Efficient Scheduling to Reduce Setup Times and Increase Utilization in a Multiple-Part Manufacturing System

by

Greg Puszko

Bachelor of Science in Mechanical Engineering
Massachusetts Institute of Technology, 2013

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN MANUFACTURING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2014

©2014 Greg Puszko. All rights reserved.

The author hereby grants MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document, in whole or in part, and to grant others permission to do so.

Signature redacted

Author

Greg Puszko
Department of Mechanical Engineering

Certified by

Stanley Gershwin
Senior Research Scientist
Thesis Supervisor

Accepted by

David E. Hardt
Ralph E. and Eloise F. Cross Professor of Mechanical Engineering
Chairman, Committee for Graduate Students
Efficient Scheduling to Reduce Setup Times and Increase Utilization in a Multiple-Part Manufacturing System

by

Greg Puszko
Bachelor of Science in Mechanical Engineering
Massachusetts Institute of Technology, 2013

Submitted to the Department of Mechanical Engineering on August 15th, 2014,
in partial fulfillment of the requirements for the degree of
Master of Engineering in Manufacturing

Abstract

Two scheduling procedures were developed to reduce the total setup time and increase utilization in a multiple-part-type manufacturing system with sequence-dependent setups. These scheduling systems focus on reducing setup time by taking advantage of part family setups. A part family is a set of parts that have similar geometries or use the same fixturing or tooling in a flexible machine, thus the labor involved in transitioning between two parts in a family is simpler, and therefore quicker, compared to the average machine setup.

One of the scheduling procedures, the setup matrix methodology, was tested through implementation in an actual manufacturing system. The setup matrix is a data-set that contains the setup time between any two parts that are fabricated on a machine. By knowing either the exact or relative values for the setup times between all parts, the setup matrix can be used to schedule a set of production orders to minimize the total setup time. The setup matrix methodology demonstrated significant reductions in total setup time, with little to no adverse effects in systems with large setup time to production time ratios (of 1:5 or larger). In systems with low setup time to production time ratios (of 1:10 or smaller), the cost of having particular orders fall behind due to this scheduling procedure is non-negligible compared to the saving generated from the removed setup time. These two values must be compared to determine whether or not it is beneficial to implement this scheduling system in those particular areas.

The other scheduling procedure, QPR Scheduling, was tested through simulation. The QPR Scheduling procedure is a derivative of the standard (Q,R) ordering policy used in supply chain design, which includes an additional reorder point to take advantage of part family setups on a machine. In the simulation, the QPR Scheduling procedure also generated considerable reductions in total setup time, without producing backorders or significant increases in inventory.

Thesis Supervisor: Stanley B. Gershwin
Title: Senior Research Scientist
This page is intentionally left blank.
Acknowledgements

Thank you to Gabriel Kelly and Kerry McNamara for their constant guidance at Waters Corporation over the course of this project in doing everything they could to see it success.

Thank you to everyone at Waters Corporation for their assistance throughout this project, including David Terricciano and Jim McPherson for making this collaboration between MIT and Waters possible.

Special thanks to Wayne Wentworth, Greg Martin, and Mike Catz for their enthusiasm and support during the implementation phases of this project.

Thank you to my project partners, Arjun Chandar and AJ Perez, for their collaboration in producing a result we can all be proud of.

Thank you to our advisor, Stanley B. Gershwin, for teaching the three of us to always hold a healthy amount of skepticism.

Thank you to David Hardt and Jose Pacheco for welcoming me into the Masters of Engineering in Manufacturing program, and for their leadership throughout the past 12 months. I also want to thank all of the other students in the MEngM program with me, for making this past year so memorable.

And finally, thank you to my family for their continuous support throughout my five years here at MIT.
This page is intentionally left blank.
# Table of Contents

Abstract ........................................................................................................................................... 3  
Acknowledgements .......................................................................................................................... 5  
Table of Contents ............................................................................................................................ 7  
List of Figures and Tables............................................................................................................... 9  

1 Introduction ..................................................................................................................................... 11  

2 Background and Problem Statement ............................................................................................ 13  

2.1 Background on Waters Corporation .......................................................................................... 13  

2.2 The Waters Advanced Manufacturing Center ......................................................................... 14  

2.3 Typical Flow of Production and Parts through Machining Center ........................................... 16  

2.3.1 Organization .................................................................................................................... 16  

2.3.2 Planning and Scheduling ............................................................................................... 16  

2.3.3 Production ...................................................................................................................... 18  

2.4 Issues with the Waters Manufacturing System ......................................................................... 20  

2.4.1 Overview of Problems Discovered .................................................................................. 21  

2.4.1.1 Low Machine Utilization ......................................................................................... 21  

2.4.1.2 Low On-Time Delivery of Parts to Assembly ......................................................... 21  

2.4.1.3 Significant Day-to-Day Adjustments to Production Schedule ................................ 22  

2.4.2 Diagnoses of Problem Causes .......................................................................................... 23  

2.4.2.1 Frequent Occurrence of Redundant Setups .............................................................. 23  

2.4.2.2 Delays in Turning and Milling Cascade to System-Wide Delays ............................... 24  

2.4.2.3 Poor Production Scheduling .................................................................................... 24  

2.4.2.4 Inefficient Data Collection and Feedback to Planning ........................................... 25  

2.5 Problem Statement ................................................................................................................. 27
Theoretical Review of System Improvements................................................................. 28

3.1 Optimized Lot Sizing ........................................................................................................ 28

3.2 Part Supermarket ............................................................................................................. 30

3.3 Production Scheduling: The Setup Matrix ........................................................................ 34
  
  3.3.1 Format of the General Setup Matrix ........................................................................ 35
  
  3.3.2 Special Cases of the Setup Matrix ........................................................................... 35
  
  3.3.3 Application of the Matrix to Scheduling ................................................................. 38

3.4 Production Scheduling: QPR Inventory Control Policy ............................................. 39

4 System Improvements Implementation ........................................................................... 45

4.1 Optimized Lot Sizing ...................................................................................................... 46

4.2 Part Supermarket ............................................................................................................ 48

4.3 Setup Matrix .................................................................................................................. 49

4.4 Automated Data Collection on Production System Performance ............................... 51

4.5 QPR Scheduling ............................................................................................................. 52

5 Results & Discussion ........................................................................................................ 60

5.1 Optimized Lot Sizing and Part Supermarket ............................................................. 60

5.2 Setup Matrix Scheduling ............................................................................................. 63

6 Conclusions and Future Work .......................................................................................... 72

References ............................................................................................................................ 77
List of Figures and Tables

Figure 1: Schematic of The Waters Machining Center of Excellence.............................15

Figure 2: Utilization of Four Machines in Milling....................................................21

Table 3: On-Time Delivery Percentage and Average Days Late from Turning and Milling......22

Figure 4: Part Supermarket.........................................................................................32

Figure 5: Illustration of the (Q, R) Policy for Inventory Review..................................40

Figure 6: A Demonstration of the QPR Scheduling system.........................................42

Table 7: The Ten Parts Selected for Implemented System Modifications......................45

Table 8: Adjusted Production Lot Sizes.........................................................................47

Table 9: Fixed Parameters of the QPR Scheduling Simulation.....................................55

Table 10: Variable Parameters of the QPR Scheduling Simulation.................................56

Table 11: Results of the QPR Scheduling Simulation....................................................57

Figure 12: Total Setup Time vs. Total Average Inventory, QPR Simulation......................58

Table 13: Production Orders Eliminated Due to New Lot Sizes..................................61

Table 14: Setup Hours Saved through Lot Sizing Methodology....................................62

Table 15: Examples of Previous and Optimized Production Schedules..........................63

Table 16: Standard Setup Time vs. Part Family Setup Time in Turning..........................64

Table 17: Standard Setup Time vs. Part Family Setup Time in Milling...........................64

Table 18: Setup Hours Saved through Setup Matrix Scheduling...................................65

Figure 19: Example System to Demonstrate Time to Scheduling Stability....................67

Figure 20: The Time to Scheduling Stability for Example System with Adjusted Lot Sizes.....68

Table 21: Time to Scheduling Stability in the Turning Department..................................69

Table 22: Time of Scheduling Stability in the Milling Department.................................70
1 Introduction

The purpose of this thesis is to illustrate how smarter scheduling procedures, designed to sequence orders to minimize the total setup time, can increase the utilization of a multiple-part-type manufacturing system with sequence-dependent setups and provide overall value to a production system.

This thesis is based off of research conducted at Waters Corporation, an analytical instrumentation company located in Milford, Massachusetts. Waters’ main product division is a line of high-performance liquid chromatography machines, assembled and partially-manufactured in their Milford facility. In 2014, the heads of global manufacturing and finance set production targets for the Milford facility to increase its total production credits by 18% from 2013 to 2014. Therefore, the purpose of this project is to develop a methodology to increase capacity in their manufacturing facility, which can be applied to a significant portion of their production line, and provide management with a framework to continuously improve the Milford manufacturing production system to reach the 18% capacity increase over the next year.

The facility is a multiple-part-type, flexible manufacturing system with sequence-dependent setups. After an initial investigation into the operations of the facility, issues were discovered with the production schedule that lead to frequent, redundant part setups on machines, creating widespread inefficiencies throughout the production line. Small production lot sizes with long setup times caused superfluous setups to occur multiple times per month. These repeated setups could be eliminated by increasing the production batch sizes for certain parts, and by holding inventory. Further, the scheduling procedure was ignorant to the sequence dependent setups that existed on the production floor; setup times can be greatly reduced if parts of the same type or geometry are scheduled back-to-back. Through changes in the production lot sizes and production schedule, part setups can be shortened or eliminated on numerous machines in various departments, allowing the production line to recapture some of its capacity.

Specifically, this thesis will focus on two scheduling procedures that are applicable to multiple-part-type manufacturing systems with sequence-dependent setups. These scheduling produces are designed to reduce the total setup time for a machine, and increase its utilization. One of the scheduling procedures, the setup matrix methodology, was applied to Waters’
manufacturing facility to test its effects on a real production system, and determine if the shorter setup times speculated can be achieved. Another scheduling procedure, dubbed the QPR Scheduling policy was tested and verified thru simulation. Both scheduling procedures, along with the altered production lot sizes, demonstrated savings in setup time that outweighed any other costs which may have incurred, such as added inventory or order tardiness, and provided overall value to the production system.

Section 2 will go into detail about the current manufacturing system at Waters Corporation, including a description of their production and assembly facility, the typical production sequence for a part fabricated in the Machining Center, and the issues leading to system-wide inefficiencies that were uncovered in our investigation. Section 3 introduces four possible system modifications intended to increase machine utilization by reducing or eliminating setups, including new scheduling procedures to take advantage of sequence-dependent setups. Section 4 will discuss the actual implementation of the system changes at Waters Corporation, and the simulation used to test the QPR Scheduling procedure. Section 5 will examine the results of the implementation phase, and the effects of the system changes on production. Finally, Section 6 will outline the recommendations for Waters Corporation going forward to achieve additional improvements in their manufacturing facility, and potential directions for future elaboration into the QPR Scheduling methodology.
2 Background and Problem Statement

The goal of this project was to develop system improvements that could be applied throughout the Machining Center at Waters Corporation to generate significant increases in productivity. The executives at Waters were interested in a methodology that could be pragmatic to a variety of areas in the Machining Center, and not just a project that focused on a microcosm of the shop to optimize it alone. Thus, the idea was established to develop system changes for the shop to improve productivity that could be applied to multiple departments, areas, machines, parts, etc., and then demonstrate their value through implementation on a portion of the Machining Center and a limited number of parts.

In order to develop system changes to improve the productivity of the Machining Center, we first had to ascertain how the production system works. Then, we had to investigate how inefficiencies emerge on the shop floor, and diagnose the causes for those inefficiencies, before we can create improvements that eliminate those causes. Hence, the project began by gaining a better understanding the production system and how and why inefficiencies matriculate on the shop floor.

This chapter was written in conjunction with Chandar [1] and Perez [2], borrowing heavily from those works.

2.1 Background on Waters Corporation

Waters Corporation is an analytical instruments company that develops test equipment used in pharmaceutical, industrial, and academic research laboratories. Their two main product divisions are their Biochemical and Chemical Analysis Division, based in Milford, Massachusetts, and their Physical Testing Division, based in Manchester, England and Wexford, Ireland. The Biochemical and Chemical Analysis Division produces liquid chromatography instruments, while the Physical Testing Division produces high-end mass spectrometry instruments. Each division develops and manufactures their standalone products, and all the consumables, chemicals, and accessories which feed or support their particular machines. Waters also maintains a global support network of authorized service centers around the world.
that manage local installation services, training, technical support, repair, and replacement part services. This thesis focuses on the development and implementation of manufacturing system improvements in Waters’ Advanced Manufacturing Center, which is located at their Biochemical and Chemical Analysis Division based in Milford, Massachusetts.

2.2 The Waters Advanced Manufacturing Center

The Waters Advanced Manufacturing Center houses a 50,000 sq. ft. Machining Center of Excellence, a 29,000 sq. ft. Advanced Instrument Assembly & Accessory Kitting Operation facility, and an 8,500 sq. ft. Class 10,000 Clean Room for Optics, Micro Valves, and Critical Parts. This thesis will focus on the operations of the Machining Center of Excellence (referred to as the “Machining Center”), which produces precision-machined metal components for the final assembly of instruments and consumables.

The Machining Center operates for 24 hours per day, 6 days per week, 52 weeks per year, producing 2.7 million parts annually covering 1500 unique SKUs. The Center is divided up into four main departments: NC Turning, NC Milling, Valve Cell, and Column Cell. The Valve Cell and Column Cell are laid out as production cells, with a variety of different machines arranged next to each other that complete different operations on the same part. Each Cell has a dedicated part type that it produces in high volumes (pump check values in the Valve Cell, consumables for the LC instruments in the Column Cell), and contains all the machines needed to complete their respective parts.

On the other hand, The NC Turning and NC Milling departments (typically referred to as simply “Turning” and “Milling”) focus on high mix production, and produce a wide variety of parts of all different types, functions, materials, and geometries. Turning and Milling are set up in a job shop format, where machines that are the same type (lathes and mills, respectively) are grouped together in the same location. The yearly production volumes of parts fabricated in these departments range between 20 and 30,000 units per year. Thus, even in these departments that focus on high-mix rather than high-volume production, there is a distinction between “high volume” and “low volume” parts and products (typically, any part is produced at 1000 units per year or more is considered to be a high volume part for these departments). Each of these departments also has its own utility area, consisting of deburring machines and simple cleaning machines used to perform secondary operations.
A majority of SKU’s in the Machining Center start their production in the Turning department, then are directed to the Milling department for continued operation, and finally are sent to another area in the Manufacturing Center for additional processes (laser marking, EDM drilling, lapping, passivation, critical cleaning, etc.). In fact, all parts that have operations in Turning and Milling must be sent another department in the Machining Center in order to be completed, so the Turning and Milling areas feed into most other departments in the Manufacturing Center. For reasons that we will delve into later, the focus of this thesis will be on improvements in the Turning and Milling departments (see Section 2.4).

