Introducing scan budgets to optimize allocation of radiology capacity

Anna Dankers

19 December 2022

Stochastic Operations Research
Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science

Supervisors:
Prof. dr. R.J. Boucherie (UT)
Dr. ir. A. Braaksma (UT)
Dr. ir. A.B. Zander (UT)
Dr. T.C. Kwee (UMCG)
Drs. J.P. Pennings (UMCG)
Dr. D. Yakar (UMCG)

Graduation committee:
Prof. dr. R.J. Boucherie (UT)
Dr. ir. A. Braaksma (UT)
Dr. A. Skopalik (UT)
Dr. ir. A.B. Zander (UT)

Abstract

We study a scan budget system as a solution to the problem of overload for a radiology department. To be able to deal with fluctuations of the demand, we introduce the possibility of *trading*: mutual exchange of budget among departments. The decentralized multi-agent problem is modeled using single-agent Markov Decision Processes that are connected via trading probabilities. We analyze different individual policies under a penalty and reward mechanism for trading and analyze the resulting collective behaviour using acceptance ratios. We find that the possibility of trading increases the acceptance ratios for all scan priorities and reduces the waste of capacity and the average annual departmental difference. When the number of trading possibilities is large, the acceptance ratios approach the acceptance ratios under centralized admission control.

Contents

Chapter 1

Introduction

Radiology departments specialize in imaging techniques. These techniques are used for many types of conditions and therefore requested by many different departments in the hospital. Hospitals often make a priori agreements on the number of scans that will be performed by each department on an annual basis. However, in practice many departments exceed this number of scans that was agreed upon. Moreover, cross-sectional imaging (CT and MRI) requests keep on rising over the years, while the radiology formation does not increase correspondingly. Therefore, there is increasing pressure on radiologists to process these scans. We investigate the consequences of rationing of cross-sectional imaging as an alternative to the current approach in which there is no limit on ordering CT and MRI scans. As a solution, we propose the introduction of scan budgets. Here, each department is assigned a budget, that can be used to request scans. Each type of scan has a certain price, which is deducted from the requesting department's budget when they request such a scan. If a department runs out of budget, it can no longer request scans. In this way, it is guaranteed that the number of requests does not exceed the budgeted capacity in the radiology department.

We make a distinction between type 1 and type 2 scans. Type 1 scans are scans that are medically necessary according to protocol. These scans are usually included in the capacity budget that is drawn up by the radiology department. Type 2 scans are additional scans that are not according to a protocol. The scan type is clear to both the radiology department and the applicant, and it can be determined objectively and quickly without the need for the scan result. A very straightforward solution to the problem described above would be to only allow requests for type 1 scans. However, because the total annual number of type 1 scan requests is a random variable, the capacity budget of the radiology department must be drawn up with sufficient margin. If only type 1 scans would be performed, there will usually be capacity left over. Not using this capacity leads to a waste of resources. Wasting capacity is not desirable and therefore requesting type 2 scans is also allowed.

A scan budget system would solve the problem of overload for the radiology department. A potential danger of a budget system, however, is that medically necessary scans cannot be performed because the responsible department ran out of budget, and thus there is a loss of health due to the system. A department may run out of budget because of poor departmental management by requesting too many type 2 scans, or because of stochasticity in demand, i.e., the number of type 1 scans is higher than expected during the year. To deal with these deviations, we introduce trading opportunities: departments can receive scan budget from and give away scan budget to other departments. In this way, individual capacity is more flexible, while overall capacity remains unchanged. Thus, departments that need extra capacity can get it from departments that do not need it during the year. We assume that departments are self-interested and want to minimize the number of scans they cannot perform. This implies that a department will not simply give away capacity if there is no reward in return. For this reason, we explore the possibility of rewarding departments through the scan budget of the following year. We can formulate the general research question as follows:

What are the effects of a scan budget system with a trading reward mechanism on the type 1 and type 2 acceptance ratios of departments?

Before we can answer this question, we need to answer the following sub-questions:

1. How can we model the acceptance and trading policy of a single department under a budget system with trading, and what does this policy look like?
2. What is the effect of a reward and penalty mechanism carried out by radiology for trading on a single department's willingness to trade?
3. How does a department's belief on other departments' willingness to trade influence the individual acceptance and trading policy?
4. What are the effects of the individual policies on the collective behaviour?
5. What are the acceptance ratios of type 1 and type 2 under the collective behaviour?

