Methods for Extension of Ground Shipment Windows through a Supplier Collaboration Initiative

by

David Barry Wible

B.S. Industrial Engineering, Northwestern University, 2001

Submitted to the MIT Sloan School of Management and the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Engineering Systems

In conjunction with the Leaders for Global Operations Program at the Massachusetts Institute of Technology

June 2011

Signature of Author _____
Engineering Systems Division, MIT Sloan School of Management
May 6, 2011

Certified by _____
Chris Caplice, Thesis Supervisor
Executive Director, Center for Transportation and Logistics

Certified by _____
Stephen Graves, Thesis Supervisor
Abraham Siegel Professor of Management, MIT Sloan School of Management

Accepted by _____
Nancy Leveson, Chair, Engineering Systems Division Education Committee
Professor, Aeronautics and Astronautics and Engineering Systems Division

Accepted by _____
Debbie Berechman, Executive Director of MBA Program
MIT Sloan School of Management

This page intentionally left blank.

Methods for Extension of Ground Shipment Windows through a Joint Supplier Collaboration Initiative

by

D. Barry Wible

Submitted to the MIT Sloan School of Management and the Engineering Systems Division
on May 6, 2011 in Partial Fulfillment of the Requirements for the Degrees of
Master of Business Administration and
Master of Science in Engineering Systems

Abstract

Amazon.com is a leading retailer and UPS is the leading package delivery company in the world. Amazon spends hundreds of millions of dollars each year shipping products to customers via parcel carriers such as UPS. Arguably, both companies rely on each other for their future success and are therefore interested in developing a closer relationship. But supplier collaboration efforts are challenging and partnerships between industry leading companies are not common. This research paper demonstrates a case study where two industry giants are working together on a process improvement projects to better serve their customers.

This project has a goal of finding opportunities to delivery and sort UPS Ground packages more efficiently so as to allow for a longer window of time each day that the UPS Ground ship method is available. Additional goals exist around implementing more consistent processes and delivering product to UPS earlier in the evening for more efficient sorting. This project demonstrated that significant transportation savings can be achieved through these process improvements. Amazon and UPS are able to work together in a productive way to strengthen their long term relationships.

Thesis Supervisor: Chris Caplice
Title: Executive Director, Center for Transportation and Logistics

Thesis Supervisor: Stephen Graves
Title: Abraham Siegel Professor of Management, MIT Sloan School of Management

This page intentionally left blank.

Acknowledgments

I would like to thank Amazon.com and UPS for sponsoring such an educational internship project. All the people that work at the Amazon IND1 fulfillment center and the UPS 81st St Hub in Indianapolis have been extremely generous. I had the opportunity to work for two world class operations and will benefit from these experiences for many years to come. In particular, I would like to thank the following people who helped guide me along the way:

- Jared Butts, LFM '09, for being my mentor and giving me excellent advice on how to manage an LGO internship and prepare for a thesis.
- Phil Landis, John Lowe, and all the other managers in Amazon operations who taught me the ropes and who put up with the adverse effects of my pilots which helped make this thesis possible.
- Dave Alperson, Matt Smith, and all the senior leadership at IND1 for being supportive of me while I was at their facility and giving me useful coaching on both a professional and personal level. This project would have made little progress without their strong support.
- The Amazon Transportation team including Bill Templeton who helped teach me about logistics.
- Dave Hehmann and all the team at 81st St who provided me generous access to their hub and who went out of their way to teach me so much about the parcel industry.
- Steve Mundy, for his willingness to work with my school and experiment with this project. I would not have had this learning opportunity without your support.

I would also like to thank my thesis advisors, Chris Caplice and Stephen Graves, for their guidance along the way. Their support was critical for letting me know I was heading in the right direction.

A thanks also goes out to the faculty and staff of MIT and the Leaders for Global Operation (LGO) program for the opportunity to be apart of such an amazing academic program and for bringing together the greatest group of classmates anyone could ask for.

Finally, a special thanks to my family for their constant love and support. Without all that you have done for me over the years, I could not be where I am today going through such a positive experience.

This page intentionally left blank.

Note on Proprietary Information

Note that much of the data presented throughout this thesis has been modified in order to maintain confidentiality. The figures and results do not represent the actual values that were measured or calculated. Many numbers have been either skewed or scaled in an effort to disguise and preserve any proprietary information of both Amazon.com and UPS while still representing the general trend or pattern the data is intended to represent.

This page intentionally left blank.

Table of Contents

List of Figures

List of Equations

List of Tables

1 Introduction

The primary object of this paper is to present a case study of two large, industry-leading companies that are working together in a coordinated fashion to improve operational processes with the goal of delivering a better overall experience to their customers.

Amazon.com (Amazon) is a leading retailer in the online space while UPS is the leading parcel delivery service in the world. Transportation expenses make up the largest expense for the Amazon fulfillment process making this a huge priority for Amazon. Likewise, the package volume that Amazon provides is quite large which is beneficial for providing UPS efficient delivery density but also potentially disruptive to their network if not properly managed. Each company is strategically important to the other's success, and therefore both firms wish to maintain positive relationships. The forces driving this desire to work well together are only getting stronger. Amazon spent over $1 billion dollars on transportation services last year, and with their 30% annual growth rate in sales, this number is expected to grow dramatically over the next few years [3]. Therefore, it makes sense for the retail and parcel delivery giants to not only maintain positive relationships but to truly collaborate to find better ways to work together and to learn from each other.

More collaborative relationships with partners are becoming increasingly common. Harris et al states, "In the future, your facility won't be competing against your competitor's facility; your supply chain will be competing with your competitor's supply chain" [12]. Yet this collaborative process is not easy, and collaboration between two industry leaders presents its own challenges. Market leaders are accustomed to using their power to have other companies tailor to their needs. In this paper we hope to show that two industry leaders can and should work closely together and that it is in their best interest to do so.

1.1 Problem Statement

The internship's goal was to extend as late as possible the window that UPS is able to receive Ground shipments from Amazon. This will allow for an increased percentage of packages to be delivered via a more efficient Ground Service as opposed to Air. To do this, the strategy was to remove waste from the processes of transitioning packages from Amazon to UPS and reduce cycle time. The belief was that Amazon would not only be able to ship packages later but that UPS would be able to begin processing more packages earlier. In addition, the amount of variability around timings of the events could be reduced to make for a more predictable process all around. The approach was to use Lean methodologies and other Operations Strategy principles to analyze and improve the process.

1.2 Background and Basis for Thesis

The basis for this thesis lies in the six month internship that the author completed at an Amazon.com distribution center and the UPS package hub in Indianapolis, Indiana. This internship and thesis are each parts in the required curriculum of the Leaders for Global Operations program at the Massachusetts Institute of Technology. Amazon.com has been a long standing partner with the LGO program. UPS is currently not a partner with the LGO program, but they are using this joint internship with Amazon.com to learn more about the program and to explore how a relationship with MIT might fit with their strategic goals.

1.3 Thesis Structure

Chapter 2: Company Backgrounds and Operational Processes

This section begins with the general background of both Amazon and UPS followed by a brief discussion on company cultures. The remainder of the section discusses details of the operational processes involved with this project.

Chapter 3: Literature Review

This section reviews findings from previous research related to the topics discussed in this paper.

Chapter 4: Research Methodology and Approach

This section describes the approach taken to analyze and complete this project

Chapter 5: The Amazon and UPS Collaboration

This section reviews details about the Amazon/UPS collaboration including what makes a good partner and a good project. It also does a detailed benefits analysis for each company for the partnership in general and the project specifically.

Chapter 6: Process Changes

This section describes details of what process changes were made to improve operations and to allow for later cutoff times for Ground shipments.

Chapter 7: Implementation and Results

This section outlines some of the challenges around implementation of the changes and reviews results found in the multiple pilots

Chapter 8: Future Work

This section lays out recommendations for future work to sustain changes and continue improvements towards the project goals.

2 Company Backgrounds and Operational Processes

The following sections describe the industry, history, and current market conditions of each of the companies that participated in this project. In addition it has a brief overview of the company cultures. This chapter ends with a description of the operational processes that occur within Amazon and UPS for Order Fulfillment. It follows the steps of an order from being assigned to an Amazon fulfillment center (Amazon FC) to being delivered to the customer by UPS.

2.1 Amazon.com Background

Amazon, headquartered in Seattle, was founded in 1994 by Jeff Bezos with the initial vision of building an online bookstore that would offer a much larger selection of books than would be possible in a traditional brick and mortar store [5][37]. The Amazon.com web site went live in 1995 and the company went public in 1997 [5]. Despite impressive sales growth for its first seven years, the company was never able to make a profit [40]. When the dot-com bubble burst in 2000 and many other online companies started to disappear, many analysts believed Amazon was doomed to the same fate using terms such as Amazon-dot-bomb and Amazon-dot-toast to describe the company [11]. But Amazon finally announced its first profitable quarter in Q4 of 2001[40]. Since that time Amazon has grown to be one of the world's leading retailers and the fifth most visited web site in the United States and 16th most visited in the world [2].

Amazon's success can be attributed to a few fundamental guiding principles. First is their continuous pursuit of customer satisfaction with an emphasis on growing the diversity of products available for sale on Amazon.com. Another is the focus on long term thinking. They believe that if they make the right decisions on the front end, the outputs on the backend (i.e. the financials) will eventually come. In pursuing these principles, Amazon has shown a repeated willingness to take chances and to reinvent itself. To better reflect these principles, Bezos soon changed Amazon's mission from being the world's largest bookstore to being "Earth's most customer-centric company where people can find and

discover anything they want to buy online [5]". These basic tenants can be represented by a drawing (see Figure 1) that has become an iconic model at Amazon and, as the story goes, was first drawn by Bezos on a napkin one day while at lunch with a colleague. Amazonian's call this drawing the Virtuous Cycle or the "fly-wheel" while M.I.T. students of System Dynamics would recognize this as a positive reinforcing loop. As Amazon grows, they gain economies of scale and can operate cheaper. This allows for lower prices which means more customers, which means more third-party sellers want to participate [14].

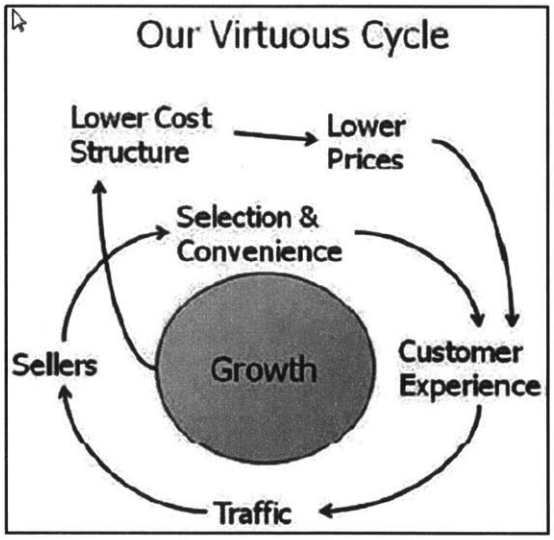

Figure 1: Amazon's Virtuous Cycle

Bezos has created a culture at Amazon which is highly tolerant of risk and Amazon has a track record of decision making which reflects this tolerance. What all these business decisions have had in common is that they were all customer-centric strategies that focused on long-term results. To describe this long-term mentality, Bezos said, "The landscape of people who do new things and expect them to be profitable quickly is littered with corpses." [24].

Below is a list of some of these decisions that were often initially met with skepticism by investors

- **Allowing competitors to sell on Amazon.com** - today over 30% of sales on the web site are by parties other than Amazon [3].
- **Allowing customers to rate products**- while this seems commonplace today, at the time many people questioned why Amazon would allow negative reviews for products it was trying to sell [24].

- **Fulfillment By Amazon (FBA)** – This is a business model where sellers can ship products to Amazon and have them store it and handle the entire fulfillment process after sell. This includes offering Amazon's steep shipping discounts to the buyer [41].
- **Kindle Reader** – In 2007, Amazon started selling an electronic book reader. Developing and marketing a piece of hardware is a bold move for a company that specializes in Fulfillment. Amazon now sells more electronic books than either hardcover or paperback books [4].
- **New Market Segments** – Amazon has relentlessly pursued new market segments in its attempts to increase selection on its web site. According to their 2009 Annual Report the number of different items available for immediate shipment grew by more than 50% in 2009 alone [4].
- **Amazon Prime** – In this strategy, Amazon offers free 2-day shipping for a fixed annual subscription fee. This was questioned by investors upon launch as they feared that consumers might take advantage of this offer by placing far more single item orders for lower margin products [43] [44]. While this could potentially drive down average order size, it has also helped drive significant growth for Amazon [42]. As Amazon grows and continues to builds more fulfillment centers, the chances are that they can ship from locations closer to customers and can deliver products to customers more efficiently.

2.2 UPS Background

UPS was founded by Jim Casey in 1907 in Seattle (just a few blocks from the current Amazon headquarters) under the name American Messenger Company. In this time before the widespread use of telephones, their business consisted mostly of message delivery with some package delivery mixed in. The company was founded on the policies of customer courtesy, reliability, round-the-clock service, and low rates – principles that still govern them today. In 1919 the company expanded into California and adopted the name United Parcel Service [27].

After decades of legal battles to overcome Federal and State limitations on interstate commerce, in the 1970's UPS became the first package delivery company to serve every address in the 48 contiguous United States. In the early 1980's, UPS began to assemble its own jet fleet and in 1988, UPS received authorization from the FAA to operate its own aircraft, thereby officially becoming an airline. Today, UPS is one of the 10 largest airlines in the United States. The 1980's were also the decade that UPS

entered the international market in earnest. Today, UPS operates a small package network in 185 countries and can reach over four billion people [27]

In 1992 UPS began tracking all packages and rolled out their handheld data tracking devices known as DIAD's. In 1995, UPS added the functionality on its web site to track packages in transit and the popularity of this functionality exceeded all expectations. Today, the UPS web site averages over 26 million daily tracking requests [27].

In 1999, 92 years after being founded, UPS offered public shares of its stock for the first time. This gave UPS access to additional capital to enter new lines of businesses and markets. In recent years, UPS has worked to grow its Supply Chain Services business which provides logistics and freight to its customers to improve their supply chains [27].

Currently headquartered in Atlanta, GA, UPS earned nearly $50 billion in revenue in 2010. They currently deliver more than 15 million packages per day using a delivery fleet of over 93,000 vehicles and have received multiple awards ranking them the #1 Logistics/Trucking organization in the world [27].

2.3 Company Cultures

Amazon and UPS each have strong company cultures, and these cultures have both some similarities and differences. Both companies are obsessed with customer satisfaction and will take great strides to ensure that no customer ever receives their packages late. Sharing this core value between them gives the two companies a common ground in which to work together.

Despite being 16 years old, we observed that it in many ways had cultural aspects similar to a startup. When an employee has an idea for a way to improve things, the general attitude is to just go do it. This can be seen in their list of six core values, one of which is "bias for action" [37]. As mentioned above, Amazon is willing to take risks in finding new ways to serve its customers. Since it is a young

company that has grown its headcount significantly in the last few years (44% in the last year alone), the average tenure at Amazon is relatively low for Fortune 100 companies [38].

UPS is a company that was founded over 100 years ago and has been a leader in its field for decades. One can often identify a company culture by the stories that are told. The one we heard the most while at UPS was how they have an exact step-by-step process (literally) for how to enter a delivery truck and drive off in less than 8 seconds. This concept of having identified and defined the proper process (which they call their methods) for most tasks is a strong aspect of UPS's culture. Another aspect of UPS culture is the strong tradition of hiring from within. UPS management is made up almost exclusively of people who were package car drivers early in their career. It would not be uncommon for a group of managers to consist of employees who have been working at UPS for an average of 30 years.

2.4 Amazon Facility

A portion of the research for this project was done at the 630,000 square-foot IND1 Fulfillment Center (FC) outside Indianapolis. This facility is in the process of being expanded to over 1 million square feet at which time it will be one of the largest warehouses in the Indianapolis area [26]. Amazon has several different categories of fulfillment centers, but the type that process the majority of customer orders are referred to as Sortable Fulfillment Centers. They process items that are capable of fitting inside a sortation tote and allow for efficient merging of multi-item orders to reduce shipping cost. The IND1 facility is a Sortable FC.

2.5 UPS Facility

The remainder of the research for this project was done at the 81st St UPS Hub located in Indianapolis approximately a twenty minute drive from the IND1 facility. This hub is designated as a 40K hub meaning that it is rated to be able to sort approximately 40,000 packages per hour. For UPS, a hub is a facility designed to unload packages out of trailers and reload them back into other trailers to

head to the next destination. A package center is a facility where the ubiquitous brown package cars are loaded up to be sent to customers. The 81st St Hub has three package centers onsite in addition to the hub.

2.6 Amazon Operations

The typical Amazon FC is divided into Inbound and Outbound operations and then further divided into discrete functions roughly according to the diagram in Figure 2.

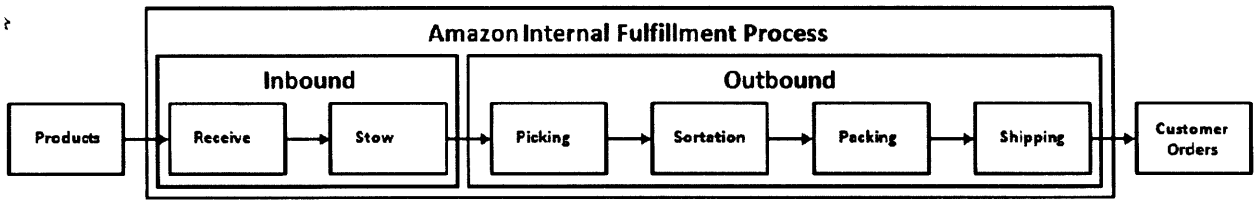

Figure 2: Amazon Operational Breakdown by Function

Inbound Operations is the process by which product is received from suppliers and stowed in the inventory bins. As product does not show available to be fulfilled until the product is in a sellable bin, both Amazon and the upstream vendors seek a fast, efficient receiving and stowing process. As this project deals with the activities after an order is assigned to the FC, Inbound operations will not be discussed further in this paper.

The Outbound Operations process begins when an order is assigned to the facility. Outbound operations includes retrieving items from inventory for a particular order and placing them in totes (Picking), grouping items into individual customer orders (Sortation), packing orders into shippable boxes, and finally shipping the products to customers. The Sortation process, where totes are grouped into larger batches and then binned into a slotted cart, is not required for single-item orders.

2.6.1 Fulfillment Center Assignment Logic

When an order is placed by an Amazon customer, a complex set of software logic decides the Amazon FC to which it should be assigned. The most obvious criterion for this decision is that the FC has the item in stock. If multiple FC's have the item in stock, there will be a check to see which facilities are capable of delivering the item by the promised ship date as meeting customer promise is one of the

highest priorities. If there are still multiple FC options, cost minimization will be the next priority. While transportation costs dominate the decision logic, minimizing operational processing costs is also a consideration. This can potentially create non-intuitive results such as choosing to ship an item to an Indianapolis customer from a different FC than the Indianapolis FC if both locations are considered to have the same Zone by the carrier (same rate charged) and the further FC has lower fulfillment costs. A similar example would be if an Indianapolis customer ordered two items, one that is stocked in Reno and one that is stocked in both Reno and Indianapolis. In this case, both shipments could ship from Reno in order to consolidate orders into a single package.

2.6.2 Ship Method Assignment

Once orders are packed, they go through an automated process where the ship method is determined and a shipping label is applied. While a preliminary ship method was calculated when the order was assigned to the FC, the moment before the application of the ship label is when the final ship method is decided. For each ship method, carriers specify to Amazon a certain time each day that after which they can no longer receive packages and still make delivery. This time is typically referred to in the industry as the Critical Pull Time (CPT). The Operations team is allocated a certain amount of time prior to that CPT to fulfill the Order.

