Eindhoven University of Technology

BACHELOR

Simulation study of a newsvendor game in a two-server fork-join queue

Dijkstra, Vivianne C.

Award date:
2021

Eindhoven University of Technology
Department of Mathematics and Computer Science

Bachelor Final Project
2WH40

Simulation study of a newsvendor game in a two-server fork-join queue

Author
V.C. Dijkstra (1237687)

Supervisors
ir. C. Schol
prof. dr. M. Vlasiou

22/04/2021

Abstract

In this report, a simulation study is done on the two-server fork-join queue model. Since every job in the queue costs the server money, in this model it is assumed that servers can produce some inventory beforehand, in order to cover some of the delay they have. To this purpose, a newsvendor game is studied, in order to minimize the total costs with respect to the service rates and the inventory levels. For the arrival and service distributions, the exponential and Erlang distributions have been used. In addition, different strategies to divide the costs among the servers are studied, where the servers solely want to minimize their own costs. The total costs obtained for the Nash equilibria are compared to the total costs of the optimal solution in order to find a Nash equilibrium that has a total costs that is close to that of the optimal solution, and thus a small price of anarchy. From this research it can be concluded that giving the backlog costs to the slowest server in many cases results in the lowest price of anarchy.

Contents

Chapter 1

Introduction

In modern society, high-level technological products, such as smart phones, smart mobility and smart lighting systems, are essential in every day lives. ASML is a high-tech manufacturer that produces chips, which can be found in each of these intelligent products. Complex machines, which are used to detect and cure illnesses in early stages, can also be found in health care nowadays. These machines are considered to be complex, since they are made up out of many different components of high technological value. In reality, each of these components is produced by specialized manufacturers spread all over the country (or maybe even all over the world), that focus on the production of one typical product. Thus, the manufacturers merely focus on the assembly of the final product out of these complex components.

As soon as an order of such a complex product arrives at the factory of the manufacturer (e.g. ASML), the manufacturing of its components is put into action simultaneously. When all the components have been produced and delivered back to the factory, the final product can be assembled and shipped out to the client. Each supplier works at his own speed, meaning that not all components are finished at the same time. Components that are finished but cannot yet be assembled are temporarily stored in a warehouse. The server who has the slowest production time causes a delay in the delivery of the final product. So naturally the question arises: how can this delay be minimized?

A fork-join queue is characterized by one single arrival stream of jobs, where each job is split up into smaller tasks that are tackled by specialized servers. So every task is completed by a unique server and when all tasks are completed, they are bundled back together to form the final product. Thus, a fork-join queue is a simplified model of the problem sketched above, where the servers represent the suppliers of the high-tech manufacturer and the jobs in the system represent the components of the product. As denoted above, all suppliers can choose their own service rate. In addition, suppliers could also choose to produce some components beforehand that can be used to cover their delay, also known as the inventory. Thus, each supplier can choose its own strategy, consisting of their service rate and inventory choice.

A server does not want to produce too much inventory beforehand, since they have to pay for every part that is stored at the warehouse. However, they also do not want to have too much delay, since they also have to pay for each part that is in delay. Furthermore, each server makes costs that are proportional to the service rate, meaning that each server has costs related to the number of components that they produce. The goal is to find the perfect balance, such that the total costs are minimized. How much inventory each server chooses is related to the newsvendor problem.

Using the model as described above, different aspects of the problem can be approximated and their behavior can be analysed. More specifically, this report focuses on minimizing the total costs in the system and finding the strategies for each of the suppliers that result in the minimal total costs in the system. The total costs in the system are built up out of the production costs of the components, the holding costs of the components that are already produced but cannot yet be assembled and the backlog costs for the components that are in delay. Our initial goal is to minimize the total costs, as defined above, with respect to the inventory levels and the service rate.

In reality, each server is considered to be selfish and thus only cares about minimizing the costs he makes himself. This means that the global optimal solution to the minimization problem in the previous paragraph, does not have to be the optimal solution for each server individually. This raises a far more interesting question: what combination of inventory value and service rate would each supplier choose when he does not care about other suppliers? This question lies in the field of game theory and more specifically Nash equilibria. One set of containing the combined strategies of all servers that can be used to answer the previously mentioned question is referred to as a Nash equilibrium.

Even so, the optimal solution for the total costs can be used as a benchmark for the Nash equilibiria that are obtained using different cost division strategies. Namely, the ratio between the total costs that are obtained for the Nash equilibrium and the total costs of the global optimal solution is an indication for the performance of the Nash equilibria. This ratio is often referred to as the price of anarchy or the price of stability. The closer this ratio is to one, the better the Nash equilibrium performs. This observation promotes the following questions: how should the costs be divided among the suppliers such that the price of anarchy reaches 1? In other words: which cost division strategy results in a Nash equilibirum that has total costs close to that of the global optimal solution?

In order to find answers to the questions stated throughout this chapter, a simulation of the fork-join queue is made. A simulation study is done, since there are only few analytical results known for the fork-join queue. In Chapter 2, the different concepts are studied in relation to the already existing literature. In Chapter 3 the basic mathematical concepts that are needed to fully understand the model that is simulated are tackled. In particular, the concepts of the fork-join queue, the newsvendor problem and Nash equilibria are explained. Furthermore, the model parameters that are used in the simulation are described in this chapter and theoretical upper and lower bounds for the total costs are determined. These upper and lower bounds are determined, since they are easy to analyze and can be computed analytically. In Chapter 4, different arrival and service distributions are considered, for which the results of the simulation are displayed in both tables and figures. In Chapter 5 the main insights of the model and the corresponding simulation are summed up. Lastly, in Chapter 6, some points of improvement are addressed that can be implemented in further research on this topic.

The following research questions will be addressed throughout this report:

1. What effect does the choice of interarrival and service time distributions have on the optimal solution for the total costs and the Nash equilibria for the cost distribution functions?

2. Can the total costs be divided among the suppliers in such a way that the global optimum and the Nash equilibrium result in approximately the same total costs?

3. In what way does the choice of the cost parameters effect the results for the optimal solution for the total costs and the Nash equilibria for the different cost distribution functions?

Chapter 2

Literature

The concepts that are used throughout this report are not new, they have been studied by many mathematicians over the last two centuries. This chapter illustrates the research that has been done.

The first concept that lies at the core of this research project is the fork-join queue. Several different researches have been done on this concept. Flatto and Hahn studied the two server fork-join queue with the arrival stream of jobs being a Poisson process with rate $\lambda = 1$ and the service times being exponentially distributed [1]. In particular, they derived the generating functions for the equilibrium probabilities. Baccelli and Makowski also studied the fork-join queue in their article *The Fork-Join Queue and Related Systems with Synchronization Constraints: Stochastic Ordering and Computable Bounds* [2].

A problem that is very similar to the one that is treated in our research problem is described in an article by Kim and Agrawala [3]. In this paper, they consider a fork-join queue as an arrival stream of jobs that are split up into m subtasks that are tackled by m servers. Each server is thus in charge of one subtask. Subtasks that have been completed have to wait in a seperate queue until they can be assembled. This definition of the fork-join queue coincides with the research that is carried out throughout our report. This article uses the waiting times of the jobs in the queue in order to determine the solutions of the fork-join queue in transient and steady state.

Another topic that plays an important role throughout our research project is game theory, and in particular the concept of Nash equilibria and the newsvendor problem. The first mathematical formulation of game theory was generated by the German economist Oskar Morgenstern and Hungarian-American mathematician John von Neumann in the book *Theory of Games and Economic Behavior* [4].

A significant contribution to the field of game theory was made by the mathematician John Nash. He introduced the notion of Nash equilibria, named after himself, which can be seen as solutions to a game. A solution is a Nash equilibrium if for each server, deviating from the optimal solution whilst ignoring the strategy of the other servers, does not benefit him. This is also called a non-cooperative game. John Nash describes the concept of Nash equilibria in his paper *Equilibrium points in n-person games* [5].

The Prisoner's Dilemma is a concept related to Nash equilibria. The main idea behind this dilemma, is that in a game setting where two prisoners only care about what will benefit them, the two prisoners will not choose the strategy that has the best outcome for the system as whole. The formulation of this concept has been described by Ross [6] and Kuhn [7] in The Stanford Encyclopedia of Philosophy.

A notion that is used to test how well a Nash equilibrium performs in comparison to the optimal solution, is the price of anarchy. This notion was studied and described by Koutsoupias and Papadimitriou [8] and the inefficiency of Nash equilibria was studied by Dubey [9].

The minimization problem that is studied in this project, is in fact a newsvendor problem. A newsvendor problem is about the question how a retailer should balance between avoiding too large production costs and making as much revenue as possible. On one hand, when a retailer produces too many products, he could in retrospect, have produced less and still made the same revenue, and thus has too large production costs. While on the other hand, when he produces too little, he cannot satisfy all the demand and misses profit. The newsvendor problem has also been studied by many mathematicians. The first mathematicians to actually use the term where Morse and Kimball according to the paper *Novel Advances in Applications of the Newsvendor Model* [10]. Not only does this paper go into detail on the researchers that have already studied this model, it also studies the new applications of this model.

Chapter 3

Methods

In this chapter, a theoretical background is provided for a better understanding of the research project. Mathematical concepts, such as the fork-join queue, Nash equilibria and the newsvendor problem, that are necessary for the simulation are introduced and discussed. In addition, the different cost distributions are introduced and a theoretical approximation of the total costs using upper and lower bounds is presented. Lastly, the methods that are used to perform the simulations are presented.

3.1 Fork-join queue

The main topic of this research project is the fork-join queue. Fork-join queues are characterised by one single arrival stream of jobs, that is split up into N different subtasks that are each served by N different unique servers. When all subtasks are completed, they are bundled back together to form the final product and leave the system as a whole. This means that in order to determine the delay of the manufacturer, one needs to look at the queue of the slowest server. Figure 3.1 gives a schematic overview of the fork-join model as described above.

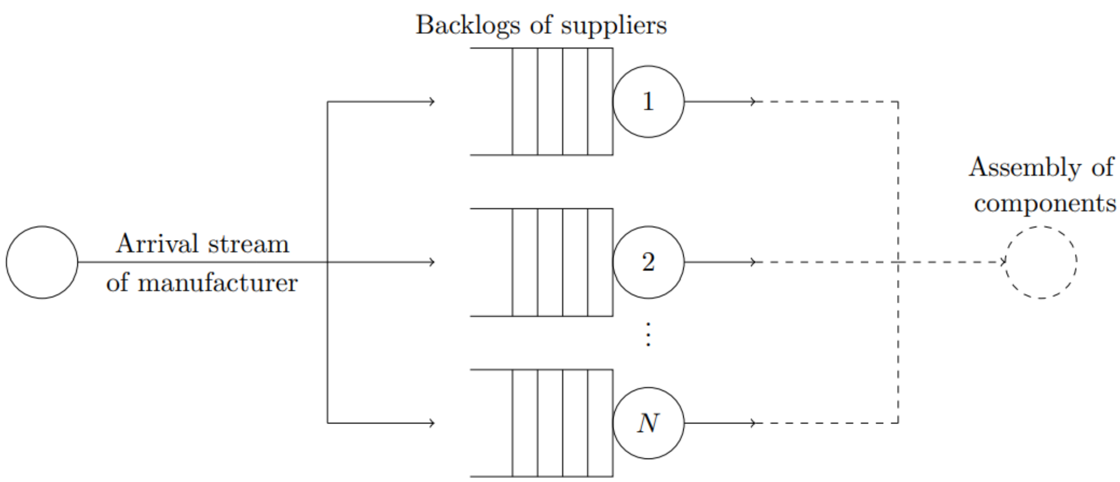

Figure 3.1: Fork-join queue with N servers. [11]

The interarrival times of the jobs in the arrival stream and the service times of the different servers can both follow any arbitrary distribution. In this report, the focus lies on the exponential distribution and the Erlang-r distribution. For the interarrival times, the exponential distribution takes a rate parameter λ, meaning that the expected interarrival time is equal to $1/\lambda$ and the variance is equal to $1/\lambda^2$. For the service times of a server i, the exponential distribution takes a rate parameter μ_i, where μ_i and μ_j for $i \neq j$ are not necessarily equal. Again, the expected service time for the exponential distribution is then equal to $1/\mu_i$ and the variance is equal to $1/\mu_i^2$. The Erlang-r distribution is equal to the sum of r exponentially distributed independent random variables with mean $1/\nu$, and therefore the mean of an Erlang-r distribution is equal to r/ν and the variance is equal to r/ν^2.

In this report, the expected interarrival times and service times are kept equal for both distributions. That means that in the case where the interarrival times are Erlang-r distributed, one takes the sum of r exponentially distributed independent random variables with mean $1/(r\lambda)$. In this case, the rate parameter ν, as mentioned earlier, is chosen to be $\nu = r\lambda$, so that the mean of the interarrival times becomes $r/(r\lambda) = 1/\lambda$, which is the same mean as for the exponential distribution. The variance in this case is equal to $r/(r\lambda)^2 = 1/(r\lambda^2)$

Something similar is done for the case where the service times are Erlang-r distributed, in this case one takes the sum of r exponentially distributed independent random variables with mean $1/(r\mu_i)$. In this case, the rate parameter ν, as mentioned earlier, is chosen to be $\nu = r\mu_i$, so that the mean of the interarrival times becomes $r/(r\mu_i) = 1/\mu_i$, which is the same mean as for the exponential distribution. The variance in this case is equal to $r/(r\mu_i)^2 = 1/(r\mu_i^2)$. This will be explained further in the upcoming sections.

All servers are considered to work independently of one another, meaning that the service of one server does not depend on the service of another server and vice versa. Therefore, the queue of each server can be considered seperately as N dependent $G/G/1$-queues, where the arrival of a job is at exactly the same time for each server. The first G stands for the distribution of the interarrival times and the second G stands for the distribution of the service times, the G can be any distribution of choice. An exponential distribution is often denoted by M and an Erlang-r distribution is often denoted by E_r. These notations are used throughout the report.

It may occur that some servers have finished components that cannot yet be assembled, because the other servers have not yet completed their components. In that case, the finished components are temporarily stored in a warehouse. In order to avoid having delays, servers can choose to create an inventory I_i beforehand. When a server has delay in comparison to the other servers, he can use the components in his inventory to make up for this delay.

3.2 Cost functions

There are different types of costs that come into play when producing a product and its components. Therefore, to determine the optimal combination of μ_i and I_i for each server i that minimizes the total costs, it is important to define the cost functions for each server i.

The total costs can roughly be split up into three components, namely the holding costs, the production costs of the components and the backlog costs. Holding costs are made when servers have to temporarily store finished components in a warehouse. For each server, the holding costs are determined by looking at the number of products that each server has stored in his warehouse. To determine this number, one has to look at the queue length of each server at the end of the simulation.

Each server has an inventory I_i that can be used to compensate for part of the delay. The variable $Q_i(\mu_i)$ stands for the queue length in front of server i with service rate μ_i and I_i is the inventory of server i. The difference $I_i - Q_i(\mu_i)$ then represents either the number of products that the server has overproduced or still has to produce. Furthermore, the maximum difference between each $(Q_j(\mu_j) - I_j)^+$ is considered, since it represents the delay that the slowest server has. The quantity $I_i - Q_i(\mu_i) + \max_{j\leq N}(Q_j(\mu_j) - I_j)^+$ represents the total number of products that server i has stored in the warehouse at the end of the simulation.

The second component of the total costs is devoted to the production costs. This component is determined by looking at the service rate of each server and multiplying it by the production costs per unit. The last component is the backlog costs. These are determined by looking at the maximum delay in the system. This is equal to $\max_{j\leq N}(Q_j(\mu_j) - I_j)^+$. The total costs for the delay are then equal to $b\max_{j\leq N}(Q_j(\mu_j) - I_j)^+$.

Combining the costs mentioned above results in the following formula for the total costs in the system:

$$\begin{aligned} C(\boldsymbol{I}, \boldsymbol{\mu}) = &\sum_{i=1}^{N} \left(\mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\leq N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i \right] \right) \\ &+ \mathbb{E}\left[b \max_{j\leq N}(Q_j(\mu_j) - I_j)^+ \right], \end{aligned} \tag{3.1}$$

where $\boldsymbol{I}$ is a vector containing the inventory levels for all servers and similarly $\boldsymbol{\mu}$ is a vector containing the service rates of all servers. Furthermore, h_i stands for the holding cost per product for server i, c_i stands for the production cost per unit of service rate for server i and b stands for the backlog cost per product in delay. Here it can be seen that the backlog costs are only paid for the delay of the slowest server, while the holding costs and production costs are paid by all the servers, where the holding and production cost parameters may differ for each component. The variable $Q_i(\mu_i)$ stands for the queue length in front of server i with service rate μ_i and I_i is the inventory of server i. We write $x^+ = \max(x, 0)$.

There are several ways in which the backlog costs can be distributed over the N servers. In the upcoming sections, six different cost division methods are described that are used in the simulation.

Symmetric choice

The first cost division method is the symmetric choice, meaning that the backlog costs are distributed evenly over all servers. The costs for server i, considering that there are N components to be produced by N servers, can be determined using the following formula:

$$\begin{aligned} C^{(i)}_{\text{symm}}(\boldsymbol{I}, \boldsymbol{\mu}) = &\,\mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\leq N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i \right] \\ &+ \mathbb{E}\left[\frac{b}{N} \max_{j\leq N}(Q_j(\mu_j) - I_j)^+ \right]. \end{aligned} \tag{3.2}$$

This choice for dividing the costs does not take into account the amount of delay that each server has. All servers have to pay the same amount of backlog costs.

Symmetric choice with punishment for delay

Another symmetric way to distribute the costs among the suppliers does take into account the amount of delay that the servers have. The amount of delay that each server has can be expressed by $(Q_i(\mu_i) - I_i)^+$. The backlog costs are divided among the servers using their amount of delay. This means that each server is accountable for $(Q_i(\mu_i) - I_i)^+ / \sum_{i=1}^N (Q_i(\mu_i) - I_i)^+$ part of the total backlog costs. The total amount of costs that each server has to pay can thus be expressed by the following formula:

$$\begin{aligned} C^{(i)}_{\text{punish}}(\boldsymbol{I}, \boldsymbol{\mu}) = \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right] \\ + \mathbb{E}\left[\frac{(Q_i(\mu_i) - I_i)^+}{\sum_{i=1}^N (Q_i(\mu_i) - I_i)^+} b \max_{j\le N}(Q_j(\mu_j) - I_j)^+\right]. \end{aligned} \tag{3.3}$$

These different ways of dividing costs among suppliers mentioned above give the same total costs. So in the first place they do not seem so important to investigate. However, when looking at Nash equilibria these different ways of dividing costs do have an influence.