We provide answers to these questions for a two-department, homogeneous pricing case and analyze the situation over several years. With homogeneous pricing, we assume that the price for a scan is always 1. Therefore, the scan budget can also be seen as the department's remaining budget. Since we intend to focus more on the policy of departments under various budget systems than on adjusting the budget system to growing demand, we assume that demand is constant over the years. The focus of this research is on examining the individual and collective effects of a budget system with a reward and penalty mechanism. The relevance of this study mainly lies in the conceptual development of an evaluation model: we develop a framework for evaluating scan budget systems.

This thesis is organized as follows. In Chapter 2 we describe the situation at the radiology department, discuss the rising number of requests and motivate the choice of investigating a scan budget system, rather than investigating, for example, a model with centralized admission control. In Chapter 3, we provide an overview of relevant literature for the topics that we study. In Chapter 4, we present the mathematical problem formulation and the models and algorithms that we use to evaluate the individual policies and collective behaviour. In Chapter 5, we show results from numerical experiments and we perform sensitivity analyses for the parameters in the model. We conclude with drawing a conclusion and discussion of results and future research in Chapter 6.

Chapter 2

Motivation

In this section, we discuss the capacity budgeting problem that motivates this research. This problem has been present in the radiology department of University Medical Center Groningen (UMCG) for several years and particularly affects the workload for radiologists. This project is done in collaboration with the Department of Radiology at UMCG, but the problem has been observed in several hospitals throughout the Netherlands. We first give the reader a short introduction to the radiology department and requesting departments. After that, we describe the capacity problem and the current capacity budget in the UMCG and address important aspects of admission control. These aspects are discussed from both a decentralized admission control and a centralized admission control perspective. We introduce problems that arise in both setting and we propose ways to solve these problems. We conclude by introducing the research topics that we study in this research.

2.1 The radiology department of UMCG

The radiology department specializes in imaging techniques. These techniques allow images of the interior of the human being to be made. These images are used to make a diagnosis, to monitor disease progression and/or to determine treatment. In this way it is possible to see if there are any disorders in the body without the need for surgery. Radiology uses X-rays (e.g., a mammogram, CT scan), sound waves (ultrasound) and magnetic fields (MRI scan). In addition to examinations, radiology also performs treatments that are done with the help of imaging studies.

In UMCG, the limiting factor in the capacity of the radiology department is mainly the radiologists, the medical specialists in the radiology department. UMCG employs 120 radiology lab technicians and 35 radiologists. The radiology lab technicians supervise patients and make the scans. A radiologist's duties include preparing protocols for requests, reviewing scans and also sometimes performing treatments (interventional radiology). In addition, a radiologist is responsible for patients in the department, which includes intervening in contrast reactions and supervising scans. Radiologists have specialties such as abdominal (abdomen), cardiothoracic (heart, vascular, lung), interventional and neuro/head-neck radiology. The radiology department at UMCG has its own scheduling department, where requested scans are scheduled. Scheduled scans are only performed during the day (8:00-16:30) and only during weekdays. The hours outside of this are called on-call hours. Acute cases, such as requests from the emergency room and requests for patients admitted to the hospital, so-called inpatients, are scheduled as soon as possible. There is always enough capacity for acute cases, even during on-call hours.

Imaging studies are used for many types of conditions. When a doctor determines that imaging studies are necessary, he refers a patient to the radiology department for examination. The type of examination that should be performed is decided by the doctor. The request for the scan that arrives at radiology includes a number of details, such as clinical data (disease patient, infection parameters), the question (e.g., disease progression, differential diagnosis) and what type of examination is to be performed. A distinction is made between acute cases and non-acute cases. Acute requests are reviewed by a radiologist as soon as possible. Non-acute requests for ultrasound and X-ray examinations are immediately forwarded to the planning department. Non-acute CT, MRI and interventional requests are listed. The radiologist writes a protocol for the requests in this list, describing which areas of the body are to be scanned, and whether, for example, contrast medium will be used. Sometimes a radiologist rejects a request, for example if the same type of scan has been done before. This decisions is always made in consultation with the requesting doctor. Once a protocol for a scan has been established, the request is scheduled and the scan is performed by radiology lab technicians according to the prepared protocol. The completed scan returns to the radiologist, after which the radiologist makes an assessment which is forwarded to the requesting doctor. Acute requests, such as requests from the emergency room and inpatients (patients admitted to the hospital), are handled as quickly as possible.