The Machining Center also houses a Model Shop which maintains its own CNC and manual mills and lathes, a fused deposition modeling 3D printer, and micro-machining capabilities, but the Model Shop is used by the New Product Division in Waters to prototype parts currently undergoing development, and it not used for the production of products to customers.

Figure 1: A Schematic of the Machining Center. 1 = Milling Department, 2 = Turning Department, 3 = Valve Cell, 4 = Column Cell, 5 = Model Shop.
2.3 Typical Flow of Production and Parts through Machining Center

Since the implementation of this work focuses on production improvements in Waters’ Turning and Milling department, this example will detail the typical process for a part which starts in the Turning department, continues on in the Milling department, and then is delivered to another department in the Manufacturing Center for additional operations.

2.3.1 Organization

Both the Turning and Milling departments consist of a number of machines, which Waters refers to as “work centers”. When performing operations or undergoing setups or teardowns, each machine will be staffed by one machinist. Depending on the machinist, they are either trained to operate on particular machines in the department, or every machine within a department. Directly supervising the machinists in each department are the section leaders. Along with the typical duties machinists have of operating the machines and producing parts, the section leaders manage the orders on the production schedule for their department, assign the machinist to work on particular machines based on those orders, and run the debrief meetings at the end of each shift. Each department will have 1-2 section leaders working per shift. Directly above the section leaders are the department supervisors, in charge of managing the operations of their specific department and making sure that production is on schedule. The supervisors are not tasked with machining, but work directly with the machinists and section leaders on a day-to-day basis.

2.3.2 Planning and Scheduling

The production schedule for all departments in the Manufacturing Center is controlled by Water’s Schedule and Planning (SAP) system. The process of creating the production scheduling for Turning and Milling starts with Sales department, which generates the sales forecast for Waters products for the next 18 months. Within this 18-month outlook, the sales department also includes a comprehensive sales forecast for the next 3 months, which outlines detailed predictions on sales volumes broken down by product. With this short-term forecast, the Assembly area in the Manufacturing Center has reliable numbers on how many units they will
build over the next quarter. By breaking down the part list and quantity for each of these items, the exact number of units that need to be produced by the Machining Center to feed the Assembly area can be determined.

Parts that are produced by the Machining Center that feed into Assembly are held in an inventory location near the Assembly area. When an assembly order begins, the necessary parts are pulled from their respective inventory locations. This depletion is recorded by the SAP system. Through the predicted sales forecast, the SAP system creates a schedule for when each unit should begin assembly, and thus it is known when parts will be needed by the assembly areas to create each product. With this information, SAP predicts what the inventory levels for different parts will be over time, and predicts when the inventory for a certain part would deplete to zero. SAP then issues a production order for that part, and schedules its start date such that production will be completed, and the inventory will be replenished, before the inventory reaches zero (e.g. if a part has a 14 day lead time, the order will be scheduled to begin 15 days before the inventory is depleted). Sometime, an order will be triggered not when the expected inventory is suppose to reach zero, but at some reorder point above zero, to keep a safety stock for assembly in case faults occur in production.

With this information, the Machining Center’s Planning Department (known as “Planning”) now has a schedule of which production orders should begin on each day for the next three months. However, in creating this schedule, SAP does not consider any limitations of the actual production system. The system does not schedule around the fact that machines have a limited production capacity, and that two orders cannot be produced on the same machine at the same time. The SAP scheduling system only considers the demands and production rates for each individual part, and constructs schedules based on those figures.

Thus, it is up to the Planning department to translate the schedule from SAP into a practical schedule that can be used by the production departments. About two weeks before an order is scheduled to begin, Planning will “cut” the order, which means that they will generate a process sheet for it. The process sheet indicates what SKU needs to be made, the number of parts to be made, the start date of the job, the expected end date of the job (to stay on schedule), and all the operations the part must undergo before being completed. Planning then creates a two-week schedule in Excel that contains all of the cut orders, broken down by each machine. Each part has a designated machine that it is produced on every time it is made; this assignment
is based on matching the geometry of the part to the capabilities of the machine. Planning organizes the schedule for each machine strictly by the “due date” for each order, which is the date that the order is supposed to be delivered to the next department for its subsequent process. If an order suddenly appears in the SAP system for any reason and is scheduled to begin in the next two weeks, Planning will cut that order as well, and place it onto the Excel spreadsheet based on its due date, even placing it ahead of all the other orders already on the schedule if need be. When the order is placed onto this two-week schedule, the process sheet will be delivered to the machine where the operation will take place, and will remain there until the order is scheduled to begin.

2.3.3 Production

When an order is set to begin, the machinist will review the process sheet for that particular order, go to his or her assigned machine, and set up the machine for the specific job. Depending on what the part is to be produced and what was already set up on the machine, a setup can take anywhere between 10 minutes and 12 hours.

The machinist will then pull the raw material for the job from the raw materials inventory, positioned in a central location in the turning department. The raw material will be loaded into the machine, and the machine will begin its operation. A job or machine may or may not need constant supervision by the machinist; this is usually dependent on the raw material that is needed to make the part. Some raw material comes in long cylindrical bars, ranging from 4 feet to 12 feet long and 1 to 2 inches in diameter, while some come in pre-cut blanks. Production from long cylindrical bar stock will utilize an automatic feeder that is built onto the lathe, so that once a part is complete, the machine will output the finished part and then autonomously feed the bar into the chuck and continue. On the other hand, pre-cut blanks have to be individually loaded into the chuck and then unloaded after the operation is finished. Typically, if a machine or job does not require constant supervision, the machinist will use this time to set up or perform basic maintenance on other machines in their department.

The machinists are also tasked with inspecting the finished parts coming off the machine, and determining which parts are good and which do not fall within specifications (known as “nonconforming”, “scrap”, or “red-tagged” parts). Depending on the part, this inspection step is done either after each part comes out of the machine, or when the entire batch is complete. The
machinist then records the number of good and scrap parts on the process sheet. Good parts are placed into a production tray, which is the main receptacle used for transporting parts around the shop. The tray is designed to hold either 24 or 30 parts, depending on the lot size of the production run (e.g., an order that calls for a batch size of 92 will use four trays of 24 parts, while an order with a lot size of 120 will use four trays of 30). Nonconforming parts are removed from the order and placed in a designated area in the turning utility station. Once a day, the supervisor will inspect the scrap parts, and determine whether or not they are acceptable, need to be reworked, or must be abandoned.

After an order is complete, the machinist will then deliver all of the trays for that order and the process sheet to the turning utility area. The process sheet will specify whether or not the parts need to undergo a secondary operation at the utility area, such as a deburring process or a simple clean. If it does, the turning utility operator will perform these necessary operations. After the secondary operations are completed, or if an order does not require any secondary operations, the utility operator will log the completed processes in the SAP system, along with the number of good and scrap parts produced. The utility operator will then deliver the parts and process sheet from the turning utility area to the milling department. The milling department has multiple incoming goods buffers, each located next to a particular milling machine. As in turning, parts have a designated machine in milling they are worked on every time an order of that part is produced, and are placed in the incoming goods buffer next to that machine.

The planning department also constructs a temporary production scheduling for the milling department, using the dates that orders are expected to arrive at the milling incoming buffer. However, this schedule is heavily dependent on when orders are actually completed by the turning department, which is subject to variability due to machine failures, scheduled machine downtime, process failures resulting in rework, lack of staff, and the day-to-day adjustments to the turning schedule. Therefore, the schedule for milling is constantly in a state of flux, and is monitored and adjusted regularly by the milling section leader and department supervisor much more than the turning schedule.

In the milling department, the section leader selects when the order should proceed through based on the milling schedule, and assigns an operator to the order. From here, the process is very similar to the process in turning: the machine is setup to produce the next part (which can take between 10 minutes and 4 hours depending on the nature of the setup), and the
A machinist begins production. Since the part has already been cut into its approximate geometry in the turning step, the machinist needs to constantly monitor the machine and load and unload parts from the vice as the order proceeds. After a part comes off the machine, the machinist will inspect the part to see if it falls within specifications, place the good parts back in the production tray, and send the non-conforming parts to the scrap bin in the milling utility station. After an order is complete, the machinist will record the total number of good and scrap parts on the process sheet, and transfer the completed parts to the milling utility station. From there, the milling utility operator will complete any secondary operations on the parts if necessary, and log the number of good and scrap parts produced into the SAP system. The milling utility operator will then bring the parts and the process sheet to the next department listed on the sheet to continue its production.

2.4 Issues with the Waters Manufacturing System

In order to establish system changes to improve the productivity of the Machining Center, we must also investigate what inefficiencies and unavailing practices exist on the shop floor, understand how they materialize as waste, and diagnose their ultimate causes. Once we recognize what the underlying sources of inefficiencies are, we can develop modifications to the manufacturing system which eliminate those issues.

This analysis began with an examination of existing production data logged in the Scheduling and Planning (SAP) system; however, it quickly became apparent that the information available on SAP was very limited. The only reliable information available through SAP were the operations necessary to complete a particular part and how many good and scrap parts were produced in an operation. These production totals can be collected over time to estimate the demand for a particular part. Nevertheless, there was no other reliable information available to investigate and model the current production system and its problems, such as actual setup times that occur on the shop floor, part processing time, machine utilization, inventory of work in progress, etc.

After moving on from the SAP system, factory analysis heavily involved conducting interviews with key stakeholders at all levels of the company. This included the head of manufacturing operations, demand planning, finance, machining department section leaders and supervisors, and machinists on the shop floor. The bulk of the characterization of factory issues
came about due to the results of these interviews, and confirmed through personal investigation of the Machining Center’s operation.

2.4.1 Overview of Problems Discovered

2.4.1.1 Low Machine Utilization

Waters finds itself unable to meet its production demands with its current capacity, due to the fact that many of its machines suffer from utilization (percentage of time spent making parts; see Section 3.1) considerably lower than the company target of 80%. After using a power monitor to track the spindle uptime (which defines machine uptime) of multiple machines in Waters’ Machining Center, the data showed that the problem of low utilization was particularly glaring in the Turning and Milling department, where utilization rates for machines producing high-volume parts were typically below 70% and below 50%, respectively. Figure 2 provides an example from July, 2014:

![Utilization of four machines in the Milling department on July 10th, 2014, for both the day and night shifts. Aggregate utilization of these machines during that week was 52%, much lower than Water's goal of 80%.

2.4.1.2 Low On-Time Delivery of Parts to Assembly

Interviews with machine shop and assembly supervisors revealed a consensus that on-time delivery rates for essentially all parts is much lower than desired between the Machining Center (responsible for all processes from raw material to finished good) and Assembly. The Machining Center seeks a service level of 95% to Assembly, as measured by on-time delivery to
the finished-goods inventory holding location in the Assembly area. However, through our investigation into the on-time delivery data available from the Planning department, those percentages, when broken down by department, are consistently under 50% and, in some weeks, have dipped below 20%. Again, this problem in especially glaring in the main upstream processes of turning and milling, as Table 3 demonstrates for the second quarter of 2014. This slow upstream delivery translates to overall low on-time delivery to Assembly.

<table>
<thead>
<tr>
<th>On-Time Delivery</th>
<th>Turning</th>
<th>Milling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 14</td>
<td>35%</td>
<td>17%</td>
</tr>
<tr>
<td>Week 15</td>
<td>14%</td>
<td>5%</td>
</tr>
<tr>
<td>Week 16</td>
<td>17%</td>
<td>15%</td>
</tr>
<tr>
<td>Week 17</td>
<td>8%</td>
<td>18%</td>
</tr>
<tr>
<td>Week 18</td>
<td>16%</td>
<td>17%</td>
</tr>
<tr>
<td>Week 19</td>
<td>17%</td>
<td>12%</td>
</tr>
<tr>
<td>Week 20</td>
<td>10%</td>
<td>13%</td>
</tr>
<tr>
<td>Week 21</td>
<td>26%</td>
<td>5%</td>
</tr>
<tr>
<td>Week 22</td>
<td>22%</td>
<td>15%</td>
</tr>
<tr>
<td>Week 23</td>
<td>30%</td>
<td>22%</td>
</tr>
<tr>
<td>Week 24</td>
<td>34%</td>
<td>33%</td>
</tr>
<tr>
<td>Week 25</td>
<td>32%</td>
<td>24%</td>
</tr>
</tbody>
</table>

Table 3: Percentage of orders delivered to the next department on-time (OTD), and the average number of days late for an order from the Turning and Milling departments during the second quarter of 2014. OTD was consistently below 40%, and the average days late for every order coming from those departments ranged between four and fourteen days.

2.4.1.3 Significant Day-to-Day Adjustments to Production Schedule

Through interviews with the department supervisor and machinists, and by regularly monitoring of two-week production schedules generated by Planning, it was revealed that these schedules change significantly on a daily basis. Any error in the SAP data for the setup or cycle time for a part causes discrepancies between the SAP schedule and actual production. These discrepancies in production data lead to issues with inventory and part availability, such as an unexpected backorder for a part that requires immediate restocking at Assembly, or an order made by an outside vendor that needs prompt attention. When these issues occur, SAP will
schedule the start of the order to be at some time in the past, such that if it had actually started at that time, its production would end before the due date. The Excel schedule produced by the Planning department is always organized strictly by the due dates of the orders to the next department, so orders that suddenly appear in SAP will jump to the front of the schedule, and production of previously scheduled parts gets moved back. In some situations we observed, this scheduling procedure will even cause two orders of the exact same part, that is currently scheduled to run back-to-back, to now split up to accommodate this sudden order. This forces machinists to run three setups for these three orders, when they could have only run two. The same problems occur when a previously completed job needs rework due to high amounts of nonconforming parts, or in the case of machine failures or other unforeseen delays.

2.4.2 Diagnoses of Problem Causes

The factory analysis showed the problems detailed above had several systematic causes leading to the various inefficiencies.

