ON THE SERVICE MODELS FOR DYNAMIC SCHEDULING OF MULTI-CLASS BASE-STOCK CONTROLLED SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BORA KAT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

AUGUST 2005

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Çağlar Güven
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Z. Müge Avşar
Supervisor

Examining Committee Members

Prof. Dr. Ömer Kırca	(METU,IE)	
Asst. Prof. Dr. Z. Müge Avşar	(METU,IE)	
Prof. Dr. Meral Azizoğlu	(METU,IE)	
Assoc. Prof. Dr. Levent Kandiller	(METU,IE)	
Asst. Prof. Dr. Osman Alp	(Bilkent Unv., IE)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname : Bora KAT

Signature :

ABSTRACT

ON THE SERVICE MODELS FOR DYNAMIC SCHEDULING OF MULTI-CLASS BASE-STOCK CONTROLLED SYSTEMS

Kat, Bora
M.Sc., Department of Industrial Engineering
Supervisor: Asst. Prof. Dr. Zeynep Müge Avşar

August 2005, 112 pages

This study is on the service models for dynamic scheduling of multi-class make-to-stock systems. An exponential single-server facility processes different types of items one by one and demand arrivals for different item types occur according to independent Poisson processes. Inventories of the items are managed by base-stock policies and backordering is allowed. The objective is to minimize base-stock investments or average inventory holding costs subject to a constraint on the aggregate fill rate, which is a weighted average of the fill rates of the item types. The base-stock controlled policy that maximizes aggregate fill rate is numerically investigated, for both symmetric and asymmetric systems, and is shown to be optimal for minimizing base-stock investments under an aggregate fill rate constraint. Alternative policies are generated by heuristics in order to approximate the policy that maximizes aggregate fill rate and performances of these policies are compared to those of two well-known Longest Queue and First Come First Served policies.

Also, optimal policy for the service model to minimize average inventory holding cost subject to an aggregate fill rate constraint is investigated without restricting the attention to only base-stock controlled dynamic scheduling policies. Based on the equivalence relations between this service model and the corresponding cost model, it is observed that the base-stock controlled policy that maximizes aggregate fill rate is almost the same as the solution to the service model and cost model under consideration, especially when backorder penalties are large in the cost model as compared to cost parameters for inventory holding or equivalently when the target fill rate is large in the service model.

Keywords: Multi-class, Dynamic scheduling, Base-stock, Fill rate, Cost and service models.

ÖZ

ÇOK SINIFLI, BAZ-STOK DENETİMİNDEKİ SİSTEMLERİN DİNAMİK ÇİZELGEMESİ İÇİN SERVİS MODELLERİ

Kat, Bora
Yüksek Lisans, Endüstri Mühendisliği Bölümü
Tez Yöneticisi: Y. Doç. Dr. Zeynep Müge Avşar

Ağustos 2005, 112 sayfa

Bu tezde iki veya daha fazla sayıda parça tipinin bulunduğu, stok için üretim yapan bir sistemin dinamik çizelgelemesi üzerine çalışılmaktadır. Talep süreçleri Poisson olan parçalar, işleme süresi üstel olan bir tesis tarafından birer birer işlenmektedir. Parça tipleri için stoklar baz-stok denetimindedir ve yoksatma mümkündür. Amaç, parça tiplerinin farklılıklarını göz önüne alarak hesaplanan gelen talebi anında karşılama oranı için bir kısıt altında baz-stoklara yapılan yatırımı veya ortalama stok taşıma maliyetini en aza indirmektir. Karşılama oranını en büyükleyen politika simetrik ve asimetrik sistemler için sayısal olarak araştırılmakta ve bu politikanın karşılama oranı için bir hedef düzey kısıtı altında, baz-stoklara yapılan yatırımı en aza indirdiği gösterilmektedir. Sözü geçen politika sezgisel yöntemlerle yaklaşık olarak belirlenmekte ve bu sezgisel yöntemler sıklıkla kullanılan, parçaları geliş sıralarında işleme ve işlenmek üzere sırada bekleyen parça tiplerinden sayısı fazla olanı öncelikle işleme politikalarıyla karşılaştırılmaktadır.

Ayrıca, sadece baz-stok denetimindeki dinamik çizelgeleme politikaları ile sınırlı kalmadan, karşılama oranı için bir hedef düzey kısıtı altında, ortalama stok taşıma maliyetini en aza indiren politika araştırılmaktadır. Servis modeli ile maliyet modelinin denkliği gözönüne alınarak, yoksatma maliyetinin veya servis modelinde karşılama oranı hedefinin yüksek olduğu durumlarda, karşılama oranı için bir hedef düzey kısıtı altında stok taşıma maliyetini en aza indiren politikanın, karşılama oranını en büyükleyen politika ile hemen hemen aynı olduğu gözlenmektedir.

Anahtar Kelimeler: Çok sınıflı, Dinamik çizelgeleme, Baz-stok, Karşılama oranı, Maliyet ve servis modelleri.

In memory of my beloved grandfathers

ACKNOWLEDGEMENT

First and foremost I would like to offer my thanks to my supervisor Asst. Prof. Dr. Zeynep Müge Avşar. I owe lots of gratitude for her support, dedication and enthusiasm. I've learned so much from her, especially about the meaning of scholarship.

I would also like to thank Prof. Dr. Nesim Erkip and Assoc. Prof. Dr. Geert-Jan van Houtum for their support and encouragement and Prof. Dr. Refik Güllü and Prof. Dr. Nur Evin Özdemirel for their advises.

It was a great pleasure for me to share the same room with Ayşe Koçlar, to tread the same corridor from end to end in desperate times with Oğuz Solyalı, to receive Melih Özlen's logistic support and to answer Fatma Kılınç's never ending questions. Thanks for being my colleagues.

And Halil Özbey. I owe much to discussions we made. Life wouldn't be so easy without his technical support.

Deniz Akhanlı and Ayhun Ünal, my housemates, contributed their friendship with sacrifice.

Especially, I would like to give my special thanks to Özlem Pınar and my family for their love and patience during this period.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This study is on the dynamic scheduling of multi-class make-to-stock systems for the case (s, S) policies with $s = S - 1$, namely base-stock policies, are employed at the stock points. Items are processed one by one by a single exponential server allowing preemption. Demands of different classes (item types) arrive according to independent Poisson processes and demands that cannot be satisfied upon arrival are backordered. The system is called symmetric if the demand rates, cost parameters and base-stock levels are equal for different types of items, and asymmetric otherwise. There are no set-up costs or change-over times to switch from processing one item type to another. It is possible to consider the setting outlined above within the context of manufacturing or spare part systems. One of the design issues to be considered for a given configuration is to determine the base-stock levels for items or spare parts of different types for the respective stock points. High base-stock levels bring extra cost; low base-stock levels, on the other hand, would cause frequent backorders (e.g., interruptions of the activities at the sites where repairable items are in use as in the case of many military applications).

It is common in the literature to work with backorder cost although it is hard, even impossible, to quantify or estimate the corresponding cost parameter. Service levels, on the other hand, represent customer satisfaction more realistically, especially when quality of the estimates of cost parameters for backorders is questioned. That is why specifying some target service levels to be achieved and working with backorder costs are considered as alternative approaches. Inventory models to minimize all relevant costs, including backorder cost, are known as the cost models whereas service models are the ones with some service level constraints introduced instead of incurring backorder cost. There is a number of service levels. Probability of not stocking out (α-service), fill rate (β-service level) and modified fill rate (γ-service level) are the most common ones. α-service level is the fraction of time that demand is satisfied immediately upon arrival, β-service level is the fraction of demand satisfied immediately upon arrival and γ-service level is one minus the ratio of the average backorders and the mean demand per period.

For each service model, there is a corresponding cost model where backorder cost is incurred in accordance with the service level definition in the service model. For example, if γ-service level is considered in the service model, then in the corresponding cost model the backorder cost is for the time needed to satisfy each backordered request. Since α and β type service levels are not for the case of penalizing time needed to satisfy a backorder, use of these in the service models may cause unfair policies to be feasible as pointed out in [3]. In order to avoid this, items of the same type are processed on a First Come First Served (FCFS) basis which is fair in the sense that there is no discrimination between customers of the same class. As clarified in [18], such related models are said to be equivalent under certain conditions.

Most of the research on multi-class make-to-stock M/M/1 queueing systems are to minimize the long-run expected average or discounted inventory holding and backorder costs. Due to the difficulty to quantify cost parameters for backorders, in this study we prefer to concentrate on service models with α or β-service levels. Since Poisson arrivals see time averages, α and β-service levels are equal for each class. Keeping this in mind, we will be calling the service level under consideration as the fill rate throughout the thesis. Instead of working with individual fill rates of different item types, a weighted average of them is introduced in this study as the aggregate fill rate to represent the overall system performance, and structure of the dynamic scheduling policies (state-dependent priorities) is investigated to maximize aggregate fill rate for given base-stock levels. It is observed that the base-stock controlled scheduling policies that maximize aggregate fill rate are characterized by (two) switching curves (for two-class systems). Depending on the weights of the individual fill rates of different item types, the curves shift accordingly. For the symmetric systems, these scheduling policies reduce to the Longest Queue (LQ) policy when there is not any backorder and to the Shortest Queue (SQ) policy when there are backorders of all classes. LQ (SQ) policy awards priority to the class with the highest (lowest) number of outstanding orders, i.e., the class with the lowest (highest) inventory level due to the employment of base-stock policies.

Optimal policies for the following formulations are in the set of policies maximizing aggregate fill rate: minimizing investment on base-stock quantities subject to a constraint on the aggregate fill rate to achieve a target level or maximizing aggregate fill rate subject to a budget constraint for the investment on base-stock quantities. Although we first focus on the base-stock controlled dynamic scheduling policies maximizing aggregate fill rate, our observations are not limited to only resolving the trade-off between base-stock investment and aggregate fill rate. Also a generalized problem formulation is considered under the set of all dynamic scheduling policies (not only of base-stock controlled policies) to resolve the trade-off between expected average inventory holding cost and aggregate fill rate. Basically, what is proposed in this study is to consider the service models with an aggregate fill rate constraint instead of the ones with an individual fill rate constraint for each item type. Replacing the individual fill rate constraints with a weighted average of them is, in fact, a relaxation but brings about the advantage of working with a single service level. Then, the question is how representative a single aggregate service level is for the overall system performance. This question is addressed relating the generalized formulation mentioned above to the corresponding cost model. It is observed that working with aggregate fill rate cannot be regarded as a relaxation when weights for the individual fill rates of different item types and target fill rate are chosen to guarantee equivalence of the generalized service and the related cost models. This way, we can claim more about the base-stock controlled scheduling policy maximizing aggregate fill rate.

Organization of this thesis is as follows: in Chapter 2, related studies are reviewed positioning our work among others in the literature. Chapter 3 is on the problem setting considered in this study and numerical investigation of the structure of the base-stock controlled dynamic scheduling policies maximizing aggregate fill rate for both symmetric and asymmetric two-class systems. Next, in Chapter 4, heuristics are proposed in order to approximate the base-stock controlled dynamic scheduling policies investigated in Chapter 3. Alternative problem formulations are studied in Chapter 5 using the relations between cost and service models so as to reveal the meaning or importance of working with aggregate fill rate for multi-class systems and the use or power of the policy investigated and approximated in Chapters 3 and 4, respectively. Finally, Chapter 6 is on the concluding remarks and further research topics.

CHAPTER 2

REVIEW OF RELATED STUDIES

Most of the studies on multi-class make-to-stock systems are to minimize long-run expected average inventory holding and backorder costs under base-stock control as in [21], [23], [19], [10] and [6]. Long-run expected discounted cost criterion, on the other hand, is considered by Ha [8]. Optimality of base-stock policies is shown in [8] only for a special case and is verified in general by extensive numerical experiments. [3], [7] and [4] on models to minimize long-run average holding costs subject to a service level constraint differ from the references mentioned above. The service level considered in [3], [7] and [4] is of α type. Based on this classification of the studies in the literature, our work can be considered in the class of [3], [7] and [4]. Different from [3], [7] and [4], the (equivalence) relations between the service model primarily focused on and the corresponding cost models are also investigated in this work. In the cost models related to the service model with aggregate fill rate, backorder cost is of either α or β type unlike the γ type backorder cost in [21], [23], [19], [10] and [6]. A reference book that covers many of the issues mentioned above is [24].

First study on multi-class make-to-stock queueing systems is by Zheng and Zipkin [22] for the base-stock controlled symmetric two-class case with independent Poisson processes for the arrivals of different item types and exponential constant processing rate of the single-server facility. The authors study the LQ policy and compare performance of this policy to that of FCFS policy. Allowing preemption under the LQ policy, Zheng and Zipkin obtain closed-form expressions for steady-state probabilities of the differences between queue lengths and for the first two moments of the marginal queue lengths. Moreover, they develop a recursive scheme for calculating the exact joint distribution of queue lengths and marginal queue length distributions. Zheng and Zipkin also show that, with a convex cost (inventory holding and backorder costs) function, LQ policy performs always better than FCFS policy with respect to long-run average cost criterion and support their analytical observations with numerical experiments to compare these two policies in terms of the fraction of demands backordered, average outstanding backorders and average inventory on hand. As an extension, allowing the demand rates to be different for the two classes of items they modify the LQ policy in such a way that even if the queue length for a class is shorter this class is given priority up to a threshold, i.e., the items of one class are processed only if the corresponding queue length is larger than the queue length of the other class by a prespecified constant, Δ. Besides this, Zheng and Zipkin propose an alternative policy for the symmetric case imposing a maximum total inventory level, S^+, while keeping individual maximum inventory (base-stock) levels, S, and argue that it is enough to consider only the case $S^+ = 2S - 1$. One of the heuristics we propose in this study to maximize the aggregate fill rate is similar to this alternative policy.

Wein [21] studies the same problem without restricting his attention only to two-class symmetric systems but allowing each class to have its own general service time distribution and gives an approximation analyzing a Brownian model under heavy traffic conditions. The scheduling policy Wein proposes is to minimize long-run expected average (inventory holding and backorder) cost for both preemptive and non-preemptive cases. According to this policy, the weighted inventory process (sum of the inventory levels weighted by the mean processing times) is monitored to identify the classes in danger of being backordered. Priority is given to the class with largest $b\mu$ index among the classes in danger of being backordered, where b is the backorder cost and μ is the mean service rate. If none of the classes is in danger of being backordered, then the class with the lowest $h\mu$ index is processed where h is the inventory holding cost. The proposed policy is compared with a policy that awards priority to the class with minimum inventory level and FCFS policy using simulation. The results imply that the proposed policy outperforms LQ policy which outperforms FCFS policy. Simulation results are also compared with Brownian approximation.

In [23], Zipkin considers the preemptive symmetric problems under base-stock control with two and more classes where processing times are restricted exponentially distributed random variables. LQ and FCFS policies are compared. He argues that σ, the sum of queue length standard deviations, represents the major behavior of performance of a wide range of systems. He supports his claim showing that an upper bound on the optimal expected average cost (inventory holding and backorder costs) and the optimal cost when normal approximation is used for the steady-state outstanding orders of each class are both product of σ and some other terms. He also proves that his claim is valid when expected average inventory holding cost is minimized subject to a service level constraint under the normal approximation mentioned above. He derives closed-form expression for σ under FCFS policy and gives an approximation for σ under LQ policy. The approximation turns out to be consistent with the exact closed-form expression given in [22] for two-class systems.

Veatch and Wein [19] propose several idleness and index policies for a single-server make-to-stock system with different classes of items allowing preemption under the long-run average (inventory holding and backorder) cost criterion. Idleness policies are to determine when to keep the server idle and index policies are to specify the class to be processed. One of the idleness policies decomposes the system into single-item subsystems having the same utilization as the whole system. The other aggregates the items into a single class, and a third one is based on Brownian approximation to the problem for the lost sales case and another one decomposes the system into single-class subproblems assuming that LQ policy is used. One of the index policies, called Service Time Look Ahead (STLA) policy, considers the expected cost rate after one service time instead of only considering the cost rate of the next state as in the case of fully myopic static priority policies. A second index policy is obtained by value function approximation and the last one is based on restless bandit analysis. Combinations of these index and idleness policies are tested for two and three-class systems, the results are then compared with the optimal values using value iteration. Numerical experiments show that the STLA index and the Brownian idleness policy combination performs best (with a suboptimality less than 7%) and the LQ switching curve combined with the Brownian idleness policy performs well.

Van Houtum et al. [17] investigate the performance of multi-class base-stock controlled symmetric systems under LQ policies using a model slightly different than the ones reviewed above although independent demand arrivals for different classes are still Poisson and the constant service rate is exponential. The authors do not allow preemption and their state description denotes the number of items of each class in the corresponding queue, not the number of items of each class in the system. They consider performance of the system in terms of stockout probability (which is directly related to the fill rate in symmetric systems), the payoff formulation we use in this study is equivalent to the one proposed by van Houtum et al. in [17] for the symmetric case and an immediate extension of it for the asymmetric case. Limiting the difference between queue lengths by a predetermined threshold (truncating the state-space), van Houtum et al. obtain lower and upper bounds on the performance measure, stockout probability, of the system with two variants: threshold rejection and threshold addition. With an arrival for the longest queue, if the difference between the longest and shortest queues exceeds the threshold, then the arrival is rejected in threshold rejection model and an item is added to shortest queue in threshold addition model. The authors prove that these variants, which they analyze by matrix geometric theory, give lower and upper bounds on the stockout probability in the original LQ model.

The numerical experiments in [17] are to compare LQ and FCFS policies with respect to the minimum base-stock levels required to achieve a target service level (one minus stockout probability).

Ha [8] studies two-class make-to-stock systems with a single server in order to investigate structure of the optimal dynamic scheduling policy under the expected discounted (inventory holding and backorder) cost criterion over an infinite horizon allowing parameters to be different for different classes and employing preemptive discipline. For the case exponential processing times are identically distributed among classes, he shows that for some initial inventory levels, the optimal policy is a hedging point (base-stock) policy characterized by two switching curves: one separates the state-space into busy ("produce") and idle ("do not produce") regions and the other specifies the regions for the class of items to be processed. Ha also shows that, for the case of unequal process rates for different classes, a static priority rule (choosing the class with larger $b\mu$ value where b is the cost per backorder per period and μ is the process rate) is optimal (when there are backorders of that class). Three heuristic policies proposed in [8] are for given base-stock levels. One of the heuristics, called $b\mu$ rule, is a static priority rule, which always chooses to process the class with the largest $b\mu$ index among the ones having inventory levels lower than their base-stock levels. Another heuristic, called modified $b\mu$ rule which is similar to the one proposed in [21], is such that index of each class is $b\mu$ except the one with the lowest $h\mu$ where h is the inventory holding cost per item per time unit. Suppose class k is the one with lowest $h\mu$ index, then, index of class k is $b_k\mu_k$ when its inventory level is below ϵ_k, a pre-specified constant, and zero otherwise. Priority is given to the class with the largest index. The last heuristic, called the switching rule, is to approximate switching curves giving priority to the class with the largest $b\mu$ index among the backordered classes. If there is not any backordered item, then the class with largest $b\mu(1-\frac{x}{S})$ index is given priority where $(1-\frac{x}{S})$ is interpreted as the proportion of unfilled base-stock, x and S denoting the current inventory and the base-stock level respectively. This heuristic policy turns out to be (state-dependent) longest queue policy for the symmetric case.

[10] is on multi-class make-to-stock systems with the objective function minimizing long-run expected average (inventory holding and backorder) cost. In this study, Peña-Perez and Zipkin propose heuristics similar to STLA in [19], based on myopic allocation. One heuristic chooses the class that improves the overall cost rate, which is cost rate change per time unit, considering the service time and the other considering the sojourn time (omitting all other classes and devoting the server to associated

class). For symmetric systems, the former is consistent with LQ policy, which is optimal for this case. The heuristic policies proposed are compared with the optimal policy and with the heuristic in [21] estimating some parameters by simulation. The experiments in [10] are for two-class problems with Poisson demand arrivals and exponential service times. The myopic heuristic working with the sojourn time performs very close to the optimal policy especially for the case of equal service rates.

de Vericourt et al. [6] investigate structure of the optimal policy for two-class systems by coupling and sample path comparison techniques. They give a closed-form expression for the switching curve which is a straight line in the region where the class with lower $b\mu$ index is backordered and also provide the condition for this line to pass through the origin which implies a policy giving priority to the backordered class and choosing the class with higher $b\mu$ index if both are backordered. This characterization allows generalizing results in [8] and helps to show that one of the myopic policies in [10] is consistent with the optimal policy in the stated region. Later as an extension of the work in [6], Veatch and de Vericourt [20], using coupling trajectories, set the conditions for zero inventory policy, which implies the hedging point at the origin. Obviously, the system becomes a make-to-order system when these conditions are satisfied, and then a full characterization of the optimal policy is provided referring also to the results in [8].

Bertsimas and Paschalidis [3] consider multi-class make-to-stock systems with any stationary stochastic demand and service processes.This allows them to accommodate autocorrelated demand and service processes, and this way to broaden their scope to cover also failure-prone systems. These constraints assure that specified levels of stockout probabilities are not exceeded for any of the classes. Using a fluid model, obtained by translating the deterministic version of the problem into continuous-time, Bertsimas and Paschalidis approximate the system in order to determine the scheduling policy; and using large deviations techniques, they obtain approximations for the hedging points. The policies they handle are based on fixed static priorities, the class with the highest priority is processed until its inventory level drops below its hedging point. Given fixed priorities, the authors find hedging points, that guarantee the stockout probability of each class to be less than a pre-specified amount using an approximate expression for the expected queue lengths, and report that the approximate hedging points turn out to be very close to the ones obtained by simulation. Trying all permutations, of priority orderings, the authors propose to choose the best priority policy with the smallest cost. The numerical experiments show that these approximations work well when compared to simulation results.

[7] is on the problem of allocating production capacity among several classes for two different settings under base-stock control: one to minimize long-run average inventory holding and backorder costs and the other to minimize expected holding costs subject to a service-level (stockout) constraint. The interarrival times of demands are i.i.d. and each demand arrival may represent a request of a number of items of each class. There is a maximum production rate defining the overall capacity of the manufacturer. Glasserman claims that the approach, allocating production capacity among classes, is realistic where there are distinct facilities. In the case of a single facility, on the other hand, the model can be considered as the allocation of the time for each class, "*in much the same way that the processor sharing-model in queueing theory is commonly taken as an approximation of round and robin discipline, but although this approach is reasonable to allocate the overall effort on each class, it may be insufficient in sequencing issue*". Besides this, the issues such as determination of base period, i.e., duration of a cycle, and adjustments to be made when base-stock level of a class is reached before time devoted to that class are not clear in the case of a single facility.

Milito and Levy [9] consider a computer communication network with stations competing for a single channel. The stations have limited storage capacity and the packets of information sent to stations from independent sources are blocked when this capacity is full. The objective is to avoid blocked queues, which represent limited capacities of stations. The system explained above is called a blocking system. A similar system with a single central server feeding stations, a starvation system, and another one that contains both blocking and starvation systems, a hybrid system, are also introduced. In starvation systems, the stations again have limited storage capacities and the objective is to avoid empty queues. In [9], starvation system is analyzed, and blocking and hybrid systems are presented as straightforward extensions. In the starvation system, the service times in stations and the time that it takes for the central server to feed a station are both exponentially distributed with station-dependent parameters. For each station there is a cost incurred per unit of time when associated queue is in starvation. The objective is to find a policy that minimizes the long-run expected discounted cost. Symmetric two-station system is analyzed in the first place. Monotonicity of the value function with respect to queue lengths, existence of monotonic switching curves, for both finite and infinite storage capacities, are shown and the relations between the cost parameters and shape of the optimal switching curves are investigated. It is conjectured that, any closed-form policy is unlike to be found, and then an index type heuristic rule is proposed using the results for single-station system. Numerical experiments for asymmetric three-station systems show that the proposed policy performs quite well, relative error is less than 0.04% and 5% for discount factors of 0.90 and 0.998, respectively. The problem considered is the reminiscent of the ones in multi-class make-to-stock queueing systems with lost sales where stations and central computer (channel in the blocking case) play the same role as the classes and the single server facility, respectively. But, note that there is not a cost term that corresponds to the inventory holding cost.

The study by Boyacı and Gallego [4] is on single-item base-stock controlled multi-stage serial systems with backorders but it is similar to ours in the sense that the objective is to minimize average inventory holding costs subject to a service-level (fill rate) constraint. The authors give a lower bound on the total system stock and upper bounds on both the system stock and on the stock level at each stage. Three heuristic procedures are proposed to determine the base-stock levels and one other to find the optimal base-stock levels. Number of stages that can hold inventory is restricted in the first heuristic. Second and third heuristics are based on majorization. Another heuristic proposed in [4] is to solve a cost model that guarantees some specified service level. Second and third heuristics are observed to outperform the first and the last heuristics.

Some other studies to be referred to are the following: Rosling [11] studies single-item inventory systems under different review types and demand structures and derives the conditions for shortage costs, which are related to β and α-service levels, to be quasi-convex. His analysis is to extend optimality of (s, S) and (R, nQ) policies to mentioned cost structures. Axsäter [1] considers a single-echelon system controlled by a continuous-review (R, Q) policy. The author proposes a simple two-step procedure to minimize holding and ordering costs under a fill rate constraint. The system in [12] is the same as in [4]. Shang and Song [12] propose effective closed-form approximations for the optimal base-stock levels at each stage. Sleptchenko et al. [14] study multi-class M/M/k systems with preemptive static priorities. They carry out an exact analysis and devise a solution procedure to calculate the steady-state probabilities. van der Heijden et al. [16], then, propose approximation procedures for the problem in [14] due to the high computational times of the exact solution procedure for systems with many servers and classes.

To wrap up, this study can be positioned somewhere among the ones on service models. However, we use aggregate fill rate as the service measure, which simplifies dynamic scheduling as compared to employing individual service level constraints for different item types. In the service models with the individual service level constraints, optimization is considered over the class of static priority rules, which are simpler to handle than the dynamic scheduling. The cost models related to the service model we consider are different than the cost models in literature. γ type backorder cost is incurred in the latter unlike α or β type costs incurred in our case in accordance with the use of (aggregate) fill rate in the related service model.

CHAPTER 3

FILL RATE MAXIMIZATION: NUMERICAL INVESTIGATION OF THE OPTIMAL POLICY

In this chapter, characteristics and mechanics of the multi-class base-stock controlled systems are presented introducing the notation and the mathematical (dynamic programming) formulation for dynamic scheduling of these systems (for finding the state-dependent priorities to process items of different types) is given to maximize aggregate fill rate (a weighted average of fill rates of different classes) for given base-stock levels. Structure of the policy with maximum aggregate fill rate is numerically investigated. Questioning this approach, i.e., maximizing fill rate for given base-stock levels, is deferred to Chapter 5. That is, the problem formulations for which maximizing aggregate fill rate would serve the propose of resolving the trade-off between base-stock investment and target fill rate are considered in Chapter 5.

3.1 MODEL FOR DYNAMIC SCHEDULING

The system with I different types of items is depicted in Figure 3.1. The case of single-server facility is considered in the first place (instead of a network of servers) as a building block of more generalized cases. Processing time for the items is exponentially distributed with mean $\frac{1}{\mu}$ independently of the item type, and preemption is allowed. S_i is the base-stock level for the inventories of type i. Demands for different item types occur according to independent Poisson processes with rates λ_i, $i = 1, ..., I$, and are met from the respective stock if there is available inventory (i.e., $\overline{n}_i > 0$). When demand of an item arrives (an item fails) in the case of a manufacturing (spare part) system, raw material (this failed item) is sent to the queue of orders (failed items) of the requested type to be processed at the manufacturing (repair) facility and to replenish the respective stock. In the case of a manufacturing system, it is assumed that there is always available raw material. n_i denotes the number of items to be processed and k_i is the number of backordered requests, $i = 1, ..., I$. Base-stock policies imply the following inventory balance equations: $n_i + \overline{n}_i = S_i + k_i$, $i = 1, ..., I$, note that $\overline{n}_i \cdot k_i = 0$. That is, $\mathbf{n} = (n_1, ..., n_I)$ fully describes the system under consideration.

Instead of working with continuous-time Markov chain formulation, we proceed with an equivalent discrete-time version of it by defining the corresponding transition matrix as $P = I + Q$, where Q is the generator matrix of the continuous-time Markov chain. Note that the steady-state probabilities for these Markov chains are the same[1].

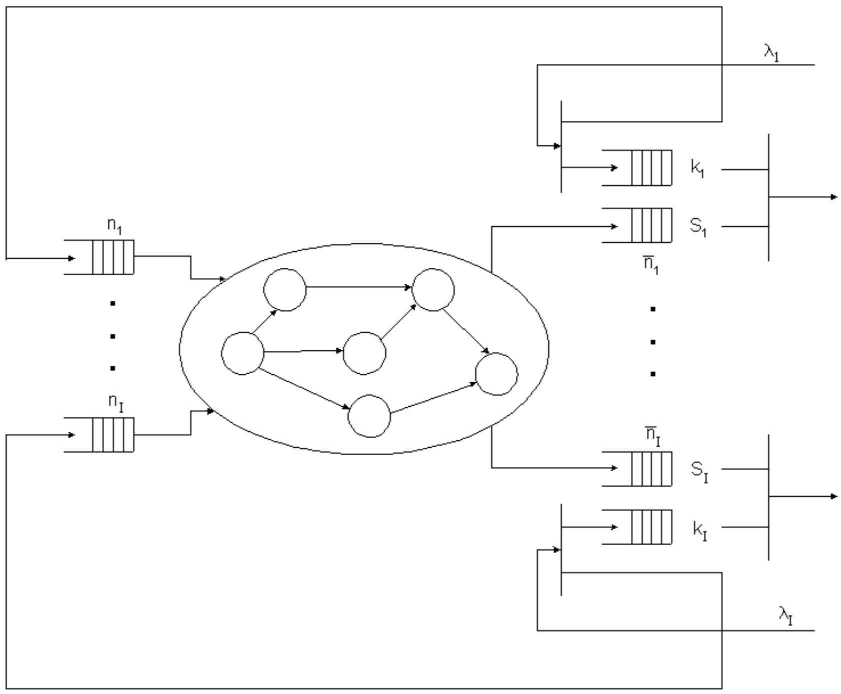


Figure 3.1: System with I item types.

Defining $f_m(\mathbf{n})$ as the minimum total cost over m periods when the initial state is $\mathbf{n}$, the recursive (multi-period) formulation given below is to determine the item type to be processed in state $\mathbf{n}$ when there are m periods to go until the end of the planning horizon.

$$
\begin{aligned}
f_m(\mathbf{n}) &= c(\mathbf{n}) + \sum_{i=1}^{I} \frac{\lambda_i}{\tau} f_{m-1}(\mathbf{n} + \mathbf{e}_i) + \frac{\mu}{\tau} v_{m-1}(\mathbf{n}) && (3.1)\\
f_0(\mathbf{n}) &= 0, && (3.2)
\end{aligned}
$$

for all $\mathbf{n}$, where $\tau = \sum_{i=1}^{I} \lambda_i + \mu$, $\mathbf{e}_i$ is the unit vector with i^{th} entry being equal to 1

[1]Let $\mathbf{p}$ and $\acute{\mathbf{p}}$ denote the steady-state distributions for the continuous-time and discrete-time Markov chains, then $\mathbf{p}Q = \mathbf{0}$ and $\acute{\mathbf{p}} = \acute{\mathbf{p}}P$, respectively. If $P = I + Q$, then $\acute{\mathbf{p}} = \acute{\mathbf{p}}(I + Q)$ which holds only if $\acute{\mathbf{p}}Q = \mathbf{0}$.

and

$$v_{m-1}(\mathbf{n}) = \begin{cases} f_{m-1}(\mathbf{n}) & \text{if } \mathbf{n} = \mathbf{0}, \\ \min_{i \ni n_i \neq 0} \{f_{m-1}(\mathbf{n} - \mathbf{e}_i)\} & \text{otherwise.} \end{cases}$$

Since it is always possible to define the time-scale, without loss of generality assume that $\tau = 1$. Then, in (3.1) λ_i stands for the probability that the first event to occur is a demand arrival of type i, $i = 1, ..., I$, and μ stands for the probability that the first event to occur is service completion for the item in service. Note that processing the item, say of type j, in service may be preempted and the server may switch to another item type, say type i, upon a demand arrival of this type based on the minimization in $v_{m-2}(\mathbf{n} + \mathbf{e}_i)$ of $f_{m-1}(\mathbf{n} + \mathbf{e}_i)$.

Let random variable $N_i(t)$ denote the number of items in the queue of type i in period t. Then,

$$f_m(\mathbf{n}) = \min_{\pi \in \overline{\Pi}_{\mathbf{S}}} \{E_\pi(\sum_{t=1}^{m} \sum_{i=1}^{I} w_i 1_{\{N_i(t) \geq S_i\}} | \mathbf{N}(0) = \mathbf{n})\},$$

where

$$1_{\{N_i(t)\}} = \begin{cases} 1 & \text{if } N_i(t) \geq S_i, \\ 0 & \text{otherwise,} \end{cases}$$

is to count the periods in stockout state for item type i and w_i is the corresponding weight and $\overline{\Pi}_{\mathbf{S}}$ is the set of dynamic scheduling policies under base-stock control given $\mathbf{S} = (S_1, ..., S_I)$. This shows that the cost function, $c(\mathbf{n})$, is a weighted average of the number of item types in stockout state, to be introduced as an extension of the one used in [17] with equal weights for the symmetric case ($\lambda_i = \lambda$ and $S_i = S$ for all $i = 1, ..., I$). That is, $c(\mathbf{n}) = \sum_{i=1}^{I} w_i 1_{\{n_i \geq S_i\}}$. Minimizing the long-run average cost working with the formulation in (3.1) and (3.2) is, then, equivalent to maximizing the weighted average of the fill rates, i.e., $1 - c(\mathbf{n}) = \sum_{i=1}^{I} w_i 1_{\{n_i < S_i\}}$, the fraction of demand satisfied upon arrival (β-service level). Note that since the demand process for each item type is a Poisson process independent of the other item types, fill rate for type i is not any different than the fraction of time stock i is not empty (α-service level), i.e., the long-run probability that random variable N_i is strictly less than S_i. Recall from the inventory balance equation implied by the base-stock policies that $N_i < S_i$ means $\overline{N}_i > 0$, i.e., there are available items in stock.

Although the use of aggregate fill rate instead of penalizing backorders in the objective function of a cost model is justified by the difficulty in estimating the cost parameters for backordered items (see Chapter 5 for details of different problem formulations: optimizing expected average inventory holding and backorder cost versus optimizing expected average inventory holding cost subject to fill rate constraint), with this approach determining the values of w_i turns out as an issue except for the symmetric case where aggregate fill rate is the same as the fill rate of each item type (just the regular average with equal weights). That is, for the symmetric systems our approach immediately comes up as an alternative one directly comparable to the ones in the literature like [17] and [22]. For the asymmetric systems, these weights can be chosen proportional to demand rates, i.e., $w_i = \frac{\lambda_i}{\sum_{j=1}^{I} \lambda_j}$, as pointed out in [18]. This means that without differentiating item types, we consider the fraction of demand satisfied upon arrival. Note that, in this case, aggregate fill rate is $\frac{\sum_{i=1}^{I} \lambda_i Pr(\overline{N}_i > 0)}{\sum_{i=1}^{I} \lambda_i}$. In addition to proposing this choice of weights proportional to the demand rates, we should note that w_i values should be determined depending on how the backorders are penalized, what characteristics of the customer classes are important in this sense. Combining all such relevant factors, analytical results are given as to how to choose w_i values in Chapter 5 based on the (equivalence) relations between cost and service models.

In order to numerically investigate the structure of the dynamic scheduling policy maximizing aggregate fill rate, two-class system is considered first for the sake of simplicity. The formulation in (3.1), (3.2) reduces to

$$\begin{aligned} f_m(n_1, n_2) &= c(n_1, n_2) + \frac{\lambda_1}{\tau} f_{m-1}(n_1 + 1, n_2) + \frac{\lambda_2}{\tau} f_{m-1}(n_1, n_2 + 1) \qquad (3.3) \\ &\quad + \frac{\mu}{\tau} v_{m-1}(n_1, n_2), \\ f_0(n_1, n_2) &= 0, \qquad (3.4) \end{aligned}$$

for all n_1, n_2, where $\tau = \lambda_1 + \lambda_2 + \mu$ and

$$v_{m-1}(n_1, n_2) = \begin{cases} min\{f_{m-1}(n_1 - 1, n_2), f_{m-1}(n_1, n_2 - 1)\} & \text{if } n_2 > 0,\ n_1 > 0, \\ f_{m-1}(n_1 - 1, n_2) & \text{if } n_2 = 0,\ n_1 > 0, \\ f_{m-1}(n_1, n_2 - 1) & \text{if } n_1 = 0,\ n_2 > 0, \\ f_{m-1}(n_1, n_2) & \text{if } n_1 = 0,\ n_2 = 0. \end{cases}$$

Regarding the analysis of the recursive formulation, intuitively immediate observations are that f is nondecreasing in n_1, n_2 and m, and it is symmetric when $\lambda_i = \lambda$ and $S_i = S$ for $i = 1, 2$. See Appendix A for proofs. These observations are not enough for proving optimality (fill rate maximization) of the policy structure investigated studying a wide range of numerical examples in this chapter.

In Section 3.2, the solution approach is presented. The succeeding two sections are for the analysis of the structure of the dynamic scheduling policies maximizing aggregate fill rate (we will call this as ς policy throughout the thesis) in both symmetric and asymmetric systems.