In this research, referring departments play a central role. Because the UMCG is a tertiary care center, there are few requests from primary care (care that does not require a referral, including general practitioners). Almost all requests come from specialties within the UMCG. Within this study, we only consider CT and MRI requests, because these two types of requests cause most of the pressure on radiologists. Figure 2.1 gives an overview of the number of CT's and MRI's that were performed for the largest departments in UMCG in 2021.

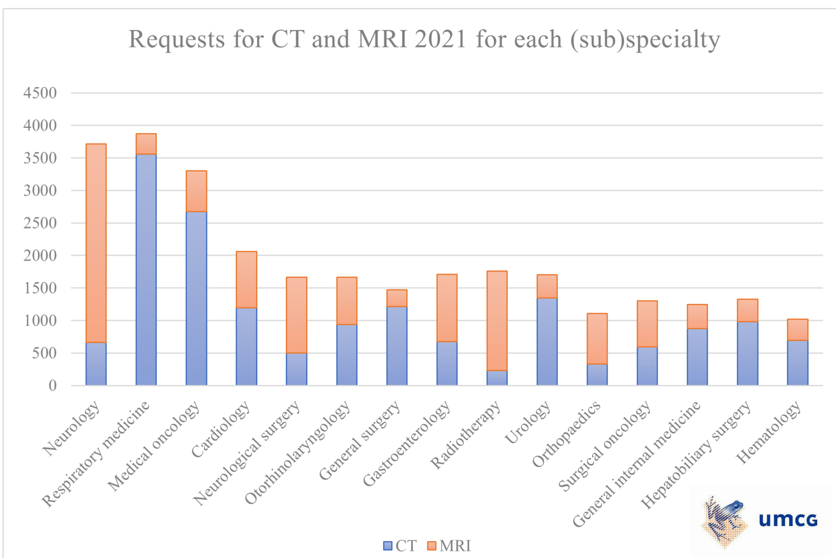

Figure 2.1: Number of CT & MRI scans that were performed per department in 2021.

2.2 Capacity budget problem

The total number of requests for CT and MRI scans in the Netherlands has been increasing for a number of years, while the number of radiologists is not increasing at the same rate. In August 2022, this was pointed out by the Nederlandse Vereniging voor Radiologie (NVvR):

More than 50 general hospitals and MSBs use the Logex benchmark for the production comparison of medical specialisms and distribution of budget and formation. For Radiology, the Logex benchmark has shown increasing production over the years with a constant formation. Because *there is a national increase, the benchmark has become increasingly higher and a 'rat race' has developed to cope with increasing production with the same formation. This has become an unsustainable system, which also does not take into account the increased care load and complexity of radiological care (shift from X-ray to CT and MRI, increase in number of images per scan). Nor does the current system take into account the increased severity of evening, night, weekend shifts despite the fact that radiology has become indispensable in the acute and emergency setting. The fact that radiological production increases without the formation growing sufficiently with it has been depicted at the national level for the period 2012-2018, see Figure 2.2. During this period, the number of so-called norm hours increased by about 240 hours per fte radiologist. This is an increase of 11%.* (source: Handreiking capaciteitsbegroting radiologie, NVvR, 24 August 2022)