2.4.2.1 Frequent Occurrence of Redundant Setups

The production floor staff indicated that the company policy and practice is to take the quarterly demand forecast for a high volume product, and divide by the number of weeks in the quarter, in order to prepare for changes in demand well in advance. In addition, interviews with various manufacturing leaders at Waters revealed an initiative in recent years to improve system “flow” and reduce inventory by aiming to produce high-volume parts in small, constant amounts every week, rather than in larger batches. This resulted in the lot sizes of certain high-volume parts being reduced. The effect of these practices, however, was to increase the total number of machine, tool and fixture setups necessary to run the same number of parts, since a part will require a new setup every time it gets produced. Since setups occurred more frequently, the amount of time a machine was being set up and not producing parts increased dramatically, which decreased the production and utilization of the machines. While the reasoning behind this endeavor from upper management was to keep inventory and inventory holding costs low, inquiries made at various levels of the company yielded no systematic method of tracking inventory holding cost, and given that high-volume parts produced in Milford generally have low
obsolescence risk (liquid chromatography machines still sell parts 20 to 25 years after first release), even conservative estimates of holding costs would be much lower than the costs of performing extra setups.

2.4.2.2 Delays in Turning and Milling Cascade to System-Wide Delays

As indicated in Section 2.2, Turning and Milling are the first two upstream processes for a majority of Waters’ critical SKUs. These two departments also feature the longest setup times for parts, have the lowest machine utilization, require more staff than any other production department in the Machining Center, and in general tend to be the busiest departments in the shop, which causes them to be bottleneck manufacturing process for all parts. If Turning and Milling fall behind on their production schedule, these delays continue throughout the entire manufacturing system, leading to reduced on-time delivery to departments further downstream and to Assembly. These setbacks are difficult for other departments to recuperate from, thus, issues in Turning and Milling end up generating inefficiencies in other departments; in fact, delays in turning and milling may even exacerbate downstream delays due to the “bullwhip effect” [3].

2.4.2.3 Poor Production Scheduling

The SAP scheduling system does not examine the entire production system as a whole, but only considers the demands and production rates for each distinct part, and constructs individual schedules based on those figures. Since the schedule output from SAP does not take a comprehensive look at the entire production system, a whole host of problems arise when trying to translate the information from SAP into a useful schedule for production. Since SAP assumes that the production system has infinite capacity, Planning has to assign the orders scheduled in SAP to particular machines in the Center, so that the production line can actually produce the orders when SAP plans them. Since the Planning department blindly follows the due date of the orders when producing this schedule, they miss opportunities to set production runs back-to-back with reduced setups. Parts with similar geometries, or parts that use the same tooling or fixturing, could be scheduled sequentially in order to reduce the setup time needed to begin
production of the second part. Nevertheless, even when orders that can undergo reduced setups are already on the two-week Excel schedule, they are never scheduled consecutively to reduce setup times and increase machine utilization (unless it happens out of pure coincidence). Further, due to the small lot sizes for high-volume products, the production schedules call for the same part to be run at least once a week, requiring a lengthy setup each time. Instead, the lot size of these orders can be increased, or orders of the same part can be scheduled to run back-to-back, to eliminate the additional setups. These issues directly hurt the utilization of the machines.

Other issues with production scheduling relate to the safety stock for available parts. Due to the initiative by Waters executives to remove inventory and work-in-progress on the production floor, safety stock numbers in SAP used to calculate inventory thresholds for older SKUs are often ignored by the planning department, which can cause an abundance of obsolete inventory for old parts, or worse, a backlog of orders for high-volume parts. This, along with the issues discussed in Section 2.4.1, leads directly to the need to alter the production schedule daily to manage crises, as well as low on-time delivery to the Assembly area.

2.4.2.4 Inefficient Data Collection and Feedback to Planning

Another issue noticed during the factory study is the lack of efficient data collection and use in the Machining Center and elsewhere at Waters. The project team learned through staff interviews that many legacy data collection practices found in the finance and accounting department cause issues in production planning. The finance and accounting staff indicated that company policy and practice is to perform a cost roll once every five years. A cost roll is a system that assigns a production cost to all parts produced in the Machining Center, based on the total production time needed to complete a specific part, and the cost per hour of operation of a particular department. The data collected during the cost roll are average process cycle time, average machine tool set up time, and average machine tool change over time. This data are collected manually for a limited set of parts and processes using handheld stop watches as timing mechanisms, and having department section leaders observing, timing, and recording the duration of these various processes. This information is then used to calculate the expected number of machinists needed and the capital required to produce parts to demand. This cost roll
methodology not only records data too infrequently, but makes broad comparative assumptions. For example, the finance department assumes that the process time and setup times are the same for all parts that have similar geometries. Further, whenever a new SKU is introduced to the Machining Center, instead of running a cost roll for that individual SKU, Waters will determine which part is most similar to the newly introduced SKU, and simply copy the information on setup and process cycle time from that similar part to the new one.

Outside of the cost roll, the only production data collection that currently exists on the shop floor is a Filemaker data system that is run by the turning and milling supervisors. In this system, the machinist who is responsible for a production order must also manually record the part SKU number, the order number, the setup time, cycle time, good and scrap parts produced, and any other comments related to the production run into a Filemaker data sheet. This not only leads to inefficiencies in the workers themselves, since they have to spend time collecting data rather than machining, but the information is often incorrect. Since the machinists consider this data collection as a secondary responsibility, they do not keep track of how long a setup takes, and instead only input an estimate to the closest whole hour. Further, the cycle time input into Filemaker is purely an output of the G-code program running on the mill or lathe, which defines how long it takes for tools to complete their tool path for each operation. This cycle time does not account for all the time in which the machine is idle during a cycle, either spent loading new material into the chuck or due to a machinist not being present at the machine continue the process. Hence, cycle times are grossly underestimated.

Another issue with this Filemaker data is that even though the machinists collect it, the accumulated data is only accessible to department supervisors and section leaders, and the software to open the data files is only available on a few computers in the shop floor. Moreover, none of this information is fed back to the Planning department to adjust their scheduling procedure, nor is it applied to any area of the shop or used at any time outside of the debrief meetings at the end of each shift. As a result, it is unclear to some people on the shop floor what the useful purpose of the data collected is.
2.5 Problem Statement

Again, the objective for this project is to develop system improvements that could be applied to multiple departments, areas, machines, and parts in the Machining Center to expand productivity, and then demonstrate their value through implementation on a portion of the Machining Center and a limited number of parts. After we identified the issues of low machine utilization and low on-time delivery of parts to the Assembly area, we diagnosed the causes of those issues as frequent, redundant setups that increase machine downtime and decrease utilization, and a lack of a data collection system to accurately model the production system and provide feedback to the Planning department. Therefore, our system modifications will focus on eliminating these two sources of inefficiencies in the Machining Center. The work by Perez [2] focuses on the benefits of automated data collection to improve production system performance, while Chandar [1] and this work considers changes to the manufacturing system to improve machine utilization through reducing the number of setups needed to produce the same number of parts. More specifically, Chandar [1] stresses the benefits of increased lot sizes to reduce setups, while this work focuses on improved scheduling procedures to take advantage of sequentially-dependent setups.

For shop location, the NC Turning and NC Milling areas were selected as our areas of implementation. The reasoning behind this decision was twofold. For one, the turning and milling areas were subjected to the same problems seen throughout the shop: high setup times resulting in long lead times for orders, low machine utilization, and low on-time delivery of parts to customers or to the assembly area. In general, the CNC turning and CNC milling areas were regarded as the bottlenecks of most manufacturing processes within the shop. Furthermore, turning and milling comprise initial operations for the vast majority of SKUs Waters produces in its Milford plant, so these delays in the upstream processes propagate down the entire line to other processes and areas of the facility.
Theoretical Review of System Improvements

The objective of the system modifications proposed here is to increase the utilization of the machines by reducing the total setup time needed to produce the same amount of parts. We can define the utilization $u$ of a workstation as the fraction of time that the station is operating and producing goods, as opposed to being setup, repaired, or idle due to lack of input materials (which can be anything that the machine needs to operate, including consumables of the machine and workers). For the purposes of the project, the utilization of a machine will be defined as the percentage of time that the machine is drawing power and its spindle is running, thereby directly doing work to produce parts, compared to the total time that machine is staffed (which in our case is 24 hours a day, 6 days a week, or 144 hours per week).

We earlier identified the Turning and Milling areas as the bottleneck processes for most parts in the Machining Center, so these modifications will focus on reducing the total setup time in those areas, thereby improving the performance of the Machining Center system-wide. These adjustments are designed such that they can be applied to all machines within the Turning and Milling departments, and be applied to other similar multiple-part, multiple-flexible-machine systems.

Improving the manufacturing system through reduced setups involves three distinct techniques: resizing the production lots for high-volume parts, establishing of a supermarket for those parts, and reordering of the production schedule to make those parts in a more efficient manner.

3.1 Optimized Lot Sizing

After discussing with multiple people in the company about how production lot sizes were determined, it became apparent that there was no methodology in place to determine an optimal size. The SAP system does have a function which generates the Economic Order Quantity (EOQ) for a part, which is an initial figure used by the company to determine final lot size. The EOQ model originally developed by Harris [4], computes lot size to minimize cost for
one part type, taking into account the demand for the part, the inventory holding cost, and the
ordering cost (analogous to setup cost for production):

\[ Q^* = \sqrt{\frac{2DK}{h}} \]  

(3.1),

where \( k \) is the ordering cost (setup cost), \( h \) is inventory holding cost and \( D \) is demand rate.

While the Economic Order Quantity model is a popular methodology for determining production
sizes or order quantities, it assumes that parts are ordered or made in a vacuum, independent of
the production or procurement of any other part, and that those orders can be placed at any time.
In our case, if we were to apply an EOQ-model for all the parts that have to go through one
machine, the resulting lot sizes would be too large for our system, as it does not include the
constraints of our multiple-part, one-machine system. This would causes issues with backorders,
as the production of certain parts would have to wait for the machine to complete other
production runs before being set up. Employees in the Planning department quickly realized this
problem, and override the EOQ-generated lot size for most items. However, there is no
methodology to these changes; only when a lot size is noted to be too big or too small are lot
sizes changed based on intuition.

The hypothesis was that production lots for most high volume parts could be increased, in
order to reduce the total number of setups needed to make the same amount of parts over a
particular time period; that time recovered from reducing number of setups could be spent
making more parts, directly increasing the productivity of the machine. Due to practice of
artificially increasing the flow of parts through the factory as discussed in Section 2.4.2, high
volume parts are produced in lot sizes that are less than one-quarter of their monthly demand,
causing them to be set up very frequently and redundantly. By increasing the lot size, the
number of setups over time can be reduced, while still producing to demand. It is understood
that this proposal is counter to the “lean” dogma put into place by Waters Corporation, which
believes that smaller production lot sizes allow for higher throughout, a smaller inventory of
parts at the end of the line, and less work-in-progress throughout the line. While increasing lot
sizes will result in higher inventory levels and increased work-in-progress, in a multiple-part-
type manufacturing system with a large setup time to production time ratio (like in Waters’ case),
the savings from limiting the number of setups done are much greater than the costs of holding
additional inventory or work-in-progress, and actually lead to a higher throughput.
The challenge is balancing the lot sizes between all the parts that have to go through a particular machine, so that all parts can be made according to their demand and without falling into backorder. Very little literature exists on lot sizing and scheduling on machines with multiple, dependent part types, but the most recent attempt to model such a system came from Pinedo [5]. The Pinedo model, which he describes as the “economic lot scheduling problem,” is more comprehensive than the EOQ model in describing machines with multiple part types and with setups, as it factors in the setup time, cycle time, demand rate and inventory holding cost for all parts going through a machine. The equation for the optimal lot size \( Q \) of part \( i \) in an \( n \)-part, single-machine case is re-written below

\[
Q_i = D_i \sqrt{\left( \sum_{j=1}^{n} h_j \left( D_j - G_j \right) \right)^{-1} \sum_{j=1}^{n} c_j}
\]

(3.2)

where \( D \) is the demand rate, \( h \) is the inventory holding cost, \( c \) is the setup cost, and \( G \) is the production rate of each part. By gathering this information for all parts that go through the machines under investigation in the milling and turning areas, the optimal lot sizes for the parts in question can be determined. If this methodology results in larger production batches for high-volume parts, then these new lot sizes can be applied on the shop floor to reduce setups.

It should be noted that the Pinedo model is not perfect either; it sizes lots assuming all parts have independent setups. However, it takes stock of the dependence of all parts with a limited capacity on a machine and sizes lots to reduce costs while preventing inventory backlogs, and is the best formulaic approach available. Part of the results of this thesis will judge the effectiveness of the optimal lot sizes on reducing setup and holding costs while preventing backorders using the new production lot sizes determined by the Equation 3.2.

3.2 Part Supermarket

One potential issue with the lot sizing methodology is that this investigation only looks at the production of parts in the turning and milling areas. The optimized lot sizes developed for these departments could be sub-optimal for processes down the line. Large batches could wreak havoc on a downstream process with large cycle times and shorter setups, causing orders of other parts that need to be processed on the same machine to fall into backorder. As the machining
area is setup as a job shop with an assortment of different processes that pull from and feed each other in a variety of combinations, it would be difficult to generate optimized lot sizes for every part and every process. If the optimized lot sizes for turning and milling are to be implemented, there must be a way to decouple the production in milling and turning from the rest of the manufacturing system, so that production in these areas can be optimized without disrupting the production in downstream processes.

With these requirements in mind, it appears that the optimal system would be one in which parts are produced upstream in Milling and Turning with a make-to-stock policy or push system, and while parts downstream are produced with a make-to-order or pull system. Push production systems are characterized by having production orders arrive at the beginning of a production line. Batch sizes and order quantities are the result of a centralized decision process, usually coming from a master production scheduler (like SAP). Decisions on lot sizes and production levels are made irrespective of current demand, and are instead based upon a demand forecast [6]. Moreover, push systems are more in-line with the practice of producing parts in larger lot sizes with fewer changes in setups, which can be more productive for a manufacturing system (provided they do not result in explosions in WIP) [7].

Push systems have several drawbacks, however. They are not very responsive to changing customer needs, which means their effectiveness is heavily dependent on the quality of the demand forecast which precipitates production. This in turn means that parts in a push system are highly susceptible to the “bullwhip effect,” in which the variability of orders increases more and more upstream in a supply chain in response to small changes in customer demands downstream. Due to this, supply chains can be susceptible to either huge surpluses of inventory or huge backlogs for popular products [3]. This bullwhip effect can be reduced however, by shorting the length of the push-based supply chain.