To further illustrate this process, we will walk through a hypothetical example. If a customer orders a package on a Tuesday and requests a two-day shipping option with a promised delivery date of Thursday, then a particular 2nd Day Air ship method might be considered. For this example, let's assume that the 2nd Day Air method has a CPT of 5:00 pm and that Amazon allocates to Operations at least three hours to fulfill any Order. With those values, Orders could be assigned to that ship method up until 2:00 pm. For any Order placed after 2:00 pm, Amazon assumes that this 2nd Day Air ship method would not depart Amazon until the following day. If this ship method was selected, the expectation is that the Order will arrive to where the ship label is applied before 5:00 pm. If Operations fails to deliver that product in

time, Amazon might have to use a more expedited ship method to maintain a positive customer experience. The Operations teams go to great strides to ensure that this rarely happens.

The preliminary ship method assignment process is intricate, but it primarily focuses on filtering for ship methods that meet the promised customer delivery date and then minimizes cost. Customers choose a ship option (e.g. Next Day or Two Day shipping) and not a ship method (e.g. Next Day Air or 2nd Day Air). Therefore, Amazon is often able to meet customers Two Day shipping requirement with services that do not require air transportation. For example, the graph from UPS's web site (Figure 3) shows that UPS is able to deliver to a large geographic region around the IND1 FC in one or two days [27].

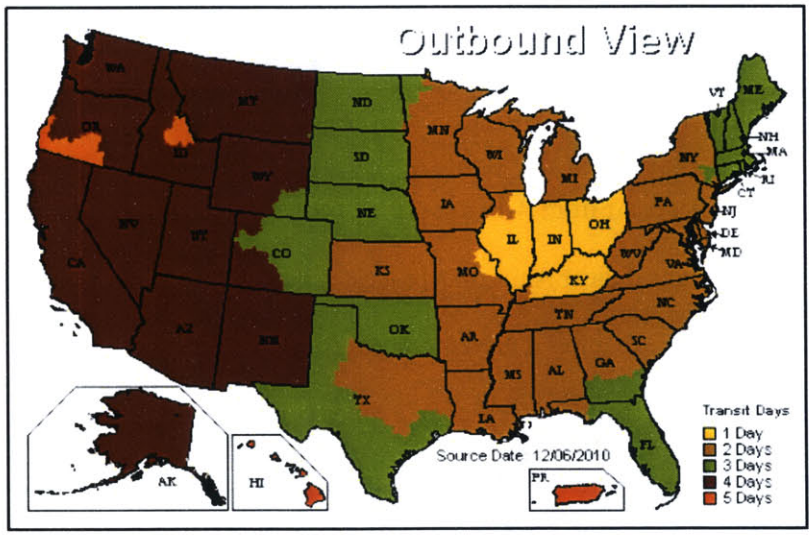

Figure 3: UPS Transit Time from Indianapolis Hub

2.6.3 Shipping Dock Processes

At the IND1 facility, once a package destined for UPS ground service (as this paper is primarily focused on UPS ground service) completes the ship label application process, it will complete one of three different shipping dock processes based on package size (see Figure 4).

- Packages larger than a certain size never leave conveyors up until they are loaded directly into a trailer in a process known as fluid loading.
- Smaller packages which are difficult to fluid load are diverted to a line where an Amazon employee will sort them into large cardboard containers known as Gaylords.

- The smallest packages, usually single items in flat envelopes, are sorted by automated equipment into separate Gaylords from the medium sized packages.

Packages that are placed in Gaylords are usually staged in an area away from the trailers. When a trailer is about to depart, the Gaylords will be loaded on the back of the trailers via pallet jacks.

Figure 4: Fluid loading of packages (left). Gaylord loading of packages (right)

UPS ground trailers depart from Amazon during roughly fixed time windows throughout the day, but UPS is able to alter the schedule at Amazon's request if demand is different than anticipated. For UPS to pick up a trailer usually requires a driver to come from the UPS hub and drop off an empty trailer and taking away the full trailer. But usually a couple times per day, UPS is able to multi-task and bring a full trailer to Amazon's Inbound Operations while carrying away a full trailer from Outbound. To depart a trailer, the conveyor's must be removed, the staged Gaylords loaded with pallet jacks, the trailer closed and sealed, the locks removed from the brake lines, and the dock lock disengaged.

All the movement of UPS trailers in and out of loading doors is performed by UPS drivers. Ideally, UPS can place the empty trailers into the Yard or into another open dock door. If the empty trailer must be placed back into the same dock door that the full trailer is being removed from, the UPS driver must set down the empty trailer in the Yard, hook up to the full trailer and move it to the Yard, hook back up to the empty trailer and move it to the dock door, and finally hook up to the full trailer and drive away.

2.7 UPS Hub Operations

The purpose of a UPS Hub is to receive inbound packages and sort them out to new destinations. They both receive in from and send out to several different types of locales. These include UPS package centers (where the brown package cars are loaded), customer pickups (CPU's) such as Amazon, as well as other UPS hubs. In addition to unloading, sorting, and loading, a hub also containerizes small packages, handles exceptions such as invalid addresses, and creates Customs documentation for international orders.

A UPS hub is staffed by part-time Teamster employees during a series of three to four sort windows that each lasts four to five hours in duration. For instance, 81st St hub has a Day Sort that runs in the early afternoon, a Twilight Sort that runs in the evening, and a Night Sort that runs in the middle of the night.

When a trailer arrives at the UPS hub and completes security checks, it is either dropped in the yard or placed directly into a free unloading door. This process of bypassing the yard is known as a "Free Spot" because it saves the step of having a Yard Jockey move the trailer. The 81st St hub has approximately 40-50 active doors from which they unload trailers. Trailers are unloaded by either one or two employees with the assistance of an automated conveyor. Any pallets will be broken down from within the trailer rather than removing the pallet.

Once a package is out of the trailer it will be diverted by another employee based on package size. Large packages will get diverted to a series of belts based on destination. This divert process happens several times, moving the package from conveyor to conveyor, until it reaches its outbound trailer. Very large and irregularly shaped packages that cannot travel down regular conveyor are offloaded onto train cars that drive them to the appropriate outbound trailer in a milk run fashion. Small packages are diverted to a conveyor that takes them to the Small Sort process.

The Small Sort process is designed for any package that is smaller than 16"x16"x7" and weighs less than eight pounds. These packages will be sorted into approximately 35 primary destinations (known

in UPS as SLIC's) and over one hundred secondary destinations. This sort process is done at the 81st St hub manually by employees that read the destination label and place packages into the correct slot in a wall divided like a grid. Once a slot is full, a person on the other end of the wall will scan each package and place them into a reusable bag. From that point on, each package is tied to the bag, and all movements through the UPS network will be tracked by scanning the bag. This containerizing process is done to reduce touches of the individual packages and to avoid small items getting stuck in conveyors. Once a bag is closed, it is placed back on the conveyors and treated similarly to a large package.

. During a given Sort, there could potentially be multiple trailers that are filled for any given destination. If that is the case, each trailer will either depart for its destination or get moved to the Yard when full. All Outbound trailers must leave the UPS hub by a specified time each day in order to reach their next destination in time to be processed on their next scheduled Sort. Some trailers might have this Critical Pull Time (CPT) that falls on a later Sort. For instance, packages unloaded during the Twilight sort might get loaded onto an outbound trailer that does not need to leave until after the Night sort to make its next destination. If that is the case, these packages can potentially sit in the outbound trailer for several hours waiting for a trailer departure. The trailers that have a CPT closest to the end of a Sort are referred to as "Hot" destinations and are the key determinant for when a Sort ends. For instance, if the Twilight sort is scheduled to run from 5:00 pm to 9:00 pm and a certain destination with a CPT of 9:30 pm must have all packages processed on the Twilight Sort, then this 9:30 pm CPT is the Hot destination that limits the Twilight sort window. If the Hub were not able to unload and sort all the trailers on the Twilight sort in time, some packages would miss the trailer when it was forced to depart at 9:30 pm and therefore, miss service to the customer.

2.8 Seasonality of Businesses

Both UPS and Amazon must be able to manage the very strong seasonality of their businesses. The Christmas holiday demand is significantly stronger than the rest of the year for both companies. For Amazon, demand during the six weeks after Thanksgiving increases dramatically compared to the rest of

the year. According to Amazon's Quarterly Reports, the fourth quarter has represented 35%, 39%, and 38% of their annual sales for 2008, 2009, and 2010 respectively [4]. In 2010, Amazon shipped over 13 million items on its largest shipping day of the year [39]. Amazon handles this volume by significantly ramping up temporary staffing during the holidays and using more Sortation areas, Packing lines, and other infrastructure to increase throughput. Because Amazon must be flexible enough to vary throughput dramatically, they prefer solutions that are easily and quickly scalable. This tends to mean that they avoid highly automated systems. They would either have to invest in automation that could handle the peak volumes and therefore go underutilized most of the year or automation that could only handle their non-peak volumes and would not have the flexibility to scale up easily.

For UPS, the volume increase during the holidays is not as dramatic as Amazon's but still very significant. UPS primarily handles their volume bursts by working one of the days of the weekend that they normally do not work and by increasing the length of each Sort window such that they are sorting more hours of the day. They also ramp up their staffing during the holidays but not by multiple factors as Amazon does. One step that UPS takes during the holidays is to send UPS employees onsite to some of their largest customers to help coordinate the heavy volume. They will help schedule the arrival and departure of trailers and control the flow of volume to UPS based on communication between the two sites. This is one example of where UPS and Amazon are already working together to some degree to smooth operational processes.

3 Literature Review

This chapter contains a review of previous research that has been conducted on topics related to this project. This includes research on different views of supplier relationships and collaboration, research on different approaches to analyze and improve operational processes, as well as research on effective ways to establish process metrics. For each topic below, we describe the research finding and include a brief summary of how this can apply to this project.

3.1 Supplier Collaboration Research

There is a wide body of research around what makes Supplier collaboration projects successful. Some of the consistent findings are that the two organizations must have the following attributes [17].

- Commitment – The willingness of partners to exert effort to maintain the relationship, to be future oriented, and to pursue joint goals rather than to pursue opportunistic behavior.
- Coordination – The definition of the tasks that each party expects the other to perform.
- Interdependence – Both parties perceive benefit from interacting.
- Trust – The belief that each party will fulfill its obligations.

Trust is the most talked about attribute in the literature. Trust allows for the parties to manage greater stress and display greater adaptability [17]. The more frequently parties successfully transact, the more likely they will build trust for subsequent transactions. Trust allows for increased information sharing without suffering the adverse affects of information asymmetry, and it makes it more likely that each party will view the information they do receive as reliable [21].

In addition to these attributes, quality of communication is fundamental to success. Timely, accurate, and relevant information is essential to achieving the desired goals of the partnership. Related to this is the willingness to share critical, often proprietary information that allows for more efficient operations. Finally, joint participation in both planning and goal setting help align expectations [17].

A final success characteristic of partnerships is the ability of the partners to develop effective conflict resolution techniques. Given the interdependence of the two partners, conflict is unavoidable. The key is to focus on joint problem solving efforts that allow for mutually satisfying solutions. The use of persuasion is much more effective than destructive conflict resolution techniques such as coercion, domination, and confrontation. Finally, smoothing over or avoiding the issues fails to identify root causes and in counterproductive to the partnerships goals of mutual gain [17].

3.1.1 Bilateral Deterrence and Bilateral Convergence Theories

Based on Bilateral Deterrence Theory, the amount of conflict one would expect to see between two firms is a function of their level of dependence on each other. A firm's dependence on a partner is traditionally defined as the firm's need to maintain a relationship with the partner to achieve its goals. This other firm's dependence is a source of power for that partner. Bilateral Deterrence Theory states that – all things equal – conflict is inversely proportional to the combined power of the two firms. Each party's desire to engage in conflict is suppressed by its fear of retaliation and expectation of attack. This effect is strongest when the power relationship between the firms is symmetric. As interdependence becomes more asymmetric, the relatively powerful firm has increasingly less motivation to avoid conflict since they can inflict proportionally more damage on their weaker. The weaker partner expects to be exploited or attacked and will attempt preemptive strikes [15].

While the absence of conflict described above permits the development of trust and commitment (critical factors for the development of long-term relationships), it does not itself generate these attributes. These will exist when there is bilateral convergence – when the interests of the partners are convergent. Minimally interdependent firms base their operations on short-term contracts, transactional pricing, and frequent switching to alternative partners. In contrast, high interdependence deters opportunistic behavior, negative tactics, or coercion, because both have much to lose. "Because both firms receive valued contributions from each other and face relatively high exit barriers, each partner has strong motivation to build, maintain, strengthen, and perhaps even deepen the relationship. These convergent

interests make it more likely that the channel partners will perceive their relationship as a 'win-win' opportunity that, through cooperation and joint action, will pay handsome dividends for both partners."[15]

3.1.2 Lean Supplier Development

A growing field of Supplier Relationship literature is developing around Lean supplier relationships modeled after Toyota's relationships with its suppliers. This field looks at an entire value stream across the supply chain and refers to all the players in this supply chain as the Lean Enterprise [31]. Many of the guiding principles of Lean Supplier Development fall counter to the traditional partnership theories stated in the previous section. They often involve a dominant firm with significantly more power that is working with a small highly dependent partner. The goal of the large firm is build a *keiretsu* or "a network of vendors that learn, improve, and prosper in sync with their parent company" [16]. They reward small players with business for doing a good job at achieving improvement goals. They never have suppliers bid for the lowest price; rather they learn what a fair price is and a fair profit for the supplier and set price accordingly. They are able to do this because they learn all that they can about their suppliers. They practice *Genchi Genbutsu* or "go and see" and spend much time at the *gemba* or "the place the work is done". They send many employees, right up to presidents, to see and understand firsthand how suppliers work. Jeffrey Liker and Bruce Choi wrote in their paper, "Building Deep Supplier Relationships" that Japanese automakers pursue six key strategies when working with suppliers [16].

1) Understand how your suppliers work – Send employees to the gemba

2) Turn supplier rivalry into opportunity – Develop several suppliers and reward business based on ability to achieve long term goals.

3) Monitor vendors closely – Keep a report card and monitor vendor performance at all times

4) Develop their capabilities – Share expertise with activities such as guest engineer programs.

5) Share information intensively but selectively – Provide the right information without overwhelming supplier

6) Conduct Joint Improvement Activities – Hold extensive *kaizen* (continuous improvement) events with suppliers.

With the success that Japanese automakers have had with this type of *keiretsu* development, many companies have attempted to replicate this type of relationship with little to moderate success. As discussed in the previous section, the short term benefits that a dominant firm can acquire through opportunistic behavior, price wars, and coercion are difficult to resist long term.

3.1.3 Applicability to this project

Some of the findings in this review help justify our explanation for why these two companies make for strategic partners. The importance of trust is emphasized in some of the proposed improvements and in the iterative approach which is designed to build trust. Finally several of the future recommendations are designed around improving communication and conflict resolution strategies. Bilateral Deterrence Theory suggests that these two firms would want to avoid conflict and bilateral convergence suggests that these two firms would want to build strong relationships. This supports our view that two large firms can be strong partners. Lean supplier development gives details of some of the most common partnerships we see in industry today. Because of the typical parent to child relationship in most of these partnerships it gives an interesting framework to compare and contrast against since the UPS and Amazon relationship does not have the same power structure.

3.2 Queuing Theory

Queuing theory is a discipline of Operations Research that takes a mathematical approach to studying queues [30]. Key metrics in this discipline include the expected number of items in queue and the average waiting time in queue. Mathematical formulas have been derived to calculate these metrics based on the distribution of the arrival and service processes. Mathematical queuing theory has many real world limitations that prevent it from modeling all processes such as assumptions of infinite queue length (no blocking), no decisions to avoid waiting in line (balking) and no ability to adjust arrival and processing rates based on system attributes (independence). For these more complex processes,

Simulation software is a better tool for analyzing these processes and developing approximate values for the above mentioned key metrics [6]. One of the most common equations tied to queuing theory is Little's Law. This is a brief but powerful equation that establishes a relationship between the average Work in Process (WIP), average arrival and processing rates, and the average cycle time. We will perform some high level queuing theory analysis to help analyze the flow of products through the UPS hub later in the paper.

3.3 Lean Manufacturing

It is nearly impossible to study Operations Strategy these days without the repeated mention of Lean Manufacturing. Almost without exception, in each manufacturing facility that we have toured over the last few years across many different countries, we have seen some version of Lean manufacturing programs in progress. Companies will often offer their own interpretation with their own branding such as GM's Global Manufacturing System (GM-GMS) or the Caterpillar Production System (CPS), but they are all based off of the same fundamental concepts found within Lean Manufacturing.

The term Lean was coined in 1988 and first published in a master's thesis at the MIT Sloan School of Management [29]. The movement has been made popular by many successful business books on the topic such as *The Machine That Changed the World*, *Lean Thinking*, and the *Toyota Way*. Lean is a management philosophy that is based primarily off of the Toyota Production System (TPS). Its primary focus is on the reduction of *muda*, or waste. The seven primary forms of waste, as defined by Taichi Ohno of Toyota, are overproduction, unnecessary transportation, inventory, motion, defects, over-processing, and waiting [31]. Many people add additional wastes to this list including the untapped creativity of their employees and the designing of products that do not meet customers' needs.

Hopp and Spearman, define Lean by saying that "Production of goods or services is lean if it is accomplished with minimal buffering costs." They go on to explain that all variability in an operation must be buffered in some way or another, and there are three types of variability buffering – having

excess inventory, excess capacity, or excess time. Therefore, Lean is focused on minimizing th cost of these buffers. Shigeo Shingo, one of the most famous historical experts on TPS, once said, "'Eliminate Waste' is a nonsensical slogan. We take it as a given that the elimination of obvious waste is obvious." He said that we should focus rather on the less obvious source of waste which is variability. This includes variability in process times, delivery times, yield rates, staffing levels, demand rates, etc. [13].

There are many tools that are associated with the Lean concepts. These include Value Stream Mapping, Five S, Kanban systems, visual factories, and *poka-yoke* (error-proofing). While many companies focus on these tools while implementing Lean, the original TPS focused much more on cultural aspects such as customer focus, continuous improvement, respect for people, and building long-term relationships with suppliers. An additional goal associated with Lean manufacturing is the implantation of single-piece flow, *heijunka* (production leveling), and pull systems that produce products at the rate demanded by customers. To achieve these goals usually requires a heavy focus on reducing setup and changeover costs to minimize batching [31].

A significant portion of the improvement strategies for this project are based off of the basic concepts of Lean manufacturing. In our analysis, we focus on identification of waste and approaches for reducing variation. Some of the common tools of Lean such as Value Stream Mapping are used.

3.4 Metrics

There has been extensive research around determining what makes for effective and ineffective metrics. This research shows that systems involving humans respond very strongly to performance measures and that people will modify their behaviors in an attempt to ensure a positive performance outcome even if this means pursuing inappropriate courses of action. Since inadequately designed performance measures will result in dysfunctional behavior, significant amount of effort needs to be put into designing the correct measures. Some of the attributes that well designed measures have include [19]

- Being Simple to Understand
- Having Visual Impact

- Focus on improvement rather than variance
- Visible to all
- Provide timely and accurate feedback
- Be part of a closed management system
- Uses data which is automatically collected whenever possible

Neely et al. defines ten elements that must be specified when generating any new metric [19].

1) Measure – The title should clearly explain the measure and why it is important
2) Purpose – Define the rationale for creating the measure
3) Business Objective – Define which business objectives the measure is designed to measure
4) Target – Determine the level of performance needed to achieve business needs
5) Formula – This is one of the key for ensuring that adverse incentives are not being created
6) Frequency – Function of the importance and data availability. Frequently of measure should be tied with frequency of action.
7) Who Measures – Who is to collect and report the data
8) Source of Data – A consistent source of data is vital for performing trend analysis over time
9) Who acts on the data – Who is to take action when the targets are not being met
10) What do they do – Probably the most important element -- without a closed management loop, the metric serves no purpose. The exact action might be context specific, but the management process needs to be defined

At both Amazon and UPS we found examples of metrics that we believed were poorly designed or missing and found this to be an opportunity for further process improvement. The above literature review provides a framework to compare existing metrics against to see where they are lacking as well as a reference point to use when designing new metrics.

