The Impact of Information on the Performance of an M/M/1 Queueing System

by

Mojgan Nasiri

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2017

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Reviews provided by previous customers contain information, which can be used by new customers. This research examines the impact of the user-generated reviews, on the performance of an M/M/1 queueing system. We assume that customers do not know the expected service time and they obtain this information by reading reviews. The results show that reading unbiased reviews can result in either a better or worse performance, depending on the parameters of the system. We also investigate the impact of the number of reviews each customer reads, on the different performance measures. We observe that if each customer reads more reviews, it does not necessarily result in a system which is more similar to a system with full information. Moreover, even with a huge pool of reviews, it may either not converge to the system with full information or converges very slowly. Finally, we show that if reviews consist of the waiting time that customers experience in the system along with the number of people that they observe upon their arrival, the rate of convergence to the system with full information is much faster.

Acknowledgements

First, I would like to express my extreme sincere gratitude to my supervisor Dr. Hossein Abouee Mahrizi for his helpful comments, bright ideas and two years of constant support on this master thesis. Although he always gave me the opportunity to work independently, he supervised every step I took and mentored my thesis to make sure I am on the right path. Furthermore, I would like to thank Ata G. Zare for his useful discussions on my mathematical model and my friend Forough Pourhossein who is the reason I started this research.

Second, I like to thank Dr. Selcuk Onay and Dr. Bon Koo, who accepted to be my thesis readers and help me to improve it.

Finally, my deep and sincere gratitude to my family especially my lovely sister Madine, and her family for their continuous and unparalleled love, help, and support. I also express my thanks to my dearest friends Nasibe, Elham, Samira, Sulmaz, Ghazal, Nastaran, and, Zoha who mean more than a family to me even though they are far away, and also Nazli and Nikpai's family who never let me feel alone.

Dedication

To my Dad.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Reviews provided by previous customers, contain information, which can be used by new customers. This research examines the impact of the user-generated reviews, on the performance of an M/M/1 queueing system, specifically, congestion. Customer reviews play an undeniable role in the sale of a product or service, even if the related information and specifications have been published in details. According to a survey done by BrightLocal [2], 92% of customers in 2015 read online reviews. Reviews are often provided about different aspects of the service or product and most of them can be translated into cost (revenue) that one spends (gains) by choosing to purchase that product or service. In a queueing system, the main cost imposed to customers is the waiting cost and the main revenue is the value of the service they gain by being served, which is also called *reward*.

One important parameter of a queue is the service time which is usually considered as a random variable. The randomness nature of the service time, is the source of uncertainty in the cost. This uncertainty about the service time leads customers to find a way to estimate it. In the literature of observable queues, originated by Naor [14] it is assumed that the expected service rate is known. In a very recent paper, Cui and Veeraraghavan [3] assume that the service rate is unknown and customers can estimate it through different procedures like their previous experience and reading reviews. They implicitly note that, if the server reveals the information about the service rate, customers just use that information to make a decision. Therefore, the problem is either to reveal the information or not, in order to improve the performance of the system.

We consider a single server queue with a *first-come first-served* (FCFS) discipline in which some previous customers randomly provide a review on the service time after being served. Each new customer picks some reviews randomly and estimates the service time based on those reviews. We assume that customers are homogeneous with respect to the number of reviews they read. However, as the reviews are written and chosen randomly, the estimations of the service time may vary among customers which results in heterogeneous customers in terms of the estimation of the service rate. The aim of this research is to determine the impact of nowadays' prevalence of customer reviews on the performance of a queueing system. Specifically, we investigate the impact of the number of reviews on the probability of idleness, the expected length of queue, and, revenue. The reviews are written based on the exact service time that the reviewer experiences, i.e., customers write the reviews based on the truth and also reviews are picked randomly. Therefore, the information about the service time in each review, follows the same distribution as the service time. However, the results of this study show that, the same distribution of the exact service time and reviews, does not result in a system that is identical to the one with the known service rate.

Chapter 2

Literature Review

Literature on queueing systems in which customers are decision makers, began with Naor [14]. Strategic customers in Naor's model make their decision about whether to join the queue or not, based on the expected costs and benefits of joining. Naor considers an observable FCFS single server queue with homogeneous customers whose decision is a function of the parameters of the system and also the length of the queue they observe upon arrival. He assumes that customers are homogeneous in terms of the waiting cost and the reward of the service completion. In addition, they all have full information about the service time. Therefore, for any given queue length observed upon arrival, all customers make the same decision.

After Naor, various streams of research have emerged, discussing different aspects of his model including: changing the assumptions and expanding his model. Among the broad literature, we can mention Huang et al. [9] which discuss the rationality assumption and Hassin [6] investigates the impact of a *last-come first-served* (LCFS) discipline on the social welfare optimization.

In the stream that is closely related to this research, the homogeneity assumption in Naor [14], has been brought into question. Hassin ans Haviv [7] provide a comprehensive review on this stream of research. Edelson and Hildebrand [4] consider that customers are heterogeneous in terms of the waiting cost per unit of time. Larsen [11] assumes that the reward, is different for each customer and it follows a uniform distribution. Zheng [19] also considers heterogeneous customers in terms of the waiting cost, by considering a uniform distribution for cost. Zheng also adds a new assumption of heterogeneity in terms of the service rate. He considers two types of customers: optimistic and pessimistic. For each type of customers he assigns a belief about the service rate. Then, he discusses the joining probabilities of optimistic and pessimistic customers.

In our model, the queue is observable and although the customers are homogeneous in the waiting cost and the reward they gain by service completion, because of the various estimations of the service rate, they may make different decisions in the same state of the system. The most closely related model to our research, is provided by Cui and Veeraraghavan [3]. They suggest a model in which the service rate is unknown for customers. Thus, each customer considers a threshold for the number of people in the system below which, she joins the queue. Then, they assume a general but finite distribution for the thresholds. They show that for pessimistic and consistent thresholds of customers; i.e., when the expected value of customers' threshold is not more than the threshold derived from Naor's model, by revealing the information about the service rate, the revenue of the service provider increases.

Revealing information about the service rate in Cui and Veeraraghavan [3] means that all customers have full information about the service rate and so the system turns to Naor's model. In other words, they assume that, if the server reveals the exact expected service time, customers forget about reviews and take the released information and the corresponding threshold, into consideration. As a result, they do not specifically investigates the impact of reviews on the queueing system. In order to fill this gap in the literature, we introduce a mechanism for writing and reading reviews. We show that although this mechanism results in constant beliefs about the service rate, it is always optimistic in terms of the joining threshold. The main difference between this study and Cui and Veeraraghavan [3], is that we assume customers rely on reviews to estimate the rate of service. Thus, the question in our research is to determine whether reading more or less reviews improves the performance of the system.

There is also some research investigating the impact of the previous experiences on a queueing system. These experiences can be obtained either from the customer, herself or from the other customers. Ho et al. [8] investigate the satisfaction of a customer from the most recent purchase of that customer and its impact on her decision. Using the experience of the other customers about the waiting time is discussed by Sankaranarayanan et al. [15]. In their agent-based model, at the beginning of each period, each agent based on its own experience and also the experience of other agents about the sojourn time, decide either to queue up or not. Veeraraghavan and Debo [18] and also Jin et al. [10] discuss how people might make a decision based on the behavior of the other customers. They discuss a case in which the reward is unknown and customers guess it by observing the other customers' behavior. In their context, a server with a longer queue indicates a service with a higher quality. However, none of these papers, consider the reviews on the service time by a focus on the impact of the amount of information obtained from reviews, which is the aim of this study.

Chapter 3

Model

We consider an observable single-server queueing system in which customers arrive to the system according to a Poisson process with parameter Λ, and they decide to join the queue or not, upon arrival. Once one chooses to join, she will not change her decision had been made upon arrival. In other words, there is no *reneging* in this system. The cost of waiting per unit of time, represented by c, and each customer receives a *reward*, R, after the service completion.

The service time for each customer follows an exponential distribution with rate μ. Each customer has a belief about the service rate which comes from the reviews provided by other customers, who have already experienced the service. We assume that customers who decide to provide a review, express the truth about the service time they encountered. We denote the belief of each customer about the service rate by $\hat{\mu}$. Consider a customer with belief $\hat{\mu}$, confronting n customers in the system upon arrival. Then, this customer decides to join if and only if:

$$\frac{(n+1)c}{\hat{\mu}} \leq R. \tag{3.1}$$

In this research, we first find the stationary distribution of the number of people in the system when each customer reads one review randomly and take the information in the review as her belief about the service time. Then, we compare such a *review-based* system with the model with full information about service rate, μ. From now on we refer to Noar's model which is a similar system with full information about μ, as the *M/M/1* system.

Then, we generalize our analysis to a system, in which, every customer reads m reviews. In this context, we analyze the impact of the number of reviews, m, on the performance of the system. In the third section, we investigate the conditions under which our model behaves similar to the M/M/1 system. Finally, we discuss reviews that reflect both waiting time and the number of people in the system, instead of the service time.

3.1 One Review

We assume that some customers randomly decide to provide a review and share their experience about the service time, which is available to all customers. Besides, strategic customers who would make a decision about queueing up, pick one review among all written reviews, randomly. As a result, the belief about the service time, $1/\hat{\mu}$, is a random variable following an exponential distribution with parameter μ. Note that, joining condition is the same as Eq. (3.1) in which $1/\hat{\mu}$ follows an exponential distribution. Thus, we can obtain the probability of joining when there are n customers in the system, $Pr(n)$, as follows:

$$Pr(n) = 1 - e^{-\frac{R\mu}{c(n+1)}}. \tag{3.2}$$

To obtain the probability of joining when there are n customers in the system upon arrival, we use a birth and death process. Let π_n^1 denote the probability that there are n customers in the system, when customers decide based on one review on the service time. Also, define the state of the process as the number of customers in the system. Then, the probability of being in state n is:

$$\pi^1_n(\rho, \frac{R\mu}{c}) = \frac{\rho^n \prod_{k=0}^{n-1} Pr(k)}{1 + \sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} Pr(j)}; n \in \{1, 2, 3, ...\}, \tag{3.3}$$

where $\rho = \Lambda/\mu$. This probability is a function of ρ and $\frac{R\mu}{c}$. Eq. (3.4) provides the probability that the system is idle, i.e., π^1_0,

$$\pi^1_0(\rho, \frac{R\mu}{c}) = \frac{1}{1 + \sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} Pr(j)}. \tag{3.4}$$

For more details about the birth and death process and the procedure to find the stationary distribution of the number of people in the system, please see Appendix A.1.

Note that, any belief about the service time corresponds to a threshold for the number of customers in the system, T, below which a customer joins the system. In the M/M/1 system, this threshold is $\lfloor \frac{R\mu}{c} \rfloor$. Therefore, in the M/M/1 system, a customer joins if and only if she encounters less than or equal to $\lfloor \frac{R\mu}{c} \rfloor$ customers in the system, upon arrival. In a system with belief $\hat{\mu}$ about the service rate, this threshold is $\lfloor \frac{R\hat{\mu}}{c} \rfloor$. As a result, Eq. (3.4) can be written in terms of the distribution of the beliefs about thresholds instead of the service time:

$$\pi^1_0(\rho, \frac{R\mu}{c}) = \frac{1}{1 + \sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} (1 - F(T))}, \tag{3.5}$$

where $F(T)$ is the *cumulative distribution function* (CDF) of thresholds. Note that, Eq. (3.5) is equivalent to the Eq.(1) in Cui and Veeraraghavan [3].

In case that the distribution of the service time is known and all customers have full information about it, Naor [14] provides the stationary distribution of the number of people in the system, which is:

$$\pi^{M/M/1}_n(\rho, \frac{R\mu}{c}) = \frac{\rho^n (1 - \rho)}{1 - \rho^{\lfloor \frac{R\mu}{c} \rfloor + 1}}; n \in \{0, 1, 2, 3, ..., \lfloor R\mu/c \rfloor\}. \tag{3.6}$$

Considering that

$$1 - x^{q+1} = (1-x)\sum_{i=0}^{q} x^i,$$

$\pi_0^{M/M/1}$ can be also written as:

$$\pi_0^{M/M/1} = \frac{1}{\sum_{i=0}^{\lfloor \frac{R\mu}{c} \rfloor + 1} \rho^i}. \tag{3.7}$$

The rest of this study about the system with one review on the service time, has been structured as follows: first, we compare the congestion in the system when customers read one review, with the congestion in the corresponding M/M/1 system. Then, we discuss the revenue in such a system, and, finally we examine whether the proposed model results in either pessimistic or optimistic beliefs defined by Cui and Veeraraghavan [3].

3.1.1 Congestion

One of the most popular and easy to use performance measures of a queue is the probability that the server is idle. From the service provider's point of view, it is more efficient to have less idle time, since with a given service time, less idleness results in serving more customers, which consequently increases the revenue. We examine, under what circumstances, the system with one review is more or less efficient than the M/M/1 system. In order to conduct this comparison, probability of idleness in both systems have been obtained numerically for $\frac{R\mu}{c} \in \{1,\ 1.1,\ 1.2, ..., \ 5\} \cup \{10, 20, ..., 100\}$ and $\rho \in \{0.1, 0.2, ..., 0.9\}$ and the results are illustrated in Figure 3.1.