In contrast to push systems, pull systems have production orders arrive at the end of a line. These orders are then “pulled” out of the system from an inventory or raw material location. The key benefit of a pull system over a pure push system is that production is triggered based on actual demand, rather than anticipated demand. Upstream processes then compensate for parts being pulled out of the WIP or raw material stock by sending orders to vendors or by triggering
production to refill the inventory. Moreover, pull systems establish a WIP cap at the various buffers in a production line, preventing the production of WIP that fails to add to throughput [8].

A way to incorporate a push-pull hybrid system and manage the tradeoff between the need to limit inventory and the need to produce high-volume parts at a high rate is through the use of supermarkets. Supermarkets are essentially decoupling buffers, storage areas placed at locations within a manufacturing system to serve as intermediate sources of inventory [9]. In this situation, production in turning and milling will operate in order to maintain the inventory of the part supermarket, and downstream processes can pull from that inventory as orders come in. With a part supermarket in place, production in turning and milling can occur at their optimized lot sizes, while not affecting the production of the downstream processes; the next step of the production process can be completed at any lot size by pulling any number of parts from the inventory, based on the order quantity. Downstream production can then operate, decoupled from upstream inventory, by simply taking its inputs from the supermarket, rather than waiting for upstream production to finish [9]. An example of a supermarket placement within the Waters manufacturing system is shown in Figure 4 below.

![Diagram](image)

Figure 4: The part supermarket, a buffer used to decouple upstream production from downstream processes.
A supermarket system can provide improvements over a pure push system because it can achieve lower WIP upstream than a push system would typically allow. It also allows for an increase in on-time delivery in cases where upstream processes can bottleneck downstream production, as turning and milling do at Waters, and can allow different production batch sizes to exist upstream and downstream due to its decoupling effect.

In this new line, whenever an order is placed in the system, the “production” of the part will begin by pulling the desired number of parts from the supermarket after milling, instead of starting at the beginning of turning. Since these parts had previously been completed in turning and milling before being held at the supermarket, they are immediately ready for the next step of the process. The inventory levels of the parts at the supermarket will be monitored, and when SAP calls for a new order based on predicted demand, production will be triggered at the beginning of turning which will refill the inventory at the supermarket. This production will occur at the optimized lot size determined in Section 4.1 for each part.

An additional benefit to the part supermarket will be to decrease the lead time of parts from order to arrival at assembly, which will increase their on-time delivery. Whenever an order is placed for a part, the order will signal the movement of parts from the supermarket to the next step in the process, instead of signaling the beginning of production at turning. This will remove the turning and milling processes from the lead time of the part from the time it is ordered to when the particular parts are completed. Currently, turning and milling comprise about 20% to 50% of the total lead time for a part’s production in the machining center, so removing those processes will significantly decrease lead time. If the quoted due date remains the same, then the on-time delivery percentage is expected to increase dramatically.

Part of the work involved in the implementation of the supermarket will involve tracking the effect of the supermarket on the on-time delivery percentages for downstream processes and into the assembly area. While the focus of this work is on the productivity and utilization improvements possible on a flexible machine due to optimized lot sizing and production scheduling, the results of the investigation into the improved lead time will be detailed briefly in this work. Chandar [1] will go into further detail on the effects of the part supermarket on on-time delivery improvements in this industrial setting.
Sections 3.3 and 3.4 present unique scheduling procedures that are designed to take advantage of the reduced setup times that ensue when producing parts within the same family back-to-back. Part families are parts that have like geometries and require similar tooling, fixturing, or setups on a machine, such that the time it takes to transition production from one part to another is shorter than the transition between parts in different families. Simply put, a part is considered to be in the same family as another part if the setup time and costs of going from one part to another, and vice versa, is much less than the setup between any other parts in the system. Thus, in order to improve utilization in a multi-part system, it is optimal to produce parts in a family sequentially, in order to reduce total setup times.

The setup matrix is a tool that enables smarter scheduling of production orders to minimize the total setup time and maximize utilization. A setup matrix is useful in situations where setups are sequentially dependent, that is, the setup time for a certain part depends on not only what the part is, but also on what part was previously set up on the machine. Even in situations where setup dependencies are not considered in planning and scheduling, they can still occur on the factory floor. At Waters, even though they have standards by which all setup times are determined, and their scheduling system considers all setups to be sequentially independent, this is not the case. For example, if there is a part family that consists of two parts, they typically use the same tools and similar settings on the machine. Going from the production run of one part to the other usually just takes a change in the G-code program and some small tooling offset adjustments. Therefore, when these parts are produced back to back on a machine, the setup time for the second part will be less than what the setup time would be if a different part (a part not in the same family) were made before it. The setup matrix allows planners to create a schedule that takes advantage of these reduced setup times. The setup matrix methodology consists of two parts: formation of the setup matrix, and application of the setup matrix to scheduling.
3.3.1 Format of the General Setup Matrix

In a situation where there are \( n \) parts under consideration (typically the number of different parts produced on a machine, or the number of orders in place on a production schedule), a setup matrix is a square \( n \)-by-\( n \) matrix such that the value in row \( i \) and column \( j \) is the setup time for part \( i \) after being run directly after part \( j \). This setup time is denoted \( s_{ij} \). The general layout for a setup matrix is shown below.

\[
\begin{bmatrix}
1 & 2 & \cdots & \cdots & n-1 & n \\
1 & s_{11} & s_{12} & \cdots & s_{1(n-1)} & s_{1n} \\
2 & s_{21} & s_{22} & \cdots & \cdots & s_{2n} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
n & s_{(n-1)1} & \cdots & \cdots & s_{(n-1)(n-1)} & s_{(n-1)n} \\
n-1 & s_{n1} & s_{n2} & \cdots & s_{n(n-1)} & s_{nn} \\
\end{bmatrix}
\]  

(3.3)

3.3.2 Special Cases of the Setup Matrix

3.3.2.1 Two Orders of the Same Part

If a machine is set up to run a particular part, it is typically assumed that the machine will be able to produce that part indefinitely without any further setups needed. In a case where there is no delay in running two orders of the same part back to back, or if the preparation or delay before starting the second order is much less than any other setup time, the setup time when \( i=j \), \( s_{ii} \), can be approximated as zero. Thus, the setup matrix would contain all zeroes across its main diagonal. A setup matrix designed with this common assumption is shown below.

\[
\begin{bmatrix}
1 & 2 & \cdots & \cdots & n-1 & n \\
1 & 0 & s_{12} & \cdots & s_{1(n-1)} & s_{1n} \\
2 & s_{21} & 0 & \cdots & \cdots & s_{2n} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
n & s_{(n-1)1} & \cdots & \cdots & 0 & s_{(n-1)n} \\
n-1 & s_{n1} & s_{n2} & \cdots & s_{n(n-1)} & 0 \\
\end{bmatrix}
\]  

(3.4)
3.3.2.2 Parts within Families

The most basic method of representing a part family in a setup matrix is to assume that setups between part families take some standard amount of time, while the setup time between any two parts in a family is some reduced figure. In a case where there are n parts produced on one machine, and there is one part family consisting of 2 parts, Part 1 and Part 2, the setup matrix would be

\[
\begin{bmatrix}
S_{\text{reduced}} & S_{\text{reduced}} & S_{\text{std}} & S_{\text{std}} & \cdots & S_{\text{std}} \\
S_{\text{std}} & S_{\text{reduced}} & S_{\text{std}} & \cdots & \cdots & S_{\text{std}} \\
S_{\text{std}} & S_{\text{std}} & S_{\text{std}} & \cdots & \cdots & \cdots & \cdots & S_{\text{std}} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
S_{\text{std}} & S_{\text{std}} & S_{\text{std}} & \cdots & S_{\text{std}} & S_{\text{std}}
\end{bmatrix}
\]

(3.5)

where \(S_{\text{std}}\) is the standard setup time, and \(S_{\text{reduced}}\) is the part family setup time.

This approach for representing part families can be extended to most cases where sequential dependent setups exist. For example, consider a case where a machine makes eight different parts, Part 1 through Part 8, and consists of two part families. The first family comprises of Part 1 through Part 4, and the second family contains Part 5 and Part 6 (Part 7 and Part 8 are not in a family with any other part). Furthermore, Part 1 and Part 2 are in their own subfamily within the first part family, such that if Part 2 is made after Part 1, or vice versa, the setup time is even further reduced, compared to going from Part 1 to Part 3, or Part 1 to Part 4, or Part 2 to Part 3, etc. This complex scenario can be modeled by the simple setup matrix below:

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
1 & a & a & b & b & s & s & s & s \\
2 & a & a & b & b & s & s & s & s \\
3 & b & b & b & b & s & s & s & s \\
4 & b & b & b & b & s & s & s & s \\
5 & s & s & s & s & c & c & s & s \\
6 & s & s & s & s & c & c & s & s \\
7 & s & s & s & s & s & s & s & s \\
8 & s & s & s & s & s & s & s & s
\end{array}
\]

(3.6)
where the base setup time (the setup time when transitioning between families) is \( s \), the setup
time within the first family is \( b \), the setup time within the second family is \( c \), and the setup time
within the subfamily is \( a \).

### 3.3.2.3 Boundary Condition: An Un-setup Machine

While the setup matrix indicates the setup times between different production runs, one
must also consider the setup time for when the machine is in its base or neutral state, with no part
currently set up on the machine. This can occur whenever a new machine is added to the factory,
a production schedule must be created from scratch, or repairs or maintenance occur which force
the teardown of the machine. To account for this scenario, an additional row and column can be
added to the matrix to represent the transition from a neutral state to a setup state and vice versa.
Thus, the setup matrix becomes an \((n+1)\times(n+1)\) matrix, where the neutral state is treated like
another part that is fabricated on the machine. An example of a setup matrix that includes the
base state of the machine is shown below

\[
\begin{bmatrix}
1 & 2 & \cdots & n-1 & n & \text{base} \\
1 & s_{11} & s_{12} & \cdots & s_{1(n-1)} & s_{1n} & s_{1\text{base}} \\
2 & s_{21} & s_{22} & \cdots & \cdots & s_{2n} & s_{2\text{base}} \\
\vdots & \vdots & \ddots & \ddots & \cdots & \vdots & \vdots \\
n-1 & s_{(n-1)1} & \vdots & \vdots & s_{(n-1)(n-1)} & s_{(n-1)n} & s_{(n-1)\text{base}} \\
n & s_{n1} & s_{n2} & \cdots & s_{n(n-1)} & s_{nn} & s_{n\text{base}} \\
\text{base} & s_{\text{base}1} & s_{\text{base}2} & \cdots & s_{\text{base}(n-1)} & s_{\text{base}n} & s_{\text{basebase}}
\end{bmatrix}
\]

(3.7)

where the subscript "base" denotes the base or neutral state of the machine, with no part set up
on it.

### 3.3.2.4 Other Cases

While common cases for the application of a setup matrix are outlined above, the matrix
applies to all scenarios in which sequentially dependent setups occur. Even in a situation where
there are \( n \) parts that run on a machine, with no part families, with setup times that are not
commutative \((s_{ij} \neq s_{ji})\), and have unique setup times for each combination \((s_{11} \neq s_{12} \neq s_{13} \ldots \neq s_{nn})\), a
setup matrix can still be formed consisting of $n^2$ unique values, and can be used fruitfully in production scheduling and planning in order to maximize productivity.

### 3.3.3 Application of the Matrix to Scheduling

A completed setup matrix can then be used to develop a schedule, based on a set of production orders, to reduce or minimize total setup time. If $n$ production orders are planned on a certain machine over a certain time period, create the $n$-by-$n$ setup matrix for those orders. Then, determine the initial state of the machine. The machine must already have a designated state from which the schedule commences: either a part currently running on the machine, or the base state of the machine with no part set up. This state is designated as $k$. From here, search the $k$ row in the setup matrix for the smallest setup time that is not on the main diagonal of the matrix. This value indicates the quickest setup possible when transitioning from the current state of the machine to another (which is likely to be another part in the same family as the current state). Therefore, that part should be scheduled next. Then, reassign $k$ to the order placed on the schedule, and continue this process throughout the entire list of production orders. The final result will be a production schedule that takes advantage of part family setups to lower the overall setup time and increase productivity.

One issue that has not been addressed is how to determine the values for the setup times that are input into the matrix. In a case where frequent setups occur on a machine, planners should have information on at least a base setup time for each part (this information is needed in order to certify that the machine will be able to produce parts according to their demand rate; if the total production rate of a machine for any part is lower than the part’s demand rate, backorders for a part will build up indefinitely and the system becomes unsustainable). From there, if sequentially-dependent setups exist, sets of setup protocols may subsist that apply to a particular group of parts. These protocols can reveal similarities in tooling, fixturing, and machine configuration, and therefore expose part families. Further, workers closest to the production line will typically have intuition in terms of which parts fall into part families or have reduced setup times when progressing between certain parts. While this methodology does not result in exact values for the setup times, a recognition that a certain setup time is less than the base case is enough to improve scheduling using the setup matrix. Furthermore, determining the
relative values of all the items in the matrix, without even knowing the actual values, is enough to optimize scheduling.

In order to achieve more exact values to populate the setup matrix, time studies can be conducted to document the setup times for different production runs. Furthermore, automated data collection systems can be used to gather data on setup times across a large range of parts quickly. Once setup times are determined, these automated systems can then populate and update the values of the setup matrix automatically. For more information on attaining production setup times using automated data collection systems in a manufacturing environment, see Perez [2]

3.4 Production Scheduling: QPR Inventory Control Policy

One issue that quickly arises in the setup matrix methodology is the difficulty in applying the process to a large number of orders or parts. If you wanted to use the setup matrix methodology to create an optimized schedule for a complete set of parts which run on a machine (instead of just the ones in line on a production schedule), every possible combination of orders must be cross-checked using the setup matrix in order to determine the optimal sequence to minimize setup time. This is an exact representation of the Traveling Salesmen Problem [10]. Thus, the number of combinations that must be tested is on the order of \( n! \), where \( n \) is the number of parts in the setup matrix (equivalent to the number of orders on the schedule, or the number of parts that are produced on a machine). As \( n \) increases, the setup matrix methodology becomes unwieldy, even when the scheduling procedure is managed by modern computers. Hence, a more generic scheduling system has been developed which delivers the same advantages of the setup matrix methodology (applicable in a multi-part, single machine system to take advantage of the reduced setup times that occur in part families), but is applicable to a large number of parts.