4 Research Methodology and Approach

The following section describes the research methodology used to understand the operating environment, develop recommendations, and pursue improvements opportunities within the UPS Ground shipping operations at Amazon.

4.1 Literature Review

We attempted to leverage existing knowledge and research within the relevant fields of study prior to pursuing any recommendations or process improvements. This included reading existing papers and books on topics such as Supplier Collaboration, Lean Methodologies, and Operations Research. In addition we leveraged all relevant information learned within the classroom setting over the last several years.

4.2 Data Collection

The process of collection data for this project came in many different forms including learning from doing, talking with workers, querying databases, and performing time studies.

4.2.1 Performing the work

One of the most effective ways of understanding the processes and identifying opportunities for improvement is to execute each role in the entire process oneself. This is a technique that Amazon strongly promotes, and we had the opportunity to perform every role from Picking to Trailer Loading within Amazon. Within UPS the opportunities for executing the work were more limited with a strong Union culture, but this was possible to a limited degree.

4.2.2 Observation

Much of the data collected for this project came from observation of different processes. These included observations of the steps that were taken as well as measurements of different cycle times and processing rates through time studies. We made attempts to avoid the observation bias known as the

Hawthorne Effect where people modify aspects of their behavior in response to the knowledge that they are being measured [36]. This was done by observing for extended periods of time in non-obvious forms as well as being very open with all participants that the observations were in no means meant to measure anyone's individual performance.

4.2.3 Paper Records

Some critical data such as the arrival and departure times of trailers were only stored on paper security logs. Therefore, months worth of these logs were scanned and re-entered into electronic systems in order to collect critical metrics on cycle times of carrier pickups.

4.2.4 Electronic Data

Both Amazon and UPS are well known for having sophisticated information systems that collect massive amounts of data. We were able to access data directly from web browsers as well as querying data out of data warehouse databases. The data warehouse retrieval required significant amounts of SQL querying as well as understanding of table structures.

4.2.5 Excel Macros

There was certain data for which historic records were not kept but time critical values were required. Therefore, this data had to be captured and stored at all times throughout the day. The only way to achieve this was to develop Excel macros written in the VBA language that would screen scrape web pages on a periodic basis and store the data

4.3 Value Stream Mapping

With all the data collected and the processes documented, our next step was to create Value Stream Maps of both the current and future states. A key benefit to participating in a Supplier Collaboration project is that we are able to view and document the entire value stream. Taking a value stream perspective means working on the big picture, not just individual processes, and improving the whole, not just optimizing the parts. The process for generating Value Stream Maps was taken from the book

Learning to See. This mapping creates a visual representation of every process in the material and information flow required to deliver the product or service [22]. In this mapping we identify each activity as either being 1) value added to the customer, 2) non-value added but required by the production system, or 3) not creating value and can be eliminated [31]. In addition we create data boxes for each process that records information such as cycle time, up time, defect rates, and number of operators. Areas of opportunity are identified on the current state map, and then a future state map is generated [22].

4.4 Data Analysis

Once we identified opportunities for improving the operational processes that occur around the end of day delivery of ground packages to UPS, our next step was to predict what type of impact a delayed Critical Pull Time would have on the flow of orders through the facilities and on the ship method assignment of different orders.

One technique for doing this was to look at the order download rates for different ship methods over time. Figure 5shows an example of what an Order download rate chart might look like for two CPT's relevant to this project.

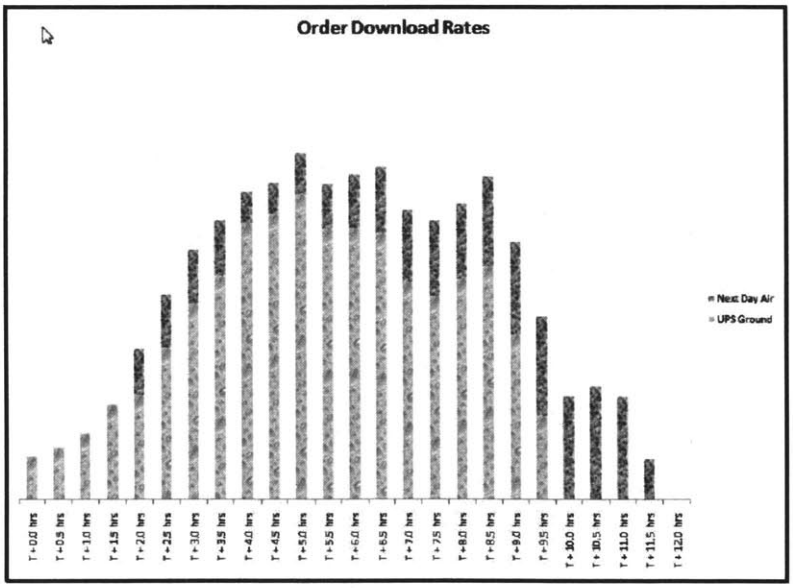

Figure 5: Order Download Rate Patterns for two ship methods relevant to project

These order download rate patterns give an idea for how strong demand for each ship method is each day and helps us predict the type of impact the change will have on the fulfillment process.

A second approach for analyzing impacts is to run ship method assignment simulations. The application inside Amazon that calculates the optimal ship method has functionality to simulate changes to the assignment logic and predict the impact of changes. It performs these simulations by allowing the user to make changes to the assignment logic in a test environment (such as move the CPT for a ship method to a different time) and then rerunning a historic set of orders (e.g. all the orders that processed through the IND1 FC last week) against the new logic. The user can then compare the newly calculated method to the actual method that was assigned historically. These calculations give an approximation for the amount of impact any logic change will have. This process is critical prior to making any changes to this assignment logic because small changes can potentially have significant changes to individual carrier volumes, and both Amazon and the carriers need to prepare and coordinate for the change in volumes.

The simulation process has a few limitations which make its impact calculation of changing ship assignment logic only approximate. Therefore, the full impact can only be understood with actually piloting the changes. These limitations are not present when calculating what the ship method would have been without the change. Therefore, using the simulations for calculating the before and after ship method once a pilot has been run allows us to gather benefit analysis data that is quite accurate.

4.5 Piloting

Once improvement opportunities were identified and the impacts analyzed, our next step was to test the results with small one-day pilots. These pilots had several purposes. The first was to validate that the benefits to the change were real and were able to be measured. Another purpose was to identify the impact of the changes on each company's operations. It was predicted that these changes would create stresses on the order fulfillment processes of both companies, but pilots were needed to see how severe the disturbances were. A third purpose was to validate that all the operational steps could be executed in

the time windows that were predicted and therefore allow for completion of the sorting of all the packages prior to the end of the UPS Sort window. These pilots were initially done on the days that were considered the least challenging to execute and least risky to customer experience to minimize any potential negative impacts.

4.6 Change Management

The most challenging phase of this project was around change management. On paper, the case for implementation seemed strong, but getting buy in from all parties that this change was worth pursuing took a significant amount of time and effort. There were many sources of these challenges including the way benefits distributed across different organizations, the metrics that different groups were held accountable for, as well as some cultural aspects. In the end, all parties did all agree that the changes were worth while.

4.7 Standardizing Work

A final approach to this project was to implement standard work for the purpose of sustaining the improvements in production. This was important due to criticality that Lean manufacturing places on the concept. Taichi Ohno, considered to be the father of TPS, is credited for having said, "where there is no standard, there can be no *kaizen*" (*kaizen* meaning improvement) [1]. Standardized work allows for changes to be sustained and provides a basis for before and after comparisons during a change process. A couple of graphical analogies explaining the benefits for standardized work can be found in the following figures. Figure 6 represents the "wedge" analogy that prevents backsliding on the way to improvement [1]. Similarly, Figure 7 shows the progress of process improvement over time with and without standard work [25].

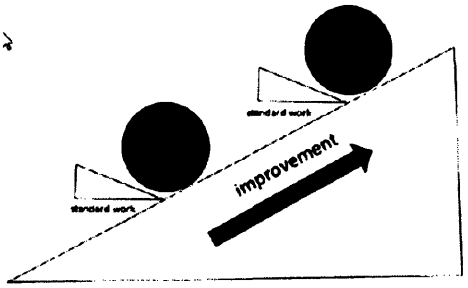

Figure 6: Standard Work as a wedge that locks improvements in place

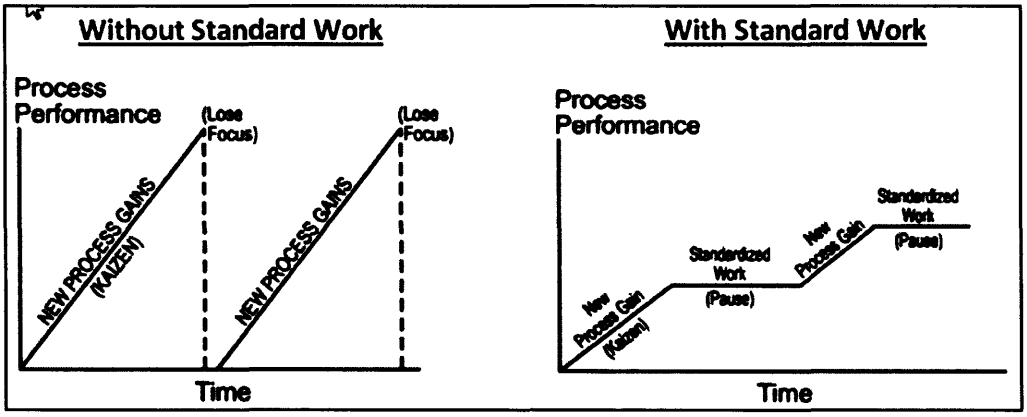

Figure 7: Graphical depiction of performance over time with and without standard work.

Amazon is a company that changes processes and evolves very quickly. This can create challenges with maintaining well documented standardized work within and across its fulfillment centers. This is in stark contrast to UPS which has very well documented processes for each of their operational steps. Amazon's version of standard work consists of a document made up of photos of each step with brief descriptions beneath each photo. The only PMV's that Amazon has on record for the IND1 shipping area are a generic description on how to palletize products and a generic description on how to load a trailer. For this project, we wanted to create a more detailed PMV specific to this process that documented the success criteria.

5 The Amazon and UPS Collaboration

In this chapter we talk about some of the specifics around the collaboration project between Amazon and UPS and how this partnership compares to many others and to findings in prior research. We also talk about the process of choosing a partner and for choosing a project. The chapter ends with a detailed analysis of the benefits that each company might see from working with each other.

5.1 Challenges with Partnerships

A supplier partnership can be defined as a, '"purposive strategic relationship between independent firms who share compatible goals, strive for mutual benefit, and acknowledge a high level of mutual interdependence" [17]. The benefits to supplier collaboration seem quite intuitive and many examples of successful partnerships can be found. Yet this collaborative approach remains the exception, especially in the United States and Europe. Some reasons given for this are that companies are simply accustomed to the more traditional methods and that the collaborative approach requires more work [32]. Drawbacks include increased complexity, loss of autonomy, and the obstacles that come with information asymmetry [17]. In his *Lean Logistics* book, Michel Baudin analyzes these challenges from the perspective of Game Theory. Once two companies are in such a relationship, each could benefit short term by either leveraging the knowledge they gained during partnering to demand price concessions or by leveraging its single-source position to stop improving. Given that the adversarial position is optimal for each side given what the other is doing, this creates a Prisoner's Dilemma scenario where the Nash equilibrium is for each side to resort back to being adversarial. [32].

5.2 The Power-Player Dynamic

The typical partnership has a power player which acts something like a maestro for the supply chain. They decide with whom they would like to partner and often approach the partnership as a development project. The power dynamics vary across industry. For instance, in the computer industry, the supplier of a key component (Intel's processors) is in a position of strength with respect to assemblers,

where as in the auto industry, the car companies maintain power over small parts suppliers [32]. This is one aspect which makes Amazon's and UPS's decision to work together unique because neither company is in a clear position of strength over the other. Both companies are leaders in their industry, and both are consistently ranked in the top 50 of the World's Most Admired Companies [5][27]. Having the size, power, and reputation that comes with being the leader in one's field can naturally lead to an expectation that all partners will accommodate your wishes and tailor their processes to your methods. Seeing how Amazon and UPS work through that dynamic will be an interesting aspect of how this partnership develops in the long term.

5.3 Finding the Right Partner

Finding the right company with which to work more collaboratively is a strategic decision that must consider many factors. When you analyze the relationship between UPS and Amazon, we see that these two companies are a good fit for pursuing collaboration. Buyers of services are historically more likely to seek longer-term partnerships than those in the market for goods. The intangibility of service attributes makes a transactional relationship more challenging because quality is typically revealed over a longer time scale [8]. In their book *Lean Supplier Development*, Harris et al. walk through a strategy for assessing a good partner. This book focuses on partners that a lead company is looking to develop, but much of the strategy is the same [12].

1) *The right attitude* – Does each company want to partner? The relationship between UPS and Amazon was propelled by strategic discussions made at the highest levels of executives at the two companies and all agree that pursuing a closer relationship is something that each wants.
2) *Quality* – Does each company have similar views and track records with regards to quality? In this context we can define quality as delivering shipments to customers quickly and as promised. Amazon and UPS are both obsessed with on time delivery and providing a positive customer experience. This joint passion is a common bond which strengthens their relationship.

3) *Capacity and growth strategy* – As Amazon and UPS look to grow, they want to make sure that each has the growth strategy that will get them to their goals. As the largest carrier with one of the most developed infrastructures in the logistics industry, UPS is well positioned to be able to handle the dramatic growth that Amazon has been and expects to continue to deliver. Amazon's growth will continue to give UPS the volume and delivery density to make their entire operations for all of their customers more efficient.

4) *Vital expertise* – Do the companies have something the other needs? Amazon and UPS potentially have much to learn from the other, as each has core competencies that the other would find useful. For instance, even though Amazon fluid loads many trailers every day and has to quickly turn trailers when they are full, this is something that UPS does much more frequently in their hundreds of facilities across the world. UPS could potentially share expertise on how Amazon efficiently delivers products to UPS with their other customers. For instance, some of the findings with this project around reducing trailer delivery cycle time are potentially applicable to other UPS customers.

5) *Location* – Partnerships require effort to implement and sustain, and they cannot be done effectively virtually over email and phone calls. They require frequent visits to each other's facilities for both face-to-face meetings and to see the processes first hand (*Genchi Genbutsu)*. Therefore, an initial location for a partner must be one where the two companies are in close proximity. UPS and Amazon strategically chose the Indianapolis locations due to the warehouses close proximity to the local ground hub and the reasonably close distance to UPS's primary Air hub.

5.4 Finding the Right Project

Once UPS and Amazon decided to pursue this relationship and also decided on a location based on proximity of key facilities, the next step was to decide on the scope for their first project. Therefore, a list of criteria was generated with which to evaluate all potential project ideas. The criteria list was as follows

- *Large enough impact to warrant a multi-month project* – The resources for this project were preset, so we wanted to ensure that the project would take advantage of the allocated time.
- *Not so large as to be infeasible* – There needed to be tangible deliveries at the end of the project, so any projects that involved dramatic infrastructure changes or large resources of time, people, or capital were not realistic.
- *No Negative Impact to their peak season* – While both companies considered this project strategically important, their entire companies' reputations revolve around flawless execution during the Holiday season. Therefore, the scope could not expose any undue risk to peak operations.
- *Benefits to both companies* – Both companies are pursuing the project with a goal of strengthening long term long term relationships, but nonetheless, the project should not have significantly asymmetrical benefits to any one side.
- *Measurable and Quantifiable Benefits* – The project needs to impact operations in a way that is tangible and measurable (e.g. must be more than "strengthened relationship" or "built trust".)
- *Not possible without having joint access to both companies* – This project is a unique opportunity to view operations across the entire Value Stream by having full access to both companies. Therefore, we want to make sure the project is not something that could be easily done by any one employee of the two companies without that type of access.

The initial time onsite at the two companies was spent identifying potential project ideas with a list of about fifteen potential ideas generated in the first few weeks. Many of these were rejected once analyzed using the criteria stated above. For instance, some of the projects were designed to improve issues that had noticeable negative impacts on operations but were considered bad candidates because they were overly tactical in nature and could potentially be resolved solely within Amazon. Other project ideas were rejected due to either being feasible without requiring full UPS access or for unclear benefits to both parties.

The agreed upon project fit these criteria better than any others. The goal of the project was to streamline operations around the UPS Ground CPT so as to both provide earlier volume availability to UPS and a later cutoff time for Amazon. The approach was to use Lean Manufacturing methodologies to identify wastes with the primary focus being on the waste of waiting that resulted in increased cycle time. By accomplishing this goal, more orders could be delivered to customers using more efficient delivery methods. The size of the project was appropriate, it had real benefits, and it required a solid understanding of both Amazon and UPS operational processes to be done well.

5.5 Finding the Right Project Lead

One of the most challenging and critical aspects of a successful project is maintaining trust throughout the relationship that each side is respectful of the other's interests. This can be especially challenging if the primary project lead has clear allegiances to one of the two partners. As the Lead is almost always an employee of one of the companies, it is hard not to have allegiances – or at least to create the perception that allegiances exist. In this case, Amazon and UPS pursued a unique approach by having the project run as a dual-funded university research project through with the Graduate Fellow being employed by both companies. In theory, this is an intelligent and effective approach that should resolve concerns over asymmetric loyalties. In practice, the implementation did not completely resolve these issues. Due to constraints, some examples of the implementation details included having the Fellow employed by only one of the companies and having full network access with only one of the companies. Therefore, despite all attempts to maintain neutrality, the perception of interest bias did exist to some degree during this project. The concept of the joint research project was nonetheless well conceived and a more symmetrical relationship on any future project would be by very effective.

5.6 Benefit Analysis

In analyzing any collaboration project, it is critical to analyze and understand the benefits. The exchanging parties must perceive the relationship as productive, worthwhile, and satisfying to continue. If benefits are not gained, the relationship may be perceived as ineffective which can result in decreased trust and/or commitment [10]. Optimally the focus of collaboration should be on finding ways to increase the total payoff for the two companies. This does not mean that each company should not negotiate on how to share the payoff, but if each company goes into the relationship focused only about getting their money back instead of on developing long-term relationships, the results can be disastrous [12][32].

The above discussion around the optimal approach to benefits as found in Supplier Collaboration literature is also supported by economic theory. One definition of economic efficiency, known as Pareto efficiency, requires that no one be worse off for the change to be efficient. As there are very rarely instances where some party was not made worse off, a stronger definition of efficiency, known as Kaldor-Hicks efficiency, states that a situation is made better off overall if the size of the gains and size of the losses are such that gainers could fully compensate losers and still be better off themselves [33]. In other words, changes where the benefits outweigh the costs should be pursued to remove cost from the system as any losers can be compensated by some type of transfer payment. A related economic concept, known as Coase theorem, states that, if transaction costs are sufficiently low, bargaining will lead to an efficient outcome regardless of initial property rights. For example, between Amazon and UPS, if a process change requires increased transportation costs between the two facilities but has benefits that outweigh the cost, the group that is responsible for covering these transportation expenses should not affect the decision to pursue the change. In reality, transaction costs and other economic frictions result in these property rights (i.e. who has to pay) impacting many decisions.

Taking the above mentioned approach of streamlining the collective system, the CPT extension project has obvious systemic benefits. A larger volume of packages are able to reach customers in time using Ground rather than Air service. This provides a more efficient and direct transportation method than flying products in jets using a hub and spoke system. That being said, this project specifically and this relationship in general have clear benefits to both Amazon and UPS, and these will be laid out in detail in the next section.

