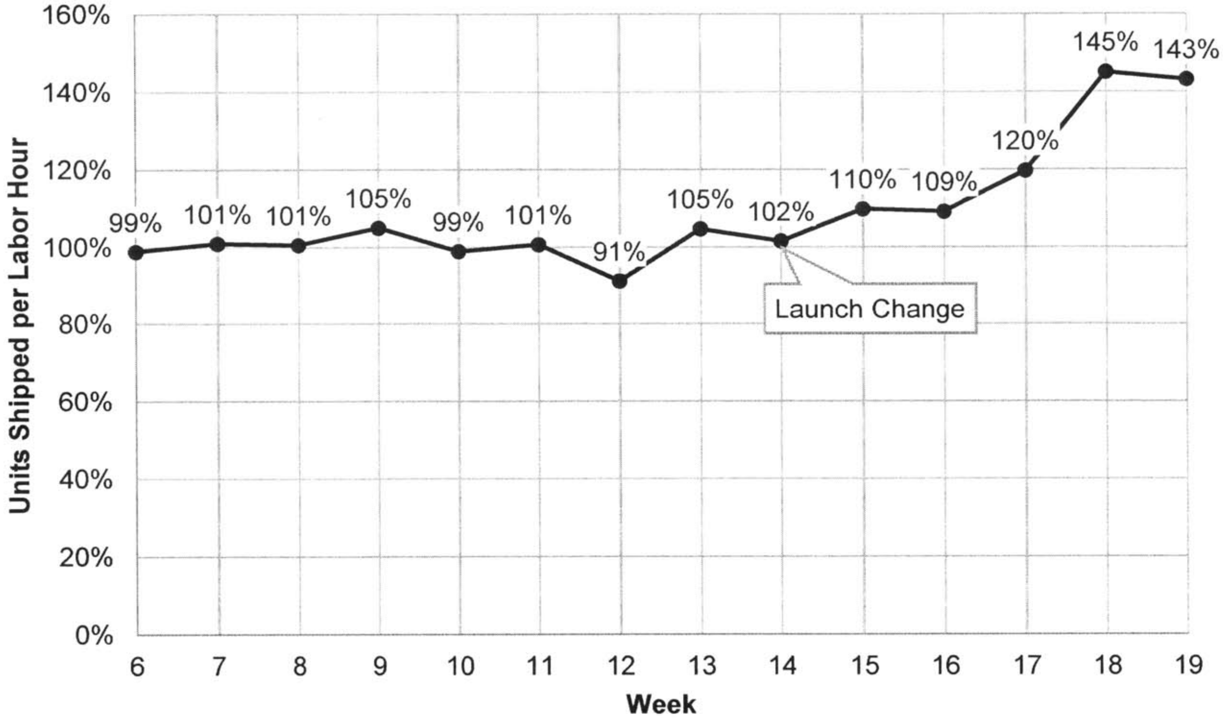

Figure 2: Productivity, Weeks 6-19 (% of baseline)

Furthermore, in these two weeks the two major workload drivers, IB cases and OB units, explained 85% of the variation in labor hours used. The increase in productivity exceeded expectations because some shift leaders sought to out-perform projections instead of simply follow them. Two shift leaders in particular took a great deal of pride in being able to run an efficient shift and beat the targets. Indeed, over time, the team consistently began to staff at or under projections and achieve performance better than during the baseline period (Figure 3, Figure 4).