This new scheduling system, dubbed the QPR Scheduling system, is a modification of the classic (Q,R) scheduling and ordering methodology used in supply chain design today. This modified policy will take advantage of the reduced setup times that stem from producing parts within the same family back-to-back.
To review, the (Q,R) policy is a continuous review model in which the inventory level at a stock location is monitored in real time. At any point the inventory level dips below a threshold known as the reorder point, R, a fixed order quantity or lot size of Q parts is ordered or produced [11]. The reorder point $R$ consists of two components: the average inventory consumption over the lead time of the product, and safety stock to cover demand variation to an acceptable service level. The reorder point $R$ can be calculated as:

$$ R = D \times L + SS $$

(3.8)

where $D$ is the demand rate, $L$ is the lead time for the order or for the production, and SS is the safety stock to account for variation in the demand [11]. The reorder point $R$ is set so that the inventory in the system can cover the demand during the ordering/production lead time, so that stockouts are minimized or near-eliminated. The value of the order quantity $Q$ is based upon the optimal preferences for the inventory manager and product distributor. Typically, the proper order quantity or lot size for a part managed by the (Q,R) policy is determined by the Economic Order Quantity model (see Section 3.1). An example of an inventory tracked with a (Q,R) control policy is shown in Figure 5.

Figure 5: Illustration of the (Q, R) policy for continuous inventory review. $L$ is the lead time for a order of quantity $Q$. The dotted line represents the total number of units in stock and in process/transit, while the solid line represents actual inventory in stock [11].
This system reduces costs by limiting both the number of orders made over a certain time period, and the amount of inventory on hold, while still maintaining enough parts to produce to demand or feed the downstream supply chain. The principal limitation of the (Q,R) policy is that, just like the EOQ-model, it only can consider the production or management of one part in a vacuum, and therefore cannot be applied to our multiple-part, single-machine system.

For the modified QPR policy, an additional inventory point, known as the part family reorder point, or P, is included in the system. This part family reorder point is established to reduce the total setup time of the system by triggering production of parts in the same family consecutively, even if one of the parts is not at their standard reorder point R.

In the QPR system, if a part within a family reaches its normal reorder point R, then production of that part is scheduled, just like with the (Q,R) policy. At the end of the production run, the inventory level of all other parts in the same family is checked. If the inventory level of another part in the family is below its part family reorder point P, this means that there are still enough parts to cover the demand for the near future, and production is not needed immediately, but will be needed soon. Therefore, instead of putting off the production of this second part in order to set up a different production run, the QPR scheduling system will instruct the machinist to produce the second part in the family immediately. Since the machine is already setup to produce parts of that family, the second part will experience a reduced setup. Thus, the QPR scheduling procedure will reduce the total setup time of all parts on the machine.

As an example, consider a part family consisting of two parts, Part 1 and Part 2, each with its own particular lot size Qi, standard reorder point Ri, and part family reorder point Pi, produced on a single machine that manages more than two parts. If the inventory level for Part 1 reaches its standard reorder point Ri, then a production order for lot size Qi is placed into the system. When the production of Part 1 is completed, the inventory level of Part 2 is checked. If the inventory level of Part 2 is below R2, then as expected, production of Part 2 is set to begin; parts need to be made immediately in order to prevent backorders of Part 2 from developing. This policy is consistent between both the (Q,R) and QPR systems. On the other hand, if instead the inventory level of Part 2 is above R2 but below P2, the part is not currently at risk of accumulating backorders, but it indicates that production of Part 2 will soon be activated. The standard (Q,R) policy would designate that production of Part 2 should not begin, and the
machine would be torn down in order to set up the next part in the schedule. However, in the QPR system, the schedule takes advantage of the already set-up machine by triggering production of Part 2, since that part will have to be produced soon anyway. Instead of tearing down the machine at the end of production of Part 1 to setup for another part (family), making the new part, and then tearing down and setting up the machine once again when production of Part 2 is triggered by $R_2$, this system takes advantage of the reduced setup times of going from one part in a family to another. This decreases the total setup time of the production system, thereby increasing utilization. A comprehensive example of a two-part family under the QPR policy is shown in Figure 6.

Figure 6: A demonstration of the QPR Scheduling system. Part 1 and Part 2 are in the same part family. When Part 1 reaches $R_1$ for the first time, the machine is set up (which takes a certain amount of time), and then a batch of Part 1 is produced. At the end of the order, the inventory level of Part 2 is checked. Since the inventory level of Part 2 is above $P_2$, production of Part 2 does not occur. The second time Part 1 reaches $R_1$, the inventory level of Part 2 is again checked at the end of Part 1’s production. Since Part 2 is below its part family reorder point, the production of Part 2 is scheduled immediately, to take advantage of the already set-up machine. Notice how after the inventory level of Part 2 is check, the production of Part 2 begins without any setup delays.
This system can be expanded to a family with more than two parts. When any part in a family reaches its standard reorder point $R_i$, then a production order of that part for lot size $Q_j$ is placed into the system. When production of that part is finished, the inventory levels of all parts in the same family are checked. If any other part is below their respective part family reorder point $P_j$, then production of that part is placed into the system and ordered to begin immediately, taking advantage of the setup already on the machine.

In the QPR system, $Q$ is determined by the optimized lot sizing methodology shown in Section 3.1, which takes into account all of the parts that run on a particular machine. Further, since we are tracking multiple parts, the standard reorder point $R$ will need to be determined for each part based on its specific demand and production lead time. For the part family reorder point, $P$, there are two competing forces that affect its value. In general, parts with a small batch size or short process time will want to have a high $P$ value, in order to take advantage of reduced setups more often, since setup times will make up the bulk of the production time. On the other hand, a $P$-value that is too large will force the system to make parts in that family too often, starving the machine’s ability to produce other parts and causing backlogs in the rest of the system. For the sake of using a conservatively low value of $P$ so that as to avoid risk in a real, industrial production system, the $P$-value in this system has been set to the standard reorder point of the part, plus the demand rate of the part times the combined total process time for all other parts in the family. Thus, for our purposes, if a family consist of $n$ parts, $j=1...n$, then the $P$ value for any part in the family $i$ will be set to:

$$P_i = R_i + D_i \times \left( \sum_{j=1}^{n} T_j \right) - T_i$$

(3.9)

where $R$ is the standard reorder point for a part, $D$ is the demand rate for a part, and $T$ is the total process time (setup time plus run time) for a part.

In our implementation phase at Waters Corporation, we are only looking to apply our methodology to a limited number of parts and machines in order to prove out alterations before applying them full-scale. Due to the select number of parts focused on in the implementation
phase, the restricted amount of time available for system implementation, and the intuitive nature of the setup matrix methodology, improved part family scheduling through the setup matrix methodology will be implemented in the Machining Center for testing, along with optimized lot sizing and the part supermarket. To demonstrate the value of QPR scheduling, a MATLAB program has been developed to simulate a multi-part, single-machine system in which production for a family of parts is scheduled by the QPR inventory control system. This program will demonstrate the reduced setup times and increased productivity that occurs under QPR scheduling, verify the P-values chosen to maintain a stable production system, and provide insights and intuition into the effects of changing P-values on the production system. More information on the QPR simulation program, along with the results of the simulation, is described in Section 4.5.
4 System Improvements Implementation

To demonstrate the benefits of implementing larger production lot sizes in Turning and Milling, employing a part supermarket to decouple upstream and downstream processes, and applying scheduling procedures to take advantage of part family setups, certain part types that are produced in the Machining Center were selected to undergo these system improvements. In total, ten of the 1500 SKU’s that are fabricated in the Machining Center experienced the changes: one family of pump heads consisting of four parts, and two families of flow cells, each with three parts. Pump heads and flow cells were selected due to their consistently high demand, their production sequence which starts in Turning and continues on into Milling before going into other departments in the shop, and the difficulties supervisors stressed in producing these parts to demand and delivering them on-time to the Assembly area. The list of parts is shown below. All of the pump heads are produced on one designed machine in the Turning area, while all of the flow cells are produced on another machine; the same is also true for the Milling area. Therefore, our modifications affect four machines in total, two from each department.

<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Part Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>Pump Head</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>Pump Head</td>
</tr>
<tr>
<td>405012748</td>
<td>P3</td>
<td>Pump Head</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
<td>Pump Head</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>Flow Cell (Family #1)</td>
</tr>
<tr>
<td>289007447</td>
<td>F2</td>
<td>Flow Cell (Family #1)</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>Flow Cell (Family #1)</td>
</tr>
<tr>
<td>405007811</td>
<td>F4</td>
<td>Flow Cell (Family #2)</td>
</tr>
<tr>
<td>405008774</td>
<td>F5</td>
<td>Flow Cell (Family #2)</td>
</tr>
<tr>
<td>405001315</td>
<td>F6</td>
<td>Flow Cell (Family #2)</td>
</tr>
</tbody>
</table>

Table 7: The 10 parts selected for implementation of the system modifications, along with their part type and part family.
4.1 Optimized Lot Sizing

To determine the proper lot sizes for our ten parts, the demand rate, production rate, setup cost, and inventory holding cost for all parts which are produced on the same machines as the pump heads and flow cells must be obtained. Even though only the lot sizes for a few parts on each machine are being adjusted, it is important to include all parts in the analysis, so that lot sizes for the parts in question will not be too large as to take away the capacity of the machine for production of other parts, and limit any backorders. This information was obtained through the company’s SAP system.

The new lot sizes that were implemented on the shop floor are shown below. Five of the ten lot sizes remained the same, while four of the lot sizes were augmented between 50% and 200%. Not surprisingly, the lot sizes that were changed were for parts that have high demands into the Assembly area, between 1000 units and 7000 units per year. For most of these parts, the monthly demand rates were at least twice their former lot size, and in some cases up to four times as much as their former lot size. Therefore, these parts were produced at least once every two weeks; in some situations we observed, certain parts were even scheduled and produced multiple times a week. These new lot sizes will ensure that these high-volume parts are manufactured less frequently, in larger production runs, which will result in fewer setup hours needed on the machine. These lot sizes were put into place by manually overriding the lot size on the individual orders produced by SAP and placed on the Turning and Milling schedule during the implementation period.
<table>
<thead>
<tr>
<th>Part #</th>
<th>Previous Lot Size</th>
<th>New Lot Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>92</td>
<td>276</td>
</tr>
<tr>
<td>405013109</td>
<td>50</td>
<td>73</td>
</tr>
<tr>
<td>405012748</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>405008452</td>
<td>60</td>
<td>76</td>
</tr>
<tr>
<td>289003308</td>
<td>110</td>
<td>165</td>
</tr>
<tr>
<td>289007447</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>WAT081193</td>
<td>250</td>
<td>524</td>
</tr>
<tr>
<td>405007811</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>405008774</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>405001315</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 8: The previous lot sizes used in production of the ten parts under investigation, and their new lot size for the implementation period.

While the new lot sizes in the implementation plan are either the same or larger than their previous size, the direct results of our analysis using the equation from Pinedo did indicate that certain lot sizes should be reduced, mainly the parts in the small-volume family of flow cells (in purple). However, these lower lot sizes were not implemented for two reasons. The first is that a conservatively high estimate for inventory holding cost was used in our analysis, as the corporation did not have an estimate or value on-hand for calculating inventory holding costs. A high estimate was purposely chosen in order to be sure that the larger lot sizes for the high volume parts would not disrupt the production of other parts in the system. As the input for inventory holding costs in the Pinedo equation is reduced, the lot sizes for parts become much larger. The second reason is that previous lot sizes were selected in order to consume a set amount of raw material. The raw material for all the parts under investigation are long cylindrical bars known as “long stock”, which are fed into a lathe using automated feeders. Thus, it is important that the lot sizes used in production are selected so that the production run consumes either a full bar or a half-bar of raw stock. Any lot sizes smaller than the ones currently set would consume less than a half-bar of material during a production run, which would eventually cause unnecessary material set up and teardown to occur. This consideration
was carried over into the new lot sizes, which are chosen to in order consume either half or whole bars of raw material.

4.2 Part Supermarket

The part supermarket decouples the production of parts in Milling in Turning from the rest of the factory. This decoupling allows for Milling and Turning to fabricate parts in larger batch sizes (indicated in Section 4.1) without disrupting the rest of the system. The part supermarket also causes production of parts downstream of the Milling department to proceed based on realized demand instead of a predictive forecast. Furthermore, by scheduling these orders to start at the supermarket, the bottleneck processes (turning and milling) are removed from the production lead time.

As stated in Section 4.1, orders that were placed on the Turning and Milling schedules during the implementation period were manually adjusted from their previous lot size to the new lot size. These orders were still being triggered by SAP though, meaning that the Assembly area required the original order quantity to refill the finished-goods inventory at Assembly. Once the production order for the new lot size reached the end of Milling and arrived at the Milling utility operator, the original quantity of parts was sent downstream to the next operation, while the remaining units were held on a designated rack in the utility operator station. This became the inventory location for the part supermarket. A virtual stock location was also created in the company’s SAP system. The parts held at the supermarket were input into the computerized system, which would keep track of the inventory levels of the parts in the supermarket. Then, when a production order for the parts under consideration appeared on the schedule, the Planning department would check the inventory levels at the supermarket. If the inventory at the supermarket was enough to cover the order quantity, that quantity is sent from the supermarket to the next downstream process. If the inventory in the supermarket is not enough to cover the order quantity, an order is placed at the beginning of turning for those parts (at the new improved lot sizes), and the process repeats.

Ideally, there would be a safety stock level that exists in the supermarket such that if an order would bring the inventory level of the part to below the safety stock level (instead of below
zero), production of that part would then be placed on the schedule to refill the inventory at the supermarket. In this situation, instead of being triggered by an order, production in Turning and Milling would be triggered by an inventory threshold at the supermarket, and there would not be a case where the order for a part would need to begin at turning (instead, the order would always begin by bringing the units from the supermarket to the next process). Currently, the corporation does not have a way to decouple the production schedules of different departments in their SAP system, and hence the system described above could not be implemented; future recommendations will include having two production schedules for parts, one that is just for milling and turning (the bottleneck operations) to fill the supermarket, and one that is for downstream processes that pull from it.

4.3 Setup Matrix

The ten parts selected for the implementation phase of this project were chosen to take advantage of setup matrix scheduling. These parts all exist in part families, where the setup time to shift from one part in a family to another is reduced, and the setup time to transition to another part outside the family is consistently longer. Furthermore, any of the setups that force the machinist to alter production from one part type or family to another all take about the same amount of time. The standard setup time for these ten parts (the setup time going from a part outside the family, or from an unsetup machine), $s_{sd}$, was determined from SAP. From there, possible reductions in setup times were determined by looking at the setup protocols for the parts under investigation, and talking with machinists on the shop floor about what factors make setups easier and quicker.