Slowest server

Another way to distribute the backlog costs among the servers is to let the slowest servers account for the backlog. In this case, all backlog costs are given to the slowest servers. If there is only one slowest server, then this server gets all the backlog costs and the remaining servers get none. If there are multiple servers that are all equally slow, then all of these servers get an equal part of the backlog costs and the remaining servers get none.

Let J be the set of indices of all slowest servers and $|J|$ denote the number of elements in J. Then the total costs for one server i are defined in the following way:

$$C^{(i)}_{\text{slowest}}(\boldsymbol{I}, \boldsymbol{\mu}) = \begin{cases} \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right], & \text{if } i \notin J \\ \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right] & \\ +\mathbb{E}\left[\frac{b}{|J|} \max_{j\le N}(Q_j(\mu_j) - I_j)^+\right], & \text{if } i \in J. \end{cases} \tag{3.4}$$

Fastest server

In contrary to the method that was mentioned previously, one can also give the fastest server all of the backlog costs and the rest none. If there are multiple servers who are all equally fast, all of these servers get an equal part of the backlog costs and the remaining servers get none. If there is only one server who is the quickest, then this server gets all of the backlog costs and the rest of the servers get none.

Let J denote the set of indices of the fastest servers and $|J|$ denote the number of elements in J. Then the total costs for one server i are defined in the following way:

$$C^{(i)}_{\text{fastest}}(\boldsymbol{I}, \boldsymbol{\mu}) = \begin{cases} \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right], & \text{if } i \notin J \\ \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j\le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right] & \\ +\mathbb{E}\left[\frac{b}{|J|} \max_{j\le N}(Q_j(\mu_j) - I_j)^+\right], & \text{if } i \in J. \end{cases} \tag{3.5}$$

Production costs

The backlog costs could also be divided based on the productions costs of each server per unit. The part of the backlog costs that each server has to pay can be determined by $c_i / \sum_{i=1}^N c_i$. This is also illustrated by the formula below:

$$\begin{aligned} C^{(i)}_{\text{prod}}(\boldsymbol{I}, \boldsymbol{\mu}) = \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j \le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right] \\ + \mathbb{E}\left[\frac{c_i}{\sum_{i=1}^N c_i} b \max_{j \le N}(Q_j(\mu_j) - I_j)^+\right]. \end{aligned} \tag{3.6}$$

Holding costs

The backlog costs can also be divided based on the holding costs of each server per unit. The part of the backlog costs that each server has to pay can be determined by $h_i / \sum_{i=1}^N h_i$. This is also illustrated by the formula below:

$$\begin{aligned} C^{(i)}_{\text{hold}}(\boldsymbol{I}, \boldsymbol{\mu}) = \mathbb{E}\left[h_i(I_i - Q_i(\mu_i)) + h_i \max_{j \le N}(Q_j(\mu_j) - I_j)^+ + c_i\mu_i\right] \\ + \mathbb{E}\left[\frac{h_i}{\sum_{i=1}^N h_i} b \max_{j \le N}(Q_j(\mu_j) - I_j)^+\right]. \end{aligned} \tag{3.7}$$

3.3 Newsvendor problem

As the title already suggests, the newsvendor problem is related to the real-life case of a newsvendor. Every day, the newsvendor goes to his stand to sell his set of newspapers he chose to take with him that day. Depending on the demand, the seller can either make some loss, since he has newspapers left that he did not manage to sell, or he does not make enough profit, since the number of newspapers he brought was too few and could not cover the demand. Either case is undesirable, since in the case that the demand is smaller than his inventory, he loses money on the unsold papers because they cannot be sold the next day, and in the case that the demand is larger than his inventory, he loses additional profit. The question is: how many newspapers should he take with him every morning in order to optimize his profit?

The profit that the seller makes can be expressed in the following equation,

$$R(I) = p \cdot \min(D, I) - q \cdot I, \tag{3.8}$$

where $R(I)$ denotes the revenue for an inventory choice I, p denotes the price of a newspaper, q denotes the production cost of a newspaper and D is a random variable that denotes the demand.

The expected profit can then be expressed by the following equation,

$$\mathbb{E}\left[R(I)\right] = \mathbb{E}\left[p \cdot \min(D, I) - q \cdot I\right]. \tag{3.9}$$

The goal of the newsvendor is to maximize his profit. In order to do so, he needs to find an inventory level I such that his expected profit is maximized. This can be expressed by the following equation,

$$\max_{I \ge 0}\left(\mathbb{E}\left[R(I)\right]\right) = \max_{I \ge 0}\left(\mathbb{E}\left[p \cdot \min(D, I) - q \cdot I\right]\right). \tag{3.10}$$

This is the solution to the newsvendor problem. For more details, see [12, Ch. 4].

In this project, a minimization problem is studied, thus the aim is to minimize the costs rather than to maximize the profit. Furthermore, the objective function is given in Equation (3.1), and has 2N variables, representing the inventory levels and service rates of the N servers.

3.4 Nash equilibrium

Since this project encounters multiple servers who each choose their own strategy and do not cooperate, the servers do not necessarily choose the service rates and inventory levels that minimize the cost function in Equation (3.1) as their service rate and inventory level. Their costs are dependent on both their own decisions, and that of other servers. So, the question what inventory level and service rate will be chosen is a game-theoretical question. The Nash equilibrium is a general solution concept for these games.

One considers a game of N players where each player chooses his own strategy. The set of strategies from all players is called the strategy profile. A strategy profile is called a Nash equilibrium if none of the players could do better by changing only his own strategy [5]. In general it means that you ask yourself the following question: if the strategy of all other players but my own stays the same, is there a better strategy that I could choose for myself that will benefit me more?

The general concept can also be explained using an example. A game with two servers is considered: server 1 and server 2. There are two strategies that both servers can choose from: strategy A or B. Strategy A is defined in the following way: the server chooses a service rate of 1.5 and an inventory of 1. Strategy B is defined in the following way: the server chooses a service rate of 2.5 and an inventory of 0. A table is provided in which the different options and their outcomes are described. The outcomes are represented as a combination (C_1, C_2) of the total costs for each individual server. So C_1 represents the total costs for server 1 and the same for server 2. This table is displayed below.

		Server 2	
		Strategy A	Strategy B
Server 1	Strategy A	(2,2)	(10,1)
	Strategy B	(1,10)	(5,5)

Table 3.1: Example of possible strategy combinations and their outcomes.

From this table the Nash equilibria can be derived. Starting at a solution (2,2) where server 1 chooses strategy A and server 2 chooses strategy A, the Nash equilibrium is determined. Firstly, server 1 is considered and it is assumed that server 2 does not change his strategy. Therefore, it can be derived from Table 3.1 that if server 1 changes his strategy to strategy B, while server 2 remains remains at strategy A, that server 1 benefits from this move, since the total costs for server 1 decrease. Therefore the selected strategy for both servers as mentioned before is not a Nash equilibrium.

The same can be done for server 2 for the resulting combination of the last step (1,10). Assuming now that server 1 remains at strategy B and server 2 can choose to either remain at strategy A or move to strategy B. It can be derived from Table 3.1 that if server 2 moves to strategy B, while server 1 remains at strategy B, the total costs for server 2 decrease. These steps can be repeated until both servers decide to not change strategy anymore. If this result has been achieved, the resulting set of strategies is called a Nash equilibrium.

In the example mentioned above, the Nash equilibrium $(5, 5)$ is a strict Nash equilibrium, since changing strategy for either one of the servers results in a higher total costs. Therefore, the servers do not want to change their strategy, since this only results in loss for them.

In Table 3.1 it can also be seen that $(2, 2)$ is the optimal solution to the Newsvendor problem on Equation (3.1). In this case, the Nash equilibrium is not the same as the optimal solution. This means that servers making decisions based on their own game does not necessarily have to result in the optimal solution of the total cost function.

This is also related to the Prisoner's dilemma, as described by Ross [6] and Kuhn [7] in The Stanford Encyclopedia of Philosophy. The Prisoner's dilemma describes a similar situation, where two prisoners can choose between two strategies: confession or denial. If both servers confess, they both get the highest jail sentence. If either one of the prisoners confess and the other denies, then the prisoner that confesses gets no jail sentence, whereas the other prisoner faces a high jail sentence, but one that is lower than in the case that both prisoners confess. The last case is the case where both prisoners deny. In this case, the jail sentence is split between the two servers, these are the lowest jail sentences that both can get that are not equal to zero. Considering this problem, both prisoners will opt to confess, because the jail sentences for both of them are the lowest, considering that they have no clue what the other prisoner will choose. In this case it can be seen that the Nash equilibrium is not equal to the optimal solution in this case, which is for both of them to deny.

To measure how well a Nash equilibrium performs, it can be compared to the optimal solution for the total cost function. This is done using the so called price of anarchy or price of stability. The price of anarchy is defined as the ratio between the total cost of the Nash equilibrium and the total costs of the optimal solution [8]. This ratio is described in the following formula:

$$PoA = \frac{C_{\text{optimum}}(\boldsymbol{I}, \boldsymbol{\mu})}{C_{\text{Nash}}(\boldsymbol{I}, \boldsymbol{\mu})}, \tag{3.11}$$

where PoA denotes the price of anarchy, $C_{\text{optimum}}(\boldsymbol{I}, \boldsymbol{\mu})$ denotes the total costs of the global optimum and $C_{\text{Nash}}(\boldsymbol{I}, \boldsymbol{\mu})$ denotes the total costs of the Nash equilibrium. This ratio can take on values larger or equal to 1. The closer the price of anarchy is to 1, the better the Nash equilibrium performs.

This principle of Nash equilibria is used throughout this report to investigate whether different cost distribution functions for servers result in different Nash equilibria. In addition, the goal is to investigate whether there is a cost distribution function that has a Nash equilibria that is equal to the solution of the minimum total costs.

3.5 Distributions for the interarrival and service times

In the sections that follow, the approach from Adan and Resing [13, Ch. 6,8] is followed.

3.5.1 $M/M/1$-queue

An $M/M/1$-queue is a queue that has exponential interarrival times with mean $1/\lambda$ and exponential service times with mean $1/\mu_i$. The 1 represents the fact that there is only one server. For stability of the queue, to make sure that the queue lengths do not converge to ∞, it is required that $\rho_i = \lambda/\mu_i < 1$. Here ρ_i is called the occupation rate and indicates the fraction of time that the server is working.

The variable p_n denotes the probability that there are n jobs in the system and is described by the following equation:

$$p_n = (1 - \rho_i)\rho_i^n. \tag{3.12}$$

Using this function, the mean number of jobs in the system L can be determined using the following formula:

$$\mathbb{E}[L] = \sum_{n=0}^{\infty} n p_n = \frac{\rho_i}{1-\rho_i}. \tag{3.13}$$

These two characteristics of the $M/M/1$-queue are used later on to determine parameters for the model.

3.5.2 $M/E_r/1$-queue

An $M/E_r/1$-queue is a queue that has exponential interarrival times with mean $1/\lambda$ and Erlang-r distributed service times with mean $r/(r\mu_i)$, where r is the shape of the distribution and $r\mu_i$ is the rate. The variable μ_i denotes the service rate of server i. The 1 indicates that this queue consists of one server. The occupation rate is equal to $\rho_i = \lambda \cdot r/(r\mu_i) = \lambda/\mu_i$ and for the stability of the queue it is required that $\rho_i < 1$.

Let p_n be the equilibrium probability that there are n phases of work in the system. The queue length distribution is obtained by looking at the equilibrium equations of the queue, by setting the flow into the state n equal to the flow out of state n [13]. These equations are shown below:

$$\begin{align} \lambda p_0 &= r\mu_i p_1 \tag{3.14}\\ (\lambda + r\mu_i)p_n &= r\mu_i p_{n+1}, \qquad n = 1, \dots r-1 \tag{3.15}\\ (\lambda + r\mu_i)p_n &= \lambda p_{n-r} + r\mu_i p_{n+1}, \qquad n = r, r+1, r+2, \dots \tag{3.16}\end{align}$$

These equations can be solved by finding a solution of the form $p_n = x^n$ with $n = 0, 1, 2, \dots$. This solution is substituted into Equation (3.16) to find an expression for x, where $|x| < 1$ and $\mu > 1$. Throughout this report, it is assumed that $\lambda = 1$ and $r = 2$. Then, the equation that needs to be solved for x is the following:

$$(1 + 2\mu_i)x^2 = 1 + 2\mu_i x^3. \tag{3.17}$$

Solving this equation for x results in the following solutions:

$$x_1 = \frac{1}{4\mu_i}\left(1 - \sqrt{1+8\mu_i}\right) \quad \text{and} \quad x_2 = \frac{1}{4\mu_i}\left(1 + \sqrt{1+8\mu_i}\right) \tag{3.18}$$

A linear combination $p_n = c_1 x_1^n + c_2 x_2^n$ of both solutions is now considered. From Adan and Resing [13] it is known that the coefficients c_1 and c_2 satisfy,

$$c_1 = \frac{1 - \frac{1}{\mu_i}}{1 - \frac{x_2}{x_1}} \quad \text{and} \quad c_2 = \frac{1 - \frac{1}{\mu_i}}{1 - \frac{x_1}{x_2}}. \tag{3.19}$$

Let q_n be the probability that there are n jobs in the system. Then q_n can be expressed by the following formula:

$$q_n = c_1 \cdot (x_1^{-1} + 1)(x_1^2)^n + c_2 \cdot (x_2^{-1} + 1)(x_2^2)^n. \tag{3.20}$$

Using this probability the value of $\mathbb{E}\left[Q_i(\mu_i)\right]$ can be determined.

3.5.3 $E_r/M/1$-queue

An $E_r/M/1$-queue is a queue that has Erlang-r distributed interarrival times with mean $r/(\lambda r)$, where r is the shape parameter and λr is the rate, and exponentially distributed service times with mean $1/\mu_i$, where μ_i is the service rate of server i. The 1 indicates that this queue consists of one server. In this report, the expected value of the interarrival times for the Erlang-r distribution is kept equal to that of the $M/M/1$-queue. That means that for the interarrival times one takes $1/r$ times the sum of r exponential random variables. Then the occupation rate is equal to $\rho_i = (r\lambda)/(r\mu_i) = \lambda/\mu_i$ and it is required that $\rho_i < 1$ for the stability of the queue.

In order to find the distribution of the number of jobs in the queue, the following equation must be solved:

$$\sigma = \tilde{A}(\mu_i - \mu_i \sigma), \tag{3.21}$$

where $\tilde{A}$ is the Laplace transform of the Erlang-r distribution, μ_i is the service rate and σ is the parameter of the geometric distribution of the number of jobs in the queue. To keep the expected interarrival time equal, as mentioned earlier, the shape paramater must be equal to the rate parameter, taking that $\lambda = 1$. For the Laplace transform the following equation holds:

$$\tilde{A}(s) = \left(\frac{r}{r+s}\right)^r. \tag{3.22}$$

Throughout this report $r = 2$. Solving Equation (3.21) using the constraint that $0 < \sigma < 1$, the following holds:

$$\sigma = \frac{-\sqrt{\mu_i\,(\mu_i + 8)} + \mu_i + 4}{2\mu_i}. \tag{3.23}$$

From this it follows that the number of jobs just before arrival can be described using the following equation:

$$\mathbb{P}(L^a = n) = a_n = (1 - \sigma)\sigma^n, \tag{3.24}$$

where $\mathbb{P}(L^a = n)$ is the probability that there are n jobs in the queue just before arrival. Since $\mathbb{P}(L^a = n)$ follows a geometrical distribution, the numerical analysis is done for L^a. Nevertheless, one is aware that $\mathbb{P}(Q_i(\mu_i) = n) \neq \mathbb{P}(L^a = n)$.

3.6 Model parameters

In order to run the simulation, various model parameters need to be specified: the rate for the interarrival times λ, the service rate μ_i, the shape parameter r for the Erlang-r distributions, the number of servers N, h_i, c_i and b, the inventory levels I_i, and the simulation time T.

Throughout this report, the number of servers N is set to be 2. This is done because the simulation time increases exponentially when N increases. In order to compute the Nash equilibria, it is necessary to simulate all possible combinations of service rates μ_i for all servers. Therefore, increasing the number of servers N, means that the number of combinations of service rates increases exponentially and thus also the simulation time increases exponentially. Furthermore, in the cases that the interarrival times are exponentially distributed, the rate λ is set to be 1. In the case where the interarrival times have an Erlang-r distribution, the rate parameter $r\lambda$ and the shape parameter r satisfy $r = 2$ and $\lambda = 1$.

In the case that the service times are Erlang-r distributed, the shape parameter r is set to be 2. In this case the rate parameter is equal to $r\mu_i = 2\mu_i$. In the case that the service times are exponentially distributed, the rate parameter is equal to μ_i. Since the goal in this report is to find the combinations of μ_i such that the total costs are minimized, the values of μ_i are ranged. The values that need to be considered are determined using upper and lower bounds for the total costs. This is done in Section 3.7.

Different values for the parameters h_i, c_i and b are tested to determine the influence of each of these parameters on the choice of μ_i and I_i for each server. The combinations of parameters that are considered are shown in Table 3.2 below.

	h_1	h_2	c_1	c_2	b
CASE 1	1	1	1	1	1
CASE 2	1	2	1	2	1
CASE 3	1	1	1	1	2
CASE 4	1	2	2	1	1

Table 3.2: Combinations of the cost parameters c_i, h_i and b that are considered.

The values for the simulation time T and the inventory I_i of a server i are determined in the following sections.

3.6.1 Simulation time

In the model where the interarrival times and the service times are exponentially distributed, the individual queues of each server i are considered to be $M/M/1-$queues. Therefore, the simulation time can be determined by looking at the stationary distribution of the $M/M/1-$queue. Since the values in the set for the parameters μ_i are larger than 1, it can be said that the queues of both servers will reach a steady state.