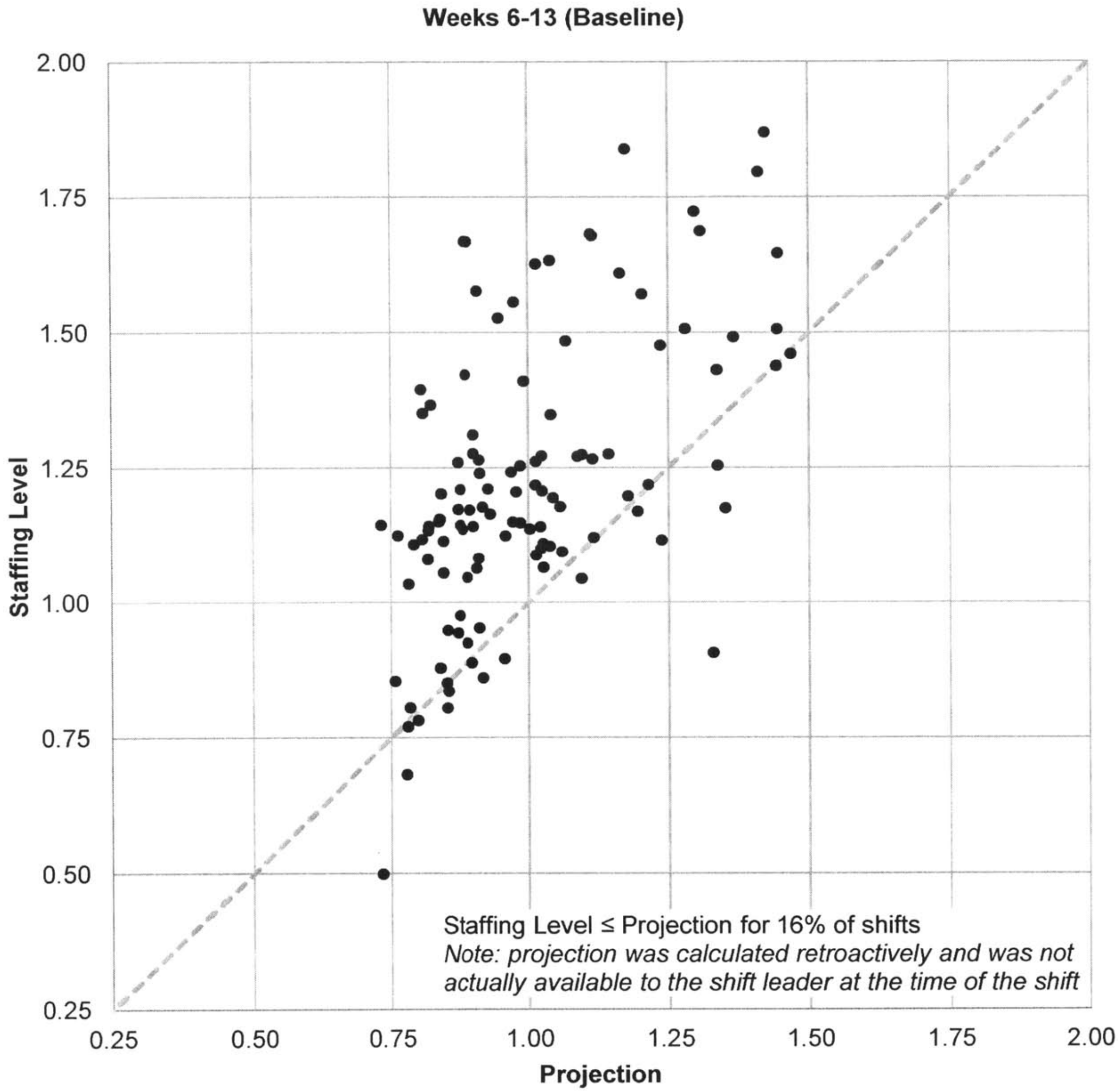

Figure 3: Staffing Level vs. Projection (Baseline); each dot represents one shift; normalized

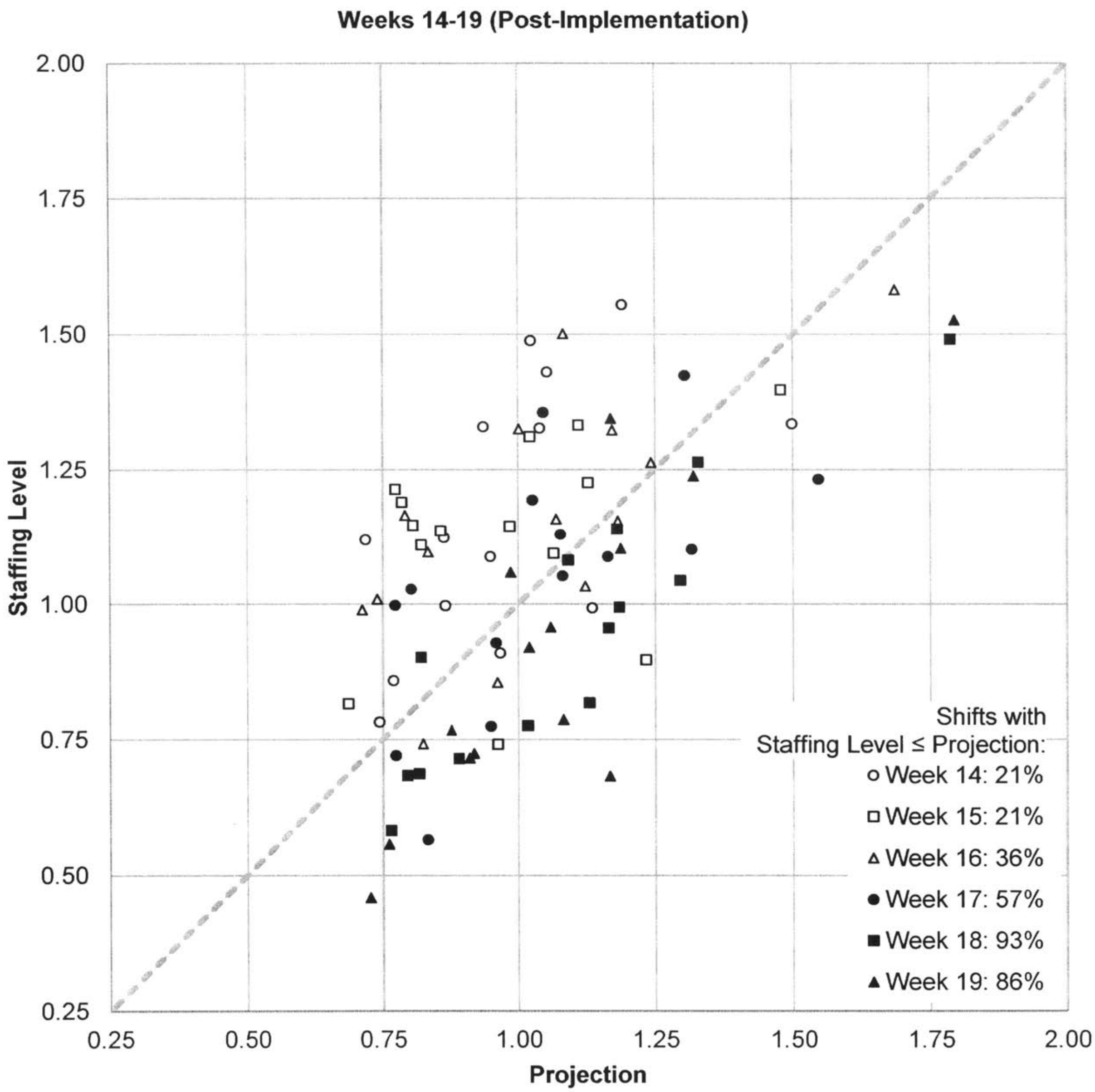

Figure 4: Staffing Level vs. Projection (Post-Implementation); each dot represents one shift; normalized

It was possible to perform better than the projections because it turned out that there was actually more slack originally present in the system after accounting more accurately for the true IB and OB rates. The projections were intentionally set using the optimization result that set a lower bound on the amount of excess labor, because it was determined that implementation would be significantly easier if the projections were initially easily achievable. This was indeed the case, and in fact, setting a target that was easily achievable but required a higher level of productivity

than before improved team morale, while an overly aggressive target would likely have worsened morale and made implementation more difficult.