During our investigation, the exact setup times for all transitions between the ten parts could not be determined; SAP had no data available regarding this information, while the machinist collected data was not comprehensive enough to populate a numerical table to a reasonable degree of certainty. However, by talking with machinists on the shop floor and learning how setups are performed, the relative difficulty of different setups was determined, along with why certain setups are quicker or easier. From this investigation, a hierarchy was developed that ranked the complexity of the different setups. The hierarchy is presented below, from the least complex setup change to the most complex. Complexity directly relates to the
time it takes to complete the setup on the machine, thus the following hierarchy is also a presentation of setup times from shortest to longest:

Setups only requires program swap, with no mechanical variation to the machine.
Setups that need adjusted tool offsets.
Setups for a part that is made out of a different raw material.
Setups that need different tooling but the same fixturing.
Setup requires a complete changeout of fixturing and tooling.

With this information, the comparative setup times for all parts under investigation could be determined, and this information was enough to initially populate the setup matrix. The setup matrix for the two machines and the parts under investigation are shown below, where the last row and column in each matrix represent all other parts that are produced on the same machine that are not in the part family. In the setup matrices below, P1 through P4 represent the four pump heads, F1 through F6 represent the six flow cells, p is the setup time for pump heads, f1 is the setup time for the high-volume family of flow cells, sf is the setup time for the sub-family within the high-volume family, f2 is the setup time for the low-volume family of flow cells, and s is the standard setup time, where p < s, sf < f1 < s, and f2 < s.

<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
</tr>
<tr>
<td>405012748</td>
<td>P3</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
</tr>
<tr>
<td>289007447</td>
<td>F2</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
</tr>
<tr>
<td>405007811</td>
<td>F4</td>
</tr>
<tr>
<td>405008774</td>
<td>F5</td>
</tr>
<tr>
<td>405001315</td>
<td>F6</td>
</tr>
</tbody>
</table>

Table 7-A: A portion of Table 7 has been reprinted here as a convenience to the reader, with part numbers, their respective part family (by color), and the part name shorthand.
All pump heads have the same reduced setup time when transitioning from one part in their family to another, and the same standard setup time with moving outside the family. This is also true for each family of flow cells. Thus, the schedule was manipulated so that all production runs for pump heads and flow cells are shifted so that they run back-to-back. Furthermore, the high-volume flow cell family has a subfamily within it. Therefore, the schedule was manipulated so that if any two parts in the family were on the schedule, their orders would be rearranged so they would run back to back, and if all three parts in the family were on the schedule, the subfamily would run first, and then the third part in the family would run. This setup matrix scheduling procedure will decrease the total setup time for all orders on the machine, and increase the machine’s utilization and effective capacity.

4.4 Automated Data Collection on Production System Performance

A few issues still remain in the implementation phase of this proposal. The first issue lies in the accuracy of the statistics used to ascertain proper lot sizes and form the setup matrices. As stated above, all of the inputs for determining the new lot sizes (demand rate, production rate, setup time, inventory holding cost) and framing the setup matrix were gathered from SAP and machinist collected data. It was noted earlier that the data from SAP can be wildly inaccurate, due to the fact that production information in SAP is based on set standards that are updated only every five years, and does not reflect what actually happens on the production floor. While our initial investigations indicate that the machinist collected data is more accurate than the
information in the SAP system, there is no set standard on how this data is amassed and processed, and typically the machinists only record estimates for setup and run time hours (machinists will not actually track how long certain processes take, so it is common that these values are rounded out to the closest hour or half-hour). There is also the issue of how to track the performance of the new production system so that it can be compared it to the old system.

To solve both of these issues, an automated data collection system was implemented in the Machining Center during the course of the project. This data collection system combines RFID tracking of parts with power monitoring of machines to capture setup time, production time, cycle time, and idle time for part production. The data collection system will allow us to gather more accurate data on the production of parts under investigation, such that lot sizes and the setup matrix can be constantly updated and optimized over time. The data collection system will also track the performance of the Machining Center to demonstrate improvements in utilization and productivity through the implemented system modifications. For more information on the design, implementation, and benefits of this data collection system, see Perez [2].

Our implementation period lasted 28 days. During this time, there were a total of 15 orders placed for the ten parts under investigation, for a total of 30 production runs in the Turning and Milling departments for the ten parts. Six of these runs did not have any adjustments in lot size or scheduling, and instead followed Waters’ current methodology. Fourteen of these runs had adjusted lot sizes, and 21 of the production runs experienced an altered schedule (11 of the runs had both the adjusted lot size and underwent an altered schedule). The affects of these changes on the production system were captured through SAP, the machinist collected data, and the newly designed automated data collection system.

4.5 QPR Scheduling

Due to the select number of parts used in the implementation phase and the limited amount of time available for execution, the QPR Scheduling procedure was not tested through implementation in Waters’ production system. Instead, the QPR Scheduling methodology was tested using a MATLAB simulation that is loosely based on a machine in the Turning
department. The purpose of the simulation is to determine how many setup hours would occur on a model system both with and without the QPR Scheduling Procedure, and what affects QPR Scheduling would have on potential backorders and average inventory levels.

Our simulation involves a machine that produces three parts: Part A, Part B, and Part C. Part A and B are individual parts that are in the same part family, and their production schedule will be managed by the QPR Scheduling Procedure. Part C is designed to represent all the other parts which are made on that particular machine. Each part has a standard setup time, production rate, and demand rate. Part A and Part B are produced in batches, while Part C is produced at a constant rate whenever Parts A and B are above their respective reorders points and do not need to be fabricated. For our simulation, we will assume that Part A and B are an ideal part family, where there is no setup time needed to go from production of Part A to Part B and vice versa, but there is a setup time necessary to go from Part A to Part C (and vice versa) and from Part B to Part C (and vice versa).

This simulation is designed to loosely model the lathe in Turning responsible for producing pump heads; Part A and B are based off of two pump heads under investigation in our actual implementation plan (405015367 and 405013109), and has the same demand rate, production rate, and setup time as the actual parts. Part C has the same setup time and demand rate as the averages of the 12 other parts manufactured on the machine. The production of Part C also takes into account that due to its high part mix (representing 12 different parts), Part C will have to undergo a setup after some amount of time, even if “Part C” is running continuously; we have determined how often a setup occurs on that machine for parts that are not the 405015367 or 405013109, and have included this factor in the simulation as well.

The simulated machine is designed to run 24 hours a day, seven days a week, and is managed through a periodic inventory review policy, with a review period of one day (24 hours). The inventory levels of Part A, B, and C are checked at the beginning of each day, and then a decision is made on what to produce during that day. At the end of each day, the respective demand is pulled from the inventories of all three parts. Initially in the simulation, the inventory levels of Part A and B are set to their production lot size (92 and 50, respectively), while the inventory level of Part C is set to zero. Thus, the machine will setup and produce Part C on the beginning of day 1. An internal counter keeps track of the total number of setup hours that occur...
in the system. The average setup time for Part C is 3 hours, and the production rate is 6 parts per hour. Therefore, on days where there is a setup for Part C, the machine can produce 126 parts in a day, while on days where there was no setup for Part C (i.e., the lathe was setup the previous day and left to run for more than one day) the machine can produce 144 parts in a day.

If the inventory level of either Part A or Part B is below their respective standard reorder point R when it is checked at the beginning of the day, the machine will transition to run that particular part, and the total setup time for the part (4 hours for either Part A and Part B) is added to the counter. The total production lead time (setup and run time) for both Part A and Part B is one day (which is within 13% of the actual lead time for 405015367 and 405013109, and chosen for simplicity), and after that time the designated lot size will be added to the particular inventory. At the end of a run of Part A or Part B (which also happens to be at the beginning of the next day), the inventory level of the other part in the family is checked. If the inventory of this other part is below their part family reorder point P, the other part will begin production immediately. For simplicity, the lead time for this order, even when running after another part in the same family, will be one day; however, no time will be added to the setup time counter (in a real system, the total production lead time for the second run would be shorter, therefore leading to increases in productivity). This simulation will only track the reduction in setup hours, without its effects of using that recouped setup time for added production. If the other part is not below their part family reorder point when this check occurs, production of Part C will be setup and continue again, and its setup hours will be added to the counter. The inventory levels of all three parts will be tracked and recorded every day by the simulation. This simulation is designed to run the machine over the course of one simulated year.

Information on the demand rate and setup times for all parts, the production lot sizes, lead time, and standard reorder points for Parts A and B, and the production rate for Part C used in this particular simulation is shown in Table 9:
Table 9: The demand rate and setup times for all parts, the production lot sizes, lead time, and standard reorder points for Parts A and B, and the production rate for Part C in the MATLAB simulation to test the QPR Scheduling procedure. Part A and B are in the same part family, and their production will be managed using QPR Scheduling.

The simulation was run eleven times, and the information in Table 9 was fixed across all runs. The only input that changed in each run was the part family reorder point, PA and PB, for the two parts in the part family. The first run was a control case, designated Test 0, where the machine runs without the QPR system in place. In this trial, the part family reorder point for each part Pi is set to its standard reorder point Ri. This is exactly as if the system is managed by the classic (Q,R) scheduling and ordering methodology used in supply chain design (This control was selected over a situation in which the P-value for each part is set to zero, because that would be a case in which a part family setup never occurs in the system, and the full setup time/cost occurs every time a setup happens. However, even in the scheduling system at Waters Corporation, part family setups do sometimes occur coincidentally, thus, the control scenario still allows for situations where part family setups occur coincidentally, when both parts happen to be at their standard reorder point within one day of each other). The next ten trials, Test 1 through 10, were run to test the QPR Scheduling system with various P-values. For each Test x, the part family reorder point P was set to:

\[ P_{A,x} = R_A + x(D_A L_B) \quad P_{B,x} = R_B + x(D_B L_A) \]

(4.3, 4.4)

where R is the standard reorder point, D is the demand rate, and L is the production lead time. Thus, when the QPR procedure is first initiated in the simulation, the part family reorder point
was set to the value that would be found using Equation 3.9, which is the standard reorder point plus the demand rate of the particular part multiplied by the lead time of all other parts in the family. In later tests, the P-value was then set to progressively higher values, to observe the effects of the changing P-value on the system.

A table of the variable inputs, $P_A$ and $P_B$, for both the control case and the various QPR test cases is shown below.

<table>
<thead>
<tr>
<th>Test Number</th>
<th>$P_i$</th>
<th>$P_A$</th>
<th>$P_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (Control)</td>
<td>$R_i$</td>
<td>9.5</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>$R_i + D_i L_j$</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>$R_i + 2D_i L_j$</td>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>$R_i + 3D_i L_j$</td>
<td>38</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>$R_i + 4D_i L_j$</td>
<td>47</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>$R_i + 5D_i L_j$</td>
<td>57</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>$R_i + 6D_i L_j$</td>
<td>66</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>$R_i + 7D_i L_j$</td>
<td>76</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>$R_i + 8D_i L_j$</td>
<td>85</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>$R_i + 9D_i L_j$</td>
<td>95</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>$R_i + 10D_i L_j$</td>
<td>104</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 10: The values for the part family reorder point for Part A and Part B for each test case.

The primary output of the simulation is the total setup time on the machine after one year. If the QPR Scheduling system is successful, there will be less setup time on the machine when QPR is in place compared to the control case. Therefore, we need to compare the final value of the setup time counter after one year for all tests. We also need to be sure that the QPR system does not take away the capacity of the machine to produce all the other parts that need to be made outside the part family. Thus, we have to examine the inventory levels of Part C, and be sure the QPR System does not cause Part C to fall into backorder. If the minimal value for the inventory of Part C does not fall below zero at any point during the year-long simulation, then we know that the system allows for enough capacity on the machine to produce Part C to demand. There is also the issue with potentially higher inventories of Part A and B, leading to
increased inventory holding costs. Therefore, we must also examine the average inventory of Parts A and B through the year, and compare their values to the control case.

The results of the simulation for the control case and each test case are shown below, where "% More Inv" denotes the percent more total average inventory for each test case compared to the control case.

<table>
<thead>
<tr>
<th>Test Number</th>
<th>Total Setup Time (hours)</th>
<th>Part C Backordered</th>
<th>Part A Average Inventory</th>
<th>Part B Average Inventory</th>
<th>Total Avg Inv</th>
<th>% More Inv</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (Control)</td>
<td>481</td>
<td>0</td>
<td>55.7</td>
<td>28.2</td>
<td>83.9</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>467</td>
<td>0</td>
<td>56</td>
<td>28.3</td>
<td>84.3</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>446</td>
<td>0</td>
<td>56.5</td>
<td>28.9</td>
<td>85.4</td>
<td>1.8</td>
</tr>
<tr>
<td>3</td>
<td>435</td>
<td>0</td>
<td>57.2</td>
<td>29.3</td>
<td>86.5</td>
<td>3.1</td>
</tr>
<tr>
<td>4</td>
<td>427</td>
<td>0</td>
<td>57.2</td>
<td>30.4</td>
<td>87.6</td>
<td>4.4</td>
</tr>
<tr>
<td>5</td>
<td>427</td>
<td>0</td>
<td>56.5</td>
<td>31</td>
<td>87.5</td>
<td>4.3</td>
</tr>
<tr>
<td>6</td>
<td>427</td>
<td>0</td>
<td>55.7</td>
<td>32.7</td>
<td>88.4</td>
<td>5.4</td>
</tr>
<tr>
<td>7</td>
<td>427</td>
<td>0</td>
<td>55.7</td>
<td>33.6</td>
<td>89.3</td>
<td>6.4</td>
</tr>
<tr>
<td>8</td>
<td>424</td>
<td>0</td>
<td>55.7</td>
<td>34.7</td>
<td>90.4</td>
<td>7.7</td>
</tr>
<tr>
<td>9</td>
<td>427</td>
<td>0</td>
<td>55.7</td>
<td>36.1</td>
<td>91.8</td>
<td>9.4</td>
</tr>
<tr>
<td>10</td>
<td>424</td>
<td>0</td>
<td>55.7</td>
<td>40</td>
<td>95.7</td>
<td>14.1</td>
</tr>
</tbody>
</table>

Table 11: The total setup time after one year, largest backorder of Part C, average inventory of Parts A and B, and the increase in inventory compared to the control point for all test cases. The total setup time for the machine is immediately reduced when the QPR system is in place, and is reduced further as the P-values are increased. Part C does not fall into backorder for any P value tested, but the total inventory of Part A and Part B does increase as P increases.
Figure 12: The total setup time after one year and the total average inventory for Parts A and B for each run of the simulation. As the P-values increase, the total setup hours decrease asymptotically to some minimal value, while inventory grows unbounded.