From Figure 3.1, we observe that:

i. For any integer value of $\frac{R\mu}{c}$, we have: $\pi_0^{M/M/1} \leq \pi_0^1$. The intuition behind this result is that, for an integer value of $\frac{R\mu}{c}$, customers who observe $\frac{R\mu}{c} - 1$ people in the

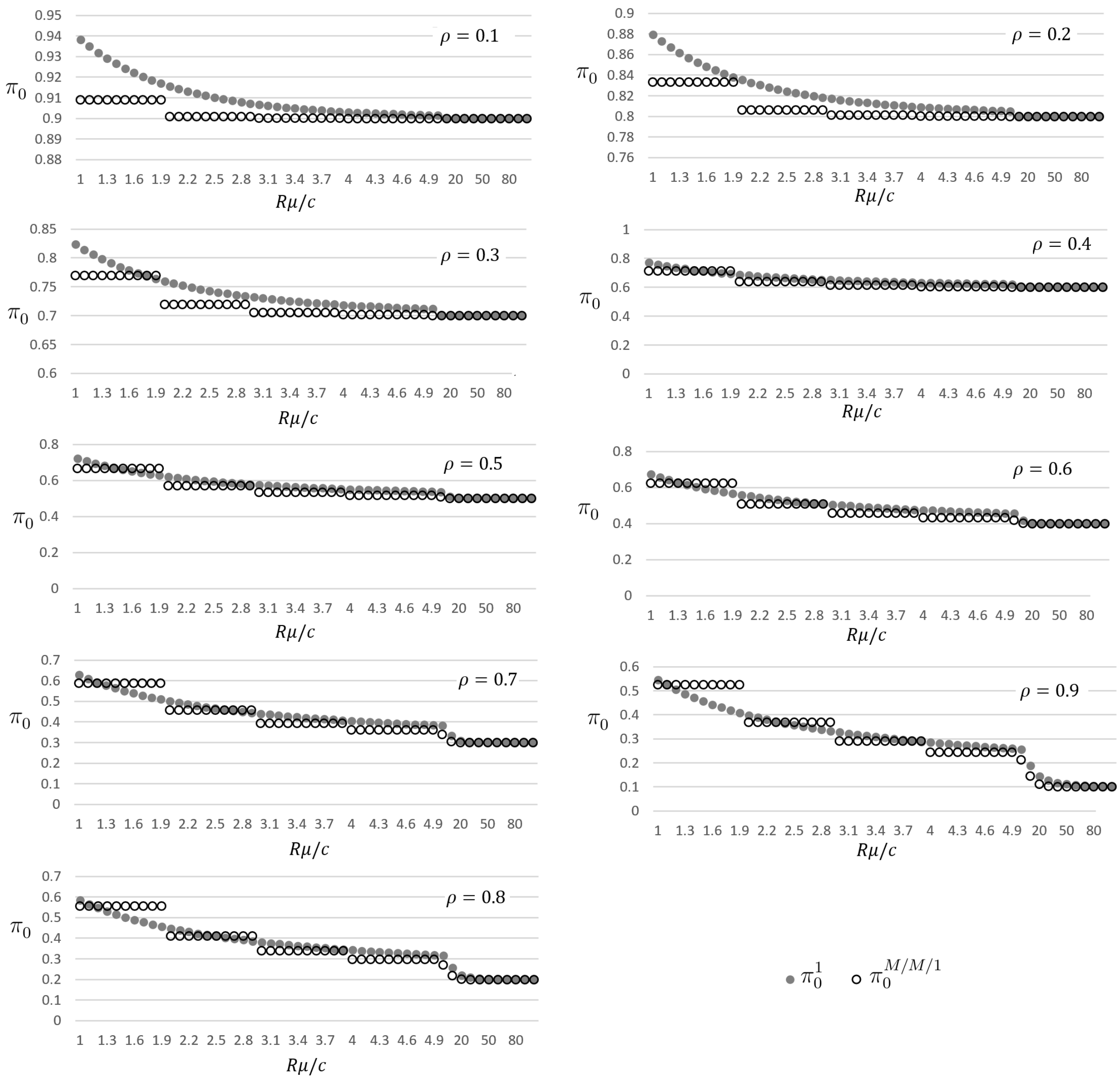


Figure 3.1: A comparison between π_0^1 and $\pi_0^{M/M/1}$ for $\frac{R\mu}{c} \in \{1,\ 1.1,\ 1.2,...,\ 5\} \cup \{10,20,...,100\}$ and $\rho \in \{0.1, 0.2, ..., 0.9\}$

system upon arrival, decide to join if they have full information about the service time. However, if customers estimate the service rate using reviews, for estimations even slightly less than the exact service rate, they do not join. For example if $\frac{R\mu}{c} = 2$, or equivalently $\mu = 2\frac{c}{R}$, in case of observing 1 person in the system upon arrival, the following happens in each system: (a) The M/M/1 system: The waiting cost is $\frac{c\times 2}{\mu} = R$. As the expected waiting cost is not more than the reward, customers who face 1 customer upon arrival, join. (b) System with one review: If $\hat{\mu} < \mu$, then, the waiting cost is $\frac{c\times 2}{\hat{\mu}} > \frac{c\times 2}{\mu} = R$, so they do not join. Otherwise, they join.

Therefore, some customers underestimate service rate. These customers, decide not to join in the review-based system with one review, while all customers with a similar situation join the M/M/1 system. We further discuss this case in Conjecture 1.

ii. If π_0^1 is less than $\pi_0^{M/M/1}$, for all values of $\frac{R\mu}{c}$ between two consecutive integers, k and $k+1$, it will remain the same for any $\frac{R\mu}{c}$ more than k. In addition, $\pi_0^{M/M/1}$ is a lower asymptote of π_0^1. We can mathematically show that $\pi_0^{M/M/1}$ is the asymptote of π_0^1 with respect to $\frac{R\mu}{c}$; i.e.,

$$\lim_{\frac{R\mu}{c}\to\infty} \pi_0^1 = \pi_0^{M/M/1}.$$

As joining probabilities approach 1 when $\frac{R\mu}{c}$ goes to infinity, the limit equals to:

$$\pi_0^{M/M/1} = \frac{1}{\sum_{i=0}^{\lfloor \frac{R\mu}{c} \rfloor +1} \rho^i}.$$

iii. The value of π_0^1 is strictly decreasing with respect to $\frac{R\mu}{c}$. Note that $\pi_0^{M/M/1}$ is a function of $\lfloor \frac{R\mu}{c} \rfloor$. Thus, it is constant between each two consecutive integer values of $\frac{R\mu}{c}$. On the other hand, $\pi_0^{M/M/1}$ is a function of $\frac{R\mu}{c}$. Based on Eq. (3.2), joining probabilities are all strictly increasing with respect to $\frac{R\mu}{c}$ which implies that π_0^1 is strictly decreasing.

The average number of people waiting in the system, including the one in the server, can be considered as another performance measure of a queueing system, which is:

$$\sum_{l=0}^{\infty} k\pi_l,$$

where π_l stands for the stationary probability of state l. Let EL^1 and $EL^{M/M/1}$ denote the expected number of customers in the system with one review and the M/M/1 system, respectively. Then,

$$EL^1 = \frac{\sum_{n=1}^{\infty} n\rho^n \prod_{k=0}^{n-1} Pr(k)}{1 + \sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} Pr(j)}, \tag{3.8}$$

and,

$$EL^{M/M/1} = \frac{\sum_{n=1}^{\lfloor \frac{R\mu}{c} \rfloor} n\rho^n}{1 + \sum_{i=1}^{\lfloor \frac{R\mu}{c} \rfloor + 1} \rho^i}. \tag{3.9}$$

Figure 3.2 provides a comparison between the expected number of people in the review-based system with one review and that of the M/M/1 system. The expected length of the queue in the review-based system, strictly increases in $\frac{R\mu}{c}$: more reward, more people decide to join, a longer queue. This figure also shows that, although using just one review to estimate the service time, leads to a system with more idle time for an integer $\frac{R\mu}{c}$, it does not necessarily induce a less congested system. In other words, when π_0^1 is greater than $\pi_0^{M/M/1}$, it does not always mean that π_1^1 is less than $\pi_1^{M/M/1}$. It can be interpreted from Figure 3.2 that when the arrival rate, Λ, is high, for smaller integer values of $\frac{R\mu}{c}$, the system with one review is more congested. We can also show that $EL^{M/M/1}$ is an asymptote of EL^1 and numerical results indicate that it is an upper asymptote. Although Figure 3.2 does not illustrate that EL^1 is asymptotically approaching $EL^{M/M/1}$, when $\rho = 0.9$, we verified that when $\lfloor \frac{R\mu}{c} \rfloor$ is large enough, the expected number of people in the

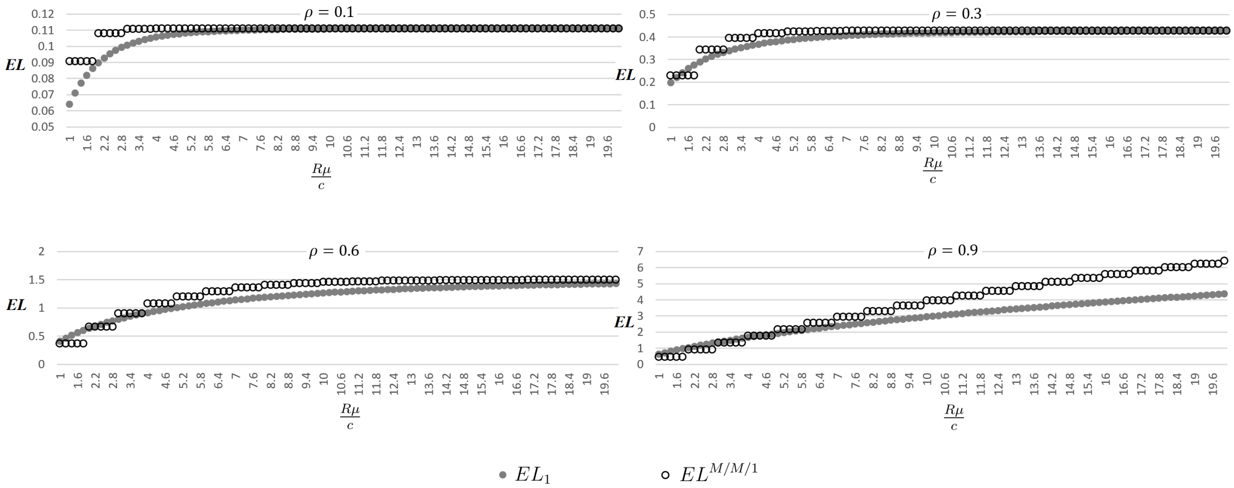


Figure 3.2: A comparison between EL^1 and $EL^{M/M/1}$ for $\frac{R\mu}{c}$ $\in$\{1, 1.2, 1.4,...,20\} and $\rho \in \{0.1, 03, 0.6, 0.9\}$

system, approaches that of the M/M/1 system with a slower rate. Table 3.1 shows the the expected number of people in both systems. The results verify that for a large enough $\frac{R\mu}{c}$, they are almost equal.

3.1.2 Revenue

In this section, we investigate the impact of reviews, on the revenue of the service provider. Assume that once customers join the system, the server charges them a service fee, denoted by p. Note that this assumption does not affect the results given in Section 3.1.1, because we can replace R with $R - p$ and conclude the same results. The effective rate of arrival, denoted by Λ_e can be defined as $\Lambda \times$ *joining probability* (see Hassin and Haviv [7] for detailed definition). Joining probability can be obtained by conditioning on the state of the system which is defined as the number of people in the system. The effective rate of

Table 3.1: Verifying if EL^1 approaches $EL^{M/M/1}$ for a large $\frac{R\mu}{c}$

ρ	$R\mu/c$=100		$R\mu/c$=1000	
	EL^1	$EL^{M/M/1}$	EL^1	$EL^{M/M/1}$
0.1	0.111111111	0.111111111	0.111111111	0.111111111
0.2	0.250000000	0.250000000	0.250000000	0.250000000
0.3	0.428571415	0.428571429	0.428571428	0.428571429
0.4	0.666666334	0.666666667	0.666666667	0.666666667
0.5	0.999994220	1.000000000	1.000000000	1.000000000
0.6	1.499909957	1.500000000	1.500000000	1.500000000
0.7	2.331871463	2.333333333	2.333333334	2.333333333
0.8	3.971160418	3.999999983	4.000000000	4.000000000
0.9	8.159853272	8.997585512	9.000000000	9.000000000

arrival can be written as:

$$\Lambda_e = \mu(1 - \pi_0), \tag{3.10}$$

(see Appendix B for details).

Eq. (3.10) and Figure 3.1 together imply that Λ_e for both the M/M/1 system and the system with one review, is increasing with respect to $\frac{R\mu}{c}$.

The expected revenue is defined as $p\Lambda_e$, when R is replaced with $R - p$ in Eq. (3.4) and Eq. (3.7). To simplify our analysis, we compare the M/M/1 system with the model with one review for a given p, rather than finding the revenue optimizer price. Revenue analysis based on a given p, is also used by Cui and Veerarghavan [3]. Here, we compare the revenues based on the prices for which, $\frac{(R-p)\mu}{c}$ is between two successive integers, k and $k + 1$.