Chapter 6:
Capturing benefit of a redesigned product with flexible labor

This chapter describes the subsequent refinement of flexible labor within the Seattle FC Produce Team. These changes led to an additional average 16% increase in productivity, measured relative to the initial baseline.

A. Observations: Banana process wasteful, team out-performing projections

After fully implementing flex labor in the Produce Team, the project focused on other ways to reduce labor use. The banana process became the next focal point for the project because all parties saw significant room for improvement.

As previously described, the banana process initially consisted of modifying and/or discarding bunches of bananas until all remaining bunches had five bananas. These bunches would be placed in trays on a special cart for the picker. The process was very wasteful. For example, consider a box of bananas that contains an equal number of bunches that have four, five, six, and seven bananas. In this scenario, the bunch of four would be discarded, the bunch of six would have one banana torn off, and the bunch of seven would have two bananas torn off. In total, seven out of the twenty-two bananas – nearly a third – would be discarded in advance. In addition to the obvious financial impact of this, many Associates found the process frustrating since it was highly repetitive, and they did not feel they were creating value for the customer.

Simultaneously, once flex labor practices and labor projections were successfully implemented, shift leaders started regularly out-performing projections, indicating the presence of additional slack in the system. Surprisingly, there was also enthusiasm from the team to improve efficiency and have more accurate projections.

B. Countermeasure: Redesigned product and updated projections

In response to these observations, the company decided to transition from selling bunches of five bananas to selling two-pound bags of bananas. By weight, this actually slightly increased the amount of product that the customer received. However, this simultaneously eliminated the need for the Produce Team to process the bananas individually, since major banana vendors were

willing to sell pre-packaged two-pound bags of bananas. The only remaining banana-related task for the Produce Team would be to receive them and ensure that they met temperature and color/ripeness specifications.

Following the product redesign, labor tracking data from after the implementation of flex labor was used to estimate the amount of time that would be saved due to the prepackaged bananas, and the Excel tool was updated to reduce the labor projections by that amount. Additionally, to account for the fact that shifts were out-performing projections more generally due to additional slack, the Excel tool's built-in assumptions for IB and OB rates were revisited and updated. The goal in doing this was to still make it possible to regularly hit the projections, but to make it difficult to significantly out-perform them.

C. Result: Additional 16% productivity improvement

Figure 5 shows the immediate impact of the redesigned banana product and the corresponding updates to the labor projections. Conventional bananas were switched to the prepackaged format on 5/14, and organic bananas were switched to the prepackaged format the next week on 5/19; the updated labor projection tool went into use on 5/21. Interestingly, for several of the days in between when the prepackaged conventional bananas arrived and when the tool was updated, the number of hours spent processing bananas did not materially go down. This suggests that the shift leaders were frequently allocating an Associate to process bananas for the same amount of time as before, even though the banana workload was reduced by more than 50%. This further suggests that the lower labor projections from the updated tool helped shift leaders realize that they could allocate less labor for banana processing.

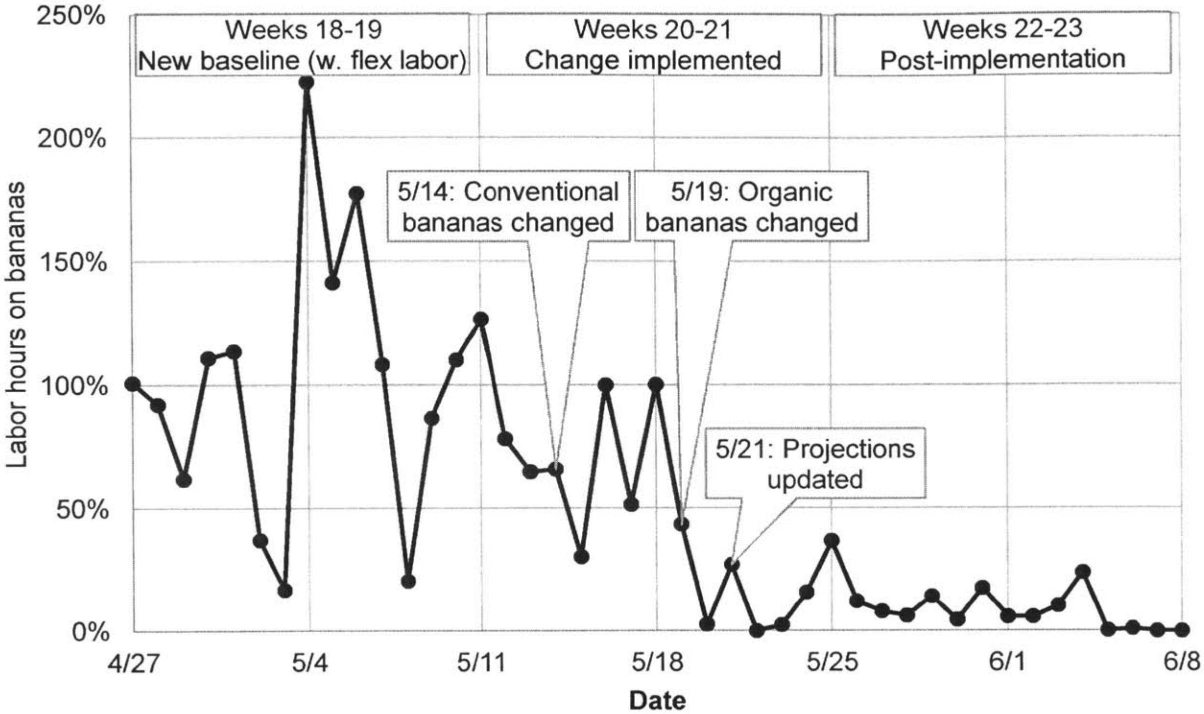

Figure 5: Banana Labor Hours (% of new baseline from Weeks 18-19)