By applying the QPR Scheduling procedure with the minimum value for P, the total setup time on the machine was reduced by about 3% with minimal increases in inventory. For the fourth test case (where $P_i$ is set to the standard reorder point plus four times the demand rate of the part multiplied by the lead time of the other part in the family, $R_i+4D_iL_j$), the number of setup hours on the machine was reduced by 11%, while the average inventory only increased 4.4%. Additionally, Part C does not fall into backorder in any of the eleven test cases.

However, as P becomes larger, the machine experiences diminished returns on the reduction in total setup time, while the average inventory in the system continues to grow unbounded. This can be seen by the trend in Figure 12, and confirmed by running the simulation at higher values for P. Further, while none of the test cases tabulated in Table 11 produce backorders of Part C, when $x = 20$ or larger in Equation 4.3 & 4.4, meaning that $P_i = R_i + 20D_iL_j$,
the system begins accumulating backorders. Thus, there is a select number for the P value that benefits the system the most.

Nevertheless, the simulation demonstrated that with a proper selection of the part-family reorder point, the QPR Scheduling procedure can reduce the total setup time in a multiple-part-type single-machine manufacturing system with part families, while still allowing the machine to produce all parts to demand and without significant increases in inventory. Future work for the development of the QPR Scheduling methodology will be discussed in Section 6.
5 Results & Discussion

In the new production system, the total setup time on each machine was reduced using two methods. The first method was to increase the production lot size for high volume parts, and implement a part supermarket. These two changes decreased the number of production orders for a part over a period of time, and thus lowered the total number of setups completed. The second method was to schedule production runs of the same part family back to back; setups between parts in the same family are simpler and easier to do, and thus are quicker for the machinists to complete, also reducing the total setup time on a machine. The effects of each system adjustment will be discussed separately.

When discussing the effects of these changes on the production system, we must discuss all the consequences which result from the changes, not just the total setup time, but also machine utilization, on-time delivery percentage, inventory, any changes in order promptness or tardiness, and anything else that could affect the total performance of the manufacturing system. If these system changes cause problems or long-term inefficiencies in other areas or aspects of the Machining Center, then these changes may be counter-productive. For example, if a modification is put into place that saves setup hours but causes a build-up of inventory such that the inventory holding cost is higher than the saved production dollars, then the system changes end up hampering the bottom line of the Machining Center.

5.1 Optimized Lot Sizing and Part Supermarket

Seven orders on the SAP schedule were adjusted to the new lot sizes shown in Table 13. Since these lot sizes went into effect in both the Turning and Milling departments, a total of 14 production runs were completed with new lot sizes in place: seven in Turning and seven in Milling. By increasing the lot size for a production run, future orders and their setups could be removed from the production schedule, and their necessary quantities can be produced in one larger run. The number of orders placed with the modified lot size for each part, along with the
number of orders eliminated from the schedule by producing at the new lot size, is shown below. Since each order has to pass through and undergo a setup in both Turning and Milling, the total number of setups eliminated from the schedule is double the number of orders eliminated by cutting orders at the new lot size:

<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Previous Lot Size</th>
<th>New Lot Size</th>
<th>Total Orders Placed w/ New Lot Size</th>
<th>Total Orders Eliminated by Cutting Order(s) at New Lot Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>92</td>
<td>276</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>50</td>
<td>73</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
<td>60</td>
<td>76</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>110</td>
<td>165</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>250</td>
<td>524</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 13: The new lot sizes for parts whose lot size had changed, along with their previous values, and the number of orders placed with the new lot size. By producing in larger batches, orders (along with their setups) were eliminated from the schedule.

Information on the standard setup time for a part in both Turning and Milling was available through SAP and previously collected machinist data. By knowing the average setup time and the number of setups eliminated from the schedule, the total production hours saved due to changing lot size can be determined. Again, each order extracted from the schedule equates to one setup removed from both Turning and Milling, so each order eliminated recaptures both the turning setup time and milling setup time into production hours. In total, the new lot sizes eliminated 33.84 hours of setup time across four machines over 28 days.
<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Total Orders Eliminated</th>
<th>Turning Setup Time (hrs.)</th>
<th>Milling Setup Time (hrs.)</th>
<th>Total Hours Saved</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>2</td>
<td>4.67</td>
<td>1.75</td>
<td>12.84*</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>1</td>
<td>4</td>
<td>0.5</td>
<td>4.5</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>1</td>
<td>3.5</td>
<td>5</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33.84</td>
</tr>
</tbody>
</table>

Table 14: The total setup hours saved through the lot sizing methodology, broken down by part.

The recouped production time ranges between 4.5 and 8.5 hours for every order removed from the schedule. Note that the start (*) indicates that this was the total setup time saved for the 405015367 from eliminating two orders during the implementation period.

The primary concern with increasing production lot sizes and establishing a part supermarket is the costs associated with holding additional inventory. With parts now being held in an additional stock location positioned on the production floor, an added cost is incurred for each part stored there. However, the inventory holding costs for parts in the supermarket is minimal compared to the savings obtained through the recouped setup/production time.

Under the current system with no safety stock, the maximum inventory level for a particular part in the supermarket would be the new production lot size of that part; this would occur right after a supermarket inventory is replenished by an order (after the inventory level had previously fallen to zero). If this replenishment occurred on all four parts simultaneously, there would be a total of 986 parts sitting in the supermarket. Parts that are kept in the supermarket are stored in production bins, which are typically used to stock and transport goods around Waters’ facility (see Section 2.3.3). In total, the 986 parts would be fill 37 production bins. While a conservatively large value for inventory holding cost was used in our lot sizing methodology, an estimate for the inventory holding cost in the Machining Center was derived by the project team (based on proprietary information from Waters Corporation, including the approximate production value per square foot of their Machining Center) to be 35.7 cents per day ($2.50 per week) for every bin that is stored in an inventory location in Machining Center. In the worst case scenario, where the supermarket is filled with the maximum possible amount of parts (37
bins) for the entire length of the implementation period (28 days), the total inventory holding cost would be only $370. The number of setup hours that incurs the same cost to Waters Corporation is in the single digits. Therefore, the 33.84 recovered production hours heavily outweigh the inventory holding costs produced by increasing the lot sizes of the high volume parts.

5.2 Setup Matrix Scheduling

In our implementation period, 21 productions orders on the Turning and Milling schedules were adjusted to take advantage of setup matrix scheduling. These were production runs that were already placed on the Planning Department’s two-week schedule for Turning and Milling, and rearranged. An example of a previous production schedule and an updated schedule, formed through the setup matrix methodology and executed by Turning and Milling, is shown below.

<table>
<thead>
<tr>
<th>Order 1</th>
<th>Order 2</th>
<th>Order 3</th>
<th>Order 4</th>
<th>Order 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old Sequence</td>
<td>289003308</td>
<td>405008774</td>
<td>WAT081193</td>
<td>405007811</td>
</tr>
<tr>
<td>New Sequence</td>
<td>289003308</td>
<td>289007447</td>
<td>WAT081193</td>
<td>405008774</td>
</tr>
</tbody>
</table>

Table 15: An example of a production schedule produced by the Planning department, and the schedule for the same five order that was formed through the setup matrix methodology.

In the 21 production runs that were affected, 7 orders in Turning and 6 orders in Milling were placed directly after a part in the same family. Information on the reduced setup time for a part (otherwise known as the “part family setup time”) was available through both machinist collected data and the automated data collection system put into place during the implementation period of the project. Information on the standard setup time for a part was available both through SAP and previously-collected machinist data. The average length of a standard setup and a part family setup for each of the ten parts in both Turning and Milling is shown below. Note that on the actualized schedule, not all of the parts were able to run directly after a part in the same family during our implementation period. Thus, this information is not available.
### Table 16: The standard setup time and part family setup time for the ten parts under investigation in Turning (where data is available). The difference between the two setup times ranges between 1.5 and 3.92 hours.

<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Turning Standard Setup Time (hrs.)</th>
<th>Turning Part Family Setup Time (hrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>4.67</td>
<td>0.75</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>4</td>
<td>0.75</td>
</tr>
<tr>
<td>405012748</td>
<td>P3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>7</td>
<td>N/A</td>
</tr>
<tr>
<td>289007447</td>
<td>F2</td>
<td>4</td>
<td>1.75</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>3.5</td>
<td>2</td>
</tr>
<tr>
<td>405007811</td>
<td>F4</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>405008774</td>
<td>F5</td>
<td>9.5</td>
<td>N/A</td>
</tr>
<tr>
<td>405001315</td>
<td>F6</td>
<td>4</td>
<td>N/A</td>
</tr>
</tbody>
</table>

### Table 17: The standard setup time and part family setup time for the ten parts under investigation in Milling (where data is available). Here the difference between the two setup times is much less drastic, between 0.25 and 1.25 hours, while some parts do not see a reduced setup time at all.

<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Milling Standard Setup Time (hrs.)</th>
<th>Milling Part Family Setup Time (hrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>1.75</td>
<td>0.5</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>405012748</td>
<td>P3</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>289007447</td>
<td>F2</td>
<td>0.5</td>
<td>N/A</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>405007811</td>
<td>F4</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>405008774</td>
<td>F5</td>
<td>2</td>
<td>N/A</td>
</tr>
<tr>
<td>405001315</td>
<td>F6</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
All parts experience either a reduced setup time or the same setup time when running after a part in the same family. This difference between the standard and part family setup time is most apparent in the Turning department. Most parts under investigation have a Turning part family setup time that is about 50% of their standard length, while two of the pump heads see an 81% and 84% reduction in Turning setup time when made directly after a part in the same family. The difference between the part family setup time and the standard setup time in Turning ranges between 1.5 to almost 4 hours.

In Milling, the contrast is less pronounced. While certain parts did have a reduction in setup time between 33% and 71% compared to their standard length, the amount of time saved in completing a part family setup in Milling is less drastic, between 15 minutes and 75 minutes, while some parts did not see a reduced setup time at all. This was likely due to the already low setup times in the Milling department compared to the Turning department. When discussing our findings with employees at Waters, they expressed their recent initiatives to reduce setup times in the Milling department through careful planning of each machine’s tool gantry. The automated tool gantry on the mills in the Milling department can hold up to 180 tools at a time, and change between any of the tools in the gantry during a production run. By carefully selecting what tools remain standard in the machine, the machine can produce many different parts with little setup in between.

Table 18 displays the total setup/production saved by shifting the 21 orders on the Turning and Milling schedule amongst the 10 parts and across four machines. In total, 23.59 hours recovered on the four machines using the setup matrix scheduling procedure.

<table>
<thead>
<tr>
<th>Department</th>
<th>Turning</th>
<th>Milling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pump Head</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow Cell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Orders Adjusted</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Total Part Family Setups</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Total Hours Saved, Machine</td>
<td>18.34</td>
<td>2.25</td>
</tr>
<tr>
<td>Total Hours saved, Dept.</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

Table 18: Total setup time hours saved through the setup matrix methodology, broken down by part type and department.
A concern with the setup matrix methodology, and for that matter another concern with increasing lot sizes, is the effect on the on-time delivery of orders. Since the schedule had been rearranged from the sequence generated by Planning, which was based strictly by the due date to the next department, some orders had moved up in the schedule, while others moved back because of our adjustments. The same phenomenon occurs when we increase lot sizes; boosting the lot size in order to remove future orders from the schedule is just like moving up production orders for the same part to the front of the schedule so they run back-to-back.

In our situation, all of the orders in the Turning and Milling department were already well past their due date to the next department, so this caused some orders to make up time, while others fell further behind. However, we have to determine how the new schedule affected the promptness or lateness of certain orders, and how much that change affected or cost Waters Corporation.

This investigation began with research into how Waters tracks the cost of an order being late. Again, in our case, most orders on the production schedule were already behind by the time they began production, meaning that orders arrived late to their destination. Therefore, we had to determine the cost of a certain order becoming even more late to the next department (for orders sent back), and the savings for an order which catches up in the schedule (for orders shifted forward). After discussing with representatives in the financial department, it was determined that they did not even have a figure available on the cost of a late order compared to an on-time order, and that deriving a figure for the cost of an order become more late or catching up on the schedule (the difference between one day and two days late, two days to three days late, etc.) would be very difficult. Thus, quantifying this cost in terms of dollars would be infeasible.

One thing to note is that even with an adjusted schedule, the net number of days late across all orders on the schedule remains that same. Thus, adjusting the schedule does not increase the total days late for all the orders, and in fact can be manipulated in order to boost the on-time percentage of orders. However, when implementing large lot sizes or adjusted orders, certain orders will be pushed back on the schedule, and these individual orders would be made more late. This is the cost that occurs when shifting orders on the schedule. This cost is countered by the benefit of increased efficiency; the total set of orders on a production schedule can now be produced faster on the machine. Since this machine is producing these sets of orders
quicker, the next set of orders start sooner. This continues for every reduced or eliminated setup that occurs, and the succeeding set of orders start even sooner than the last. Eventually, we would reach a point where the saved setup hours would move the entire production schedule so far ahead, that even when an order is “pushed back”, that order starts at the same time as it would on the original schedule. This is known as the “breakeven point.” Before the breakeven point, the system would be producing all parts on the schedule with greater efficiency, and even completing certain parts before their previously scheduled due date. After this breakeven point, not only does the increased machine utilization persist, but now every order is completed before their previously scheduled due date.

Therefore, the best approach to determine the effects of the increased lot sizes and scheduling adjustments on order tardiness was to determine when this system would reach stability, or when all orders on the schedule would stop being more late, and no orders on the schedule will individually incur additional lateness.

Consider a system with three parts, Part 1 through Part 3, representing parts with both high and low demand rates. Part 1 has a demand rate of 200 units every four days, while Part 2 and Part 3 have a demand rate of 100 units every four days. Each of these orders has a setup time of 8 hours, and a production rate of 100 parts every 16 hours (therefore in one 24-hour day, an order can be setup and run to produce 100 parts). Currently, all parts are made in lot sizes of 100 units, and are produced in the following schedule according to their due dates.