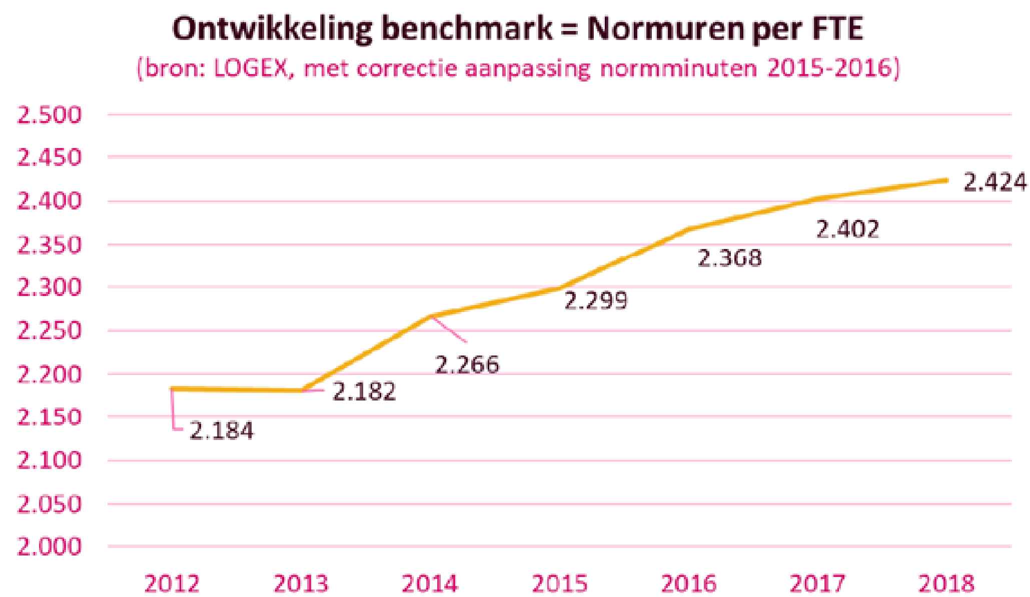

Figure 2.2

As a result of this, the pressure on radiologists to process scans is increasing. In this chapter, we describe how capacity is established and discuss the rising demand in more detail.

Every year, a sales budget is drawn up to determine the required capacity in the radiology department. This is done on the basis of care products and care activities. The Dutch healthcare system uses dbc care products. The dbc (diagnosis-treatment combination) is a 9-digit code that contains information about the diagnosis and treatments. A dbc care product, in turn, consists of care activities. These are components of a patient's treatment, such as a surgery, consultation, or MRI scan. All claimable services are expressed in so-called dbc care products. The dbc is thus a means of expressing costs of treatment, as not all treatments in a hospital are billed separately. With the dbc care product, a fixed amount can be declared. This amount is calculated on the basis of the costs of the care activities associated with a particular diagnosis.

The annual sales budget for radiology is not based on the number of care activities in previous years, but rather on the numbers of dbc's realized per specialty (department). The corresponding theoretical number of care activities that follows from the care profiles of the dbc's is then used to draw up the budget. However, doctors are not bound to the care activities included in the care product (dbc), so they can deviate from these numbers within their treatment plan. This means that doctors can request more scans than are included in the dbc. However, the sales budget and thus the capacity is based on the care activities in the dbc, so in practice, this capacity is often exceeded.

The fact that doctors request more scans than are included in the dbc on average and the resulting rising number of requests over the years can partly be explained by recent developments within imaging techniques. Because imaging is becoming increasingly sophisticated and is often an efficient alternative to other forms of examination, an increasing number of doctors are using imaging and it is used more often and more widely. However, there are indications that many imaging tests could be inappropriate. European studies that focused on a limited set of imaging tests found rates of inappropriate examinations between 15% and 35% [2], [10], [15]. A study that was carried at a tertiary referral academic institution in Ireland, including several radiograph subtypes, found that a significant amount of inappropriate radiographs are requested and performed, exposing patients to needless ionizing radiation and wasting staff members time at a financial cost [13]. Hence, it might be justifiable to reject some scan request, not only for capacity reasons but also from a medical perspective.

2.3 Centralized admission control vs. decentralized admission control

When capacity is limited such that not all scans requested can be performed, decisions must be made about which requests are rejected. This is referred to as admission control. In this subsection, we discuss two forms of admission control decision making: centralized and decentralized admission control.

In both cases the decision process consists of two consecutive steps. First, the priority of a request is determined using clinical information about the patient and the condition. Then, the priority is used to decide if the request is accepted. The agent that makes decisions about accepting and rejecting scans is called the *decision maker*. For the capacity problem that we address in this study there are two possibilities: the role of decision maker is fulfilled by the radiology department, or by requesting doctors. The latter case is called decentralized admission control. Imposing a scan budget system for departments is a form of decentralized admission control. Here, departments are given a scan budget, whereby they can prioritize scans themselves and use this priority to decide which scans will or will not be requested with their budget. The requesting department is thus the decision maker here. When the radiology department acts as the decision maker, we speak of centralized admission control. In this case, there is no limit for departments to request scans. Instead, radiology serves as a gatekeeper and evaluates the referral appropriateness and priority of a scan and performs or does not perform a scan, depending on the remaining capacity and the priority of a scan. In this chapter, we address several aspects of the admission control decision process and we discuss the implications of these aspects for both forms of admission control.