Using Proposition 5.3 by Robert [14], the simulation time T can be determined that is needed to reach steady state. The formula as denoted in this proposition is displayed in the equation below:

$$||\mathbb{P}_x(L(T) \in \cdot) - G_\rho||_{tv} = \frac{1}{2}\sum_{n=0}^{+\infty} |\mathbb{P}_x(L(T) = n) - \rho_i^n(1-\rho_i)| \leq \left(\sqrt{\frac{\mu_i}{\lambda}}^{x} + 1\right) e^{-(\sqrt{\lambda}-\sqrt{\mu_i})^2 T}, \quad (3.25)$$

where G_ρ is the geometric distribution on $\mathbb{N}$ with parameter $\rho_i = \lambda/\mu < 1$, $x \in \mathbb{N}$ is the starting length of the queue and $T \geq 0$ and $\mathbb{P}_x(L(T) \in \cdot)$ is the distribution of the variable $L(T)$, the distribution of the number of jobs in the system, at time T. This inequality is an upper bound for the convergence rate of the total variation distance between the system at time T, and the system in steady state. Thus, this inequality gives us a measure how far the system is from steady state, at time T.

In our case, $x = 0$ and the values for μ_i are determined using the upper and lower bounds in Section 3.7. The value for the simulation time T is found by trying different values of T such that the distance to equilibrium $||\mathbb{P}_x(L(T) \in \cdot) - G_\rho||_{tv}$ of the distribution of $L(T)$ is as small as possible.

The simulation time T for a combination (μ_1, μ_2) that is used to simulate is determined by looking at the minimum of both μ_i's and using its value in the formula denoted above. The reason for this is that the queue with the smallest service rate has a longer simulation time before it reaches steady state, since it is closer to 1.

To determine the simulation time T, the right hand side of the formula is set equal to some parameter s that denotes the desired accuracy. This parameter s is chosen to be 0.01 in this simulation. The general formula that is used to determine the simulation time can then be denoted by, using that $\lambda = 1$:

$$2e^{-(1-\sqrt{\mu_i})^2 T} = 0.01. \tag{3.26}$$

This equation can be rewritten to obtain an expression for t:

$$T = \frac{\ln(200)}{(1-\sqrt{\mu_i})^2}. \tag{3.27}$$

Observe that when $\mu \downarrow 1$, $T \to \infty$.

The assumption is made that the simulation times as calculated for the $M/M/1$-queue can also be used for the $E_r/M/1$-queue and the $M/E_r/1$-queue.

3.6.2 Inventory level

In theory, the inventory level can take on any value in the set $\mathbb{N}$. However, it is impossible to consider all these values since this would take too many computations. Therefore, a limited set of inventory values is determined for each value of μ_i.

The expected queue length of a server is used to determine such a set. It follows from an $M/M/1$-queue that the queue length has a geometric distribution. Therefore the probability distribution for the queue length is:

$$\mathbb{P}(Q_i(\mu_i) = n) = (1-\rho_i)\rho_i^n, \quad n = 0, 1, 2, \ldots. \tag{3.28}$$

Consider this probability distribution and look at the probability that $\mathbb{P}(Q_i(\mu_i) \geq n) = 0.01$ for some $n \in \mathbb{N}$. This is the same as checking the following:

$$\mathbb{P}(Q_i(\mu_i) \geq n) = \sum_{j=n}^{\infty}(1-\rho_i)\rho_i^j = 0.01. \tag{3.29}$$

This value of n is determined for each value of μ_i. For each service rate μ_i the inventory level I_i is chosen from a set $[0, n]$.

The assumption is made that the way of calculating the inventory values for the $M/M/1$-queue as shown above also applies to the $E_r/M/1$-queue and the $M/E_r/1$-queue.

3.7 Theoretical bounds for the total costs

In the sections that follow, the approach from Adan and Resing [13, Ch. 6,8] is followed.

3.7.1 $M/M/1$-queue

To the best of our knowledge, the total costs can only be determined using a simulation, because the term $\mathbb{E}[\max_{j \leq N}(Q_j(\mu_j) - I_j)^+]$ is difficult to compute, since $Q_i(\mu_i)$ and $Q_j(\mu_j)$ have the same arrival process. However, easy to analyze upper and lower bounds are available.

Therefore, the total costs are approximated theoretically by considering upper and lower bounds. To determine such an lower and upper bound, each component of the total cost function is considered separately. To that extent, the total costs formula can be rewritten in the following way:

$$\begin{aligned} C(\boldsymbol{I},\boldsymbol{\mu}) = &\sum_{i=1}^{N}\left(\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right] + \mathbb{E}\left[h_i \max_{j\leq N}(Q_j(\mu_j) - I_j)^+\right] + \mathbb{E}\left[c_i\mu_i\right]\right) \\ &+ \mathbb{E}\left[b \max_{j\leq N}(Q_j(\mu_j) - I_j)^+\right]. \end{aligned} \tag{3.30}$$

To find a general expression for the upper and lower bounds, the cost parameters h_i, c_i and b remain undefined. Firstly, the expected value $\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right]$ can be determined exactly by the following:

$$\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right] = h_i\mathbb{E}\left[I_i - Q_i(\mu_i)\right] = h_i(\mathbb{E}\left[I_i\right] - \mathbb{E}[Q_i(\mu_i)]) = h_i(I_i - \mathbb{E}[Q_i(\mu_i)]). \tag{3.31}$$

Furthermore, it is known that $\mathbb{E}[Q_i(\mu_i)] = \rho_i/(1-\rho_i)$, with $\rho_i = \lambda/\mu_i$ and $\lambda = 1$. Substituting this back into Equation (3.31) results in the following:

$$\begin{aligned} \mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right] &= h_i\left(I_i - \mathbb{E}[Q_i(\mu_i)]\right) = h_i\left(I_i - \frac{\rho_i}{1-\rho_i}\right) \\ &= h_i\left(I_i - \frac{\frac{1}{\mu_i}}{1-\frac{1}{\mu_i}}\right) = h_i\left(I_i - \frac{1}{\mu_i - 1}\right). \end{aligned} \tag{3.32}$$

The expected value $\mathbb{E}\left[c_i\mu_i\right]$ can also be easily determined, since c_i and μ_i are both constants, so this results in the following:

$$\mathbb{E}\left[c_i\mu_i\right] = c_i\mu_i. \tag{3.33}$$

Now all that is left to be determined is $\mathbb{E}\left[\max_{j\leq N}(Q_j(\mu_j) - I_j)^+\right]$. This cannot be determined exactly, so an upper and lower bound is used to estimate this expected value. Since it is known that $(Q_j(\mu_j) - I_j)^+$ is always non-negative, it can be said that for all $n \in \{1,\ldots,N\}$, the following holds:

$$(Q_n(\mu_n) - I_n)^+ \leq \max_{j\leq N}(Q_j(\mu_j) - I_j)^+ \leq \sum_{j=1}^{N}(Q_j(\mu_j) - I_j)^+. \tag{3.34}$$

It can be said that for $Q_i(\mu_i) \in \{0,1,\ldots,I_i\}$ the value of $(Q_j(\mu_j) - I_j)^+$ is equal to 0. Therefore, the expected value of $(Q_j(\mu_j) - I_j)^+$ can be determined in the following way:

$$\begin{aligned} \mathbb{E}\left[(Q_j(\mu_j) - I_j)^+\right] &= \sum_{n=I_j}^{\infty}(n - I_j)\mathbb{P}(Q_j(\mu_j) = n) = \sum_{n=I_j}^{\infty}(n - I_j)(1-\rho_j)\rho_j^n \\ &= \frac{\rho_j^{1+I_j}}{1-\rho_j} = \frac{\mu_j^{-I_j}}{\mu_j - 1}. \end{aligned} \tag{3.35}$$

Using this result, the expected value $\mathbb{E}\left[\sum_{j=1}^{N}(Q_j(\mu_j) - I_j)^+\right]$ can also be easily determined, since the expected value of the sum of random variables is equal to the sum of the expected values of each random variable. So it follows that:

$$\mathbb{E}\left[\sum_{j=1}^{N}(Q_j(\mu_j) - I_j)^+\right] = \sum_{j=1}^{N}\mathbb{E}\left[(Q_j(\mu_j) - I_j)^+\right] = \sum_{j=1}^{N}\frac{\mu_j^{-I_j}}{\mu_j - 1}. \tag{3.36}$$

Combining the results obtained above, the following holds for the lower bound, for all $j \in \{1, \dots, N\}$:

$$\sum_{i=1}^{N} \left(h_i \left(I_i - \frac{1}{\mu_i - 1} + \frac{\mu_j^{-I_j}}{\mu_j - 1} \right) + c_i \mu_i \right) + b \frac{\mu_j^{-I_j}}{\mu_j - 1}. \tag{3.37}$$

For the upper bound the following holds:

$$\sum_{i=1}^{N} \left(h_i \left(I_i - \frac{1}{\mu_i - 1} + \sum_{j=1}^{N} \frac{\mu_j^{-I_j}}{\mu_j - 1} \right) + c_i \mu_i \right) + b \sum_{j=1}^{N} \frac{\mu_j^{-I_j}}{\mu_j - 1}. \tag{3.38}$$

For $N = 2$, the upper and lower bounds as defined above can be minimized in order to find the combination of service rates and inventory levels (μ_1, I_1, μ_2, I_2) for which the total cost function is minimal. All inventory values I_i as defined in Section 3.6.2 are considered. The service rate μ_i can take on all values in the real numbers that are larger than 1, this condition is necessary for the stability requirement of the queue. The upper and lower bounds as computed in this section are used to determine the values of μ_i that are considered in this simulation.

In this report, four different combinations of cost parameters are considered. For each of these combinations, the upper and lower bounds combinations have been determined. The results are displayed in the table below (μ_i rounded to one decimal).

Cost parameter combinations	Lower bound		Upper bound	
	Strategy	Total costs	Strategy	Total costs
$h_1 = 1, h_2 = 1, c_1 = 1, c_2 = 1, b = 1$	(1.7, 0, 1.7, 0)	4.8284	(2.1, 1, 2.1, 1)	6.9789
$h_1 = 1, h_2 = 2, c_1 = 1, c_2 = 2, b = 1$	(1.6, 0, 1.6, 0)	6.4641	(2.3, 1, 1.8, 1)	9.7453
$h_1 = 1, h_2 = 1, c_1 = 1, c_2 = 1, b = 2$	(1.8, 1, 1.8, 1)	5.8769	(2.3, 1, 2.3, 1)	7.7370
$h_1 = 1, h_2 = 2, c_1 = 2, c_2 = 1, b = 1$	(1.5, 1, 1.7, 0)	6.3952	(1.8, 2, 2.4, 0)	9.7212

Table 3.3: Upper and lower bounds for different cost parameter combinations for the $M/M/1$-queue.

The upper and lower bounds that have been computed above become less sharp for for $N > 2$, since the $\sum_{j=1}^{N} \frac{\mu_j^{-I_j}}{\mu_j - 1}$ term becomes larger. In these cases, it is possible to fins sharper upper and lower bounds.

3.7.2 $M/E_2/1$-queue

The second case is the case where the service times are Erlang-2 distributed, with shape parameter 2. Using a similar approach to the section above, an expression can be found for the upper and lower bounds for this case. Each of the components in the total cost function are looked at separately. In order to find a general expression for the upper and lower bound, the parameters h_i, c_i and b remain undefined.

An exact expression is determined for $\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right]$ similarly to what was done for the $M/M/1$-queue. Firstly an expression for $\mathbb{E}\left[Q_i(\mu_i)\right]$ is determined, this is done using the probability $\mathbb{P}(Q_i(\mu_i) = n) = q_n$ from Equation (3.20) that is defined in Section 3.5.2. For $\mathbb{E}\left[Q_i(\mu_i)\right]$ the following expression is obtained:

$$\mathbb{E}\left[Q_i(\mu_i)\right] = \sum_{n=0}^{\infty} n q_n = \sum_{n=0}^{\infty} n \left(c_1 (x_1^{-1} + 1)(x_1^2)^n + c_2 (x_2^{-1} + 1)(x_2^2)^n \right). \tag{3.39}$$

The exact value of this expression results in the following:

$$\mathbb{E}\left[Q_i(\mu_i)\right] = \frac{4\mu_i^2 - 1}{8(\mu_i - 1)\mu_i^2}. \tag{3.40}$$

Substituting this back into Equation (3.31) results in the following:

$$\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right] = h_i\left(I_i - \frac{4\mu_i^2 - 1}{8(\mu_i - 1)\mu_i^2}\right). \tag{3.41}$$

The expected value $\mathbb{E}\left[c_i\mu_i\right]$ is determined in the same way as for the $M/M/1$-queue, so for this expected value, Equation (3.33) still holds.

For $\mathbb{E}\left[\max_{j\leq N}(Q_j(\mu_j) - I_j)^+\right]$ the bounds in Equation (3.34) are applied. An expression for $\mathbb{E}\left[(Q_i(\mu_i) - I_i)^+\right]$ can be determined using the following,

$$\begin{aligned}\mathbb{E}\left[(Q_i(\mu_i) - I_i)^+\right] &= \sum_{n=I_i}^{\infty}(n - I_i)q_n \\ &= \sum_{n=I_i}^{\infty}(n - I_i)\left(c_1(x_1^{-1}+1)(x_1^2)^n + c_2(x_2^{-1}+1)(x_2^2)^n\right).\end{aligned} \tag{3.42}$$

The closed expression for this expectation is too long to display. Therefore, it has been left out of this report. Combining the formulas found above in the same way as was done for the $M/M/1$-queue results in expression for the upper and lower bounds.

The results for the upper and lower bounds for the four cases considered in the report are displayed in the table below.

Cost parameter combinations	Lower bound		Upper bound	
	Strategy	Total costs	Strategy	Total costs
$c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 1$	(1.6, 0, 1.6, 0)	4.6038	(2.0, 1, 2.0, 1)	6.4938
$c_1 = 1, c_2 = 2, h_1 = 1, h_2 = 2, b = 1$	(1.5, 0, 1.5, 0)	6.1662	(2.1, 1, 1.7, 1)	9.0774
$c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 2$	(1.7, 1, 1.7, 1)	5.4826	(2.1, 1, 2.1, 1)	7.1898
$c_1 = 2, c_2 = 1, h_1 = 1, h_2 = 2, b = 1$	(1.4, 1, 1.6, 0)	6.0544	(1.7, 2, 2.3, 0)	9.1313

Table 3.4: Upper and lower bounds for different cost parameter combinations for the $M/E_2/1$-queue.

3.7.3 $E_2/M/1$-queue

The third and final case is the case where the interarrival times are Erlang-2 distributed, with shape parameter 2. Using a similar approach to the sections above, an expression can be found for the upper and lower bounds for this case. Each of the components in the total cost function are looked at separately. In order to find a general expression for the upper and lower bound, the parameters h_i, c_i and b remain undefined.

An exact expression is determined for $\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right]$ similarly to what was done for the $M/M/1$-queue. Firstly, an expression for $\mathbb{E}\left[Q_i(\mu_i)\right]$ is computed, using a similar approach to the sections above. This is done using the distribution $\mathbb{P}(Q_i(\mu_i) = n)$ that was found in Section 3.5.3, from which it follows that,

$$\mathbb{E}\left[Q_i(\mu_i)\right] = \sum_{n=0}^{\infty} n\mathbb{P}(Q_i(\mu_i) = n) = \frac{\mu_i - \sqrt{\mu_i(\mu_i+8)} + 4}{\mu_i + \sqrt{\mu_i(\mu_i+8)} - 4}. \tag{3.43}$$

Substituting this back into Equation (3.31) results in the following,

$$\mathbb{E}\left[h_i(I_i - Q_i(\mu_i))\right] = h_i\left(I_i - \frac{\mu_i - \sqrt{\mu_i(\mu_i+8)} + 4}{\mu_i + \sqrt{\mu_i(\mu_i+8)} - 4}\right). \tag{3.44}$$

The expected value $\mathbb{E}\left[c_i\mu_i\right]$ is determined in the same way as for the $M/M/1$-queue, so for this expected value, Equation (3.33) still holds.

For $\mathbb{E}\left[\max_{j\leq N}(Q_j(\mu_j) - I_j)^+\right]$ the bounds in Equation (3.34) are applied. An expression for $\mathbb{E}\left[(Q_i(\mu_i) - I_i)^+\right]$ can be found in a similar manner as for the cases above. It follows that,

$$\mathbb{E}\left[(Q_i(\mu_i) - I_i)^+\right] = \sum_{n=I_i}^{\infty}(n - I_i)\mathbb{P}(Q_i(\mu_i) = n) = \frac{2^{-I_i}\mu_i\left(\frac{\mu_i - \sqrt{\mu_i(\mu_i+8)}+4}{\mu_i}\right)^{I_i+1}}{\mu_i + \sqrt{\mu_i(\mu_i+8)} - 4}. \tag{3.45}$$

Combining the formulas found above in the same way as was done for the $M/M/1$-queue results in expression for the upper and lower bounds.

The results for the upper and lower bounds for the four cases considered in the report are displayed in the table below.

Cost parameters	Lower bound		Upper bound	
	Strategy	Total costs	Strategy	Total costs
$c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 1$	(1.6, 0, 1.6, 0)	4.3062	(1.9, 1, 1.9, 1)	6.1208
$c_1 = 1, c_2 = 2, h_1 = 1, h_2 = 2, b = 1$	(1.5, 0, 1.5, 0)	5.8531	(2.0, 1, 1.9, 0)	8.5181
$c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 2$	(1.9, 0, 1.9, 0)	5.1926	(2.0, 1, 2.0, 1)	6.6511
$c_1 = 2, c_2 = 1, h_1 = 1, h_2 = 2, b = 1$	(1.4, 1, 1.6, 0)	5.8585	(1.8, 1, 2.2, 0)	8.4184

Table 3.5: Upper and lower bounds for different cost parameter combinations for the $E_2/M/1$-queue.

3.7.4 $M/M/1$ vs. $M/E_2/1$ vs. $E_2/M/1$

There are a couple of observations that can be made already based on the upper and lower bounds for the different types of interarrival and service time distributions. To make the comparison between the results for the upper and lower bounds for the different cases and distributions, a table has been made of the total costs. This is Table 3.6 below.