In the M/M/1 system, when $\frac{(R-p)\mu}{c} = k$, the probability of idleness can be obtained by replacing n with zero in Eq. (3.6) which yields:

$$\pi_0^{M/M/1} = \frac{1-\rho}{1-\rho^{k+1}}.$$

Considering Eq. (3.10), revenue in the M/M/1 system can be written as:

$$Revenue = p\mu(1 - \pi_0^{M/M/1}).$$

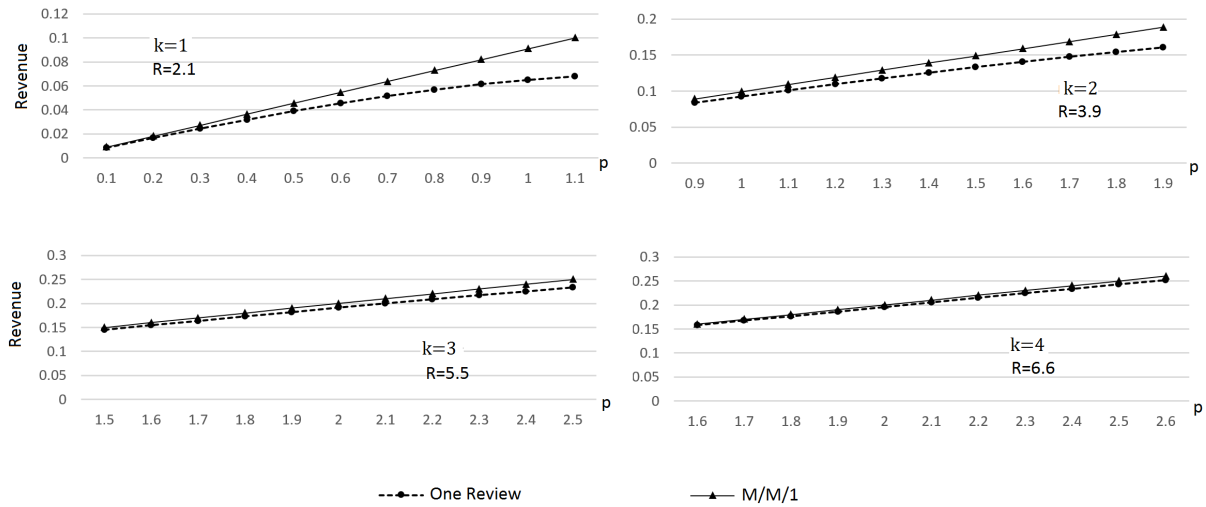
For a given k, it is straight-forward to show that the revenue increases in p. Therefore, in order to increase the revenue in the given interval for p, the service provider should charge as high as possible. Since, $\frac{(R-p)\mu}{c}$ is between k and $k+1$ by assumption, the maximum price for a given k is:

$$p = R - \frac{ck}{\mu}.$$

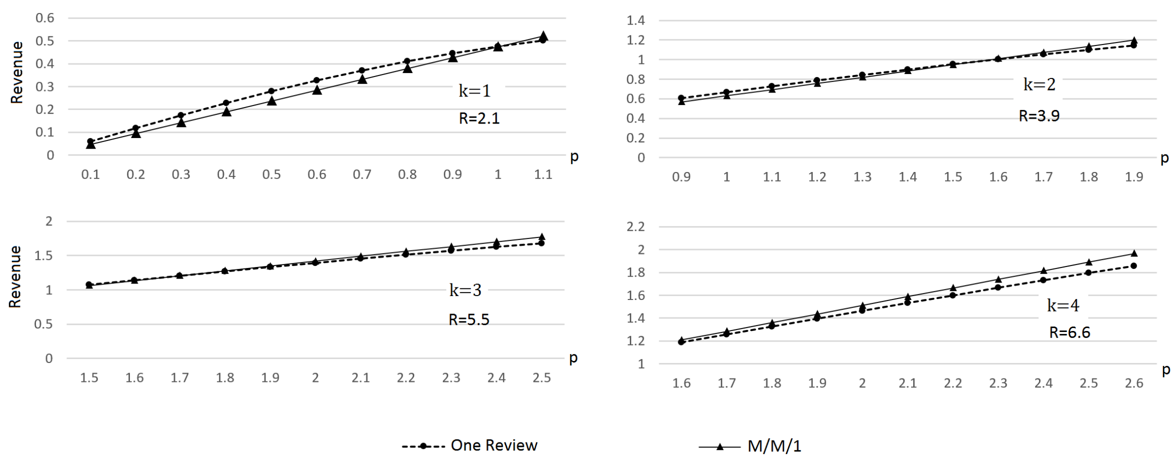
Unlike the M/M/1 system, in the review-based model we already discussed that π_0^1 is strictly decreasing with respect to $\frac{(R-p)\mu}{c}$, which implies that it is increasing with respect to p. As a result, Λ_e is decreasing with respect to p. Therefore, it is not straight-forward to determine if the revenue is either increasing or decreasing in p.

We next, examine the effect of the price on the revenue, numerically. In the examples illustrated in Figure 3.3, we set $\mu = c = 1$ and vary k from 1 to 4. Then, we obtained the revenue of both the M/M/1 system and the system with one review for some values of R and ρ, for the prices such that $\frac{(R-p)\mu}{c}$ varies between k and $k+1$.

Figure 3.3 shows that although Λ_e is decreasing with respect to p, revenue is increasing between each two integers for $\frac{(R-p)\mu}{c}$. However, we cannot say if the revenue of the M/M/1



(a)



(b)

Figure 3.3: The Revenue of review-based and M/M/1 systems for various values of R and k such that $\frac{(R-p)\mu}{c} \in [k, k+1)$. (a)$\rho = 0.1$ and (b)$\rho = 0.9$

system is always more or less than the one in the review-based model. The numerical results suggest that for a small ρ, the M/M/1 system always gain more revenue for a given p. However, when ρ is high, this result does not always hold.

3.1.3 Optimistic vs. Pessimistic Customers

In the review-based system, we assume that reviewers provide the exact service time they have experienced. In addition, the potential customers read a review randomly which results in the beliefs that follow an exponential distribution with the same parameter as the exact service time. As a result, customers are neither optimistic nor pessimistic regarding the service time. However, Cui and Veeraraghavan [3] define pessimism and optimism based on the threshold, rather than the service rate.

As we mentioned in Section 3.1, each customer with a certain belief about the service time, considers a threshold, T, for the number of customers in the system. Therefore, if the customer encounters less than or equal to T customers in the system, she joins. This threshold is equal to $\lfloor R\hat{\mu}/c \rfloor$. Cui and Veeraraghavan [3], consider the population to be optimistic (pessimistic), if the expected threshold, $E(T)$ is more (less) than the one in the M/M/1 system. Then, they show that a system with pessimistic and even constant beliefs about joining thresholds, T, results in a lower revenue than the M/M/1 system. In order to use their result, we should examine if in the system with one review, customers are pessimistic or optimistic about their joining threshold. The probability that a customer has a threshold less than or equal to i, is:

$$Pr\{T \leq i\} = Pr(\lfloor R\hat{\mu}/c \rfloor \leq i) = Pr(\tfrac{R\hat{\mu}}{c} < i+1) = Pr(\tfrac{1}{\hat{\mu}} > \tfrac{R}{c(i+1)}) = e^{-\frac{R\mu}{c(i+1)}}.$$

Similarly, the probability that a customer has a threshold less than or equal to $i-1$ is:

$$Pr\{T \leq i-1\} = e^{-\frac{R\mu}{c\times i}}.$$

Therefore, the probability that a customer considers i as her threshold, can be written as follows:

$$Pr\{T \leq i\} = Pr\{T \leq i\} - Pr\{T \leq i-1\} = e^{-\frac{\mu R}{c(i+1)}} - e^{-\frac{\mu R}{c\cdot i}}.$$

As a result, the expected value of the customer beliefs about the threshold which is denoted by $E(T)$ is:

$$E(T) = \sum_{i=1}^{\infty} i \cdot \left(e^{-\frac{\mu R}{c(i+1)}} - e^{-\frac{\mu R}{c(i)}}\right). \tag{3.11}$$

We numerically approximated $E(T)$ for integer values of $\frac{R\mu}{c} \in \{1, 2, ..., 200\}$. Figure 3.4 shows that in this system, although the population is consistent about the service time (neither optimistic nor pessimistic), they are defined significantly optimistic according to Cui and Veeraraghavan [3], i.e., ($E(T) > \frac{R\mu}{c}$). Therefore, we cannot use their result about the revenue in this model.

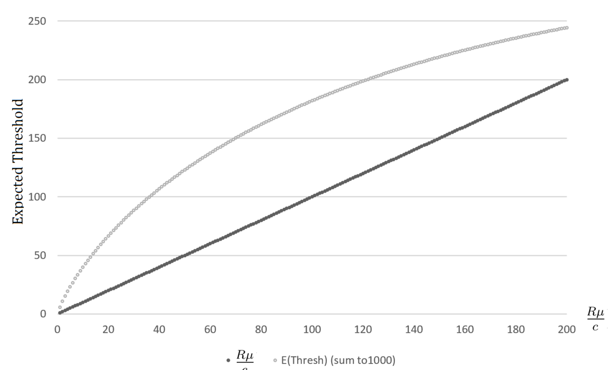


Figure 3.4: Approximated expected threshold vs. the threshold in the M/M/1 system.

3.2 General Number of Reviews

In Section 3.1, we discussed the impact of one review on the performance of the queueing system. Now, the question is, what happens to the performance of the system, if each customer reads more reviews. Does it necessarily result in a system with a performance closer to that of the M/M/1 system? In other words, we investigate the impact of the amount and the accuracy of information about the service time, on the performance measures of the system. In this section we show that sometimes the performance of the model with less number of reviews, is more similar to that of the system with full information.

3.2.1 Congestion

Assume that some of the previous customers have provided a review on the service time they experienced. Meanwhile, new customers pick m reviews randomly and they consider the average of the service times of the chosen reviews as an estimation for the service time, $\frac{1}{\hat{\mu}}$. We can show that in the steady state, this system is in state n, with probability π_n^m that can be obtained from Eq.(3.3) with a different joining probability, $Pr(.)$. The details are given in Appendix A.2.

We next obtain the probability of joining, when there are n customers in the system. As the service time in each review is a random variable following an exponential distribution, the summation of the service times in the reviews, follows an Erlang distribution with the shape parameter, m and the rate, μ (see Evans et. al [5]). As a result, $Pr(n)$ can be defined as:

$$Pr(n) = 1 - \sum_{l=0}^{m-1} \frac{1}{l!} e^{\frac{-mR\mu}{c(n+1)}} (\frac{mR\mu}{c(n+1)})^l. \tag{3.12}$$

We numerically obtain $\pi_0^m(\rho, \frac{R\mu}{c})$, for various values of $m,\ \rho,\ and,\ \frac{R\mu}{c}$, from two points of view: First, we examine if the observations about the congestion of the system with one review, holds for a general number of reviews. Second, we investigate the impact of the number of reviews on the performance of the system.

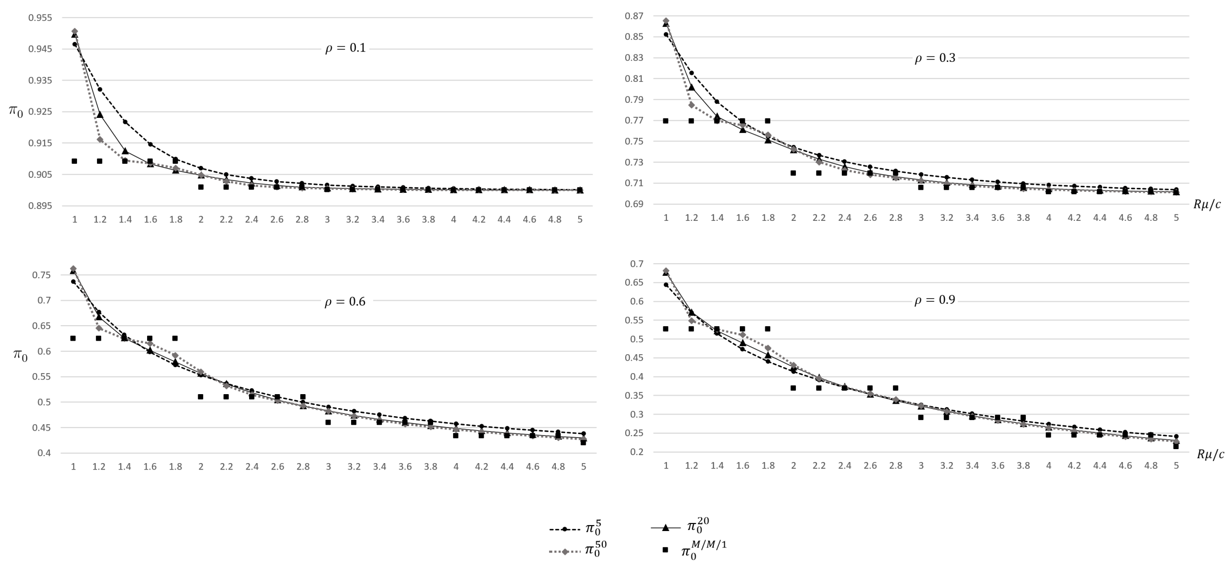


Figure 3.5: A comparison between π_0^m and $\pi_0^{M/M/1}$ with respect to $\frac{R\mu}{c}$ for various numbers of reviews.

In order to compare the system with m reviews with the M/M/1 system, we plotted π_0^m, for $m \in \{5, 20, 50\}$ and $\rho \in \{0.1, 0.3, 0.6, 0.9\}$ as a function of $\frac{R\mu}{c} \in \{1, 1.2, 1.4, ..., 5\}$, in Figure 3.5. Numerical examples reveal that:

i. For any integer value of $\frac{R\mu}{c}$, we have: $\pi_0^{M/M/1} \leq \pi_0^m$. The intuition behind this result is the same as what discussed in Section 3.1.1.

ii. Like π_0^1, π_0^m is always decreasing in $\frac{R\mu}{c}$.

iii. If π_0^m is less than $\pi_0^{M/M/1}$, for all values of $\frac{R\mu}{c}$ between two consecutive integers, k

and $k+1$, it remains the same for any $\frac{R\mu}{c}$ more than k. In addition, $\pi_0^{M/M/1}$ is a lower asymptote of π_0^m.