The first aspect is information access. To decide on the priority and appropriateness of a scan, information about the patient, specific conditions and the possible consequences for the patient is necessary. A referring doctor has more knowledge about these topics than a radiologist. With centralized admission control, this information is passed on to the radiologist, who then makes a decision based on the information obtained. In this process there is risk of information loss and other forms of noise. With decentralized admission, the decision maker (the referring doctor) has direct access to this information. Furthermore, because shared and limited capacity gives rise to a competitive situation, a possible consequence of the 'information gap' is the risk of unfair information provision. Much information is needed to prioritize a scan, and most of this information comes from the referring doctor. Because in centralized admission control all requests arrive at one central location, requests from different specialties are compared. In this context, it is conceivable that doctors will try to do everything possible to have their application given the highest possible priority when passing on information. This makes for less reliable information. With decentralized admission control, departments prioritize their own scans, so only applications within their own specialty are compared. As a result, self-interest plays less of a role and applications are compared more objectively.

We also consider the practical implementation and time. Under centralized admission control, requests are written by a referring doctor, after which the request is reviewed by the radiology department. In some cases, the request will be rejected. During the time it takes to submit a request and having it reviewed by the radiology department, the patient does not know whether the request will be accepted. With decentralized admission control, the referring doctor directly decides if the scan can be requested. Thus with decentralized admission control, waiting times are avoided. Moreover, decentralized admission control also avoids double and unnecessary work, as with centralized admission control, referring doctors write scan requests that may not be accepted by the radiology department. Moreover, centralized admission control places an additional burden on radiologists, because incoming requests from all departments must be reviewed and prioritized. With decentralized admission control, this work is divided over the departments. The decentralized approach, where requesting specialties function as gatekeepers is therefore more time-efficient.

Finally, we consider fluctuations in the demand. With scan budgets, fluctuations of the demand can be problematic, especially when the number of scans that arrives at a department during a year is much higher than expected and far exceeds the department's scan budget. In that case, medically necessary scans could possibly not be performed, which results in health loss. This is very undesirable. A system with centralized admission control will also experience fluctuations of the demand, but these fluctuations are relatively small, because the total demand is much larger. The scan capacity that is used for departments that have a larger demand than expected is balanced by the scan capacity that was not used by departments that have a smaller demand than expected. This is possible under centralized admission control, because the departments have shared capacity and there is a central agent (the radiology department) that coordinates the acceptance of scan requests. Under decentralized admission control, there is no shared capacity and hence no 'balancing' effect. Next to departments that experience an unexpected high demand and have the risk of rejecting scans that are medically necessary, there are also departments with a smaller demand than expected, that do not need all capacity they have. This results in waste of capacity. Hence, capacity allocation using a scan budget system can be unfair, because there can be large departmental differences in acceptance ratios.

Considering the access to information and the information gap between referring doctors and the radiology department, decentralized admission control is preferable over centralized admission control. Also regarding the original problem that we study, the workload for radiologists, decentralized admission control seems to be a natural solution, as it distributes the necessary work of prioritizing scans over all departments, rather than placing an extra burden on radiologists by shifting this responsibility to radiologists, which is the case with centralized admission control. This makes a centralized admission control model undesirable and the decentralized admission control model most suitable. A big disadvantage of decentralized admission control is the risk of rejecting medically necessary scans and on the other hand capacity waste because of the unflexible capacity. However, if there would be a mechanism that redistributes capacity during the year to account for these departmental fluctuations, this issue could be largely solved. Therefore, we consider a trading model where mutual exchange of scan budgets among departments is allowed.

2.4 Conclusions and research direction

So far, we discussed the rising demand, the appropriateness and priority of scans, the deficiencies of the current system and the necessity of a new system. We concluded that decentralized admission control in the form of a scan buget system is a better alternative than centralized admission control with regard to the workload problem of radiologists. A disadvantage of such a system is the possible unfair division of budget: some departments may waste part of their budget, while others might experience the problem of medically necessary scans that cannot be performed/requested because the department's budget is zero, due to stochasticity of the demand. This problem motivates our research. We study if mutual exchange of capacity ('trading') among departments can preserve equity between departments in terms of departmental acceptance ratios, i.e., the proportion of scans that can be performed for a department. This research includes several relevant topics.