	CASE 1		CASE 2		CASE 3		CASE 4	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
$M/M/1$	4.8284	6.9789	6.4641	9.7453	5.8769	7.7370	6.3952	9.7212
$M/E_2/1$	4.6038	6.4938	6.1662	9.0774	5.4826	7.1898	6.0544	9.1313
$E_2/M/1$	4.3062	6.1208	5.8531	8.5181	5.1926	6.6511	5.8585	8.4184

Table 3.6: Upper and lower bounds for the optimal total costs for $M/M/1$ vs. $M/E_2/1$ vs. $E_2/M/1$.

In this table it can be observed that the total costs for the upper and lower bounds for the four cases are the highest for the $M/M/1$-queue. The costs decrease for the $M/E_2/1$-queue and they decrease even further for the $E_2/M/1$-queue. From this it can already be concluded that when the variance of the distribution decreases, so do the total costs.

A similar table to Table 3.6 has been made to compare the chosen service rates for the upper and lower bounds for each type of queue and each case that is considered. The results are shown in Table 3.7

	CASE 1		CASE 2		CASE 3		CASE 4	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
$M/M/1$	$(1.7, 1.7)$	$(2.1, 2.1)$	$(1.6, 1.6)$	$(2.3, 1.8)$	$(1.8, 1.8)$	$(2.3, 2.3)$	$(1.5, 1.7)$	$(1.8, 2.4)$
$M/E_2/1$	$(1.6, 1.6)$	$(2.0, 2.0)$	$(1.5, 1.5)$	$(2.1, 1.7)$	$(1.7, 1.7)$	$(2.1, 2.1)$	$(1.4, 1.6)$	$(1.7, 2.3)$
$E_2/M/1$	$(1.6, 1.6)$	$(1.9, 1.9)$	$(1.5, 1.5)$	$(2.0, 1.9)$	$(1.9, 1.9)$	$(2.0, 2.0)$	$(1.4, 1.6)$	$(1.8, 2.2)$

Table 3.7: Upper and lower bounds for the service rates for $M/M/1$ vs. $M/E_2/1$ vs. $E_2/M/1$.

In Table 3.7 it can be observed that the service rates in the different cases are the highest for the $M/M/1$-queue. The service rates decrease for the $M/E_2/1$-queue and decrease even further for the $E_2/M/1$-queue. This is the same observation as for the total costs.

The service rates that result in the upper and lower bounds for the total costs are used as upper and lower bounds for the service rates that are considered in the simulation. This is done to limit our search.

Chapter 4

Results

In this section, the results of the models with two servers are presented. Each server i has his own service rate μ_i, with $i \in \{1, 2\}$. The values for the service rates μ_i that are considered for each simulation are determined using the theoretical upper and lower bounds that were computed in Section 3.7.

The computation time of the simulations is a limitation for the number of runs and therefore the number of runs for all simulations is set to be 1000. The maximum distance between the simulation and the theoretical distribution is set to be 0.01. The simulation times are then computed using Equation (3.27) in Section 3.6.1. In order to be able to compare the influence of the arrival and service distributions on the obtained results, the same simulation times are used for $M/E_2/1$-queue and the $E_2/M/1$-queue, as for the $M/M/1$-queue. Different combinations of the parameters h_i, c_i and b are considered for each model.

The simulation determines the minimum total costs for each combination of service rates (μ_1, μ_2) based on the empirical mean of 1000 observations, considering all possible combinations of inventory values. The inventory levels that are considered for each pair are determined using Equation (3.29) in Section 3.6. For each pair of service rates (μ_1, μ_2) that is considered, the simulation determines the pair of inventory levels (I_1, I_2) that results in the minimum total costs for that combination of service rates (μ_1, μ_2). For each combination of service rates (μ_1, μ_2) the simulations thus give a pair of inventory levels (I_1, I_2) that results in the minimum total costs for that combination of service rates (μ_1, μ_2).

Out of these results, the global minimum of all minimum combinations of service rates and inventory levels (μ_1, I_1, μ_2, I_2) is determined to give the optimal combination of service rates and inventory levels $(\tilde{\mu}_1, \tilde{I}_1, \tilde{\mu}_2, \tilde{I}_2)$ that results in the global minimum of the total costs.

4.1 $M/M/1$-queue

The first model that is considered is the model where the interarrival times, as well as the service times, are exponentially distributed. The queue of each individual server can then be considered as an $M/M/1$-queue that has the same arrival stream as the other servers.

In Section 3.7.1, the upper and lower bounds for the total costs are computed for each combination of cost parameters that is considered. The service rates that result in these minimum total costs are displayed in Table 3.7. From this table, it can be seen that the service rates for this type of queue range from 1.5 to 2.4. To determine if this is indeed the case, the values of the service rates μ_i in the simulation are ranged from 1.4 to 2.4 with steps of 0.1. The simulation is run for all combinations of service rates (μ_1, μ_2). A step size of 0.1 is chosen since taking an even smaller step size did not show more accuracy.

4.1.1 CASE 1: $h_1 = 1$, $h_2 = 1$, $c_1 = 1$, $c_2 = 1$ and $b = 1$

Optimal strategy

In the initial case, the holding, backlog and production cost parameters are set to be 1.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds that were calculated in Section 3.7.1 suggest that the minimum of the total costs for this combination of the parameters is obtained for the service rates μ_i between 1.7 and 2.1.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(2.1, 0, 2.1, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 6.1170. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.1 and displayed in Table 3.3, it can be concluded that the optimal strategy resulting from the simulation is considerably close to the upper bound that was calculated for the total costs.

The difference between the two results is the choice for the inventory levels. The optimal inventory level choice for the theoretical upper bound is equal to $(1, 1)$ while the optimal inventory level choice from the simulation is $(0, 0)$.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.1 below. The inventory levels that assure these minimal costs are left out of the plot.

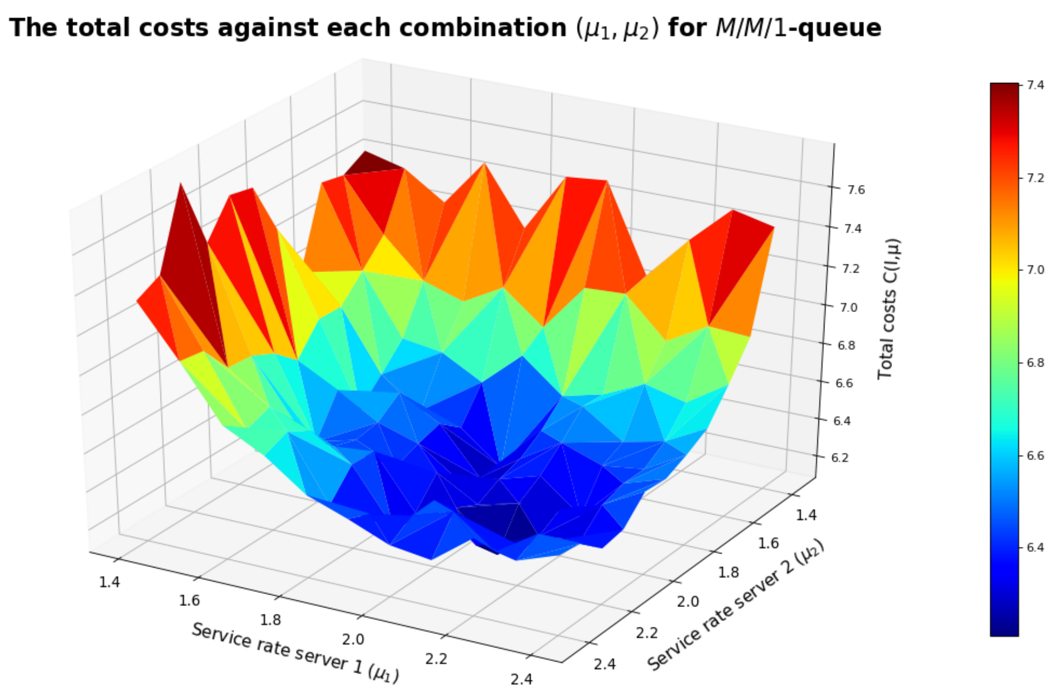

Figure 4.1: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/M/1$-queue for CASE 1.

In Figure 4.1 it can be seen that for the combinations of service rates (μ_1, μ_2) where one of the two service rates μ_i is small (close to 1.4) that the total costs are high. Combinations with both service rates μ_i closer to a value of 2.0 have lower total costs. Another observation is that for the combinations of service rates μ_i that have both service rates μ_i higher than 2.0 the total costs increase again. The optimum solution is clearly visible at $(2.1, 2.1)$ in the dark blue.

Furthermore, the plot seems to be symmetric around the plane $\mu_1 = \mu_2$. In addition, the figure clearly shows that the minimum is attained in the interior of the set of values that is observed. Thus, it can be said that the actual minimum of $C(\boldsymbol{I}, \boldsymbol{\mu})$ is very well approximated.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.1.

		Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution		6.1170	2.1	0	2.1	0
Symmetric	Nash equilibrium	7.9760	1.4	0	1.4	0
	Theoretical value	9.0717	1.7	0	1.7	0
Symmetric with Punishment	Nash equilibrium	6.2820	1.9	0	1.7	0
Slowest server	Nash equilibrium	6.1170	2.1	0	2.1	0
Fastest server	Nash equilibrium	7.9760	1.4	0	1.4	0
Production costs	Nash equilibrium	7.9760	1.4	0	1.4	0
Holding costs	Nash equilibrium	7.9760	1.4	0	1.4	0

Table 4.1: Results $M/M/1$-queue for $c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.4, 0, 1.4, 0)$ with a total cost of 7.976. In this case, the price of anarchy is the highest, in comparison to the other Nash equilibria. It must be noted that the service rate value $\mu_i = 1.4$ is at the boundary of the service rates μ_i that have been considered in this simulation. This could mean that the actual Nash equilibrium is at a lower value of μ_i.

Comparing this result to the theoretical value for the Nash equilibrium (as computed using the upper and lower bounds in Section 3.7.1), shows that there is a slight difference between the two. The theoretical result for the Nash equilibrium of the symmetric choice is $(1.7, 0, 1.7, 0)$ with a total cost of 9.0717. The theoretical Nash equilibrium is obtained for a combination of μ_i whose values are higher than those of the simulation result.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay is, also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.9, 0, 1.7, 0)$ with a total cost of 6.2820. Again it can be seen that the Nash equilibrium has an inventory choice of 0 for both servers. In this case, the service rates μ_i that both servers choose are different from one another.

Looking at the results in Table 4.1 it can be seen that assigning the backlog costs to the slowest server give a Nash equilibrium at the combination of service rates and inventory levels $(2.1, 0, 2.1, 0)$ with a total costs of 6.1170. This is in line with the results obtained for the optimal solution, which is also displayed in Table 4.1. Furthermore, assigning the backlog costs to the slowest server give the lowest price of anarchy out of all Nash equilibria. In this case the price of anarchy is equal to 1.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.4, 0, 1.4, 0)$ with a total costs of 7.9760. In these cases, the price of anarchy is the highest. Again, it has to be noted that the values of these service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.4]$.

4.1.2 CASE 2: $h_1 = 1$, $h_2 = 2$, $c_1 = 1$, $c_2 = 2$ and $b = 1$

In the second case, the holding cost parameters are set to be $h_1 = 1$ and $h_2 = 2$ and the same holds for the production cost parameters which are set to be $c_1 = 1$ and $c_2 = 2$. Lastly, the parameter for the backlog costs for all servers is set to be $b = 1$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds for the total costs can be calculated using the formulas in Section 3.7.1. It follows that the choice of service rates and inventory levels that results in the optimal costs for the theoretical lower bound is $(1.6, 0, 1.6, 0)$ and for the theoretical upper bound is $(2.3, 1, 1.8, 1)$.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.9, 0, 1.7, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 8.3420. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.1, it can be concluded that the optimal strategy resulting from the simulation falls between the two theoretical bounds.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.2 below. The inventory levels that assure these minimal costs are left out of the plot.

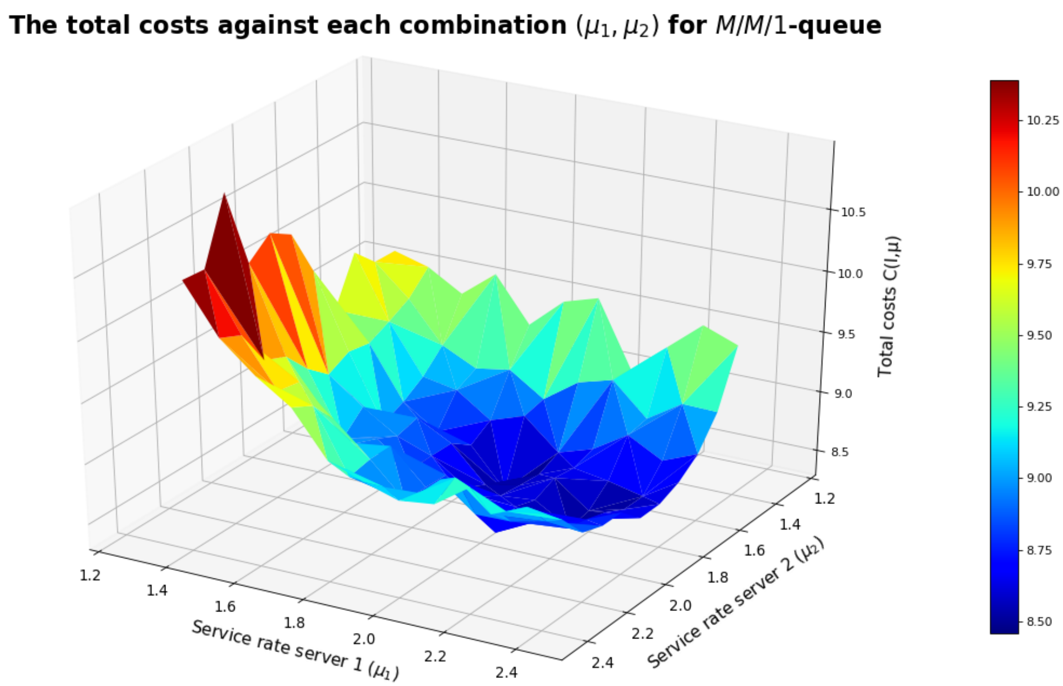

Figure 4.2: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/M/1$-queue for CASE 2.

In Figure 4.2 it can be seen that the plot is not symmetric, as was the case Section 4.1.1. The reason why the symmetry is lost here, is that the costs parameters differ per server, whereas for the previous case the cost parameters were the same. The highest total costs are obtained for combinations of service rates (μ_1, μ_2), where μ_1 is close to 1.4 and μ_2 is close to 2.4. The optimal combination is clearly visible in the color dark blue.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.2.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	8.3420	1.9	0	1.7	0
Symmetric	10.3370	1.4	0	1.4	0
Symmetric with Punishment	8.3420	1.9	0	1.7	0
Slowest server	8.3420	1.9	0	1.7	0
Fastest server	10.3370	1.4	0	1.4	0
Production costs	10.3370	1.4	0	1.4	0
Holding costs	10.3370	1.4	0	1.4	0

Table 4.2: Results $M/M/1$-queue for $c_1 = 1, c_2 = 2, h_1 = 1, h_2 = 2, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.4, 0, 1.4, 0)$ with a total cost of 10.3370. In this case, the price of anarchy is the largest, in comparison to the other Nash equilibria. It must be noted that a service rate of $\mu_i = 1.4$ is at the boundary of the service rates μ_i that have been considered in this simulation. This could mean that the actual Nash equilibrium is at a lower value of μ_i.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i $(1.9, 0, 1.7, 0)$ with a total cost of 8.3420. Again it can be seen that the Nash equilibrium has an inventory choice of 0 for both servers. In this case, the service rate μ_i that both servers choose are different from one another, which can be clarified due to the difference in holding and production cost parameters for both servers. Furthermore, this Nash equilibrium gives the same result as the optimal solution, so the price of anarchy in this case is equal to 1.

Looking at the results in Table 4.2 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium at the combination of service rates and inventory levels $(1.9, 0, 1.7, 0)$ with a total costs of 8.3420. This is in line with the results obtained for the optimal solution, which are also displayed in Table 4.2. In this case, the price of anarchy is equal to 1.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.4, 0, 1.4, 0)$ with a total costs of 10.3370. In these cases, the price of anarchy is the largest, in comparison to the other Nash equilibria. Again, it has to be noted that the service rates μ_i are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.4]$.

4.1.3 CASE 3: $h_1 = 1$, $h_2 = 1$, $c_1 = 1$, $c_2 = 1$ and $b = 2$

In the third case, the holding and production cost parameters are set to be 1. Lastly, the parameter for the backlog costs for all servers is set to be $b = 2$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.1. It follows that the choice of service rates and inventory levels that results in the optimal costs for the theoretical lower bound is $(1.8, 1, 1.8, 1)$ and for the theoretical upper bound $(2.3, 1, 2.3, 1)$.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(2.1, 1, 2.1, 1)$ is the optimal strategy for both servers and gives a minimum total costs equal to 6.9070. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.1, it can be concluded that the optimal strategy resulting from the simulation falls between the two theoretical bounds.

When comparing this optimal solution to that of case 1 in Section 4.1.1, it can be seen that the service rate choices for both servers has stayed the same, but the inventory values for both servers has increased to 1. This is in line with the expectations, since the holding and production cost parameters are the same for both servers.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.3 below. The inventory levels that assure these minimal costs are left out of the plot.