Once the amount of time spent on banana processing went down to a new steady state, it stayed there (Figure 6). Within a week, the updated labor projections, together with the flex labor practices, made it possible to realize nearly the full labor savings from the redesigned product, which resulted in an incremental 4% net productivity increase across all produce.

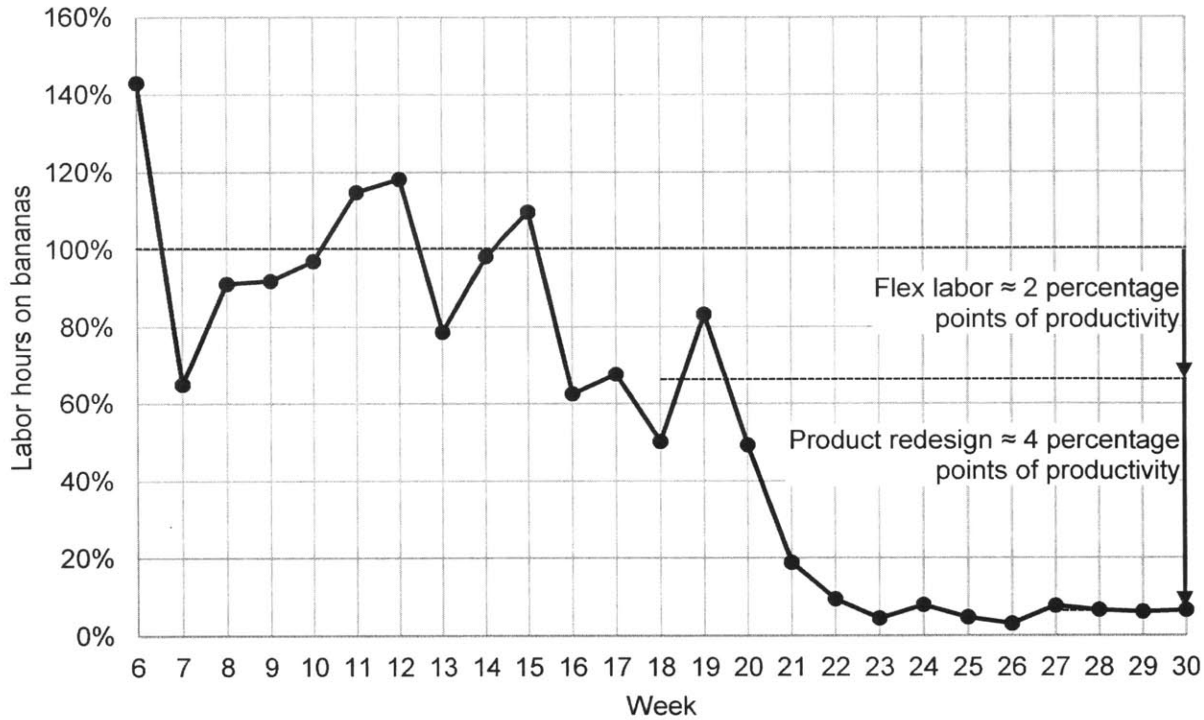

Figure 6: Banana Labor Hours (% of original baseline from Weeks 6-13)

The success of prepackaged bananas prompted the beginning of a broader transition towards prepackaged produce. At the same time, the Excel tool's built-in IB and OB rates were updated to better align with actually observed work rates. As shift leaders adjusted their staffing behavior to meet the new projections, productivity increased an additional 12%, relative to the original baseline period (Figure 7).