The first topic is the individual policy of a department. The concept of rejecting and accepting scans is based on the fact that some requests have a higher priority than others. When the capacity is limited and not all scans can be performed, the ones with a higher priority should take precedence. To calculate the departmental acceptance ratios it is important to know how departments make decisions regarding the acceptance of type 1 and type 2 scans under a scan budget system. Moreover, there is also collaboration with other departments involved. To be able to evaluate this collaboration, the first step is to investigate if and under which circumstances individual departments are willing to collaborate. Therefore, one of our research topics is modeling the individual acceptance and trading policy of departments.

The second research topic is the analysis of collective behaviour. Analyzing the individual policies of individual departments is not enough to evaluate the collective behaviour, because a successful trade needs at least two departments that are willing to trade at the same time. Therefore, the collective behaviour must be modeled using a multi-agent model. A multi-agent model considers multiple stakeholders and connects the individual decisions to a collective policy, such that the collective state and reward can be observed. The stakeholders might have conflicting goals, because they are self-interested. However, in our case they are not fully competitive, as multiple players might benefit from collaboration. We investigate how we can model these dynamics between departments and the resulting acceptance ratios of the departments using a multi-agent model.

Chapter 3

Literature review

In this section, we present relevant literature related to our research aims. The problem we study includes many topics, including shared resources, capacity division and allocation, multi-priority models, decentralized admission control, resource constraints, multiple stakeholders, multi-agent systems with self-interested agents and sequential decision making. We start by providing an overview of literature about multi-priority models that are used for diagnostic resources. We then focus in a broader context on decision making with shared resources, capacity allocation and solution methods for these types of capacity problems. In the final subsection, we consider literature on multi-agent reinforcement learning, a framework for sequential decision making problems with multiple agents. We conclude by presenting our contribution, both in practice and our contribution to literature.

3.1 Operations research on radiology department

Multi-priority models for diagnostic resources have been extensively studied in operations research literature. Most studies in this area are performed from a scheduling perspective, aiming at minimization of waiting times for patients using a centralized approach. Operation research methods as dynamic programming and MDPs are often used to solve problems on radiology departments. In Green et al. [8], several diverse patient groups in a diagnostic facility are considered. The problem is formulated as a finite-horizon dynamic program with costs including a revenue per patient, a waiting cost if a service request is delayed and a penalty function associated with patients not served at the end of the day. In Patrick et al. [11], a similar problem is studied, considering patients with different priorities. Available capacity is allocated dynamically to incoming demand to achieve wait-time targets in a cost-effective manner. In Abtahi et al. [1], a scheduling problem of outpatients in a radiology center with an emphasis on priority is discussed. Stochastic programming and a mixed integer linear program (MILP) are used to minimize outpatients' total spent time in the radiology center. Benedito et al. [6] formulate an MDP to allocate capacity in a radiology department that serves different types of patients. The objective is to minimize the total cost of care, consisting of waiting, overtime, and penalty costs for not serving patients over a business day of service. Gao et al. [7] approach the problem of improving the utilization of diagnostic imaging resources from a game theoretic perspective. In their work, patient individual preferences are coordinated through automated negotiation, where patients and radiologists are represented by agents.

3.2 Shared capacity & capacity allocation

As we intend to focus on high-level and long-term capacity allocation and consequences rather than on detailed day-to-day scheduling, we focus in this section on the literature that is closely related to our study, including capacity allocation under resource constraints, maintaining equity, self-interested agents and policy evaluation. The study that is closest to our work is the work by Zonderland et al. [21]. Here the allocation of MRI scanning capacity among competing departments with incomplete information is studied from a game-theoretic perspective, aiming for fairness of the capacity allocation. In allocation mechanisms, various methods are used. Zonderland et al. argue that a suitable allocation mechanism should satisfy the following properties: 1. each agent should receive a nonnegative amount, 2. all capacity is allocated, and 3. each agent receives at most the amount that it requests. Literature on bankruptcy problems and fair division provides multiple allocation mechanisms that satisfy these conditions. In Thomson [14], various division rules are discussed, of which the most commonly used one is the proportional rule. This rules prescribes division of the total capacity proportional to the claims of the agents. Fairness of capacity allocation is also studied by Zhou et al. [19], [20]. In these works, inpatient room allocation with the objective of maximizing revenue under the constraints of maintaining equity is considered. To maintain fairness, predetermined acceptance and reject ratios are used to ensure equity among different patient groups.