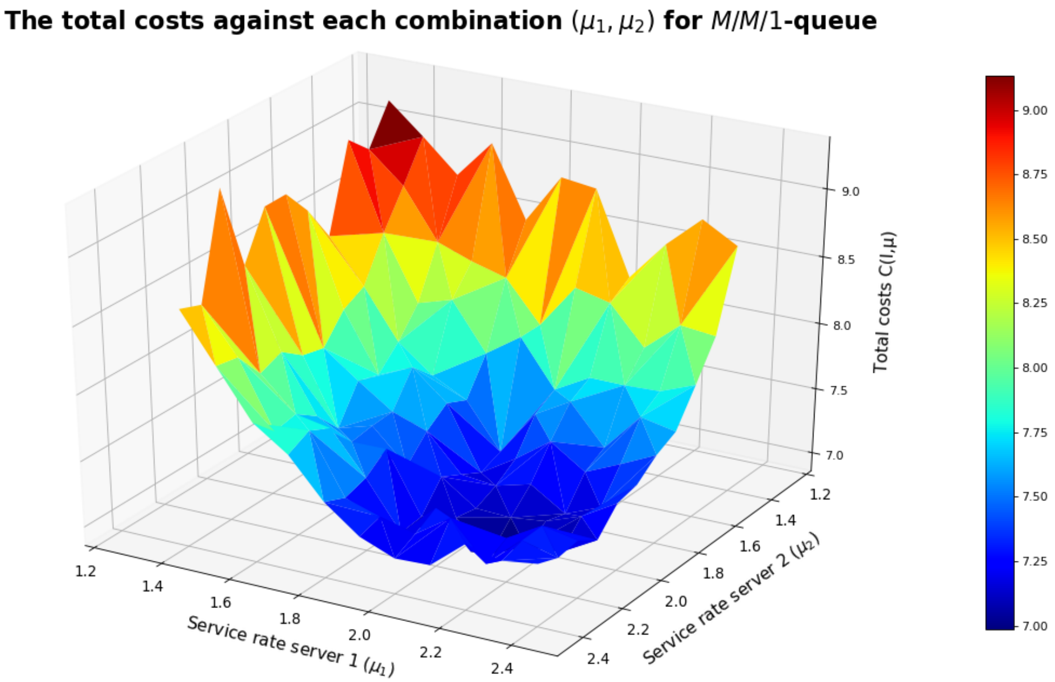

Figure 4.3: Plot of all combinations of service rates(μ_1, μ_2) and their minimal total costs for the $M/M/1$-queue for CASE 3.

In Figure 4.3 it can be seen that the plot is symmetric, similarly to the plot in Figure 4.1 in Section 4.1.1. The highest total costs are obtained for the combinations of service rates (μ_1, μ_2), where μ_1 and μ_2 are equal to 1.3. The optimal combination is clearly visible in the color dark blue. Again it can be noted clearly that the minimum is attained in the interior of the set of values that is observed.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.3.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	6.9070	2.1	1	2.1	1
Symmetric	7.9640	1.9	0	1.7	0
Symmetric with Punishment	6.9070	2.1	1	2.1	1
Slowest server	6.9070	2.1	1	2.1	1
Fastest server	11.3110	1.4	0	1.4	0
Production costs	7.9640	1.9	0	1.7	0
Holding costs	7.9640	1.9	0	1.7	0

Table 4.3: Results $M/M/1$-queue for $c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 2$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.9, 0, 1.7, 0)$ with a total cost of 7.9640. Both servers have a different service rate.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(2.1, 1, 2.1, 1)$ with a total cost of 6.9070. This Nash equilibrium gives the same result as the optimal solution. The price of anarchy in this case is equal to 1. Furthermore, it can be seen that the inventory choice of both servers is not equal to 0. This can be clarified due to the increase of the backlog cost parameter b.

Looking at the results in Table 4.3 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium at the combination of service rates and inventory levels $(2.1, 1, 2.1, 1)$ with a total costs of 6.9070. This is in line with the results obtained for the optimal solution, which is also displayed in Table 4.3. The price of anarchy in this case is equal to 1.

For the case where the backlog costs are assigned to the fastest server, the Nash equilibrium is obtained for the combination of service rates and inventory values $(1.4, 0, 1.4, 0)$ with a total costs of 11.3110. In this case, the price of anarchy is the highest out of all Nash equilibria. Again, it has to be noted that service rates μ_i are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibrium for this cost distribution is obtained for service rates $\mu_i \in (1, 1.4]$.

Lastly, it can be seen that the cost division strategies based on the holding and production costs have the same Nash equilibrium as the symmetric cost distribution. This can be clarified, since the holding and production cost parameters are the same for both servers, meaning that the costs are distributed symmetrically over the servers. So when the holding and production cost parameters are the same for both servers, the cost distribution functions that use the holding or productions costs are actually equal to the symmetric cost distribution.

4.1.4 CASE 4: $h_1 = 1$, $h_2 = 2$, $c_1 = 2$, $c_2 = 1$ and $b = 1$

In the fourth case, the holding cost parameters are set to be $h_1 = 1$ and $h_2 = 2$ and the production cost parameters are set to be $c_1 = 2$ and $c_2 = 1$. Lastly, the parameter for the backlog costs for all servers is set to be $b = 1$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.1. It follows that choice of service rates and inventory levels that results in the optimal costs for the theoretical lower bound is $(1.5, 1, 1.7, 0)$ and for the theoretical upper bound is $(1.8, 2, 2.4, 0)$. In the upper and lower bounds it is visible that the chosen inventory level for server 1 is higher than for server 2.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.9, 1, 2.2, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 8.5030. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.1, it can be concluded that the optimal strategy resulting from the simulation falls between the two theoretical bounds. Furthermore, it can be seen that the inventory level for server 1 is not chosen to be 0 which is the same as for the upper and lower bounds.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.4 below. The inventory levels that assure this minimal cost are left out of the plot.

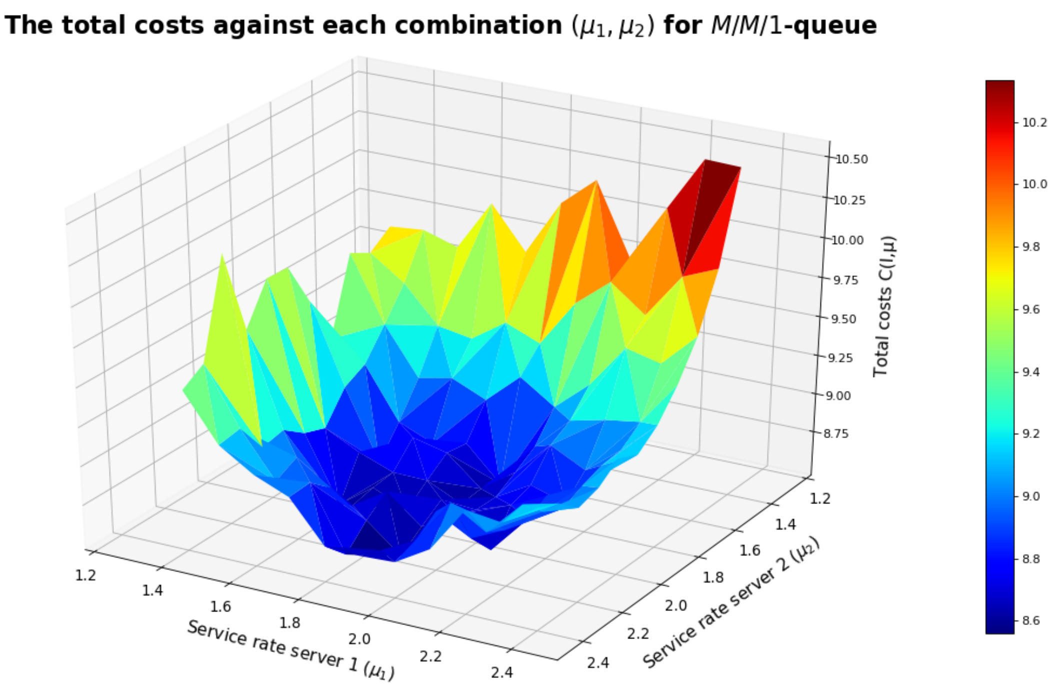

Figure 4.4: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/M/1$-queue for CASE 4.

In Figure 4.4 it can be seen that the plot is not symmetric, similarly to the plot in Figure 4.2 in Section 4.1.2. The highest total costs are obtained for the combinations of service rates (μ_1, μ_2), where μ_1 is equal to 2.4 and μ_2 are equal to 1.3. The optimal combination is clearly visible in the color dark blue.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.4.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	8.5030	1.9	1	2.2	0
Symmetric	10.3370	1.4	0	1.4	0
Symmetric with Punishment	10.3370	1.4	0	1.4	0
Slowest server	8.9290	1.5	1	1.9	0
Fastest server	10.3370	1.4	0	1.4	0
Production costs	10.3370	1.4	0	1.4	0
Holding costs	10.3370	1.4	0	1.4	0

Table 4.4: Results $M/M/1$-queue for $c_1 = 2, c_2 = 1, h_1 = 1, h_2 = 2, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.4, 0, 1.4, 0)$ with a total cost of 10.3370. It must be noted that a service rate of $\mu_i = 1.4$ is at the boundary of the service rates μ_i that have been considered in this simulation. This could mean that the actual Nash equilibrium is at a lower value of μ_i.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. This gives the same result as for the symmetric choice of division of costs. Both ways of dividing the costs result in a high price of anarchy.

Looking at the results in Table 4.4 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium for the combination of service rates and inventory levels $(1.5, 1, 1.9, 0)$ with a total cost of 8.9290. This Nash equilibrium does not give the same result as the optimal solution as seen before in the other cases. The price of anarchy in this case is the lowest in comparison to the other Nash equilibria. Furthermore, it can be seen that the inventory level of server 1 is not equal to 0. This can be clarified due to the increase of the production cost parameter c_1. Both servers choose different service rates, this can be clarified by the difference in cost and production parameters between the two servers.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.4, 0, 1.4, 0)$ with a total costs of 10.3370. In these cases, the price of anarchy is the highest in comparison to the other Nash equilibria. Again, it has to be noted that the service rates μ_i are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.4]$.

4.2 $M/E_2/1$-queue

The second model that is considered is the model where the interarrival times are exponentially distributed, but the service times follow an Erlang-2 distribution. The queue of each individual server can then be considered as $M/E_2/1$-queues with a joint arrival stream.

In Section 3.7.2, the upper and lower bounds for the total costs are computed for each combination of cost parameters that is considered. The service rates that result in these minimum total costs are displayed in Table 3.7. From this table, it can be seen that the service rates for this type of queue range from 1.4 to 2.3. To determine if this is indeed the case, the values of the service rates μ_i in the simulation are ranged from 1.3 to 2.4 with steps of 0.1. The simulation is run for all combinations of service rates (μ_1, μ_2). A step size of 0.1 is chosen since taking an even smaller step size did not show more accuracy.

4.2.1 CASE 1: $h_1 = 1$, $h_2 = 1$, $c_1 = 1$, $c_2 = 1$ and $b = 1$

Optimal strategy

In the initial case, the holding, production and backlog cost parameters are set to be $b = 1$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds that were calculated in Section 3.7.2 suggest that the minimum of the total costs for this combination of the parameters is obtained for the service rates μ_i between 1.6 and 2.0.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.9, 0, 1.9, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 5.6500. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.2, it can be concluded that the optimal strategy resulting from the simulation is considerably close to the upper bound that is calculated for the total costs.

The difference between the two results is the choice for the inventory levels. The optimal inventory choice for the theoretical upper bound is equal to $(1, 1)$ while the optimal inventory choice from the simulation is $(0, 0)$. Furthermore, there is a slight difference in the service rates. For the theoretical results the pair of service rates $(2.0, 2.0)$ gives the optimal upper bound for the total costs, while the optimal solution of the simulation had $\mu_1 = 1.9$.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.5 below. The inventories that assure these minimal costs are left out of the plot.

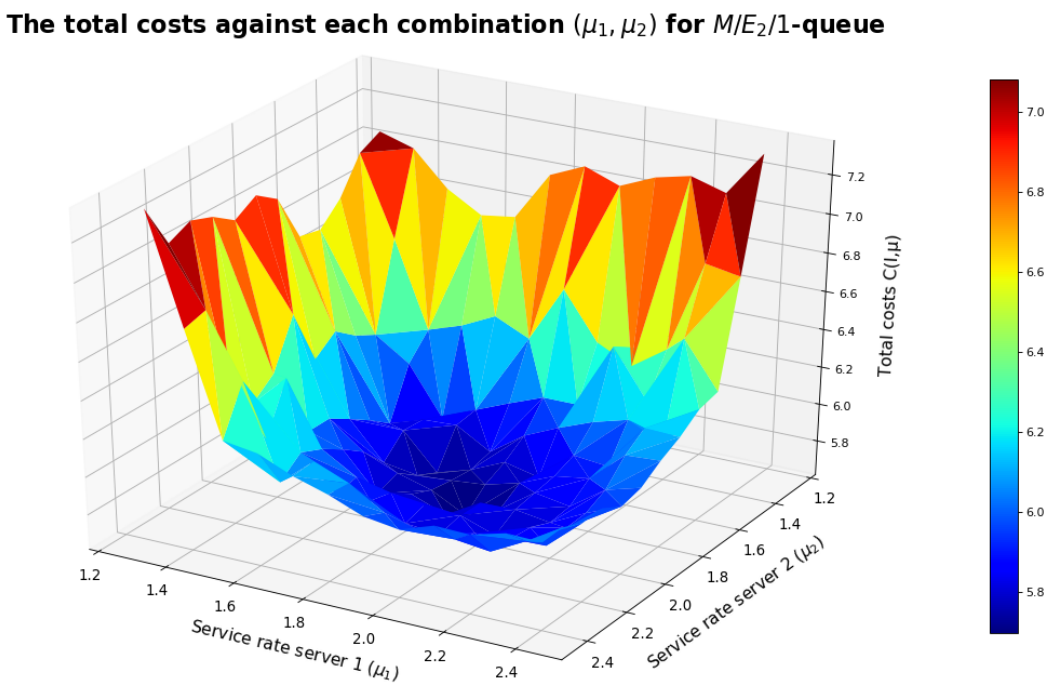

Figure 4.5: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/E_2/1$-queue for CASE 1.

In Figure 4.5 it can be seen that for the combinations of service rates (μ_1, μ_2) where one of the two service rates μ_i is small that the total costs are high. There is a rapid decrease of the total costs visible when either one of the service rates μ_i deviates from 1.4. Combinations with both service rates μ_i closer to a value of 1.9 have lower total costs. Another observation is that for the combinations of service rates that have both service rates μ_i higher than 1.9 the total costs increase again but only slightly.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.5.

		Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution		5.6500	1.9	0	1.9	0
Symmetric	Nash equilibrium	7.7240	1.3	0	1.3	0
	Theoretical value	8.6512	1.6	0	1.6	0
Symmetric with Punishment	Nash equilibrium	5.7610	1.7	0	1.7	0
Slowest server	Nash equilibrium	5.6500	1.9	0	1.9	0
Fastest server	Nash equilibrium	7.7240	1.3	0	1.3	0
Production costs	Nash equilibrium	7.7240	1.3	0	1.3	0
Holding costs	Nash equilibrium	7.7240	1.3	0	1.3	0

Table 4.5: Results $M/E_2/1$-queue for $c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total cost of 7.7240. In this case, the price of anarchy is the smallest out of all Nash equilibira. It must be noted that the service rate $\mu_i = 1.3$ is at the boundary of the service rates μ_i that have been considered in this simulation. This could mean that the actual Nash equilibrium is at a lower value of μ_i.

Comparing this result to the theoretical value for the Nash equilibrium (as computed using the upper and lower bounds in Section 3.7.2), shows that there is a slight difference between the two. The theoretical result for the Nash equilibrium of the symmetric choice is at a combination of service rates and inventory levels $(1.6, 0, 1.6, 0)$ with a total cost of 8.6512. The theoretical Nash equilibrium is obtained for a combination of service rates whose values are higher than those of the simulation result.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.7, 0, 1.7, 0)$ with a total cost of 5.7610. Again it can be seen that the Nash equilibrium has an inventory choice of 0 for both servers. In terms of the price of anarchy, this way of dividing the costs come close to the costs of the optimal solution, so the price of anarchy is small, but not yet equal to 1.

Looking at the results in Table 4.5 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium at the combination of service rates and inventory levels $(1.9, 0, 1.9, 0)$ with a total costs of 5.6500. This is in line with the results obtained for the optimal solution, which are also displayed in Table 4.5. In this case, the price of anarchy is equal to 1 and thus this way of dividing the costs among the servers is the best.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total costs of 7.7240. In these cases, the price of anarchy is the highest in comparison to the other Nash equilibria. Again, it has to be noted that the service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

4.2.2 CASE 2: $h_1 = 1$, $h_2 = 2$, $c_1 = 1$, $c_2 = 2$ and $b = 1$

Optimal strategy

In the second case, the holding cost parameters are set to be $h_1 = 1$ and $h_2 = 2$ and the production cost parameters are set to be $c_1 = 1$ and $c_2 = 2$. Lastly, the parameter for the backlog costs is set to be $b = 1$.

As stated earlier, there are two servers that each have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.2. It follows that the optimal combination for the theoretical lower bound is obtained for the combination of service rates and inventory levels $(1.5, 0, 1.5, 0)$ and for the theoretical upper bound is obtained for the combination of service rates and inventory levels $(2.1, 1, 1.7, 1)$.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(2.0, 0, 1.7, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 7.6750. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.2, it can be concluded that the optimal strategy resulting from the simulation is considerably close to the upper bound that is calculated for the total costs.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.6 below. The inventory levels that assure these minimal costs are left out of the plot.

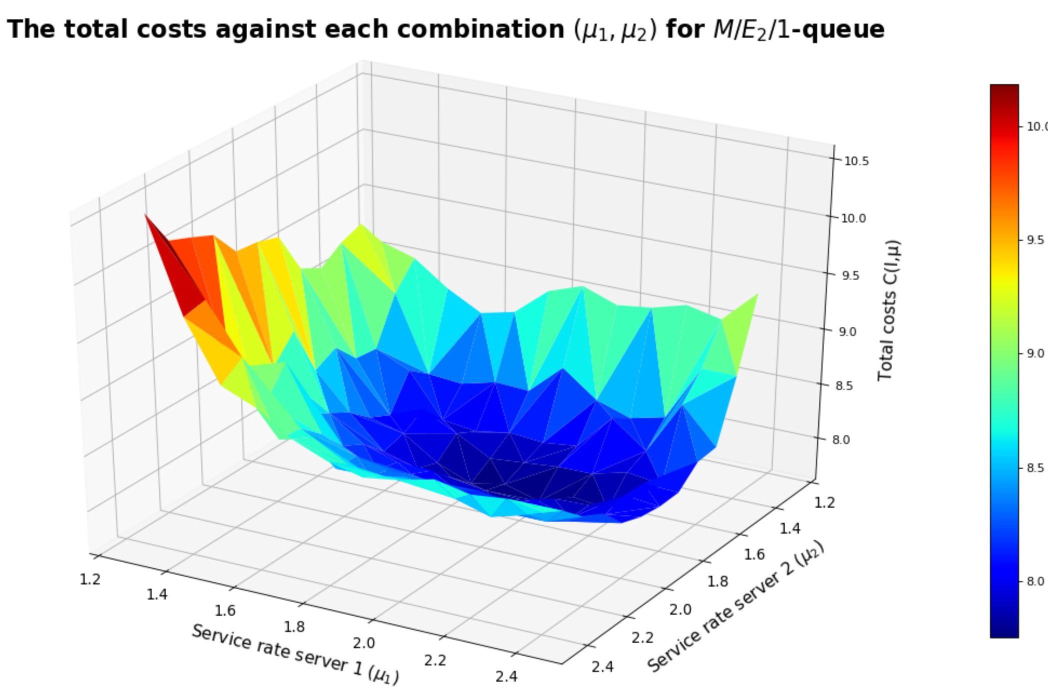

Figure 4.6: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/E_2/1$-queue for CASE 2.