In the context of this research, we are more interested in studying the impact of the number of reviews everyone reads, on the performance of the system. In other words, we examine if it is more beneficial to have a larger or smaller pool of reviews.

Here, the question is that how the probability of idleness changes, if each customer reads more reviews, before making a decision. One possible response may be as follows: by increasing the number of reviews, the estimation of the service approaches the exact expected service time and, the customers decide more similar to the case when they have full information. As a result, the stationary distribution of the review-based system becomes closer to that of the M/M/1 system. Although this explanation might be correct for a large enough m, we show that it is not true in general. In other words, we show that having more accurate information does not necessarily leads to a more similar system to a system with full information, in terms of the stationary distribution of the number of people in the system. We clarify it by two examples.

Consider Figure 3.6a and assume that each customer reads 3 reviews. In this case, the probability of idleness in a system with three reviews is higher than the one in the M/M/1 system. However, it increases for more than three reviews and so the gap between the review-based system and the M/M/1 system, increases. In other words, by increasing the number of reviews we can see that π_0^m does not necessarily become closer to $\pi_0^{M/M/1}$. Figure 3.6b also illustrates another case in which with more number of reviews, the system not only does not become more similar to the M/M/1 system, but also, the gap between them increases. In case (a), it is more efficient to have just one review and keep the pool of reviews as small as possible. In contrast, in case (b), we can observe that it is more efficient to have 5 reviews rather than 1 to 4 reviews.

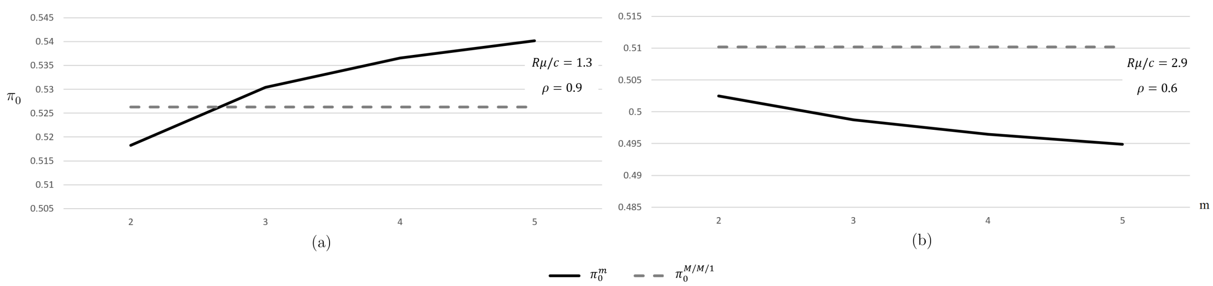


Figure 3.6: The probability of idleness in two systems with 1 to 5 reviews vs. the M/M/1 system

We obtained π_0^m for m from 1 to 15 and various amounts of $\frac{R\mu}{c}$ and ρ. Figure 3.7 depicts the result for three values of $\frac{R\mu}{c}$ and ρ.

In general, we observe many different patterns for π_0^m with respect to the number of observations. Thus, if the service provider wants to decrease the idle time by affecting the number of reviews each customer reads, he needs to determine if the probability of idleness is increasing or decreasing with respect to m, for his system. We also find that when $\frac{R\mu}{c}$ is large enough, π_0^m is decreasing in m. Therefore, in systems with a large $\frac{R\mu}{c}$, more reviews results in a less probability of idleness.

3.3 Approaching the M/M/1 System

It may sound trivial that when the number of reviews each customer reads, go to infinity, the belief about the service rate approaches the exact rate and so π_0^m approaches $\pi_0^{M/M/1}$. However, it is not always true. Consider Figure 3.5. When everyone reads 50 reviews, which is a large number, π_0^{50} looks far away from $\pi_0^{M/M/1}$ for some values of $\frac{R\mu}{c}$. We realized

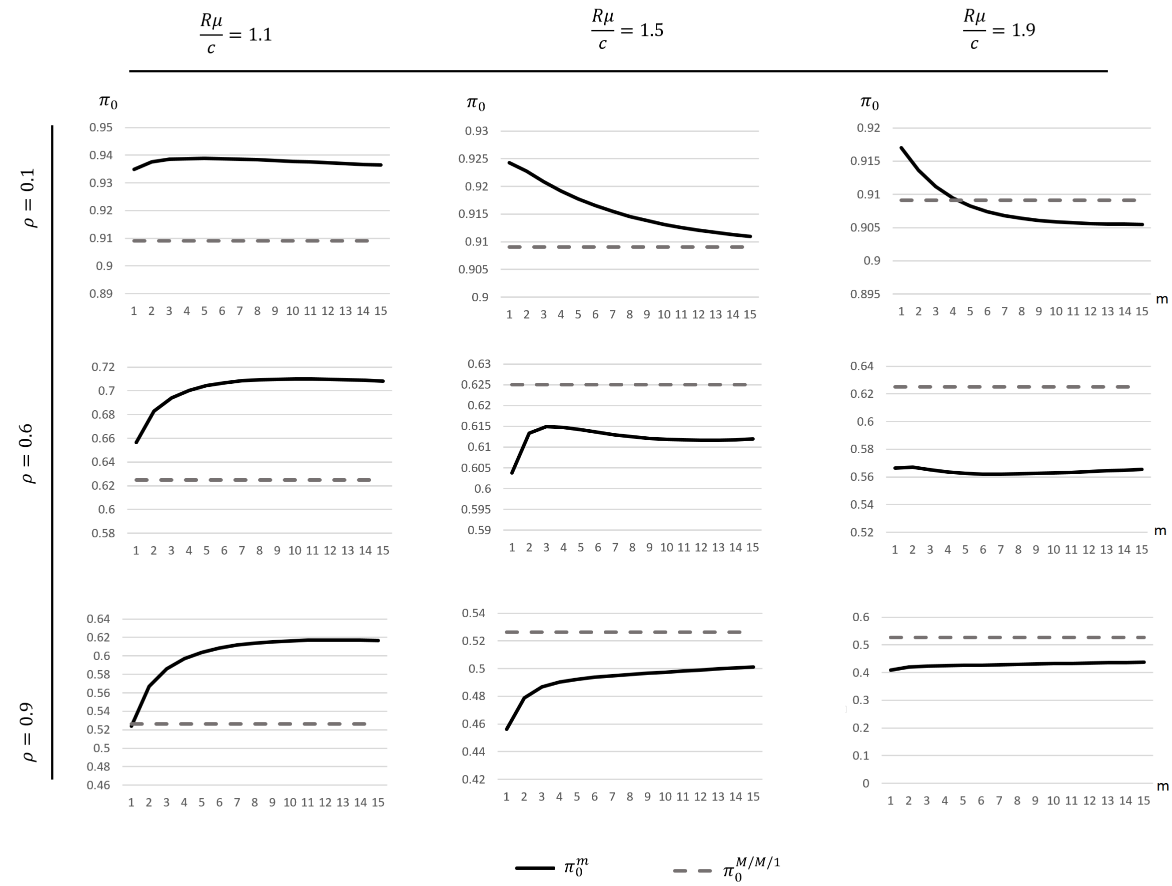


Figure 3.7: Probability of idleness in the systems with various $\frac{R\mu}{c}$ and ρ for 1 to 15 reviews vs. the M/M/1 system

that not all those differences have the same reason. For some values of $\frac{R\mu}{c}$, although π_0^m approaches $\pi_0^{M/M/1}$ asymptotically, the rate is too low. On the other hand, for some other values of $\frac{R\mu}{c}$, the probability of idleness in the review-based system never approaches the one in the M/M/1 system. Below, we discuss each case, briefly.

3.3.1 Conditions under which π_0^m approaches $\pi_0^{M/M/1}$

In this section, we discuss the necessary and sufficient condition for approaching the M/M/1 system when the number of reviews everyone reads goes to infinity. In Conjecture 1, we show that when $\frac{R\mu}{c}$ is an integer, no matter how many reviews everyone reads, π_0^m never approaches $\pi_0^{M/M/1}$. In other words, $\pi_0^{M/M/1}$ is not an asymptote for π_0^m for integer values of $\frac{R\mu}{c}$.

Conjecture 1. *Let* $\pi_0^m = \frac{1}{1+\sum_{i=0}^{\infty}\rho^i\prod_{k=0}^{m-1}(1-\sum_{k=0}^{m-1}(\frac{1}{k!}e^{-\frac{mR\mu}{c(j+1)}}(\frac{mR\mu}{c(j+1)})^k))}$ *for* $R, \mu, c > 0$ *and* $0 \leq \rho < 1$*. Also, let* $\pi_0^{M/M/1} = \frac{1-\rho}{1-\rho^{\lfloor\frac{R\mu}{c}\rfloor+1}}$*. Then,* $\lim_{m\to\infty}\pi_0^m = \pi_0^{M/M/1}$ *if and only if* $\frac{R\mu}{c}$ *is not an integer.*

The intuition behind proof. In order to find the asymptote, since the number of reviews just appears in the joining probabilities, we need to find the limit of the joining probabilities. Based on Eq. 3.12, we should find the limit of the CDF of an Erlang distribution. We can verify that the limit of Eq. 3.12 when m goes to infinity is:

$$\lim_{m\to\infty}(1-\sum_{k=0}^{m-1}(\frac{1}{k!}e^{-\frac{mR\mu}{c(j+1)}}(\frac{mR\mu}{c(j+1)})^k) = \begin{cases} 0, & if\,\frac{mR\mu}{c(j+1)} < m \\ 0.5, & if\,\frac{mR\mu}{c(j+1)} = m \\ 1, & if\,\frac{mR\mu}{c(j+1)} > m \end{cases}.$$

Thus, we face two cases:

One: There is a j, for which the second condition holds, i.e., there is an integer, j, for which $\frac{R\mu}{c} = j+1$ or equivalently, $\frac{R\mu}{c}$ is integer. Then,

$$\lim_{m\to\infty}\left(1+\sum_{i=0}^{\infty}\rho^i\prod_{k=0}^{m-1}\left(1-\sum_{k=0}^{m-1}(\frac{1}{k!}e^{-\frac{mR\mu}{c(j+1)}}(\frac{mR\mu}{c(j+1)})^k)\right)\right) = 1+\sum_{i=0}^{\frac{R\mu}{c}-1}\rho^i + 0.5\rho^{\frac{R\mu}{c}}.$$

Thus,

$$\pi_0^{int} = \lim_{m\to\infty} \pi_0^m = \frac{1-\rho}{1-0.5\rho^{\frac{R\mu}{c}}-0.5\rho^{\frac{R\mu}{c}+1}},$$

where π_0^{int} denotes the asymptote of π_0^1, for integer values of $\frac{R\mu}{c}$. It is straight-forward to show that pi_0^{int} is not equal to $\pi_0^{M/M/1}$ when $\rho \neq 1$.

Two: $\frac{R\mu}{c} \neq j+1, \ \forall j \in \mathbb{Z}$, so:

$$\lim_{m\to\infty} \left(1 + \sum_{i=0}^{\infty} \rho^i \prod_{k=0}^{m-1} \left(1 - \sum_{k=0}^{m-1} \left(\frac{1}{k!} e^{-\frac{mR\mu}{c(j+1)}} \left(\frac{mR\mu}{c(j+1)}\right)^k\right)\right)\right) = 1 + \sum_{i=0}^{\frac{R\mu}{c}-1} \rho^i,$$

and then, $\lim_{m\to\infty} \pi_0^m = \pi_0^{M/M/1}$.

□

Conjecture 1 indicates that when $\frac{R\mu}{c}$ is an integer, even in the case of reading a very large number of reviews, the system has still a significant difference with the M/M/1 system. The source of this phenomena is the decision of customers who face $j = \frac{R\mu}{c} - 1$ customers in the system upon arrival. If n is replaced with j in Eq. (3.1), we can verify that these customers join the system if $c.(\frac{1}{\hat{\mu}}).(\frac{R\mu}{c}) \leq R$ or equivalently $\mu \leq \hat{\mu}$. Therefore, if the estimation of the service rate is higher than the exact service rate, even though the difference converges to zero, they decide not to join. In contrast, they decide to join if they know the exact service rate.

We numerically obtained π_0^m, for integer values of $\frac{R\mu}{c}$ and a large m given in Table 3.2. Numerical results supports the idea that for an integer $\frac{R\mu}{c}$, π_0^m approaches π_0^{int}, which is different from $\pi_0^{M/M/1}$, when m is large enough.

3.3.2 The Rate and the Direction of Approaching Asymptote

In Section 3.3.1, we show that for the integer values of $\frac{R\mu}{c}$, π_0^m does not approach $\pi_0^{M/M/1}$, when m increases. In this section, we examine the rate at which, π_0^m approaches $\pi_0^{M/M/1}$ for

Table 3.2: Comparison between the values of π_0^m for $m \in \{100, 1000\}$, $\pi_0^{M/M/1}$, and, π_0^{int} for $\rho \in \{0.1, 0.5, 0.9\}$ and $\frac{R\mu}{c} \in \{2, 3, 4\}$

		π_0^1		$\pi_0^{M/M/1}$	π_0^{int}
ρ	$R\mu/c$	m=100	m=1000		
0.1	2	0.904868	0.904943	0.900901	0.904977
	3	0.900484	0.900492	0.90009	0.900495
	4	0.900049	0.900049	0.900009	0.90005
05	2	0.614126	0.614987	0.571429	0.615385
	3	0.551186	0.551564	0.533333	0.551724
	4	0.524335	0.524518	0.516129	0.52459
0.9	2	0.431816	0.433199	0.369004	0.433839
	3	0.324116	0.324932	0.290782	0.325256
	4	0.264568	0.265265	0.244194	0.26546

non-integer values of $\frac{R\mu}{c}$. We also discuss either it approaches the asymptote from above or below. Referring back to Figure 3.5, we observe that even for 50 reviews, there are still some non-integer values of $\frac{R\mu}{c}$, which are not close enough to $\pi_0^{M/M/1}$. In order to find if there is a pattern, we have numerically obtained π_0^m for various values of ρ and $\frac{R\mu}{c}$, and, for some large numbers of reviews.