3.2 SOLUTION APPROACH

In order to solve the recursive formulation presented above for the long-run average cost, i.e., for the probability of stockout, value-iteration algorithm is used (see Tijms [15] for an overview of the value-iteration algorithm). Note that the Markov chain formulated is irreducible and aperiodic (when n_1=n_2=0, there is a probability of staying in the current state for one more period) for all stationary dynamic scheduling policies under base-stock control, and value-iteration algorithm works for m being sufficiently large to approximate the long-run average payoff for every initial state (n_1, n_2), i.e., $\lim_{m\to\infty}[f_m(n_1,n_2) - f_{m-1}(n_1,n_2)] = 1 - FR$ where FR is the aggregate fill rate for the given λ_i, μ and w_i values. Recalling the definition of $f_m(\mathbf{n})$ in Section 3.1, aggregate fill rate under policy π, to be denoted by $FR(\pi)$ is

$$\begin{aligned}
&\lim_{m\to\infty}\frac{1}{m}E_\pi(\sum_{t=1}^{m}\sum_{i=1}^{I}w_i 1_{\{N_i(t)<S_i\}}|\mathbf{N}(0)=\mathbf{n})\\
&=\lim_{m\to\infty}\frac{1}{m}\sum_{i=1}^{I}w_i E_\pi(\sum_{t=1}^{m}1_{\{N_i(t)<S_i\}}|\mathbf{N}(0)=\mathbf{n})\\
&=\sum_{i=1}^{I}w_i FR_i(\pi)
\end{aligned}$$

where $FR_i(\pi)$ is the fill rate for items of type i under policy π.

Since we cannot work with infinite state space, one may think of some kind of truncation which would mean approximation. Instead, we proceed with the idea of narrowing the state space (not evaluating the functional value at some states) as the number of remaining periods decreases, this way eliminating any requirement for truncation. A sketch of this approach is given in Figure 3.2 and a pseudocode of the procedure is in Appendix B. If the functional value is to be evaluated using (3.3), (3.4) over the state space {0,1,...,No_1} $\times$ {0,1,...,No_2} with m'

periods to go until the end of the planning horizon, we need the functional values for the states in $\{0,1,...,No_1 + 1\} \times \{0,1,...,No_2 + 1\}$ with $(m' - 1)$ periods to go due to the recursive nature of the formulation. This means that, initially we need the functional values for state space $\{0,1,...,No_1 + m'\} \times \{0,1,...,No_2 + m'\}$ with no periods to go (at the end of the planning horizon). We start with state space $\{0,1,...,No_1 + m'\} \times \{0,1,...,No_2 + m'\}$, $f_0(n_1, n_2)$ being zero for every (n_1, n_2) in this state space, then continue with $\{0,1,...,No_1 + m' - 1\} \times \{0,1,...,No_2 + m' - 1\}$ to calculate $f_1(n_1, n_2)$ for every (n_1, n_2) of this smaller state space. Proceeding this way, we calculate $f_{m'}(n_1, n_2)$ for every (n_1, n_2) in the state space $\{0,1,...,No_1\} \times \{0,1,...,No_2\}$. m' should be sufficiently large to observe convergence and No_1, No_2 values should be large enough to figure out structure of the optimal (ς) policy. Complexity of the procedure is $O(m'^3)$.

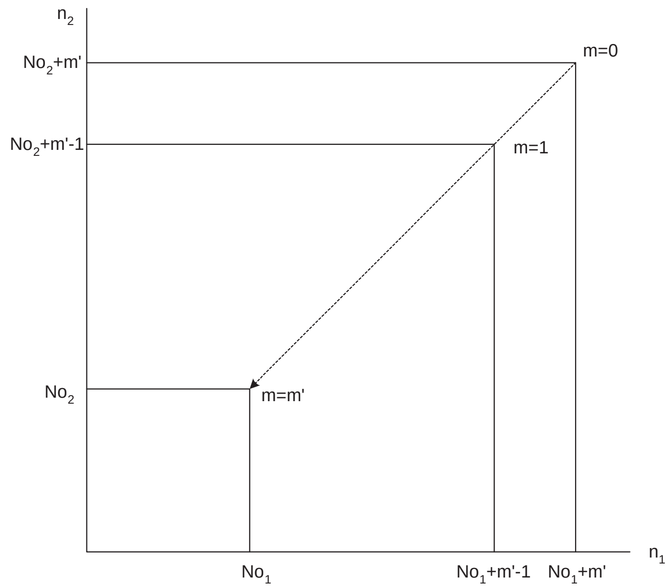


Figure 3.2: Sketch of the solution approach for value-iteration algorithm.

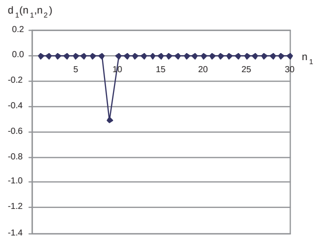
From the numerical experiments, it is observed that the convergence time is increasing in traffic intensity, $\rho = \frac{\lambda_1+\lambda_2}{\mu}$, which is obvious. When ρ is small, the processing rate is much higher than the demand rates and the process is restricted to fewer states, compared to the cases of larger ρ values. For large ρ values such as 0.95, complete convergence could not be observed for the asymmetric cases with the specified ranges to stop although the successive f_m values turn out to be still very close to each other. We have the increasing and decreasing rates of average costs for different initial states. By using these rates, we approximate the convergence points where these costs would meet. This approximation is tested for the experiments which had already converged and it is observed to perform well. In all the experiments in this thesis instead of given μ and λ_i, $i = 1, ..., I$, values, $\frac{1}{1+\rho}$ and $(\frac{\lambda_i}{\sum_{j=1}^{I} \lambda_j}\rho)/(1+\rho)$, $i = 1, ..., I$, are used respectively, so that $\tau = 1$.

3.3 SYMMETRIC SYSTEMS

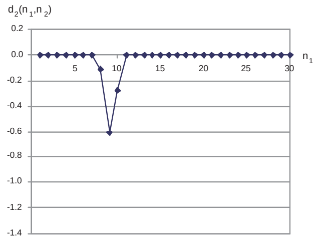
In this section, symmetry considered is the following: $S_i = S$ and $\lambda_i = \lambda$ for $i = 1,\ 2$. That is, aggregate fill rate is the same as the fill rates of each class regardless of the (w_1, w_2) values.

Since the given recursive formulation is to determine which item type to process in each state (n_1, n_2) when there are m periods to go (recall the definition of v_{m-1}, the alternatives for optimization are $f_{m-1}(n_1 - 1, n_2)$ and $f_{m-1}(n_1, n_2 - 1)$), let $d_m(n_1, n_2)$ and $D_m(n_1, n_2)$ denote the differences $f_{m-1}(n_1 - 1, n_2) - f_{m-1}(n_1, n_2 - 1)$ and $d_m(n_1, n_2) - d_m(n_1 - 1, n_2)$, respectively. In Figure 3.3, $d_m(n_1, n_2)$ is plotted against n_1 for $n_2 = 5$, $\rho = 0.80$, $S = 9$ and $m = 1,\ 2,\ 3,\ 5,\ 10,\ 50,\ 100,\ 1000$ and graphs for $D_m(n_1, n_2)$ can be seen in Figure 3.4. Some other example graphs, for different values of n_2 and ρ, can be seen in Appendix C. Our observations from these numerical experiments are as follows:

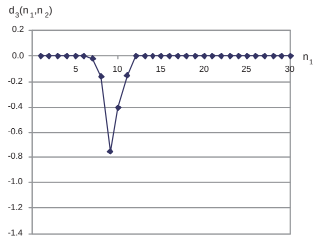
- $d_m(n_1, n_2)$ is non-increasing (non-decreasing) in n_1 (n_2) when $n_1 < S_1$ ($n_2 < S_2$) and the rate of decrease is 0 for low values of n_1 (n_2),
- $d_m(n_1, n_2)$ is non-decreasing (non-increasing) in n_1 (n_2) when $n_1 \geq S_1$ ($n_2 \geq S_2$) and rate of increase converges to 0 as n_1 (n_2) goes to infinity,
- Negative (positive) values of $d_m(n_1, n_2)$ implies that $f_m(n_1 - 1, n_2)$ is less (larger) than $f_m(n_1, n_2 - 1)$, then type 1 (2) should be chosen to process.
- For a given n_2, $d_m(n_1, n_2)$ intersects $n_1 - axis$ at two points, say $\tilde{n}_1$ and $\hat{n}_1$, $\tilde{n}_1 = n_2$. For $n_2 < S_2$ ($n_2 \geq S_2$), processing type 2 is optimal when $n_1 < \tilde{n}_1$ ($n_1 < \hat{n}_1$). Between $\tilde{n}_1$ and $\hat{n}_1$, it is optimal to process type 1 and beyond $\hat{n}_1$ ($\tilde{n}_1$), it is optimal to process type 2. Note that $\tilde{n}_1 < \hat{n}_1$ ($\hat{n}_1 < \tilde{n}_1$) when $n_2 < S_2$ ($n_2 \geq S_2$).



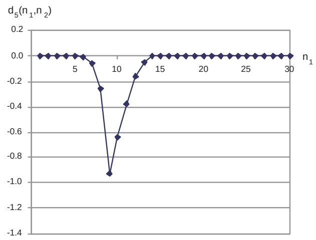
(a) $m = 1$.



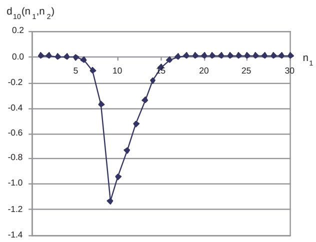
(b) $m = 2$.



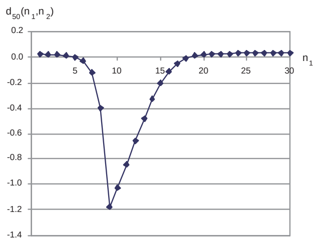
(c) $m = 3$.



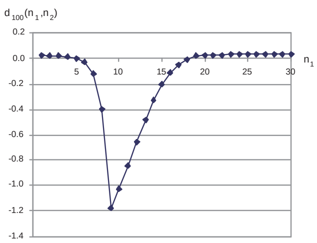
(d) $m = 5$.



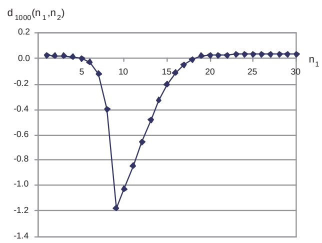
(e) $m = 10$.



(f) $m = 50$.

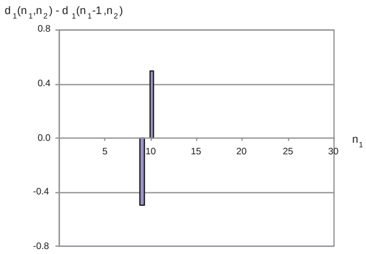


(g) $m = 100$.

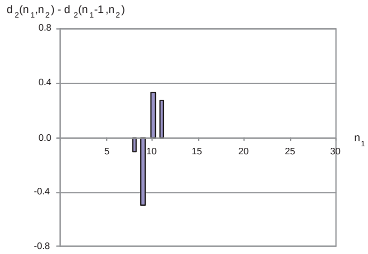


(h) $m = 1000$.

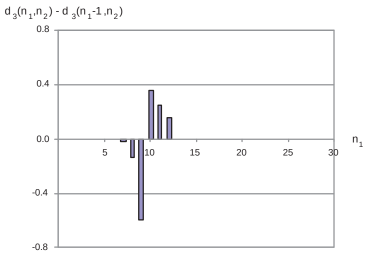
Figure 3.3: $d_m(n_1, n_2)$ versus n_1 for $n_2 = 5$, $\rho = 0.80$, $S = 9$.



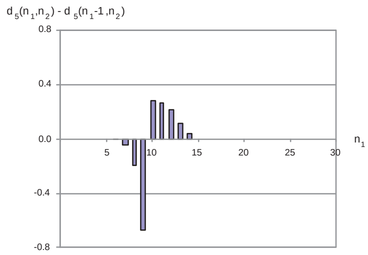
(a) $m = 1$.



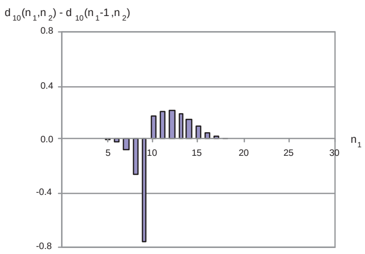
(b) $m = 2$.



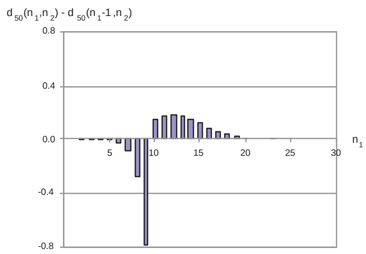
(c) $m = 3$.



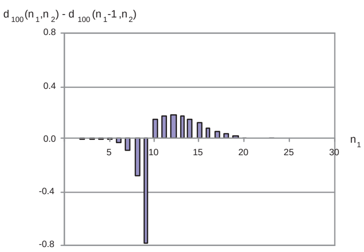
(d) $m = 5$.



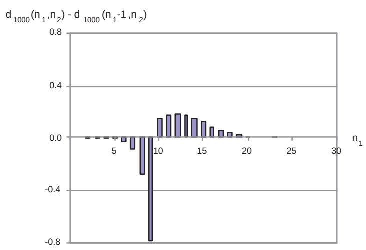
(e) $m = 10$.



(f) $m = 50$.



(g) $m = 100$.



(h) $m = 1000$.

Figure 3.4: $d_m(n_1, n_2) - d_m(n_1 - 1, n_2)$ versus n_1 for $n_2 = 5$, $\rho = 0.80$, $S = 9$.

In Figure 3.5, optimal decisions are shown for each initial state when there are m periods to go, $m = 2, 3, 5, 1000$. Such figures for some other ρ values are given in Appendix D. As seen for three example cases (three different ρ values) in Figure 3.6, once m is sufficiently large, policies converge to the ones characterized by two switching curves intersecting at $\mathbf{n} = (S_1, S_2)$. These curves for m periods to go are defined as follows in a way similar to the ones Ha uses in [8]:

$$B_m(n_1) = \begin{cases} max\{n_2 : d_m(n_1, n_2) \geq 0\} & \text{if } n_1 < S_1, \\ min\{n_2 : d_m(n_1, n_2) \geq 0\} & \text{if } n_1 \geq S_1, \end{cases}$$

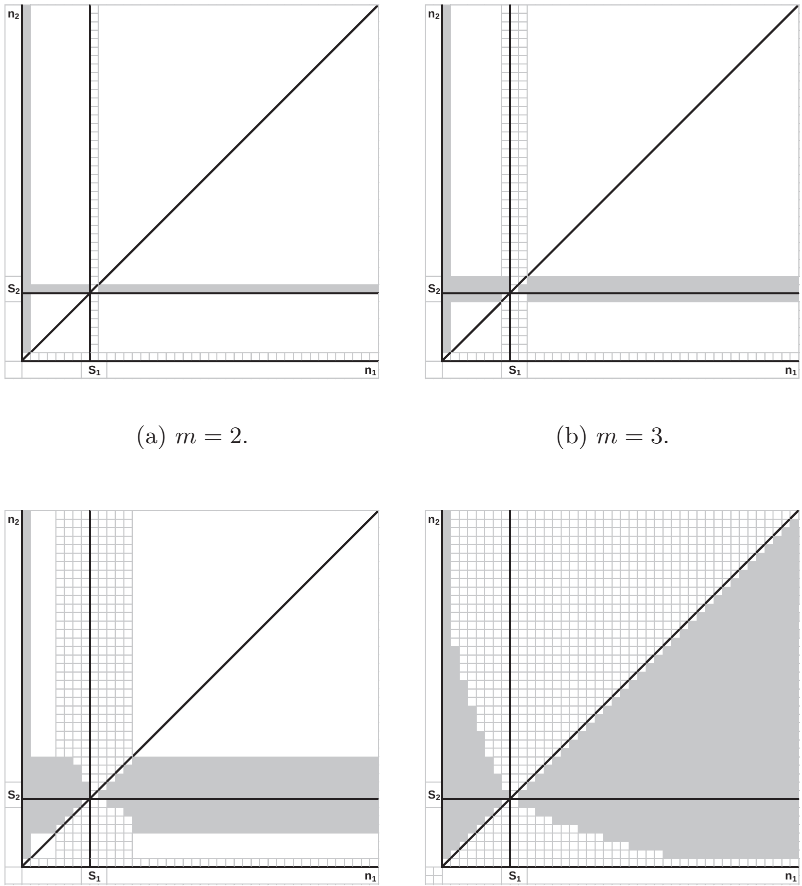
$$A_m(n_1) = \begin{cases} min\{n_2 : d_m(n_1, n_2) \geq 0\} & \text{if } n_1 < S_1, \\ max\{n_2 : d_m(n_1, n_2) \geq 0\} & \text{if } n_1 \geq S_1. \end{cases}$$

Let the state space be separated into the following four regions: $n_1 < S_1$ and $n_2 < S_2$ (region 1), $n_1 \geq S_1$ and $n_2 < S_2$ (region 2), $n_1 \geq S_1$ and $n_2 \geq S_2$ (region 3), $n_1 < S_1$ and $n_2 \geq S_2$ (region 4). When m takes very large values, LQ and SQ policies are optimal in regions 1 and 3, respectively. In the remaining regions 2 and 4, optimal priorities are determined by switching curve $B_m(n_1)$ that converges to $B(n_1)$ as m tends to infinity. Note that $A_m(n_1)$ is the diagonal for all m, so is $A(n_1)$. Hence, the ς policy shows different characteristics over four regions. This policy structure makes sense as clarified below.

- In region 1, none of the items is in stockout. LQ policy is followed in order to avoid stockout for the item with higher risk of falling into the stockout region, i.e., the one with higher n_i.

- In region 3, both of the items are in stockout. SQ policy is followed in order to eliminate stockout for the promising item to reach non-stockout region sooner, i.e., the one with smaller n_i.

- In regions 2 and 4, one of the items is in stockout. $B(n_1)$ is the threshold level to be sufficiently away from region 3, while trying to reach region 1. For example, if the current state is in region 2, say $\mathbf{n}' = (n_1, S_2 - 1)$ such that $n_1 > S_1$, is reached from (n_1, S_2), we do not immediately start processing items of type 1 although, in region 2, type 1 is in stockout but type 2 is not. Instead, we continue processing type 2 until threshold level is reached and then we switch to type 1. This is not to take the risk of falling into more costly region 3 while processing type 1 towards the end of avoiding the cost of region 2 by reaching region 1, i.e., if we immediately start processing type 1 when the current state is $\mathbf{n}'$, most probably the system will be in region 3 before reaching region 1.

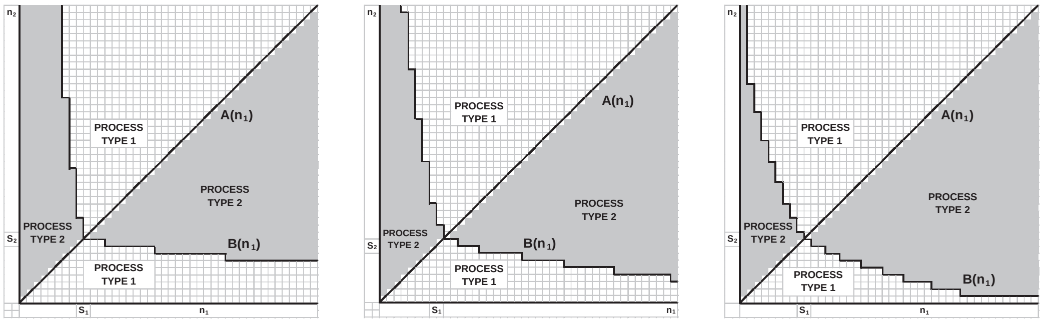


(a) $m = 2$. (b) $m = 3$.

(c) $m = 5$. (d) $m = 1000$.

Figure 3.5: ς policy for $\rho = 0.80$, $S = 9$.

To a certain extent, FCFS would reflect the behavior of LQ policy because it is more probable that the next item to be processed would be of type i if $n_i > n_j$, $j \neq i$. As a matter of fact, LQ and FCFS policies perform almost equally well for a wide range of parameters, especially when they are compared to the ς policy, as seen in Tables 3.1. These two policies are compared to the ς policy in terms of (aggregate) fill rate for constant base-stock levels in Chapter 4. LQ policy is optimal in region 1. When ρ is small, probability of observing only a few items in the queues would be high, i.e., the system would mostly be in region 1. This explains why performance of the ς policy is not strikingly dominating FCFS and LQ policies when ρ is small. On the other hand, for higher ρ values, visiting states outside region 1 is more probable making the difference between the ς policy and the other policies significant. The difference



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure 3.6: ς policy for $S = 9$.

is due to handling states outside region 1 optimally (i.e., to maximize aggregate fill rate), instead of persisting with LQ or FCFS policies.

In this section, performance of the ς policy investigated above is compared with those of LQ and FCFS policies. The comparison is to see the improvement in the required base-stock levels to achieve a given target fill rate under the ς policy compared to LQ and FCFS policies as in the way LQ policy is compared to FCFS in [17]. But, note that our results are not consistent with the ones in [17] because there are several differences between the models; n_i in [17] is for the number of items of type i waiting in the respective queue and the cases with no items in the queue with idle and busy server are distinguished defining $(\mathbf{0}; i)$ and $(\mathbf{0}; b)$, where i stands for *idle* and b stands for *busy*, respectively, and $\mathbf{0}$ is vector of size I with zeros. They do not allow preemption and assume the item in service is considered to be ready for use, either to be put in stock or to be sent to the customer. On the other hand, as in [22], preemption is allowed in our case and the states represent the number of items of different types in the system, i.e., including the one being processed. Thus, our numerical results are directly comparable to the ones in [22].

In Table 3.1, the minimum base-stock levels required in order to achieve different target service levels (aggregate fill rate) under each of the ς, LQ and FCFS policies can be seen. The values for LQ are calculated by using the recursive scheme developed in [22] for finding the steady-state probabilities and the values for FCFS[2] are obtained

[2] $Pr\{N_i = n_i\} = (1 - \widehat{\rho}_i)\widehat{\rho}_i^{n_i}$ is the marginal probability for queue length of class i and $\widehat{\rho}_i =$

Table 3.1: Comparison of the ς, LQ and FCFS policies with minimum base-stock levels to satisfy the given target fill rates.

		ς		FCFS			LQ		
Target FR	ρ	S	FR(%)	S	FR(%)	Error(%)	S	FR(%)	Error(%)
0.90	0.40	**2**	94.35	**2**	93.75	0.60	**2**	94.03	0.32
	0.60	**3**	93.64	**3**	92.13	1.51	**3**	92.68	0.96
	0.80	**5**	90.87	6	91.22	-0.35	6	91.91	-1.04
	0.90	**9**	90.04	12	91.00	-0.96	12	91.50	-1.46
	0.95	**18**	90.90	24	90.95	-0.05	23	90.28	0.62
0.95	0.40	**3**	98.73	**3**	98.44	0.29	**3**	98.68	0.05
	0.60	**4**	97.51	**4**	96.63	0.88	**4**	97.17	0.34
	0.80	**7**	96.22	8	96.10	0.12	8	96.67	-0.45
	0.90	**13**	95.71	15	95.07	0.64	15	95.48	0.23
	0.95	**24**	95.10	30	95.03	0.07	30	95.26	-0.16
0.99	0.40	**4**	99.72	**4**	99.61	0.11	**4**	99.72	0.00
	0.60	**5**	99.05	6	99.38	-0.33	6	99.59	-0.54
	0.80	**10**	99.01	12	99.23	-0.22	11	99.13	-0.12
	0.90	**20**	99.02	23	99.01	0.01	23	99.16	-0.14
	0.95	**40**	99.08	47	99.09	-0.01	46	99.08	0.00

by using the fact that the conditional distribution of $\mathbf{n}$ given $|\mathbf{n}| = n$ is binomial, i.e., $Pr(\mathbf{N} = \mathbf{n}|\ |\mathbf{N}| = n) = \binom{n}{n_1} (\frac{\lambda_1}{\lambda_1+\lambda_2})^{n_1} (\frac{\lambda_2}{\lambda_1+\lambda_2})^{n_2}$ for all $\mathbf{n}$ such that $|\mathbf{n}| = n$, and the steady-state probability of $|\mathbf{N}| = n$ is $(1-\rho)\rho^n$ due to M/M/1 nature of the system when item types are not differentiated.

Under the ς policies, the desired service levels are achieved with lower base-stock levels as compared to the other two policies. The decrease in investment for stock keeping units (sku) is more for higher ρ values. This is explained by the increase in the probability of being in stockout (visiting the states in regions 2, 3 and 4 more frequently) when ρ is high. Recall that handling the stockout case in regions 2, 3 and 4 optimally puts forward the difference between the ς and the other two policies. When $\rho = 0.90$ and $S = 8$, the difference between the fill rates under ς and LQ policies is around 8%. This difference increases to 17% for $S = 1$, because smaller S values increase probability of stockout and cause the policies to be distinguishable (LQ and ς policies are the same as long as the system is in region 1). Performances of the LQ and FCFS policies in terms of the required sku investment turn out to be almost the same (not only for the example cases in Table 3.1, but in general for the extensive numerical experiments) in accordance with the intuitive comparison of LQ and FCFS policies to behave similarly on the average as pointed out in one of the preceding paragraphs.

$\frac{\lambda_i}{\mu-\sum_{j\neq i}\lambda_j}$. See Buzacott and Shanthikumar [5] for details.

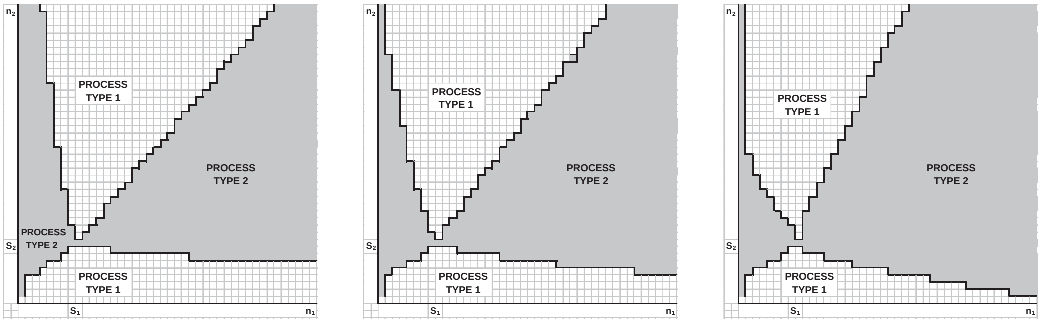
Some error values are negative which means lower (aggregate) fill rate under the ς policy, but note that these cases are observed when the ς policy achieves the target service level with a smaller base-stock level.

3.4 ASYMMETRIC SYSTEMS

Asymmetric two-class systems are considered in this section allowing demand rates and base-stock levels to be different for the item types. When $w_1 = w_2$, both item types are equally important. When $w_1 \neq w_2$, on the other hand, immediate delivery for the item type with higher weight is more important. Note that these two cases cannot be differentiated for the symmetric systems in Section 3.3.

In this section, without loss of generality we assume that $\lambda_1 > \lambda_2$. Figure 3.7 shows structure of the ς policy for $\lambda_1 = 2\lambda_2$, $S_1 = S_2 = S$ and $w_1 = w_2$ and figures for $\lambda_1 = 4\lambda_2$ can be seen in Appendix G. This time, A is not the diagonal and B is not symmetric with respect to the diagonal as one would expect. In figure 3.7, curve A has a smaller (larger) slope in region 1 (3) compared to the symmetric case. Curve B, on the other hand, is not that steep in region 2 but steeper in region 4. The explanation below for such shifts of the switching curves as compared to the symmetric case is immediate when $S_1 = S_2 = S$ and $w_1 = w_2$.

- In region 1, there is not stockout for any of the item types. In addition to the states such that $n_1 > n_2$, at some others such that $n_1 \leq n_2$ it would be more probable for type 1 to be in stockout before type 2 because $\lambda_1 > \lambda_2$. That is, A shifts to favour (to give priority to) the type with higher demand rate more.
- In region 3, there is stockout for both of the item types. In addition to the states such that $n_1 > n_2$, at some others with $n_1 \leq n_2$ it would be more probable to eliminate stockout of type 2 before that of type 1 because $\lambda_2 < \lambda_1$. That is, A shifts to favour (to give priority to) the type with lower demand rate more.
- In regions 2 and 4, one of the item types is in stockout. B is the threshold to be sufficiently away from region 3 while trying to reach region 1. But, since the demand rates are different, B is not symmetric with respect to the diagonal. When the type with higher (lower) demand rate is in stockout and the other is not, then at any state it is more (less) probable as compared to the symmetric case to reach region 1 before the other type steps up to region 3 while processing the type in stockout. That is, B shifts to favour (to give priority to) the type with higher demand rate more.



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure 3.7: ς policy: equally weighted cost function, $w_1 = w_2$, $S = 8$, $\lambda_1 = 2\lambda_2$.

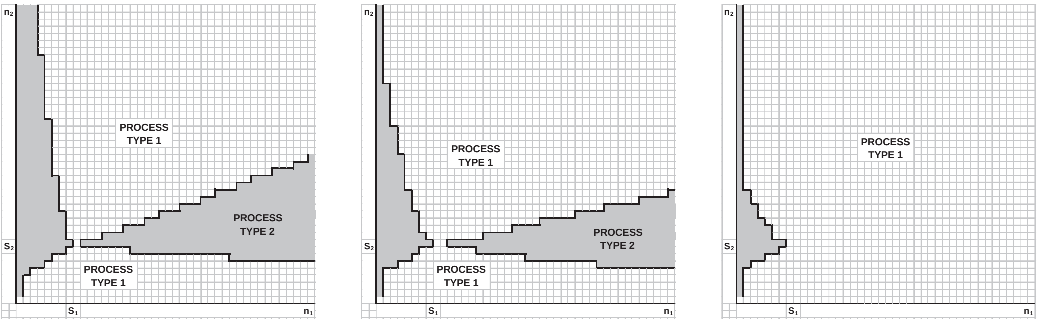
Figures 3.8, 3.9, 3.10 show structures of the ς policy for three different weight vectors. Some other figures can be seen in Appendix G. These figures imply that the weights are dominant in determining structure of the ς policy. Especially for high values of ρ, the ς policy chooses to process the class with higher weight in most of the states even if that class has lower demand rate. The observations are itemized below.

- In region 1, in addition to states such that $n_1 > n_2$, at some others with $n_1 \leq n_2$ it would be more probable for type 1 to be in stockout before type 2 because $\lambda_1 > \lambda_2$. Since region 2 is more costly than region 4 when $w_1 > w_2$, A shifts to favor (to give priority to) type 1 even more compared to the case $w_1 = w_2$.

- In region 3, it is possible to have more cost saving by reaching region 4 than reaching region 2 when $w_1 > w_2$, then A shifts to favor (to give priority to) type 1 even more compared to the case $w_1 = w_2$.

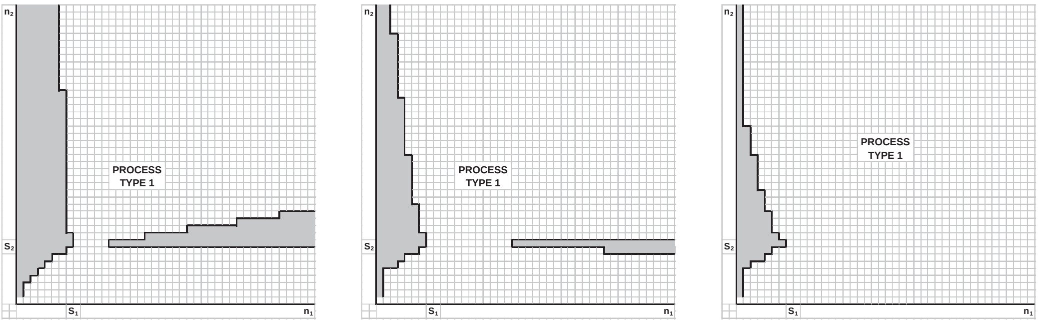
- In region 2 (4) in addition to states in which it would be more probable for type 1 to eliminate stockout (to step up to region 3) than for type 2 to step up to region 3 (to eliminate stockout), at some others type 1 is given priority because the cost saving by reaching region 1 is larger (less) than the one avoided by being sufficiently away from region 3.

- When demand rates and/or weights for the item types differ, one may wish to permit unequal base-stock levels. Since the structure of the ς policy is highly dependent on the distances to the base-stock levels from any given state, the general structure of the ς policy stays the same in the case of unequal base-stock levels.



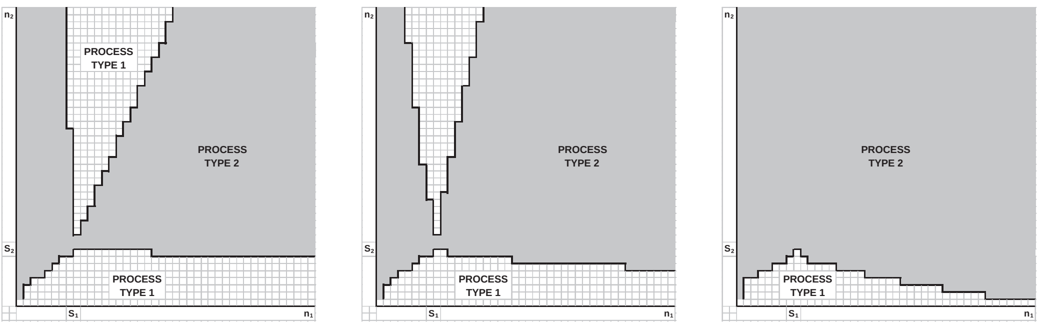
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure 3.8: ς policy: cost function weighted by demand rates, $w_1 = 2w_2$, $S = 8$, $\lambda_1 = 2\lambda_2$.



(a) $\rho = 0.1$. (b) $\rho = 0.40$. (c) $\rho = 0.60$.

Figure 3.9: ς policy for $w_1 = 5w_2$, $S = 8$, $\lambda_1 = 2\lambda_2$.



(a) $\rho = 0.1$. (b) $\rho = 0.40$. (c) $\rho = 0.90$.

Figure 3.10: ς policy for $w_2 = 2w_1$, $S = 8$, $\lambda_1 = 2\lambda_2$.

CHAPTER 4

HEURISTICS FOR FILL RATE MAXIMIZATION

In this chapter, heuristics are proposed to approximate the ς policy of which structure is investigated based on numerical experiments for two-class systems in Sections 3.3 and 3.4. In devising these heuristics the approach is to approximate the switching curve B (and also A in the asymmetric case) because a closed-form expression cannot be derived for B (and A). Section 4.1 is on the heuristics proposed for the symmetric case. Some of the heuristic approaches in Section 4.1 are then extended in a natural way to the asymmetric case with two classes in Section 4.2 and more than two classes (numerical experiments being for the three-class case) in Section 4.3.

4.1 SYMMETRIC TWO-CLASS SYSTEMS

Five heuristics are presented in this section for $S_i = S$ and $\lambda_i = \lambda$. Recall that B characterizes the ς policy together with A, the latter of which turns out to be the diagonal $n_2 = n_1$ for symmetric systems. The heuristics are for m being sufficiently large to approximate the long-run average behavior. Heuristics 3 and 4 are proposed with rather rough approximations of B_m for large m compared to the other heuristics; on the other hand, heuristic 3 has the advantage of having balance equations which we can solve recursively unlike others for which we have to continue using time consuming value-iteration to compute the steady-state aggregate fill rates (or any other measures).

All alternative policies are of LQ and SQ types in regions 1 and 3, respectively, as observed numerically for the ς policy. The reasoning behind the first two heuristics is given next. When the system is in state (n_1, n_2) in region 2, with rate λ_2 the system moves to the states closer to the more costly region 3; on the other hand, the system gets away from the zero-cost region 1 with rate λ_1. In order to approximate B, the choice for each state in region 2 would be between trying to get away from the more costly region 3 by processing items of type 2 (while at the same time getting away from region 1 by arrivals of type 1) and trying to get closer to zero-cost region 1 by processing type 1 items (while at the same time getting closer or even into region 3 by arrivals of type 2 items). Similar arguments can be raised also for the states in region 4. Heuristics 1 and 2 are based on comparing the expected times required to cover the distance from a state in region 2 (4) to regions 1 and 3 while processing type 1 (2). For each state in region 2 and 4, indices proposed for both of the item types are to represent the expected times to cover these distances as shown in Figure 4.1 for an example state, and heuristics 1 and 2 are to choose processing the item type with lower index in that state. Note that the system reduces to single-class M/M/1 when the server is devoted to processing one of the classes, say class 1. Then, the expected time for the server to become idle, when there is one item of type 1 in the system, is $\frac{1}{\mu-\lambda_1}$, which is the so-called expected length of a busy period. With a similar reasoning, the expected time required to move from $n_1 = S_1$ to $n_1 = S_1 - 1$ (to process an item of type 1, say S_1^{st} item, and all the items of type 1 arriving during the processing time of the S_1^{st} item under consideration) is $\frac{1}{\mu-\lambda_1}$. Then, the expected time required to reach $n_1 = S_1 - 1$ from any $\mathbf{n}'$ in region 2 (to process $n' - (S_1 - 1)$ items of type 1 and all the items of type 1 arriving during the processing of the $n' - (S_1 - 1)$ items under consideration) is equal to $\frac{n_1-(S_1-1)}{\mu-\lambda_1}$. On the other hand, the expected time required to reach $n_2 = S_2$ from $\mathbf{n}'$ while processing type 1 is equal to the expected time for $(S_2 - n_2')$ items of type 2 to arrive, i.e., $\frac{S_2-n_2'}{\lambda_2}$.