5.6.1 Benefits to Amazon of a General Partnership

Shipping expenses make up the largest expense to fulfilling an order for Amazon. The cost of picking, sorting, and packing an order is only a fraction of the expense that they pay carrier to deliver the product from the Amazon facility to the customer. The net cost to Amazon for shipping activities exceeded $1 billion for the first time in 2010 [3]. Amazon explicitly calls out in the Risks section of

their Annual Report that there are only a limited number of shipping companies, and negotiating favorable terms with these companies is critical for their continued success [3].

5.6.2 Benefits to Amazon for Specific Project

The most straightforward benefits from this project are the transportation savings that Amazon realizes due to a later Critical Pull Time. These savings result from Amazon being able to ship orders that have premium shipping (such as Amazon Prime 2-day shipping) via Ground methods rather than via Air. As shown in Figure 3 earlier in this paper, UPS is able to deliver to a large section of the country around an Amazon FC in two days or less via Ground service if they receive the packages prior to the CPT. For the IND1 Fulfillment center, the Ground CPT is in the early evening and the 2nd Day Air CPT is much later in the night. Orders stop being assigned to a ship method a fixed period prior to a CPT to provide Operations time to fulfill the order. Orders placed between the Ground and Air cutoffs ship via Air. The types of orders that would realize savings are Orders that require two day shipping and are in the two day Ground range as well as Orders that require one day shipping and are in the one day Ground range. Orders that require two day shipping but are in the one day Ground range can be assigned Ground even after the Ground CPT because they can ship out the following day and still make service. Orders outside the two day range are going to require Air even if they are placed prior to the Ground cutoff.

The extension of the Ground CPT is particularly beneficial because the affected time window (time between the old cutoff and the new one) is during a period of heavy order activity. Extending the CPT of a ship method that fell at a different time (e.g. in the middle of the night) would affect a smaller number of orders. Figure 8 shows approximate order download patterns throughout the day.

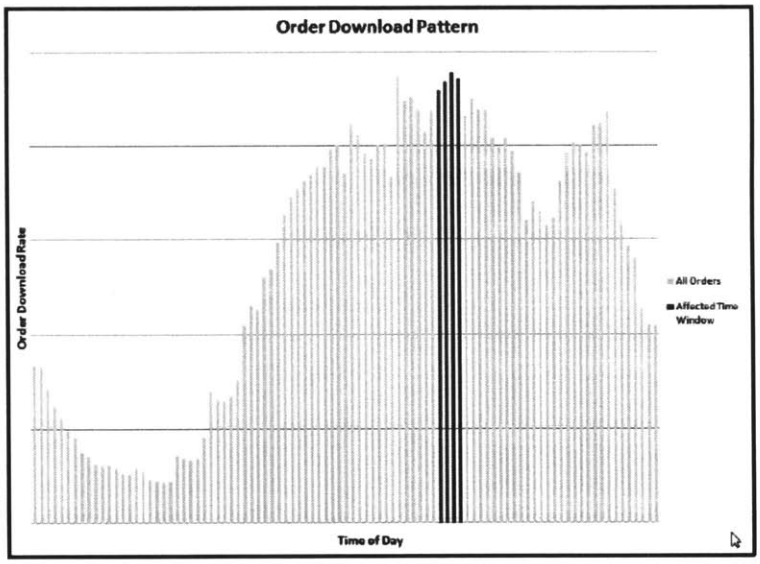

Figure 8: Order download rates throughout the day

In addition to switching the Ship Method for orders already destined for the IND1 facility, a shifted CPT can change the facility assignment for orders. For instance, if an order would get Air shipment at either the IND1 facility or another nearby facility and the other facility had cheaper overall fulfillment costs, the other facility would likely get assigned the order prior to the changes but IND1 would receive the order after Ground shipments were permitted. Therefore, during an extended order download window, the IND1 facility is likely to see more overall orders and not just a change in the ship method of existing orders.

5.6.3 Benefits to UPS of a General Partnership

Partnering with Amazon to find better ways to operate and to continuously improve their processes is one way of ensuring that UPS will remain on top of their industry for the long term. The alternative is to become complacent as the market leader and risk becoming vulnerable to new competition. The idea of constantly having to improve and evolve just to maintain status quo in the market is known as Red Queen competition in the academic field of Strategy. This name comes from the *Alice in Wonderland* quote where the Red Queen says, "Here, it takes all the running you can do, to keep in the same place." History offers many examples of one time industry leaders that cease to exist despite a successful industry including Bethlehem Steel and PanAm Airlines [7]. One way this can happen is by

assuming that as market leader all parties should adapt to your processes. One notable example of this was when McDonald's approached food processing giants, such as Kraft and Heinz, and attempted to get them to adapt products to their specifications. When McDonald's was forced to turn to much smaller suppliers, they helped create a new set of major vendors [28].

One of UPS's key strategic advantages is the barrier to entry that new competitors face when trying to amass the scale needed to efficiently provide last mile delivery service. This is a fact that Mike Jones of UPS' Investor Relations explicitly calls out to investors in an interview with Morningstar. He states, "The firm's competitive advantage is principally its barriers to entry, which are staggering from its massive network of trucks, aircrafts, sorting facilities, information technology and know how." He also points out how by investing over $1 billion in technology and raising customer expectations, they have created a wide moat and raised the bar higher for other competitors [23]. Research supports the advantage of these barriers. When the UK opened up the delivery market to competition last decade, Royal Mail only lost ¼ of one per cent of its delivery business to competition in the first year. Figure 9, taken from a study of the UK postal market, shows that a competitor would have to have significant market share to affordably replicate Royal Mails services [18].

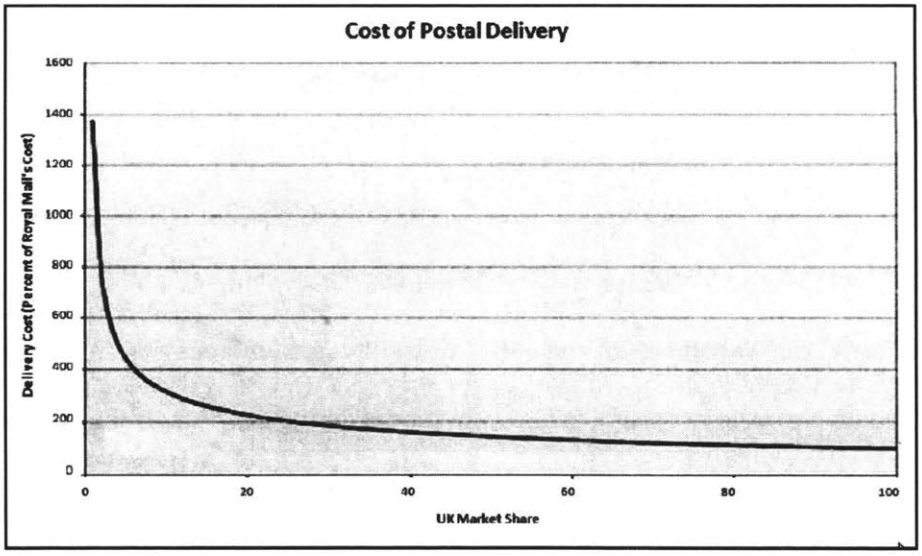

Figure 9: Cost to replicate Royal Mail's Service 6 Days a Week

Given these steep barriers to entry, one of the few ways that a fledgling delivery service could compete with the established players is to work with a few companies that have the package volume of a company like Amazon. Acquiring volume from many smaller players would be unlikely for a new entrant as branding and reputation is another barrier to entry that the established players possess. Even then, a new company would have to initially focus on a small urban market. The same UK study shows that a much smaller share of the market would be needed to compete solely in the urban market (Figure 10) [18]. By maintaining strong relationships with its largest customers, UPS is able to protect one of its strongest competitive advantages – its strong barriers to entry.

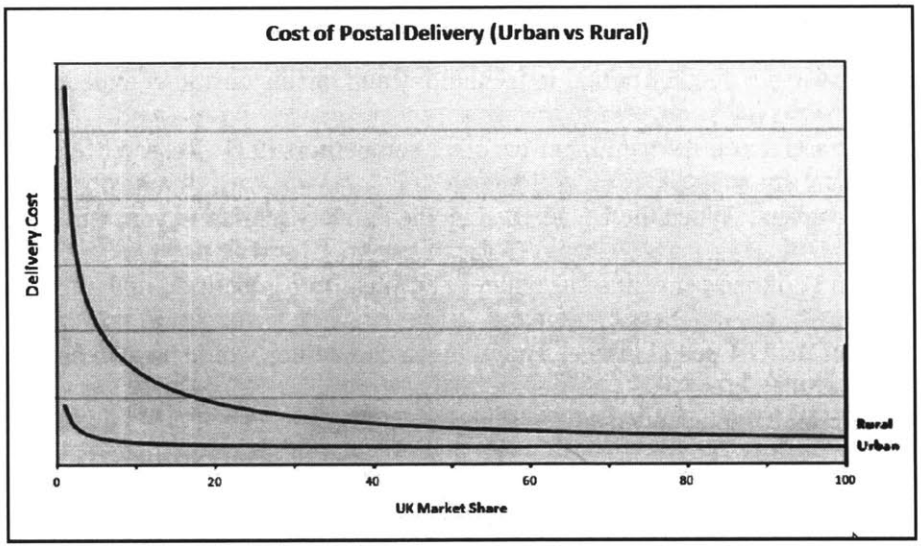

Figure 10: Cost of Urban vs. Rural

5.6.4 Benefits to UPS of the Specific Project

UPS prefers to have trailers arrive to their hub as early as possible. This "volume availability" gives them the flexibility to run their sorts more efficiently. This volume is a form of inventory that, when it arrives early, can sit in their Lot and buffer variability in their processes. At the very least, UPS needs enough volume so as to not starve their sort resources during the shift. If the volume arrives too late, underutilized resources at the beginning of the shift cannot be made back up later in the shift and they cannot finish their sort in time. A common graphical way that UPS represents this is can be found in

Figure 11. In this graph, the top line, which steps up each time a trailer arrives, must remain above the straight line in order to prevent resource starvation.

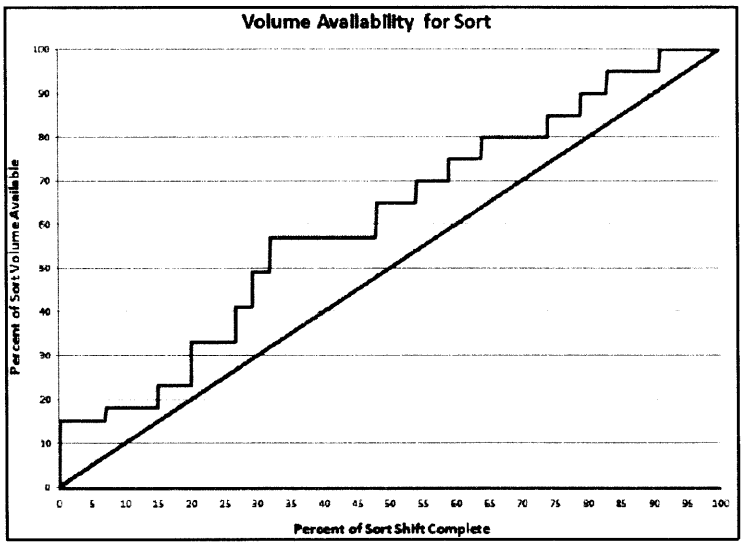

Figure 11: Volume Availability Stair Stepper Chart

Like Amazon, all of UPS's customers would prefer to deliver their trailers to the hub as late as possible. Unlike Amazon, many customers do not have the volume to send multiple trailers per day to UPS so the one trailer that arrives as late as possible arrives at the tail end of UPS's sort capacity. In that sense, the multiple Amazon trailers arriving throughout the day help subsidize the sort by providing sufficient volume to keep the resources busy early in the shift. The more early volume that UPS can get from Amazon, the better off UPS will be. The process changes associated with this project helps provide that earlier volume availability. Even though the last bit of Amazon volume arrives later, the mass of the arrival distribution is shifted earlier due to the bulk of the volume being able to depart Amazon sooner.

Many of the other benefits to UPS are associated with reducing the variability in the process and having more predictable operations. Today, UPS drivers can spend a significant amount of time onsite at Amazon (Figure 12) partly because of the large variation in times they are able to depart Amazon (Figure 13).

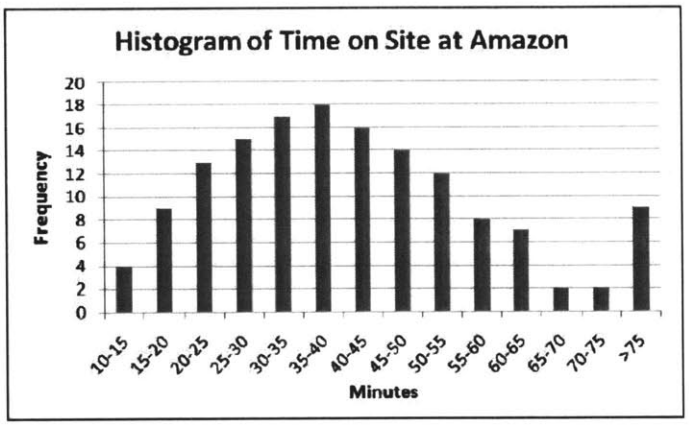

Figure 12: Historical Time UPS driver spends in the Amazon yard

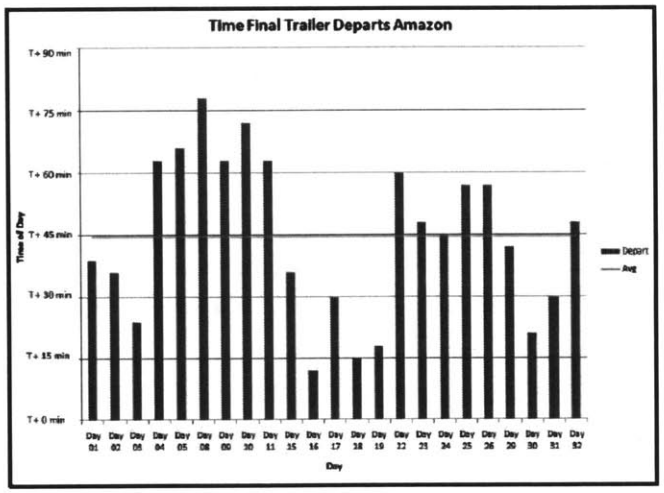

Figure 13: Departure time from Amazon of final UPS trailer over the course of one month

In addition to reduced driver wait time, a more consistent departure from Amazon will allow UPS to plan their hub operations better. A trailer that consistently arrived within a few-minute window every day could have an empty door reserved, thus avoiding having to make a pit stop in the yard. There are many other operational efficiencies that come with consistent behavior. For instance, UPS has a software tool that assigns trailers to specific doors at specific times in an attempt to optimize the shift. This tool is for the most part ignored by the control room due to the typical trailers variance in arrival time from what is scheduled.

A final benefit to UPS of a later CPT is additional package volume from Amazon. While it is true that some of this volume that is now getting shipped UPS Ground is cannibalized from UPS Air

volume, much of the volume comes from other carriers as well. Figure 14 shows the split between different alternative ship methods that would be used if UPS Ground is not available.

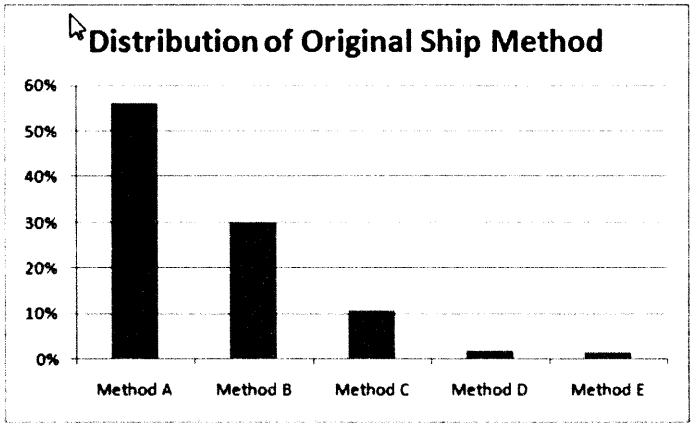

Figure 14: Original Ship Method for Orders had UPS Ground not been available

While a later ship cutoff will add more volume to UPS Ground than it takes away from UPS Air, the difference in margins between Ground and Air methods could make the changes either revenue positive or negative. The full impact to revenue is unclear due to insufficient margin data.

6 Ground Shipment Extension

6.1 Current State Analysis

UPS has a specific time each night that they must complete their sort process to allow their trailers to be released to the next destination. From this, they work backwards to set the time that their city pickups such as Amazon must cease to ship packages via UPS. This time is often set with substantial padding to allow for unexpected variances in the process. They are inherently risk averse to not being able to complete the sort due to the substantial impact that this would have on customer experience.

In this time between the ship cutoff and the completion of the UPS sort a series of process steps must take place. The trailers must be completely loaded, checked out, and released. The trailer must be picked up by a driver and transported to the hub. The trailer must be docked at the hub and all the packages unloaded. The packages must be sorted to the correct destination with some packages being containerized into bags. Finally the packages must be reloaded into outbound trailers. This process typically takes several hours.

In analyzing the current state to document the exact steps and timing of each event, we noted that there were several opportunities to reduce the overall cycle time in this process. Some of these opportunities were the result of identifying unnecessary steps or wasteful processes such as excess waiting. Other opportunities were the result of buffering needed due to high variability.

6.2 Process Improvement Opportunities

There were five key process improvement opportunities that we identified that we felt could be implemented by Amazon and UPS without a significant commitment of time, resources, or capital investment. Each of these improvements would contribute to a reduced cycle time for the overall process. These were all changes that could be tested during single day pilots without major disruptions to everyday activities. In the following section we will describe what each of these process improvement

opportunities are. In addition we provide our relative assessment on a scale of 1 to 5 of the benefit as well as the cost for both Amazon and UPS to implement and sustain (where 5 is the highest benefit and highest cost).

6.2.1 Batch Size Reduction of Final Trailer

The final trailer that Amazon sends UPS is often completely full. By choosing to reduce the number of packages that are on the final trailer, we can reduce the amount of time it takes to unload the packages from the trailer at UPS. Since it can take UPS well over one hour to unload a full trailer, this can have a noticeable impact

6.2.1.1 Lean Viewpoint on Batching

In Lean production, the ideal means of processing a product is via continuous, single-piece flow. In this type of process, the production steps are arranged in a sequence and the product moves from one step to the next, with no buffer of work-in-process or batching of goods. Batching always leads to long wait times as the product sits awaiting the changeover to the type of activity the product needs next [31]. But batching is often unavoidable due to economies of scale effect that come from fixed costs such as order costs or setup costs. Batch size reduction becomes feasible when the fixed costs associated with creating each batch can be reduced [6]. This was the exact approach that Taichi Ohno took in the early days of Toyota. They were able to achieve "one-piece flow" by reducing setup times on machines from 2-3 hours in 1945 to 3 minutes in 1971. Today, Single-Minute Exchange of Die (SMED) is a popular lean production discipline [13].

6.2.1.2 Continuous-Flow Opportunity at Similar Amazon Site

The impact of batching in the context of this project can be demonstrated with an example from another Amazon FC very similar to the Amazon facility in Indianapolis. At this facility, the UPS hub is only a few hundred yards away from the Amazon shipping docks (as shown in Figure 15).