![Production Schedule Diagram](image-url)

Figure 19: The production schedule for our system outlined above. Part 1 is in blue, Part 2 is in red, and Part 3 is in grey. The shaded areas represent setup time (8 hours) while the fully-colored areas represent production time (16 hours). Notice how every order takes one day to complete, and the production cycle (Part 1, Part 2, Part 1, Part 3) repeats every four days.
By increasing the lot size of Part 1 to 200 units, we eliminate a setup from the schedule, at the cost of pushing future orders back by the length of the production window (16 hours). However, instead of taking four days to complete all the orders, it took only 3.67 days, meaning the next production cycle can start 8 hours early. As the schedule continues through subsequent production cycles, the new schedule starts moving further and further ahead. Eventually, it reaches a point where all the orders have a finish date either before or the same as the previous schedule system. This occurs after two cycles, when the total saved setup time (2 setups of 8 hours each) equals the production time of the order shifted up on the schedule (16 hours).

Figure 20: The previous production schedule (above), and our new schedule that results when the lot size for Part 1 is increased from 100 units to 200 units. Notice that while certain orders (specifically orders for Part 2) are initially pushed back in the schedule to make room for the larger lot size, by day 10 all of the orders have due dates that are the same or ahead of their previous schedule. During this time, this system was able to produce the same amount of in parts in 11 days, while the old system would need 12 days, and from this point forward, all future orders will be ahead of their previous schedule.

Therefore, for our lot sizing methodology, the breakeven point occurs after \( g \) orders for a part, where \( g \) is the production time of the previous lot size divided by the setup time. Similarly, if a part is routinely pushed up in the schedule to take advantage of sequence-dependent setups, the breakeven point occurs when the total saved setup time is equal to or greater than the production time for one order of that part; this occurs after \( h \) orders for a part, were \( h \) is the production time at the old lot size divided by the standard setup time minus the part family setup time. The equations for \( g \) and \( h \) are:
\[ g = \frac{\text{Production Time}_{\text{old lot size}}}{\text{Standard Setup Time}} \]

\[ h = \frac{\text{Production Time}}{\text{Standard Setup Time} - \text{Part Family Setup Time}} \]

(5.1)

(5.2)

Therefore, we can determine how many orders need to go through the system, and how long it would take, before the system reaches stability and no orders have any additional late days. Since we have only complete information on the standard setup time (and not the part family setup time) for all ten parts under investigation, for comparison we will focus on the values for \( g \), the approximate number of orders that need to be placed under a new production lot size before the system reaches stability. The values for \( g \) are shown for each part in both Turning and Milling.

<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Production Time, Turning (hrs)</th>
<th>Setup Time, Turning (hrs)</th>
<th>Breakeven Point (orders)</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>23</td>
<td>4.67</td>
<td>5</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>19</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>405012748</td>
<td>P3</td>
<td>9.5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
<td>22.8</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>9.13</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>289007447</td>
<td>F2</td>
<td>4.05</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>20.75</td>
<td>3.5</td>
<td>6</td>
</tr>
<tr>
<td>405007811</td>
<td>F4</td>
<td>4.425</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>405008774</td>
<td>F5</td>
<td>18.92</td>
<td>9.5</td>
<td>2</td>
</tr>
<tr>
<td>405001315</td>
<td>F6</td>
<td>3.75</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 21: The breakeven point (in number of orders at the new lot size) where the scheduling system reaches stability for Turning. Notice how it takes as few as one order for the system to reach stability, with the longest time being six orders, or about five months.
<table>
<thead>
<tr>
<th>Part #</th>
<th>Part Name Shorthand</th>
<th>Production Time, Milling (hrs)</th>
<th>Setup Time, Milling (hrs)</th>
<th>Breakeven Point (orders)</th>
</tr>
</thead>
<tbody>
<tr>
<td>405015367</td>
<td>P1</td>
<td>16.56</td>
<td>1.75</td>
<td>10</td>
</tr>
<tr>
<td>405013109</td>
<td>P2</td>
<td>21</td>
<td>0.5</td>
<td>42</td>
</tr>
<tr>
<td>405012748</td>
<td>P3</td>
<td>6.65</td>
<td>0.5</td>
<td>14</td>
</tr>
<tr>
<td>405008452</td>
<td>P4</td>
<td>14.76</td>
<td>0.5</td>
<td>30</td>
</tr>
<tr>
<td>289003308</td>
<td>F1</td>
<td>28.6</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>289007447</td>
<td>F2</td>
<td>7.8</td>
<td>0.5</td>
<td>16</td>
</tr>
<tr>
<td>WAT081193</td>
<td>F3</td>
<td>25</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>405007811</td>
<td>F4</td>
<td>12.5</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>405008774</td>
<td>F5</td>
<td>25</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>405001315</td>
<td>F6</td>
<td>9</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 22: The breakeven point (in number of orders at the new lot size) where the scheduling system reaches stability for Milling. In contrast to Turning, it can take as many as 42 orders for the system to reach stability for a certain part, while even the fastest time to stability would be approximately four to five months.

In the Turning department, the system would quickly reach stability, as it would take at most six orders of any part before the schedule reaches the breakeven point, and for some orders as few as one or two. For the five parts that were recommended to change lot size, the longest time to stability would occur for the 405013109; due to its paucity in demand compared to the other parts which had adjusted lot sizes, this part would take approximately five months to reach its breakeven point.

For the Milling department, the time to stability is generally much longer. The number of orders required to reach a breakeven point is much more dispersed, ranging between 3 and 42 orders for a part. Compared to the Turning department, the time of stability for the five parts that were recommended to change lot size is much more severe. The longest time to stability again occurs for the 405013109, which would take approximately three and a half years, while even the fastest time to stability would take approximately four to five months. Just like in the results of the setup matrix methodology, this prolonged time to stability in the Milling department is due to the short setup times and the small setup time to production time ratio in Milling compared to...
Turning. Therefore, while the time of stability for the Turning department supports the implementation of the increased lot size and the setup matrix methodology, the same cannot be said for the Milling department. To justify the cost of tardiness added onto specific orders when implementing these system modifications, a true cost of comparison between the setup time hours saved and the cost incurred from the late orders must be executed.

In conclusion, both the modified lot sizes and the setup matrix methodology eradicate a significant amount of setup hours within the Turning and Milling departments. These setup hours can be directly translated back to production hours, increasing the utilization of the machines in these departments. These changes generate minimal inventory holding cost, and have minor effects on the promptness or tardiness of particular orders in areas with large setup time to production time ratios like the Turning department. For departments with small setup time to production time ratios like Milling, a cost comparison between the recovered production hours and the cost of order lateness must be completed to determine whether or not implementation of these changes is justified in those areas.
6 Conclusions and Future Work

Modifying production lot sizes and the setup matrix scheduling procedure are effective methodologies to reduce the total setup time, and thereby increase utilization, in a multiple-part-type manufacturing system with sequence dependent setups. These system modifications also generate a negligible increase in inventory holding cost, lower the average number of days late for all orders, and quickly dissipate any issues with specific order tardiness in situations where there are high setup time to production time ratios. In systems with low setup time to production time ratios, the cost of particular order tardiness must be calculated and compared to the financial returns from the reduction in setup times, in order to determine if these presented methodologies improve or exacerbate the overall bottom line of those systems. Further, while not tested in a true manufacturing system, the QPR Scheduling system demonstrated comparable diminutions in setup times and increases in productivity in a simulated manufacturing environment, while preventing backorders and adding minimal inventory holding costs.

For Waters Corporation, the conclusions and recommendations focus on both the continuation of the system improvements implemented during the course of this project, and other adjustments to their current manufacturing procedure which can aid in further increasing utilization and improving their bottom line. While both the lot sizing & part supermarket tandem and the setup matrix methodology lead to a significant increase in productivity on the machines in question, Waters Corporation should concentrate transient, immediate projects on extending the lot sizing methodology to other machines in the Turning and Milling area, and other machines throughout the facility. In general, with the production times, setup times, demand rates, and inventory holding costs for all parts processed on one machine, the Pinedo equation used in our implementation can be applied to any machine to determine optimal lot sizes; this can be used as a starting point to indicate certain parts that should increase in production size. Specifically in the Turning and Milling areas in the Machining Center, each mill and lathe has at least one part that is made in very high volume, on the order of 4000 or more units per year. By adjusting the lot size for these very high volume parts, either according to the Pinedo equation or to a value closer to the monthly demand (rather than the weekly demand) for that item, setup times can be reduced and utilization can be increased on the machines in these departments. Once the new lot sizes are in place on a machine, the total setup time can be tracked to determine
if the reduction in setup time befalls as expected; if it does, similar approaches can be applied to other high volume parts on those machines (between 1000 and 4000 units per year), potentially with lot sizes higher than what is indicated by Pinedo, and to other departments in Waters’ Machining Center of Excellence.

While changing the production lot sizes to decrease the number of setups required can be done in Turning and Milling easily to see immediate returns, the setup matrix methodology is harder to implement. The setup matrix methodology requires extensive knowledge and research into the compatibility of various parts within the manufacturing system; not only does it require information on the setup time for a particular part, but information on the setup time for a particular part after transitioning from another particular part, for all parts and combinations on a machine.

Through discussions with machinists on the shop floor, an intuition into what parts fall into part families, and therefore see a reduced setup time when running back-to-back on a machine, can be obtained. Those who are directly involved with the setup process will have at least a basic understanding as to what makes certain setups easier, and thus quicker, to complete, potentially allowing other part families to be identified as well. This would allow for a development of a setup matrix using relative, rather than exact, values, much like the methodology used during our testing period. More exact values for setups can be determined using time studies or an automated data collection system much like the one developed by Perez [2].

However, in the case of Waters Corporation, where there are over 1500 SKU’s produced through the Machining Center, gaining enough data to confidently populate a setup matrix can be a very time intensive task. In order to develop a fully optimize scheduling system, this entire process may require employees and support from various departments outside of Manufacturing, such as Planning, Finance, and Engineering. Further, these setup matrixes are databases which would constantly need to be maintained with the latest data on production processes and setup times, and updated whenever novel parts are added to the Machining Center (to allow for their scheduling). The difficulty of and prolonged time for implementation, along with potential issues already addressed about the setup matrix methodology itself (such as its problems when being applied to a large number of parts and its potential effects on areas with low setup time to
production time ratios) do not make it a suitable candidate for further exploration, especially compared to the lot sizing methodology.

Other changes to the Machining Center’s planning and production system that can complement the lot sizing and part supermarket methodologies include the implementation of transfer lots, or protocols which allow for the movement of a single bin of an order to the next process or department. Currently, there is no protocol in place which allows the machinist to move a single production bin of parts to the next station after it is completed, since all bins have to stay with their process sheet; instead, the entire production lot size has to be completed before being moved on to the next process (even in situations where the current production lot sizes call for over 350 parts and 12 bins for a production order). Initiating transfer lots in the Machining Center would prevent situations where a downstream machine is left idle because it is scheduled to work on a production order that has not arrived yet. With transfer lots in place, once a portion of that order is completed in the upstream department, it can be sent to the next process immediately, allowing that downstream process to begin, which increases overall throughput. Permitting production orders (especially with large lot sizes) to transfer to the next station one bin at a time will also remove another potential issue with increasing lot sizes in the Machining Center.

Furthermore, if the part supermarket is to remain after the Milling Department, the production schedules for the Turning and Milling department should be decoupled from the rest of the production system. During our implementation phase, the production of Turning and Milling were not scheduled to manage the inventory of parts in the supermarket, but rather by the strict demand forecast based in SAP. Instead, orders in Turning and Milling should be triggered by an inventory level (or reorder point) at the supermarket, and parts that are stocked in the supermarket should commence their downstream production from that inventory location. In the current manufacturing organization, this requires that the two schedules be separated on Waters’ computerized SAP scheduling system.

Outside of the system improvements instigated in the Turning and Milling departments as part of this thesis, final recommendations to improve productivity in the Machining Center include discerning other areas on the manufacturing floor where it could be beneficial to apply a supermarket. Departments or processes that have a long production lead time or a high part mix that would benefit from producing or handling larger batches would be good candidates to
implement supermarkets after. There is also the potential for establishing a part supermarket after Turning, either to replace or compliment the one introduced after Milling. Another technique that may assist in decreasing overall setup times is a protocol that allows for simple adjustments in the Turning and Milling Excel schedules that can be made by the machinists to take advantage of part family setups. Instead of developing a pervasive, complex methodology to generate an optimized schedule, adjusting certain orders already on the production schedule to take advantage of part families could produce financial returns simply and quickly. On the other hand, a procedure to lock down the production schedule on a machine a certain number of days in advance, and prohibit any changes to it, could also help the system by preventing orders from coming in and disrupting a set of orders in queue, causing wasteful setups.

Outside of the recommendations for Waters Corporation, the equation developed by Pinedo to generate lot sizes in an actual multiple-part, single-machine system to minimize cost while preventing backorders was successful for our purposes. Further development of the Pinedo equation should be conducted to account for randomness in a manufacturing system, but the Pinedo methodology should be considered in implementation in other multiple-part, single machine systems. Further, the setup matrix methodology was shown to be successful in a multiple-part-type manufacturing system with sequence dependent setups, especially in areas with high setup time to production time ratios. Limits of this system, including its moderated success in areas with low setup time to production time ratios, and issues with its application to a large number of parts (greater than 10) should be investigated.

Finally, we also observed how the QPR Scheduling procedure can reduce the total setup time in a simulated multiple-part-type manufacturing system, without causing parts to fall into backorder or the accumulation of large inventories. While the QPR Scheduling methodology is a promising way to decrease setup times through scheduling in a multi-part manufacturing environment with part families, additional investigation and testing must be completed in order to validate the QPR system before it can be implemented into an actual production line. While the QPR Scheduling procedure was designed for use in a machine with part families, a more strict definition of the characteristics of a production system which make QPR Scheduling a useful solution should be determined, before it can be applied to all real-world situations where part families exist (this is related to the concerns with the setup matrix methodology, where it was discovered that its benefits to machines with low setup time to production time ratios were
mitigated). Further, while a conservative value for the part family reorder point P was considered for possible implementation in Water's manufacturing system, various P-values were tested through simulation in order to gain some intuition as to what are appropriate P-values for the QPR system; in the future, the development of an equational, rather than analytical or iterative method, to determine the optimal P-value would make the QPR Scheduling system more empirical and effective.

Additional, more rigorous simulation testing should be conducted on the QPR system. First, future simulations of a similar multiple-part, single-machine system with part setups should be refined to include randomness in demand, setup times, failures, and repairs, and take into account more parts and multiple part families. Moreover, future simulations should investigate the success of the QPR Scheduling system on machines or production systems unlike the machine tested, such as ones with low setup time to production time ratios, less frequent setups, or varying demands of parts not in a part family. Finally, if the QPR Schedule system looks promising in a simulation comparable to an actual system, testing of the QPR Scheduling methodology in small, real-world systems should be conducted to further validate its potential benefits.
References