Figure 3.8 illustrates the *relative deviation* of π_0^m from $\pi_0^{M/M/1}$, i.e.,

$$\frac{\pi_0^m - \pi_0^{M/M/1}}{\pi_0^{M/M/1}} \times 100,$$

for $m \in \{100, 200, ..., 1000\}$, $\rho \in \{0.1, 0.3, 0.6, 0.9\}$, and, $\frac{R\mu}{c} \in \{i+0.1, i+0.5+i+0.9\} for\ i \in \{1, 3, 5\}$. Note that we do not use the absolute value of relative deviation, as it is important to see if the probability of the idleness is less or more than the M/M/1 system.

We observe that $\frac{R\mu}{c}$ and ρ play the main role in determining the rate of approaching the M/M/1 system. However, the main factor in determining whether it approaches the asymptote from below or above, is the difference between $\frac{R\mu}{c}$ and the closest integer. Let $frac(\frac{R\mu}{c})$ denote the fractional part of $\frac{R\mu}{c}$, i.e., Let $frac(\frac{R\mu}{c}) = \frac{R\mu}{c} - \lfloor \frac{R\mu}{c} \rfloor$. Then,

i. When $frac(\frac{R\mu}{c})$ is less than *0.5*, $\pi_0^{M/M/1}$ is the lower asymptote of π_0^m. In other words, for a large enough number of reviews, when $frac(\frac{R\mu}{c}) < 0.5$, π_0^m is decreasing with respect to m and approaches $\pi_0^{M/M/1}$, from above. As a result, when a service provider encounters a large number of reviews, every step taken to increase the number of reviews, results in a lower idle time which consequently increases the effective arrival rate and revenue. In addition, as $frac(\frac{R\mu}{c})$ becomes closer to zero, π_0^m approaches its asymptote faster while it may decrease at a lower rate. Figure 3.9a provides an example to support this result.

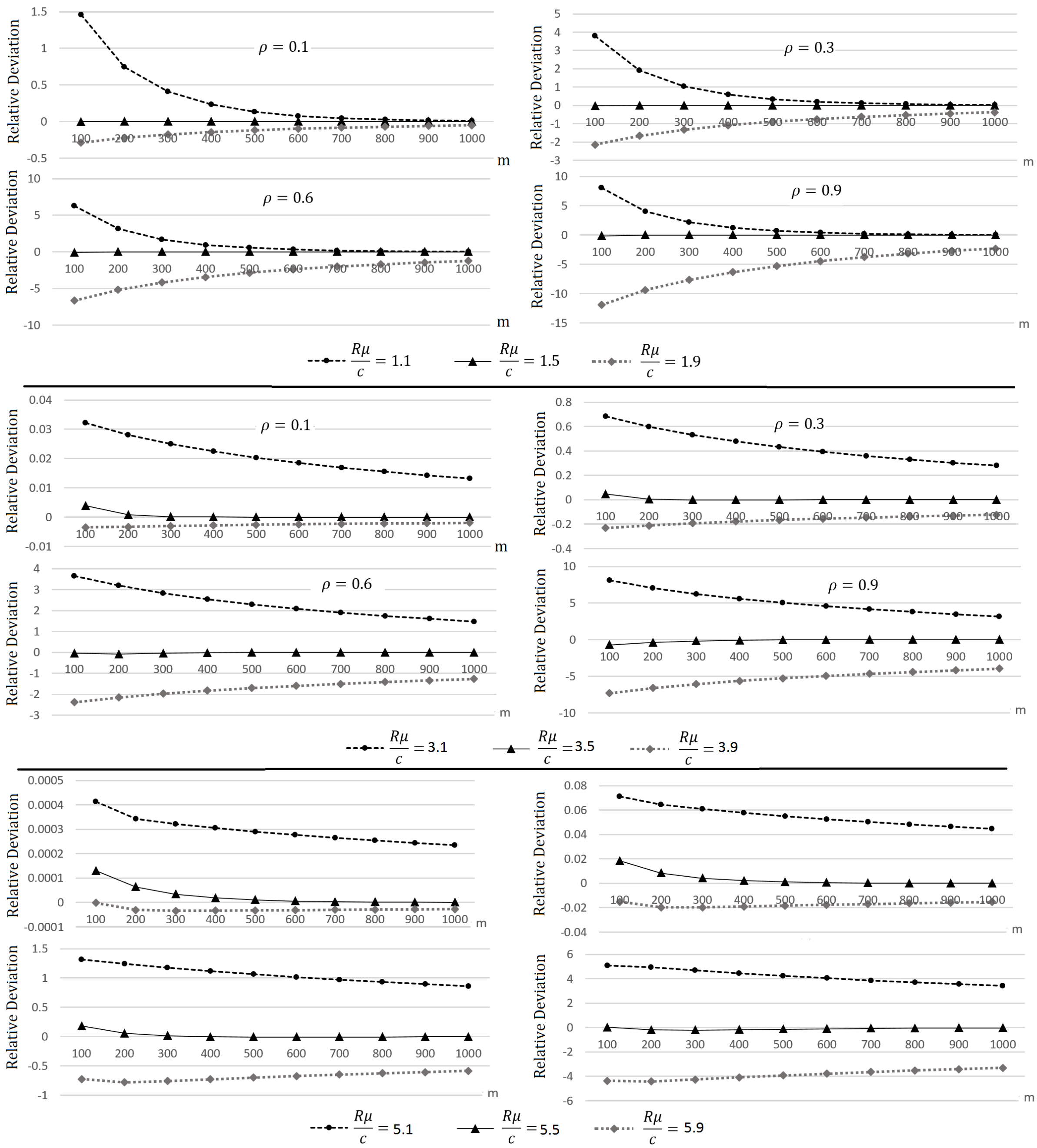


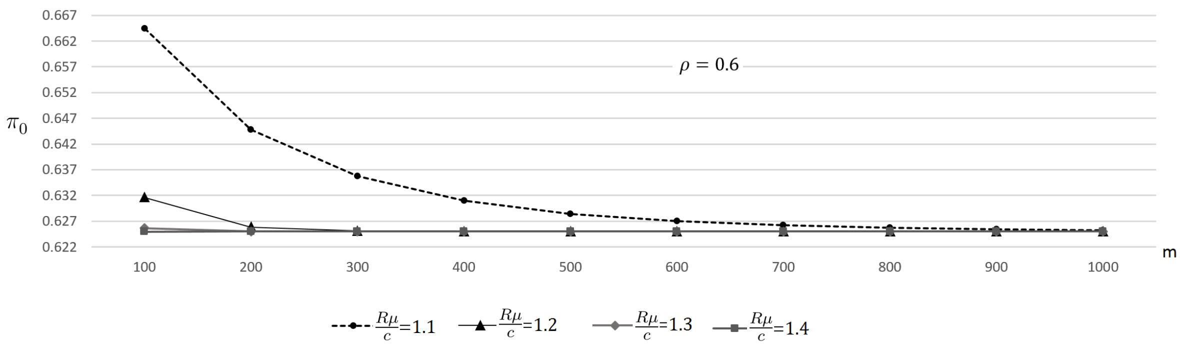
Figure 3.8: The relative deviation of π_0^m from $\pi_0^{M/M/1}$

ii. When $frac(\frac{R\mu}{c})$ is more than or equal to *0.5*, $\pi_0^{M/M/1}$ is the upper asymptote of π_0^m. In other words, for a large enough number of reviews, when $frac(\frac{R\mu}{c}) > 0.5$, π_0^m is increasing with respect to m and approaches $\pi_0^{M/M/1}$, from below. As a result, when a service provider encounters a large number of reviews, every step taken to decrease the number of reviews, will result in a less idle time and consequently increases the effective arrival rate and revenue. In addition, as $frac(\frac{R\mu}{c})$ becomes closer to 1, π_0^m approaches its asymptote slower while it may increase at a higher rate. Figure 3.9b provides an example to support this result.

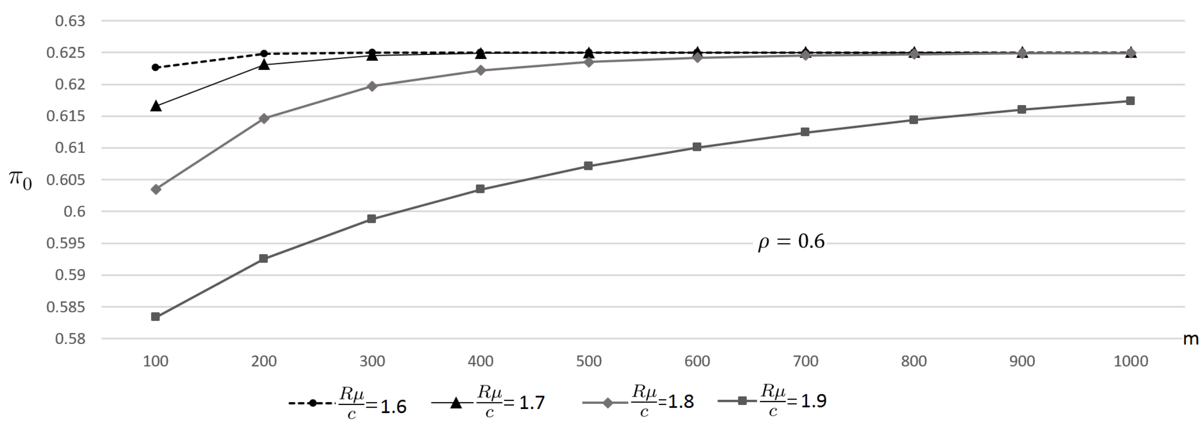
We discussed the impact of the fractional part of $\frac{R\mu}{c}$ on the direction and also the rate at which π_0^m approaches $\pi_0^{M/M/1}$ as m increases. However, the value of ρ and $\frac{R\mu}{c}$ also affects the rate at which π_0^m approaches $\pi_0^{M/M/1}$.

In order to investigate the impact of ρ, we plot the relative deviation of π_0^m from $\pi_0^{M/M/1}$ for different values of ρ and $\frac{R\mu}{c}$. The results are shown in Figure 3.10. According to these results, a system with a lower ρ approaches the M/M/1 system faster, while it may change at a lower rate. For example, when $\rho = 0.1$, π_0^m is closer to $\pi_0^{M/M/1}$ for a large enough m, however, it changes less with respect to m.

In order to discuss the impact of $\frac{R\mu}{c}$ on the rate of approaching asymptote, we plot the relative deviation of π_0^m from $\pi_0^{M/M/1}$ for various values of $\frac{R\mu}{c}$ with the same fractional part. The results are illustrated in Figure 3.11. We observe that when $\frac{R\mu}{c}$ is small, π_0^m approaches its asymptote faster and it changes at a higher rate, too (e.g. see Figure 3.11 for $\frac{R\mu}{c} = 1.1$). As a result, when $\frac{R\mu}{c}$ is low, and there is still a difference between the idle time of the review-based system and the one in the M/M/1 system, a change in the number of reviews, affects the idle time more. When $\frac{R\mu}{c}$ is not small, as it increases, π_0^m approaches its asymptote faster while it may change at a lower rate. For example, in Figure 3.11, for



(a) Value of π_0^m for $\frac{R\mu}{c} \in \{1.1, 1.2, 1.3, 1.4\}$



(b) Value of $\pi_0^m for \frac{R\mu}{c} \in \{1.6, 1.7, 1.8, 1.9\}$

Figure 3.9: The rate of approaching $\pi_0^{M/M/1}$ for different values of $frac(\frac{R\mu}{c})$. $m \in \{100, 200, .., 100\}$ and $\rho = 0.6$

$\frac{R\mu}{c} > 7$, π_0^m is closer to $\pi_0^{M/M/1}$ in compare with a lower $\frac{R\mu}{c}$, while its slope is less.

Overall, we observe that there are three factors affecting the rate of approaching asymptote: (a)$\frac{R\mu}{c}$, (b)ρ, and, (c) fractional part of $\frac{R\mu}{c}$. However, the value of $frac(\frac{R\mu}{c})$ is enough to determine the direction that π_0^m approaches $\pi_0^{M/M/1}$.