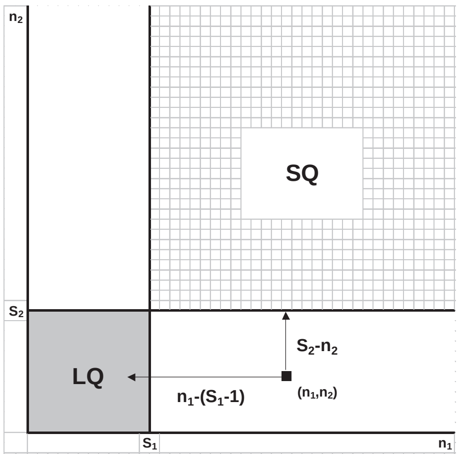
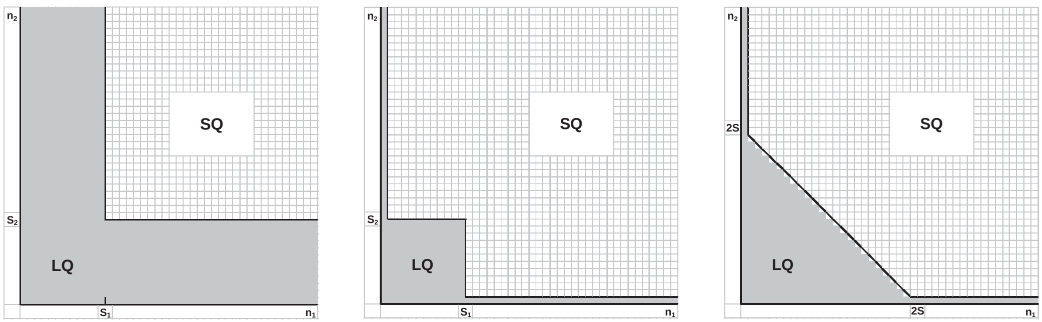


Figure 4.1: Distances from a state in region 2 to regions 1 and 3.

Heuristic 1: $\frac{n_1-(S_1-1)}{\mu-\lambda_1}$ and $\frac{S_2-n_2}{\lambda_2}$ are the indices for item types 1 and 2, respectively, for state (n_1, n_2) in region 2. The former is the expected time to reach region 1 from state (n_1, n_2) (to cover distance $n_1 - (S_1 - 1)$) while type 1 is being processed. Similarly, the latter is the expected time to reach region 3 (to cover distance $S_2 - n_2$) while type 1 is being processed. If the former is smaller, then we can process type 1 items because we expect to reach (zero-cost) region 1 before stepping up the more costly region 3. Otherwise, if $\frac{S_2-n_2}{\lambda_2}$ is smaller, then it is more probable that we will step up the more costly region 3 before reaching region 1 by processing type 1 items, so going away from the more costly region is preferable for the state under consideration. Note that $\frac{S_1-n_1}{\lambda_1}$ and $\frac{n_2-(S_2-1)}{\mu-\lambda_2}$ are the indices to be compared for any state (n_1, n_2) in Region 4.

Heuristic 2: This heuristic is just a variation of heuristic 1 obtained revising indices as $\frac{n_1-(S_1-1)}{\mu}$ and $\frac{S_2-n_2}{\lambda_2}$ for item types 1 and 2, respectively, for state (n_1, n_2) in Region 2 and as $\frac{S_1-n_1}{\lambda_1}$ and $\frac{n_2-(S_2-1)}{\mu}$ for state (n_1, n_2) in Region 4.



(a) Heuristic 3. (b) Heuristic 4. (c) Heuristic 5.

Figure 4.2: Heuristics.

Although heuristics 1 and 2 are easy to implement, use of time-consuming (especially for problems with high traffic intensity) value-iteration is unavoidable to compute the steady-state performance measures. This is why three other heuristics are proposed (to capture general structure of the ς policy) so that we may either derive a closed-form solution or at least devise a recursive scheme to calculate the steady-state probabilities easily. As a matter of fact, as noted in the first paragraph of this section, we come up with a recursive algorithm for heuristic 3 but unfortunately not for heuristics 4 and 5.

Heuristic 3: SQ policy in region 3 and LQ policy in the remaining regions are followed as seen in Figure 4.2(a). The algorithm in [22] is used to find the steady-state probabilities in the region $\{(i,j)|\ i < S_1 - 1 \text{ or } j < S_2 - 1\} \cup (i = S_1 - 1,\ j = S_2 - 1)$ where LQ policy is followed, then the other steady-state equations can be solved recursively. Introducing the notation and some preliminaries next, the algorithm in [22] is extended for this heuristic.

State-transition diagram for the LQ model studied in [22] is seen in Figure 4.3. Arcs for the arrivals are not shown in this figure. Each node has two arcs with rate λ (because $\lambda_i = \lambda$), one upward and one to the right. Let p_{ij} be the steady-state probability of being in state $(n_1, n_2) = (i, j)$. Note that $p_{ij} = p_{ji}$ for the symmetric case. Also, p_{ij} is set to zero if either i or j is less than zero. Then, the balance equations for the system can be written as

$$(1+\rho)p_{ij} = \frac{1}{2}\rho(p_{i-1,j} + p_{i,j-1}) + \begin{cases} p_{i+1,j} & \text{if } i \geq j+2, \\ p_{i,j+1} & \text{if } j \geq i+2, \\ p_{i+1,j} + \frac{1}{2}p_{i,j+1} & \text{if } i = j+1, \\ \frac{1}{2}p_{i+1,j} + p_{i,j+1} & \text{if } j = i+1, \\ p_{i+1,j} + p_{i,j+1} & \text{if } i = j, \end{cases} \tag{4.1}$$

for $(i,j) \neq (0,0)$ (note that $\rho p_{00} = p_{01} + p_{10}$). On the other hand, since total number of items in the system, $n_1 + n_2$, is the state description of M/M/1 queue with parameter ρ when item types are not differentiated, we have

$$\begin{aligned} p_{00} \quad &= Pr(n_1 + n_2 = 0) &&= 1 - \rho, \\ p_{01} = p_{10} \quad &= \tfrac{1}{2}Pr(n_1 + n_2 = 1) &&= \frac{1}{2}(1-\rho)\rho. \end{aligned}$$

The algorithm in [22] is summarized next. Letting $x_i(k) = p_{i,i+k}$, $k \geq 0$, the balance equations for $j \geq i+2$ can be rewritten as

$$\begin{aligned} x_i(k+1) - (1+\rho)x_i(k) + \frac{1}{2}\rho x_i(k-1) \\ = -\frac{1}{2}\rho x_{i-1}(k+1) \qquad\qquad & for\ k \geq 2,\ i \geq 0. \end{aligned} \tag{4.2}$$

If $x_{i-1}(\cdot)$ is known, then (4.2) becomes a linear second-order difference equation with unknowns $x_i(\cdot)$. Note that, for $i = 0$, $x_{i-1}(\cdot) = 0$, (4.2) is a homogeneous difference equation with the characteristic equation $z^2 - (1+\rho)z + \frac{1}{2}\rho = 0$. Let ζ be the root of the characteristic equation. Then, $x_0(k)$, $k \geq 0$, can be obtained by using $x_0(1) = p_{01} = \frac{1}{2}(1-\rho)\rho$, and the algorithm proceeds by increasing i. Marginal probabilities of the queue lengths under LQ policy can be calculated by using equation

(4.1) as follows

$$\begin{aligned} p_i &= \frac{1}{2}\rho p_{i-1} + \frac{1}{2}p_{ii} + \frac{1}{\rho}p_{i+1,j+1} \\ &= \frac{1}{2}\rho p_{i-1} + \frac{1}{2}x_i(0) + \frac{1}{\rho}x_{i+1}(0) \qquad\qquad for\ i \geq 1. \end{aligned}$$

where $p_i = \sum_{j=0}^{\infty} p_{i,j}$ for $i \geq 0$ and $p_0 = (1-\rho)[1+\frac{1}{2}\rho/(1-\zeta)]$. Details of the algorithm can be found in [22].

State-transition diagram for heuristic 3 can be seen in Figure 4.4. Figure 4.5 is to show how we proceed in a recursive manner to calculate steady-state probabilities in SQ region for heuristic 3. Black nodes in Figure 4.5(a) represent the states for which steady-state probabilities are determined by Zheng and Zipkin's algorithm in [22], shaded nodes are the ones for which steady-state probabilities are to be determined in the order of Figures 4.5(a), 4.5(b), 4.5(c), 4.5(d). The nodes become black once the steady-state probabilities are calculated. The recursive scheme is given below to achieve an accuracy less than ϵ for some given ϵ.

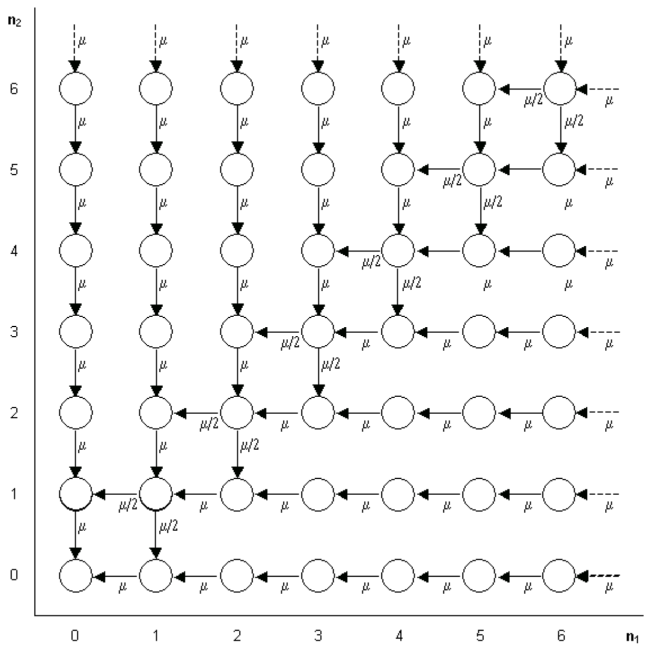


Figure 4.3: State-transition diagram for the LQ model (arrival transitions are not shown).

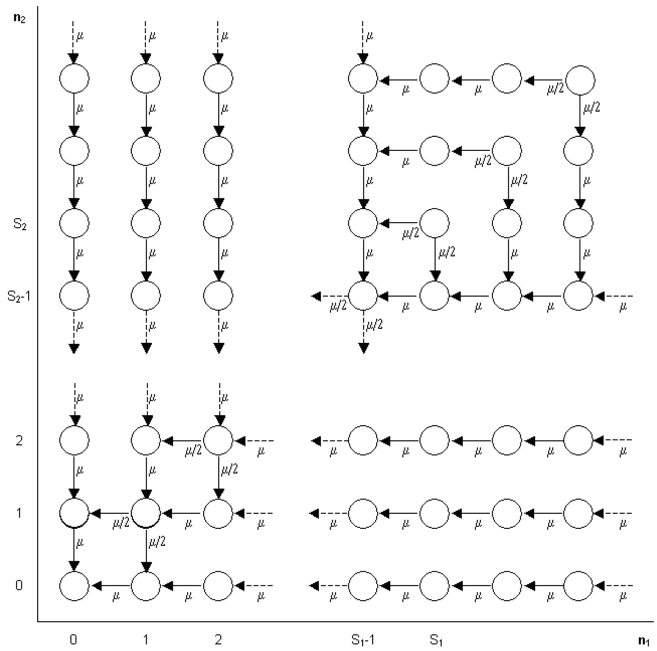


Figure 4.4: State-transition diagram for Heuristic 3 (arrival transitions are not shown).

Recursive Scheme to Calculate Steady-State Probabilities:

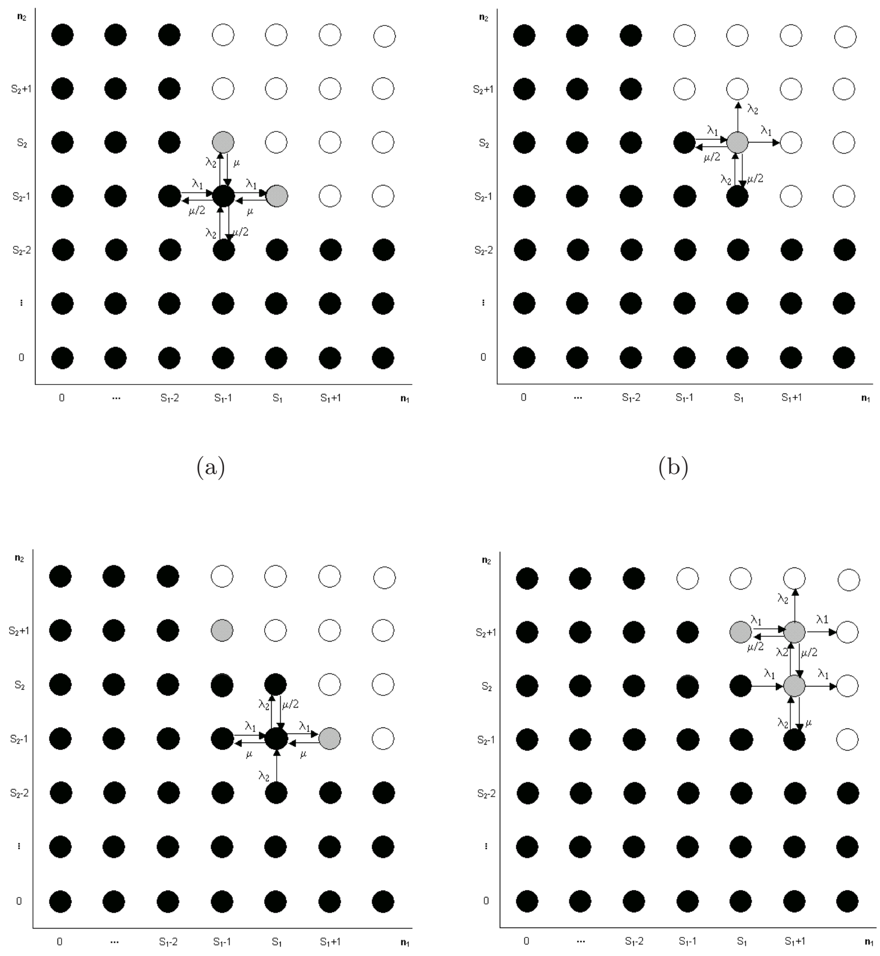
Step 1. *Use algorithm in [22] to find* p_{ij} *for* (i, j) *in* $\{(i, j)|\ i < S_1 - 1$ *or* $j < S_2 - 1\}$ $\cup$ $(i = S_1 - 1,\ j = S_2 - 1)$. *Set* $\omega = 0$.

Step 2. *Solve the balance equation of state* $(S_1 - 1 + \omega, S_2 - 1)$ *for* $p_{s_1+\omega,s_2-1}$.

Step 3. *Solve the balance equations of states* $(S_1 + \omega, S_2 + \upsilon)$, $\upsilon = 0,\ ...,\ \omega$, *(*$\omega + 1$ *unknowns and* $\omega + 1$ *linearly independent equations) for* $p_{s_1+\omega,s_2+\upsilon}$, $\upsilon = 0,\ ...,\ \omega$.

Step 4. *If total probability calculated is greater than* $1 - \varepsilon$*, then stop; else, increase* ω *by 1 and go to step 2.*

Table 4.1 is to compare the efforts spent for the recursive scheme above and value iteration. CPU time for value-iteration is not any different for the heuristics and for finding the ς policy. The number of iterations required for the recursive scheme to converge when $\varepsilon = 0.00001$ is also given in this table. As seen in the table, the recursive scheme gives the opportunity to calculate aggregate fill rates almost instantaneously. With the given ε value and the required number of iterations listed, aggregate fill rates calculated using the recursive scheme turn out to be the same (upto and including the second decimal digit) as the ones calculated by value-iteration for heuristic 3.



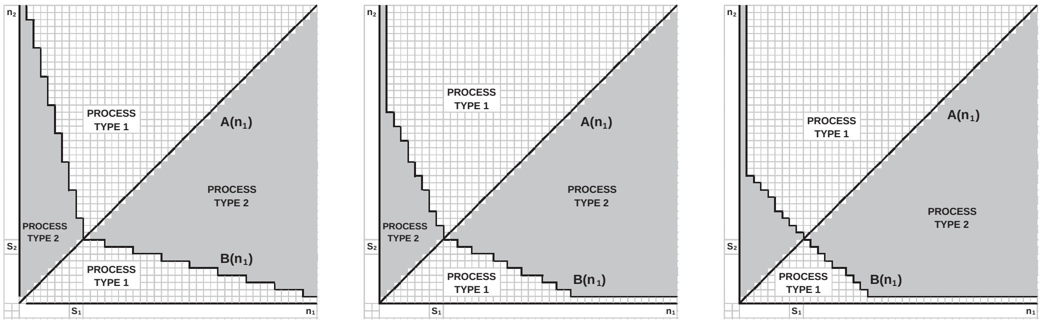
(a) (b) (c) (d)

Figure 4.5: Recursion for Heuristic 3.

Heuristic 4: LQ Policy in region 1 and SQ Policy in the remaining regions are followed as seen in Figure 4.2(b). Note that for the states where one of the queues is empty LQ policy is applied.

Heuristic 5: LQ policy in region $\{(n_1, n_2)|\ n_1 + n_2 < 2S\}$ and SQ policy in the remaining regions are followed as seen in Figure 4.2(c).

Note that the heuristics are the same for the case $S = 1$; SQ policy is employed when both n_1 and n_2 are positive and LQ is employed when n_1 or $n_2 = 0$. The observations that result from the numerical experiments given in Table 4.2 and Figure 4.6 and in



(a) $\rho = 0.40$, $S = 9$.

(b) $\rho = 0.60$, $S = 9$

(c) $\rho = 0.90$, $S = 9$

Figure 4.6: Heuristic 2.

Table 4.1: Performance of the recursive scheme for Heuristic 3.

ρ	number of iterations for the recursive scheme								CPU time for	
ρ	$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$	value-iteration	recursive scheme
0.10	4	2	1	1	1	1	1	1	0.88982	**0.00739**
0.20	6	4	2	1	1	1	1	1	3.65946	**0.00738**
0.30	8	6	4	3	1	1	1	1	8.37664	**0.00739**
0.40	10	9	7	5	2	1	1	1	17.99361	**0.00740**
0.50	14	13	11	9	5	2	1	1	43.66775	**0.00752**
0.60	20	18	16	15	11	7	2	1	115.50373	**0.00782**
0.70	30	28	26	24	20	16	11	3	414.66782	**0.00979**
0.80	50	48	46	44	40	36	30	22	1501.28810	**0.02788**
0.90	108	106	104	102	98	94	88	80	5196.20532	**0.57421**

Appendix E:

- Heuristics 1 and 2 perform very well since the actions chosen are the same as the ς policy in regions 1 and 3 and in most of the states in regions 2 and 4. Heuristic 2 works better than heuristic 1 for almost all ρ and S combinations (parameter sets) considered.

- Heuristics 3 and 4 perform worse (better) than heuristic 1 and 2 (LQ and FCFS policies) because the actions chosen are very different than (the same as) the ς policy (only) in regions 2 and 4 (1 and 3).

- To compare the approximate switching curve B with the exact curve, see Figure 4.6 for heuristic 2 (this figure given for heuristic 2 because it turns out to be

Table 4.2: Comparison of the aggregate fill rates (%) of the ς, LQ, FCFS policies and the heuristics, $\lambda_i = \lambda$, $w_i = \frac{1}{2}$, $i = 1, 2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$
	ς	94.76	99.73	99.99	100.00	100.00	100.00
	LQ	94.72	**99.73**	**99.99**	**100.00**	**100.00**	**100.00**
	FCFS	94.74	99.72	**99.99**	**100.00**	**100.00**	**100.00**
0.1	Heuristic 1	**94.76**	**99.73**	**99.99**	**100.00**	**100.00**	**100.00**
	Heuristic 2	**94.76**	**99.73**	**99.99**	**100.00**	**100.00**	**100.00**
	Heuristic 3	**94.76**	**99.73**	**99.99**	**100.00**	**100.00**	**100.00**
	Heuristic 4	**94.76**	99.70	99.98	**100.00**	**100.00**	**100.00**
	Heuristic 5	**94.76**	**99.73**	**99.99**	**100.00**	**100.00**	**100.00**
	ς	89.07	98.85	99.88	99.99	100.00	100.00
	LQ	88.79	98.84	**99.88**	**99.99**	**100.00**	**100.00**
	FCFS	88.89	98.77	99.86	99.98	**100.00**	**100.00**
0.2	Heuristic 1	**89.07**	**98.85**	**99.88**	**99.99**	**100.00**	**100.00**
	Heuristic 2	**89.07**	**98.85**	**99.88**	**99.99**	**100.00**	**100.00**
	Heuristic 3	**89.07**	**98.85**	**99.88**	**99.99**	**100.00**	**100.00**
	Heuristic 4	**89.07**	98.64	99.84	99.98	**100.00**	**100.00**
	Heuristic 5	**89.07**	98.82	**99.88**	99.99	**100.00**	**100.00**
	ς	82.94	97.15	99.55	99.93	100.00	100.00
	LQ	82.04	97.07	99.54	**99.93**	**100.00**	**100.00**
	FCFS	82.35	96.89	99.45	99.90	**100.00**	**100.00**
0.3	Heuristic 1	**82.94**	**97.15**	**99.55**	**99.93**	**100.00**	**100.00**
	Heuristic 2	**82.94**	**97.15**	**99.55**	**99.93**	**100.00**	**100.00**
	Heuristic 3	**82.94**	**97.15**	**99.55**	**99.93**	**100.00**	**100.00**
	Heuristic 4	**82.94**	96.59	99.32	99.87	**100.00**	**100.00**
	Heuristic 5	**82.94**	97.06	99.53	**99.93**	**100.00**	**100.00**
	ς	76.38	94.35	98.73	99.72	99.99	100.00
	LQ	74.31	94.03	98.68	**99.72**	**99.99**	**100.00**
	FCFS	75.00	93.75	98.44	99.61	99.98	**100.00**
0.4	Heuristic 1	**76.38**	**94.35**	**98.73**	**99.72**	**99.99**	**100.00**
	Heuristic 2	**76.38**	**94.35**	**98.73**	**99.72**	**99.99**	**100.00**
	Heuristic 3	**76.38**	**94.35**	**98.73**	**99.72**	**99.99**	**100.00**
	Heuristic 4	**76.38**	93.37	98.10	99.47	99.97	**100.00**
	Heuristic 5	**76.38**	94.17	98.67	99.71	**99.99**	**100.00**
	ς	69.35	90.14	96.99	99.11	99.93	99.99
	LQ	65.45	89.13	96.74	99.05	99.92	**99.99**
	FCFS	66.67	88.89	96.30	98.77	99.86	99.98
0.5	Heuristic 1	**69.35**	90.13	**96.99**	**99.11**	**99.93**	**99.99**
	Heuristic 2	**69.35**	**90.14**	**96.99**	**99.11**	**99.93**	**99.99**
	Heuristic 3	**69.35**	90.07	96.97	**99.11**	**99.93**	**99.99**
	Heuristic 4	**69.35**	88.77	95.70	98.38	99.80	99.98
	Heuristic 5	**69.35**	88.90	96.84	99.06	99.92	**99.99**
	ς	61.82	84.21	93.64	97.51	99.64	99.95
	LQ	55.32	81.55	92.68	97.17	99.59	99.94
	FCFS	57.14	81.63	92.13	96.63	99.38	99.89
0.6	Heuristic 1	**61.82**	**84.21**	**93.63**	**97.50**	**99.64**	**99.95**
	Heuristic 2	**61.82**	84.15	93.62	**97.50**	**99.64**	**99.95**
	Heuristic 3	**61.82**	83.83	93.49	97.46	99.63	**99.95**
	Heuristic 4	**61.82**	82.61	91.58	95.92	99.11	99.83
	Heuristic 5	**61.82**	83.95	93.39	97.36	99.61	**99.95**

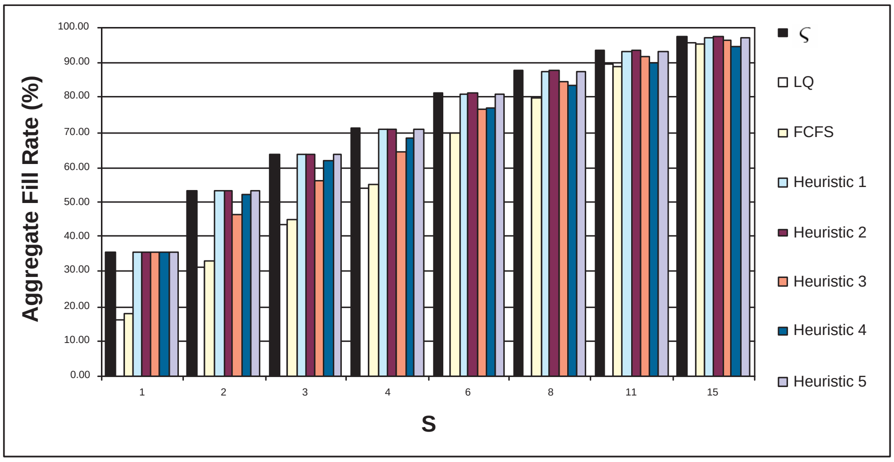
Table 4.2 Continued.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
	ς	53.72	76.26	87.73	93.75	98.43	99.61	99.95	100.00
	LQ	43.81	70.25	84.70	92.26	98.07	99.53	99.94	**100.00**
	FCFS	46.15	71.01	84.39	91.59	97.56	99.29	99.89	99.99
0.7	Heuristic 1	**53.72**	76.06	87.60	93.67	98.40	**99.61**	**99.95**	**100.00**
	Heuristic 2	**53.72**	**76.26**	**87.72**	**93.74**	**98.43**	**99.61**	**99.95**	**100.00**
	Heuristic 3	**53.72**	75.03	87.03	93.40	98.35	99.59	**99.95**	**100.00**
	Heuristic 4	**53.72**	74.68	85.11	91.09	96.89	98.98	99.83	99.99
	Heuristic 5	**53.72**	76.06	87.43	93.48	98.31	99.57	**99.95**	**100.00**
	ς	44.97	66.00	78.14	85.86	94.12	97.58	99.36	99.89
	LQ	30.81	53.95	69.88	80.48	91.91	96.67	99.13	99.85
	FCFS	33.33	55.56	70.37	80.25	91.22	96.10	98.84	99.77
0.8	Heuristic 1	**44.97**	65.92	77.95	85.69	93.98	97.49	99.34	**99.89**
	Heuristic 2	**44.97**	**66.00**	**78.14**	**85.85**	**94.12**	**97.58**	**99.36**	**99.89**
	Heuristic 3	**44.97**	62.94	75.62	84.15	93.41	97.29	99.29	99.88
	Heuristic 4	**44.97**	64.74	75.57	82.52	90.94	95.41	98.43	99.65
	Heuristic 5	**44.97**	65.92	77.95	85.57	93.83	97.39	99.29	99.88
	ς	35.46	53.20	63.86	71.24	81.26	87.71	93.47	97.19
	LQ	16.23	31.12	43.80	54.31	69.93	80.26	89.50	95.48
	FCFS	18.18	33.06	45.23	55.19	70.00	79.92	89.00	95.07
0.9	Heuristic 1	**35.46**	**53.20**	**63.79**	71.08	80.97	87.37	93.17	97.01
	Heuristic 2	**35.46**	53.17	63.74	**71.13**	**81.19**	**87.66**	**93.44**	**97.17**
	Heuristic 3	**35.46**	46.65	56.36	64.48	76.61	84.64	91.83	96.48
	Heuristic 4	**35.46**	52.48	62.11	68.50	77.24	83.38	89.74	94.72
	Heuristic 5	**35.46**	**53.20**	**63.79**	71.08	80.90	87.25	93.05	96.92

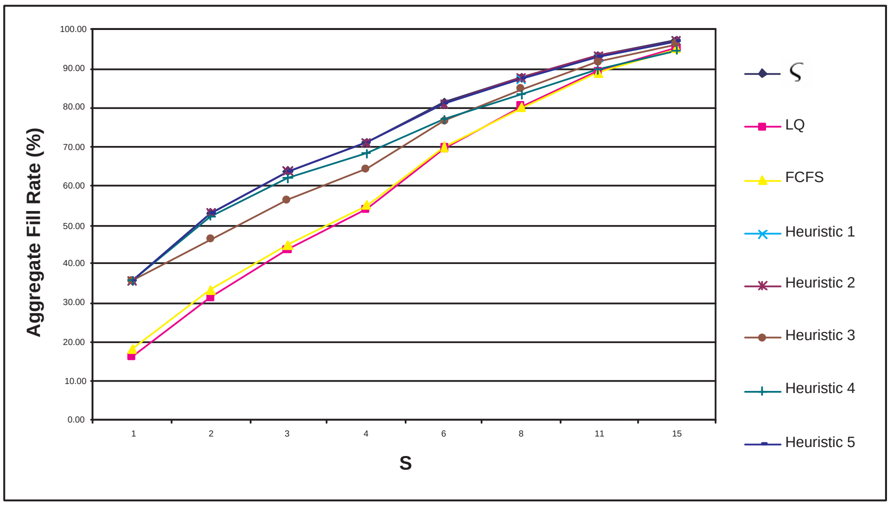
the best for almost all parameter sets) referring to Figure 3.6 as for the ς policy with the same ρ and S values.

4.2 ASYMMETRIC TWO-CLASS SYSTEMS

For the symmetric systems, it is enough to approximate only curve B, A is the diagonal. However, A also needs to be approximated for the asymmetric systems. To that end, i.e., to approximate A, the idea used in heuristics 1 and 2 for the symmetric case, which is based on comparing the expected time required to reach regions 1 and 3 from a state in regions 2 or 4, is extended for the states in regions 1 and 3. With a revision of heuristic 1, indices in each region are given as in Table 4.3. First, consider the equally weighted cost function i.e., $w_1 = w_2$. Then, while comparing the indices w_i's will cancel out, i.e., our explanation below for $w_1 = w_2$ is by disregarding w_i's in Table 4.3 as if they are already cancelled out. Then, for a state in region 1, the index for type 1 (2) is the time to reach region 2 (4) from this state under consideration while type 2 (1) is being processed. For a state in region 3, the index for type 1 (2) is the expected time to reach region 4 (2) from this state while type 1 (2) is being processed. Indices for the states in regions 2 and 4 are the same as introduced for the



(a) $\rho = 0.90$.



(b) $\rho = 0.90$.

Figure 4.7: Comparison of the ς, LQ, FCFS Policies and the heuristics.

symmetric case in Section 4.1 to approximate curve B. See Figure 4.8 for the distances from given states in different regions. For the cases of arbitrary weights w_1 and w_2, indices are adjusted dividing them by the corresponding weights, i.e., reflecting the effects of the savings gained by reaching a less costly region or by postponing to fall into a more costly region to the indices in a way proportional to the weights.

Revision of heuristic 2 in Section 4.1 is then obtained by using μ instead of $\mu - \lambda_1$

Table 4.3: Indices for heuristic 1.

Region 4	Region 3
$\frac{(S_1-n_1)}{\lambda_1 w_1}, \frac{(n_2-(S_2-1))}{(\mu-\lambda_2)w_2}$	$\frac{(n_1-(S_1-1))}{(\mu-\lambda_1)w_1}, \frac{(n_2-(S_2-1))}{(\mu-\lambda_2)w_2}$
Region 1	**Region 2**
$\frac{(S_1-n_1)}{\lambda_1 w_1}, \frac{(S_2-n_2)}{\lambda_2 w_2}$	$\frac{(n_1-(S_1-1))}{(\mu-\lambda_1)w_1}, \frac{(S_2-n_2)}{\lambda_2 w_2}$

and $\mu - \lambda_2$ for calculating the first index in region 2 and the second one in region 4, respectively. Note that these heuristics are generalizations of heuristics 1 and 2 in Section 4.1, i.e., when demand rates and weights are equal they turn into the respective ones in Section 4.1.

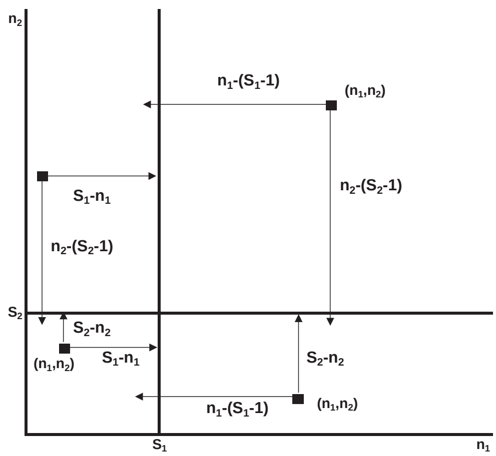
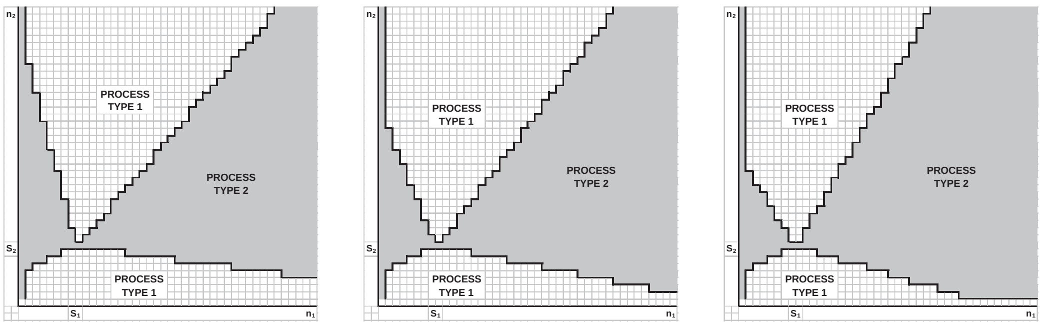


Figure 4.8: Distances from a state in one region to neighboring regions.

For the equally weighted cost function, the heuristic policies are good at representing the general structure of the optimal policy as can be seen comparing Figure 3.7 to 4.9 and Figure G.1 to G.2 in Appendix G. In this case, demand rates determine behaviour of the switching curves. Otherwise, i.e., when the weights are not equal, these weights become dominant in this sense as previously noted in Section 3.4 and the indices in Table 4.3 fail in capturing the general structure of the ς policy especially in region 3. When ρ is high (and possibly when base-stock levels are low), the system would be visiting region 3 frequently and the failure mentioned above would become apparent. Based on numerical observations, at least for capturing the general behaviour of the switching curve in region 3, index for the item type with higher weight is adjusted multiplying it by $(1-\rho)$. Performance of heuristic 2 with this adjustment, heuristic 2 (M), can be seen in Figures 4.10, 4.11 and 4.12 to be compared with Figures 3.8, 3.9, 3.10, respectively, and Figures G.4 to G.8 in Appendix G. In fact, we should note that the aggregate fill rates are approximated very accurately although approximate curve A may turn out to be very different than the one under ς policy. For example, see figures 3.8(c) and 4.10(c) for the ς policy and heuristic 2, respectively. There is a significant difference between the actions taken by the policies but the aggregate fill rates, 89.48% and 89.34%, respectively, in Table 4.5 are very close.



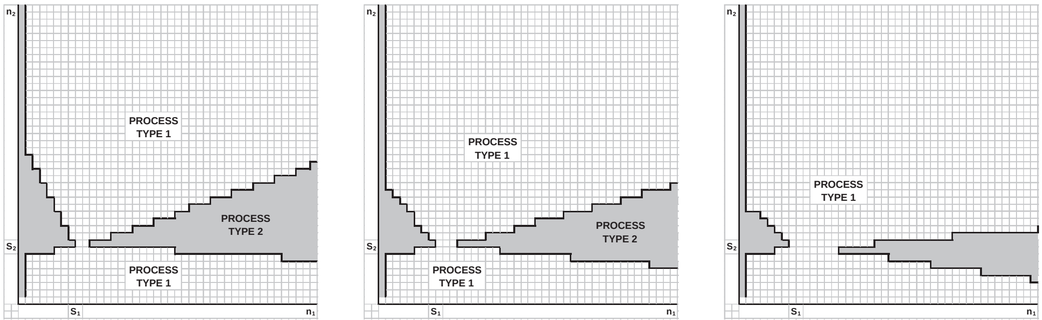
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure 4.9: Heuristic 2: equally weighted cost function, $S = 8$, $\lambda_1 = 2\lambda_2$.

In Tables 4.4 to 4.6 and Tables H.1 to H.7 the heuristics are compared with the ς and FCFS policies. There is another heuristic introduced in [22] and used here for comparisons: serve a customer of type 1 when $n_1 - n_2 < \Delta$ with Δ being some predetermined constant. Note that the relationship between Δ and λ_i is not investigated in [22], but the recursive scheme developed for LQ policy is extended. Tables 4.7 and F.1 are to determine the best Δ value(s) for some given base-stock levels. As seen in these tables, best Δ may correspond to a static priority rule as in the cases of $\mathbf{S} = (6, 3)$ and $\mathbf{S} = (9, 4)$ in Tables 4.7 and F.1 and $\mathbf{S} = (8, 7)$ in Table F.1. Although considerably good performances can be attained by setting Δ to its best value, heuristic 2 (M) is still good in approximating the ς policy, even for unequal base-stock levels.