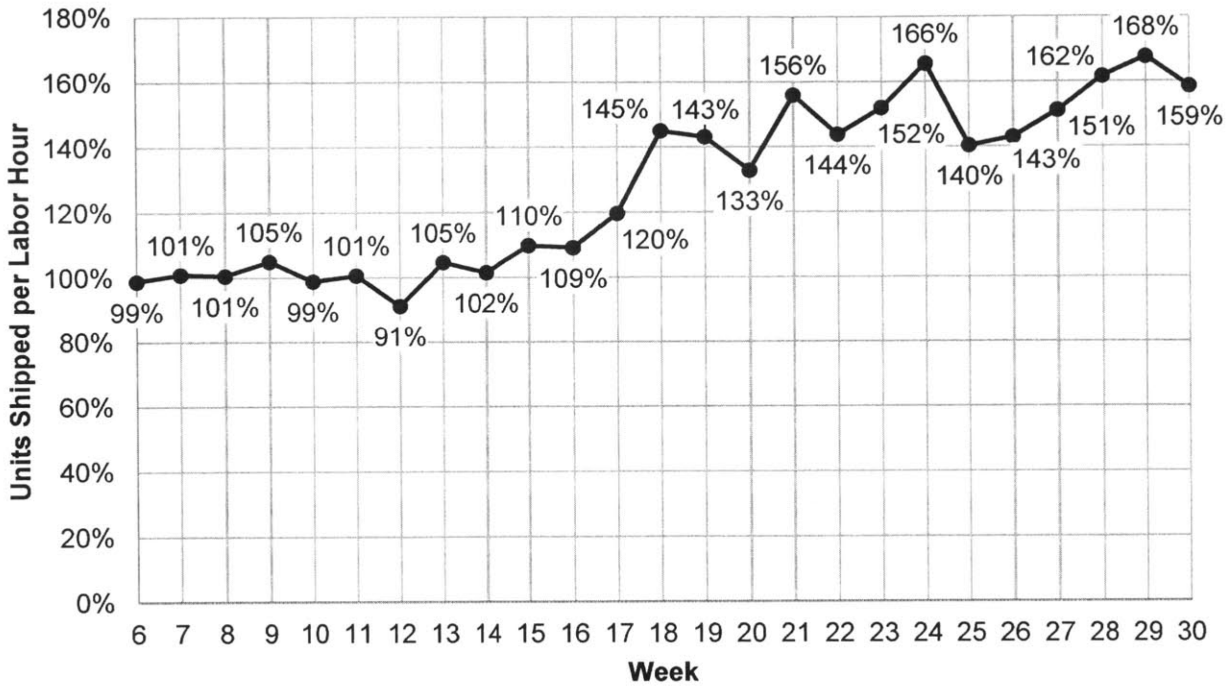

Figure 7: Productivity, Weeks 6-30 (% of baseline)

Put another way, the average productivity at the end of the project (Weeks 27-30) was 160% of the productivity during the baseline period (Weeks 6-13); 44% of this increase was observed during the initial implementation of flex labor, 4% was attributed to changes in the Banana process, roughly 11% was attributed to the changes in the Produce Pack process, and the remaining 1% was due to other initiatives. In general, shift leaders were able to hit the new projections. Figure 8 contains an overlay of the shift-by-shift performance data for this period and the original projections (for ease of comparison to Figure 3 and Figure 4). It is worth noting that the gains from flex labor came disproportionately from flexing Produce Associates (as opposed to Produce Specialists). Specialists preferred VTO over flexing into the FC, so shift leaders flexed Associates into the FC but offered VTO to Specialists if there was slack. However, if there was a significant excess of Specialist labor available, and Specialists did not

want to take VTO, then shift leaders had difficulty flexing Specialists into the FC. In these cases, the shift leader generally exceeded the labor projection and attributed the miss to "too much labor available." Change resistance began to manifest itself towards the end of the project as Produce Specialists started to feel that they would soon need to be flexed into the FC if productivity improvements continued.

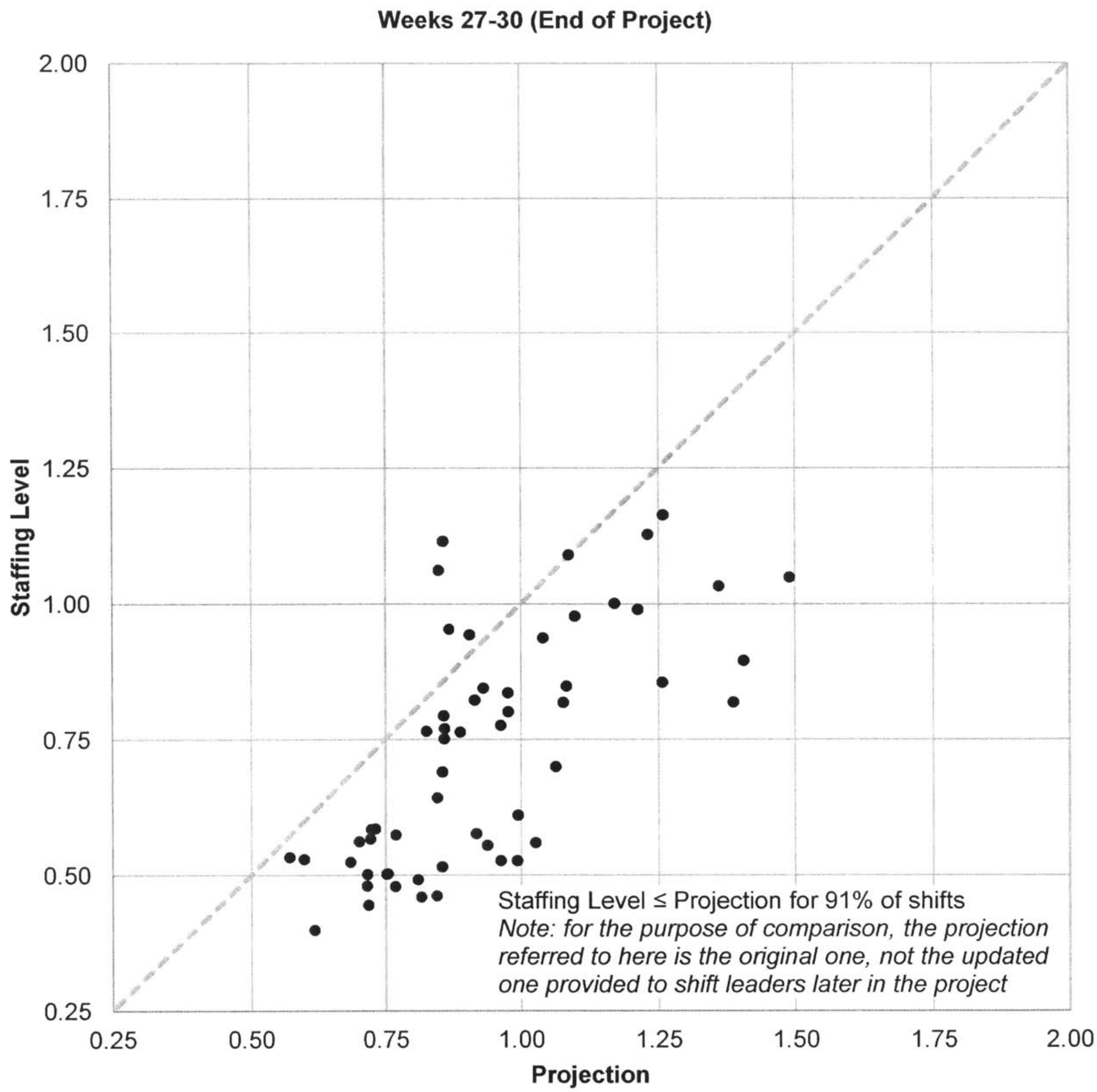

Figure 8: Staffing Level vs. Projection (End of Project); each dot represents one shift; normalized