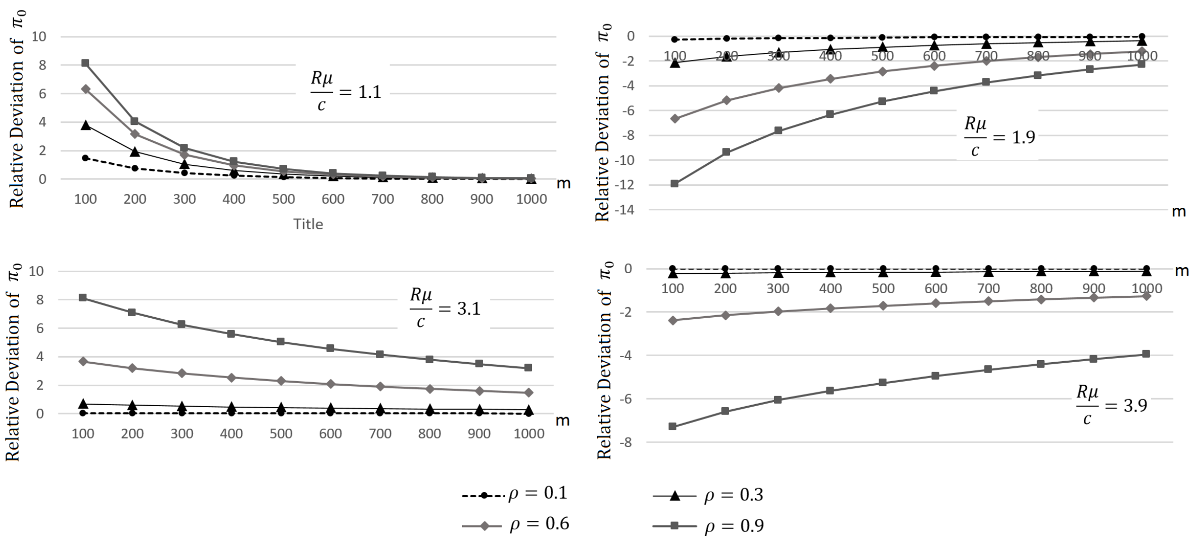


Figure 3.10: The relative deviation of π_0^m from $\pi_0^{M/M/1}$ for various values of ρ.

3.4 Reviews on Waiting Time and the Number of People in the System

So far, we assumed that some customers provide reviews just about the service time they have experienced. What happens if they provide more information including the whole waiting time and the number of people they observed upon arrival? An estimation based on each review can be obtained by dividing the waiting time by the number of people upon arrival, i.e.,

$$\frac{1}{\hat{\mu}} = \frac{w}{N+1},$$

where, w and N denote the total waiting time and the number of customers upon arrival, respectively. Consider an exponential distribution with the service rate, μ. Then,

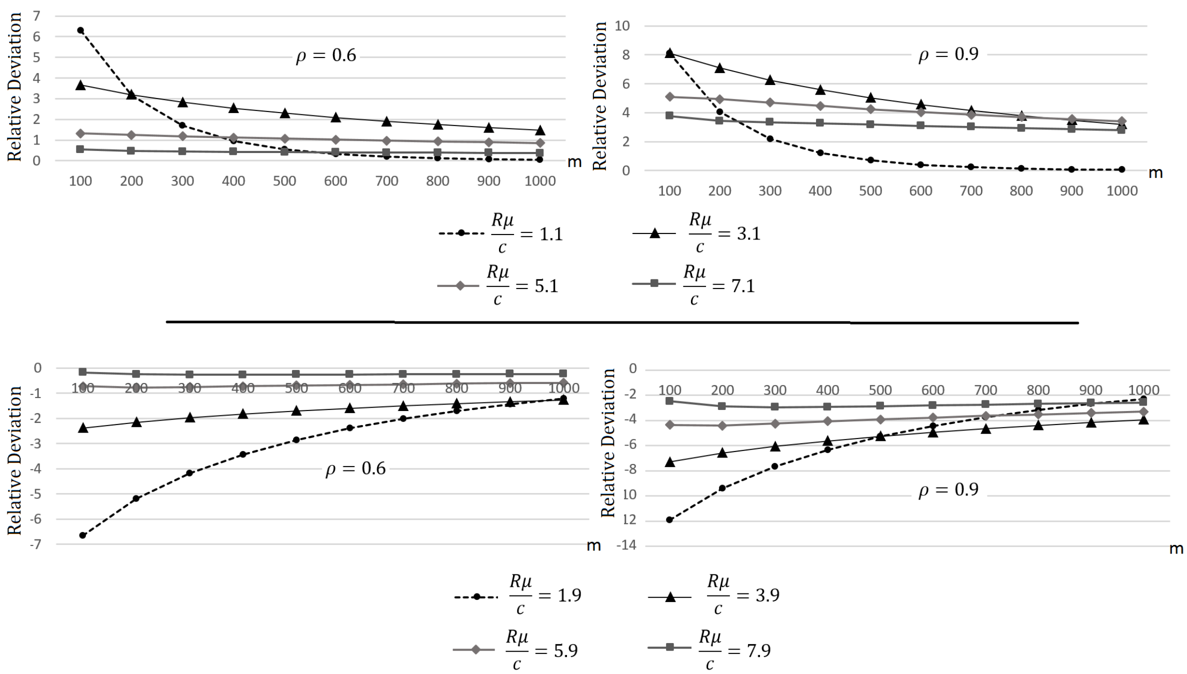


Figure 3.11: The relative deviation of π_0^m from $\pi_0^{M/M/1}$ for various values of $\frac{R\mu}{c}$.

w follows an Erlang distribution with shape parameter $N+1$ and rate μ (see [5]). Note that the shape parameter of this distribution is a random variable, too. If the system is not empty upon arrival, each review may contain the information of the service time of more than one customer. Thus, depending on the congestion of the system, we expect the people who read reviews on w and N to behave like the people who read more reviews on the service rate. Specifically, when the system is more congested, each review contains the information about the service time of more customers which may result in a higher rate of approaching asymptote.

3.4.1 One Review on the Waiting Time and the Number of People in the System

First, assume that each customer, before making a joining decision, picks one review randomly and estimates the service rate as $\hat{\mu} = \frac{N+1}{w}$. In order to obtain the stationary distribution of the number of people in the system, we need to obtain the joining probability when there are i customers in the system. The difference between the joining probability of this system and the one with one review on the service time is the definition of $\hat{\mu}$, which results in a different distribution. Using Bayes' theorem and conditioning on N, the probability of joining in state i, $Pr(i)$, can be written as:

$$Pr(i) = \sum_{q=0}^{\infty} Pr(w \leq \frac{(q+1)R}{(i+1) \cdot c} | N = q) \cdot Pr(N = q), \tag{3.13}$$

where w follows an Erlang distribution with shape parameter, $q + 1$ and rate μ. Note that the sum of exponential random variables with the same rate, follows an Erlang distribution which is a special case of Gamma distribution (see [5]). Let $\pi_q^{1w/n}$ denote the probability that there are q customers in the system, in the steady state. The notation of the superscript indicates that everyone reads one review and each review contains the information about w and N.

Now we have two sets of equations related to each other. The stationary distribution of the number of people in the system is a function of joining probability, and, joining probability itself is a function of the stationary distribution. These two sets of equations are as follows:

$$\pi_q^{1w/n} = \pi_0^{1w/n} \rho^i \prod_{i=0}^{q-1} Pr(j), \tag{3.14a}$$

$$Pr(i) = \sum_{q=0}^{\infty} \pi_q^{1w/n} Pr(w \leq \frac{(q+1)R}{(i+1) \cdot c} | N = q). \tag{3.14b}$$

In order to solve these equations, we begin with an initial value for the stationary distribution. A recommended initial point can be:

$$\pi_q^{1w/n} = \begin{cases} 1, & for\ q = 0 \\ 0, & otherwise. \end{cases}$$

Then, we substitute these initial values in Eq. (3.14b). Next, we find the values of joining probabilities in different states. Then, we substitute the results in Eq. (3.14a). We repeat this procedure till the difference between two consecutive steps, is insignificant. We ran this procedure for $\frac{R\mu}{c} \in \{1, 1.1, 1.2, ..., 5\} \cup \{10, 20, ..., 100\}$ and $\rho \in \{0.1, 0.2, ..., 0.9\}$ to find the probability of idleness. Figure 3.16a illustrates the results.

Figure 3.16a illustrates that when reviewers provide information about both waiting time and then number of people in system upon her arrival, we still can observe results, similar to the one in the system with reviews on the service time. For example, $\pi_0^{1w/n}$ is greater than $\pi_0^{M/M/1}$ whenever $\frac{R\mu}{c}$ is integer.

By comparing Figure 3.16a with Figure 3.1, we observe the same pattern in both $\pi_0^{1w/n}$ and π_0^1. However there are also some differences. In order to focus on differences, we plot the relative deviation of the probability of idleness in each system from the one in the M/M/1 system. Figure 3.13 shows that for a given ρ, $\pi_0^{1w/n}$ is closer to $\pi_0^{M/M/1}$ for a higher $\frac{R\mu}{c}$. This result, also supports the idea of the similar patterns in $\pi_0^{1w/n}$ and π_0^1.

The other performance measure for congestion is the expected number of customers in the system. Let $EL^{1w/N}$ denote the expected number of people in the system when everyone reads one review on w and N. Here, we examine if, for more congested systems,

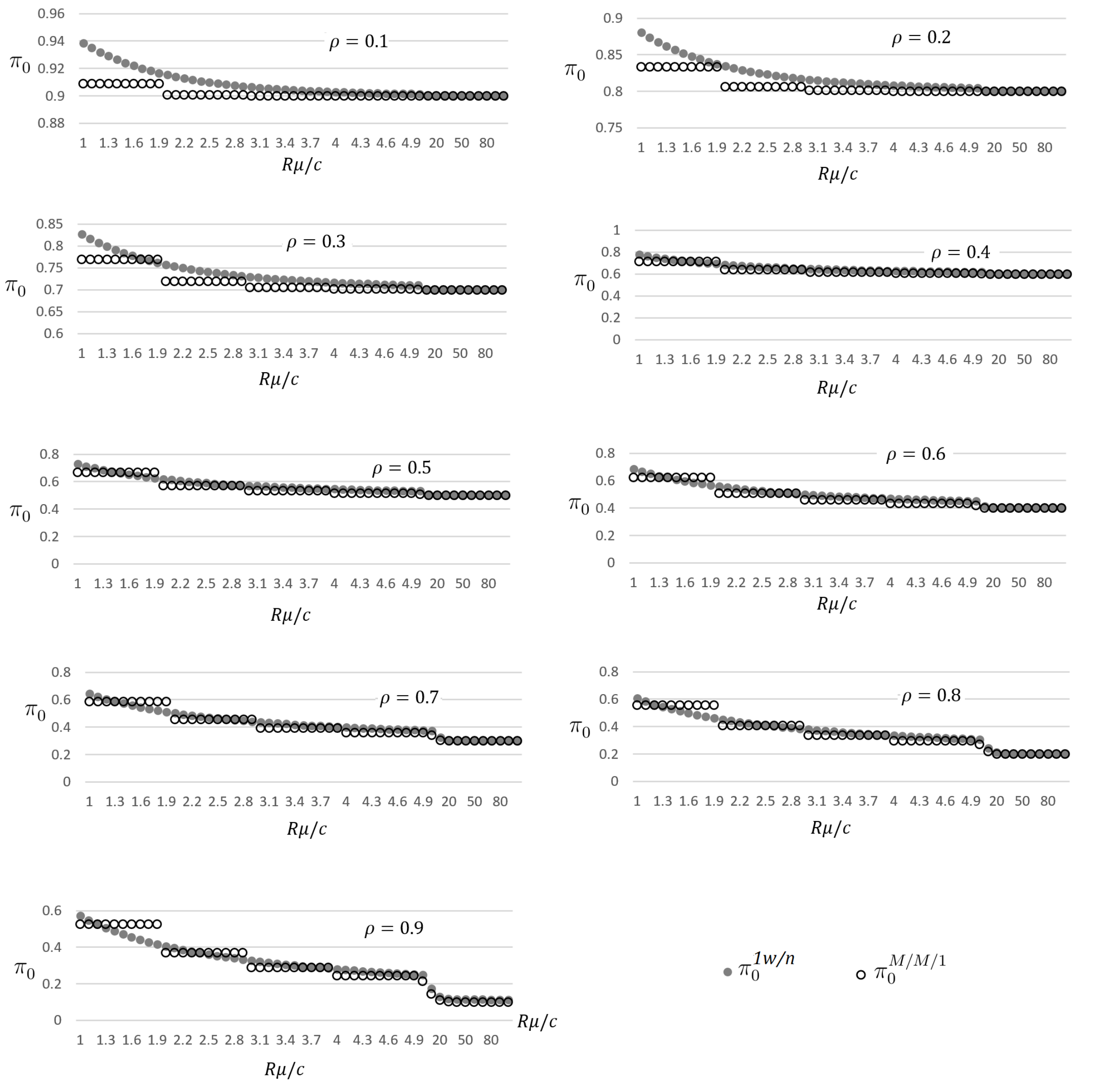


Figure 3.12: A comparison between $\pi_0^{1w/n}$ and $\pi_0^{M/M/1}$ for $\frac{R\mu}{c} \in \{1, 1.1, 1.2, ..., 5\} \cup \{10, 20, ..., 100\}$ and $\rho \in \{0.1, 0.2, ..., 0.9\}$.