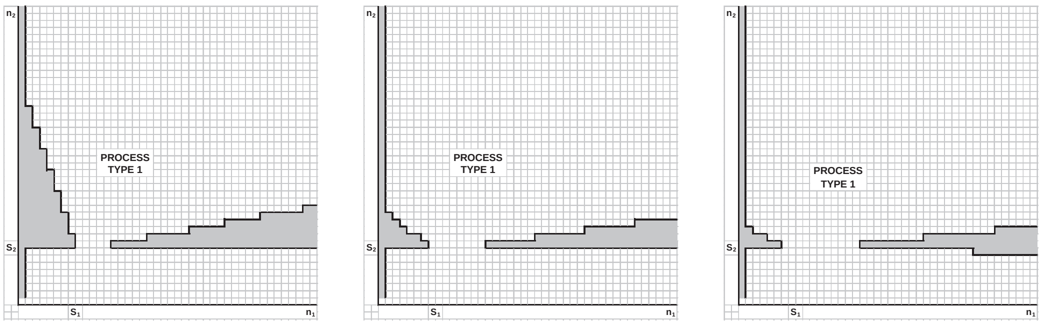
Table 4.4: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and the heuristics: equally weighted case, $\lambda_1 = 2\lambda_2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.10	ς	94.79	99.71	99.98	100.00	100.00	100.00	100.00	100.00
	FCFS	94.77	99.70	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1	**94.79**	**99.71**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2	**94.79**	**99.71**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
0.20	ς	89.20	98.78	99.86	99.98	100.00	100.00	100.00	100.00
	FCFS	89.01	98.68	99.83	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1	**89.20**	**98.78**	**99.85**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2	**89.20**	**98.78**	**99.85**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
0.30	ς	83.25	97.06	99.48	99.90	100.00	100.00	100.00	100.00
	FCFS	82.64	96.75	99.35	99.87	99.99	**100.00**	**100.00**	**100.00**
	Heuristic 1	**83.25**	**97.06**	**99.47**	**99.90**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2	**83.25**	**97.06**	**99.47**	**99.90**	**100.00**	**100.00**	**100.00**	**100.00**
0.40	ς	76.96	94.31	98.61	99.66	99.98	100.00	100.00	100.00
	FCFS	75.52	93.61	98.24	99.50	99.96	**100.00**	**100.00**	**100.00**
	Heuristic 1	**76.96**	**94.31**	**98.59**	**99.65**	**99.98**	**100.00**	**100.00**	**100.00**
	Heuristic 2	**76.96**	**94.31**	**98.59**	**99.65**	**99.98**	**100.00**	**100.00**	**100.00**
0.50	ς	70.34	90.23	96.86	99.00	99.90	99.99	100.00	100.00
	FCFS	67.50	88.88	96.02	98.52	99.78	99.97	**100.00**	**100.00**
	Heuristic 1	**70.34**	**90.23**	**96.81**	**98.98**	**99.90**	**99.99**	**100.00**	**100.00**
	Heuristic 2	**70.34**	90.22	**96.81**	**98.98**	**99.90**	**99.99**	**100.00**	**100.00**
0.60	ς	63.42	84.56	93.59	97.38	99.58	99.94	100.00	100.00
	FCFS	58.33	81.94	91.90	96.26	99.15	99.80	99.98	**100.00**
	Heuristic 1	**63.39**	84.44	93.45	97.32	**99.57**	**99.93**	**100.00**	**100.00**
	Heuristic 2	**63.39**	**84.51**	**93.49**	**97.34**	**99.57**	**99.93**	**100.00**	**100.00**
0.70	ς	56.21	77.12	87.94	93.72	98.37	99.59	99.95	100.00
	FCFS	47.69	71.90	84.54	91.30	97.11	98.99	99.78	99.97
	Heuristic 1	**56.16**	**77.00**	**87.71**	93.57	98.31	99.57	**99.95**	**100.00**
	Heuristic 2	**56.16**	76.90	**87.71**	**93.60**	**98.32**	**99.58**	**99.95**	**100.00**
0.80	ς	48.76	67.89	78.93	86.17	94.19	97.60	99.37	99.90
	FCFS	35.06	57.23	71.44	80.68	90.86	95.52	98.39	99.57
	Heuristic 1	**48.72**	**67.65**	78.51	85.81	93.93	97.46	99.33	**99.89**
	Heuristic 2	**48.72**	67.54	**78.54**	**85.93**	**94.06**	**97.54**	**99.35**	**99.89**
0.90	ς	41.12	56.93	65.78	72.25	81.72	88.00	93.64	97.27
	FCFS	19.64	35.14	47.42	57.19	71.27	80.43	88.71	94.38
	Heuristic 1	**41.07**	**56.76**	**65.31**	71.54	80.88	87.22	93.07	96.96
	Heuristic 2	**41.07**	56.45	65.20	**71.78**	**81.40**	**87.78**	**93.51**	**97.21**



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure 4.10: Heuristic 2: cost function weighted by demand rates, $S = 8$, $\lambda_1 = 2\lambda_2$.

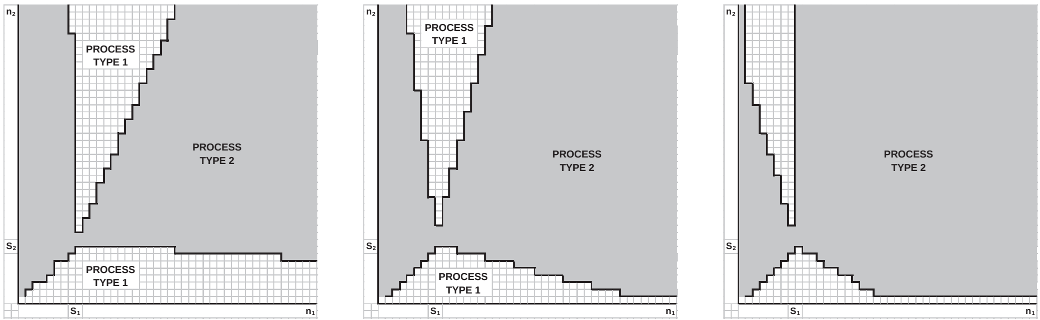


(a) $\rho = 0.10$. (b) $\rho = 0.40$. (c) $\rho = 0.60$.

Figure 4.11: Heuristic 2 (M) for $w_1 = 5w_2$, $S = 8$, $\lambda_1 = 2\lambda_2$.

4.3 SYSTEMS WITH MORE THAN TWO CLASSES

Extension of heuristic 1 with indices in Table 4.3 and heuristic 2 as its variation to the cases with more than two classes is, in fact, immediate. In each region there will be one index to be written for each class depending on whether it is in stockout or not. But, the difficulty arises due to the increase in the number of conditions, which are specified for each arrival, to decide whether the item in the server should be preempted or not. These conditions are to take into account all possible orderings of the indices and the class in service. Around 100 conditions are specified for the three-class system, and the number of conditions increase in the number of classes. A systematic



(a) $\rho = 0.10$. (b) $\rho = 0.40$. (c) $\rho = 0.90$.

Figure 4.12: Heuristic 2 (M) for $w_2 = 2w_1$, $S = 8$, $\lambda_1 = 2\lambda_2$.

way of generating all conditions for a given number of classes is not considered within the scope of the this thesis and left as a further work to handle more than three classes. All these point out the difficulty of implementing ς policy and heuristics 1 and 2, unless the conditions mentioned above are systematically generated.

In this study, we consider the symmetric three-class system to test accuracy of the proposed heuristics. Preliminary numerical experiments show that value-iteration cannot be employed due to computational restrictions. Starting with a very limited state-space, it is possible to obtain convergence only for low values of ρ. To overcome this limitation, simulation (Arena) is used. Minimum number of items processed is 200000 for each class in a single replication and 15 replications are made for each S and ρ combination. 95% confidence intervals are considered to ensure that the relative precisions of any half-width (ratio of half of the confidence interval to the mean) is not greater than 0.0005 for two-class systems and not greater than 0.002 for three-class systems. Simulation results are compared with the ones obtained by value-iteration for both two-class and three-class systems, but only with low traffic intensity cases for the three-class systems. Almost all simulation results are consistent with the ones obtained by value-iteration algorithm. Table I.1 in Appendix I shows this comparison for two-class systems.

Comparison of the aggregate fill rates under ς, FCFS, LQ policies and Heuristic 2 can be seen in Table 4.8 for symmetric three-class systems. The values for ς and LQ and heuristic 2 are calculated by value-iteration for ρ = 0.10, 0.25, 0.50. For

$\rho = 0.75$, 0.90 aggregate fill rates under ς policy could not be found due to the computational restrictions, and under LQ policy and for heuristic 2 simulation results are referred to. All the values for FCFS policy are calculated using the closed-form steady-state distribution. Accuracy of the proposed heuristics is still good when there are three classes.

Table 4.5: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristics: weighted by demand rates, $\lambda_1 = 2\lambda_2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.10	ς	94.29	99.66	99.98	100.00	100.00	100.00	100.00	100.00
	FCFS	**94.29**	**99.66**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1	**94.29**	**99.66**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1 (M)	**94.29**	**99.66**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2	**94.29**	**99.66**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	94.21	99.64	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
0.20	ς	88.24	98.58	99.82	99.98	100.00	100.00	100.00	100.00
	FCFS	**88.24**	**98.58**	**99.82**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1	88.23	**98.58**	**99.82**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1 (M)	**88.24**	**98.58**	**99.82**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2	88.23	**98.58**	**99.82**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	87.91	98.44	99.79	99.97	**100.00**	**100.00**	**100.00**	**100.00**
0.30	ς	81.81	96.63	99.37	99.88	100.00	100.00	100.00	100.00
	FCFS	**81.81**	**96.63**	**99.36**	**99.88**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1	81.80	**96.63**	**99.36**	**99.88**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 1 (M)	**81.81**	**96.63**	**99.36**	**99.88**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2	81.80	**96.63**	**99.36**	**99.88**	**100.00**	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	81.02	96.19	99.20	99.83	99.99	**100.00**	**100.00**	**100.00**
0.40	ς	74.99	93.59	98.37	99.59	99.97	100.00	100.00	100.00
	FCFS	**74.99**	**93.59**	**98.36**	**99.57**	**99.97**	**100.00**	**100.00**	**100.00**
	Heuristic 1	74.95	**93.59**	**98.36**	**99.57**	**99.97**	**100.00**	**100.00**	**100.00**
	Heuristic 1 (M)	**74.99**	**93.59**	**98.36**	**99.57**	**99.97**	**100.00**	**100.00**	**100.00**
	Heuristic 2	74.95	**93.59**	**98.36**	**99.57**	**99.97**	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	73.43	92.59	97.86	99.37	99.94	**100.00**	**100.00**	**100.00**
0.50	ς	67.72	89.18	96.40	98.83	99.88	99.99	100.00	100.00
	FCFS	**67.72**	**89.17**	**96.40**	**98.79**	**99.87**	**99.99**	**100.00**	**100.00**
	Heuristic 1	67.49	89.15	96.39	**98.79**	**99.87**	**99.99**	**100.00**	**100.00**
	Heuristic 1 (M)	**67.72**	**89.17**	**96.40**	**98.79**	**99.87**	**99.99**	**100.00**	**100.00**
	Heuristic 2	67.49	89.15	96.39	**98.79**	**99.87**	**99.99**	**100.00**	**100.00**
	Heuristic 2 (M)	65.00	87.25	95.21	98.16	99.72	99.96	**100.00**	**100.00**
0.60	ς	59.97	83.09	92.88	97.07	99.52	99.93	100.00	100.00
	FCFS	**59.96**	82.76	92.73	96.93	99.48	**99.91**	**99.99**	**100.00**
	Heuristic 1	59.61	82.77	92.74	96.93	99.48	**99.91**	**99.99**	**100.00**
	Heuristic 1 (M)	**59.96**	83.00	92.85	96.98	**99.49**	**99.91**	**99.99**	**100.00**
	Heuristic 2	59.61	**83.05**	**92.87**	**96.99**	**99.49**	**99.91**	**99.99**	**100.00**
	Heuristic 2 (M)	55.56	79.63	90.43	95.42	98.91	99.73	99.97	**100.00**
0.70	ς	51.66	74.98	86.96	93.26	98.24	99.56	99.95	100.00
	FCFS	44.84	68.92	82.17	89.63	96.38	98.70	99.71	99.96
	Heuristic 1	**51.54**	74.41	86.63	92.98	98.11	99.49	**99.93**	99.99
	Heuristic 1 (M)	50.98	74.71	86.86	93.09	98.14	99.49	**99.93**	99.99
	Heuristic 2	**51.54**	74.52	86.66	93.00	98.13	99.50	**99.93**	**100.00**
	Heuristic 2 (M)	50.98	**74.92**	**86.95**	**93.15**	**98.16**	**99.51**	**99.93**	**100.00**
0.80	ς	42.69	64.62	77.84	85.89	94.20	97.62	99.38	99.90
	FCFS	32.47	53.85	68.14	77.80	88.97	94.40	97.92	99.43
	Heuristic 1	**42.50**	63.39	76.67	84.99	93.71	97.32	99.25	99.86
	Heuristic 1 (M)	41.80	64.46	77.75	85.74	93.99	97.42	99.27	99.87
	Heuristic 2	**42.50**	63.36	76.71	85.04	93.76	97.37	99.28	99.87
	Heuristic 2 (M)	41.80	**64.52**	**77.84**	**85.80**	**94.06**	**97.48**	**99.30**	**99.88**
0.90	ς	32.95	52.01	64.79	73.24	83.56	89.48	94.48	97.64
	FCFS	17.86	32.27	43.95	53.47	67.63	77.24	86.36	92.95
	Heuristic 1	32.73	48.98	60.70	69.59	81.19	87.90	93.52	97.12
	Heuristic 1 (M)	32.22	51.79	64.42	72.82	83.17	89.08	94.04	97.31
	Heuristic 2	**32.74**	49.29	61.30	70.10	81.41	88.03	93.65	97.22
	Heuristic 2 (M)	32.22	**52.01**	**64.78**	**73.21**	**83.53**	**89.34**	**94.30**	**97.48**

Table 4.6: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): $\lambda_1 = 2\lambda_2$, $w_1 = 5w_2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	74.05	93.12	98.20	99.53	99.97	100.00	100.00	100.00
	FCFS	71.33	91.56	97.47	99.23	99.93	99.99	**100.00**	**100.00**
	Heuristic 2 (M)	**74.05**	**93.12**	**98.19**	**99.52**	**99.97**	**100.00**	**100.00**	**100.00**
0.60	ς	59.66	83.35	93.06	97.11	99.52	99.92	100.00	100.00
	FCFS	52.78	77.31	88.97	94.59	98.68	99.67	99.96	**100.00**
	Heuristic 2 (M)	**59.66**	**83.35**	**93.00**	**97.08**	**99.48**	**99.90**	**99.99**	**100.00**
0.80	ς	44.21	68.03	81.14	88.60	95.56	98.23	99.55	99.93
	FCFS	29.87	50.48	64.83	74.91	87.09	93.29	97.46	99.29
	Heuristic 2 (M)	**44.21**	**68.03**	**81.14**	**88.60**	**95.46**	**98.02**	**99.34**	**99.86**
0.90	ς	36.11	58.01	71.48	79.92	88.98	93.33	96.61	98.57
	FCFS	16.07	29.40	40.49	49.75	63.99	74.05	84.01	91.52
	Heuristic 2 (M)	**36.11**	**58.01**	**71.48**	**79.91**	**88.94**	**93.11**	**96.09**	**98.10**

Table 4.7: Comparison of the aggregate fill rates (%) of the ς, Δ policies and heuristic 2 (M): equally weighted case, $\lambda_1 = 2\lambda_2$.

		$\mathbf{S}=(6,3)$	$\mathbf{S}=(9,4)$	$\mathbf{S}=(11,11)$
ς		73.692	82.832	86.422
Heuristic 2		**73.590**	**82.766**	**86.399**
Δ	-1000	73.436	81.266	85.165
	-100	73.435	81.265	85.165
	-50	73.309	81.160	85.141
	-40	73.072	80.962	85.095
	-25	71.671	79.791	84.823
	-15	68.375	77.036	84.184
	-9	63.911	73.305	84.419
	-8	62.851	72.804	85.199
	-7	61.674	72.797	86.013
	-6	60.364	72.924	86.787
	-5	59.440	72.997	87.484
	-4	59.171	72.917	88.085
	-3	59.053	72.626	88.580
	-2	58.824	72.088	88.965
	-1	58.346	71.278	89.238
	0	57.544	70.182	89.399
	1	56.524	68.846	89.444
	2	55.736	67.459	89.373
	3	55.996	66.354	89.186
	5	57.985	66.745	88.479
	6	58.822	67.239	87.969
	7	59.570	67.683	87.373
	8	60.239	68.082	86.722
	9	60.838	68.439	86.080
	10	61.375	68.760	85.571
	15	63.336	69.939	86.058
	25	65.155	71.039	86.808
	40	65.922	71.504	87.125
	50	66.051	71.582	87.179
	100	66.119	71.624	87.208
	1000	66.120	71.624	87.208

Table 4.8: Comparison of the aggregate fill rates (%) of ς, FCFS, LQ policies and heuristic 2 for symmetric three-class systems.

ρ	S	ς	FCFS	LQ	H2
0.10	1	96.444	96.429	96.421	**96.444**
	2	99.880	99.872	**99.880**	**99.880**
0.25	1	90.262	90.000	89.863	**90.262**
	2	99.128	99.000	99.109	**99.119**
	3	99.925	99.900	99.924	**99.925**
0.50	1	77.480	75.000	73.766	**77.480**
	2	95.160	93.750	94.294	**94.819**
	3	99.002	98.438	98.867	**98.938**
	4	99.806	99.609	99.786	**99.796**
	6	99.993	99.902	**99.993**	**99.993**
0.75	1	-	50.000	46.345	**60.743**
	2	-	75.000	74.302	**81.533**
	3	-	87.500	88.339	**91.487**
	4	-	93.750	94.861	**96.208**
	6	-	98.438	99.043	**99.288**
	8	-	99.609	99.826	**99.871**
	11	-	99.951	99.987	**99.990**
0.90	1	-	25.000	21.390	**47.903**
	2	-	43.750	40.905	**66.445**
	3	-	57.813	56.349	**76.584**
	4	-	68.359	67.996	**83.162**
	6	-	82.202	82.944	**91.071**
	8	-	89.989	90.941	**95.263**
	11	-	95.776	96.490	**98.159**
	15	-	98.664	99.003	**99.482**

CHAPTER 5

RELATIONS BETWEEN COST AND SERVICE MODELS

This chapter is to question the problem formulations which justify the maximization of aggregate fill rate for multi-class base-stock controlled systems. Some questions addressed are the following: What kind of trade-offs can be resolved directly (or indirectly) with (implications of) the use of dynamic scheduling policy maximizing fill rate? Which trade-offs are missing concentrating on maximization of aggregate fill rate? Relevant service measure being fill rate (equivalently, α-service due to demand arrivals according to Poisson process), i.e., penalizing each backordered item regardless of how long it is backordered (the fraction of time with backorders regardless of the number of items backordered and the time required to satisfy each), how meaningful is it to work with base-stock controlled policies? Clearly, different formulations lead to different types of optimal policies. So, our analysis in this chapter would be to evaluate the performances of optimal policies for alternative formulations considered in the literature as compared to the ς policy proposed and approximated in Chapters 3 and 4, respectively, for maximization of aggregate fill rate. Section 5.1 is on the formulations for which the base-stock controlled dynamic scheduling policy investigated in Chapter 3 is optimal. Section 5.2 is to figure out differences between the policies proposed in Chapters 3 and 4 and other base-stock controlled policies in terms of relevant performance measures. Finally, in Section 5.3, the last question raised above is answered to a certain extent referring to some (equivalence) relations between different problem formulations.

5.1 FORMULATIONS FOR FILL RATE MAXIMIZATION

Let $\mathbf{S} = (S_1, ..., S_I)$ and let the base-stock controlled dynamic scheduling policy that maximizes aggregate fill rate for given $\mathbf{S}$ be denoted as $\varsigma_{\mathbf{S}}$ and consider the following problems:

$$\begin{aligned}
P_1(C): \ & Maximize\ FR(\pi_{\mathbf{S}}) \\
& subject\ to \\
& \sum_{i=1}^{I} c_i S_i \leq C, \\
& \pi_{\mathbf{S}} \in \overline{\Pi}, \\
& S_i \geq 0\ and\ integer,\ \ i = 1, ..., I,
\end{aligned}$$

$$\begin{aligned}
P_2(\beta): \ & Minimize\ \sum_{i=1}^{I} c_i S_i \\
& subject\ to \\
& FR(\pi_{\mathbf{S}}) \geq \beta, \\
& \pi_{\mathbf{S}} \in \overline{\Pi}, \\
& S_i \geq 0\ and\ integer,\ \ i = 1, ..., I,
\end{aligned}$$

where $\overline{\Pi}$ is the set of all stationary dynamic scheduling policies under base-stock control, C is the budget limitation and β is the target level for aggregate fill rate and $FR(\pi_{\mathbf{S}})$ is the aggregate fill rate under policy $\pi_{\mathbf{S}}$. Proposition 1 given below with its immediate proof is to show that $\varsigma_{\mathbf{S}}$ is optimal for problems $P_1(C)$ and $P_2(\beta)$.

Proposition 1.

a. Let $\vartheta_{\mathbf{S}^*}$ be optimal for problem $P_1(C)$. Then, $\vartheta_{\mathbf{S}^*} = \varsigma_{\mathbf{S}^*}$.

b. Let $\vartheta_{\mathbf{S}^*}$ be optimal for problem $P_2(\beta)$. Then, $\vartheta_{\mathbf{S}^*}$ and $\varsigma_{\mathbf{S}^*}$ are the same policies or alternative optima.

Proof.

a. In problem $P_1(C)$, for any feasible $\mathbf{S}$, aggregate fill rate FR is maximized under $\varsigma_{\mathbf{S}}$. That is among all $FR(\varsigma_{\mathbf{S}})$ with feasible $\mathbf{S}$, the maximizing $\mathbf{S}^*$ is optimal.

b. For problem $P_2(\beta)$, suppose $\vartheta_{\mathbf{S}^*}$ is different than $\varsigma_{\mathbf{S}^*}$. Since $\varsigma_{\mathbf{S}^*} = argmax_{\pi_{\mathbf{S}^*} \in \overline{\Pi}}\{FR(\pi_{\mathbf{S}^*})\}$ is feasible, i.e., $FR(\varsigma_{\mathbf{S}^*}) \geq FR(\vartheta_{\mathbf{S}^*}) \geq \beta$, $\varsigma_{\mathbf{S}^*}$ and $\vartheta_{\mathbf{S}^*}$ would be alternative optima.

$\square$

In the literature, we come across the use of $\sum_{i=1}^{I} c_i S_i$ to represent the inventory cost in spare part management. This quantity makes sense as the investment required for the spare parts to be kept in stock. But, in manufacturing systems use of expected average inventory holding cost is preferred. So, even if it would not be possible in general to justify working with $\sum_{i=1}^{I} c_i S_i$ in problems $P_1(C)$ and $P_2(\beta)$ which direct us to maximize aggregate fill rate and come up with $\varsigma_{\mathbf{S}^*}$ policies, there is at least the following indirect advantage of ς policy over other base-stock controlled policies: consider problem $P_2(\beta)$ for a given β, and let $\mathbf{S}$ be the best solution for $P_2(\beta)$ under some base-stock controlled policy π other than ς policy, e.g., LQ for the symmetric systems or Δ policies for the asymmetric systems or FCFS, (i.e., when π is restricted to the set of base-stock controlled LQ or FCFS policies). This advantage gets striking as the traffic intensity increases as explained in the next section referring to extensive numerical results. It is possible to put forward similar arguments to show the advantage of ς considering problem $P_1(C)$ for a given C value.

Note that, it is a common practice at managerial level to figure out the behaviour of efficient frontier for the investment required and the (aggregate) fill rate while trying to resolve the trade-off. That is, one can show performance of any $\mathbf{S}$ on a graph of the required investment versus one minus the aggregate fill rate under the corresponding optimal policy ($\varsigma_{\mathbf{S}}$ in this case). This is to be repeated for all possible $\mathbf{S}$. Then, the decision makers can see which $\mathbf{S}$ would be preferable on the efficient frontier depending on the budget limitation and the target service level.

5.2 COMPARISONS WITH ALTERNATIVE BASE-STOCK CONTROLLED POLICIES

The generic problem formulations considered in the literature are either to minimize expected average inventory holding and backorder costs or to minimize expected average inventory holding costs subject to a service level constraint. These are the cost and service models, respectively. The service model is preferred when there is the difficulty of estimating penalties (cost parameters) for backorders. Depending on the way backorder costs are incurred, the corresponding service levels are defined. Cost and service models with consistent backorder costs and service level constraints, respectively, are related in the literature for a number of rather simple single-item cases through the backorder penalties and the target service levels to be satisfied. Such related cost and service models are said to be equivalent, i.e., optimal policies for these models are the same. Section 5.3 is to elaborate on the equivalence relationships between the cost and service models of multi-class systems to place the policies maximizing aggregate fill rate (ς policies) among others in the literature.

With this perspective, problem $P_2(\beta)$ is generalized as follows:

$$\begin{aligned} P_3(\beta): \ & Minimize\ r(\pi_{\mathbf{S}}) \\ & subject\ to \\ & FR(\pi_{\mathbf{S}}) \geq \beta, \\ & \pi_{\mathbf{S}} \in \overline{\Pi}, \\ & S_i \geq 0\ and\ integer,\ i = 1, ..., I, \end{aligned}$$

where $r(\pi_{\mathbf{S}})$ is a cost function. A variation of problem $P_2(\beta)$ that we can think of in the first place is obtained letting $r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$, where h_i is the inventory holding cost per item of type i per time unit and $Inv_i(\pi_{\mathbf{S}})$ is the expected average inventory of type i under policy $\pi_{\mathbf{S}}$, i.e., $E_{\pi_{\mathbf{S}}}(\overline{N}_i)$ recalling the notation introduced in Chapter 3. Note that the arguments in the proof of Proposition 1 do not work for problem $P_3(\beta)$, i.e., when $\sum_{i=1}^{I} c_i S_i$ in problem $P_2(\beta)$ is replaced with $r(\pi_{\mathbf{S}})$, because r is a function of π unlike $\sum_{i=1}^{I} c_i S_i$.

As mentioned in Chapter 2, most of the studies on multi-class make-to-stock systems in the literature are to minimize expected average inventory holding and backorder costs. This problem is as below.

$$P_4: \ \min_{\pi \in \Pi} \{ \sum_{i=1}^{I} h_i Inv_i(\pi) + \sum_{i=1}^{I} b_i B_i(\pi) \}$$

where Π is the set of all feasible policies, $B_i(\pi_{\mathbf{S}})$ is the expected average backorders of type i under policy $\pi_{\mathbf{S}}$, i.e., $E_{\pi_{\mathbf{S}}}(K_i)$ recalling the notation introduced in Chapter 3, and b_i is the penalty charged for each time unit a request is backordered. For the symmetric systems with $\lambda_i = \lambda$, $h_i = h$ and $b_i = b$ for all i operating under base-stock policies (i.e., when Π is replaced with $\overline{\Pi}$), LQ policy with $S_i = S$ is optimal as pointed out by Veatch and Wein in [19]. Ha [8] later shows that base-stock controlled LQ policies are optimal for symmetric systems. Then, it is enough to concentrate on $\overline{\Pi}$ in the problem above instead of Π when the system is symmetric. Note that $\overline{\Pi} \subset \Pi$.

The cost model above and problem $P_3(\beta)$ with $r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$ (even replacing $\overline{\Pi}$ by Π and $\pi_{\mathbf{S}}$ by π in P_4) cannot be related directly in the sense mentioned in the first paragraph of this section. In other words, backorder penalties are not consistent in these two models. In problem P_4, it is charged for each time unit needed to satisfy a backordered request whereas fill rate in the service model is to represent the case penalty is charged per backordered item regardless of the backordering duration. In the literature, for symmetric systems alternative policies FCFS and LQ are tried for problem P_4 in [22] and for fill rate being the relevant performance measure in [17], without questioning (equivalence) relations between the cost and service models. Similarly, instead of getting into equivalence relations, here we compare these alternative policies studied in the literature and ς policy. The comparisons are for the symmetric systems (otherwise, instead of LQ policy we should maybe proceed with Δ policies) with equal demand rates, cost figures and fill rate weights, and to see the performances of LQ, FCFS and ς policies as feasible alternative solutions of problem $P_3(\beta)$ with their respective minimum base-stock levels required to achieve target level β and when $r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$.

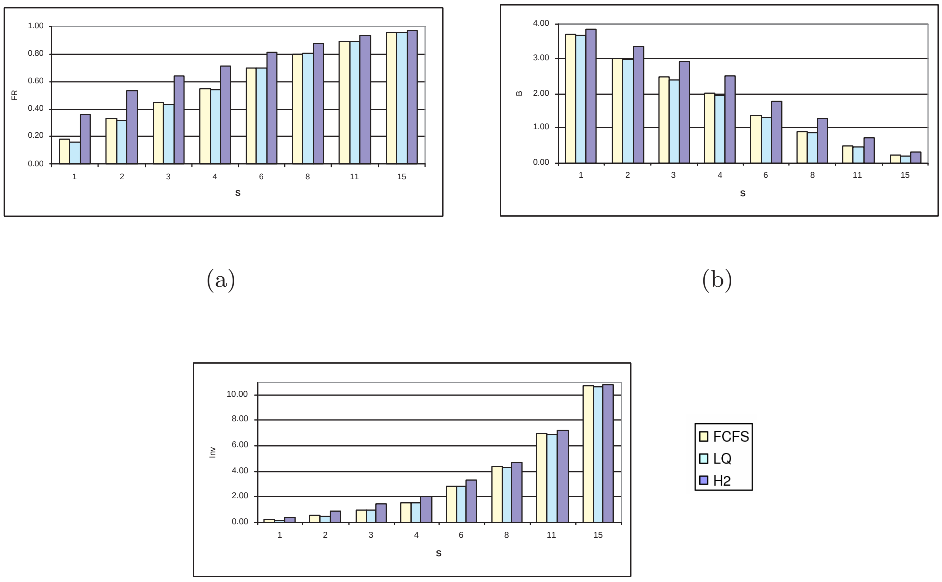
Note that a feasible policy π satisfying the fill rate constraint of problem $P_3(\beta)$ with minimum base-stock levels is, in fact, the solution of problem $P_2(\beta)$ under the set of base-stock controlled π policies when $c_i = c$. By optimality of ς policy in problem $P_2(\beta)$, $\varsigma_{\mathbf{S}^*}$ gives the base-stock level combination with minimum $\sum_{i=1}^{I} c_i S_i$ (minimum $\sum_{i=1}^{I} S_i$ when $c_i = c$). If the minimum base-stock levels required to achieve target level β are considerably smaller under ς than the minimum levels required under LQ and FCFS policies, then the expected average inventory holding cost will be expected to be smaller under $\varsigma_{\mathbf{S}*}$ as compared to other policies. This sort of an indirect or implied advantage of ς policy over LQ and FCFS is observed numerically although, in fact, it is known that the proposed ς policy that maximizes the aggregate fill rate is worse than LQ with respect to the expected average inventory holding and backorder costs (considering the case $b_i = 0$) when base-stock levels are the same for both of these policies.

In Table 5.1, for the symmetric two-class systems with $\lambda_i = \lambda$, $h_i = h$, $w_i = w$ and $c_i = c$, FCFS and LQ policies and heuristics 2 and 3 are compared. Only heuristics 2 and 3 are considered in this table to approximate the behaviour of ς policy because heuristic 2 is the best as reported in Chapter 4 and heuristic 3 is the only one for which the steady-state probabilities can be calculated instantaneously using the recursive scheme devised in Section 4.1 for symmetric systems. For any given β and ρ combination in this table, S_i values are the minimum base-stock levels to satisfy target fill rate under the considered policy and $Inv_i = Inv$, $i = 1,\ 2$, is the corresponding expected average inventory given these base-stock levels. The following observations are immediate from Table 5.1.

- Heuristic 2 attains target fill rates with minimum base-stock levels. This is an expected result because heuristic 2 is the best approximation for the optimal ς policy of problem $P_2(\beta)$.
- When the traffic intensity, ρ, is small, minimum S_i values required for given β are almost always the same under the alternative policies and LQ gives the minimum expected average inventory among the three policies. However, as ρ increases advantage of heuristic 2 and 3 over other policies becomes apparent to guarantee fill rate with smaller S_i values thus to give and with smaller expected average inventory values.

In order to further elaborate on the use of aggregate fill rate, one can refer to Figure 5.1 and the tables in Appendix J. This figure and the tables are for the symmetric systems with $\lambda_i = \lambda$, $h_i = h$, $b_i = b$ and $S_i = S$. As the LQ policy is optimal to minimize expected average inventories and/or expected average backorders in such systems operating under base-stock policies as previously noted referring to [19], (letting h and b take values of zero and one alternatively) LQ dominates other policies in Table J.1 with respect to these performance measures but not fill rate. Our explanation for this is that, for the sake of avoiding stockout of a class, under $\varsigma_{\mathbf{S}}$ for given $\mathbf{S}$ we may allow the number of backorders of the other class to reach higher values if it is already in stockout, causing the expected average backorders to be high as compared to LQ policy. This disadvantage may be overcome considering γ type service level constraints, maybe one constraint for each type of items instead of an aggregate service level, but then an increase in the required base-stock levels and so in the expected average inventory would be expected. Having an insight into all such trade-offs, one should work with the appropriate formulation.

Note that above we think of the comparison of policy ς, which is supposed to be a good solution for P_3, with others, especially with LQ, in terms of expected average backorders because LQ minimizes expected average backorders in P_4 when $h = 0$ and $b = 1$. But, in fact, P_4 and P_3 are not related. The way backorder penalty is charged in problem P_4 where expected average backorders appear in the objective function is not in accordance with the service level constraint (aggregate fill rate) in problem $P_3(\beta)$. This is further clarified in the following section.



(a) (b)

(c)

Figure 5.1: Performance measures for $\lambda_i = \lambda$, $w_i = \frac{1}{2}$, $S_i = S$ for $i = 1, 2$ and $\rho = 0.90$.

Next, for the symmetric two-class systems, some analytical results are given with Propositions 2 and 3 and a remark in order to guide one in enumerating $\mathbf{S}$ to find the smallest base-stock levels under one of the considered policy as in the case of filling out Table 5.1. Although the anaytical results may not hold under each of ς, FCFS and LQ policies, sometimes supporting our arguments with numerical observations we proceed with the guidelines for all of these policies. Hence, we somehow mean solving problem $P_3(\beta)$, but not only solving $P_2(\beta)$ (with $c_i = c$) for the "best" $\mathbf{S}$ under the given scheduling policy. Optimality of this $\mathbf{S}$ may not be guaranteed when the considered policy is ς or LQ, but FCFS, as clarified below.

Table 5.1: Comparison of the alternative policies for problem $P_3(\beta)$.

		FCFS				LQ				Heuristic 2				Heuristic 3			
β	ρ	S_1	S_2	$FR(\%)$	Inv	S_1	S_2	$FR(\%)$	Inv	S_1	S_2	$FR(\%)$	Inv	S_1	S_2	$FR(\%)$	Inv
0.90	0.25	**2**	**1**	91.837	1.347	**2**	**1**	91.807	**1.346**	**2**	**1**	92.462	**1.346**	**2**	**1**	92.462	**1.346**
	0.50	3	2	92.593	2.037	3	2	92.936	2.030	**2**	**2**	90.139	1.562	**2**	**2**	90.070	**1.555**
	0.60	**3**	**3**	92.128	2.309	**3**	**3**	92.681	**2.296**	**3**	**3**	93.616	2.313	**3**	**3**	93.493	2.304
	0.75	5	5	92.224	3.617	5	4	90.259	3.127	**4**	**4**	90.458	**2.745**	5	4	92.136	3.138
	0.90	12	11	90.001	7.450	12	11	90.502	7.405	**9**	**9**	90.000	**5.545**	11	10	90.924	6.523
	0.95	24	23	90.470	14.905	23	23	90.284	14.399	**18**	**17**	90.093	**10.684**	21	20	90.427	12.195
0.95	0.25	**2**	**2**	97.959	1.837	**2**	**2**	98.085	**1.836**	**2**	**2**	98.115	**1.836**	**2**	**2**	98.115	**1.836**
	0.50	**3**	**3**	96.296	2.519	**3**	**3**	96.741	**2.513**	**3**	**3**	96.989	2.517	**3**	**3**	96.971	2.516
	0.60	4	4	96.626	3.275	4	4	97.167	3.267	**4**	**3**	95.837	2.789	**4**	**3**	95.764	**2.784**
	0.75	6	6	95.334	4.570	6	6	95.990	4.552	**6**	**5**	95.864	4.108	**6**	**5**	95.532	**4.078**
	0.90	15	15	95.071	10.722	15	15	95.482	10.693	**13**	**12**	95.214	**8.533**	14	13	95.177	9.278
	0.95	30	30	95.034	20.972	30	29	95.006	20.462	**24**	**24**	95.075	**15.850**	27	27	95.081	18.112
0.99	0.25	3	3	99.708	2.834	3	3	99.755	2.834	**3**	**2**	99.025	**2.335**	**3**	**2**	99.025	**2.335**
	0.50	5	4	99.177	4.004	**4**	**4**	99.052	**3.504**	**4**	**4**	99.113	3.505	**4**	**4**	99.109	**3.504**
	0.60	6	6	99.380	5.255	6	5	99.258	4.754	**5**	**5**	99.042	4.259	**5**	**5**	99.026	**4.257**
	0.75	10	9	99.194	8.012	9	9	99.278	7.509	**8**	**8**	99.007	**6.526**	9	8	99.194	7.014
	0.90	23	23	99.010	18.545	23	22	99.065	18.040	**20**	**20**	99.014	**15.617**	21	21	99.007	16.557
	0.95	47	46	99.046	37.091	46	45	99.033	36.090	**40**	**40**	99.073	**30.810**	43	43	99.047	33.619

Proposition 2. *Given a policy $\pi_{\mathbf{S}} \in \overline{\Pi}$ for symmetric two-class systems with $\lambda_i = \lambda$, $h_i = h$ and $w_i = w$ for $i = 1,\ 2$, let the following conditions be satisfied.*
Condition 1. $FR(\pi_{\mathbf{S}})$ is concave in $\mathbf{S}$,
Condition 2. $r(\pi_{\mathbf{S}})$ is convex in $\mathbf{S}$,
Condition 3. $FR(\pi_{\mathbf{S}})$ is increasing in $\mathbf{S}$,
Condition 4. $r(\pi_{\mathbf{S}})$ is increasing in $\mathbf{S}$.
Suppose $FR(\pi_{(s,s)}) \geq \beta$ and $FR(\pi_{(s-1,s-1)}) < \beta$. Then, the solution for problem $P_3(\beta)$ under $\pi_{\mathbf{S}}$ is either $\mathbf{S} = (s,s)$ or $\mathbf{S} = (s, s-1)$ or $\mathbf{S} = (s, s-1)$.