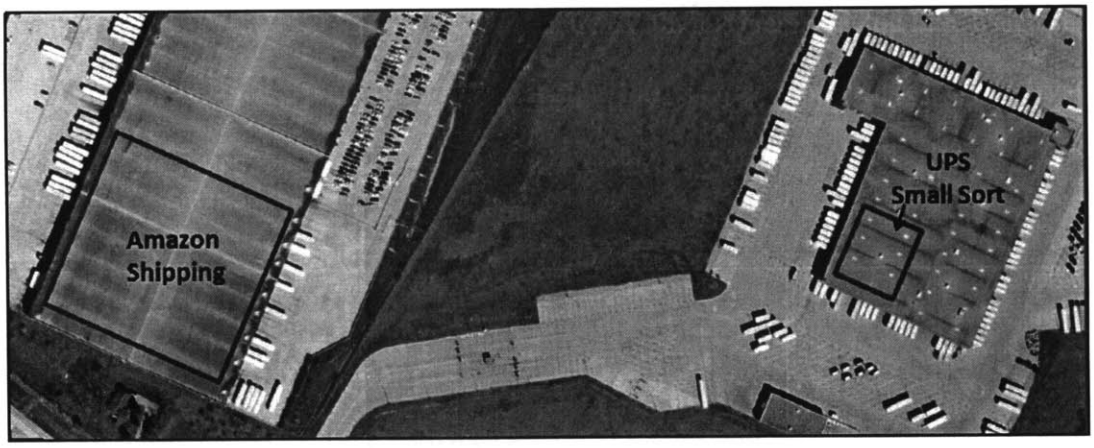

Figure 15: Map showing proximity of one specific Amazon and UPS facility [34]

At this particular facility, one could imagine that a continuous, single-piece flow operation could be created if the conveyor taking packages to the Amazon trailers could some how be extended directly into the UPS processing line. If that were done, the packages would arrive at the UPS hub only a few minutes after they were packed at Amazon and would be sorted as they arrived.

Instead, what actually happens at the above facility is that the packages are loaded into trailers. When the trailers are full, a UPS driver will pick the trailer up and bring it to the UPS facility. Despite the short distance, UPS's goal is still to fill trailers close to 100% full (largest possible batch) to keep the transportation costs down. That is because the current process has significant fixed costs despite the close proximity. To get to the proper entrances, the driver has to drive down three long blocks. Because they drive on public streets, they use full trucks rather than the yard mules which have a significantly reduced trailer hookup time. They have to go through a security process which is sometimes arduous. Finally they sometimes have to wait long periods of time before trailers are released by Amazon. The round trip process could potentially take over an hour for a UPS driver. Therefore, these large fixed costs drive the desire for large batches. The large batches result in a Critical Pull Time for Amazon which is over three hours prior to the end of the UPS sort as compared to something a little more than the travel time on a conveyor in the hypothetical flow example.

In many real world contexts, continuous, single-piece flow is a theoretical ideal which cannot be fully achieved since fixed costs cannot be entirely eliminated. But any reduction in fixed costs can lead to reduced batch sizes which will then lead to reduced product waiting time. In this example, small changes such as reducing security delays and better coordinating pickup schedules would help. An example of a more extreme measure might be to connect the driveways of the two buildings currently separated by the strip of grass (see Figure 15). This could potentially allow for bypassing security, use of yard mules instead of trucks, etc. Any of these improvements would reduce pressure to deliver 100% full trailers.

6.2.1.3 IND1 Final Trailer Strategy

At IND1, standard practice is to send the final UPS trailer of the day that is very close to full. It is not uncommon for them to need two trailers at the critical pull time because Amazon has processed more than one trailer's volume since the last delivery to UPS. This large batch puts pressure on UPS to require an early CPT for two reasons. First, it takes a significant amount of time to unload a full trailer once it arrives to UPS (see Appendix B for more details). Second, the material that can be sitting at Amazon for two-plus hours after it is packaged is not available to UPS to be sorted. This increases the likelihood that they starve their sort resources early in the shift (i.e. falls below the volume availability line in Figure 11). A better solution is to send a mostly full trailer prior to the ship cutoff and a partially full trailer right at the ship cutoff. The exact timing of the second to last trailer can be adjusted to get the delivery profile that best meets UPS's needs. If the primary objective is to have very little volume in the last trailer, then the second to last trailer would need to leave close to the ship cutoff. If the primary objective is to have more volume arrive at UPS earlier, then the second to last trailer can leave Amazon earlier at the expense of having more to unload in the final trailer.

Having a final trailer with reduced volume sometimes results in an extra trailer being delivered to UPS each day. This is not always the case because two trailers would sometimes be sent at the ship cutoff time regardless due to the need for slightly more than one trailer in capacity. The cost of an additional trailer can be partially mitigated by reduced fixed costs as a result of some of the other changes

discussed later in this chapter (e.g. driver time on site). It can also be viewed as just a more accurate assessment of optimal batch size analysis. The batch size reduction savings outweigh the batch setup costs and justified an additional transport.

This particular process change would have the largest benefit to cycle time reduction of any of the ones listed in this section due to the large reduction in unload time at UPS. Amazon is required to check in and check out an additional trailer as well as move conveyors between trailers. This change is most costly to UPS as they would on some days have to provide a driver to make an additional round trip between the two sites. *Cycle Time Benefit: 4 Amazon Cost: 2 UPS Cost: 4*

6.2.2 Batching Trailer Release Work

Another opportunity to speed up the delivery of products to UPS is to avoid batching the trailer release operations of multiple ship methods that share the same ship cutoff time. For many valid reasons, the Operations group prefers to have multiple ship methods that all have the same cutoff time rather than having many CPT's spread throughout the day. As a result, there are at least four ship methods that have the same cutoff time as UPS Ground. There are typically a limited number of people who checkout, close, and release trailers. For efficiency, they will batch operations for these multiple ship methods together. For instance, they will virtually create the manifests for all the ship methods at once, load all the Gaylords onto the trailers at once, close and seal all the trailers at once, and finally remove the brake locks and dock locks at once. The end result is a process that can take two to three times as long from beginning to end as it would take to release any single trailer. A root cause of this behavior is the metrics that the shipping associates are measured against. They are given a maximum amount of time to perform these operations, and as long as they are within this time, their objective then is to try to minimize their labor costs. By having a binary metric (i.e. am I fast enough or not) rather than a continuous metric (i.e. how quick can I complete this), there is no incentive to find faster ways to complete this work. In this situation, batching is the expected behavior. More analysis on ways that the metrics can be improved can be found in Chapter 8.

Allocating more people to the trailer release process when many ship methods share the same critical pull time will reduce the time it takes from when the final package reaches the ship dock to when the trailer is released to the UPS driver. This can be implemented short term with work instructions stating that the steps must be done by a certain time or with management oversight ensuring that batching doesn't happen. A more sustainable solution would be to adjust the metrics so that workers are incentivized to minimize the cycle time. This step does not require any changes for UPS.

Cycle Time Benefit: 2 Amazon Short Term Cost: 2 Amazon Sustainable Cost: 3 UPS Cost: 1

6.2.3 Doing work in parallel

In the pursuit of trying to release trailers more quickly, there are many tasks that can be done prior to the ship cutoff rather than waiting until all packages have arrived at the ship dock. Today, all of the steps required to release a trailer typically do not start until after the ship cutoff has been reached. There are many steps that can happen in the time leading up to this time. Gaylords can be moved from their staging areas to right outside the trailers in the few minutes prior to the ship cutoff. An exception area (known as Jackpot) can be sorted ahead of time. Extra Gaylords from the two small package lines can be removed from the line and brought to the trailer. Finally, the rolling conveyor inside the trailer can be removed and the Gaylords start to be loaded while the last few packages are arriving. Once all Gaylords are on the trailer, these last few packages that arrived after the conveyor was removed can be loaded on the tail end of the trailer.

Doing this work in parallel allows for the trailer to be sealed shortly after the arrival of the final package to the shipping dock. Similar to the previous improvement, this can be implemented via work instructions and management oversight but would be more sustainable with improved metrics. This does not require any changes for UPS. *Cycle Time Benefit: 2 Amazon Short Term Cost: 2 Amazon Sustainable Cost: 3 UPS Cost: 1*

6.2.4 Trailer Jockeying

Once the final trailer is sealed, the speed with which the UPS driver is able to leave Amazon is highly dependent on what type trailer jockeying the driver is asked to do. Since IND1 does not have a yard jockey, carriers perform all trailer moves for Amazon. If the carrier has to put an empty trailer back in the same door as they are picking up from, this is a time consuming process (see section 2.6.3 for details on this process). Since IND1 does not have fixed doors assigned to each carrier, Amazon has the flexibility to avoid reusing the same door by allocating an extra door that the carriers use to hop between. It would be equally effective to have a second UPS Ground trailer already placed in an extra door by an earlier driver so that the UPS driver that picks up the final trailer can set down their empty trailer in the yard prior to the CPT. A final method to avoid reusing the same door would be for IND1 to hire a yard jockey and yard mule so that the carriers spend no time moving trailers around. This is what is done at some other Amazon facilities and works well.

Implementing one of the options above can speed up the process by the amount of time it takes to fully hook up and then release a trailer two times. The primary cost to Amazon is the extra effort it takes to plan out a logical door utilization strategy in advance. This could potentially require Amazon to use doors that are in suboptimal locations for short periods of time as well. Implementing this change would reduce UPS's workload. *Cycle Time Benefit: 2 Amazon Cost: 2 UPS Cost: 0*

6.2.5 Consistent Process Timing

The current process has significant variability with regards to the timing of events such as the release of the trailer by Amazon (see Figure 13 for a chart). This variability results in UPS drivers not knowing exactly when they need to arrive at Amazon, and therefore they have to potentially wait extra time in the Amazon yard. In addition, the variability in their departure time means that the time that trailers arrive at the UPS hub fluctuates from day to day. A byproduct of all the above changes is that the Amazon trailer should be released within a tight window of just a few minutes each day and therefore arrives at the hub consistently. Once this happens consistently, UPS should be able to anticipate and plan

for the arrival of this trailer. The control room at UPS can reserve a door so that the UPS driver could release the trailer directly to the Unload area rather than setting the trailer down in the yard.

Trailers can sit in the UPS yard for 10-15 minutes prior to being assigned an Unload door. Being able to bring a trailer directly to a door and bypassing the yard actually saves UPS the work of having to unattach and reattach a trailer. The downside is that UPS is exposing itself to the risk of leaving a door unutilized for a period of time while the Amazon trailer arrives. Therefore the implementation challenge of this improvement is dependent on the predictability of arrival and the level of trust they have in the process. As the track record of consistent performance grows, this should be an easy feature to implement.

Cycle Time Benefit: 3 Amazon Cost: 1 UPS Cost: 2

6.2.6 Process Improvement Summary

Looking at the summary table found in Table 1, none of the strategies listed are rated as being difficult for Amazon to implement. The first strategy is the most costly for UPS to implement, but it also generates the largest benefit and is critical to the success of the project. Therefore, the outcome of these changes is largely dependent on UPS's ability and willingness to alter the final trailer strategy as described in Section 6.2.1.

Table 1: Summary of Process Improvement Strategies with Cost/Benefit Scale

Section	Improvement	Cycle Time Benefit	Amazon Cost	UPS Cost
6.2.1	IND1 Final Trailer Strategy	4	2	4
6.2.2	Batching Trailer Release Work	2	2	1
6.2.3	Doing work in parallel	2	2	1
6.2.4	Trailer Jockeying	2	2	0
6.2.5	Consistent Process Timing	3	1	2

6.3 Piloting process

Many of the improvement opportunities mentioned in the above section were tested in isolation to validate that the operations teams could implement the changes and that the cycle time reductions were in line with expectations. For instance, there were several days where the Batching Trailer Release Work

(section 6.2.2) and Doing Work in Parallel (section 6.2.3) sections were implemented to validate that we could release a trailer within a fixed number of minutes after the critical pull time. There were other days where we implemented a door assignment strategy that ensured that there was no extra trailer jockeying required (as in section 6.2.4). None of these small trials resulted in benefits that could be measured because they were all performed without changing the critical pull time. For instance, closing the trailer quicker just meant that the closed trailer would sit waiting for a driver. Eliminating the jockeying would result in the trailer getting to UPS earlier, but it would not result in any packages changing their ship method.

The most effective way to validate all of the process improvements was to perform a series of single-day trials or pilots where we actually moved the critical pull time to later in the evening. By extending the shipping window, we reduced the time allocated to complete the sort process forcing a reduction in the cycle time. In each of these pilots we implemented all of the changes in section 6.2 with the exception of reserving the door at UPS (section 6.2.5) as we had not yet established our ability to consistently deliver trailers at the same time.

6.3.1 Individual Pilot Details

Over the course of the project, we extended the ship window on seven different days. To build confidence, these pilots started with days that were predicted to have the smallest impact on Operations and therefore the most likely to succeed. From these pilots, we were able to validate whether trailers could depart according to the predicted schedule, whether the expected volume of packages were on each trailer and whether UPS could finish processing the loads by the end of their sort.

Pilot 1 - In the initial pilot, we were testing whether we could execute to all the changes, get the trailer released to UPS in time, and have UPS complete the unload process prior to the end of their sort. For the initial test, we ran on the lightest volume day of the week. The results went quite well with the trailers arriving at UPS much quicker than normal and with significantly less volume to unload. There were two key findings from the first pilot. The first was that even though Amazon could make the

changes spelled out in Section 6.2 without too much difficulty, they had a much harder time ensuring that all packages arrived to the ship dock in time with the new ship cutoff time (details of these challenge can be found in Section 6.4). The second was that we realized that we had insufficient means for collecting data on the transportation savings benefits.

Pilot 2 – There was a large gap in time between this and the previous pilot as we implemented better data collection techniques and better communicated with Amazon management about the process changes. This pilot was done on a different day of the week that was slightly more challenging to implement. Again, the trailer timings and volume went according to plan. This was the first pilot where we could start to measure the benefits as well.

Pilot 3 – This pilot was on the same day of the week as the first pilot with no additional process changes. We were simply trying to gain confidence and momentum for full time sign off. This pilot taught us the importance of change management as we implemented this test on the same day that another major process change was to be implemented at Amazon. This other process did not implement cleanly which resulted in downtime in the facility for much of the afternoon and invalidated any results from this pilot.

Pilot 4 and Pilot 5 – These two pilots were executed in the same week. One of their primary goals was to test how well the Operations teams could work with increased autonomy. As opposed to the first few pilots where we individually walked each worker through the steps they needed to take, in these pilots we reviewed the process at the beginning of the shift and tested how well the process could run on its own. The trailers were not released as quickly which showed that more work was needed in implementing sustainable solutions.

Pilot 6 and Pilot 7 – These final two pilots were executed in the same week as well. These pilots were implemented with even more autonomy than the previous two. As we were approaching making a decision about full time deployment, these two pilots were primarily an opportunity for additional practice and to continue to collect data on the impact to ship method assignments.

6.4 Amazon Fulfillment Challenges

As detailed in the pilot descriptions above, being able to implement the process changes that resulted in reduced cycle time was not the greatest impediment to implementation. We were actually able to make most of these changes fairly successfully during the first pilot. Had that been the only obstacle, this project would have gone from piloting to full implementation very quickly. A key success criteria was ensuring that Amazon Outbound operations was able to consistently get all Orders through the ship label application process by their target ship time. The timing of these changes could make this challenging and hence full implementation was delayed. Below we spell out some of the obstacles that delayed full implementation.

Network Load – Amazon's flexibility in being able to implement process changes is dependent on the amount of stress their network is experiencing due to demand. This is why little process change happens during the holiday season. During the time that we were pursuing the first changes, the back to school demand combined with the delayed startup of Amazon's newest fulfillment center strained the Amazon fulfillment network and reduced their ability to experiment with new processes.

Shift Changeovers - As mentioned earlier in this paper, Amazon Outbound operations is allocated a certain amount of time from when orders stop charging to the FC until when the orders must reach the ship label application process. All processes in the FC are designed with that allocation window in mind. If the UPS Ground CPT is moved back, then a shift change falls during that allocation window which takes away processing time from Amazon. This lost processing time might be acceptable for some ship methods, but with UPS Ground being a very large ship method, this creates many challenges for Outbound. To get all the orders to shipping in time without creating unplanned shipping upgrades, they must be extremely diligent and potentially implement less efficient methods to speed up the process.

Proximity to Other CPT's – There are other large CPT's a couple hours after the UPS Ground CPT. As the Ground CPT is moved later in the evening, it starts to move two ship methods with very high volume

close together. This opens up the possibility that Amazon can spend too many resources focused on the earlier ship method and not leave themselves enough resources to execute to the later ship method.

The challenges created by the proximity of these two ship methods vary by day of the week since the volume of each ship method can fluctuate heavily throughout the week. Figure 16 shows order download rate patterns for different days of the week with the dark bars representing the earlier UPS Ground ship method and the light bars representing the later Next Day ship method that falls shortly afterwards. In these examples, Monday's would be a challenging day to implement a later Ground ship cutoff because the UPS Ground volume is so heavy. With this amount of Ground volume, the lost time as a result of the shift change is critical. Thursday is a challenging day to implement a later Ground ship cutoff because of the very high Next Day volume. If Amazon focuses too much on UPS Ground prior to switching their attention to the Next Day volume, they could end up missing many orders. Friday would be the easiest day to pilot due to the low volume of both Ground and Next day volume.

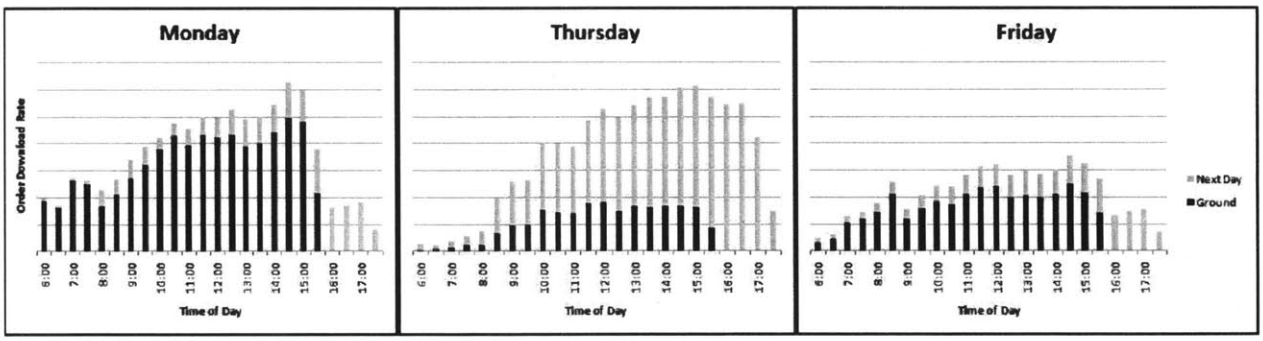

Figure 16: Order Download Rate Patterns for each day of week

Organizational Alignment and Metrics – While the project might have net benefits for Amazon as a whole, the changes would result in transportation benefits at the expense of operational metrics. At Amazon, the Transportation and Operations groups were recently realigned to be more in sync, but they are still working to synergize their goals. This project could potentially impact two metrics for which Operations is held closely accountable. The first is variable labor cost per unit. To position themselves to fulfill all the UPS Ground orders in a reduced time (due to shift change), Operations must attack their backlog more aggressively than normal. While it is hard to quantify, the belief is that this can have a

noticeable impact on labor efficiency. The second metric is unplanned upgrades. Amazon has a culture of doing everything possible to prevent orders from getting to shipping late and therefore upgrading their ship method. The Operations managers are held strongly accountable for this and therefore are strongly opposed to any changes which increase the risk of upgrades.

Leadership changes – Amazon is a capable enough company to overcome the organizational alignment challenges mentioned in the previous section, but it often requires top down leadership to help drive the change. A critical leadership change that took place during when many of these changes were being deployed set back the deployment timeline for by at least a few weeks.

6.5 Validation of Transportation Savings

One of the key reasons for running the pilots was to collect data that could be used to validate that the ship methods assignment logic would behave as anticipated and that we would see the corresponding transportation savings. Prior to running the pilots, we ran a series of simulations using the ship method assignment software (details about simulations in section 4.4) that allowed us to predict the effects of moving the UPS Ground CPT to a later time. In the pilots, we were able to validate that the ship method assignment changes were indeed in line with our predictions.