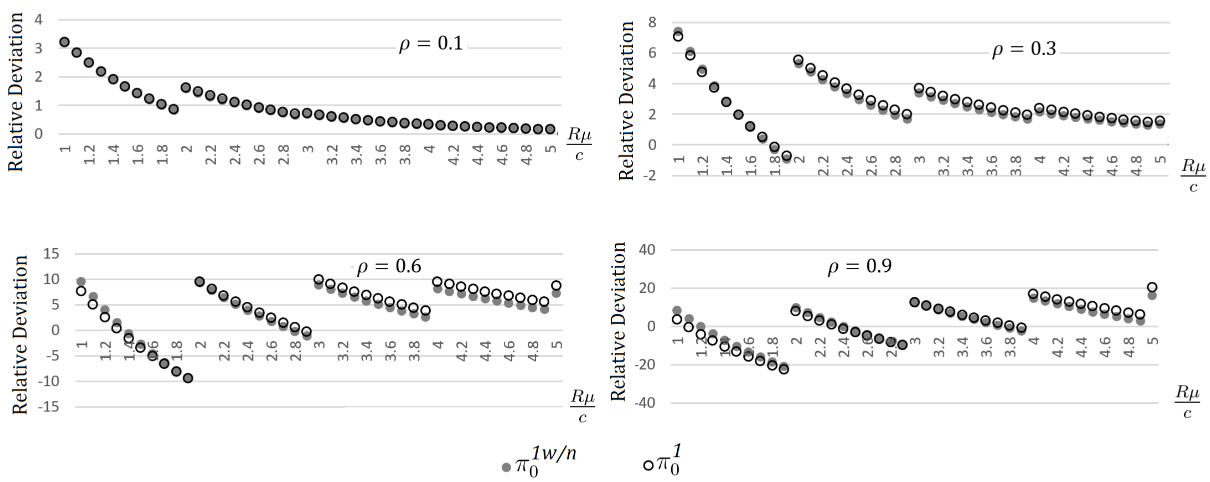


Figure 3.13: A comparison between the relative deviation of $\pi_0^{1w/n}$ and π_0^1 from $\pi_0^{M/M/1}$.

the difference between $EL^{1w/N}$ and $EL^{M/M/1}$ is less than the difference between EL^1 and $EL^{M/M/1}$. For this purpose, we plot the relative deviation obtained by:

$$\frac{EL^{1w/N} - EL^{M/M/1}}{EL^{M/M/1}} \times 100$$

, and we compare that to the relative deviation of EL^1 from $EL^{M/M/1}$. Not that, if the graph is closer to zero, it means that $EL^{1w/N}$ is less deviated from the expected number of customers in the M/M/1 system.

As we expected, Figure 3.14 indicates that when ρ is high, the difference between $EL^{1w/n}$ and EL^1 is higher. The intuition behind this observation is that for more congested systems, reviewers, on average, encounter a longer queue upon arrival. Thus, the estimation of the service rate based on their reviews, is obtained from a larger sample and consequently, the estimation is closer to the exact service rate.

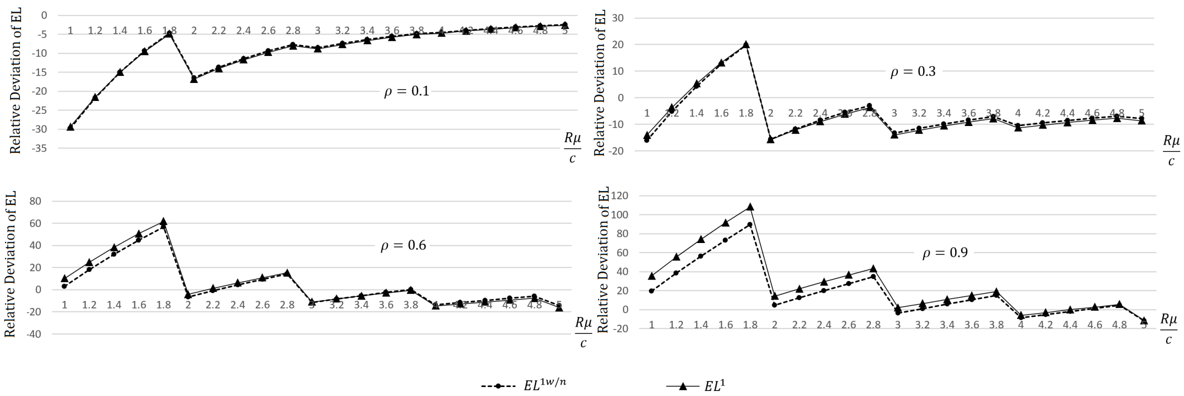


Figure 3.14: A comparison between the relative deviation of $EL^{1w/n}$ and EL^1 from $EL^{M/M/1}$.

3.4.2 Two Reviews on the Waiting Time and the Number of People in the System

We discussed the case in which each customer reads one review on w and N. Here, we investigate how the congestion of the system can be affected, if customers read more reviews. Assume that each customer picks two reviews randomly, and let w_i and N_i denote the waiting time and the number of customers upon arrival indicated in those two reviews, respectively, for $i \in \{1, 2\}$. Then, assume this customer estimates service time by taking the average of the estimations from reviews, i.e.,

$$\frac{1}{\hat{\mu}} = \frac{1}{2}\left(\frac{w_1}{N_1 + 1} + \frac{w_2}{N_2 + 1}\right). \tag{3.15}$$

In other words, the new customer gives the same weight to reviews. We can obtain the joining condition of this new customer, by replacing $\frac{1}{\hat{\mu}}$ in Eq. (3.1) with its value in Eq. (3.15), when she faces i customers upon arrival. Then, the joining probability would be:

$$Pr(i) = Pr\Big(w_1(N_2+1) + w_2(N_1+1) \leq \tfrac{2(N_1+1)(N_2+1)R}{(i+1)c}\Big).$$

The stationary distribution of the number of people in the system, is the same as Eq. (3.14a). Let $\pi_0^{mw/n}$ denote the probability that the system with m reviews on w and N, is in state n in the steady state. In this system, joining probabilities do not follow a Gamma distribution anymore (both parameters are different now). Thus, we try to numerically obtain the stationary distribution using Bayes' theorem and by conditioning on N_1, N_2 *and* w_2. Therefore, the joining probability can be written as follows:

$$Pr(i) = \sum_{N_1=0}^{\infty}\sum_{N_2=0}^{\infty} \pi_{N_1}^{2w/n}\pi_{N_2}^{2w/n}\int_0^{\infty} F_{w_1}\Big(\frac{\frac{2(N_1+1)(N_2+1)R}{(i+1)c} - w_2(N_1+1)}{N_2+1}; N_1{+}1, \mu\Big)\cdot f_{w_2}(w_2, N_2{+}1, \mu)dw_2, \tag{3.16}$$

where $F_x(X, \alpha, \beta)$ and $f_x(X, \alpha, \beta)$ indicate CDF and *probability density function* (PDF) of an Erlang distribution with shape parameter α and rate β at point X, respectively.

In order to solve these sets of equations numerically, we assume that the length of the queue is not more than 20, which seems reasonable for the assumed range of parameters. In addition, we assume that μ is normalized to 1.

We discussed that the probability of idleness in the system with reviews on N and w, may be closer to probability of idleness in the system with more reviews, m, on the service time. We also showed that by increasing m, π_0^m does not necessarily become closer to $\pi_0^{M/M/1}$ when $\frac{R\mu}{c}$ is not large enough. This is the reason that $\pi_0^{2w/n}$ is closer to $\pi_0^{M/M/1}$, just for a large $\frac{R\mu}{c}$.

In order to support the idea that for congested systems, the probability of idleness in the systems with reviews on N and w are close to the one in the system with more reviews

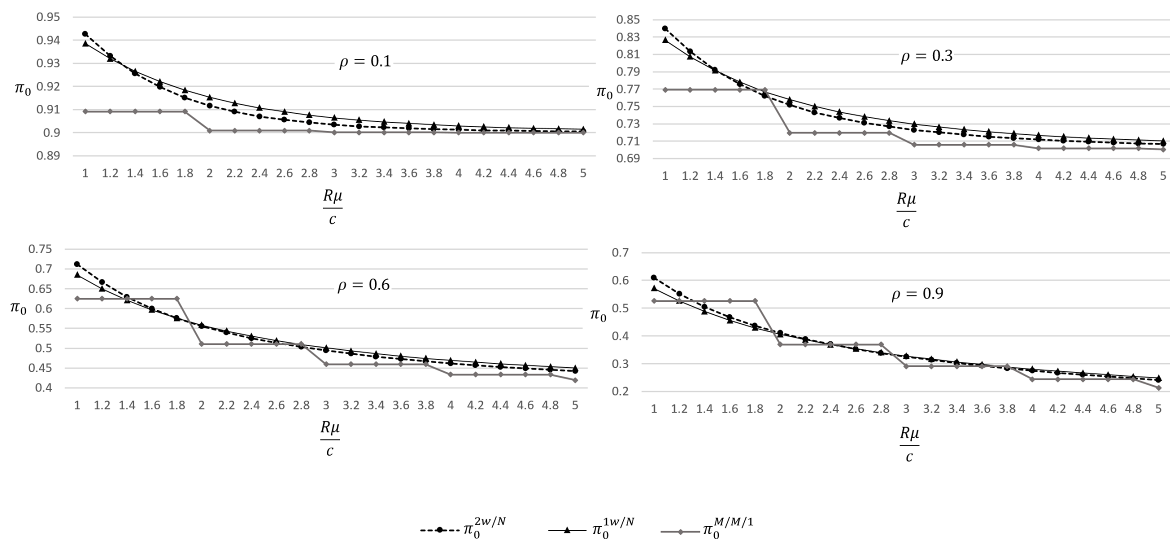
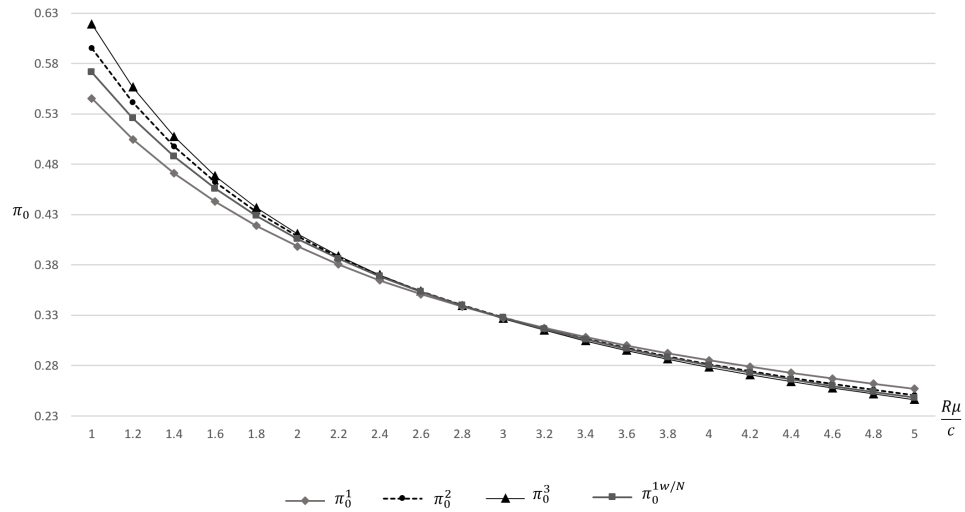


Figure 3.15: The probability of idleness for systems with one and two reviews on w and N for various amounts of ρ and $\frac{R\mu}{c}$.

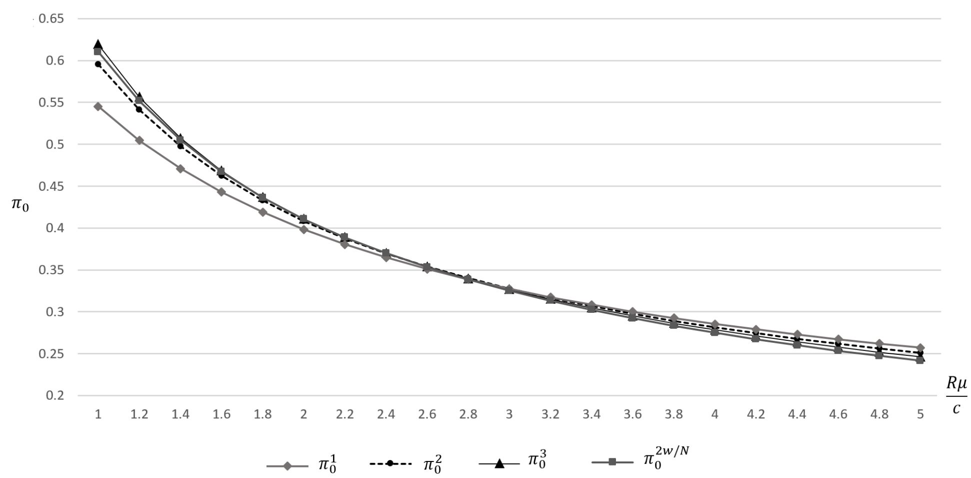
on service time, we plot the probability of idleness for systems with one review and 2 reviews on w and N for $\rho = 0.9$.

Figure 3.16a illustrates that the probability of idleness in the system with one review on w and N, is between the one in the systems with one and two reviews on the service time. For a large $\frac{R\mu}{c}$, it even tends to the one in the system with three reviews on the service time. Figure 3.16b also shows that 2 reviews on w and N yields a probability of idleness, closer to that of the system with 3 reviews on the service time. Based on these results observed in Figure 3.16, when the number of reviews on w and N increases, $\pi_0^{mw/n}$ approaches $\pi_0^{M/M/1}$, with a rate, faster than what π_0^m does.

We also investigate the impact of reading two reviews on w and N, on the expected number of customers in the system, denoted by $EL^{2w/N}$. For this purpose we obtained the relative deviation of expected number of customers from that of the M/M/1 system.



(a) 1 review on w and N



(b) 2 reviews on w and N

Figure 3.16: Comparison of the probability of idleness in systems with different numbers of reviews on the service time and also on w and N

Figure 3.17 illustrates the relative deviation of the expected number of customers in different systems. This figure shows that for a large $\frac{R\mu}{c}$, $EL^{2w/N}$ is the closest to $EL^{M/M/1}$.