Proof. Since $FR(\pi_{\mathbf{S}})$ is concave, it satisfies the following inequality for any $0 \leq \kappa \leq 1$, $\mathbf{S}'$ and $\mathbf{S}''$:

$$FR(\pi_{\kappa\mathbf{S}'+(1-\kappa)\mathbf{S}''}) \geq \kappa FR(\pi_{\mathbf{S}'}) + (1-\kappa) FR(\pi_{\mathbf{S}''}).$$

a. Then, letting $\mathbf{S}' = (s+k, s-k)$ and $\mathbf{S}'' = (s-k, s+k)$, with $-s \leq k \leq s$, and $\kappa = 0.5$ and noting that $FR(\pi_{(s+k,s-k)}) = FR(\pi_{(s-k,s+k)})$ because the system is symmetric with $\lambda_1 = \lambda_2$ and $w_1 = w_2$, we obtain

$$FR(\pi_{(s,s)}) \geq FR(\pi_{(s+k,s-k)}) = FR(\pi_{(s-k,s+k)})$$

for all s and $-s \leq k \leq s$, which implies $FR(\pi_{(s,s)}) \geq FR(\pi_{(S_1,S_2)})$ for all S_1 and S_2 such that $S_1 + S_2 = 2s$. Similarly, since $\lambda_i = \lambda$, $h_i = h$, $r(\pi_{(s,s)}) \leq r(\pi_{(S_1,S_2)})$ for all S_1 and S_2 such that $S_1 + S_2 = 2s$.

b. Also, letting $\mathbf{S}' = (s+k, s-1-k)$ and $\mathbf{S}'' = (s-1-k, s+k)$ with $0 \leq k \leq s-1$, and $\kappa = \frac{1+k}{1+2k}$, we obtain

$$FR(\pi_{(s,s-1)}) \geq FR(\pi_{(s+k,s-1-k)}) = FR(\pi_{(s-1-k,s+k)})$$

for all s and $0 \leq k \leq s-1$. Then, $FR(\pi_{(s,s-1)}) \geq FR(\pi_{(S_1,S_2)})$ for all S_1 and S_2 such that $S_1 > S_2$ and $S_1 + S_2 = 2s-1$. Note that $FR(S_1, S_2) = FR(S_2, S_1)$ because the system is symmetric with $\lambda_1 = \lambda_2$ and $w_1 = w_2$, then the $FR(\pi_{(s,s-1)}) = FR(\pi_{(s-1,s)}) \geq FR(\pi_{(S_1,S_2)})$ holds for all S_1 and S_2 such that $S_1 + S_2 = 2s-1$. Similarly, $r(\pi_{(s,s-1)}) = r(\pi_{(s-1,s)}) \leq r(\pi_{(S_1,S_2)})$ for all S_1 and S_2 such that $S_1 + S_2 = 2s-1$

It is given that $FR(\pi_{(s,s)}) \geq \beta$ and $FR(\pi_{(s-1,s-1)}) < \beta$. From Condition 3, $FR(\pi_{(s-1,s-1)}) \leq FR(\pi_{(s,s-1)})$. Then, we are to consider the following cases.
If $FR(\pi_{(s,s-1)}) < \beta$, then (s,s) is the best for π in problem $P_3(\beta)$. This is shown comparing (s,s) with every $\mathbf{S}$.

- Let $\mathbf{S} \in \{\mathbf{S} | S_1 + S_2 \geq 2s\}$.
  If $S_1 + S_2$ is even, then $r(\pi_{(\frac{S_1+S_2}{2}+k, \frac{S_1+S_2}{2}-k)}) \geq r(\pi_{(\frac{S_1+S_2}{2}, \frac{S_1+S_2}{2})}) \geq r(\pi_{(s,s)})$ for all $-\frac{S_1+S_2}{2} \leq k \leq \frac{S_1+S_2}{2}$ where the first and second inequalities result from (a) and Condition 4, respectively.
  If $S_1 + S_2$ is odd, then $r(\pi_{(\frac{S_1+S_2+1}{2}+k, \frac{S_1+S_2-1}{2}-k)}) \geq r(\pi_{(\frac{S_1+S_2+1}{2}, \frac{S_1+S_2-1}{2})}) \geq r(\pi_{(s,s)})$ for all $0 \leq k \leq \frac{S_1+S_2-1}{2}$ where the first and second inequalities result from (b) and Condition 4, respectively.
  Then, even if $\mathbf{S}$ satisfying $|\mathbf{S}| \geq 2s$ is feasible, it is not optimal for $P_3(\beta)$.

- Let $\mathbf{S} \in \{\mathbf{S} | S_1 + S_2 < 2s\}$.
  If $S_1 + S_2$ is even, then $FR(\pi_{(\frac{S_1+S_2}{2}+k, \frac{S_1+S_2}{2}-k)}) \leq FR(\pi_{(\frac{S_1+S_2}{2}, \frac{S_1+S_2}{2})}) \leq FR(\pi_{(s-1,s-1)}) < \beta$ for all $-\frac{S_1+S_2}{2} \leq k \leq \frac{S_1+S_2}{2}$ where the first and second inequalities result from (a) and Condition 3 respectively.
  If $S_1 + S_2$ is odd, then $FR(\pi_{(\frac{S_1+S_2+1}{2}+k, \frac{S_1+S_2-1}{2}-k)}) \leq FR(\pi_{(\frac{S_1+S_2+1}{2}, \frac{S_1+S_2-1}{2})}) \leq FR(\pi_{(s,s-1)}) < \beta$ for all $0 \leq k \leq \frac{S_1+S_2-1}{2}$ where the first and second inequalities result from (b) and Condition 3, respectively.
  Then, $\mathbf{S}$ satisfying $|\mathbf{S}| < 2s$ is not feasible for problem $P_3(\beta)$.

If $FR(\pi_{(s,s-1)}) \geq \beta$, then $(s, s-1)$ or equivalently $(s-1, s)$ is the best for π in $P_3(\beta)$.

- Let $\mathbf{S} \in \{\mathbf{S} | S_1 + S_2 \geq 2s - 1\}$.
  If $S_1 + S_2$ is even, then $r(\pi_{(\frac{S_1+S_2}{2}+k, \frac{S_1+S_2}{2}-k)}) \geq r(\pi_{(\frac{S_1+S_2}{2}, \frac{S_1+S_2}{2})}) \geq r(\pi_{(s,s-1)})$ for all $-\frac{S_1+S_2}{2} \leq k \leq \frac{S_1+S_2}{2}$ where the first and second inequalities result from (a) and Condition 4, respectively.
  If $S_1 + S_2$ is odd, then $r(\pi_{(\frac{S_1+S_2+1}{2}+k, \frac{S_1+S_2-1}{2}-k)}) \geq r(\pi_{(\frac{S_1+S_2+1}{2}, \frac{S_1+S_2-1}{2})}) \geq r(\pi_{(s,s-1)})$ for all $0 \leq k \leq \frac{S_1+S_2-1}{2}$ where the first and second inequalities result from (b) and Condition 4, respectively.
  Then, even if $\mathbf{S}$ satisfying $|\mathbf{S}| \geq 2s - 1$ is feasible, it is not optimal for $P_3(\beta)$.

- Let $\mathbf{S} \in \{\mathbf{S} | S_1 + S_2 < 2s - 1\}$.
  If $S_1 + S_2$ is even, then $FR(\pi_{(\frac{S_1+S_2}{2}+k, \frac{S_1+S_2}{2}-k)}) \leq FR(\pi_{(\frac{S_1+S_2}{2}, \frac{S_1+S_2}{2})}) \leq FR(\pi_{(s-1,s-1)}) < \beta$ for all $-\frac{S_1+S_2}{2} \leq k \leq \frac{S_1+S_2}{2}$ where the first and second inequalities result from (a) and Condition 3, respectively.
  If $S_1 + S_2$ is odd, then $FR(\pi_{(\frac{S_1+S_2+1}{2}+k, \frac{S_1+S_2-1}{2}-k)}) \leq FR(\pi_{(\frac{S_1+S_2+1}{2}, \frac{S_1+S_2-1}{2})}) \leq FR(\pi_{(s-1,s-1)}) < \beta$ for all $0 \leq k \leq \frac{S_1+S_2-1}{2}$ where the first and second inequalities result from (a) and Condition 3, respectively.
  Then, $\mathbf{S}$ satisfying $|\mathbf{S}| < 2s - 1$ is not feasible for problem $P_3(\beta)$.

$\square$

Note that if the policy employed is independent of $\mathbf{S}$ as in the case of FCFS and LQ policies, then the steady-state distribution depends on only the policy employed and is independent of the choice of base-stock levels. Then, the performance measures for an item type, under such a policy, depend on only the base-stock level of this type, i.e., $FR(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} w_i FR_i(\pi_{S_i})$ and $r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} r_i(\pi_{S_i})$ if r is separable ($r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} r_i(\pi_{\mathbf{S}})$). We will call such policies as $\mathbf{S}$-independent policies. For policies like $\varsigma_{\mathbf{S}}$, on the other hand, the steady-state distribution depends on $\mathbf{S}$, then the performance measures of an item type are also dependent on the base-stock levels of the other types of items.

Proposition 3. *If $r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$ and $\lambda_i = \lambda$ and $h_i = h$, then Condition 2 is satisfied for problem $P_3(\beta)$under an $\mathbf{S}$-independent policy $\pi_{\mathbf{S}}$.*

Proof. Since $\pi_{\mathbf{S}}$ is $\mathbf{S}$-independent in the sense mentioned above, $Inv_i(\pi_{\mathbf{S}}) = Inv_i(\pi_{S_i})$. Then, dropping subscript i for the sake of keeping notation simple, and letting $\Delta_S = Inv(\pi_S) - Inv(\pi_{S-1})$,

$$\begin{aligned}
\Delta_{S+1} &= \sum_{n=0}^{S+1} (S+1-n)Pr(N=n) - \sum_{n=0}^{S} (S-n)Pr(N=n) \\
&= (\sum_{n=0}^{S} (S-n)Pr(N=n) + \sum_{n=0}^{S} Pr(N=n) - \sum_{n=0}^{S} (S-n)Pr(N=n) \\
&= \sum_{n=0}^{S} Pr(N=n).
\end{aligned}$$

Then, $\Delta_{S+1} - \Delta_S = \sum_{n=0}^{S} Pr(N=n) - \sum_{n=0}^{S-1} Pr(N=n) = Pr(N=S) \geq 0$, which shows that $Inv(\pi_S)$ is convex in S. Since sum of convex functions is also convex, the result follows. $\square$

Remark. *Consider problem $P_2(\beta)$ with $\lambda_1 = \lambda_2$, $c_1 = c_2$, $w_1 = w_2$ when $I = 2$. If Conditions 1 and 3 are satisfied under policy $\varsigma_{\mathbf{S}}$, then Proposition 2 holds for problem $P_2(\beta)$ under $\varsigma_{\mathbf{S}}$. Note that Conditions 2 and 4 are already satisfied because $\sum_{i=1}^{I} c_i S_i$ is linear in $\mathbf{S}$. (Recall that, for any $\mathbf{S}$, if there is a feasible policy for problem $P_2(\beta)$, then $\varsigma_{\mathbf{S}}$ is feasible with the best aggregate fill rate. This explains why we can restrict our arguments in this remark to ς policies.)* $\square$

When π is FCFS policy, $FR(\pi_{\mathbf{S}}) = 1 - \frac{\hat{\rho}_1^{S_1} + \hat{\rho}_2^{S_2}}{2}$, where $\hat{\rho}_i = \frac{\lambda_i}{\mu - \sum_{j \neq i} \lambda_j}$ for $i = 1,\ 2$, (see the reference to [5] in Section 3) is concave in $\mathbf{S}$ and $Inv_i(\pi_{S_i})$ is convex in $\mathbf{S}$ (from Proposition 3). Thus, having analytically observed that $FR(\pi_{\mathbf{S}})$ is concave and $\sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$ is convex in $\mathbf{S}$ when π is FCFS policy, Proposition 2 holds for this case. When π is LQ policy, convexity of $\sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$ is guaranteed analytically (from Proposition 3) but concave behaviour of $FR(\pi_{\mathbf{S}})$ is observed only numerically. And when π is ς policy, then the same behaviours for $FR(\pi_{\mathbf{S}})$ and $\sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$ can only be observed numerically. See figures in Appendix K. Based on all these, referring to Proposition 2 and 3, we fill out Table 5.1 with the "best" $\mathbf{S}$ checking $S_1 = S_2 = s$ values to satisfy target fill rate β for any given ρ until $FR(\pi_{(s,s)}) \geq \beta$ but $FR(\pi_{(s-1,s-1)}) < \beta$ is observed. Then, we check also $\mathbf{S} = (s, s-1)$ or equivalently $\mathbf{S} = (s-1, s)$.

5.3 ON THE OPTIMAL POLICY OF THE SERVICE MODEL

In this section, optimal policy for the generalized version of problem $P_3(\beta)$ is investigated using the equivalence relations between cost and service models. Recall that cost models are to optimize any kind of costs such as inventory holding costs, ordering costs and penalties for backorders. In service models, on the other hand, a constraint on the service level is introduced instead of handling backorder costs in the objective function. The reader is referred to [18] for the relations between cost and service models of general inventory systems. For a cost and a service model to be related in the case of a single-item system, Condition 1 below given in [18] needs to be satisfied. $a_{pen}(\pi)$ is defined as the expected number of times penalty b is paid per time unit and $\theta(\pi)$ is the service level under policy π, and Π is the set of all feasible policies. $ba_{pen}(\pi)$ appears in the objective function of the cost model as a part of the expected average cost. In the service model, $\theta(\pi)$ is bounded below with a target service level θ instead of incurring $ba_{pen}(\pi)$ but keeping all the other cost terms in the objective function.

Condition 1. *There exists a constant $K > 0$ such that $a_{pen}(\pi) = K(1 - \theta(\pi))$, or equivalently $\theta(\pi) = 1 - \frac{a_{pen}(\pi)}{K}$, for all $\pi \in \Pi$.*

As explained in [18], cost and service models can be related to one another for different combinations of the backorder penalties and service measures. Review type, continuous or periodic, and whether backordering is allowed or not, i.e., the way penalty is incurred and service level is measured, determine these combinations. When the system is continuously reviewed and backordering is allowed as in the case we consider in this study, the following three combinations of the penalties and service measures arise to satisfy Condition 1 above for the single-item systems:

i. a penalty of b is paid for the fraction of time with backorders, the related service measure is of type α;

ii. a penalty of b is paid for each backordered item, the related service measure is of type β (fill rate);

iii. a penalty of b is paid for each time unit an item is backordered, the related service measure is of type γ (modified fill rate).

More on these service levels can be found in Silver et al. [13]. Based on the equivalence relations, van Houtum and Zijm [18] give the following theorem.

Theorem 1. *If*

i. *Condition 1 is satisfied,*

ii. π^* *is an optimal policy for the cost model,*

iii. $\theta(\pi^*) = \theta_0$,

then π^* *is also an optimal policy for the service model with target service level being* θ_0.

Now, consider the following generalized version of problem $P_3(\beta)$ for a given target service level β:

$$\begin{aligned} P_5(\beta): \ & Minimize \sum_{i=1}^{I} h_i Inv_i(\pi) \\ & subject\ to \\ & FR(\pi) \geq \beta, \\ & \pi \in \Pi. \end{aligned}$$

Problem $P_5(\beta)$ is a service model where an overall (aggregate) fill rate is considered as service measure of the multi-class systems we focus on. Note that $FR(\pi) \geq \beta$ is, in fact, a relaxation of $FR_i(\pi) \geq \beta_i$, $i = 1, ..., I$ if β is chosen as $\sum_{i=1}^{I} w_i \beta_i$. (Multiplying the fill rate constraint of each class by w_i and then summing up these constraints, we end up with the single constraint on aggregate fill rate.) That is, when $\beta = \sum_{i=1}^{I} w_i \beta_i$, the feasible region in the formulations we consider in this thesis is larger than the ones in [3] and [7].

The cost model corresponding to problem $P_5(\beta)$ turns out to be one of two types depending on the backorder costs incurred, for the fraction of demand backordered or the fraction of time with backorders. Then, a related cost model can be constructed referring to combination (ii) above if FR is regarded as the weighted average of fill rates (β-service levels) of different types of items and another cost model would be for (i) regarding FR as the weighted average of α-service levels of different item types. Denote the weighted average of α-service levels, i.e., aggregate α-service level, by $\alpha(\pi)$. Note that $FR(\pi)$ and $\alpha(\pi)$ are the same for a given system due to the independent Poisson arrivals of different item types. Definition of $a_{pen}(\pi)$ is generalized to handle different types of items letting $a_i(\pi)$ denote the expected number of times penalty cost for item type i, i.e., b_i, is paid per time unit, $i = 1, ..., I$, under policy π. Then, $ba_{pen}(\pi)$ in Condition 1 is $\sum_{i=1}^{I} b_i a_i(\pi)$ for multi-class systems. Note that $b_i a_i(\pi)$ is $b_i\lambda_i(1-FR_i(\pi))$ and $b_i(1-FR_i(\pi))$ for cases (ii) and (i), respectively. The cost model mentioned above is

$$P_6(\mathbf{b}): \ \min_{\pi\in\Pi}\{\sum_{i=1}^{I} h_i Inv_i(\pi) + \sum_{i=1}^{I} b_i a_i(\pi)\}$$

where $\mathbf{b} = (b_1, ..., b_I)$. Similarly, let $\mathbf{h} = (h_1, ..., h_I)$ to be used later.

Next, Condition 2 is given for cost and service models (problems $P_6(\mathbf{b})$ and $P_5(\beta)$, respectively) of the multi-class systems to be related when backorder penalties and service levels are of α or β type.

Condition 2. *There exists a constant $K > 0$ such that $\sum_{i=1}^{I} b_i a_i(\pi) = K(1-FR(\pi))$, for all $\pi \in \Pi$.*

Proposition 4. *If a penalty is paid for each backordered request (for the fraction of time with backorders) and w_i, $i = 1, ..., I$, in problem P_5 are chosen to satisfy $\frac{w_i}{w_j} = \frac{b_i\lambda_i}{b_j\lambda_j}$ ($\frac{w_i}{w_j} = \frac{b_i}{b_j}$) for all i, j, then Condition 2 holds for aggregate fill rate (aggregate α-service level) with $K = \frac{b_i\lambda_i}{w_i}$ ($K = \frac{b_i}{w_i}$).*

Proof. If a penalty is paid for each backordered request,

$$\begin{aligned}
\sum_{i=1}^{I} b_i a_i(\pi) &= \sum_{i=1}^{I} b_i\lambda_i(1 - FR_i(\pi)) \\
&= \sum_{i=1}^{I} \frac{b_i\lambda_i}{w_i}(w_i - w_i FR_i(\pi)) \\
&= \frac{b_1\lambda_1}{w_1}(1 - FR(\pi)) \quad when\ \frac{b_1\lambda_1}{w_1} = \frac{b_i\lambda_i}{w_i} \text{ for all } i.
\end{aligned}$$

Proof similarly follows when a penalty is paid for the fraction of time with backorders.

$\square$

Corollary. *Suppose a penalty is paid for each backordered request (for the fraction of time with backorders), and* w_i*,* $i = 1, ..., I$*, in problem* P_5 *are chosen to satisfy* $\frac{w_i}{w_j} = \frac{b_i\lambda_i}{b_j\lambda_j}$ $(\frac{w_i}{w_j} = \frac{b_i}{b_j})$ *for all* i*,* j*.*

i. Then, Condition 2 is equivalent to $FR(\pi) = 1 - \sum_{i=1}^{I} w_i(\frac{a_i(\pi)}{\lambda_i})$ $(\alpha(\pi) = 1 - \sum_{i=1}^{I} w_i a_i(\pi))$ *where* $w_i = \frac{b_i\lambda_i}{\sum_{j=1}^{I} b_j\lambda_j}$ $(w_i = \frac{b_i}{\sum_{j=1}^{I} b_j})$*.*

ii. If also $b_i = b_j$ *for all* i*,* j*, then* $FR(\pi) = 1 - \frac{\sum_{i=1}^{I} a_i(\pi)}{\sum_{i=1}^{I}\lambda_i}$ $(\alpha(\pi) = 1 - \frac{\sum_{i=1}^{I} a_i(\pi)}{I})$ *where* $w_i = \frac{\lambda_i}{\sum_{j=1}^{I}\lambda_j}$ $(w_i = \frac{1}{I})$*.*

Note that all these make sense. Since we relate cost models to a service model with an aggregate fill rate, the relation between the models is expressed in terms of weights chosen. In the most general form (case (i) in the corollary) of the systems with backorder costs being incurred in accordance with α-service (β-service) level, the weights of the individual α-service levels (fill rates) of different item types should be proportional to backorder costs (weighted by demand rates) and $1 - a_i(\pi)$ $(1 - \frac{a_i(\pi)}{\lambda_i})$ represents $FR_i(\pi)$. If b_i is the same for all i (case (ii) in the corollary), then we do not differentiate item types with respect to b_i to calculate aggregate service level in the service model. In this case, $FR(\pi)$ turns out to be $\frac{1}{I}\sum_{i=1}^{I} FR_i(\pi)$ (regular average of the fraction of time with backorders) when the backorder penalty is of type α and $\frac{\sum_{i=1}^{I}\lambda_i FR_i(\pi)}{\sum_{i=1}^{I}\lambda_i}$ (fraction of demand satisfied upon arrival) when the backorder penalty is of type β. Then, Theorem 1 given in [18] can be revised as follows for multi-class systems when aggregate fill rate is considered as the service level in the service model. Proof of Theorem 2 follows as the one in [18].

Theorem 2. *If*

i. $\frac{w_i}{w_j} = \frac{b_i\lambda_i}{b_j\lambda_j}$ $(\frac{w_i}{w_j} = \frac{b_i}{b_j})$ *for all* i*,* j*,*

ii. π^* *is an optimal policy for the cost model* P_6*,*

iii. $FR(\pi^*) = \beta_0$*,*

then π^* *is also an optimal policy for the service model* $P_5(\beta_0)$*.*

van Houtum and Zijm [18] state that, under Condition 1 for single-class systems (or when there is a service level constraint for each class in the service model unlike our case), the cost model is a kind of Lagrangean relaxation of the service model and it is a straightforward Lagrangean relaxation of the service model if $a_{pen}(\pi)$ is $\theta_0 - \theta(\pi)$. To clarify this statement for the cost and service models we study, let $P_5^L(\beta)$ be the Lagrangean relaxation of $P_5(\beta)$, i.e.,

$$P_5^L(\beta,\psi): \min_{\pi\in\Pi}\{\sum_{i=1}^{I} h_i Inv_i(\pi) + \psi(\beta - FR(\pi))\},$$

for any $\psi \geq 0$. It is possible to rewrite $(\beta - FR(\pi))$ as $((1 - FR(\pi)) - (1-\beta))$ and $(1 - FR(\pi))$ as $\sum_{i=1}^{I} w_i - \sum_{i=1}^{I} w_i FR_i(\pi)$. Then, $P_5^L(\beta,\psi)$ becomes

$$P_5^L(\beta,\psi): \min_{\pi\in\Pi}\{\sum_{i=1}^{I} h_i Inv_i(\pi) + \sum_{i=1}^{I} \psi w_i(1 - FR_i(\pi)) - \psi(1-\beta)\}$$

for any $\psi \geq 0$.

For the case a penalty is paid for each backordered request (for the fraction of time with backorders), immediate observations regarding the Lagrangean relaxation of $P_5(\beta)$ are as follows.

- The optimal policies for $P_5^L(\beta,\psi)$ and $P_6(\frac{\psi w_1}{\lambda_1}, ..., \frac{\psi w_I}{\lambda_I})$ ($P_6(\psi w_1, ..., \psi w_I)$) are the same for given β (irrespective of the value of β). Note that given any ψ, $\psi(1-\beta)$ is constant.

- Problem $P_5^L(\beta,\psi)$ provides a lower bound, called the Lagrangean lower bound, on the optimal objecive value of the original problem $P_5(\beta)$ for any $\psi \geq 0$. We are seeking for the values of multiplier ψ that give the maximum lower bound which is closest to the optimal solution of the original problem $P_5(\beta)$. The problem of finding the maximizing ψ, i.e., $\max_{\psi\geq 0}\{P_5^L(\beta,\psi)\}$, is called the Lagrangean dual problem to be denoted by $P_5^{LD}(\beta)$. Ideally, the optimal objective function values of problem $P_5(\beta)$ and the corresponding Lagrangean dual problem are the same. A duality gap is said to exist if these two values are different (see [2] for an overview of Lagrangean relaxation).

- Let ψ^* and ϑ^* be the optimal Lagrange multiplier and the optimal policy for $P_5^{LD}(\beta)$, respectively. Then, ϑ^* is also the optimal policy for $P_5(\beta)$ and $P_6(\frac{\psi^* w_1}{\lambda_1}, ..., \frac{\psi^* w_I}{\lambda_I})$ ($P_6(\psi^* w_1, ..., \psi^* w_I)$ if a duality gap does not exist.

- Let π^* be the optimal policy for $P_6(\mathbf{b})$ and $FR(\pi^*) = \beta_0$, then π^* is also an optimal policy for $P_5(\beta_0)$ according to Theorem 2.

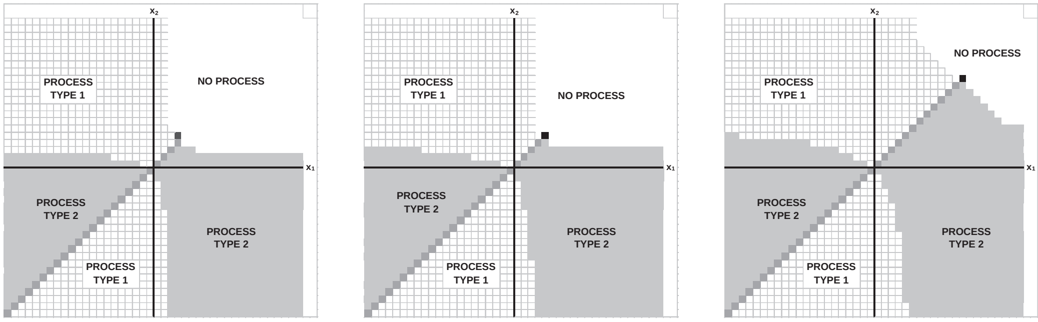
- For $P_5(\beta)$ and $P_6(\mathbf{b})$ to be related, we should have $\frac{\psi w_i}{\lambda_i} = b_i$ ($\psi w_i = b_i$), i.e., $\psi = \frac{b_i \lambda_i}{w_i}$ ($\psi = \frac{b_i}{w_i}$), for all i. Recall Proposition 4.

Note that π is not limited to dynamic scheduling policies under base-stock control, neither in problem $P_5(\beta)$ nor in $P_6(\mathbf{b})$. Optimal policy for problem $P_6(\mathbf{b})$ can be determined by value-iteration algorithm by adding a "no process" option for each state. For the two-class systems, state description n_i, being the number of items of class i in queue or in service as in chapters 3 and 4 is convenient when the system is analyzed under base-stock control. However, for problem $P_6(\mathbf{b})$, another state description is required: x_i denotes the inventory level of class i, where a negative value of x_i shows backorder(s) for that class. Let $\mathbf{x} = (x_1, x_2)$. Recursive (multi-period) formulation below is to determine the optimal dynamic scheduling policy for problem $P_6(\mathbf{b})$.

$$\begin{aligned} g_m(x_1, x_2) &= c(x_1, x_2) + \frac{\lambda_1}{\tau} g_{m-1}(x_1 - 1, x_2) + \frac{\lambda_2}{\tau} g_{m-1}(x_1, x_2 - 1) \\ &\quad + \frac{\mu}{\tau} min\{g_{m-1}(x_1 + 1, x_2), g_{m-1}(x_1, x_2 + 1), g_{m-1}(x_1, x_2)\}, \quad (5.1) \\ g_0(x_1, x_2) &= 0, \quad (5.2) \end{aligned}$$

for all (x_1, x_2) where $\overline{c}(\mathbf{x}) = h_1 x_1^+ + h_2 x_2^+ + b_1 \lambda_1 1_{\{x_1 \leq 0\}} + b_2 \lambda_2 1_{\{x_2 \leq 0\}}$ if $a_i(\pi) = \lambda_i(1 - FR_i(\pi))$ and $\overline{c}(\mathbf{x}) = h_1 x_1^+ + h_2 x_2^+ + b_1 1_{\{x_1 \leq 0\}} + b_2 1_{\{x_2 \leq 0\}}$ if $a_i(\pi) = 1 - FR_i(\pi)$, and $x_i^+ = max\{x_i, 0\}$ and $\tau = \lambda_1 + \lambda_2 + \mu = 1$. All the numerical experiments in the remaining part of this chapter are for the case $a_i(\pi) = \lambda_i(1 - FR_i(\pi))$, i.e., backorder penalties charged are of type β. Figure 5.2 shows the optimal policies for a symmetric two-class system with three different values of ρ. As seen, the optimal policy is of base-stock type with two switching curves. Note that under the optimal policy, once the inventory level of class i drops below its base-stock level then it can never exceed its base-stock level. So, the states outside the region $\{(x_1, x_2) : x_1 \leq S_1, x_2 \leq S_2\}$ are transient. See Appendix L for some other examples.

In Section 5.2, problem $P_5(\beta)$ is considered over the set of base-stock controlled policies, recall problem $P_3(\beta)$ with $r(\pi_{\mathbf{S}}) = \sum_{i=1}^{I} h_i Inv_i(\pi_{\mathbf{S}})$, and performances of some alternative scheduling policies are compared for symmetric two-class systems in Table 5.1. In other words, given one of the scheduling policies, "best" base-stock levels, $\mathbf{S}$, are determined to solve problem P_3: Recall that optimality of $\mathbf{S}$ cannot be guaranteed for each scheduling policy but the arguments regarding being close to the optimal $\mathbf{S}$ are supported with numerical observations. It is observed that in general ς policy (which is approximated by heuristic 2) works better in solving P_3 than LQ and FCFS policies. Observing these numerically, the problem we note in Section 5.2 is that base-stock controlled LQ policy is known to be optimal for problem P_4 if the system is symmetric, but in fact problems P_3 and P_4 are not related in terms of the



(a) $\rho = 0.40$.　　(b) $\rho = 0.60$.　　(c) $\rho = 0.90$.

Figure 5.2: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 100$, $h_i = 1$ for $i = 1, 2$.

way backorder penalties are paid. However, using the relation between P_6 and P_5, it is now possible to compare $\varsigma_{\mathbf{S}}$ with the optimal policy of P_6 (optimal policy of P_6 will be optimal also for P_5 under the conditions of Theorem 2 if there is not any duality gap). Note that this comparison is only for ς because we have already observed that ς is almost always better than LQ and FCFS policies for the symmetric systems under consideration.

Let π^* be the optimal policy for $P_6(\mathbf{b})$ and $FR(\pi^*) = \beta_0$. Under the conditions of Theorem 2, π^* is optimal also for $P_5(\beta_0)$. Then, it is possible to compare π^* and $\varsigma_{\mathbf{S}^*}$, where $\mathbf{S}^*$ denotes the "optimal" base-stock levels that solves $P_5(\beta_0)$ under policy ς, but not only $P_2(\beta_0)$. (Note once more that optimality of $\mathbf{S}^*$ here is not in strict analytical sense, but referring to Propositions 2 and 3 based on some numerical experiments for ς policy in case the system is symmetric and has two classes.) Figures 5.3 and 5.5 show the optimal policies of $P_6(\mathbf{b})$ for the symmetric two-class systems with $\lambda_i = \lambda$, $b_i = b$, $h_i = h$, $w_i = \frac{1}{2}$ for all i. Note that $\mathbf{n}$ is used as the state description instead of $\mathbf{x}$ once the base-stock levels are determined for $P_6(\mathbf{b})$. Switching from $\mathbf{x}$ to $\mathbf{n}$ according to $x_i = S_i - n_i$ for all i (e.g., from Figure 5.2 to 5.3) makes the comparison of π^* and $\varsigma_{\mathbf{S}^*}$ easier. Figures 5.4 and 5.6 show the "optimal" policies for $P_5(\beta_0)$ under ς, where target fill rates (β_0 values), are set as the aggregate fill rates in Figures 5.3 and 5.5, respectively and choosing $w_i = \frac{1}{2}$, $i = 1, 2$. These figures show that, base-stock levels are the same for both policies and the policy structures are very close to each other, which explains the success of using ς in solving $P_5(\beta)$ or the equivalent cost model $P_6(\mathbf{b})$. The equivalence relation needs to be expressed in terms of β

and b. Some other numerical comparisons can be seen in Table 5.2 and in Appendix M.

Table 5.3 to 5.5 and N.1, and Figures 5.7 to 5.9 and the ones in Appendix O are for the asymmetric systems. Since we cannot refer to Propositions 2 and 3 for finding the "best" $\mathbf{S}$ and claiming optimality of it for problem $P_5(\beta_0)$ under policy ς when the system is asymmetric, we proceed as follows: we solve problem $P_6(\mathbf{b})$ for π^*, and let $\beta_0 = FR(\pi^*)$ and $\mathbf{S}_0$ be the optimal $\mathbf{S}$. Then, we choose w_i values as in Proposition 4 to guarantee equivalence of $P_6(\mathbf{b})$ and $P_5(\beta_0)$, and observe that inequality $FR(\varsigma_{\mathbf{S}_0}) \geq \beta_0$ is satisfied for all the numerical experiments. We search around $\mathbf{S}_0$ for better base-stock levels in $P_5(\beta_0)$ under ς, and we come across such an instance just once (in Table N.1 $\mathbf{S} = (6, 2)$ is optimal for $P_6(100, 50)$ when $\frac{\lambda_1}{\lambda_2} = 2$ whereas $\varsigma_{(5,3)}$ satisfies target service level β_0 with lower objective function value). As in the case of symmetric systems, the "best" $\mathbf{S}$ that we can numerically identify for $P_5(\beta_0)$ under policy ς is denoted by $\mathbf{S}^*$. Except for the instance noted above, $\mathbf{S}^* = \mathbf{S}_0$ in all the figures and tables. Based on all the numerical experiments, the following points are noteworthy.

- In symmetric systems, both π^* and $\varsigma_{\mathbf{S}^*}$ are LQ policy in region 1 and SQ policy in region 3.

- $\varsigma_{\mathbf{S}^*}$ approximates π^* better for high values of $\mathbf{b}$. The reason is that as b_i values increase in problem $P_6(b)$, the difference between the costs to be incurred in states in region 1 and regions 2, 3 and 4 becomes more apparent as in the case of fill rate maximization (recall that $c(\mathbf{n})$ is 0 in region 1 and w_1 (w_2) in region 2 (4) and 1 in region 3). Note that as b_i values increase, the trade-off between inventory holding cost and backorder cost is resolved in $P_6(\mathbf{b})$ avoiding backorders more, i.e., backorder cost becomes dominant in resolving the trade-off. Then, in the service model P_5, considering the ς policy (which is obtained maximizing the fill rate but not working with inventory holding cost) makes sense and it is not surprising that the optimal policy is similar to ς.

- $\varsigma_{\mathbf{S}^*}$ is better in approximating π^* for low values of ρ. The reason for this is obvious in the symmetric systems: as ρ increases, the steady-state probabilities of being in regions 2 and 4, where these policies are not exactly the same, increase.

- We have investigated structure of the optimal policy for $P_5(\beta)$ using the related model $P_6(\mathbf{b})$. But, since the base-stock levels are discrete, any scheduling policy, e.g., LQ, FCFS, ς or some other policy, can be the optimal solution of $P_5(\beta)$ for some β values. Note that LQ is the best for P_3 for low values of ρ among the policies considered in Table 5.1. If integrality constraint for base-stock levels were relaxed, then for any β it would always be possible to find some b such that π^* is the optimal policy for $P_6(\mathbf{b})$ and $FR(\pi^*) = \beta$, which implies that π^* is an optimal solution also for $P_5(\beta)$ due to Theorem 2.

We can claim good performance of the ς policies for problem P_6 (or for P_5 under the equivalence conditions in Theorem 2) when b_i values are relatively higher than the h_i values (when β is high). Otherwise, optimal base-stock levels for problem P_6 would be very small, even equal to zero, causing fill rate to be zero and turning the system into a make-to-order system. A related reference is [20]. The following proposition is to give a sufficient condition for the optimality of zero base-stock levels.

Proposition 5. *If $b_i\lambda_i < h_i$, then the optimal base-stock level for item type i is zero for problem P_6. If $b_i\lambda_i = h_i$, base-stock levels of zero and one are alternative optima.*

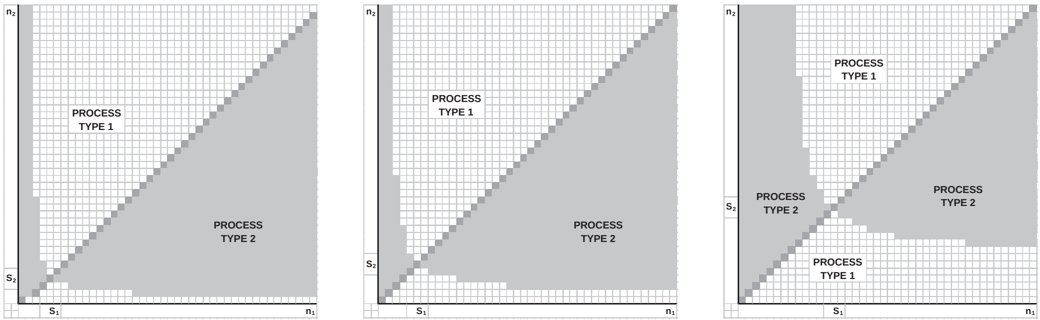
Proof. The result is intuitive. Since $\overline{c}(x_1, ..., x_i', ..., x_I) < \overline{c}(x_1, ..., x_i'', ..., x_I)$ for all $x_i' \leq 0$ and $x_i'' > 0$ when $b_i\lambda_i < h_i$, then it is not possible to improve long-run average cost by visiting states with $x_i > 0$. It is straightforward to extend the result for $b_i\lambda_i = h_i$. $\square$

Examples for Proposition 5 can be seen in Table 5.5, Figures 5.7 to 5.9 and in Appendix O. Immediate results and observations are as follows.

- If base-stock level of one of the item types is zero, then the optimal policy, π^*, is characterized by a static priority rule to choose the item type with zero base-stock level only if there is no outstanding order for the other item types. The result is intuitive because there is no way to eliminate stockout, i.e., fill rate is zero, for the item type with zero base-stock level, then the main concern is to minimize the expected average cost for the other item types. See Figure 5.8 for an example.

- If base-stock levels of both of the item types are zero, then processing any of the item types is optimal when there are outstanding orders for both of the items. Examples can be seen in Figures 5.9 and O.3.