The first validation point was to see that orders continue to be assigned to UPS Ground later in the day (Figure 17).

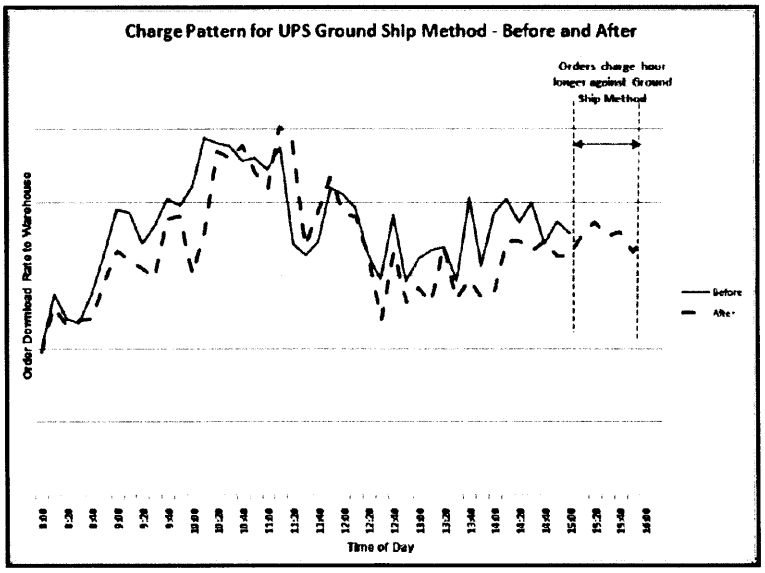

Figure 17: UPS Ground ship method continues for an extended duration

Once UPS Ground is no longer a viable ship method, the number of orders that get assigned to other 2nd Day Air methods spikes. The second validation point was to confirm that the spike in 2nd day air shipments was also delayed. Figure 18 shows that, in the pilots (dotted line) the download rate does not spike up until an hour later.

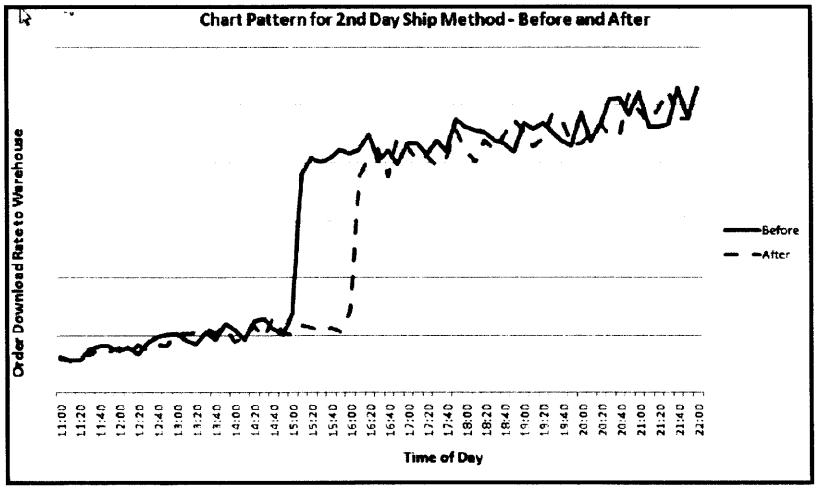

Figure 18: 2nd Day Air spike delayed during pilots

Another validation point was to confirm that the average transportation savings per package were in line with predictions. The actual savings per packages varied wildly from just a few cents per package when the alternative ship method was a similar ground carrier to many dollars per package for some of the

larger packages that would have otherwise gotten shipped via Next Day Air. The histogram in Figure 19 shows the distribution of savings (scaled to a dollar base of 10).

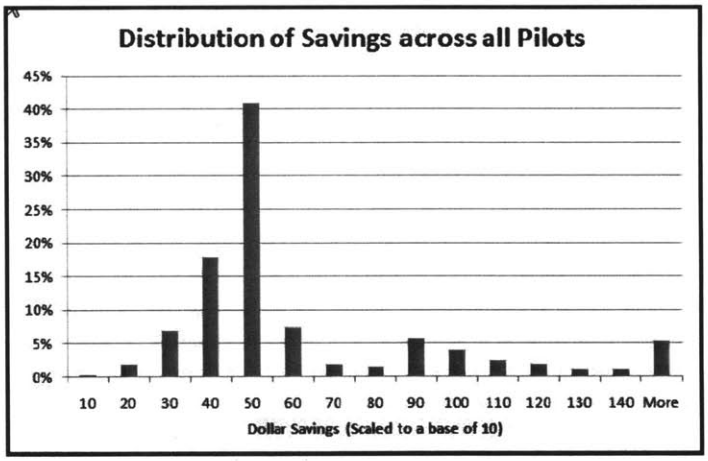

Figure 19: Distribution of Transportation Savings across all pilots

6.6 Project Implementation Conclusions

After seven pilot attempts and several months of implementation discussions, all parties agreed that full time implementation was the right decision. One of the key factors in this decision was the data comparing the transportation savings to unplanned upgrade costs. Amazon Outbound was never able to execute during any of the pilots to their traditionally excellent track record of nearly zero unplanned upgrades. But, as shown in figure 20, those upgrade costs were negligible compared to the transportation savings.

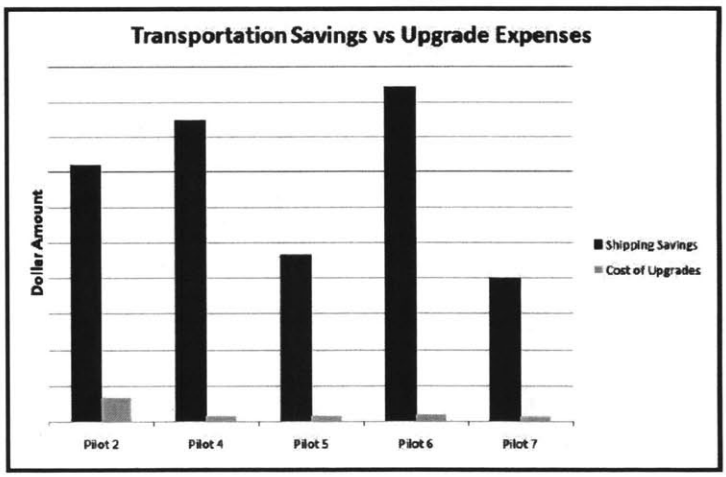

Figure 20: Transportation Savings vs. Unplanned Upgrade Expenses

6.6.1 Iterative Process Improvement

One drawback to the drawn out implementation process was that it did not tailor itself well to an iterative approach to process improvement. In an iterative approach, one would make small changes, see how well they worked, and then gradually make incremental improvements. This is a superior strategy in a supplier collaboration project because it accommodates some of the trust issues that are often found in these partnerships. With a single "big bang" implementation, individual firms are going to be conservative with the promises that they commit to because they are not confident that the other side can or will be able to execute to their half of the commitments. Iterative strategies are also the preferred approach to Lean implementation projects. In Lean this is often referred to as the Plan-Do-Check-Act (PDCA) cycle.

In this project, the amount of time that the ship cutoff was moved back was not directly tied to the impact of the process improvements. The extension was more directly a result of what UPS was willing to commit to. It is quite possible that had all the changes been implemented in a steady state that additional slack could have been identified in the schedule. In addition, once the above mentioned changes had been implemented, it would have provided a foundation from which we could have attempted additional process improvements. The next chapter outlines some of the future work that we believe can potentially result in even further reductions in the overall cycle time.

7 Future Work

In the course of this project, we identified other opportunities which could either improve the speed of product delivery to UPS, better sustain the current project, or strengthen the ongoing partnership between the two companies. While none of these ideas were fully implemented during the project, they are all areas worth exploring further in the future.

7.1 Next Generation Small Sort

The big drawback to sending full trailers of product to UPS (large batches) is that some product has to wait at Amazon for hours after they arrive at the Ship dock with no value added work being done to them. The small packages then have to be sorted and bagged once they arrive at UPS's Small Sort area. A more efficient solution would be to use this waiting time at Amazon to do some of this sorting and bagging onsite at Amazon. This is a proposal strongly supported by UPS, one that has many benefits.

In this process, UPS employees would work onsite at Amazon using a small footprint of floor space. Using the Next Generation Small Sort (NGSS) equipment (shown in Figure 21), they scan labels at a scanner in front of them and place the package into the appropriate bin that lights up. Baggers, on the opposite side of the bins, empty the bins when full. As opposed to traditional UPS Small Sorts, packages are scanned prior to being placed in the bins. The bagger does not need to scan each package [20].

Figure 21: Photo of Next Generation Small Sort Equipment

Using some of the trailer unload calculations found in Appendix B, we approximate that if a trailer that was 100% full of small packages were replaced instead with nearly 1000 bags of Smalls, that the time it would take to unload a trailer at UPS would be reduced by over one hour. In addition, the bags would skip the Small Sort process at UPS, reducing the processing time even more. Other benefits to Amazon of having onsite sorting are reduced missorts (since packages are electronically scanned at Amazon) and the environmentally friendly reduction of up to 10,000 gaylord containers per year.

7.2 Improved Process Metrics for Amazon Shipping

Well designed metrics are critical for driving the desired behavior in any process as research has shown that they have a strong influence on human behavior. To ensure that Amazon is within compliance of their trailer release policies and to help drive continuous improvement, there are three data points that are critical for Amazon to measure. None of these are currently being tracked in a reliable way today

- *Amount of product on each trailer* – Amazon accurately tracks every package that leaves their facilities, but they don't accurately track which trailer it is on when there are multiple trailers for the same ship method. There is no electronic means of knowing today whether a departing trailer has 1,000 or 5,000 packages on it. This data is critical for determining compliance with maximum package limits.

- *Time that Amazon releases a trailer* – The time that a trailer is physically released to a carrier is not currently recorded. The only time that is recorded is when the trailer is virtually released by the Ship Clerk in the shipping software. There is often not a strong correlation between when the trailer virtually departs and when it physically departs. If targets are put in place that state that a trailer must depart by a certain number of minutes after a CPT, then the ship dock workers will make sure that this virtually happens regardless of the physical timing of the event.

- *Time that carriers arrive and leave the Amazon property* – This is accurately tracked today but only via paper security logs. In an extreme event, this data can be recorded by traveling to the security shack and searching through binders of logs, but this is not an effective way for determining if processes are in compliance.

A final issue with the metrics is that they are designed to read something like "Percent of time we depart trailers in XX minutes or less." This is a binary metric which only ensures that the event

happens in at least the maximum amount of time. If the operators are hitting 100% compliance, this metric provides no additional incentive for continuous improvement.

7.3 Increased Information Sharing

One of the most often mentioned traits of strong partnerships between companies is increased information sharing that comes along with increased trust. There are several data points that make strategic sense for Amazon to share with their carriers in some type of real time electronic fashion. The first are the compliance metrics mentioned in the previous section. We not only recommend that Amazon track volume per trailer, release time of trailer, and arrival and departure time of drivers; but we recommend that this data be available electronically to the individual carriers so that they can both use this data for planning and help hold Amazon accountable that they are doing what they say. Another useful category of data that could be shared with carriers is demand data such as real time order backlog for a ship method or at least the volume already on the ship dock at any given time. This would help carriers prepare for unusual demand such as the day of a major release of an anticipated book or game.

7.4 The Virtual Andon Cord

One of the biggest hurdles to collaboration projects that streamline processes across companies is building and maintaining trust. A typical cross-company process will have slack built in on both sides because one company cannot be sure what the other is going to do. In a collaboration project, the goal might be to say something to the effect of "If you do A, B, and C and if I do X, Y, and Z then we do not need all this slack." But there is often not a clear (or at least efficient) process for escalating when the other company is not doing A, B, or C. In this project, we propose that this process be straight forward and formalized in the form of a supplier contract. We compare it to the Andon Cord in TPS where any worker can stop the line when there is a defect and seek immediate root cause rather than continuing to work through the defects. In the partner agreement contract that we propose, it is clearly spelled out what the agreed upon conditions are. It then spells out what a violation is (i.e. how much do you have to be off

the target and how many times can you be off). The next section specifies the escalation process and the expected response by the violating party. The final section spells out what qualifies as egregious enough violations to warrant unraveling any process changes.

7.5 UPS Hub Cycle Time Reductions

Appendix A details some Operations analysis within the UPS hub and proposes some potential ideas that could be implemented to reduce WIP and cycle time within their hub. Some ideas proposed include implementing more Pull systems and switching from inventory to capacity buffering to manage process variations. See Appendix A for more details.

7.6 Removing Hot Destinations from Final Shipments

The UPS Ground ship cutoff is determined by working backwards from the end of their sort shifts and ensuring that all packages that are scheduled to be processed on that sort are completed. The sort shift is scheduled to end based off of several "hot" destinations that must depart the hub shortly after the end of the sort. But many other destinations do not have a pull time from the UPS hub until much later in the night, and these packages that get processed on the Twilight sort will often sit in outbound trailers for many hours at the UPS hub. We believe the possibility exists for Amazon to create a new Ship Method which removes all "hot" destinations from shipping out of Amazon in the last hour or so. Amazon would then be able to send a trailer which arrives to the UPS hub towards the end of the Twilight shift. The idea being that UPS should be able to process the entire trailer before the end of shift, but if they could not, the remaining trailer would be processed in the first few minutes of the Night Sort. With no "hot" packages on the trailer, all packages would still reach the Outbound trailers before they had to depart.

7.7 Quarterly Business Review

Much of the literature on Supplier Collaboration discusses the importance of the Quarterly Business Review process for continuing the momentum of ongoing projects. Classic supplier meetings are negative, one-sided, and are focused on cost reductions and quality problems. But true collaboration

efforts should focus on information flow, learning opportunities, and sharing of best practices. In these meetings, upper management is present and engaged to send messages that leadership is serious about the process and the relationship. The meeting place will ideally rotate locations so that all parties can see the end-to-end process. When each partner can see their products being handled and used in the others facility, it can spark improvement ideas [12].

Our belief is that the Amazon and UPS relationship would benefit significantly from the increased interaction that comes with a Quarterly Business Review process. Our experience has shown that despite the significant amount of business that the two companies do together, the interaction between the employees is fairly limited. Even basic language issues such as what does a Critical Pull mean varied wildly between the companies. Most employees from each company had not seen the operations inside the other's facilities – even at the sites where the facilities are adjacent to each other. More interaction would spark more continuous improvement ideas, more efficiently resolve conflicts, and strengthen the overall relationship.

7.8 Batch Reduction at other UPS Customers

As mentioned before, UPS's desire for early cutoff times is not just to ensure that they are able to process all the packages before the end of the sort but also to ensure that they have enough volume early on in the sort to keep their resources from going idle. Compared to many other UPS customers, the bulk of Amazon's volume arrives relatively early in the day and helps keep the Day sort and early Twilight sort busy. To help with volume availability, UPS should explore what other customers besides Amazon that it might make sense to pick up mostly full trailers earlier in the evening and pick up a less full trailer at the final ship cutoff.

8 References

[1] Abilla, Pete. shmula. 23 April 2010. 14 March 2011 <http://www.shmula.com/no-standard-then-no-kaizen/2035/>

[2] Alexa Web Site Rankings. 15 March 2011 <http://www.alexa.com/topsites>

[3] Amazon.com. 2010 Annual Report. 2011.

[4] Amazon.com. 2010 Q4 Quarterly Report. 2011.

[5] Amazon.com Internet homepage. 15 March 2011 <www.amazon.com>.

[6] Anupindia, Ravi, et al. Managing Business Process Flows: Principles of Operations Management. Upper Saddle River: Pearson Prentice Hall, 2006.

[7] Barnett and William P. The Red Queen Among Corporations: How Competiveness Evolves. Princeton: Princeton University Press, 2008.

[8] Barry, James and Tamara Terry. "Empirical Study of Relationship Value in Industrial Services." Journal of Business & Industrial Marketing, Vol. 23, No. 4 (2008): 228-241.

[9] Chaneski, Wayne. "Hurry Up and Wait." Modern Machine Shop, Vol. 79, No. 9 (2007): 50-52.

[10] Friman, Margareta, et al. "An Analysis of International Business-to-Business Relationships Based on the Commitment-Trust Theory." Industrial Marketing Management, Vol. 31 (2002): 403-409.

[11] Hamilton, D. "Amazon.com." Searcher, Vol. 12, No. 6 (2004): 42-45.

[12] Harris, Chris, Rick Harris and Chuck Streeter. Lean Supplier Development: Establishing Partnerships and True Costs Throughout the Supply Chain. New York: Productivity Press, 2010.

[13] Hopp, Wallace J and Mark L Spearman. "To Pull or Not to Pull: What Is the Question." Manufacturing and Service Operations Management (2004): 133-148.

[14] James, Andrea and Nick Eaton. Seattle Pi. 10 December 2008. 15 March 2011 <http://blog.seattlepi.com/amazon/2008/12/10/amazon-talks-strategy-international-growth/>.

[15] Kumar, Nirmalya, Lisa K. Scheer and Jan-Benedict E.M. Steenkamp. "The Effects of Perceived Interdependence on Dealer Attitudes." American Marketing Association, Vol. 32, No. 3 (1995): 348-356.

[16] Liker, Jeffrey K and Thomas Y Choi. "Building Deep Supplier Relationships." Harvard Business Review December 2004.

[17] Mohr, Jakki and Robert Spekman. "Characteristics of Partership Success: Partnership Attributes, Communication Behavior, and Conflict Resolution Techniques." Strategic Management Journal, Vol.15, No. 2 (1994): 135-152.

[18] Moriarty, Richard and Paul Smith. "Barriers to Entry in Post and Regulatory Responses." Crew, Michael A and Paul R Kleindorfer. Regulatory and Economic Challenges in the Postal and Delivery Sector. New York: Springer Science, 2005. 101-119.

[19] Neely, Andy, Huw Richards, John Mills, Ken Platts, and Mike Bourne. "Designing Performance Measures: A Structured Approach." International Journal of Operations and Production Management, Vol. 17, No. 11 (1997): 1131.

[20] pemanager. The Brown Cafe. 08 January 2010. 14 March 2011 <http://www.browncafe.com/forum/f6/bullfrog-scanners-automated-small-sort-297628/>.

[21] Ring, Peter Smith and Andrew H. Van De Ven. "Structuring Cooperative Relationships between Organizations." Strategic Management Journal, Vol. 13, No. 7 (1992): 483-498.

[22] Rother, Mike and John Shook. Learning to See: Value Stream Mapping to Create Value and Eliminate Muda. Cambridge: Lean Enterprise Institute, 2009.

[23] Schoonmaker, Keith. Morningstar. 08 November 2010. 14 March 2011 <http://www.morningstar.com/cover/videocenter.aspx?id=359077>.

[24] Spector, Robert. Amazon.com: Get Big Fast. New York: HarperBusiness, 2002.

[25] Steward, Mark. ReliablePlant Blogs. 11 December 2008. 14 March 2011 <http://blogs.reliableplant.com/86/standard-work/>.

[26] Swiatek, Jeff. Indy.com. 21 September 2010. 14 Mar 2011 <http://www.indy.com/posts/132746>.

[27] UPS Web Site. 2011. 15 March 2011 <http://www.ups.com/content/us/en/about/index.html>.

[28] Upton, David. McDonald's Corporation (Abridged) 9-603-041. HBS Case. Cambridge: Harvard Business Publishing, 2005.

[29] Wikipedia Article – Lean Manufacturing. <http://en.wikipedia.org/wiki/ Lean_manufacturing >.

[30] Wikipedia Article - Queueing Theory. <http://en.wikipedia.org/wiki/Queueing_theory>.

[31] Womack, James P and Daniel T Jones. Lean Thinking. New York: Free Press, 2003.

[32] Baudin, Michel. Lean Logistics: The Nuts and Bolts of Delivering Materials and Goods. New York: Productivity Press, 2004.