In Figure 4.6 it can be seen that the plot is not symmetric, similarly to Plot 4.2 for the $M/M/1$-queue. The highest total costs are obtained for combinations of service rates (μ_1, μ_2), where μ_1 is equal to 1.4 and μ_2 are equal to 2.4. The optimal combination is clearly visible in the color dark blue. In addition, this figure clearly shows that the minimum is attained in the interior of the set of service rates that is considered.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.6.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	7.6750	2.0	0	1.7	0
Symmetric	9.8350	1.3	0	1.3	0
Symmetric with Punishment	7.9600	1.6	0	1.5	0
Slowest server	7.9600	1.6	0	1.5	0
Fastest server	9.8350	1.3	0	1.3	0
Production costs	9.8350	1.3	0	1.3	0
Holding costs	9.8350	1.3	0	1.3	0

Table 4.6: Results $M/E_2/1$-queue for $c_1 = 1, c_2 = 2, h_1 = 1, h_2 = 2, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total cost of 9.8350. It must be noted that the service rate $\mu_i = 1.3$ is at the boundary of the service rates that have been considered in this simulation. This could mean that the actual Nash equilibrium is at a lower value of μ_i.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.6, 0, 1.5, 0)$ with a total cost of 7.9600. Again it can be seen that the Nash equilibrium has an inventory choice of 0 for both servers. In this case, the service rates μ_i that both servers choose are different from one another, which can be clarified due to the difference in holding and production costs for both servers. Furthermore, this Nash equilibrium does not give the same result as the optimal solution as was seen for the similar case in the $M/M/1$-queue. In this case, the price of anarchy is the lowest in comparison to the other Nash equilibrium, but it is larger than 1.

Looking at the results in Table 4.6 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium for the combination of service rates and inventory levels $(1.6, 0, 1.5, 0)$ with a total costs of 7.9600. Again, this Nash equilibrium does not give the same result as the optimal solution as was seen for the similar case in the $M/M/1$-queue. In this case, the price of anarchy is the same as for the symmetric choice of cost distribution with punishment for the delay.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total costs of 9.8350. In these cases, the price of anarchy is the highest in comparison to the other Nash equilibria. Again, it has to be noted that the service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

4.2.3 CASE 3: $h_1 = 1$, $h_2 = 1$, $c_1 = 1$, $c_2 = 1$ and $b = 2$

Optimal strategy

In the third case, the holding and production cost parameters for both servers are set to be 1. Lastly, the parameter for the backlog costs for all servers is set to be $b = 2$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.2. It follows that the optimal total costs for the theoretical lower bound are obtained at the combination of service rates and inventory values $(1.7, 1, 1.7, 1)$ and for the theoretical upper bound are obtained at the combination of service rates and inventory levels $(2.1, 1, 2.1, 1)$. In this case, it can be seen that the upper and lower bounds both have inventory values that are larger than 1 for both servers.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.9, 1, 1.9, 1)$ is the optimal strategy for both servers and gives a minimum total costs equal to 6.3970. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.2, it can be concluded that the optimal strategy resulting from the simulation is between the upper and lower bounds that are calculated for the total costs.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.7 below. The inventory levels that assure these minimal costs are left out of the plot.

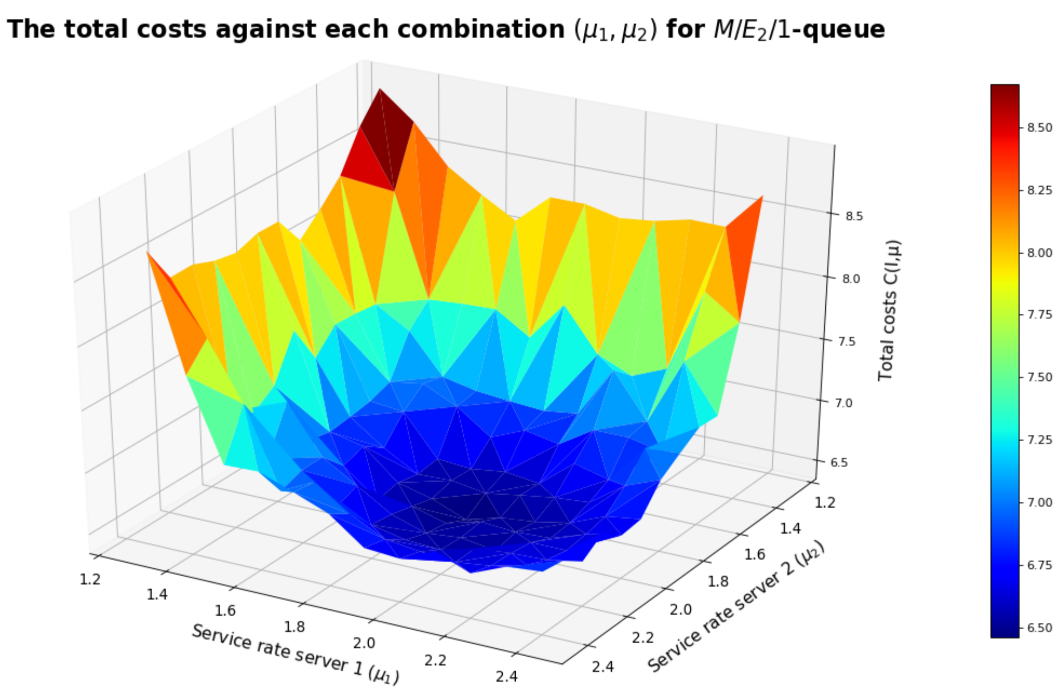

Figure 4.7: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/E_2/1$-queue for CASE 3.

In Figure 4.7 it can be seen that the plot is symmetric, similarly to Figure 4.5. The highest total costs are obtained for the combinations of service rates (μ_1, μ_2), where μ_1 and μ_2 are equal to 1.3. The optimal combination is clearly visible in the color dark blue. Furthermore, it can be seen that the plot has a very similar shape to the plot in Figure 4.5. In addition, the figure clearly shows that the minimum is attained in the interior of the set of service rates that are considered.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.7.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	6.3970	1.9	1	1.9	1
Symmetric	8.0170	1.5	0	1.6	0
Symmetric with Punishment	6.3970	1.9	1	1.9	1
Slowest server	6.3970	1.9	1	1.9	1
Fastest server	11.2070	1.3	0	1.3	0
Production costs	8.0170	1.5	0	1.6	0
Holding costs	8.0170	1.5	0	1.6	0

Table 4.7: Results $M/E_2/1$-queue for $c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 2$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory values $(1.5, 0, 1.6, 0)$ with a total cost of 8.0170. For this combination the service rates μ_i are different from one another. This does not coincide with the expectations based on the parameters in the cost function.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.9, 1, 1.9, 1)$ with a total cost of 6.3970. It can be seen that the inventory levels are none zero. Furthermore, this Nash equilibrium gives the same result as the optimal solution. Therefore, the price of anarchy in this case is equal to 1.

Looking at the results in Table 4.7 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium for the combination of service rates and inventory levels $(1.9, 1, 1.9, 1)$ with a total costs of 6.3970. Again, this Nash equilibrium gives the same result as the optimal solution. So in this case, the price of anarchy is equal to 1.

For the case where the backlog costs are assigned to the fastest server, the Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total costs of 11.2070. In this case, the price of anarchy is the highest in comparison to the other Nash equilibria. Again, it has to be noted that the service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the Nash equilibrium for this cost distribution is obtained for service rates $\mu_i \in (1, 1.3]$.

Lastly, it can be seen that the cost division strategies based on the holding and production costs have the same Nash equilibrium as the symmetric cost distribution. This can be clarified, since the holding and production cost parameters are the same for both servers, meaning that the costs are distributed symmetrically over the servers. So when the holding and production cost parameters are the same for both servers, the cost distribution functions that use the holding or productions costs are actually equal to the symmetric cost distribution.

4.2.4 CASE 4: $h_1 = 1$, $h_2 = 2$, $c_1 = 2$, $c_2 = 1$ and $b = 1$

In the fourth case, the holding cost parameters for each server are set to be $h_1 = 1$ and $h_2 = 2$ and the production cost parameters for each server are set to be $c_1 = 2$ and $c_2 = 1$. Lastly, the parameter for the backlog costs for all servers is set to be $b = 1$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.2. It follows that the optimal total costs for the theoretical lower bound is obtained for the combination of service rates and inventory levels $(1.4, 1, 1.6, 0)$ and for the theoretical upper bound is obtained for the combination of service rates and inventory levels $(1.7, 2, 2.3, 0)$. In the upper and lower bounds it is visible that the chosen inventory for server 1 is higher than server 2.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.6, 1, 1.8, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 7.7820. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.2, it can be concluded that the optimal strategy resulting from the simulation falls between the two theoretical bounds. Furthermore, it can be seen that the inventory for server 1 is not chosen to be 0 which is the same as for the upper and lower bounds.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.8 below. The inventory levels that assure these minimal costs are left out of the plot.

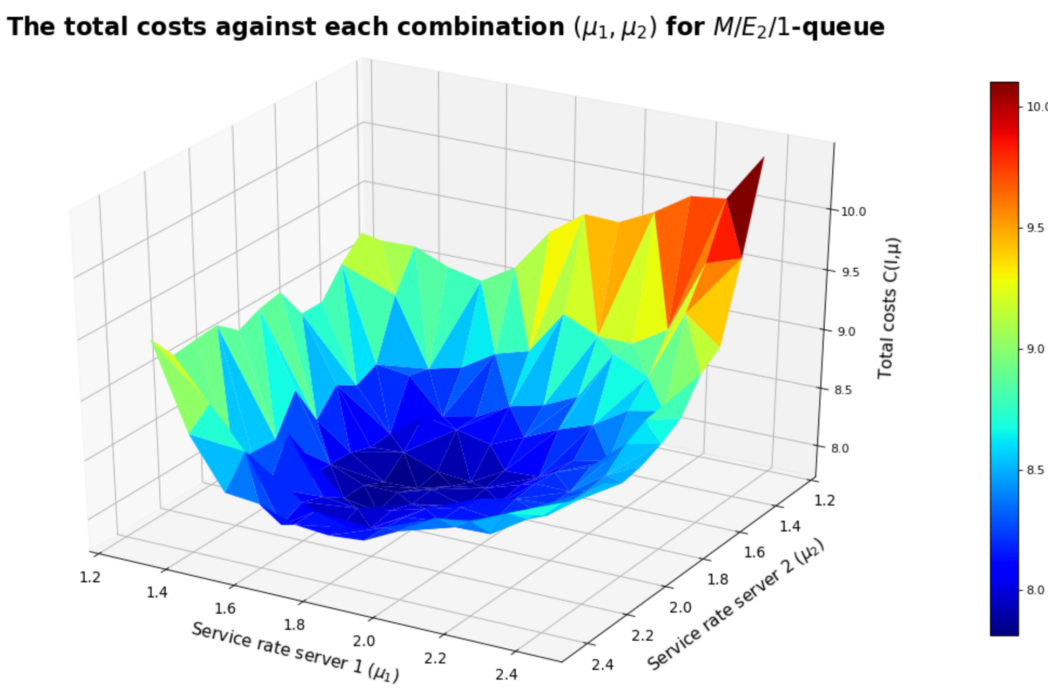

Figure 4.8: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $M/E_2/1$-queue for CASE 4.

In Figure 4.8 it can be seen that the plot is not symmetric, similarly to the plot in Figure 4.6. The highest total costs are obtained for combinations of service rates (μ_1, μ_2), where μ_1 is equal to

2.4 and μ_2 are equal to 1.3. The optimal combination is clearly visible in the color dark blue. In addition, this figure shows that the minimum is attained in the interior of the set of service rates that is considered.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.8.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	7.7820	1.6	1	1.8	0
Symmetric	9.8350	1.3	0	1.3	0
Symmetric with Punishment	8.4810	1.5	0	1.4	0
Slowest server	8.1510	1.5	0	1.6	0
Fastest server	9.8350	1.3	0	1.3	0
Production costs	9.8350	1.3	0	1.3	0
Holding costs	9.8350	1.3	0	1.3	0

Table 4.8: Results $M/E_2/1$-queue for $c_1 = 2, c_2 = 1, h_1 = 1, h_2 = 2, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total cost of 9.8350. It must be noted that the service rate $\mu_i = 1.3$ is at the boundary of the service rates that have been considered in this simulation. This could mean that the actual Nash equilibrium is obtained for a lower service rate μ_i.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay is, also computed using the simulation. This gives a Nash equilibrium at the combination of service rates and inventory levels $(1.5, 0, 1.4, 0)$ with a total cost of 8.4810. This Nash equilibrium does not give the same result as the optimal solution as seen before in the other cases. The price of anarchy is low, but not equal to 1. Furthermore, it can be seen that the service rates of both servers differ from one another.

Looking at the results in Table 4.8 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium for the combination of service rates and inventory levels $(1.5, 0, 1.6, 0)$ with a total cost of 8.1510. This Nash equilibrium does not give the same result as the optimal solution as seen before in the other cases. The price of anarchy is the lowest of all Nash equilibria, but not equal to 1. Both servers choose different service rates, this can be clarified by the difference in cost and production parameters between the two servers.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with a total costs of 9.8350. In these cases, the price of anarchy has the highest value of all Nash equilibria. Again, it has to be noted that the service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

4.3 $E_2/M/1$-queue

The second model that is considered is the model where the interarrival times follow an Erlang-2 distribution, whereas the service times are exponentially distributed. The queue of each individual server can then be considered as an $E_2/M/1$-queues that has the same arrival stream as the other servers.

In Section 3.7.1, the upper and lower bounds for the total costs are computed for each combination of cost parameters that is considered. The service rates that result in these minimum total costs are displayed in Table 3.7. From this table, it can be seen that the service rates for this type of queue range from 1.4 to 2.2. To determine if this is indeed the case, the values of the service rates μ_i in the simulation are ranged from 1.3 to 2.4 with steps of 0.1. The simulation is run for all combinations of service rates (μ_1, μ_2). A step size of 0.1 is chosen since taking an even smaller step size did not show more accuracy.

4.3.1 CASE 1: $h_1 = 1$, $h_2 = 1$, $c_1 = 1$, $c_2 = 1$ and $b = 1$

Optimal strategy

In the initial case, the holding, production and backlog cost parameters are set to be 1.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds that are calculated in Section 3.7.3 suggest that the minimum of the total costs for this combination of the parameters is obtained for the service rates μ_i between 1.6 and 1.9.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.8, 1, 1.8, 1)$ is the optimal strategy for both servers and gives a minimum total costs equal to 5.7440. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.3, it can be concluded that the optimal strategy resulting from the simulation is considerably close to the upper bound that is calculated for the total costs.

The difference between the two results is the choice for the service rates. The optimal service rate choice for the theoretical upper bound is equal to $(1.9, 1.9)$ whereas the optimal service rate choice from the simulation is $(1.8, 1.8)$.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.5 below. The inventory levels that assure this minimal cost are left out of the plot.

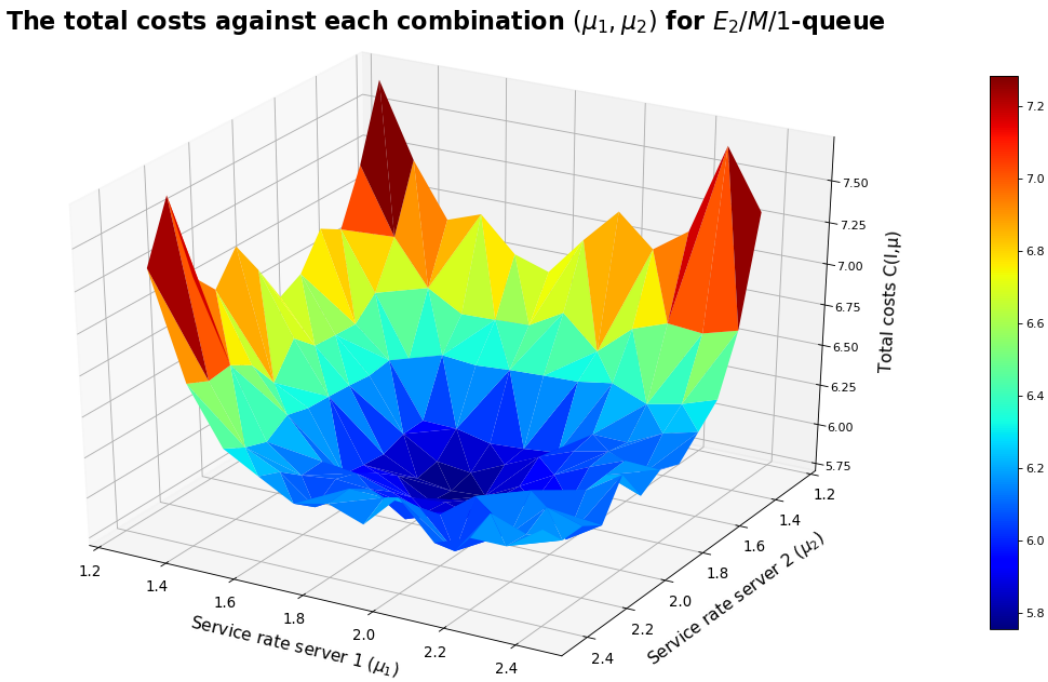

Figure 4.9: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $E_2/M/1$-queue for CASE 1.