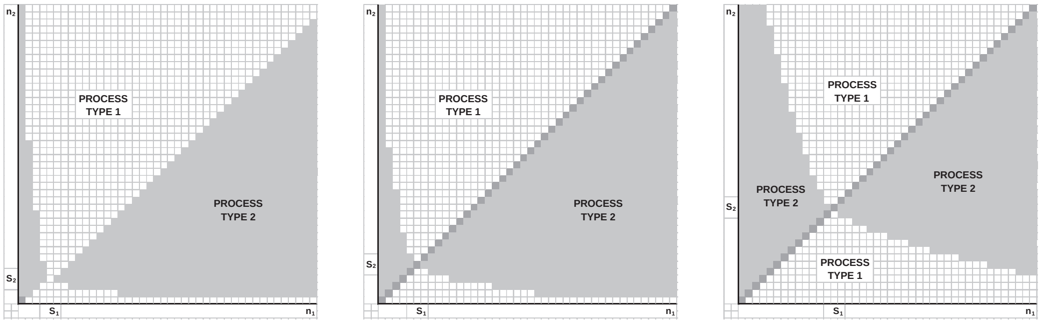
ς policy is optimal for problem P_2. On the other hand, optimal policy π^* for P_6 or its equivalent $P_5(\beta_0)$ with $\beta_0 = FR(\pi^*)$ is very similar to ς policy when β_0 is large. That is why we should note that when β is large, "best" $\mathbf{S}$ vectors for problems P_2 and P_5 are almost always the same, meaning that the objective functions in these problems



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure 5.3: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 100$, $h_i = 1$ for $i = 1,\ 2$.

are almost equally good in resolving the trade-off between inventories and aggregate fill rate. h_i is mostly approximated by c_i multiplied by the interest rate; thus, these coefficients appearing in the objective functions of P_2 and P_5 are proportional, which to a certain extent explains finding the same "best" $\mathbf{S}$. This argument is supported by checking possible base-stock level combinations numerically when $h_i = h$, but its validity cannot be fully guaranteed for the cases with unequal h_i values due to large number of possible base-stock level combinations to be checked numerically. Beyond this relationship between h_i and c_i, working with S_i instead of Inv_i we can apparently resolve the trade-off between inventories and aggregate fill rate with almost the same base-stock levels for all the numerical experiments we consider as long as β is large enough.

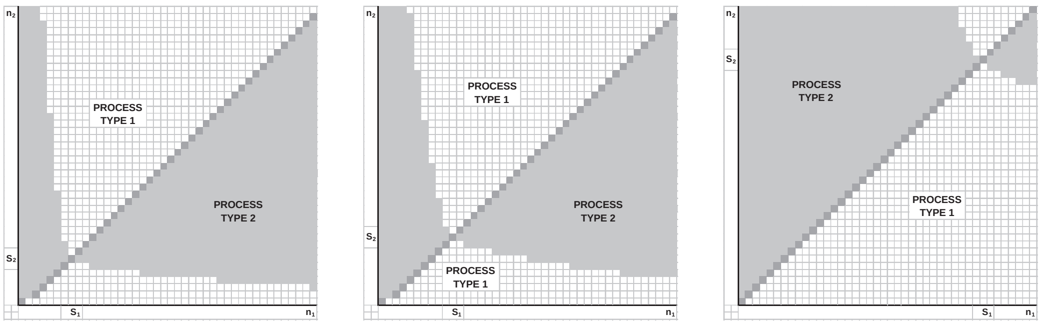


(a) $\rho = 0.40$.

(b) $\rho = 0.60$.

(c) $\rho = 0.90$.

Figure 5.4: Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$, where $\beta_0 = FR(\pi^*)$ in Figure 5.3.

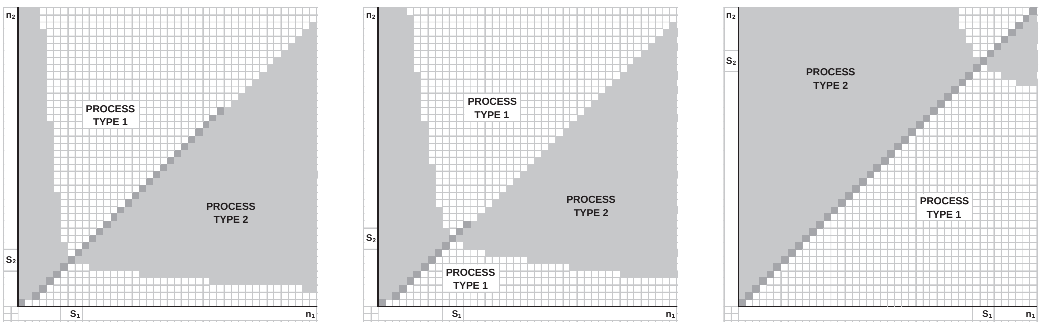


(a) $\rho = 0.40$.

(b) $\rho = 0.60$.

(c) $\rho = 0.90$.

Figure 5.5: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 10000$, $h_i = 1$ for $i = 1,\ 2$.



(a) $\rho = 0.40$.

(b) $\rho = 0.60$.

(c) $\rho = 0.90$.

Figure 5.6: Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$, where $\beta_0 = FR(\pi^*)$ in Figure 5.5.

Table 5.2: Comparison of the policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\lambda_i = \lambda$, $b_i = b$, $h_i = 1$, $w_i = \frac{1}{2}$ for $i = 1$, 2.

			FR			Inv		
$b\lambda$	ρ	$\mathbf{S}$	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)
1	0.40	(0,0)	0	0	0.000	0	0	0.000
	0.50	(0,0)	0	0	0.000	0	0	0.000
	0.60	(0,0)	0	0	0.000	0	0	0.000
	0.70	(0,0)	0	0	0.000	0	0	0.000
	0.80	(0,0)	0	0	0.000	0	0	0.000
	0.90	(0,0)	0	0	0.000	0	0	0.000
5	0.40	(2,1)	85.869	85.881	0.014	2.43293	2.43892	0.246
	0.50	(2,2)	90.070	90.139	0.077	3.11042	3.12350	0.421
	0.60	(2,2)	83.914	84.208	0.350	2.79096	2.84715	2.013
	0.70	(2,2)	75.808	76.257	0.593	2.44430	2.51329	2.822
	0.80	(3,2)	71.691	72.463	1.077	2.67845	2.87496	7.337
	0.90	(3,3)	61.330	63.862	4.127	2.40954	2.88831	19.869
10	0.40	(2,2)	94.349	94.355	0.006	3.37335	3.37653	0.094
	0.50	(2,2)	90.113	90.139	0.028	3.11676	3.12350	0.216
	0.60	(3,2)	89.346	89.391	0.050	3.70256	3.72628	0.641
	0.70	(3,3)	87.533	87.726	0.220	4.06954	4.13673	1.651
	0.80	(4,3)	81.853	82.282	0.524	4.12177	4.29827	4.282
	0.90	(5,4)	73.119	74.180	1.451	4.19346	4.73767	12.978
100	0.40	(4,3)	99.310	99.310	0.000	6.33771	6.33772	0.000
	0.50	(4,4)	99.113	99.113	0.000	7.00951	7.00955	0.001
	0.60	(5,5)	99.045	99.045	0.000	8.51862	8.51909	0.006
	0.70	(6,6)	98.429	98.429	0.001	9.72288	9.72566	0.029
	0.80	(9,8)	98.058	98.060	0.002	13.14259	13.15381	0.085
	0.90	(13,13)	95.698	95.712	0.015	17.81306	17.91816	0.590
1000	0.40	(5,5)	99.941	99.941	0.000	9.33369	9.33369	0.000
	0.50	(6,6)	99.928	99.928	0.000	11.00071	11.00071	0.000
	0.60	(8,7)	99.916	99.916	0.000	13.50153	13.50156	0.000
	0.70	(10,9)	99.866	99.866	0.000	16.67157	16.67157	0.000
	0.80	(14,13)	99.791	99.791	0.000	23.01640	23.01654	0.001
	0.90	(24,23)	99.531	99.531	0.000	38.09554	38.10130	0.015
10000	0.40	(7,6)	99.994	99.994	0.000	12.33337	12.33337	0.000
	0.50	(8,8)	99.994	99.994	0.000	15.00005	15.00005	0.000
	0.60	(10,9)	99.989	99.989	0.000	17.50020	17.50021	0.000
	0.70	(13,12)	99.984	99.984	0.000	22.66724	22.66724	0.000
	0.80	(19,18)	99.978	99.98	0.000	33.00177	33.00178	0.000
	0.90	(34,34)	99.949	99.949	0.000	59.01083	59.01105	0.000

Table 5.3: Comparison of the policies π^* and $\varsigma_{\mathbf{S}^*}$ for $b_1 = 10000$, $h_i = 1$ for $i = 1,\ 2$.

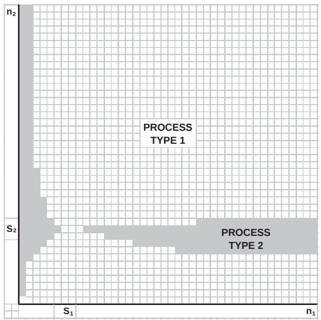
				FR			$\sum_i h_i Inv_i$		
b_2	$\frac{\lambda_1}{\lambda_2}$	$\frac{w_1}{w_2}$	ρ	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)
20000	2	1	0.40	99.971	99.971	0.000	10.33353316	10.33353316	0.000
			0.60	99.951	99.951	0.000	14.50091021	14.50091032	0.000
			0.80	99.918	99.918	0.000	27.00634857	27.0063621	0.000
20000	4	2	0.40	99.942	99.942	0.000	9.333675927	9.333675927	0.000
			0.60	99.945	99.945	0.000	14.50091153	14.50091868	0.000
			0.80	99.908	99.908	0.000	27.00658284	27.00660517	0.000
5000	2	4	0.40	99.946	99.946	0.000	9.333692917	9.333692917	0.000
			0.60	99.928	99.928	0.000	13.50191942	13.5019511	0.000
			0.80	99.881	99.881	0.000	24.01883222	24.01909785	0.001
2500	3	12	0.40	99.959	99.959	0.000	9.333783914	9.333783915	0.000
			0.60	99.940	99.940	0.000	13.50267803	13.50268445	0.000
			0.80	99.866	99.866	0.000	22.0439803	22.04536569	0.006

Table 5.4: Comparison of the policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\rho = 0.70$, $b_i = 1000$, $\lambda_i = \lambda$ for $i = 1,\ 2$.

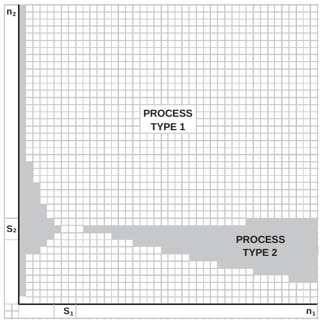
		FR			$\sum_i h_i Inv_i$		
h	$\mathbf{S}^*$	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)
(0.2,8)	(14,3)	97.692	97.692	0.000	22.25294474	22.25466379	0.008
(0.5,11)	(11,5)	99.418	99.422	0.005	13.40970335	13.43469052	0.186
(1,3)	(10,4)	98.666	98.667	0.000	18.51346979	18.5150628	0.009
(1,10)	(10,3)	97.144	97.148	0.004	32.74109323	32.76239365	0.065
(2,5)	(8,4)	98.031	98.032	0.001	29.4596704	29.46898136	0.032

Table 5.5: Comparison of the policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\rho = 0.80$.

					FR			$\sum_i h_i Inv_i$		
$\frac{\lambda_1}{\lambda_2}$	**b**	$(b_1\lambda_1,b_2\lambda_2)$	**h**	$\mathbf{S}^*$	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)	π^*	$\varsigma_{\mathbf{S}^*}$	Error(%)
2	(200,100)	(59.26,14.81)	(2,1)	(5,5)	92.326	92.473	0.159	10.41715	10.59913	1.747
	(20,30)	(5.93,4.44)	(8,4)	(0,1)	31.429	31.429	0.000	2.93333	2.93333	0.000
	(10,5)	(2.96,0.74)	(3,5)	(0,0)	0	0	0.000	0	0	0.000
1.5	(300,1000)	(80.00,177.78)	(3,1)	(5,10)	97.353	97.456	0.106	18.51970	18.87665	1.927
	(100,50)	(26.67,8.89)	(6,9)	(2,0)	57.720	57.720	0.000	7.73760	7.73760	0.000
	(15,45)	(4.00,8.00)	(4,8)	(0,0)	0	0	0.000	0	0	0.000
1.2	(1000,800)	(242.42,161.62)	(3,2)	(6,10)	97.558	97.564	0.006	29.29224	29.36301	0.242
	(40,10)	(9.70,2.02)	(1,3)	(3,0)	75.882	75.882	0.000	2.29013	2.29013	0.000
	(5,10)	(1.21,2.02)	(2,3)	(0,0)	0	0	0	0	0	0.000
3	(2000,400)	(666.67,44.44)	(8,3)	(8,4)	96.809	96.989	0.187	56.40573	58.32497	3.403
	(250,100)	(83.33,11.11)	(5,4)	(5,1)	84.492	85.767	1.510	17.66350	19.24278	8.941
	(20,40)	(6.67,4.44)	(7,5)	(0,0)	0	0	0.000	0	0	0.000

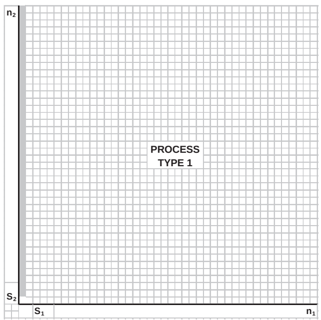


(a) Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$.

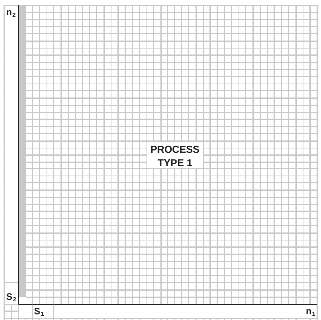


(b) Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$ where $\beta_0 = FR(\pi^*)$.

Figure 5.7: Comparison of the Policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\mathbf{b} = (1000, 800)$, $\mathbf{h} = (3, 2)$, $\rho = 0.80$ and $\lambda_1 = 1.2\lambda_2$.

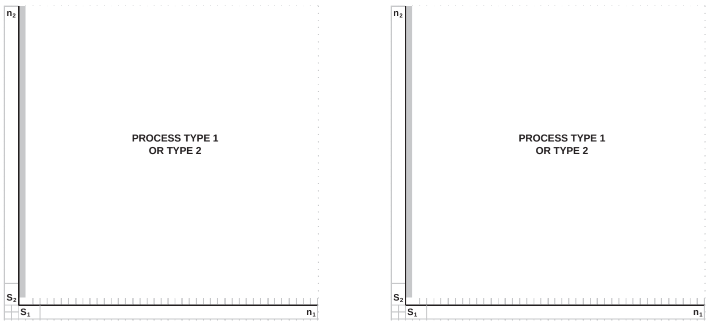


(a) Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$.



(b) Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$ where $\beta_0 = FR(\pi^*)$.

Figure 5.8: Comparison of the Policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\mathbf{b} = (40, 10)$, $\mathbf{h} = (1, 3)$, $\rho = 0.80$ and $\lambda_1 = 1.2\lambda_2$.



(a) Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$.

(b) Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$ where $\beta_0 = FR(\pi^*)$.

Figure 5.9: Comparison of the Policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\mathbf{b} = (5, 10)$, $\mathbf{h} = (2, 3)$, $\rho = 0.80$ and $\lambda_1 = 1.2\lambda_2$.

CHAPTER 6

CONCLUSION

This study can be considered among a few others in the literature on the service models for dynamic scheduling of multi-class make-to-stock systems. The main difference here is the use of aggregate fill rate instead of having a fill rate constraint for each item class as in [3] and [7]. Replacing a number of fill rate constraints with a single one is, in fact, a relaxation of the service models in [3] and [7]. But, this drawback can be handled choosing weights of the individual fill rates and the target level for the aggregate fill rate to relate the service model we work with to the corresponding cost model and, this way, to guarantee optimality of the solution of the cost model also for the service model with aggregate fill rate constraint under some equivalence conditions. Note that the same cost model is also related to the service model with individual fill rate constraints under some other conditions. Furthermore, having the total base-stock investment (a linear function of the base-stock levels) as the objective function of the service model under consideration, unlike the usual service models with exponential inventory holding cost being the objective function, allows us to restrict our attention to the set of base-stock controlled dynamic scheduling policies maximizing aggregate fill rate (ς policies). Due to optimality of the ς policy for this somewhat simpler service model as compared to the ones in [3] and [7], the minimum base-stock investment (lowest base-stock levels) required to satisfy any given target fill rate can be found using accurate heuristics we propose to approximate ς policies. Working with the lowest base-stock levels, then, implies a considerable advantage of decreasing also the average inventory holding cost over the other well-known policies FCFS, LQ and its variations in the symmetric and asymmetric cases, respectively, as numerically shown in this study.

Although our focus is primarily on the service models, further investigations are to relate the service model minimizing average inventory holding cost under a constraint on the aggregate fill rate to the corresponding cost model where backorder cost is incurred in accordance with the service measure considered in the service model. These investigations lead to the following: We numerically observe that the optimal policy of the cost model is almost the same as the ς policy with the minimum base-stock investment to achieve the target fill rate implied by the optimal solution of the cost model. At least to a certain extent, we also answer the questions that would be raised about how to determine the weights for fill rates of different classes. These, with some reservations due to the lack of the equivalence relation between the backorder penalties and the target level for aggregate fill rate, suggest the employment of ς policy instead of using value-iteration for the cost model or solving the lagrangean dual for the service model.

Any progress in the literature on the relation between the backorder penalties in the cost model and the target level of aggregate fill rate in the service model as for the equivalence of these two models would reveal the use of ς policies. For the time being, in the current status of the literature, practical use of our observations in this study is limited. We should also note the following: the difficulty regarding the progress mentioned above is that the optimal dynamic scheduling policy, which turns out to be of base-stock type for the cost and service models under consideration, are dependent on the base-stock levels (i.e., they are not $\mathbf{S}$-independent policies as base-stock controlled FCFS and LQ policies) as numerically shown in Chapter 5.

All these position this thesis somewhere among the studies on service models. However, it should be pointed out that the cost model related to the service models we study under aggregate fill rate constraint are different than the cost models in literature. γ type backorder cost are incurred in the latter unlike α or β type costs incurred in our case in accordance with the use of (aggregate) fill rate in the service model.

An immediate generalization of the heuristics proposed in Chapter 4 is for systems more than two classes. Since evaluating the performance of heuristics 1 and 2 using value-iteration is not computationally efficient even for three-class systems when the traffic intensity is high, we cannot avoid proceeding with simulation. Recall that heuristic 3, for which the steady-state probabilities can be calculated recursively, is only for the symmetric two-class systems. Not only to evaluate performance of heuristics 1 and 2, but even to employ them, in systems with more than two classes, there are implementation difficulties mentioned in Section 4.3 for these heuristics, unlike easily implementable state-dependent LQ policy and its variations (e.g., Δ policy). Then, coming up with a systematic way of handling this difficulty for any given number of classes and writing its code appears as a further work.

Another important issue is to determine the best base-stock levels while solving service models, under some specified policy. The set of base-stock levels searched is restricted in the symmetric case based on the analytical results presented for certain policies. These results can be supported only by some numerical observations under some other policies. Unfortunately, an extensive (maybe enumeration type of) search is still required for the asymmetric systems. So, devising or adopting search procedures towards that end would be useful.

An extension of this study can be considered relaxing the assumption of constant service rate but allowing this rate to be dependent on the item type being processed. [8] has some results on the partial characterization of the optimal policy in this case for the long-run discounted inventory holding and backorder costs. Replacing the single server facility with a network of servers, maybe with different routings for different item types, would make the problem more realistic and challenging.

Incorporating set-up times to switch from processing one item type to another can be considered as another extension. But, then, the use of base-stock policies cannot be justified, and the problem should be studied within the context of stochastic lot scheduling.

BIBLIOGRAPHY

[1] Axsäter, S. 2005. A simple procedure for determining order quantities under a fill rate constraint and normally distributed lead-time demand. *European Journal of Operational Research.* In Press.

[2] Beasley, J. 1993. Lagrangean relaxation. *Modern Heuristic Techniques for Combinatorial Problems, edited by C.R. Reeves.* Blackwell Scientific Publications, 243-303.

[3] Bertsimas, D., I. CH. Paschalidis. 2001. Probabilistic service level guarantees in make-to-stock manufacturing systems. *Oper. Res.* **49**, 119-133.

[4] Boyacı, G., G. Gallego. 2001. Serial production/distribution systems under service constraints. *Manufacturing and Service Operations Management.* **3**, 43-50.

[5] Buzacott, J.A., J.G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems.* Prentice Hall, Englewood Cliffs, NJ.

[6] de Vericourt, F., F. Karaesmen, Y. Dallery. 2000. Dynamic scheduling in a make-to-stock system: A partial characterization of optimal policies. *Oper. Res.* **48**, 811-819.

[7] Glasserman P. 1996. Allocating production capacity among multiple products. *Oper. Res.* **44**, 724-734.

[8] Ha, A. 1997. Optimal dynamic scheduling policy for a make-to-stock production system. *Oper. Res.* **45**, 42-53.

[9] Milito, R.A., H. Levy. 1989. Modeling and dynamic scheduling of a queueing system with blocking and starvation. *IEEE Transactions on Communications.* **37**, 1318-1329.

[10] Peña-Perez, A., P. Zipkin. 1997. Dynamic scheduling rules for a multiproduct make-to-stock queue. *Oper. Res.* **45**, 919-930.

[11] Rosling, K. 2002. Inventory cost rate functions with nonlinear shortage costs. *Operations Research.* **50**, 1007-1017.

[12] Shang, K.H., J.S. Song. 2004. Analysis of serial supply systems with a service constraint. Working Paper. The Fuqua School of Business, Duke University.

[13] Silver, E.A., D.F. Pyke, R. Peterson. 1998. *Inventory management and production planning and scheduling* Wiley, NY.

[14] Sleptchenko, A., A. van Harten, M. van der Heijden. 2002. Analyzing multi-class, multi-server queueing systems with preemptive priorities. Working paper. Faculty of Technology and Management, University of Twente.

[15] Tijms, H.C. 1994. *Stochastic Models; An Algorithmic Approach.* John Wiley & Sons, Chichester.

[16] van der Heijden, M., A. van Harten, A. Sleptchenko. 2004. Approximations for Markovian multi-class queues with preemptive priorities. *Operations Research Letters.* **32**, 273-282.

[17] van Houtum, G.J., I.J.B.F. Adan, J. van der Wal. 1997. The symmetric longest queue system. *Communications in Statistics: Stochastic Models.* **13**, 105-120.

[18] van Houtum, G.J., W.H.M. Zijm. 2000. On the relations between cost and service models for general inventory systems. *Statistica Neerlandica.* **54**, 127-147.

[19] Veatch, M., L. M. Wein. 1996. Scheduling a make-to-stock queue: Index policies and hedging points. *Oper. Res.* **44**, 634-647.

[20] Veatch, M., F. de Vericourt. 2003. Zero inventory policy for a two-part-type make-to-stock production system. *Queueing Systems.* **43**, 251-266.

[21] Wein, L. M. 1992. Dynamic scheduling of a multiclass make-to-stock queue. *Oper. Res.* **40**, 724-735.

[22] Zheng, Y., P. Zipkin 1990. A queueing model to analyze the value of centralized inventory information. *Oper. Res.* **38**, 296-307.

[23] Zipkin, P. 1995. Performance analysis of a multi-item production-inventory system under alternative policies. *Mgmt. Sci.* **41**, 690-703.

[24] Zipkin, P. 2000. *Foundations of inventory management.* McGraw-Hill, NY.

APPENDIX A

Lemma A1. *$f_m(n_1, n_2) = f_m(n_2, n_1)$ for all $\lambda_1 = \lambda_2$, $S_1 = S_2 = S$ and all m, n_1, n_2.*

Proof. Proof is by induction on m. For $m = 1$, proof is immediate because $c(n_1, n_2) = c(n_2, n_1)$ for all n_1, n_2. Next assuming that $f_{m-1}(n_1, n_2) = f_{m-1}(n_2, n_1)$ for all n_1, n_2, (2) can be rewritten as

$$\begin{aligned} f_m(n_1, n_2) &= c(n_2, n_1) + \frac{\lambda_1}{\tau} f_{m-1}(n_2 + 1, n_1) + \frac{\lambda_2}{\tau} f_{m-1}(n_2, n_1 + 1) \\ &\quad + \frac{\mu}{\tau} v_{m-1}(n_2, n_1) \end{aligned}$$

where the right hand side is $f_m(n_2, n_1)$ for all n_1 and n_2. □

Lemma A2. *$f_m(n_1, n_2)$ is nondecreasing in n_1 (n_2) for all m, n_2 (n_1).*

Proof. Proof is by induction on m. $f_1(n_1, n_2) = c(n_1, n_2)$ is nondecreasing in n_1 as seen in Figure A.1. Assume that $f_{m-1}(n_1, n_2)$ is nondecreasing in n_1. Then, the terms on the right hand side of 3.1 are all nondecreasing in n_1, and the result follows. With similar arguments it can be shown that $f(n_1, n_2)$ is nondecreasing in n_2 for all m and n_1.

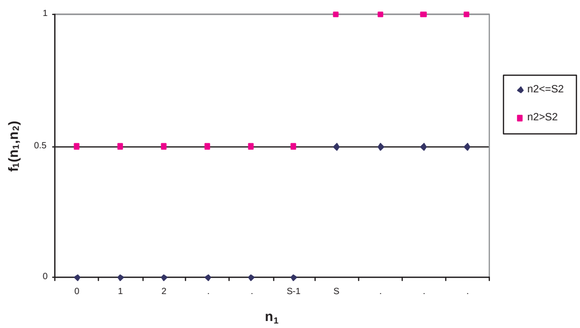


Figure A.1: $f_1(n_1, n_2)$ versus n_1.

□

Lemma A3. *$f_m(n_1, n_2)$ is nondecreasing in m for all n_1, n_2.*

Proof. Proof is by induction on m. Since $c_1(n_1, n_2) \geq 0$, $f_1(n_1, n_2) \geq f_0(n_1, n_2)$. Then, the assumption that $f_m(n_1, n_2) \geq f_{m-1}(n_1, n_2)$ leads to $f_{m+1}(n_1, n_2) \geq f_m(n_1, n_2)$ from (2) for all n_1 and n_2. $\square$

APPENDIX B

Step 1: Initialization.

For $n_1 := 0$ *to* $No_1 + m'$
For $n_2 := 0$ *to* $No_2 + m'$
$f_0(n_1, n_2) = 0$

$m := 1$, $Min := \infty$, $Max := 0$

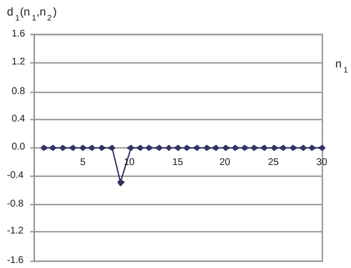
Step 2.

For $n_1 := 0$ *to* $(No_1 + m' - m)$
For $n_2 := 0$ *to* $(No_2 + m' - m)$
begin
$f_m(n_1, n_2) := c(n_1, n_2) + \lambda_1 f_{m-1}(n_1 + 1, n2) + \lambda_2 f_{m-1}(n_1, n_2 + 1) + \mu v_{m-1}(n_1, n_2)$
end

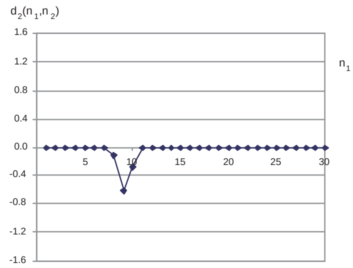
For $n_1 := 0$ *to* No_1
For $n_2 := 0$ *to* No_2
begin
if $(f_m(n_1, n_2) - f_{m-1}(n_1, n_2) > Max)$ *then*
$Max := f_m(n_1, n_2) - f_{m-1}(n_1, n_2)$
if $(f_m(n_1, n_2) - f_{m-1}(n_1, n_2) < Min)$ *then*
$Min := f_m(n_1, n_2) - f_{m-1}(n_1, n_2)$
end

Step 3. *If* $(m = m')$ *or* $(\frac{Max-Min}{Min} \leq \epsilon)$, *then STOP, else* $m := m + 1$ *goto Step 2.*

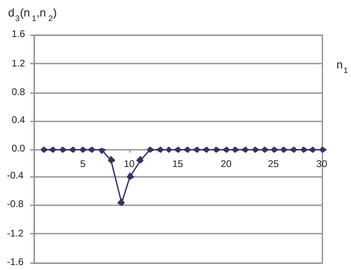
APPENDIX C



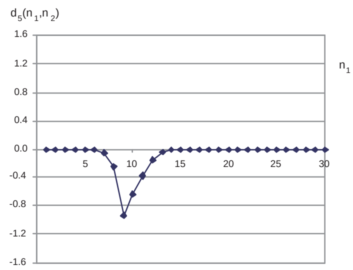
(a) $m = 1$.



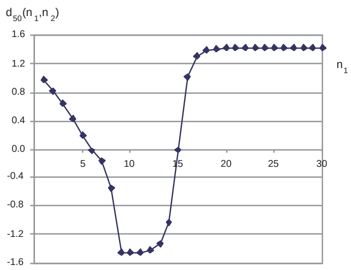
(b) $m = 2$.



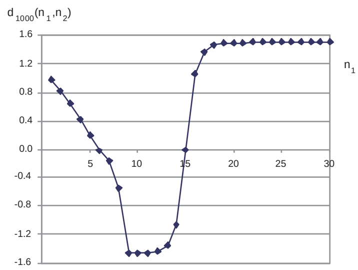
(c) $m = 3$.



(d) $m = 5$.

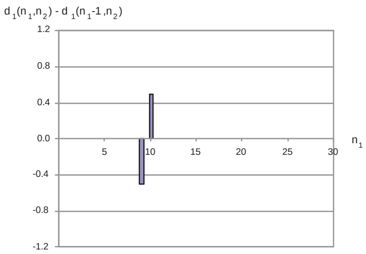


(e) $m = 50$.

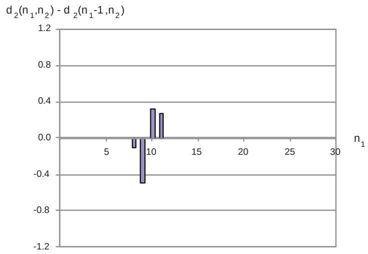


(f) $m = 1000$.

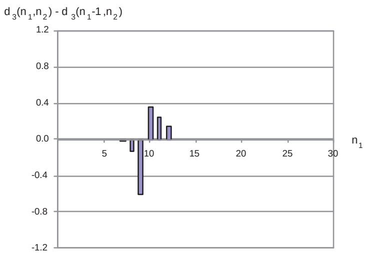
Figure C.1: $d_m(n_1, n_2)$ versus n_1 for $n_2 = 15$, $\rho = 0.80$, $S = 9$.



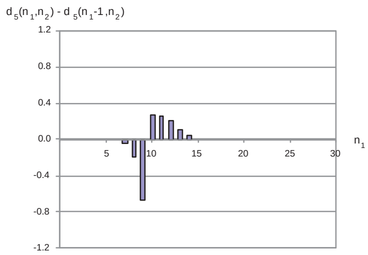
(a) $m = 1$.



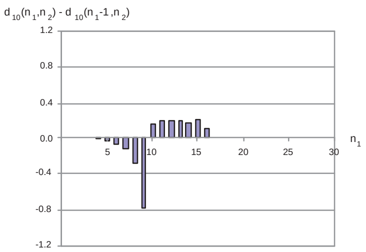
(b) $m = 2$.



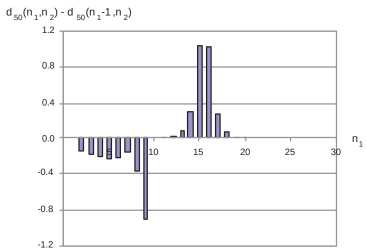
(c) $m = 3$.



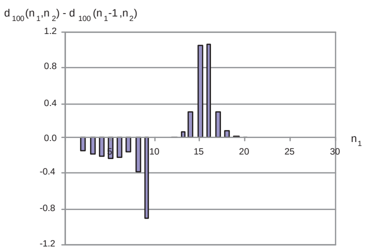
(d) $m = 5$.



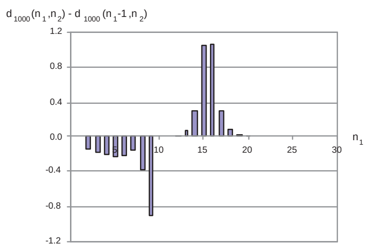
(e) $m = 10$.



(f) $m = 50$.

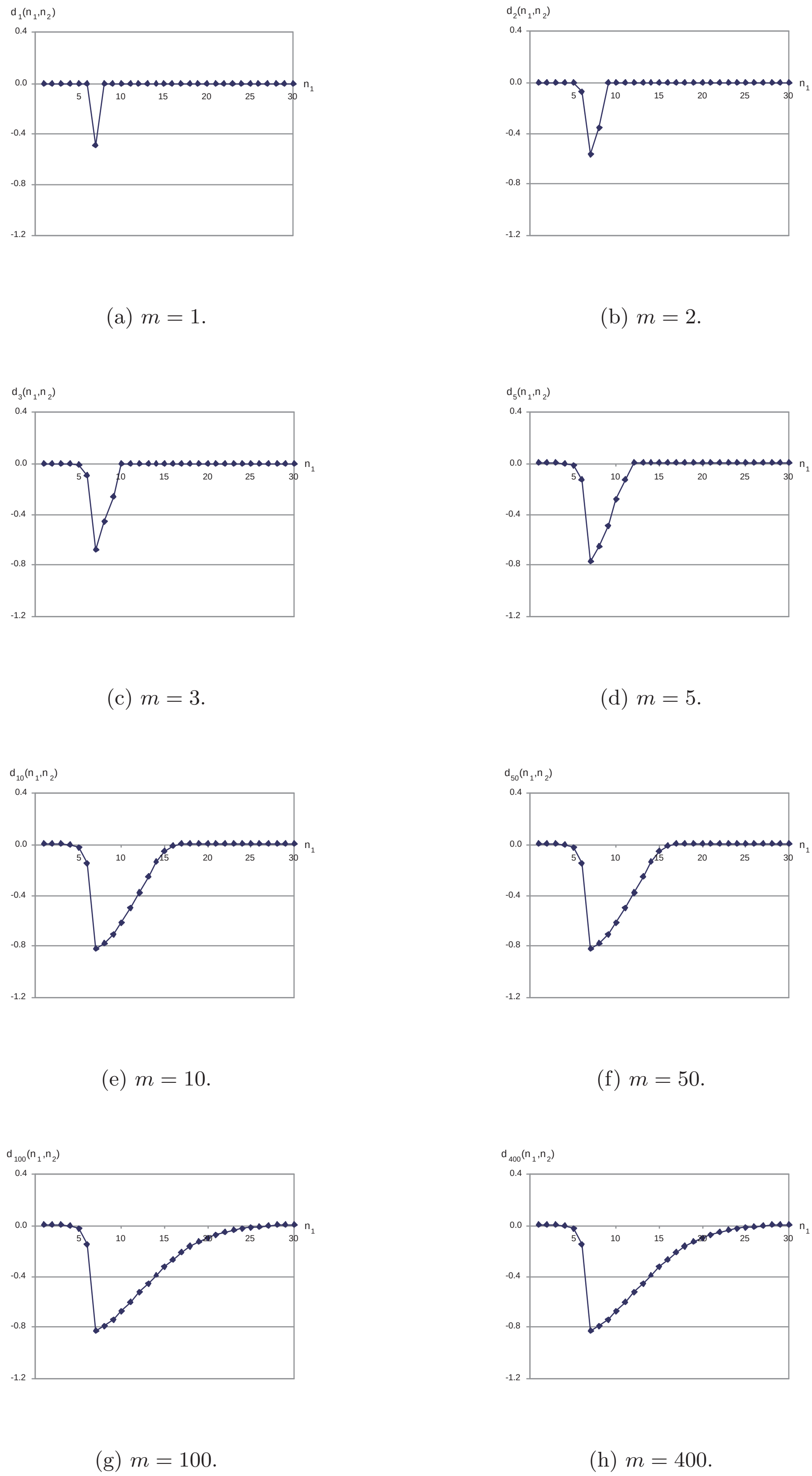


(g) $m = 100$.



(h) $m = 1000$.

Figure C.2: $d_m(n_1, n_2) - d_m(n_1 - 1, n_2)$ versus n_1 for $n_2 = 15$, $\rho = 0.80$, $S = 9$.



(a) $m = 1$.

(b) $m = 2$.

(c) $m = 3$.

(d) $m = 5$.

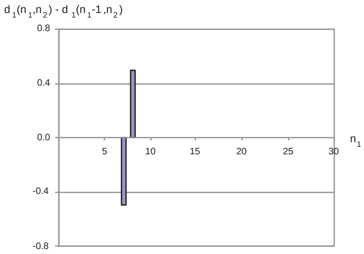
(e) $m = 10$.

(f) $m = 50$.

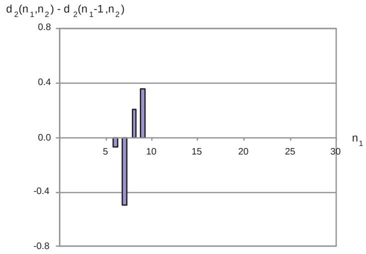
(g) $m = 100$.

(h) $m = 400$.

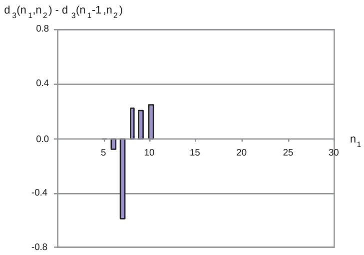
Figure C.3: $d_m(n_1, n_2)$ versus n_1 for $n_2 = 4$, $\rho = 0.40$, $S = 7$.