[33] Stavins, Robert N. Harvard University Econ 1661 Lecture. "Fundamentals of Net Present Value Analysis". 2010

[34] Google Maps. 15 March 2011 http://maps.google.com/maps?f=q&source=s_q&hl=en&geocode=&q=1850+Mercer+Road,+40511&aq=&sll=38.020586,-84.451597&sspn=0.008029,0.018818&g=422+Codell+Dr,+40509&ie=UTF8&hq=&hnear=1850+Mercer+Rd,+Lexington,+Kentucky+40511&ll=38.074523,-84.531738&spn=0.004012

[35] Greg. The Toyota System. 16 March 2009. 15 March 2011 <http://www.thetoyotasystem.com/lean_concepts/lean_philosophy_idea_behind.php>.

[36] Landsberger, Henry. Hawthorne revisited: Management and the worker, its critics, and developments in human relations in industry. Ithaca: Cornell University, 1958.

[37] Byers, Ann. Jeff Bezos: The Founder of Amazon.com. New York: Rosen Publishing, 2007.

[38] Engleman, Eric. Amazon Headcount Grows. 21 October 2010 <http://www.techflash.com/seattle/2010/10/amazon-headcount-grows.html>.

[39] Duryee, Tricia. Amazon's Holiday Season Soars By 44 Percent at Peak. 27 December 2010 < http://allthingsd.com/20101227/amazons-holiday-season-soars-by-44-percent-at-peak/>

[40] Dignan, Larry. Amazon posts its first net profit. 22 January 2002 < http://news.cnet.com/2100-1017-819688.html>

[41] Stone, Brad. Sold on eBay, Shipped by Amazon.com. 27 April 2007 http://www.nytimes.com/2007/04/27/technology/27amazon.html

[42] Terry, Phil. The Promise of Amazon Prime. 24 February 2011 <http://dailyartifacts.com/the-promise-of-amazon-prime>

[43] Kohler, Ed. Technology Evangelist. 17 February 2006 <http://www.technologyevangelist.com/2006/02/does_amazon_prime_ch.html>

[44] Hof, Rob. Amazon's Prime Challenge. 12 February 2005 <http://www.businessweek.com/the_thread/techbeat/archives/2005/02/amazons_prime_c.html>

9 Appendix A – UPS Cycle Time Analysis

In Appendix A, we take a closer look at the operations inside of the UPS Hub and propose approaches for managing the cycle time of packages through the hub.

9.1 Cycle Time Overview

If our goal is to achieve later critical pull times for UPS customers given a fixed completion time for processing through the UPS hub, then we must either eliminate unnecessary actions or reduce the cycle time of components of this process. In the following chapter we analyze ways to manage the cycle time within the UPS hub. This includes the time from when a package is unloaded from a trailer at the Inbound doors to when it is reloaded on trailers at the Outbound doors. For a more detailed analysis of the trailer Unloading process see Appendix B.

We can break this process down into its components and group the tasks by the standard Value Stream Mapping buckets of Value Added, Non-Value Added but Necessary, and Waste.

Value Added

- Sorting a package across conveyors
- Scanning a package
- Bagging a small package

Non-Value Added But Necessary

- Package traveling on conveyors to destination

Waste

- Waiting in Queues for an available resource

The value added steps in this process make up only a few seconds of the total cycle time for each package. The non-value added but necessary step of traveling from one trailer to the next makes up a few minutes of the overall cycle time. The greatest contributor by far to the overall processing time is the time packages spend waiting in queues either on a stopped conveyor or waiting in a Small Sort bin

waiting to be bagged. For that reason, the remaining analysis will focus on aspects that affect waiting times. As the Small Sort process is the primary bottleneck in the 81st St Hub and typically has the largest queues of products waiting for resources, most of our focus will be on this process.

9.2 Little's Law

One of the most famous equations in Operations Research is Little's Law (Equation 1 below). Despite its simplicity, it is quite useful for analyzing the flow of products through a process.

Equation 1: Little's Law

$$\text{Average Inventory (I)} = \text{Throughput (R)} \times \text{Average Flow Time (T)}$$

where Throughput is defined as the minimum of the Arrival Rate of packages into the process and the Processing Rate of packages through the process $\rightarrow R = \min(R_i, R_p)$

This formula states that there is a direct correlation between the average inventory inside a system (amount of packages inside the building) and the average time for a product to flow through the process (time spent in the building). Using this formula, we are able to calculate anyone of the three variables given that we know the other two. We also see that, given a constant Throughput, if we are able to reduce the WIP by a certain percentage, then we will also reduce average Cycle Time by an equivalent percentage [6]. For this reason, in much of the remainder of this chapter we will be discussing means to manage the amount of work in process inside a Hub at any time.

9.3 Queues and WIP Management

Since queues add to cycle time, the logical question is why they must exist. Queues are a byproduct of independent (unsynchronized) variability in a process. As previously discussed in section 3.3 of this paper, queues are a form of buffering against variability. Other means of buffering against variability besides excess inventory are having excess capacity or allowing for excess time. At its core, the management practice of Lean is focused on reducing the amount of buffering required.

One of the most popular Lean analogies is the River and Rocks example. An operational process might have many operational challenges (as represented by the Rocks in the analogy). Examples of these can include equipment failure and delays, scheduling issues with delivery of raw materials, labor management, and stock outs of supplies. Excess inventory (as represented by the water in the analogy), is a means to protect against these problems. The goal in Lean is to gradually reduce the inventory until you expose some of the rocks. This inventory reduction will result in blocking and starving of the lines if done too quickly. The goal is remove the rocks and then reduce the water level as opposed to letting the ship crash amongst the rocks. A term often used in Lean to identify the rocks is Exploratory Stress. An example of exploratory stress would be to gradually remove slots from a *Kanban* system (hence putting stress on the system) until issues start to appear [13] [35].

There are many benefits to inventory reduction in addition to the previous mentioned reduction in cycle time. For instance, short queues can reduce the time between the creation of a problem upstream and the detection of the problem downstream, hence reducing rework. Another benefit is that when a downstream process breaks down, it will minimize the amount of WIP that needs to be removed or diverted (e.g. product diverting after conveyor belt breaks). A third benefit is that it can reduce damage that is induced on WIP (e.g. weight of thousands of packages piled up on conveyors and slides can crush packages). In general, being forced to address problems rather than masking them with inventory results in a more efficient solution long term. This has been seen directly in the auto industry where there is close correlation between the amount of inventory that a car company maintains and there overall efficiency levels [13].

9.4 Queue Length Calculation

To help understand what impacts the size of a queue, we can turn to the field of Operations Research. This field has developed mathematical formulas for approximating the expected number of items waiting for a resource given the average utilization, arrival rates and processing rates. A basic

version of this equation that works for a general service time distribution and for multiple servers (M/G/C) can be found in equation 2 [6].

Equation 2: Queue Length Formula

$$Avg.Queue\ Length = \frac{\rho^{\sqrt{2(c+1)}}}{1-\rho} \times \frac{C_i^2 + C_p^2}{2}$$

Where

ρ = Utilization of Servers

c = Number of Servers

C_i = Coefficient of Variation[1] of Arrival Rate

C_p = Coefficient of Variation[1] of Processing Rate

There are many aspects of the process at UPS that prevent the above formula from providing an exact calculation for expected in process inventory. Some of these aspects include the blocking of arrivals due to limited queue size as well as arrival and processing rates that are not independent of the existing queue length (e.g. processing speeds up or arrivals slow down when queue grows). But analyzing the components of this formula provides insight into the primary drivers of queue length. The formula can be broken down into a component that deals with server utilization and a component that deals with process variability (see Figure 22).

Server Utilization		Process Variability
$\frac{\rho^{\sqrt{2(c+1)}}}{1-\rho}$	X	$\frac{C_i^2 + C_p^2}{2}$

Figure 22: Breakdown of Queue Length Formula

As utilization approaches 100%, the left side starts to grow exponentially. As process variability approaches zero, the queue length approaches zero. This is expected as a deterministic system (no

[1] Coefficient of Variation = Standard Deviation / Mean

stochastic variability) with utilization below 100% should have no queue. This can be seen graphically in Figure 23 which shows the impact of changing utilization and variability on queue length [6].

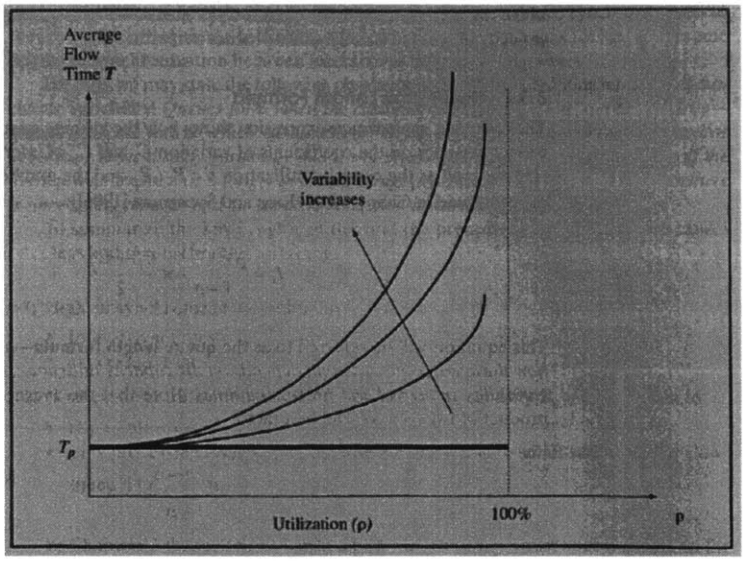

Figure 23: Effect of Utilization and Variability on Queue Length

9.5 Recommendations

Now that we have established some of the benefits to WIP reduction as well as identified some of the drivers that impact WIP, we now propose four recommendations that will help manage long term WIP inside the UPS Hub.

9.5.1 Reduce Utilization of Small Sort

A common misconception by many managers is that an operation is only running efficiently when all the equipment is running 100% of the time. In practice, it is much more efficient to design a process with spare capacity as average queue length and wait time grows dramatically as equipment approaches 100% utilization. The optimal utilization target is dependent on the amount of variability, but targets around 80-90% are common. One of the most revolutionary and overlooked steps taken by Toyota during the journey to a more efficient factory was a conscious shift from inventory buffering to capacity buffering [13]. At the 81st ST Hub, the Small Sort process is overly burdened and adding additional capacity will dramatically improve the operational efficiency of the hub. As the hub moves to

newer generation small sort equipment, they should maintain enough excess capacity to allow for 10-20% more theoretical throughput than is required.

9.5.2 Reduce Variability in Arrival and Processing Rates

As discussed above, fluctuations in the arrival rate and processing rates increase average queue length. An example of variation in arrival rate can be seen in Figure 32 in Appendix B where the unload rate of different employees can vary greatly. Fluctuation in employees' sort rate is an example of processing rate variation. With the traditional Small Sort equipment, UPS does not systemically collect sort rate data, but Figure 24 shows fluctuations in bagging rates across days and across 12 employees.

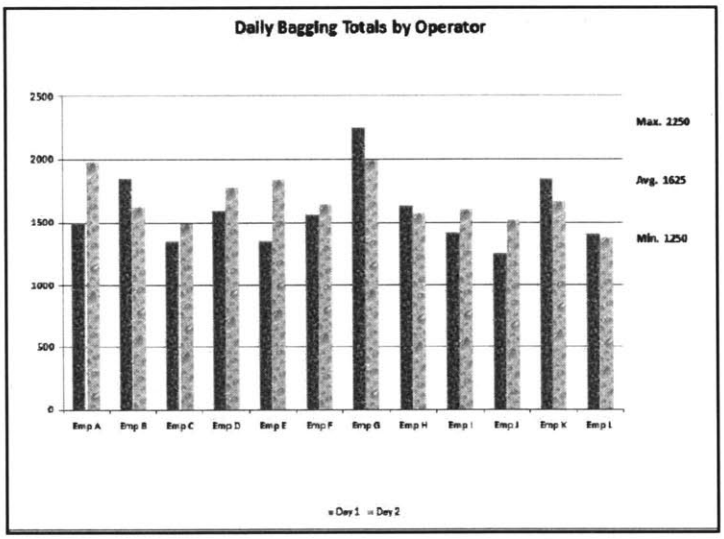

Figure 24: Fluctuations in Bagging Rates of Employees

One of the most effective ways to drive down variations in arrival and processing rates is to use the exploratory stress approach described in section 10.3. One can put stress on the process by gradually reducing the amount of inventory permitted to be maintained in the system. The next section describes a means for accomplishing this.

9.5.3 Implement a Pull System for Small Sort

There is much confusion about the definition of a Pull process, but the simplest definition is any process that explicitly limits the amount of work in process that is allowed to be in the system. Once a pull system has reached the WIP limit, it only allows inventory to move from one operation to the next

when the subsequent operation is ready to process it. Theoretically, there is always some physical constraint on WIP (i.e. all conveyors in building completely full), but the WIP limit in practical pull systems is explicitly stated and in generally small [13].

We propose the implementation of a Pull system on the conveyors leading up to the Small Sort process. We can better explain what this might look like by looking at some diagrams representing a rough approximation of the Hub layout. Figure 25 represents a rough generalization of the UPS Hub layout. In this diagram, Inbound doors are at the bottom with conveyors leading out of the trailers. Employees there sort packages either to the lighter colored belt that leads to the Small sort or to the darker belts that lead directly to the Outbound doors. The darker belts have a series of intersections which allow packages to get from any Inbound door to any Outbound door. Packages come into the Small sort where they are placed into bags. The bags eventually get to the same darker colored belts that lead to Outbound.

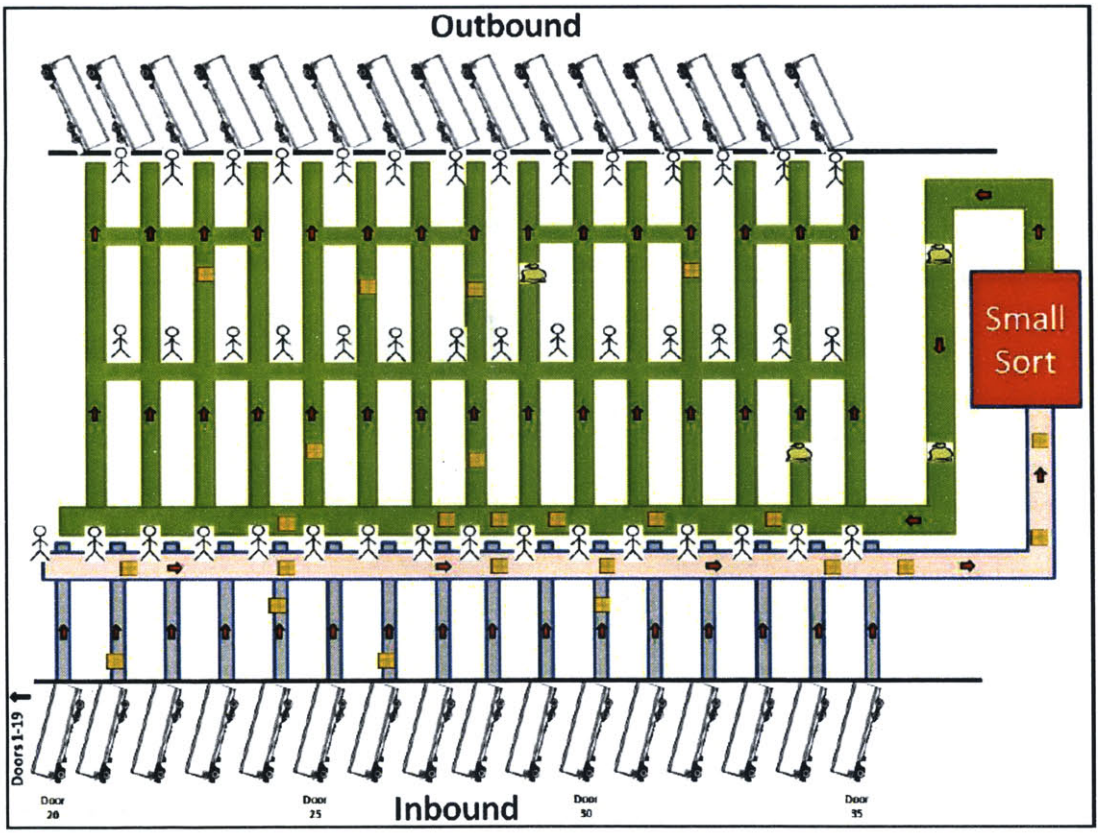

Figure 25: Generalization of UPS Hub layout

In practice today, what we often observe is the case where the arrival rate of Small packages dramatically exceeds the processing rate of the Small sort and therefore, the conveyor leading up to the Small Sort completely backs up. This is a massive conveyor that is five feet wide and over 200 yards long containing thousands of packages. The end result is that the employees unloading the trailers often are forced to pause unloading as they wait for the belts to clear up. A representation of this scenario is depicted in Figure 26. This is a result of what some people refer to as the "hurry up and wait" syndrome where inventory is rushed to where it needs to be and then waits [9].

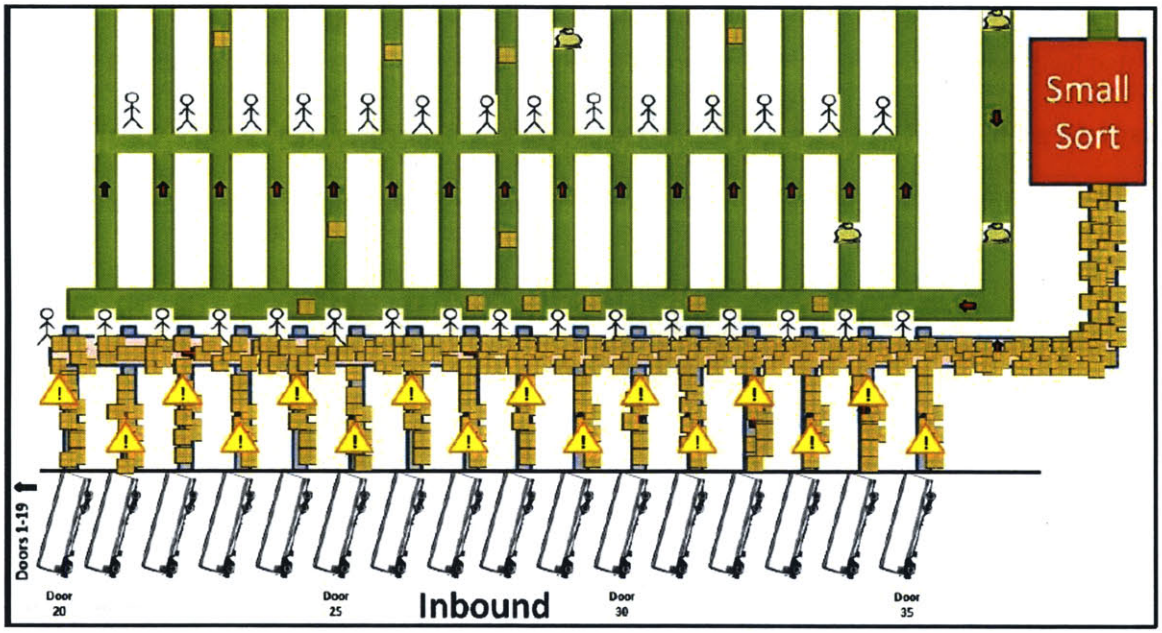

Figure 26: Hub representation with Small Sort conveyor completely full

If a Pull system were implemented it would explicitly limit the amount of material permitted on the conveyors to a level much lower than what is sometimes seen today. An example of what this might look like can be found in Figure 27.