In Figure 4.9 it can be seen that for the combinations of service rates (μ_1, μ_2) where one of the two service rates μ_i is small that the total costs are high. In comparison to the other two plots, it can be observed that there are two points for which $\mu_1 = 1.4$ and the total costs are extremely high. There is a dip visible between these two points, which had not been observed in the plot of the $M/M/1$-queue and the $M/E_2/1$-queue. There is a rapid decrease of the total costs visible when either one of the μ_i deviates from 1.4. Combinations with both service rates μ_i closer to a value of 1.9 have lower total costs. Another observation is that for the combinations of service rates μ_i that have both service rates μ_i higher than 1.9 the total costs increase again but only slightly. In addition, the figure clearly shows that the optimal total costs are attained in the interior of the service rates that are considered.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.9.

		Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution		5.7440	1.8	1	1.8	1
Symmetric	Nash equilibrium	8.0030	1.3	0	1.4	0
	Theoretical value	7.5796	1.6	0	1.6	0
Symmetric with Punishment	Nash equilibrium	6.0740	1.7	0	1.7	0
Slowest server	Nash equilibrium	6.0740	1.7	0	1.7	0
Fastest server	Nash equilibrium	8.0030	1.3	0	1.4	0
Production costs	Nash equilibrium	8.0030	1.3	0	1.4	0
Holding costs	Nash equilibrium	8.0030	1.3	0	1.4	0

Table 4.9: Results $E_2/M/1$-queue for $c_1 = 1, c_2 = 1, h_1 = 1, h_1 = 1, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total cost of 8.0030. It must be noted that the service rate $\mu_i = 1.3$ is at the boundary of the service rates that have been considered in this simulation. This could mean that the actual Nash equilibrium is obtained for a lower service rate μ_i.

Comparing this result to the theoretical value for the Nash equilibrium (as computed in Section 3.7), shows that there is a slight difference between the two. The theoretical result for the Nash equilibrium of the symmetric choice is obtained for a combination of service rates and inventory levels $(1.6, 0, 1.6, 0)$ with a total cost of 7.5796. The theoretical Nash equilibrium is obtained for a combination of service rates whose values are higher than those of the simulation result.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.7, 0, 1.7, 0)$ with a total cost of 6.0740. Again it can be seen that the Nash equilibrium has an inventory choice of 0 for both servers.

Looking at the results in Table 4.9 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium at the combination of service rates and inventory levels $(1.7, 0, 1.7, 0)$ with a total costs of 6.0740. This is the same Nash equilibrium as for the symmetric cost distribution with punishment. In this case and in the case of the symmetric cost distribution with punishment, the price of anarchy is the lowest of all Nash equilibria, but it is not equal to 1.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total costs of 8.0030. in these cases, the price of anarchy is the highest in comparison to the other Nash equilibria. Again, it has to be noted that the service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

4.3.2 CASE 2: $h_1 = 1$, $h_2 = 2$, $c_1 = 1$, $c_2 = 2$ and $b = 1$

Optimal strategy

In the second case, the holding cost parameters are set to be $h_1 = 1$ and $h_2 = 2$ and the production cost parameters are set to be $c_1 = 1$ and $c_2 = 2$. Lastly, the parameter for the backlog costs for all servers is set to be $b = 1$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.3. It follows that the optimal total costs for the theoretical lower bound are obtained for the combination of service rates and inventory values $(1.5, 0, 1.5, 0)$ and for the theoretical upper bound are obtained for the combination of service rates and inventory levels $(2.0, 1, 1.9, 0)$.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(2.0, 0, 1.6, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 7.9620. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.3, it can be concluded that the optimal strategy resulting from the simulation is considerably close to the upper bound that was calculated for the total costs. Furthermore, it can be noted that this result is similar to the one in case 2 for the $M/E_2/1$-queue.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.10 below. The inventory levels that assure these minimal costs are left out of the plot.

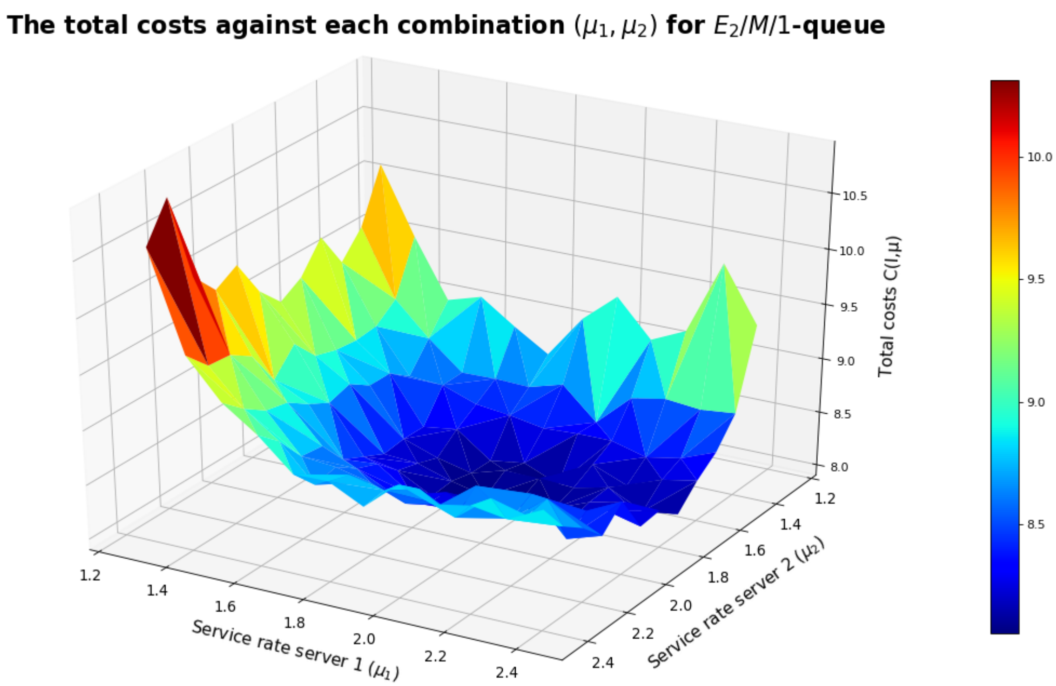

Figure 4.10: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $E_2/M/1$-queue for CASE 2.

In Figure 4.10 it can be seen that the plot is not symmetric, similarly to the plot Figure 4.2. The highest total costs are obtained for combinations of service rates (μ_1, μ_2), where μ_1 is equal to 1.3 and μ_2 are equal to 2.4. The optimal combination is clearly visible in the color dark blue. In addition, the figure clearly shows that the minimum is attained in the interior of the set of service rates that is considered.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.6.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	7.9620	2.0	0	1.6	0
Symmetric	10.6760	1.3	0	1.4	0
Symmetric with Punishment	8.8260	1.6	0	1.4	0
Slowest server	8.8250	1.5	0	1.5	0
Fastest server	11.2210	1.3	0	1.3	0
Production costs	10.6760	1.3	0	1.4	0
Holding costs	10.6760	1.3	0	1.4	0

Table 4.10: Results $E_2/M/1$-queue for $c_1 = 1, c_2 = 2, h_1 = 1, h_2 = 2, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total cost of 10.6760. It must be noted that the service rate $\mu_i = 1.3$ is at the boundary of the service rates that have been considered in this simulation. This could mean that the actual Nash equilibrium is obtained for a lower value of μ_i.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.6, 0, 1.4, 0)$ with a total cost of 8.8260. Again it can be seen that the Nash equilibrium has an inventory choice of 0 for both servers. In this case, the service rates μ_i that both servers choose are different from one another, which can be clarified due to the difference in holding and production costs for both servers. Furthermore, this Nash equilibrium does not give the same result as the optimal solution as was seen for the similar case in the $M/M/1$-queue. In this case, the price of anarchy is low in comparison to the other Nash equilibria but not equal to 1.

Looking at the results in Table 4.10 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium for the combination of service rates and inventory levels $(1.5, 0, 1.5, 0)$ with a total costs of 8.8250. Again, this Nash equilibrium does not give the same result as the optimal solution as was seen for the similar case in the $M/M/1$-queue. In this case, the price of anarchy is low in comparison to the other Nash equilibria but not equal to 1.

Assigning the backlog costs to the fastest server results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.3, 0)$ with total costs equal to 11.2210. The last two choices for the cost division among the servers have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total costs of 10.6760. Again, it has to be noted that the service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

4.3.3 CASE 3: $h_1 = 1$, $h_2 = 1$, $c_1 = 1$, $c_2 = 1$ and $b = 2$

Optimal strategy

In the third case, the holding and production cost parameters for both servers are set to be 1. Lastly, the parameter for the backlog costs for all servers is set to be $b = 2$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.3. It follows that the optimal total costs for the theoretical lower bound are obtained for the combination of service rates and inventory levels $(1.9, 0, 1.9, 0)$ and for the theoretical upper bound are obtained for the combination of service rates and inventory levels $(2.0, 1, 2.0, 1)$. In this case, it can be seen that the upper bound has inventory levels larger than 1 for both servers.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(2.0, 1, 2.0, 1)$ is the optimal strategy for both servers and gives a minimum total costs equal to 6.3470. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.3, it can be concluded that the optimal strategy resulting from the simulation is considerably close to the theoretical upper bound.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.11 below. The inventory levels that assure these minimal costs are left out of the plot.

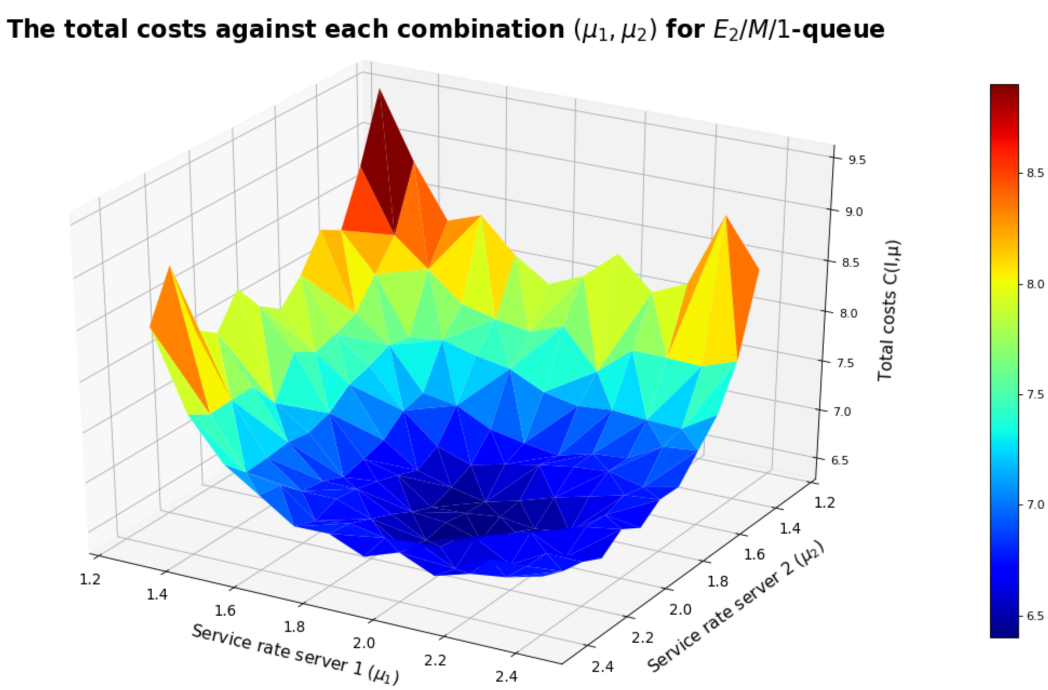

Figure 4.11: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $E_2/M/1$-queue for CASE 3.

In Figure 4.11 it can be seen that the plot is symmetric, similarly to the plot in Figure 4.9. The highest total costs are obtained for the combinations of service rates (μ_1, μ_2), where μ_1 and μ_2 are equal to 1.3. The optimal combination is clearly visible in the color dark blue. Furthermore, it can be seen that the plot has a very similar shape to the plot in Figure 4.9. In addition, the figure clearly shows that the minimum is attained in the interior of the set of service rates that are considered.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.11.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	6.3470	2.0	1	2.0	1
Symmetric	8.1490	1.6	0	1.6	0
Symmetric with Punishment	6.3950	1.8	1	1.9	1
Slowest server	6.3950	1.9	1	1.8	1
Fastest server	11.2180	1.3	0	1.4	0
Production costs	8.1490	1.6	0	1.6	0
Holding costs	8.1490	1.6	0	1.6	0

Table 4.11: Results $E_2/M/1$-queue for $c_1 = 1, c_2 = 1, h_1 = 1, h_2 = 1, b = 2$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.6, 0, 1.6, 0)$ with a total cost of 8.1490. The values for the service rates in this Nash equilibrium are much smaller than that of the optimal solution. In this case, the price of anarchy is the highest out of all Nash equilibria.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. For this cost distribution the Nash equilibrium is obtained for a combination of service rates μ_i and inventory levels I_i of $(1.8, 1, 1.9, 1)$ with a total cost of 6.3950. It can be seen that the inventory levels are none zero. Furthermore, this Nash equilibrium is quite close to the optimal solution. In this case, the price of anarchy is the lowest of all Nash equilibria, but it is not equal to 1. The service rates for both servers differ from one another. This is in contrary to the expected results, since the production and holding cost parameters are equal for both servers.

Looking at the results in Table 4.11 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium at the combination of service rates and inventory levels $(1.9, 1, 1.8, 1)$ with a total costs of 6.3950. Again, this Nash equilibrium is quite close to the optimal solution. In this case, the price of anarchy is the smallest of all Nash equilibria, but it is not equal to 1. The service rates for both servers differ from one another. This is in contrary to the expected results, since the production and holding cost parameters are equal for both servers.

For the case where the backlog costs are assigned to the fastest server, the Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total costs of 11.2180. In these cases, the price of anarchy is the high, but still smaller than in for the symmetric cost distribution where the service rates are higher. Again, it has to be noted that the service rates are at the boundary of all values that were considered in this simulation. Thus, it is likely that the Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

Lastly, it can be seen that the cost division strategies based on the holding and production costs have the same Nash equilibrium as the symmetric cost distribution. This can be clarified, since the holding and production cost parameters are the same for both servers, meaning that the costs are distributed symmetrically over the servers. So when the holding and production cost parameters are the same for both servers, the cost distribution functions that use the holding or productions costs are actually equal to the symmetric cost distribution.

4.3.4 CASE 4: $h_1 = 1$, $h_2 = 2$, $c_1 = 2$, $c_2 = 1$ and $b = 1$

In the fourth case, the holding cost parameters are set to be $h_1 = 1$ and $h_2 = 2$ and the production cost parameters are set to be $c_1 = 2$ and $c_2 = 1$. Lastly, the parameter for the backlog costs for all servers is set to be $b = 1$.

As stated earlier, there are two servers that both have their own service rate μ_i. The theoretical upper and lower bounds can be calculated using the formulas in Section 3.7.3. It follows that the optimal total costs for the theoretical lower bound are obtained for the combination of service rates and inventory levels $(1.4, 1, 1.6, 0)$ and for the theoretical upper bound are obtained for the combination of service rates and inventory levels $(1.8, 1, 2.2, 0)$. In the upper and lower bounds it is visible that the chosen inventory for server 1 is higher than server 2.

Using the simulation for this model it is determined that the combination of service rates and inventory levels $(1.8, 1, 1.9, 0)$ is the optimal strategy for both servers and gives a minimum total costs equal to 7.9760. Comparing the results of the simulation to the upper and lower bounds that are computed in Section 3.7.3, it can be concluded that the optimal strategy resulting from the simulation falls between the two theoretical bounds. Furthermore, it can be seen that the inventory level for server 1 is not chosen to be 0 which is the same as for the upper and lower bounds.

A plot is made of all combinations of service rates (μ_1, μ_2) and their minimal costs. This plot is shown in Figure 4.12 below. The inventory levels that assure these minimal costs are left out of the plot.

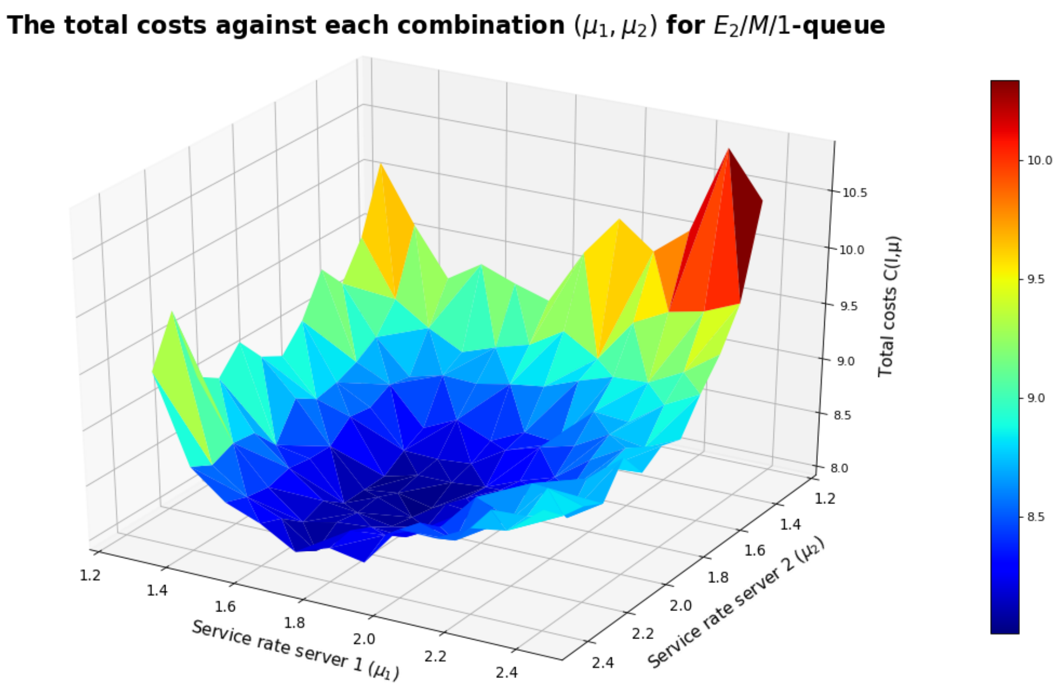

Figure 4.12: Plot of all combinations of service rates (μ_1, μ_2) and their minimal total costs for the $E_2/M/1$-queue for CASE 4.