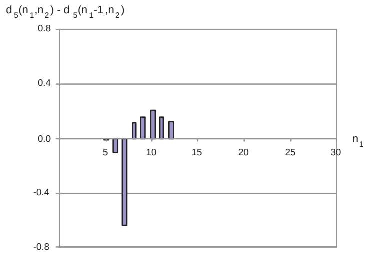
(a) $m = 1$.



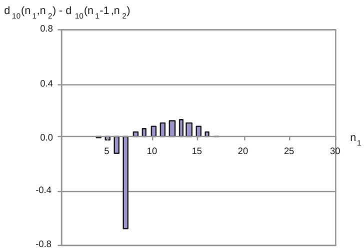
(b) $m = 2$.



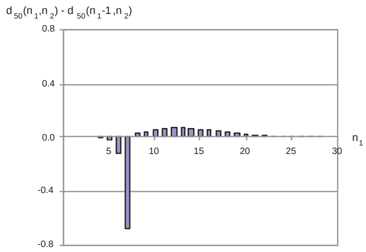
(c) $m = 3$.



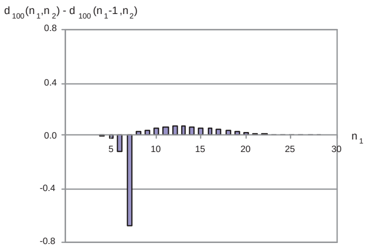
(d) $m = 5$.



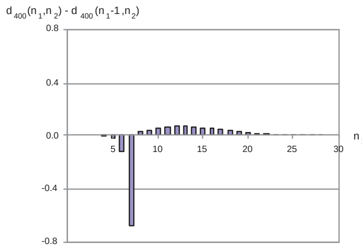
(e) $m = 10$.



(f) $m = 50$.

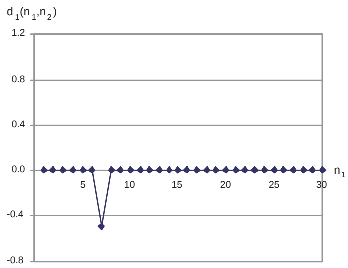


(g) $m = 100$.

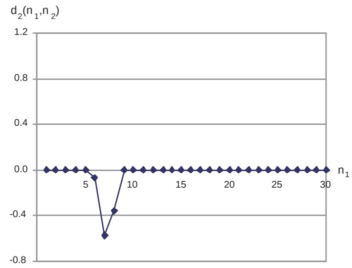


(h) $m = 400$.

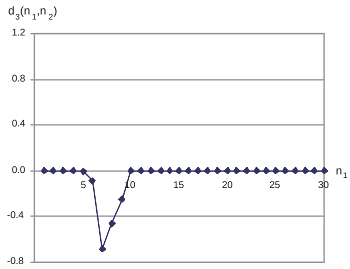
Figure C.4: $d_m(n_1, n_2) - d_m(n_1 - 1, n_2)$ versus n_1 for $n_2 = 4$, $\rho = 0.40$, $S = 7$.



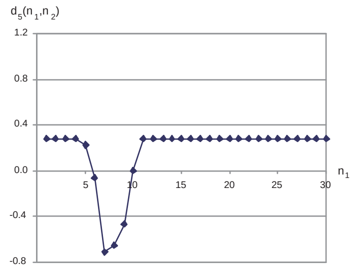
(a) $m = 1$.



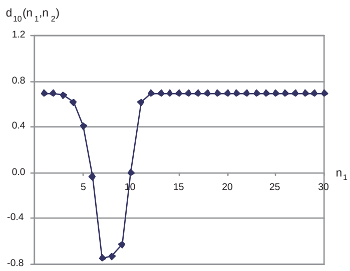
(b) $m = 2$.



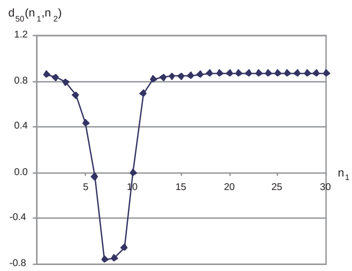
(c) $m = 3$.



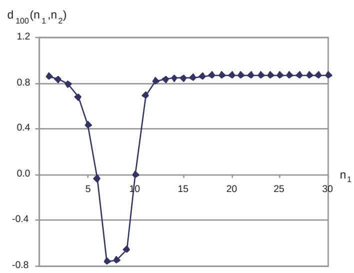
(d) $m = 5$.



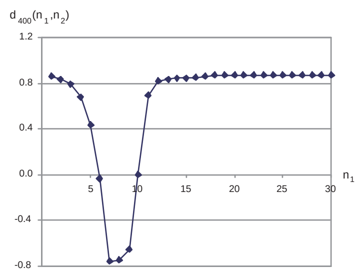
(e) $m = 10$.



(f) $m = 50$.

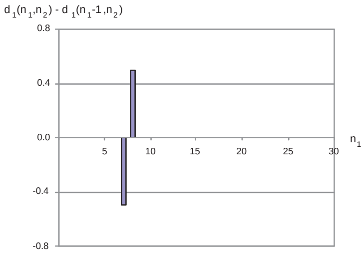


(g) $m = 100$.

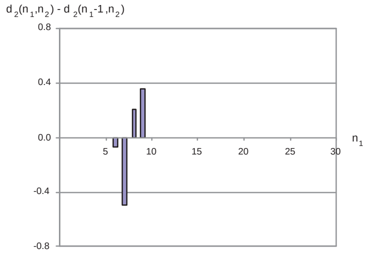


(h) $m = 400$.

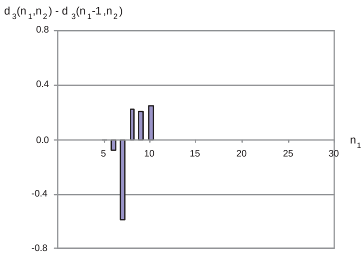
Figure C.5: $d_m(n_1, n_2)$ versus n_1 for $n_2 = 10$, $\rho = 0.40$, $S = 7$.



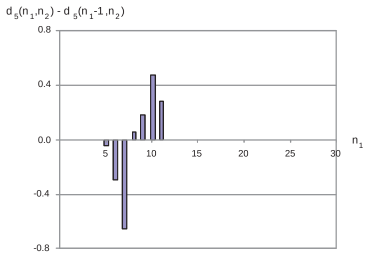
(a) $m = 1$.



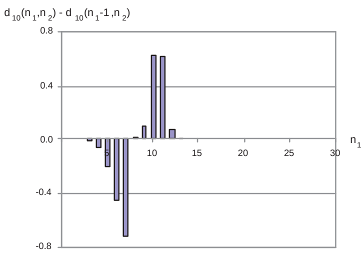
(b) $m = 2$.



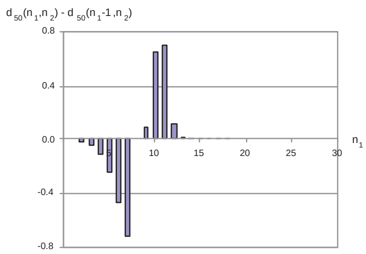
(c) $m = 3$.



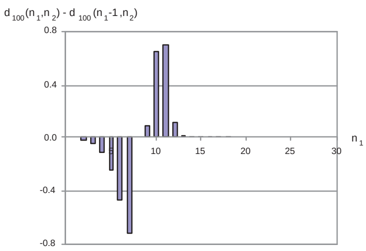
(d) $m = 5$.



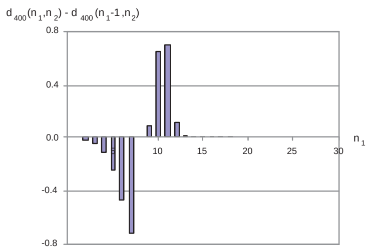
(e) $m = 10$.



(f) $m = 50$.



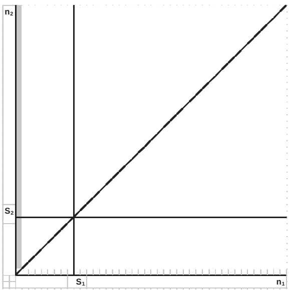
(g) $m = 100$.



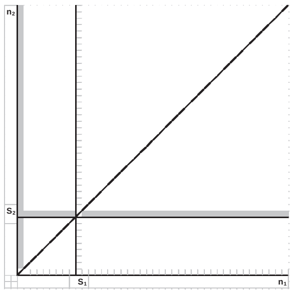
(h) $m = 400$.

Figure C.6: $d_m(n_1, n_2) - d_m(n_1 - 1, n_2)$ versus n_1 for $n_2 = 10$, $\rho = 0.40$, $S = 7$.

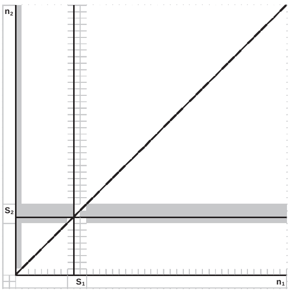
APPENDIX D



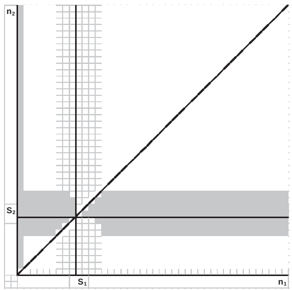
(a) $m = 1$.



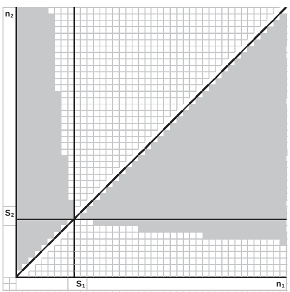
(b) $m = 2$.



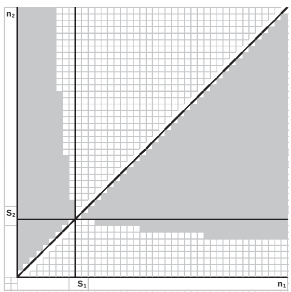
(c) $m = 3$.



(d) $m = 5$.

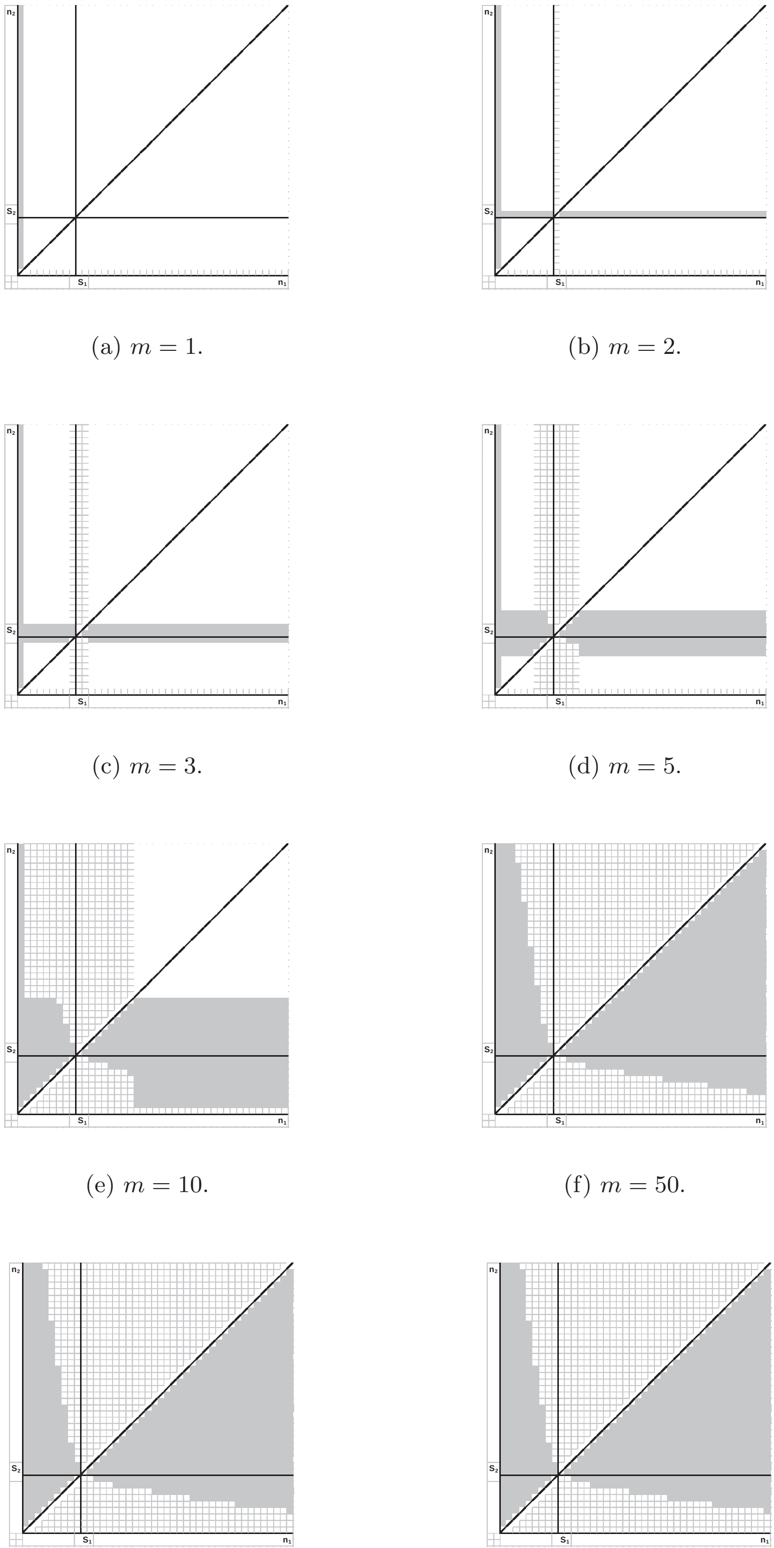


(e) $m = 50$.



(f) $m = 1000$.

Figure D.1: ς policy for $\rho = 0.40$, $S = 9$.



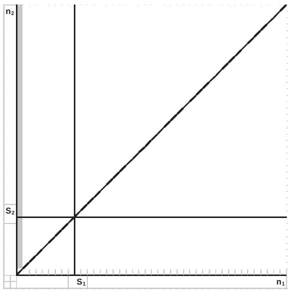
(a) $m = 1$. (b) $m = 2$.

(c) $m = 3$. (d) $m = 5$.

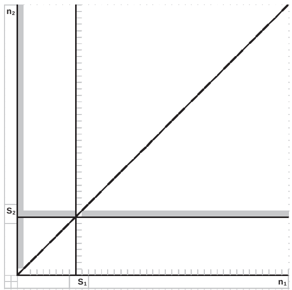
(e) $m = 10$. (f) $m = 50$.

(g) $m = 100$. (h) $m = 1000$.

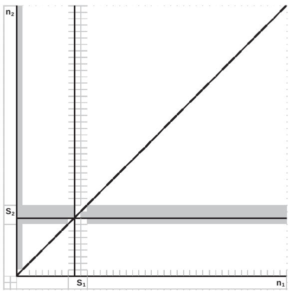
Figure D.2: ς policy for $\rho = 0.60$, $S = 9$.



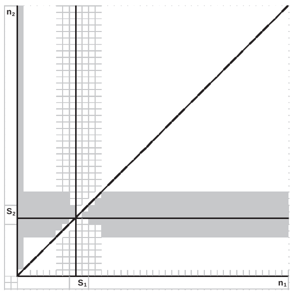
(a) $m = 1$.



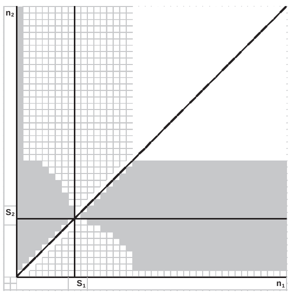
(b) $m = 2$.



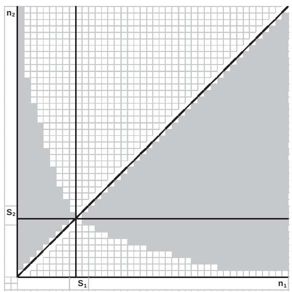
(c) $m = 3$.



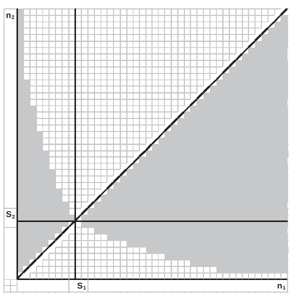
(d) $m = 5$.



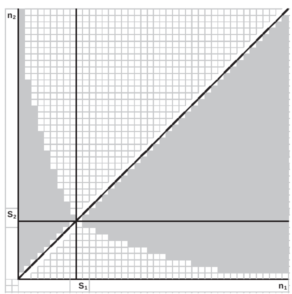
(e) $m = 10$.



(f) $m = 50$.



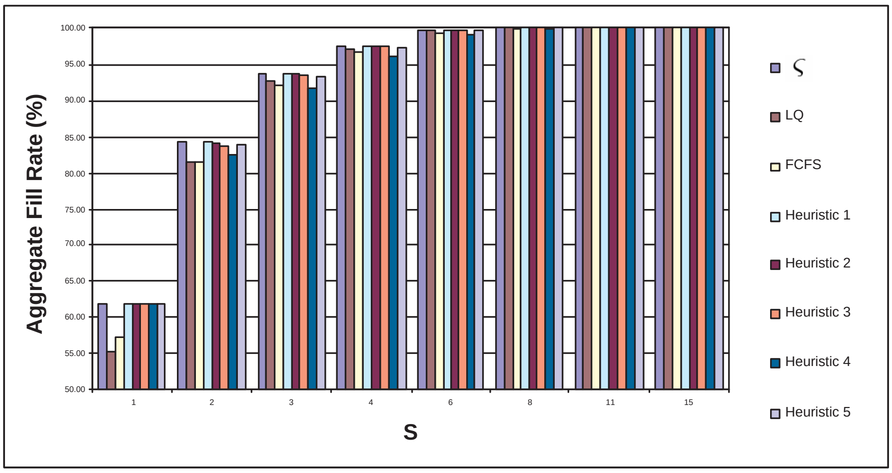
(g) $m = 100$.



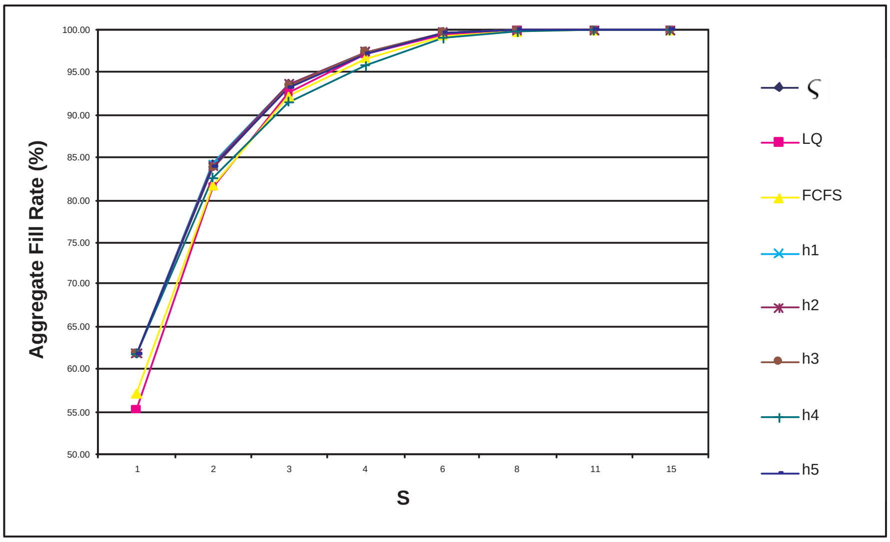
(h) $m = 1000$.

Figure D.3: ς policy for $\rho = 0.90$, $S = 9$.

APPENDIX E



(a) $\rho = 0.60$.



(b) $\rho = 0.60$.

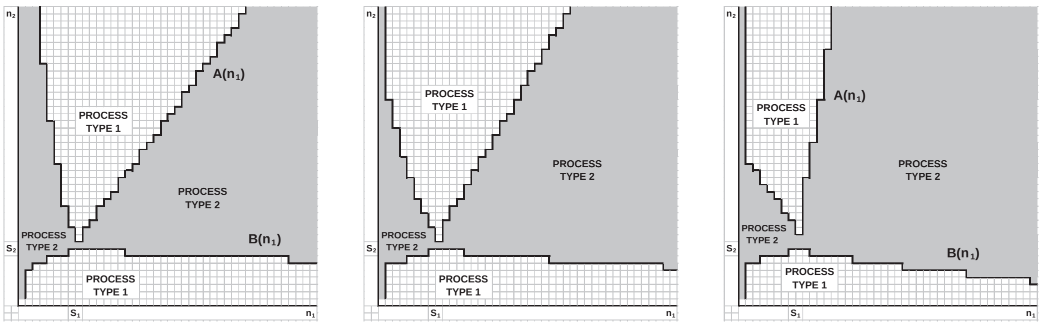
Figure E.1: Comparison of the ς, LQ, FCFS policies and the heuristics.

APPENDIX F

Table F.1: Comparison of the aggregate fill rates (%) of the ς, Δ policies and heuristic 2 (M): equally weighted case, $\lambda_1 = 1.2\lambda_2$.

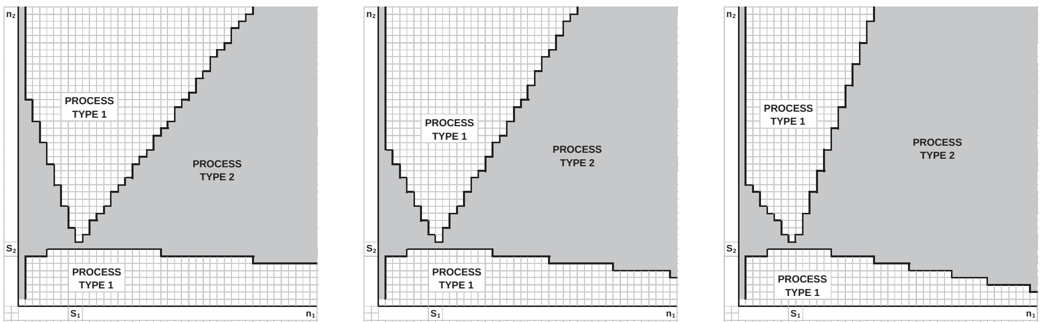
		$\mathbf{S} = (6,3)$	$\mathbf{S} = (9,4)$	$\mathbf{S} = (8,7)$
ς		73.692	82.832	86.422
Heuristic 2		**73.590**	**82.766**	**86.399**
Δ	-1000	72.173	80.919	80.229
	-100	72.173	80.919	80.229
	-50	72.051	80.812	80.169
	-40	71.823	80.612	80.055
	-25	70.469	79.429	79.384
	-15	67.282	76.644	77.804
	-10	63.880	73.673	76.121
	-9	62.954	72.865	75.664
	-8	61.924	72.385	75.156
	-6	59.501	72.565	75.578
	-5	58.663	72.641	76.291
	-4	58.510	72.536	76.980
	-3	58.499	72.192	77.537
	-2	58.341	71.575	77.906
	-1	57.881	70.663	78.060
	0	57.039	69.436	77.987
	1	55.976	67.968	77.680
	2	55.182	66.460	77.149
	3	55.483	65.267	76.418
	4	56.559	65.082	75.544
	5	57.525	65.615	74.646
	6	58.393	66.094	73.961
	7	59.172	66.525	73.958
	8	59.872	66.913	74.433
	9	60.501	67.262	74.860
	10	61.066	67.576	75.244
	15	63.146	68.732	76.660
	25	65.094	69.816	77.987
	40	65.920	70.277	78.551
	50	66.060	70.355	78.646
	100	66.134	70.396	78.697
	1000	66.135	70.397	78.697

APPENDIX G



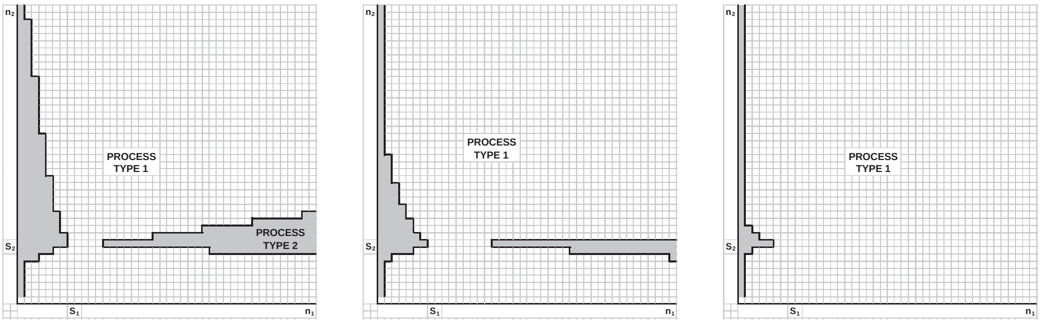
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure G.1: ς policy: equally weighted cost function, $S = 8$, $\lambda_1 = 4\lambda_2$



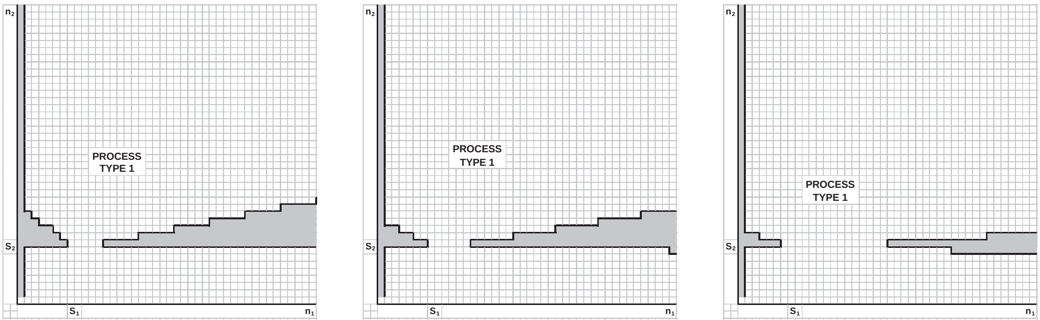
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure G.2: Heuristic 2: equally weighted cost function, $S = 8$, $\lambda_1 = 4\lambda_2$.



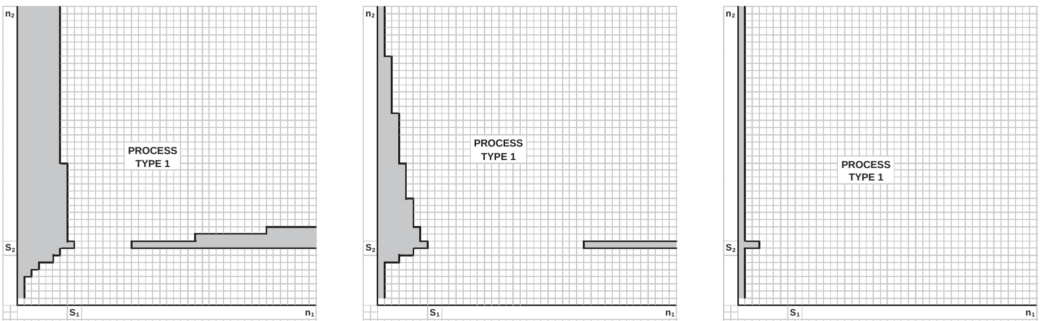
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure G.3: ς policy: cost function weighted by demand rates, $w_1 = 4w_2$, $S = 8$, $\lambda_1 = 4\lambda_2$.



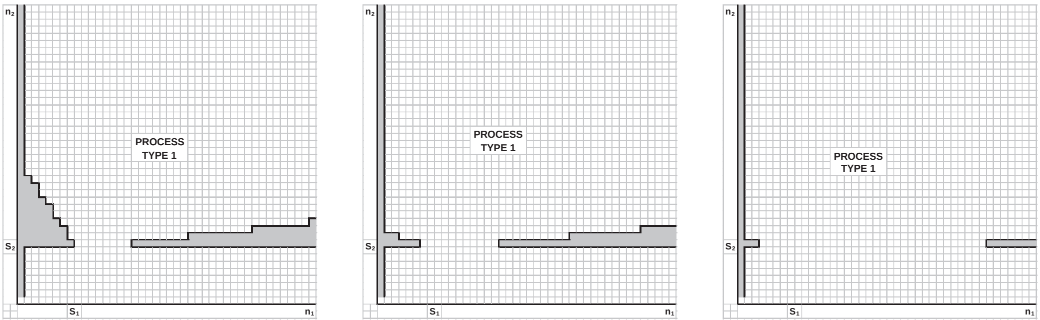
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure G.4: Heuristic 2: cost function weighted by demand rates, $w_1 = 4w_2$, $S = 8$, $\lambda_1 = 4\lambda_2$.



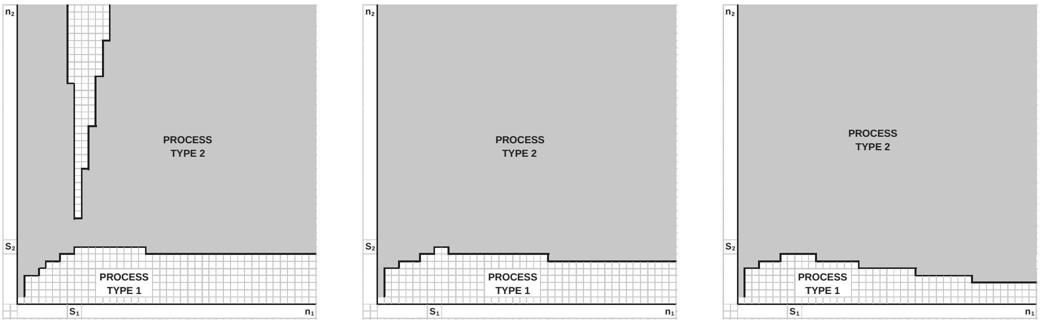
(a) $\rho = 0.10$. (b) $\rho = 0.40$. (c) $\rho = 0.90$.

Figure G.5: ς policy for $w_1 = 8w_2$, $S = 8$, $\lambda_1 = 4\lambda_2$.



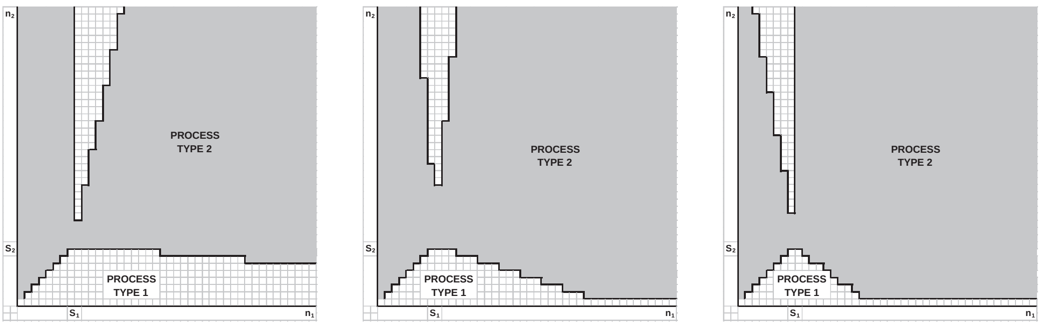
(a) $\rho = 0.10$. (b) $\rho = 0.40$. (c) $\rho = 0.90$.

Figure G.6: Heuristic 2 (M) for $w_1 = 8w_2$, $S = 8$, $\lambda_1 = 4\lambda_2$.



(a) $\rho = 0.10$, $S = 8$, $\lambda_1 = 4\lambda_2$. (b) $\rho = 0.40$, $S = 8$, $\lambda_1 = 4\lambda_2$. (c) $\rho = 0.90$, $S = 8$, $\lambda_1 = 4\lambda_2$.

Figure G.7: ς policy for $w_2 = 4w_1$, $S = 8$, $\lambda_1 = 4\lambda_2$.



(a) $\rho = 0.10$. (b) $\rho = 0.40$. (c) $\rho = 0.90$.

Figure G.8: Heuristic 2 (M) for $w_2 = 4w_1$, $S = 8$, $\lambda_1 = 4\lambda_2$.

APPENDIX H

Table H.1: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): $\lambda_1 = 2\lambda_2$, $w_2 = 2w_1$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	80.07	95.31	98.90	99.74	99.98	100.00	100.00	100.00
	FCFS	77.62	94.64	98.63	99.63	99.97	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	**80.07**	**95.26**	**98.88**	**99.73**	**99.98**	**100.00**	**100.00**	**100.00**
0.60	ς	68.78	87.72	94.96	97.97	99.68	99.95	100.00	100.00
	FCFS	61.11	84.26	93.36	97.09	99.39	99.86	99.98	**100.00**
	Heuristic 2 (M)	**68.78**	**87.51**	**94.86**	**97.89**	**99.65**	**99.94**	**100.00**	**100.00**
0.80	ς	56.85	75.72	84.39	89.81	95.70	98.22	99.53	99.92
	FCFS	37.66	60.60	74.74	83.57	92.75	96.63	98.86	99.70
	Heuristic 2 (M)	**56.85**	**75.52**	**84.28**	**89.66**	**95.55**	**98.12**	**99.49**	**99.91**
0.90	ς	50.70	68.03	75.59	80.41	87.10	91.52	95.50	98.06
	FCFS	21.43	38.01	50.88	60.91	74.92	83.61	91.07	95.81
	Heuristic 2 (M)	**50.70**	**68.02**	**75.58**	**80.30**	**86.84**	**91.22**	**95.26**	**97.94**

Table H.2: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): $\lambda_1 = 2\lambda_2$, $w_2 = 5w_1$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	83.37	96.64	99.27	99.83	99.99	100.00	100.00	100.00
	FCFS	79.72	95.67	99.01	99.76	**99.98**	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	**83.37**	**96.64**	**99.21**	**99.81**	**99.98**	**100.00**	**100.00**	**100.00**
0.60	ς	74.39	91.67	96.76	98.71	99.80	99.97	100.00	100.00
	FCFS	63.89	86.57	94.83	97.93	99.63	**99.92**	**99.99**	**100.00**
	Heuristic 2 (M)	**74.39**	**91.67**	**96.65**	**98.58**	**99.66**	**99.92**	**99.99**	**100.00**
0.80	ς	65.09	84.21	90.81	94.11	97.52	98.97	99.73	99.95
	FCFS	40.26	63.97	78.04	86.45	94.63	97.75	99.32	99.84
	Heuristic 2 (M)	**65.09**	**84.21**	**90.81**	**93.98**	**96.93**	**98.47**	**99.47**	**99.88**
0.90	ς	60.35	79.48	86.19	89.33	93.04	95.42	97.57	98.95
	FCFS	23.21	40.88	54.35	64.64	78.56	86.80	93.42	97.24
	Heuristic 2 (M)	**60.35**	**79.48**	**86.19**	**89.19**	**92.37**	**94.48**	**96.68**	**98.37**

Table H.3: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): equally weighted case, $\lambda_1 = 4\lambda_2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	77.95	93.90	98.19	99.45	99.95	99.99	100.00	100.00
	FCFS	76.73	93.26	97.81	99.26	99.91	**99.99**	**100.00**	**100.00**
	Heuristic 2 (M)	**77.95**	**93.90**	**98.19**	**99.44**	**99.95**	**99.99**	**100.00**	**100.00**
0.60	ς	65.76	84.56	92.89	96.73	99.30	99.85	99.98	100.00
	FCFS	61.19	82.46	91.27	95.43	98.68	99.61	99.94	99.99
	Heuristic 2 (M)	**65.76**	**84.51**	**92.87**	**96.68**	**99.29**	**99.84**	**99.98**	**100.00**
0.80	ς	53.05	69.22	78.93	85.76	93.64	97.22	99.22	99.86
	FCFS	39.68	61.10	73.50	81.20	89.83	94.25	97.48	99.15
	Heuristic 2 (M)	**53.04**	**68.76**	**78.63**	**85.44**	**93.46**	**97.13**	**99.19**	**99.85**
0.90	ς	46.56	59.30	66.53	72.44	81.55	87.80	93.52	97.23
	FCFS	23.95	40.79	52.87	61.74	73.56	80.88	87.65	92.83
	Heuristic 2 (M)	**46.56**	**58.87**	**66.12**	**71.99**	**81.17**	**87.51**	**93.33**	**97.15**

Table H.4: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): weighted by demand rates, $\lambda_1 = 4\lambda_2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	71.28	91.21	97.27	99.14	99.91	99.99	100.00	100.00
	FCFS	69.82	90.04	96.60	98.83	99.86	99.98	**100.00**	**100.00**
	Heuristic 2 (M)	**71.27**	**91.21**	**97.27**	**99.14**	**99.91**	**99.99**	**100.00**	**100.00**
0.60	ς	55.11	78.91	90.05	95.29	98.95	99.76	99.97	100.00
	FCFS	51.75	75.13	86.77	92.86	97.89	99.37	99.90	99.99
	Heuristic 2 (M)	**55.02**	**78.91**	**90.04**	**95.29**	**98.94**	**99.76**	**99.97**	**100.00**
0.80	ς	37.01	58.92	73.09	82.31	92.27	96.64	99.05	99.83
	FCFS	30.16	49.61	62.86	72.26	84.20	90.89	95.98	98.65
	Heuristic 2 (M)	**36.65**	**58.91**	**73.06**	**82.28**	**92.24**	**96.58**	**98.98**	**99.79**
0.90	ς	26.96	45.27	58.76	68.62	81.38	88.58	94.30	97.67
	FCFS	16.90	30.06	40.53	49.03	61.93	71.15	80.71	88.60
	Heuristic 2 (M)	**26.59**	**45.26**	**58.75**	**68.61**	**81.33**	**88.52**	**94.16**	**97.43**

Table H.5: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): $\lambda_1 = 4\lambda_2$, $w_1 = 8w_2$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	69.80	90.56	97.02	99.06	99.90	99.99	100.00	100.00
	FCFS	67.77	89.09	96.24	98.70	99.84	99.98	**100.00**	**100.00**
	Heuristic 2 (M)	**69.80**	**90.56**	**97.02**	**99.05**	**99.90**	**99.99**	**100.00**	**100.00**
0.60	ς	53.63	78.04	89.52	95.00	98.86	99.74	99.97	100.00
	FCFS	48.95	72.96	85.44	92.10	97.66	99.30	99.89	99.99
	Heuristic 2 (M)	**53.63**	**78.04**	**89.52**	**94.98**	**98.86**	**99.74**	**99.97**	**100.00**
0.80	ς	36.34	58.97	73.40	82.67	92.57	96.80	99.10	99.84
	FCFS	27.34	46.21	59.71	69.61	82.53	89.89	95.53	98.50
	Heuristic 2 (M)	**36.34**	**58.97**	**73.40**	**82.67**	**92.54**	**96.77**	**99.05**	**99.80**
0.90	ς	27.22	46.55	60.50	70.62	83.39	90.33	95.46	98.23
	FCFS	14.81	26.88	36.87	45.27	58.48	68.27	78.65	87.35
	Heuristic 2 (M)	**27.22**	**46.55**	**60.50**	**70.62**	**83.39**	**90.33**	**95.41**	**98.10**

Table H.6: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): $\lambda_1 = 4\lambda_2$, $w_2 = 4w_1$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	86.36	96.91	99.19	99.76	99.98	100.00	100.00	100.00
	FCFS	83.63	96.47	99.03	99.69	99.96	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	**86.36**	**96.87**	**99.16**	**99.75**	**99.98**	**100.00**	**100.00**	**100.00**
0.60	ς	79.09	92.41	96.71	98.53	99.70	99.94	99.99	100.00
	FCFS	70.63	89.79	95.77	98.00	99.46	99.84	99.97	**100.00**
	Heuristic 2 (M)	**79.09**	**92.20**	**96.59**	**98.44**	**99.66**	**99.92**	**99.99**	**100.00**
0.80	ς	71.61	85.67	90.59	93.67	97.18	98.77	99.65	99.94
	FCFS	49.21	72.59	84.13	90.14	95.47	97.61	98.98	99.66
	Heuristic 2 (M)	**71.61**	**85.53**	**90.48**	**93.42**	**96.88**	**98.55**	**99.56**	**99.91**
0.90	ς	67.82	81.51	85.66	88.26	92.13	94.78	97.21	98.81
	FCFS	31.01	51.52	65.21	74.45	85.19	90.60	94.60	97.05
	Heuristic 2 (M)	**67.82**	**81.38**	**85.56**	**88.05**	**91.62**	**94.21**	**96.74**	**98.53**

Table H.7: Comparison of the aggregate fill rates (%) of the ς, FCFS policies and heuristic 2 (M): $\lambda_1 = 4\lambda_2$, $w_2 = 8w_1$.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.40	ς	88.87	97.95	99.50	99.86	99.99	100.00	100.00	100.00
	FCFS	85.68	97.43	99.39	99.82	**99.98**	**100.00**	**100.00**	**100.00**
	Heuristic 2 (M)	**88.87**	**97.95**	**99.47**	**99.83**	**99.98**	**100.00**	**100.00**	**100.00**
0.60	ς	83.05	95.09	97.95	99.12	99.82	99.96	100.00	100.00
	FCFS	73.43	91.96	97.10	98.76	99.69	99.91	**99.99**	**100.00**
	Heuristic 2 (M)	**83.05**	**95.03**	**97.87**	**98.97**	**99.72**	**99.92**	**99.99**	**100.00**
0.80	ς	77.12	90.90	94.39	96.27	98.34	99.28	99.80	99.96
	FCFS	52.03	75.99	87.28	92.79	97.14	98.60	99.43	99.81
	Heuristic 2 (M)	**77.12**	**90.90**	**94.33**	**95.96**	**97.80**	**98.81**	**99.55**	**99.88**
0.90	ς	74.12	88.29	91.63	93.24	95.47	96.99	98.39	99.31
	FCFS	33.10	54.70	68.86	78.21	88.63	93.48	96.65	98.30
	Heuristic 2 (M)	**74.12**	**88.29**	**91.62**	**93.04**	**94.79**	**96.11**	**97.56**	**98.74**

APPENDIX I

Table I.1: Aggregate fill rates (%) of symmetric three-class systems for Heuristic 2.

		simulation				
ρ	S	Lower	Mean	Upper	Relative Precision	Value-Iteration
0.10	1	94.75	94.75	94.76	0.00008	94.76
	2	99.73	99.73	99.74	0.00001	99.73
	3	99.99	99.99	99.99	0.00000	99.99
0.20	1	89.06	89.07	89.08	0.00013	89.07
	2	98.85	98.85	98.85	0.00003	98.85
	3	99.88	99.88	99.88	0.00001	99.88
0.30	1	82.93	82.95	82.97	0.00020	82.94
	2	97.14	97.15	97.16	0.00008	97.15
	3	99.55	99.55	99.55	0.00003	89.55
	4	99.93	99.93	99.93	0.00001	99.93
0.40	1	76.36	76.38	76.41	0.00028	76.38
	2	94.35	94.36	94.37	0.00012	94.35
	3	98.72	98.73	98.73	0.00006	98.73
	4	99.72	99.72	99.72	0.00002	99.72
0.50	1	69.34	69.36	69.38	0.00025	69.35
	2	90.12	90.13	90.15	0.00018	90.14
	3	96.98	96.99	97.00	0.00009	96.99
	4	99.11	99.12	99.12	0.00006	99.11
	6	99.93	99.93	99.93	0.00001	99.93
0.60	1	61.79	61.82	61.85	0.00041	61.82
	2	84.13	84.15	84.17	0.00024	84.21
	3	93.59	93.61	93.64	0.00026	93.64
	4	97.50	97.51	97.51	0.00009	97.51
	6	99.63	99.64	99.64	0.00005	99.64

APPENDIX J

Table J.1: Performance measures for $\lambda_i = \lambda$, $w_i = \frac{1}{2}$, $S_i = S$ for $i = 1, 2$.