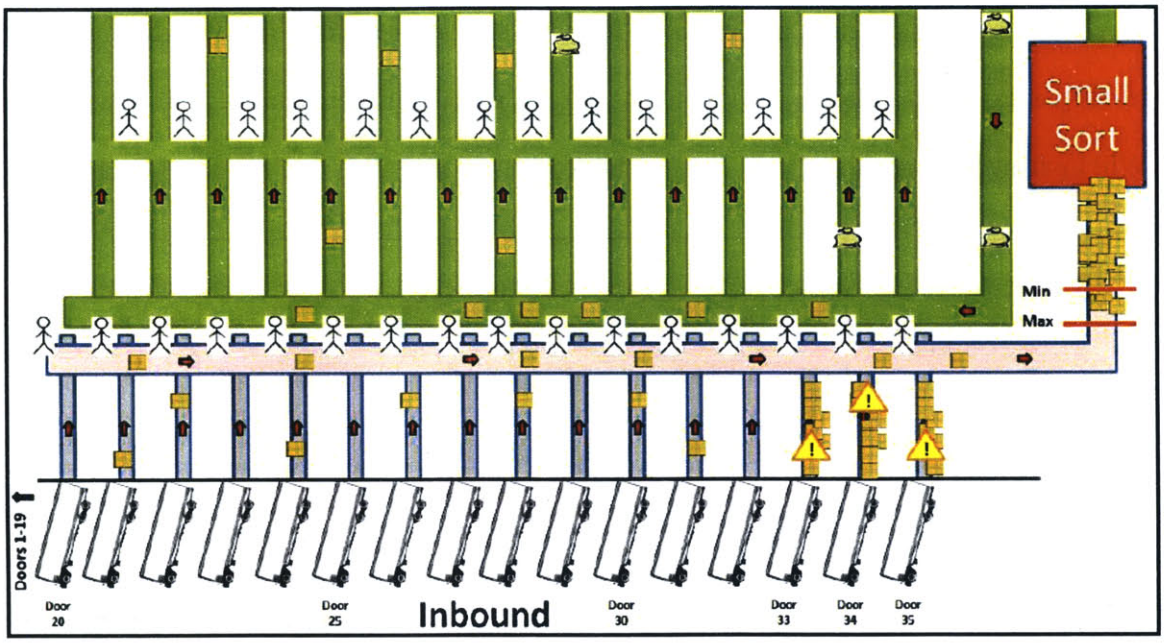

Figure 27: Hub with a Pull System limiting WIP

One way to implement this process would be to reserve a few doors for trailers that contain primarily Small packages (several of UPS's City Pickup customers create trailers that are almost exclusively small packages). In Figure 27, we represent this with doors 33-35. If the unloading of doors 1-32 generates an arrival rate that is less than the Small Sort processing rate and the unloading of all doors generates an arrival rate that is greater than the processing rate, then UPS would have complete control over the queue length. Sensors on the conveyor at the spots marked Min and Max would trigger signals to workers at Doors 33-35 whether or not they need to process. This signal could either be some type of light or a direct control of the motors running the conveyors leading to those doors.

With less WIP on the conveyors, there will be reduced cycle time for the packages in doors 1-32 along with all the other previously mentioned benefits of reduced WIP. Having the product in trailers 33-35 just sit longer in the trailers rather than on the belts is just another example of the "Hurry Up and Wait" syndrome, but ideally having full trailers untapped for longer into the shift will provide more visual signals about when the trailers are actually needed for Volume Availability.

Once a Pull system is implemented, it creates an easy means of gradually reducing inventory levels (i.e. exploratory stress). This would be done by initially implementing a pull system with a high WIP limit and then gradually reducing it as the "rocks" are identified (see Figure 28 for a depiction).

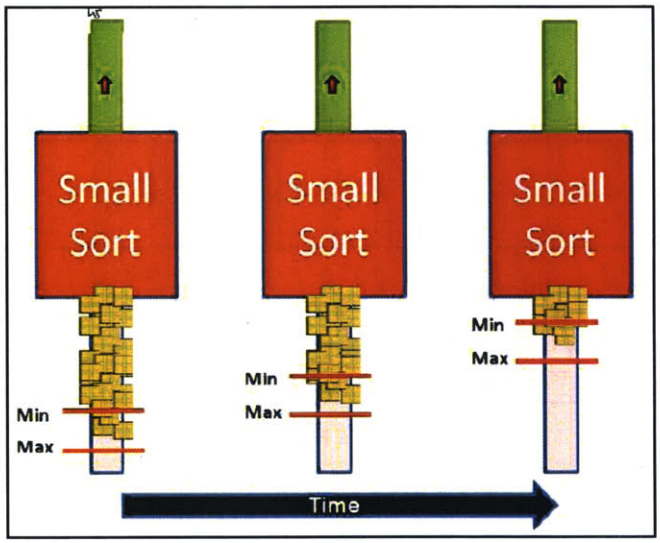

Figure 28: Gradual reduction of WIP limits

9.5.4 Implementation of Cycle Time and WIP Metrics

As an organization only gets the behaviors that they measure, our final recommendation is for UPS to implement an automated Cycle Time metric. This should be fairly straight forward given the data that is already available in UPS systems. Overhead bar code readers currently scan all packages as they are unloaded from the trailers and are able to read over 90% of the packages as they are unloaded. All packages are also scanned as they are loaded back into the Outbound trailers. For these 90% plus packages an accurate cycle time can be calculated. Since there should be no bias as to which packages are unsuccessfully scanned upon entry into the hub, the average cycle time of the 90% of packages should accurately reflect the average cycle time of all packages. Because of the missing scans, an accurate measurement of the number of packages inside the facility (WIP levels) is not directly measurable, but an approximation can be made via Little's Law using the cycle time and throughput.

For packages that go to the Small Sort process, we recommend breaking the cycle time metric down into the average time from Unloading to being sorted at the Small Sort, average time waiting in the

Small Sort bins, and average time it takes bags to travel from the Small Sort to the Outbound trailers. Because no scan happens today at the time that small packages are sorted, we cannot systemically separate out the travel time to the Small sort and the time packages spend waiting in a bin. If newer technology Small Sort equipment is installed, then this breakdown will be feasible as this equipment does require a scan upon sorting.

To sample what this data might look like, we manually collected data for a few hundred packages and plotted the cycle times as both a histogram and a time series (see Figure 29). We expected to see a trend in the cycle time over time since the queue length leading up to the Small Sort was changing throughout the shift. Instead we see simultaneously unloaded packages leaving the Small Sort area at dramatically different times. This suggests that the measurements are dramatically skewed as a result of the time that packages spend waiting in bins. Therefore, without being able to record the sort time, breaking cycle times down into component parts is not extremely useful.

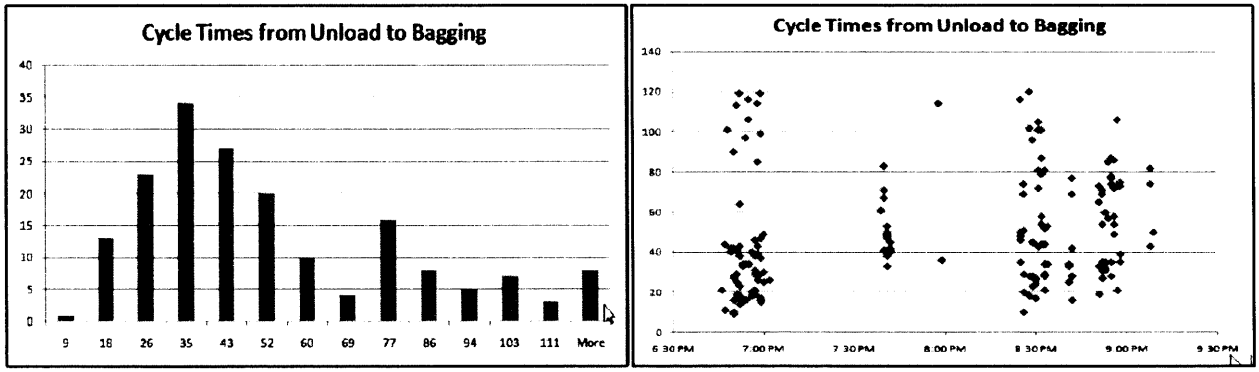

Figure 29: Manual Cycle Time Measurements from Unloading to Bagging

With manual collection of cycle time data, we were not able to collect enough data to break down cycle time by destination. With an automated collection process, metrics can be easily broken down by destination. By filtering the metrics to only destinations with Hot Pull times, the data would be much more insightful. For instance, when we manually measured the cycle times from when Small Bags were created to the time the bags were scanned on the trailer, we saw a wide distribution with a heavy right tail (see Figure 30). This suggest that many of the bags were destined for trailers that were not scheduled to

depart on that shift. Therefore, the bags likely waited on the conveyors outside the trailers for extended periods of time waiting to be scanned.

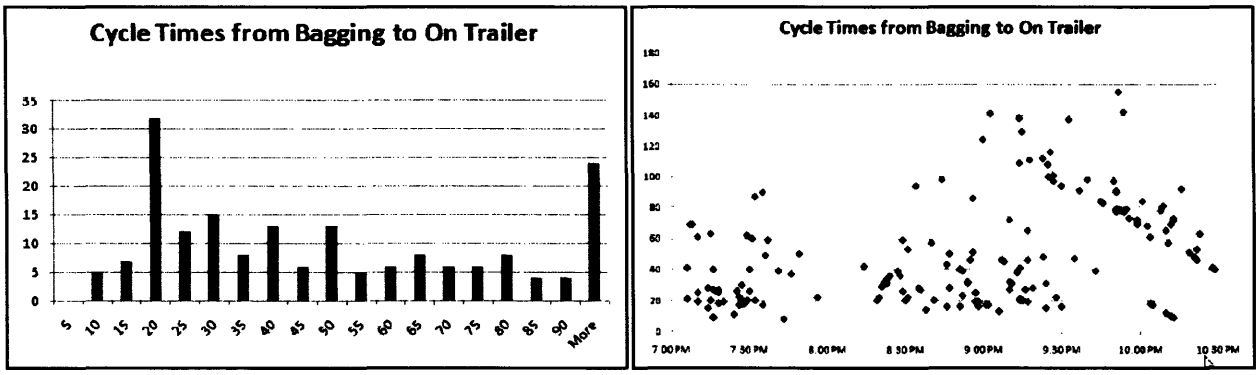

Figure 30: Manual Cycle Time Measurements from Bagging to Loading

10 Appendix B – UPS Trailer Unload Analysis

Note: Some of the numbers in this appendix have been adjusted to protect confidential information.

In Appendix B, we compare the theoretical amount of time it should take to unload a full Amazon trailer (based on time studies of individual tasks) to the actual times recorded during the most recent holiday season.

10.1 Actual Recorded Unload Times

Figure 31 shows a histogram representing the amount of time taken to unload Amazon trailers at the UPS hub. This data was collected from UPS electronic systems and represents operations on the Twilight sort during a three week window during the holiday season. Given the challenges to complete all sorting during the assigned shift in the holiday season, UPS is forced to turn trailers as quickly as possible. Therefore, the histogram should represent typical trailer processing times for the UPS hub given strong schedule pressure.

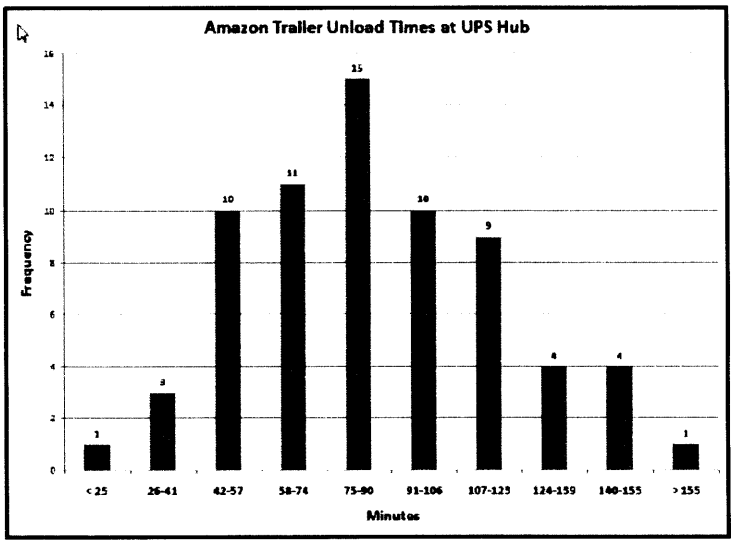

Figure 31: Historic Trailer Unload Times at UPS Hub

In this data we are unable to accurately determine the percent that each trailer was full upon the start of the unload process. From the left tail of the histogram it is clear that not all trailers were 100% full when arriving to UPS. This skews the histogram compared to a data set that only included full

trailers. From the data that we have, we see that the average unload time was 87 minutes, 41% of the trailers take over 90 minutes to unload, and 15% took over 120 minutes to unload.

10.2 Theoretical Unload Times

For comparison, it is insightful to compare the above times to the theoretical amount of time that we calculate it should take to unload a full trailer. To accomplish this, we must 1) breakdown the tasks involved in unloading a trailer (both value and non-value added), 2) collect some time study data, and 3) understand the product mix that makes up a typical Amazon trailer.

The major steps to unloading a trailer include:

1. Checking In Trailer / Opening Trailer / Deploying Conveyor
2. Unloading Packages onto extended conveyor
3. Tearing down Gaylords and Moving Pallets
4. Retrieving extended Rollers to unload last few shelves of trailer

Time studies of each of these tasks were performed to determine average time each task takes.

1. Checking In the trailer, etc. takes about 4 minutes
2. Package unloading varies by product type (see below)
3. Gaylord and Pallet clearing take about 90 seconds to break down two pallets
4. Rollers take about 4.5 minutes to retrieve and hook up

The time that it takes to unload packages depends on the type of package being unloaded. For Amazon, we have broken the packages down into three types. 1) Large packages that are Fluid loaded and do not need to go to the Small Sort , 2) Medium packages that are in Gaylords and do need to go to the Small Sort, and 3) Flat packages that are in Gaylords and need to go to the Small Sort. Figure 32 shows results from a time study of unload times for large packages. We observed three different teams of two Unloaders and saw rates that averaged between 70 and 100 seconds per 100 large packages unloaded based on the team. In the later analysis we will use a time of 95 seconds for a large package unload rate.

We performed similar analysis for the other types of packages and saw average unload times of 54 seconds for unloading 100 medium packages and 21 seconds for unloading 100 flat packages. To be conservative, in later analysis we will use 60 seconds and 25 seconds for the respective unload times.

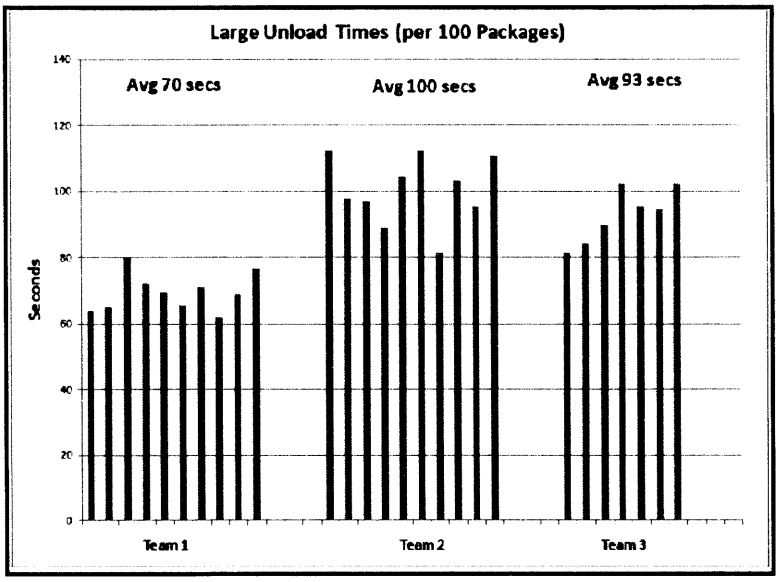

Figure 32: Unload times for Large Packages

10.3 Theoretical Calculations

In the following calculations, we will assume the values below

- A trailer that is 100% fluid loaded holds approximately 4,000 large packages
- A 48' trailer can hold 24 pallets and therefore 24 Gaylords
- A Gaylord holds approximately 275 medium packages
- A Gaylord holds approximately 750 flat packages
- Mix of packages is 15% flat, 45% medium, and 40% large packages

This mix of packages is relatively consistent from day to day with only small fluctuation. But not all trailers are loaded with the exact mix as it comes out of the Amazon Outbound processes. Since some product is staged in the Amazon shipping area prior to being loaded on trailers, it is possible to see trailers that are 100% fluid loaded with large packages or 100% loaded with Gaylords of small packages. Therefore, in the below analysis, we will analyze three scenarios – 1) the historical mix, 2) a trailer that consists of all smalls, and 3) a trailer that is completely fluid loaded.

Typical Product Mix

Table 2 shows the time it should take to unload a 100% full trailer given the typical Amazon product mix. About 61 minutes are spent unloading packages. When including the opening of trailers, teardown of Gaylords, and retrieval of rollers, the time is approximately 79 minutes.

Table 2: Trailer with a typical Product Mix

Package Type	Product Mix	Number Packages	Gaylords Used	% Space	Unload Rate per 100 pkg	Step 1 Trailer Opening	Step 2 Unload Time	Step 3 Pallet Clearing	Step 4 Rollers
Flat	15%	816	2	8%	25		204	90	
Small	45%	2447	9	38%	60		1468	405	
Fluid	40%	2175		54%	95		2066		
Fixed Time						240			270
						Value Add Time:	**62:31**	**Total Time:**	**79:04**

Trailer that is 100% Fluid Loaded (All large packages)

Table 3 shows the time it should take to unload a 100% full trailer given that there are no small packages in Gaylords. About 64 minutes are spent unloading packages. When including the opening of trailers and retrieval of rollers, the time is approximately 72 minutes.

Table 3: Trailer that is 100% Fluid loaded

Package Type	Product Mix	Number Packages	Gaylords Used	% Space	Unload Rate per 100 pkg	Step 1 Trailer Opening	Step 2 Unload Time	Step 3 Pallet Clearing	Step 4 Rollers
Flat	0%	0	0	0%	25		0	0	
Small	0%	0	0	0%	60		0	0	
Fluid	100%	4000		100%	95		3800		
Fixed Time						240			270
						Value Add Time:	**63:21**	**Total Time:**	**71:50**

Trailer that is 100% Small and Medium packages in Gaylords

Table 4 shows the time it should take to unload a 100% full trailer given that the trailer consists of 24 Gaylords (no fluid loaded packages) and that the ratio between Flat and Medium packages is the same as in the typical split. About 64 minutes are spent unloading packages. When including the opening of trailers, teardown of Gaylords, and retrieval of rollers, the time is approximately 91 minutes. Note that while this is the slowest type trailer to unload and takes about 12 minutes longer to process, this time is almost exclusively dedicated to tearing down Gaylords and moving pallets and not to value added work.

Table 4: Trailer that is 100% full of small packages in Gaylords

Package Type	Product Mix	Number Packages	Gaylords Used	% Space	Unload Rate per 100 pkg	Step 1 Trailer Opening	Step 2 Unload Time	Step 3 Pallet Clearing	Step 4 Rollers
Flat	25%	1053	2	8%	25		263	90	
Small	75%	6000	22	92%	60		3600	990	
Fluid	0%	0		0%	95		0		
Fixed Time						240			270
						Value Add Time:	**64:48**	**Total Time:**	**90:53**

10.4 Comparisons of Theoretical to Historical times

While a perfect comparison cannot be made because we are unable to determine which of the historic trailer were full upon arrival, some interesting observations can be made. About 41% of the trailers unloaded during this holiday season took longer than 91 minutes (the worst case scenario) to unload. As this unload time is critical to determining the critical pull time that UPS needs, it is worthwhile to understand what can drive some of these variances. Also, given that 22% of time is dedicated to non-value add tasks in the typical case (30% in the case of all Smalls), it is worth exploring ways to reduce the need or time it takes for these tasks. Struggling with cardboard and pallets is a large bulk of this time. Having to use manual rollers because of insufficient length of the automated rollers is also source of lost time.