Chapter 7:
Enabling product quality improvement with standard work

We turn now to the topic of product quality, which was the focus for much of the second half of the project; in particular, the objective was to understand what is meant by produce product quality and how to improve it. A key finding early on was that vendor performance was the primary root cause of product quality issues; that is, these units were already defective when they arrived at the FC. Still, through inspections, the Produce Team could catch and eliminate these defective units to ensure that customers only received good quality produce. This chapter discusses the application of standard work to enable better inspections in the FC.

A. Observation: High aggregate variability makes it difficult to prioritize or improve

The challenge of defining "quality" in the context of produce became apparent early on. The aggregate metric tracked for the project was units refunded per units sold, where the reason for the refund was product quality related; however, there were several caveats. First, for information system reasons, it was not always possible to measure the number of units refunded with 100% accuracy. Second, the call center and website encoding of the reason for the refund was not always entirely reliable; for example, a customer could easily describe an overripe unit of produce as "rotten." Third, the metric did not capture other dimensions of quality that customers perceived but were unlikely to complain about: size, shape, color/bruising, firmness/texture, and taste. Fourth, overall, there were few complaints, so there was not an abundance of data available for the analysis. For these reasons, in addition to the aggregate "quality" metric described, the customer's or call center representative's long-form comments were also monitored to gain a better understanding.

Customers complained about a wide range of issues and articulated their complaints in many different ways, most often complaining that produce was "moldy" or "rotten." Some categories of produce, such as berries, received relatively many complaints, while other products, such as dried fruit, received none. When a customer complained, a "trouble ticket" was created for the Produce Team, and a Produce Specialist was dispatched to the relevant bin in the warehouse to address any issues if more defective produce is in the same bin. However, there was no longer-term tracking at the FC for quality.

In addition to following up on customer complaints, the Produce Team also inspected produce as part of the normal process, assigning it both a "quality rating" and a "longevity rating." The quality rating was out of 5 (5: fantastic, 4: great, 3: good, 2: average, 1: below average), while the longevity rating was used to determine when the produce would be put back into the inspection queue. However, an analysis of these ratings found several opportunities for improvement. By far, nearly all produce was assigned a quality rating of "3." Furthermore, the same batch of produce was occasionally given differing ratings (even by the same Specialist) within a short period of time, reflecting a high level of subjectivity. Moreover, there were often quality defects found in bins that had been recently inspected, suggesting that inspections often missed important defects; Specialists conceded that sometimes they missed defects because they were not looking closely since they had "just checked that bin yesterday." There were also opportunities to improve the longevity ratings; an analysis found that nearly all produce was assigned the shortest longevity rating (regardless of SKU), which meant that the Produce Team had to do very frequent inspections for almost all products. Finally, each Produce Specialist had a unique way of doing their ratings: some would taste the produce, others would simply cut it open and look at it, others would pick it up and feel it, and others would just look at it as they walked by. From conversations, it emerged that many of the Specialists had had no formal training in inspecting produce; in fact, the FC training team did not have the capability to train on this process but rather would delegate that responsibility to the Produce Team.

B. Countermeasure: Standardize inspection process, rationalize inspection frequency

These observations collectively led to two hypotheses: (1) documenting and standardizing the inspection process would improve the consistency of inspection outcomes, and (2) quality could be improved by doing inspections more carefully, even if frequency of inspection was reduced (or eliminated).

To standardize the inspection process, the various Specialists were interviewed about the processes that they had been using, and a list of best practices was compiled. Quality standards published by the United States Department of Agriculture were used as references during these conversations. Then, a workshop was arranged with the Produce AM to establish an initial draft of the "standard work" for the inspection process. The list compiled from the team members was

used as an input in the conversation. This initial draft was then shared with various groups of internal experts, and in a few cases, some updates were made based on feedback. Concurrently, the Produce AM's Corporate counterpart set up a weekly "Produce University" call, which included a representative from each FC's Produce Team. The call was used to educate about quality across sites, and the members of the call also worked on in-depth flash cards by produce product category to train their respective teams on "what to look for."

With regard to the second hypothesis about the trade-off between inspection frequency and thoroughness, a "no rating pilot" was planned beginning 7/3/2014; historical data was used to identify categories of produce that potentially did not need to be inspected after being inspected initially upon receipt, allowing Specialists to spend the time saved on the remaining categories, which had a history of more quality issues. The categories selected for the pilot were: Beans & Peas, Cabbages, Dried Fruit, Eggplant, Fresh herbs (excluding basil), Mushrooms, Packaged Carrots, Potatoes and Yams, Root Vegetables, Specialty Peppers, and Sprouts & Grasses.

C. Result: Successful pilot, but stalled standard work implementation

Overall, the "no rating pilot" was successful along certain dimensions, with one exception. Across all categories in the pilot (except Potatoes and Yams), the frequency of complaints did not materially differ from the pre-pilot frequency, supporting the hypothesis that these categories did not need daily inspections. There were many more complaints for the "potatoes and yams" category during the pilot; however, the vendor for this product category was also changed during the pilot, and it was obvious during visits to the FC that the new vendor was sending the FC product that was of much lower quality than the original vendor.