		FR			*B*			*Inv*		
ρ	S	FCFS	LQ	H2	FCFS	LQ	H2	FCFS	LQ	H2
0.10	1	0.9474	0.9472	**0.9476**	0.0029	**0.0028**	0.0032	0.9474	**0.9472**	0.9476
	2	0.9972	**0.9973**	**0.9973**	0.0002	**0.0001**	**0.0001**	**1.9446**	**1.9446**	**1.9446**
	3	**0.9999**	**0.9999**	**0.9999**	**0.0000**	**0.0000**	**0.0000**	**2.9445**	**2.9445**	**2.9445**
	4	**1.0000**	**1.0000**	**1.0000**	**0.0000**	**0.0000**	**0.0000**	**3.9444**	**3.9444**	**3.9444**
0.25	1	0.8571	0.8553	**0.8606**	0.0238	**0.0220**	0.0273	0.8571	**0.8553**	0.8606
	2	0.9796	0.9809	**0.9812**	0.0034	**0.0028**	0.0031	1.8367	**1.8361**	1.8364
	3	0.9971	0.9975	**0.9976**	0.0005	**0.0004**	**0.0004**	2.8338	**2.8337**	**2.8337**
	4	0.9996	**0.9997**	**0.9997**	0.0001	**0.0000**	**0.0000**	**3.8334**	**3.8334**	**3.8334**
	6	**1.0000**	**1.0000**	**1.0000**	**0.0000**	**0.0000**	**0.0000**	**5.8333**	**5.8333**	**5.8333**
0.50	1	0.6667	0.6545	**0.6935**	0.1667	**0.1545**	0.1935	0.6667	**0.6545**	0.6935
	2	0.8889	0.8913	**0.9014**	0.0556	**0.0458**	0.0618	1.5556	**1.5458**	1.5618
	3	0.9630	0.9674	**0.9699**	0.0185	**0.0132**	0.0174	2.5185	**2.5132**	2.5174
	4	0.9877	0.9905	**0.9911**	0.0062	**0.0037**	0.0048	3.5062	**3.5037**	3.5048
	6	0.9986	0.9992	**0.9993**	0.0007	**0.0003**	0.0004	5.5007	**5.5003**	5.5004
	8	0.9998	**0.9999**	**0.9999**	0.0001	**0.0000**	**0.0000**	7.5001	**7.5000**	**7.5000**
	11	**1.0000**	**1.0000**	**1.0000**	**0.0000**	**0.0000**	**0.0000**	**10.5000**	**10.5000**	**10.5000**
0.60	1	0.5714	0.5532	**0.6182**	0.3214	**0.3032**	0.3682	0.5714	**0.5532**	0.6182
	2	0.8163	0.8155	**0.8415**	0.1378	**0.1187**	0.1611	1.3878	**1.3687**	1.4111
	3	0.9213	0.9268	**0.9362**	0.0590	**0.0456**	0.0631	2.3090	**2.2956**	2.3131
	4	0.9663	0.9717	**0.9750**	0.0253	**0.0172**	0.0237	3.2753	**3.2672**	3.2737
	6	0.9938	0.9959	**0.9964**	0.0046	**0.0024**	0.0032	5.2546	**5.2524**	5.2532
	8	0.9989	0.9994	**0.9995**	0.0009	**0.0003**	0.0004	7.2509	**7.2503**	7.2504
	11	0.9999	**1.0000**	**1.0000**	0.0001	**0.0000**	**0.0000**	10.2501	**10.2500**	**10.2500**
	15	**1.0000**	**1.0000**	**1.0000**	**0.0000**	**0.0000**	**0.0000**	**14.2500**	**14.2500**	**14.2500**
0.75	1	0.4000	0.3750	**0.4943**	0.9000	**0.8750**	0.9943	0.4000	**0.3750**	0.4943
	2	0.6400	0.6281	**0.7143**	0.5400	**0.5031**	0.6648	1.0400	**1.0031**	1.1648
	3	0.7840	0.7838	**0.8344**	0.3240	**0.2869**	0.4097	1.8240	**1.7869**	1.9097
	4	0.8704	0.8759	**0.9046**	0.1944	**0.1628**	0.2451	2.6944	**2.6628**	2.7451
	6	0.9533	0.9599	**0.9690**	0.0700	**0.0520**	0.0816	4.5700	**4.5520**	4.5816
	8	0.9832	0.9872	**0.9901**	0.0252	**0.0165**	0.0263	6.5252	**6.5165**	6.5263
	11	0.9964	0.9977	**0.9982**	0.0054	**0.0029**	0.0047	9.5054	**9.5029**	9.5047
	15	0.9995	**0.9998**	**0.9998**	0.0007	**0.0003**	0.0005	13.5007	**13.5003**	13.5005
0.90	1	0.1818	0.1623	**0.3546**	3.6818	**3.6622**	3.8546	0.1818	**0.1623**	0.3546
	2	0.3306	0.3112	**0.5317**	3.0124	**2.9734**	3.3569	0.5124	**0.4734**	0.8569
	3	0.4523	0.4380	**0.6374**	2.4647	**2.4114**	2.9097	0.9647	**0.9114**	1.4097
	4	0.5519	0.5431	**0.7113**	2.0166	**1.9545**	2.5008	1.5166	**1.4545**	2.0008
	6	0.7000	0.6993	**0.8119**	1.3499	**1.2830**	1.7961	2.8499	**2.7830**	3.2961
	8	0.7992	0.8026	**0.8766**	0.9037	**0.8419**	1.2566	4.4037	**4.3419**	4.7566
	11	0.8900	0.8950	**0.9344**	0.4949	**0.4474**	0.7138	6.9949	**6.9474**	7.2139
	15	0.9507	0.9548	**0.9717**	0.2218	**0.1926**	0.3238	10.7218	**10.6926**	10.8238

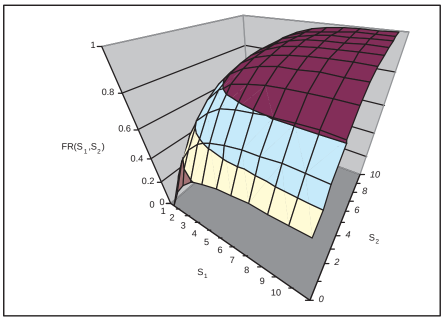
Table J.2: Differences in performance measures with respect to Heuristic 2 when $\lambda_i = \lambda$, $w_i = \frac{1}{2}$, $S_i = S$ for $i = 1, 2$.

		FR		$B)$		Inv	
ρ	S	FCFS	LQ	FCFS	LQ	FCFS	LQ
0.10	1	0.00023	0.00036	-0.00023	-0.00036	-0.00023	-0.00036
	2	0.00011	0.00000	0.00001	0.00000	0.00001	0.00000
	3	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000
	4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.25	1	0.00345	0.00530	-0.00345	-0.00530	-0.00345	-0.00530
	2	0.00156	0.00030	0.00029	-0.00030	0.00029	-0.00030
	3	0.00048	0.00002	0.00011	-0.00002	0.00011	-0.00002
	4	0.00011	0.00000	0.00002	0.00000	0.00002	0.00000
	6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.50	1	0.02680	0.03896	-0.02680	-0.03896	-0.02680	-0.03896
	2	0.01250	0.01008	-0.00620	-0.01593	-0.00620	-0.01593
	3	0.00693	0.00248	0.00117	-0.00412	0.00117	-0.00412
	4	0.00348	0.00061	0.00139	-0.00103	0.00139	-0.00103
	6	0.00065	0.00004	0.00033	-0.00006	0.00033	-0.00006
	8	0.00010	0.00000	0.00005	0.00000	0.00005	0.00000
	11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	1	0.04672	0.06491	-0.04672	-0.06491	-0.04672	-0.06491
	2	0.02515	0.02597	-0.02334	-0.04235	-0.02334	-0.04235
	3	0.01487	0.00935	-0.00403	-0.01752	-0.00403	-0.01752
	4	0.00876	0.00335	0.00161	-0.00646	0.00161	-0.00646
	6	0.00258	0.00043	0.00141	-0.00085	0.00141	-0.00085
	8	0.00063	0.00006	0.00042	-0.00011	0.00042	-0.00011
	11	0.00006	0.00000	0.00005	-0.00001	0.00005	-0.00001
	15	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.75	1	0.09435	0.11935	-0.09435	-0.11935	-0.09435	-0.11935
	2	0.07433	0.08620	-0.12484	-0.16172	-0.12484	-0.16172
	3	0.05042	0.05062	-0.08570	-0.12278	-0.08570	-0.12278
	4	0.03418	0.02869	-0.05072	-0.08231	-0.05072	-0.08231
	6	0.01563	0.00908	-0.01158	-0.02956	-0.01158	-0.02956
	8	0.00687	0.00287	-0.00113	-0.00980	-0.00113	-0.00980
	11	0.00185	0.00051	0.00072	-0.00178	0.00072	-0.00178
	15	0.00029	0.00005	0.00023	-0.00018	0.00023	-0.00018
0.90	1	0.17277	0.19232	-0.17273	-0.19232	-0.17277	-0.19232
	2	0.20111	0.22052	-0.34448	-0.38346	-0.34451	-0.38346
	3	0.18516	0.19948	-0.44500	-0.49829	-0.44501	-0.49829
	4	0.15938	0.16818	-0.48427	-0.54637	-0.48428	-0.54637
	6	0.11189	0.11257	-0.44619	-0.51314	-0.44620	-0.51314
	8	0.07740	0.07401	-0.35291	-0.41470	-0.35293	-0.41470
	11	0.04435	0.03931	-0.21889	-0.26641	-0.21890	-0.26641
	15	0.02102	0.01691	-0.10198	-0.13117	-0.10199	-0.13117

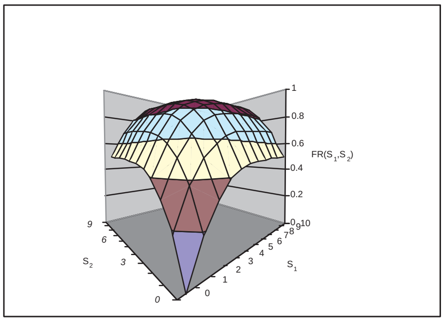
Table J.3: Differences (%) in performance measures with respect to Heuristic 2, when $\lambda_i = \lambda$, $w_i = \frac{1}{2}$, $S_i = S$ for $i = 1, 2$.

		FR(S)		B(S)		I(S)	
ρ	S	FCFS	LQ	FCFS	LQ	FCFS	LQ
0.10	1	0.02476	0.03786	-8.02204	-12.80664	-0.02476	-0.03786
	2	0.01137	0.00031	6.87859	-2.22256	0.00054	-0.00016
	3	0.00126	0.00000	13.08706	-0.40342	0.00004	0.00000
	4	0.00010	0.00000	17.69887	-0.07505	0.00000	0.00000
0.25	1	0.40257	0.61994	-14.49264	-24.14781	-0.40257	-0.61994
	2	0.15954	0.03087	8.48560	-10.78510	0.01571	-0.01650
	3	0.04801	0.00179	23.23886	-5.04438	0.00398	-0.00063
	4	0.01072	0.00011	34.17065	-2.41544	0.00062	-0.00003
	6	0.00036	0.00000	50.31778	-0.57332	0.00001	0.00000
0.50	1	4.02041	5.95270	-16.08164	-25.21602	-4.02041	-5.95270
	2	1.40636	1.13069	-11.15334	-34.76937	-0.39833	-1.03062
	3	0.71932	0.25641	6.29649	-31.16357	0.04630	-0.16405
	4	0.35215	0.06202	22.58085	-27.51856	0.03976	-0.02943
	6	0.06531	0.00381	48.50990	-22.23413	0.00605	-0.00117
	8	0.00972	0.00024	66.58008	-18.67285	0.00068	-0.00005
	11	0.00045	0.00000	83.08327	-15.07203	0.00002	0.00000
0.60	1	8.17631	11.73315	-14.53566	-21.40637	-8.17631	-11.73315
	2	3.08105	3.18484	-16.94470	-35.66826	-1.68201	-3.09435
	3	1.61445	1.00852	-6.83448	-38.46061	-0.17475	-0.76321
	4	0.90671	0.34482	6.38147	-37.49631	0.04930	-0.19771
	6	0.25915	0.04342	30.28715	-35.30302	0.02679	-0.01609
	8	0.06357	0.00560	49.12140	-33.74742	0.00578	-0.00151
	11	0.00646	0.00026	68.89561	-32.37821	0.00045	-0.00005
	15	0.00026	0.00000	84.17979	-31.47548	0.00001	0.00000
0.75	1	23.58719	31.82634	-10.48320	-13.63986	-23.58719	-31.82634
	2	11.61381	13.72392	-23.11926	-32.14291	-12.00423	-16.12152
	3	6.43068	6.45826	-26.45121	-42.79215	-4.69857	-6.87104
	4	3.92746	3.27604	-26.09169	-50.55544	-1.88251	-3.09110
	6	1.63976	0.94593	-16.54058	-56.85630	-0.25330	-0.64946
	8	0.69833	0.29078	-4.47737	-59.36054	-0.01729	-0.15046
	11	0.18564	0.05119	13.18235	-60.36552	0.00755	-0.01871
	15	0.02919	0.00511	32.78729	-60.60506	0.00171	-0.00133
0.90	1	95.02356	118.52000	-4.69152	-5.25147	-95.02407	-118.52008
	2	60.83636	70.86582	-11.43542	-12.89637	-67.23474	-80.99431
	3	40.93710	45.54728	-18.05496	-20.66414	-46.13030	-54.67288
	4	28.87914	30.96894	-24.01454	-27.95493	-31.93295	-37.56459
	6	15.98392	16.09696	-33.05276	-39.99632	-15.65668	-18.43855
	8	9.68488	9.22179	-39.05315	-49.25917	-8.01439	-9.55121
	11	4.98289	4.39193	-44.22413	-59.54143	-3.12943	-3.83459
	15	2.21111	1.77126	-45.97671	-68.10494	-0.95124	-1.22675

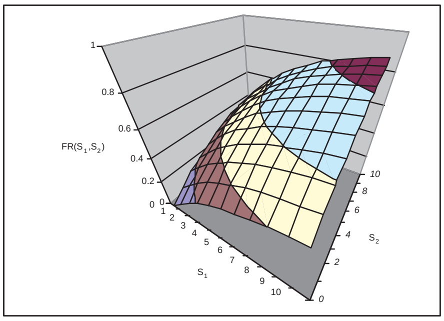
APPENDIX K



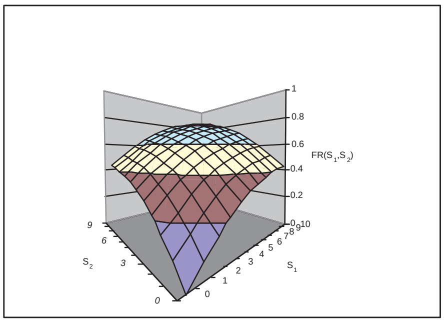
(a) $\rho = 0.75$.



(b) $\rho = 0.75$.

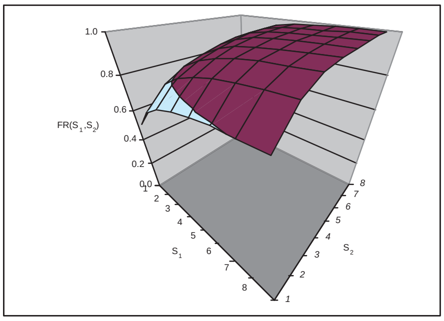


(c) $\rho = 0.90$.

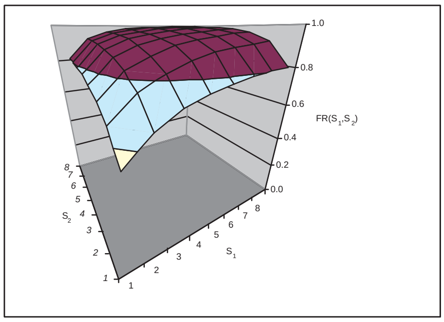


(d) $\rho = 0.90$.

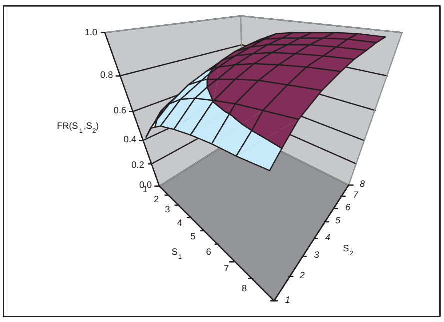
Figure K.1: Concavity of FR in $\mathbf{S}$ under LQ policy.



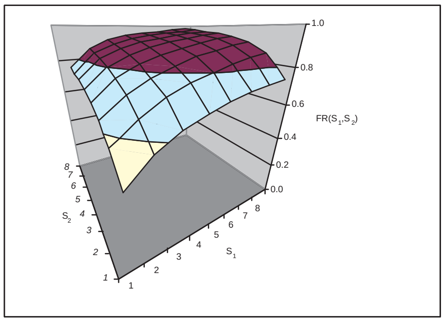
(a) $\rho = 0.70$.



(b) $\rho = 0.70$.

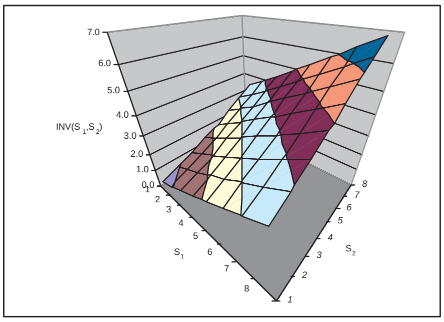


(c) $\rho = 0.80$.

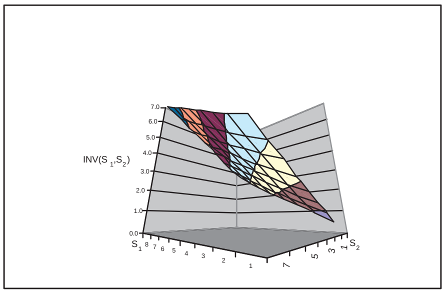


(d) $\rho = 0.80$.

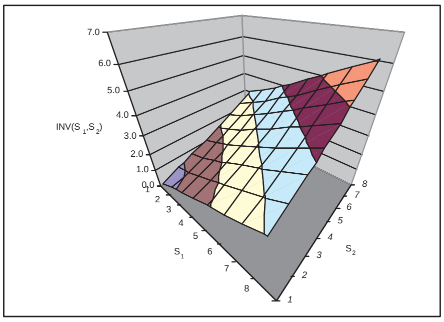
Figure K.2: Concavity of FR in $\mathbf{S}$ under ς policy.



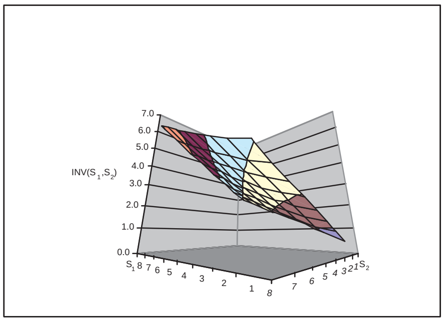
(a) $\rho = 0.70$.



(b) $\rho = 0.70$.



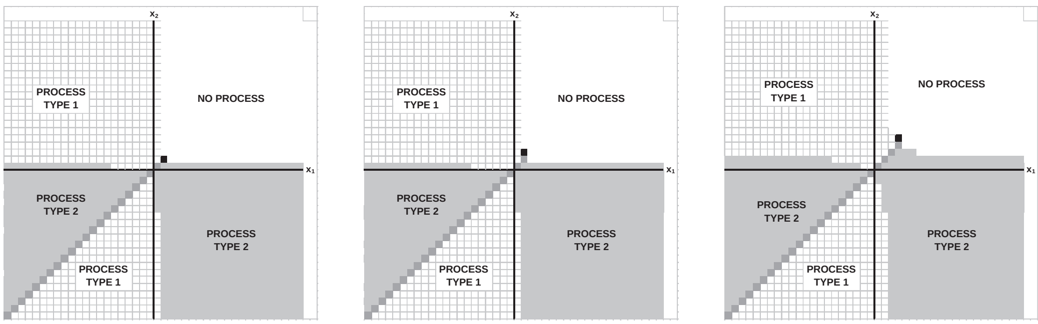
(c) $\rho = 0.80$.



(d) $\rho = 0.80$.

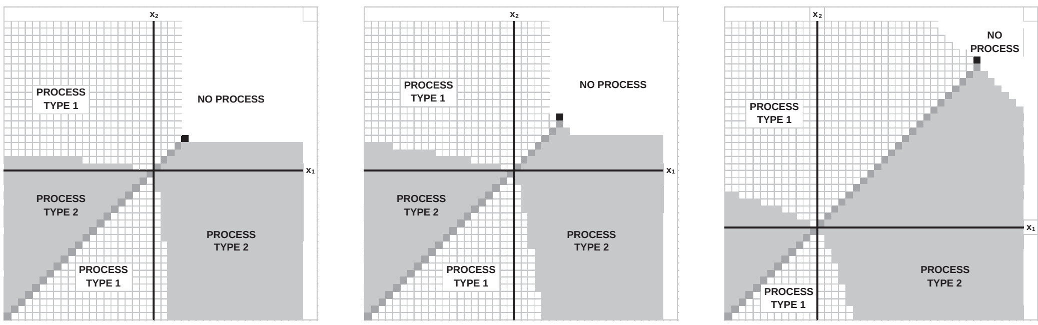
Figure K.3: Convexity of Inv in $\mathbf{S}$ under ς policy.

APPENDIX L



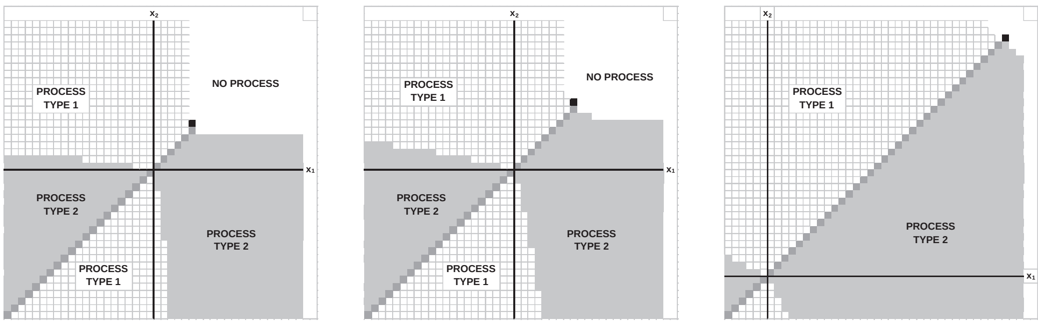
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure L.1: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 10$, $h_i = 1$ for $i = 1, 2$.



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

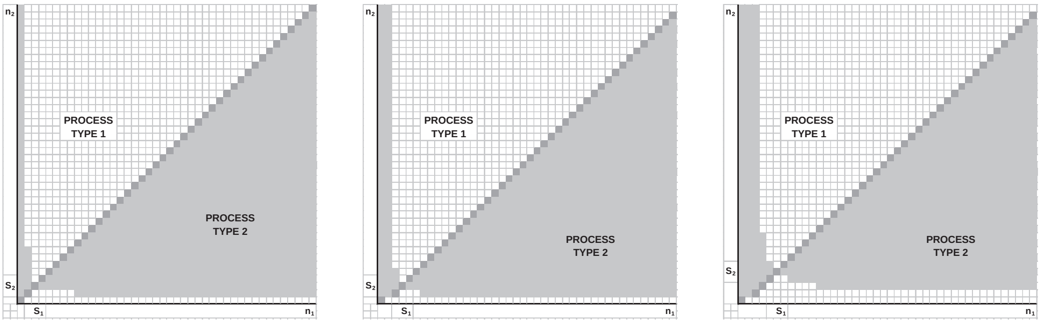
Figure L.2: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 1000$, $h_i = 1$ for $i = 1, 2$.



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

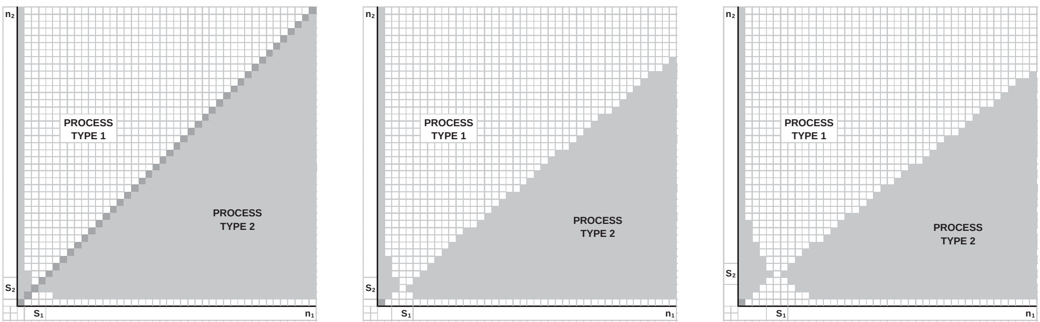
Figure L.3: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 10000$, $h_i = 1$ for $i = 1, 2$.

APPENDIX M



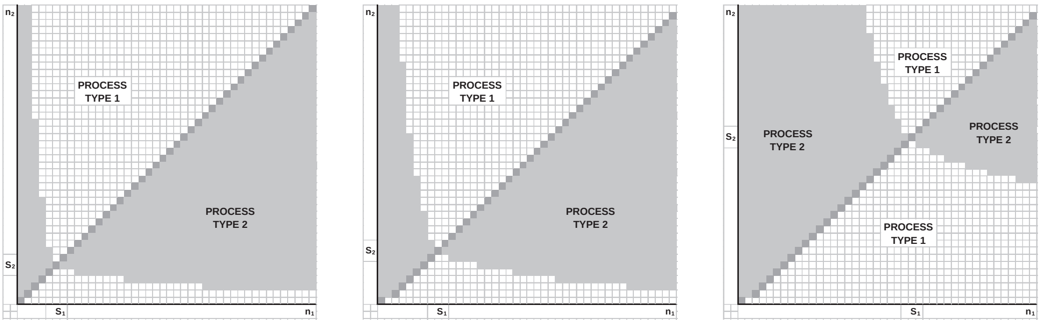
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure M.1: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 10$, $h_i = 1$ for $i = 1,\ 2$.



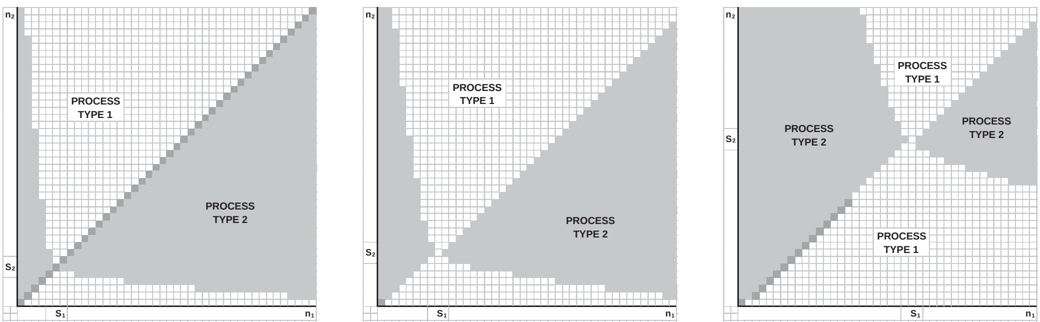
(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure M.2: Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$, where $\beta_0 = FR(\pi^*)$ in Figure M.1 and $w_i = \frac{1}{2}$, $i = 1,\ 2$.



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

Figure M.3: Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$, $\lambda_i = \lambda$, $b_i = 1000$, $h_i = 1$ for $i = 1,\ 2$.



(a) $\rho = 0.40$. (b) $\rho = 0.60$. (c) $\rho = 0.90$.

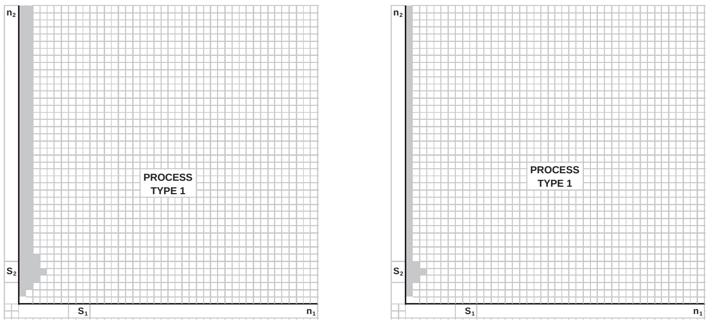
Figure M.4: Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$, where $\beta_0 = FR(\pi^*)$ in Figure M.3 and $w_i = \frac{1}{2}$, $i = 1,\ 2$.

APPENDIX N

Table N.1: Comparison of the policies π^* and $\varsigma_{\mathbf{S}^*}$ for $b_1 = 100$, $h_i = 1$ for $i = 1,\ 2$.

				FR			$\sum_i h_i Inv_i$		
b_2	$\frac{\lambda_1}{\lambda_2}$	$\frac{w_1}{w_2}$	ρ	π^*	$\varsigma^*_{\mathbf{S}}$	Error(%)	π^*	$\varsigma^*_{\mathbf{S}}$	Error(%)
200	2	1	0.40	97.414	97.415	0.001	2.17574	2.17592	0.008
			0.60	96.151	96.155	0.004	2.78954	2.79103	0.053
			0.80	91.198	91.233	0.038	3.31111	3.33869	0.833
200	4	2	0.40	97.135	97.137	0.002	2.17723	2.17750	0.013
			0.60	95.954	95.956	0.002	2.78685	2.78801	0.042
			0.80	92.151	92.202	0.055	3.74790	3.77337	0.679
50	2	4	0.40	94.744	94.749	0.006	1.69255	1.69383	0.076
			0.60	94.623	94.691	0.072	2.31985	2.33431	0.624
			0.80	89.566	89.697	0.147	2.57354	2.63387	2.344
25	3	12	0.40	96.119	96.119	0.000	1.69577	1.69577	0.000
			0.60	93.432	93.432	0.000	1.91009	1.91009	0.000
			0.80	90.465	90.887	0.466	2.34365	2.44084	4.147

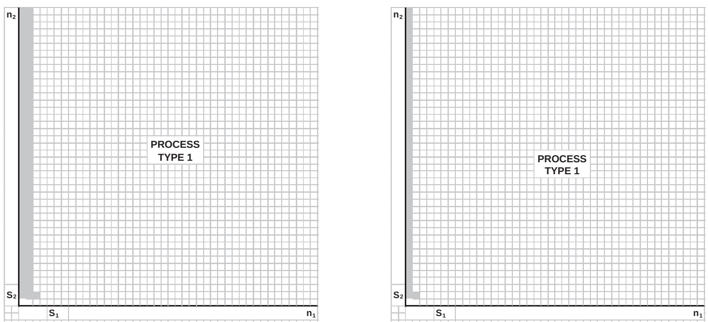
APPENDIX O



(a) Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$.

(b) Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$ where $\beta_0 = FR(\pi^*)$.

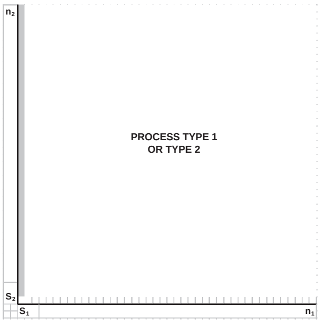
Figure O.1: Comparison of the Policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\mathbf{b} = (2000, 400)$, $\mathbf{h} = (8, 3)$, $\rho = 0.80$ and $\lambda_1 = 3\lambda_2$.



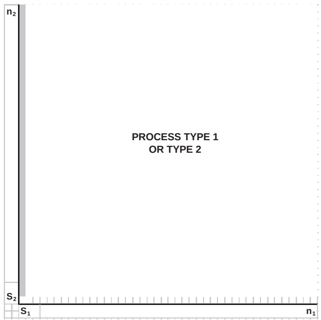
(a) Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$.

(b) Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$ where $\beta_0 = FR(\pi^*)$.

Figure O.2: Comparison of the Policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\mathbf{b} = (250, 100)$, $\mathbf{h} = (5, 4)$, $\rho = 0.80$ and $\lambda_1 = 3\lambda_2$.



(a) Optimal Scheduling Policy, π^*, for $P_6(\mathbf{b})$.



(b) Policy $\varsigma_{\mathbf{S}^*}$ for $P_5(\beta_0)$ where $\beta_0 = FR(\pi^*)$.

Figure O.3: Comparison of the Policies π^* and $\varsigma_{\mathbf{S}^*}$ for $\mathbf{b} = (20, 40)$, $\mathbf{h} = (7, 5)$, $\rho = 0.80$ and $\lambda_1 = 3\lambda_2$.