In Figure 4.12 it can be seen that the plot is not symmetric, similarly to the plot in Figure 4.10. The highest total costs are obtained for combinations of service rates (μ_1, μ_2), where μ_1 is equal to 2.4 and μ_2 are equal to 1.3. The optimal combination is clearly visible in the color dark blue.

Nash equilibria

The simulation is also used to determine the Nash equilibria for different cost functions as described in Section 3.2. The results for the different ways of dividing costs among servers are shown in Table 4.12.

	Total costs	μ_1	I_1	μ_2	I_2
Optimal Solution	7.9760	1.8	1	1.9	0
Symmetric	10.5760	1.3	0	1.4	0
Symmetric with Punishment	8.4890	1.6	0	1.6	0
Slowest server	8.4610	1.6	0	1.7	0
Fastest server	10.5760	1.3	0	1.4	0
Production costs	10.5760	1.3	0	1.4	0
Holding costs	10.5760	1.3	0	1.4	0

Table 4.12: Results $E_2/M/1$-queue for $c_1 = 2, c_2 = 1, h_1 = 1, h_2 = 2, b = 1$.

When looking at the results, it can be seen that the symmetric choice of cost division results in a Nash equilibrium for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total cost of 10.5760. It must be noted that the service rate $\mu_i = 1.3$ is at the boundary of the service rates that have been considered in this simulation. This could mean that the actual Nash equilibrium is obtained for a lower value of μ_i. In this case, the price of anarchy is the highest of all Nash equilibria.

The Nash equilibrium for a symmetric choice of cost distribution, taking into account that servers are punished for delay, is also computed using the simulation. This gives a Nash equilibrium for the combination of service rates and inventory levels $(1.6, 0, 1.6, 0)$ with a total cost of 8.4890. This Nash equilibrium does not give the same result as the optimal solution as seen before in the other cases. In this case, the price of anarchy is small in comparison to the other Nash equilibria, but not equal to 1, as seen before in similar cases. Furthermore, it can be seen that the service rates of both servers are equal.

Looking at the results in Table 4.12 it can be seen that assigning the backlog costs to the slowest server gives a Nash equilibrium for the combination of service rates and inventory levels $(1.6, 0, 1.7, 0)$ with a total cost of 8.4610. This Nash equilibrium does not give the same result as the optimal solution as seen before in the other cases. In this case, the price of anarchy is the smallest out of all Nash equilibrua, but not equal to 1 as was seen in similar cases. Both servers choose different service rates, this can be clarified by the difference in cost and production parameters between the two servers.

The last three choices for the cost division among the servers all have the same Nash equilibrium. This Nash equilibrium is obtained for the combination of service rates and inventory levels $(1.3, 0, 1.4, 0)$ with a total costs of 10.5760. In these cases, the price of anarchy is the smallest out of all Nash equilibria. Again, it has to be noted that service rates are at the boundary of all values that are considered in this simulation. Thus, it is likely that the actual Nash equilibria for these cost divisions are obtained for service rates $\mu_i \in (1, 1.3]$.

4.4 $M/M/1$ vs. $M/E_2/1$ vs. $E_2/M/1$

This section provides an overview of the most important results mentioned in the previous sections. The goal of this section is to compare the results for the optimal total costs of the different types of queues that are considered.

In Table 4.13 the combinations of service rates and inventory levels that result in optimal total costs are displayed for all four cases and all three models.

	CASE 1		CASE 2	
	(μ_1, I_1, μ_2, I_2)	$C(\boldsymbol{I}, \boldsymbol{\mu})$	(μ_1, I_1, μ_2, I_2)	$C(\boldsymbol{I}, \boldsymbol{\mu})$
$M/M/1$	$(2.1, 0, 2.1, 0)$	6.1170	$(1.9, 0, 1.7, 0)$	8.3420
$M/E_2/1$	$(1.9, 0, 1.9, 0)$	5.6500	$(2.0, 0, 1.7, 0)$	7.6750
$E_2/M/1$	$(1.8, 1, 1.8, 1)$	5.7440	$(2.0, 0, 1.6, 0)$	7.9620

	CASE 3		CASE 4	
	(μ_1, I_1, μ_2, I_2)	$C(\boldsymbol{I}, \boldsymbol{\mu})$	(μ_1, I_1, μ_2, I_2)	$C(\boldsymbol{I}, \boldsymbol{\mu})$
$M/M/1$	$(2.1, 1, 2.1, 1)$	6.9070	$(1.9, 1, 2.2, 0)$	8.5030
$M/E_2/1$	$(1.9, 1, 1.9, 1)$	6.3970	$(1.6, 1, 1.8, 0)$	7.7820
$E_2/M/1$	$(2.0, 1, 2.0, 1)$	6.3470	$(1.8, 1, 1.9, 0)$	7.9760

Table 4.13: The optimal total costs and corresponding combination of service rates and inventory levels for $M/M/1$ vs. $M/E_2/1$ vs. $E_2/M/1$ determined using the simulation.

In Table 4.13 it is clearly visible that the total costs for the $M/M/1$-queue are the highest for all cases. In addition, the service rates μ_i that assure these minimal costs are the highest for almost all cases for the $M/M/1$-queue.

In addition, it can be seen that the total costs for the $M/E_2/1$-queue are the lowest in almost all cases. This is different to what could be seen from the theoretical upper and lower bounds, where the total costs for the $E_2/M/1$-queue are the lowest. A similar result can be seen for the service rates that results in these optimal total costs. They are the lowest for the $M/E_2/1$-queue and the highest for the $M/M/1$-queue.

Chapter 5

Conclusion

In this report, a two-server fork-join queue is studied. A newsvendor game is analyzed using theoretical upper and lower bounds for the total costs and using a simulation study. Furthermore, Nash equilibria on different types of cost distributions have been studied. These Nash equilibria are compared to the optimal solutions in order to find a Nash equilibrium that has total costs close to that of the optimal solution. Based on the results of both the theoretical analysis and the simulation study, several conclusions can be made. This section briefly reflects on the results in Chapter 4 and summarizes the most important conclusions that can be drawn out of this research.

In this report, three cases are considered with different distribution choices for the interarrival times and the service times. From the theoretical upper and lower bounds in Section 3.7 it can already be observed that if either the interarrival time or the service time is set to be Erlang-2 distributed, the total costs for the same combination of cost parameters decreases as well as the service rates of both servers. This can be clarified, since the mean for both distributions is kept equal, but the variance for the Erlang-2 distribution becomes smaller. The same can be concluded from the results in the simulation. Overall it can be concluded that if the variance of the distribution of either the interarrival times or the service times decreases, so does the choice of μ_i and the total costs.

From the results in Chapter 4 several other conclusions can also be made. Firstly, it is noticeable that when the production, holding and backlog cost parameters (c_i, h_i and b) are set to be equal to 1 the optimal costs are obtained for a combination of service rates (μ_1, μ_2), where $\mu_1 = \mu_2$ and inventory levels (I_1, I_2) where $I_1 = I_2$. The same holds for the Nash equilibria. This can be clarified, since the holding cost and production cost parameters are equal for both servers, thus the cost function $C(\boldsymbol{I}, \boldsymbol{\mu})$ is symmetric with respect to the variables $\boldsymbol{I}$ and $\boldsymbol{\mu}$. In addition, when the holding cost and production cost parameters for both servers are not kept equal, it can be seen that both servers choose a service rate μ_i and an inventory value I_i that is distinct to the other server.

In case 2, where the holding and production cost parameter for server 2 is increased, it can be noticed that the service rate for server 2 decreases and that the inventory levels remain 0. This can be clarified, since an increase in holding costs means that any product in the warehouse will cost the server more money, therefore having zero inventory is cheaper. Since the production costs are also increased, the server chooses to work less hard, since producing more also costs him more money than for server 1.

In case 4, where the holding costs for server 2 are higher than for server 1 and the production costs for server 1 are higher than for server 2, it can be seen that server 1 chooses to work slower, so his service rate decreases in comparison to case 1, and in response server 1 chooses an inventory level of 1. This can be clarified, since the holding cost parameter for server 1 is low but the production cost parameter is high. The opposite can be noted for server 2. Server 2 chooses a higher service rate, since his production cost parameter is low, but chooses an inventory level of 0, since his holding cost parameter is high.

In case 3 of all three models, all parameters are set equal to 1 except for the backlog cost parameter $b = 2$. In this case, it can be seen that the optimal combination has the same service rates μ_i as that of case 1, but the inventory values I_i for both servers increase symmetrically. Therefore, it can be concluded that increasing the backlog costs is paired with an increase in the choice of inventory values I_i and naturally an increase of the total costs, but not an increase of the service rates μ_i.

Furthermore, for case 3 it could be seen that the symmetric cost distribution function as well as the cost distribution functions based on the holding or production cost parameters, have a Nash equilibrium that comes closer in terms of service rates and total costs to the optimal solution than in any of the other cases. In all the other three cases, these cost distribution functions resulted in a Nash equilibrium at the bounds of the service rates that are considered.

In the cases 1 an 3 that are considered, the holding and production cost parameters are equal for both servers. Therefore, in these cases the Nash equilibria for the symmetric cost distribution function and for the cost distribution functions based on the holding or production cost parameters have the same outcome. This can be clarified, because of the fact that the cost distribution functions based on the holding or production cost parameters in these specific cases are equal to the symmetric cost distribution.

Another conclusion that can be drawn based on the results, is that the cost distribution where the backlog costs are assigned to the slowest server results in a Nash equilibrium that is equal to the optimal solution for the first three cases for the $M/M/1$-queue. It can be seen that for the $M/E_2/1$-queue and the $E_2/M/1$-queue that the price of anarchy is still the smallest among all the cost distributions.

Lastly, it can be seen that for the $M/E_2/1$-queue and the $E_2/M/1$-queue the cost distribution that distributes the costs symmetrically among the servers taking into account the amount of delay that each server has, often has the same Nash equilibrium as that in the case where all backlog costs are assigned to the slowest server.

Chapter 6

Discussion

In this chapter, a critical reflection is made on the approach as presented throughout this report. Furthermore, the main findings of the results in Chapter 4 are discussed and based on this, suggestions for further research are presented.

One of the bottlenecks for the simulations is the simulation time. Using the formula as described in Section 3.6.1, it can be seen that the simulation time increases as the value for the service rate μ_i decreases. This is due to the fact that the service rate μ_i gets closer to the value of the arrival rate λ and thus it takes longer before the queue reaches its steady state. Using the upper and lower bounds that are computed in Section 3.7, a selection of possible service rates has been made to shrink the size of possible combinations that need to be evaluated. It can be seen that when the parameters h_i and c_i are chosen such that $h_1 \neq h_2$ and $c_1 \neq c_2$, the upper and lower bounds are found for smaller values of the service rates μ_i. The simulation time for these smaller values of the service rates $\mu_i < 1.4$ increases exponentially and therefore running a simulation for these combinations was left out. To be able to generate results for service rates below 1.4, other methods are needed.

The equations that are used to compute the simulation time in Section 3.6.1 apply to the $M/M/1$-queue. Throughout this report, it is assumed that the simulation time as denoted in Section 3.6.1 also applies to the $E_2/M/1$-queue and $M/E_2/1$-queue. This assumption has been made in order to compare the results of the $E_2/M/1$-queue and $M/E_2/1$-queue to the $M/M/1$-queue. For further research, it is suggested to find an approximation for the time it takes to reach steady state for the $E_2/M/1$-queue and $M/E_2/1$-queue.

A similar assumption has been made concerning the inventory levels that are considered in the simulation. These inventory levels are based on the $M/M/1$-queue. For the same reasons as before, these inventory levels have also been used for the $E_2/M/1$-queue and $M/E_2/1$-queue. To obtain a more accurate approximation of the inventory levels of the $E_2/M/1$-queue and $M/E_2/1$-queue that need to be considered, it is wise to determine these using the distributions of the number of jobs in the system for both queues.

In Section 3.7, the assumption has been made for the $E_2/M/1$-queue that $\mathbb{P}(Q_i(\mu_i) = n) = \mathbb{P}(L^a = n)$. Here $\mathbb{P}(L^a = n)$ represents the probability that there are n jobs in the queue just before arrival. This is not exactly the same as the definition for $\mathbb{P}(Q_i(\mu_i) = n)$. This assumption is made because $\mathbb{P}(L^a = n)$ follows a geometric distribution with respect to σ which makes further computations a lot easier. In theory the difference between $\mathbb{P}(Q_i(\mu_i) = n)$ and $\mathbb{P}(L^a = n)$ is not that big and therefore the assumption can be backed. For further research, this is a point that could be looked into in more detail.

Throughout the report, the computation of the optimal costs for each combination of service rates (μ_1, μ_2) is based on the empirical mean of 1000 simulation runs. The choice for 1000 simulation runs was based on the computation power and the number of combinations of service rates (μ_1, μ_2) that needed to be done. In order to get an even more accurate result, more simulation runs can be done.

In this report, two additional models to the $M/M/1$-queue are considered, namely the $E_2/M/1$-queue and $M/E_2/1$-queue. The Erlang distribution with shape parameter r and rate λr is in distribution the same as the average of r independent exponentially distributed random variables with mean $1/\lambda$. Beforehand, the shape parameter r as denoted in Sections 3.5.2 and 3.5.3 is chosen to be 2. From Chapter 4, it can be seen that the optimal total costs with the corresponding combination of service rates (μ_1, μ_2) decreases when choosing an Erlang-2 distribution for either the interarrival times or the service times.

The expectation is that these total costs with corresponding combination of service rates (μ_1, μ_2) will decrease even further when the shape parameter r increases. This is due to the fact that the variance of the Erlang-r distribution decreases when r increases, keeping in mind that the mean of the Erlang-r distribution is the same as for the exponential distribution. Using the Law of large numbers, it can be said the if $r \to \infty$ the systems $M/E_r/1$ and $E_r/M/1$ converge to $M/D/1$ and $D/M/1$ systems respectively, where D stands for a deterministic distribution. This is because the variance $1/(r\lambda^2)$ of the Erlang-r distribution converges to 0 as $r \to \infty$, as is the case for the deterministic distribution where the variance is equal to 0. In this case, the future can be predicted better, which means that the inventory choice becomes higher and in response, the service rates become lower. A recommendation for further research would be to simulate the $E_r/M/1$-queue and $M/E_r/1$-queue for larger values of r to check whether the conjecture as denoted above is indeed true.

The upper and lower bounds suggest that the optimal solutions that are found using the simulation are the global optimal solutions. It may be possible that there are also more local optimal solutions beyond the service rates that are considered in this report. It would therefore be interesting to investigate the total costs for service rates $\mu_i > 2.4$. It must also be noted that the Nash equilibria that are found in this research are not necessarily the only Nash equilibria. This could be investigated in further research.

Lastly, this report only addresses two different types of distributions for the arrivals and the services. To be able to implement this simulation in real-life cases, the following is suggested. If data for the number of arrivals for orders and the service times of each supplier are available, it would be wise to make plots of this data in order to find the actual distribution for the interarrival times and service times. In this way, the model can be applied in real-life situations, where in many cases, the interarrival times and service times are not necessarily exponentially or Erlang distributed.

Acknowledgments

I would like to thank Dennis Schol for supervising me during this bachelor project. Inspiration for the simulation of the fork-join queue is obtained from the course Stochastic Simulation by dr. ir. M.A.A. Boon [15].

Bibliography

[1] Flatto, L., Hahn, S. (1984). Two Parallel Queues Created by Arrivals with Two Demands I. *Siam Journal on Applied Mathematics*, 44(5):13–13.

[2] Baccelli, F., Makowski, A., Shwartz, A. (1989). The Fork-Join Queue and Related Systems with Synchronization Constraints: Stochastic Ordering and Computable Bounds. *Advances in Applied Probability*, 21(3):629-660.

[3] Kim, C., Agrawala, A. K. (1989). Analysis of the Fork-Join Queue. *IEEE Transactions on Computers*, 38(2):250-255

[4] Von Neumann, J., Morgenstern, O., Rubinstein, A. (1944). *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)*. Oxford: Princeton University Press, Princeton.

[5] Nash, J. F. (1950). Equilibrium Points in n-Person Games. *Proceedings of the National Academy of Sciences*, 36(1):48–49.

[6] Ross, D. (2019). Game Theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy (Winter 2019 Edition)*, Metaphysics Research Lab, Stanford University, Winter 2019 edition.

[7] Kuhn, S. (2019). Prisoner's Dilemma. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philisophy*. Metaphysics Research Lab, Stanford University, Winter 2019 edition.

[8] Koutsoupias, E., Papadimitriou, C. (2009). Worst-case equilibria. *Computer Science Review*, 3(2):65–69.

[9] Dubey, P. (1986). Inefficiency of nash equilibria. *Mathematics of Operations Research*, 11(1):1–8.

[10] Chen, R. R., Cheng, T. C. E., Choi, T. M., Wang, Y. (2016). Novel advances in applications of the newsvendor model. *Decision Sciences*, 47(1):8–10.

[11] Schol, C., Vlasiou, M., Zwart, B. (2019). Large fork-join networks with nearly deterministic service times. ArXiv.

[12] Snyder, L. V., Shen, Z.-J. M. (2019). *Fundamentals of supply chain theory (Second)*. John Wiley & Sons, 2nd edition.

[13] Adan, I., Resing, J. (2015). Queueing systems [Lecture notes]. Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, The Netherlands.

[14] Robert, P. (2003). *Stochastic networks and queues* (Ser. Applications of mathematics, stochastic modelling and applied probability, 52). Springer Berlin Heidelberg. 1st edition.

[15] Boon, M. A. A., Van der Boor, M., Van Leeuwarden, J., Mathijsen, B., Van der Pol, J., Resing, J. (2020) Stochastic Simulation using Python [Lecture notes]. Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, The Netherlands.

[16] Meijer, M., Schol, C., Van Jaarsveld, W., Vlasiou, M., Zwart, B. (2020). Extreme value theory for dependent brownian motions, with an application to high-tech supply chains