While the data from the pilot supported removing most of these produce categories from the inspection queue, there was not a noticeable drop in the number of complaints on the other products. This may be in large part due to challenges faced during the implementation of standard work, which is discussed next.

Implementation of standard work for quality inspections was much less successful than implementation of flex labor; surprisingly, instead of reallocating the time saved to inspect the

remaining products more thoroughly, the team simply reduced the amount of hours they spent on inspection. Furthermore, upon closer examination, it appeared that, even though standard work was now defined for all the SKUs, the Specialists still continued to inspect as before instead of following the defined procedure.

There are a few hypotheses for why this implementation was less successful. First, there was a perception within the FC Produce Team that Corporate did not actually care about produce quality. In the Produce Specialists' minds, this perception was reinforced by vendor choice, communications from the Corporate Retail team, lack of metrics, and lack of links to incentives. Second, there was much less in-person on-site support during implementation of standard work for quality inspections. During the implementation of flex labor, there were two people from Corporate fully dedicated to supporting the Produce AM during the change; in contrast, the implementation of the standard work for quality inspections was largely delegated to the FC team and in particular to the Produce AM. Finally, it is possible that messaging from Corporate became confusing; this quality initiative seemed to be at odds with several productivity initiatives that were still ongoing.

Chapter 8:
Conclusions and Recommendations

In summary, this project improved the Seattle FC Produce Team productivity 44% by implementing flexible labor, improved the productivity an additional 16% by using flexible labor to capture the labor savings from a redesigned product, identified several produce categories that did not require inspection, and established a starting point to standardize inspection processes. In terms of broader knowledge creation, this thesis has demonstrated the applicability and utility of flexible labor and standard work in the online grocery fulfillment industry. To wrap up, we reflect on possible next steps for the Seattle FC, for AmazonFresh more broadly, and for future research.

A. Recommendations on produce processes

Once the project focus shifted to quality, it became evident that the majority of quality issues were not arising at the FC itself. Rather, defective produce was delivered to the FC, and the best thing the FC could do was to reject a delivery or discard the portion of a delivery that was defective. Consequently, the first recommendation is to take a closer look at the Produce Receive process; this is the first step in the FC's produce processes and serves as the "first line of defense" against quality issues. Once defective produce makes it into the pickable bins, it is much more difficult to find and address issues.

Second, the effort to redesign products to reduce the downstream labor requirement should continue. In particular, there are several "trim" products (such as celery and leafy greens) that may be available in alternative forms of packaging. This would reduce or eliminate the need to directly cut and wash produce within the FC, thereby reducing food safety risk; additionally, this would further simplify processes within the FC.

Third, the company should continue to implement standard work in inspections and not allow the Produce Team to be exempt from clearly documented processes. Produce already presents many complications and requires a higher degree of training than other FC functions. However, the difficulty of tackling these challenges is increased when the Produce Team becomes a black box to the rest of the FC.

Fourth, the Produce Team should continue to experiment with reducing bin inspection frequencies on certain categories of produce. The "no rating pilot" demonstrated the value of inspecting different categories of produce at different frequencies; however, this pilot was binary in the sense that it called for produce to either be inspected or not. Once more sophisticated tools are available to calibrate inspection frequency, it would be worthwhile to experimentally determine the optimal inspection frequency for each produce category.

Fifth, to sustain the progress that has been made, it is critical to embed the quality and productivity metrics that have been used for this project into the weekly metrics that are reported across the network. This will help ensure that leadership quickly becomes aware of any changes and can become involved if there are any new barriers that arise.

Finally, the changes demonstrated at the Seattle FC should be adapted as appropriate and rolled out to all other AmazonFresh sites. During the project, due to time constraints, the tools developed in Seattle were shared with all sites, but there was not a focused effort to ensure that all sites were taking steps to adopt these new best practices.

B. Recommendations for AmazonFresh leadership

Additionally, there are several changes that can be made centrally to support the change effort that this project has started.

First, it is important to create an incentive system that supports productivity and quality improvements. In particular, this could include ensuring job security for Associates who participate in improvement initiatives. Additionally, this could include establishing a direct link between Associates' compensation and their team's productivity and work quality.

Second, in terms of organizational design, roles should be created within the central organization to provide support for process improvements. The stalled implementation of standard work demonstrated the importance of in-person support at the FC during the implementation. These individuals could also assist in instilling Amazon's culture of continuous improvement across the AmazonFresh organization.

Third, in terms of site design, volume and service level requirements should be taken into account when designing new sites. This will, for example, ensure enough bin space so that it becomes unnecessary to have a separate back-stock and a non-value-added "replen" process. Also, standardizing the site design will make changes more easily transferable across sites; currently, the different designs at each site require a great deal of adaptation and customization of changes.

C. Remaining questions for further research

Last, but not least, this project has answered some questions but surfaced many more. There are a couple in particular that may be rich sources of future research.

First, as mentioned previously, it was not possible to truly modulate inspection frequency and observe the impact on the consistency of quality inspections. However, due to the human behavior dynamics, this seems to be a particularly interesting area for future research. Amazon would be uniquely positioned to assign different inspection frequencies to different sites or different shifts at the same site and then observe differences in performance.

Second, this project generally did not exploit recent advances in technology. However, there are several technologies including machine learning, optical and chemical sensors, robotics, and RFID that could be applied to, for example, improve the consistency of quality processes, reduce human handling and conveyance, reduce cycle times, and ensure chilled chain compliance. A study that holistically examines the range of technologies available to improve or replace existing produce processes would likely yield some very interesting findings.