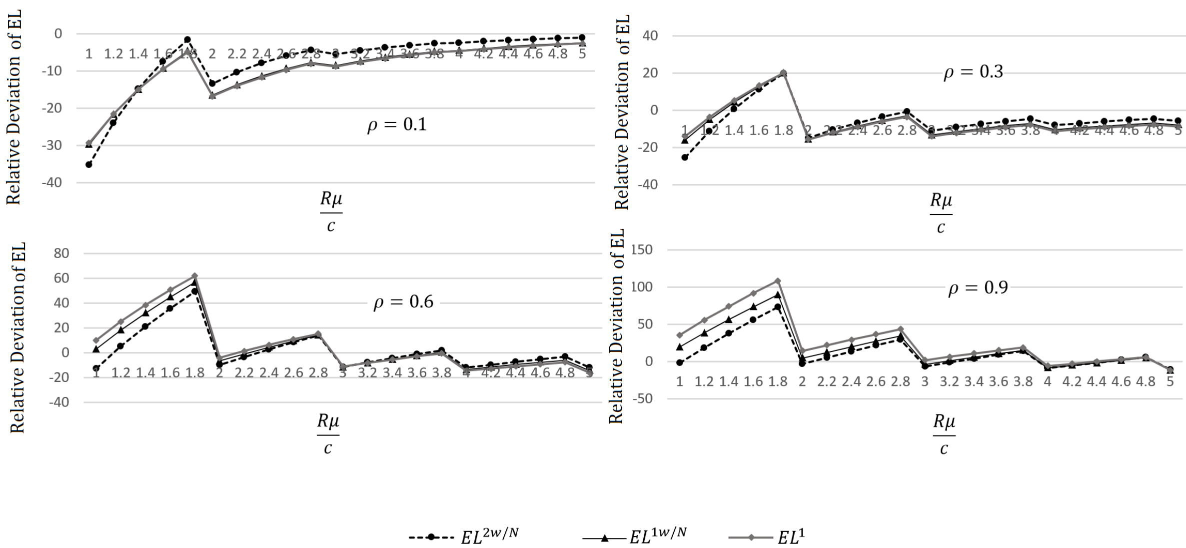


Figure 3.17: A comparison between the relative deviation of $EL^{mw/N}$ and EL^1 from $EL^{M/M/1}$.

Chapter 4

Conclusions and Future Research

Because of the prevalence of using online reviews, it is necessary to determine its impact on almost every system, including queueing systems. In a queueing system, one of the important factors in determining the customers' decision, is their perception of the waiting time. Thus, this question may arise that how this progressing behavior of using reviews of the previous customers, affects a queueing system and under what conditions it enhances the performance of a system.

In this research, we show that reading reviews, can make the performance of the queueing system, either better or worse, depending on the parameters of the system. For example, when the reward that customers gain from service completion, is large enough comparing to the expected cost of waiting for each customer ahead in the queue, i.e., $R \gg c \times \frac{1}{\mu}$, reading reviews results in more idle time. Also when the system is not congested enough, it decreases the revenue.

This study contains some limitations. First, the suggested models could not be compared mathematically. Second, we assume that customers choose among reviews randomly.

This may not be close to reality. For example, customers may just choose among the recent reviews or at least give more weight to recent reviews. In this situation, the method of sorting the reviews matters which should be studied. (See e.g. Mudambi and Schuff [13] and Marley et al. [12]).

We assume that customers reflect their exact service time. The perception of waiting time might be affected by the condition they are waiting in. Baker and Cameron [1], Thompson and Yarnold [16], and, Thompson et al. [17] discuss the impact of the environment on the perception of customers of the waiting time. This perception might lead to biased reviews and consequently, affect the performance of the system.

Bibliography

[1] Baker J., Cameron M. (1996) *The effects of the service environment on affect and consumer perception of waiting time: An integrative review and research propositions* Journal of the Academy of Marketing Science, 24(4), 338-349.

[2] BrightLocal Local Consumer Review Survey 2015, Retrieved from `https://www.brightlocal.com/learn/local-consumer-review-survey/`

[3] Cui S., Veeraraghavan S (2016) *Blind Queues: The Impact of Consumer Beliefs on Revenues and Congestion*. Management Science.

[4] Edelson N.M., Hildebrand D.K. (1975) *Congestion tolls for Poisson queueing processes* Econometrica 43, 81-92

[5] Evans M., Hastings N., Peacock B. (2000). *Erlang distribution.* Ch, 12, 71-73.

[6] Hassin R. (1985) *On the Optimality of First Come Last Served Queues* Econometrica 53, 201-202

[7] Hassin R., Haviv M (2003) *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems* Vol. 59, Springer International Series in Operations Research and Management Science (Springer, New York).

[8] Ho T. H., Park Y. H., Zhou Y. P. (2006) *Incorporating satisfaction into customer value analysis: Optimal investment in lifetime value* Marketing Science, 25(3), 260-277.

[9] Huang T., Allon G.,Bassamboo, A. (2013) *Bounded Rationality in Service Systems* Manufacturing & Service Operations Management, 15(2), 263-267. doi:10.1287/msom.1120.0417

[10] Jin C., Debo L., Iravani S., Kremer M. (2015) *Observational Learning in Environments with Multiple Choice Options: The Wisdom of Majorities and Minorities* Chicago Booth Research Paper, (15-08).

[11] Larsen C. (1998) *Investigating Sensitivity and the Impact of Information on Pricing Decisions in an $M/M/1/\infty$ Queueing Model* International Journal of Production Economics 56-57, 365-377.

[12] Markey R., Reichheld F., Dullweber A. (2009). *Closing the customer feedback loop* Harvard Business Review, 87(12), 43-47.

[13] Mudambi S. M., Schuff D. (2010). *What makes a helpful review? A study of customer reviews on Amazon. com* MIS quarterly, 34(1), 185-200.

[14] Naor P. (Jan., 1969) *The Regulation of Queue Size by Levying Tolls.* Econometrica, 37(1): 15-24.

[15] Sankaranarayanan K., Delgado C., van Ackere, A., Larsen E.R.(2014) *The micro-dynamics of queueing: understanding the formation of queues* Journal of Simulation 8(4), 304–313. doi:10.1057/jos.2014.10

[16] Thompson D. A., Yarnold P. R. (1995) *Relating patient satisfaction to waiting time perceptions and expectations: the disconfirmation paradigm* Academic Emergency Medicine, 2(12), 1057-1062.

[17] Thompson D. A., Yarnold P. R., Williams D. R., Adams S. L. (1996) *Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department.* Annals of emergency medicine, 28(6), 657-665.

[18] Veeraraghavan S. K., Debo L. G. (2011) *Herding in queues with waiting costs: Rationality and regret* Manufacturing & Service Operations Management, 13(3), 329-346.

[19] Zheng W. (2015) *Analysis of Customer Behavior Under Uncertainty on Distribution of Service Time Parameter in M/M/1 Queueing System* 12th International Conference on Service Systems and Service Management (ICSSSM), Guangzhou, pp. 1-5. doi: 10.1109/ICSSSM.2015.7170243, ISSN=161-1890.

APPENDICES

Appendix A

Calculating the Stationary Distribution

A.1 The Stationary Distribution of the Number of People in the System with One Review on Service Time

In order to find the stationary distribution of the number of people in the system with one review on the service time, we need to draw the birth and death process. This process is shown in Figure A.1.

In Figure A.1, let $Pr(i)$ denote the probability of joining in state i. Therefore, in the state i, the fraction of customers who arrive to the system and decide to join is $Pr(i)$. As the arrival rate is denoted by Λ, the rate of going from state i to $i+1$ can be written as: $\Lambda Pr(i)$. On the other hand, in all states, the probability of going from state i to $i-1$ is

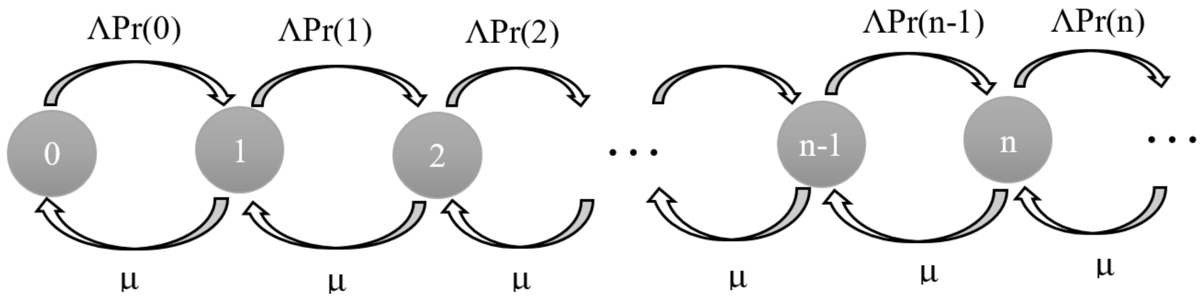


Figure A.1: The birth and death process of the system with one review on service time.

equal to the service rate, μ.

Now, based on the provided birth and death process, we can write the detailed balance conditions as follows:

$\pi_0^1 \times \Lambda \times Pr(0) = \pi_1^1 \mu,$

$\pi_1^1 \times (\Lambda \times Pr(1) + \mu) = \pi_0^1 \times (\Lambda \times Pr(0)) + \pi_2^1 \mu,$

...

$\pi_n^1 \times (\Lambda \times Pr(n) + \mu) = \pi_{n-1}^1 \times (\Lambda \times Pr(n-1)) + \pi_{n+1}^1 \mu.$

Solving for π_1^1 in the first equation, substituting in the second one and continuing this process, and, using induction, we have:

$$\pi_n^1 = \frac{\Lambda^n}{\mu^n} Pr(0).Pr(1)...Pr(n-1)\pi_0^1, \quad for\ n \geq 1. \tag{A.1}$$

The summation of the probabilities of all states should be equal to 1. Then,

$$\sum_{i=0}^{\infty} \pi_i^1 = 1.$$

$\therefore \pi_0^1(1 + \sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} Pr(j)) = 1.$

$\therefore \pi_0^1 = \frac{1}{1+\sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} Pr(j)}$.

$\therefore \pi_n^1 = \frac{\rho^n \prod_{k=0}^{n-1} Pr(k)}{1+\sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} Pr(j)}, \ n \geq 1.$

We already discussed in Eq. (3.2), that:

$$Pr(n) = 1 - e^{-\frac{\mu R}{c(n+1)}}.$$

Thus, we can write the probability of idleness and also the other states, as follows:

$$\pi_0^1 = \frac{1}{1+\sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} [1-e^{-\frac{\mu R}{c(j+1)}}]}, \ n \geq 1.$$

$$\pi_n^1 = \frac{\rho^n \prod_{k=0}^{n-1} [1-e^{-\frac{\mu R}{c(k+1)}}]}{1+\sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} [1-e^{-\frac{\mu R}{c(j+1)}}]}, \ n \geq 1.$$

A.2 The Stationary Distribution of the Number of People in the System with m Reviews on the Service Time

Here, we find the stationary distribution of the number of people in the system, when everyone reads m reviews on the service time and they consider the average as an estimation of the service time.

Let t_f denote the service time provided in review $f \in \{1, .., m\}$. Then, the estimation of the service rate,$\hat{\mu}$ can be written as follows:

$$\hat{\mu} = \frac{m}{\sum_{f=1}^{m} t_f}.$$

In order to find the stationary distribution of the number of people in the system, we just follow the steps in the previous section, as the joining rule and also the death and birth processes, are still the same. The only difference here, is the definition of joining probability when there are n customers in the system, $Pr(n)$. In this case we can write $Pr(n)$ as follows:

$$Pr(n) = Pr\Big(\tfrac{1}{\hat{\mu}} \leq \tfrac{R}{c(n+1)}\Big) = Pr\Big(\tfrac{\sum_{f=1}^{m} t_f}{m} \leq \tfrac{R}{c(n+1)}\Big) = Pr\Big(\sum_{f=1}^{m} t_f \leq \tfrac{mR}{c(n+1)}\Big).$$

As t_f's are Independent and identically distributed random variables from an exponential distribution with parameter μ, $\sum_{f=1}^{m} t_f$ follows an Erlang distribution with shape parameter, m, and scale parameter, μ. Therefore, we can write $Pr(n)$ as:

$$Pr(n) = F_{ER(m,\mu)}\big(\tfrac{mR}{c(n+1)}\big),$$

where, $F_{ER(m,\mu)}(.)$ denotes the CDF of an Erlang distribution with parameters m and μ.

Appendix B

Obtaining the Effective Arrival Rate

In order to find Λ_e, we should obtain the joining probability. In the system with one review on the service time, the joining probability can be obtained by conditioning on the state of the system. As a result:

$$\Lambda_e = \Lambda(Pr(0)\pi_0^1 + Pr(1)\pi_1^1 + ...). \tag{B.1}$$

Substituting the values of the probabilities in different states from Eq. (A.1) in Eq. (B.1), yields:

$$\Lambda_e = \Lambda \times \pi_0^1(Pr(0) + \rho Pr(0).Pr(1) + \rho^2 Pr(0).Pr(1).Pr(2) + ...,$$

$$= \tfrac{\Lambda}{\rho}(Pr(0)\pi_0^1 + \rho Pr(0).Pr(1)\pi_0^1 + \rho^2 Pr(0).Pr(1).Pr(2)\pi_0^1 + ...,$$

$$= \tfrac{\Lambda}{\rho}(\pi_1^1 + \pi_2^1 + \pi_3^1 + ...) = \tfrac{\Lambda}{\rho}(1 - \pi_0^1).$$

$$\therefore \Lambda_e = \mu(1 - \pi_0^1).$$