References

Amazon Annual Reports (1997, 2007, 2010, 2013). Accessed December 16, 2014 at http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-reportsAnnual.

"Annual Report for the 52 weeks ended 1 December 2013." Accessed December 16, 2014 at http://www.ocadogroup.com/investors/reports-and-presentations/2013.aspx.

Cachon, Gerard and Christian Terwiesch. Matching Supply with Demand: An Introduction to Operations Management (Second Edition). *McGraw-Hill*, 2009.

Chanil, Debra. "Seizing the Day | Progressive Grocer." May 1, 2014. Accessed December 16, 2014 at http://www.progressivegrocer.com/node/68197.

"Delivery areas - Express Help." Accessed December 16, 2014 at https://support.google.com/shoppingexpress/answer/4559799?hl=en.

Easton, Fred F. "Cross-training performance in flexible labor scheduling environments." *IIE Transactions* (2011) 43, 589–603. Accessed on December 20, 2014 at http://www.tandfonline.com/doi/pdf/10.1080/0740817X.2010.550906.

Francas, David, et al. "Machine and Labor Flexibility in Manufacturing Networks." February 11, 2010. Accessed December 20, 2014 at https://plis.univie.ac.at/fileadmin/user_upload/abt_logistic/Working_Papers/Francas__Loehndorf__Minner__2011___Labor_Flexibility.pdf.

"FreshDirect - Help – FAQs." Accessed December 16, 2014 at https://www.freshdirect.com/help/faq_home.jsp.

Glasner, Joanna. "Why Webvan Drove Off a Cliff." July 10, 2001. Accessed December 16, 2014 at http://archive.wired.com/techbiz/media/news/2001/07/45098.

Harris, Craig and John Cook. "Amazon starts grocery delivery service." August 1, 2007. Accessed December 16, 2014 at http://www.seattlepi.com/business/article/Amazon-starts-grocery-delivery-service-1245445.php.

Hopp, Wallace J. and Mark P. Van Oyen. "Agile Workforce Evaluation: A Framework for Cross-training and Coordination." *IIE Transactions* (2004) **36**, 919-940. Accessed on December 20, 2014 at http://webuser.bus.umich.edu/whopp/reprints/Agile%20Workforce%20Evaluation.pdf.

Horsey, Julian. "AmazonFresh Grocery Delivery Service Now Available In Los Angeles." June 10, 2013. Accessed December 16, 2014 at http://www.geeky-gadgets.com/amazonfresh-grocery-delivery-service-now-available-in-los-angeles-10-06-2013/

"Instacart - Same-day Grocery Delivery in Atlanta, Austin, Boston, Chicago, Denver, Los Angeles, New York City, Philadelphia, San Francisco Bay Area, Seattle, and Washington, D.C." Accessed December 16, 2014 at https://www.instacart.com/faq.

Lunce, Stephen E., et al. "Success and failure of pure-play organizations: Webvan versus Peapod, a comparative analysis." *Industrial Management & Data Systems*, Vol. 106, No. 9, pp. 1344-1358, 2006.

Martinez, Amy. "AmazonFresh set to expand?" January 26, 2013. Accessed December 16, 2014 at http://seattletimes.com/html/businesstechnology/2020208438_amazonfreshxml.html.

Ohno, Taiichi. Toyota Production System. Chapter 2, pp. 21-22. *Productivity Press*, 1988. English translation of Toyota seisan hoshiki, published by *Diamond, Inc.*, 1978.

Rigby, Darrell, et al. "Digital Darwinism: Winning with the best of digital and physical." *Bain Retail Holiday Newsletter*, Issue 4, December 19, 2013.

Sandoval, Greg. "Webvan to close Sacramento operations." April 24, 2001. Accessed December 16, 2014 at http://news.cnet.com/Webvan-to-close-Sacramento-operations/2100-1017_3-256370.html.
Senauer, Ben and John Seltzer. "The Changing Face of Food Retailing." *Choices: The Magazine of Food, Farm, and Resource Issues*, a publication of the Agricultural & Applied Economics Association. 4th Quarter 2010 | 25(4). Accessed December 16, 2014 at http://core.kmi.open.ac.uk/download/pdf/6220630.pdf.
Slomp, Jannes, et al. "Cross-training in a cellular manufacturing environment." *Computers & Industrial Engineering* 48 (2005) 609–624.
Soper, Taylor. "Amazon Fresh launches in San Francisco with $299/year 'Prime Fresh' membership." December 11, 2013. Accessed December 16, 2014 at http://www.geekwire.com/2013/amazon-fresh-san-francisco/.
"Standardized Work: The Foundation for Kaizen." *Lean Enterprise Institute*, 2009. Accessed December 20, 2014 at http://www.lean.org/Workshops/WorkshopDescription.cfm?WorkshopId=20.
"The stores on our site - Express Help." Accessed December 16, 2014 at https://support.google.com/shoppingexpress/answer/4562192?hl=en.
"6.3.3. What are Attributes Control Charts?" NIST/SEMATECH e-Handbook of Statistical Methods. Created June 1, 2003; updated October 30, 2013. Accessed December 21, 2014 at http://www.itl.nist.gov/div898/handbook/pmc/section3/pmc33.htm.
Whole Foods Market, Inc. Form 10-K, 2014. Accessed December 16, 2014 at http://assets.wholefoodsmarket.com/www/company-info/investor-relations/annual-reports/2014-WFM-10K.pdf.
Womack, James P. and Daniel T. Jones. Lean Thinking. Chapter 6, p. 113. *Simon & Schuster*, 1996.