In addition to capacity allocation, our research aims to study the long-term effects of a reward mechanism that stimulates mutual exchange of capacity. To study these effects, policies of departments must be evaluated. Besides well known sequential decision making frameworks that can be used to find policies such as MDPs, also game theory is a suitable candidate to perform policy analysis [9]. In Valkenburg [16], game theoretic methods are used to study budget allocation among competing defence departments. The model is set up as a dynamic non-cooperative budget allocation game with static and dynamic structure as well as with complete and incomplete information. Although the setting is different, the characteristics of this problem are similar to the characteristics of our problem. However, the situation is described as a 'zero-sum situation', in which the gain of one stakeholder is experienced as the loss of another, meaning that the game is fully competitive. Because in our case departments could, but do not necessarily have to benefit from trading, our situation is neither fully cooperative (where all departments share the same reward function) nor fully competitive. This is called a general-sum game. General-sum games that include sequential decision making are widely studied in the area of multi-agent reinforcement learning, which is discussed in the next section.

3.3 Multi-agent reinforcement learning

We consider a number of review articles that provide a clear overview of the terminology and methods used in reinforcement learning for multi-agent systems. The review by Busoniu et. al. [4] gives an overview of algorithms for fully cooperative, fully competitive and more general (neither competitive nor cooperative) tasks. It provides a useful introduction to terminology: the generalization of a single-agent Markov Decision Process (MDP) to multi-agent case is a *stochastic game* (SG). In fully competitive games, agents have opposing goals, which implies that the reward of the first player is the loss of the second player. In fully cooperative games, all agents share the same reward function. *Mixed games* or *general-sum games* are games that are neither fully competitive nor fully cooperative. In mixed games, each agent is self-interested, as in our case. Static games are stochastic games with no states and no transition function, so the reward of the players is fully determined by the joint action. A stage game is the static game that arises in a certain state of a stochastic game.

A challenge of MARL that is often mentioned in literature is the curse of dimensionality. The curse of dimensionality is more severe in MARL than in single-agent RL. Specifiying a good MARL goal in the general stochastic game is a difficult challenge, because the agents' returns are correlated and cannot be maximized independently. Furthermore, a significant number of algorithms for mixed games are designed only for static tasks (i.e., repeated, general-sum games). Even in repeated games, the learning problem is still nonstationary due to the dynamic behavior of the agents playing the repeated game. This is why most of the methods in this category focus on adaptation to the other agents. [4] This is also pointed out by a review paper by Zhang et. al. [18]. In this paper, it is discussed how a learning agent is required to account for how the other agents behave and adapt to the joint behavior accordingly, and how this invalidates the stationarity assumption for establishing the convergence of single-agent RL algorithms. Besides, the combinatorial nature of MARL is addressed: having a large number of agents complicates the theoretical analysis, especially the convergence analysis, of MARL. Zhang et. al. mention that this argument is substantiated by the fact that theories on MARL for the two-player zero-sum setting are much more extensive and advanced than those for general-sum settings with more than two agents. The authors extensively discuss the mixed-game, noting that the mixed setting is notoriously challenging and thus rather less well understood. Thus, it is concluded that additional structures on either the games or the algorithms need to be exploited, to ascertain provably convergent MARL in the mixed setting.

Yang et. al. [17] provide insights on the solutions of general-sum games. Game theory plays an essential role in multi-agent learning by offering so-called solution concepts that describe the outcomes of a game by showing which strategies will finally be adopted by players. Many types of solution concepts exist for MARL, among which the most famous is probably the Nash equilibrium (NE) in non-cooperative game theory (Nash, 1951). The word "non-cooperative" does not mean agents cannot collaborate or have to fight against each other all the time, it merely means that each agent maximises its own reward independently and that agents cannot group into coalitions to make collective decisions. Solutions to stochastic games are summarized by the master equation *normal-form game solver + MDP solver = stochastic game solver*, which was first summarized by Bowling et. al. [3]. The first term refers to solving an equilibrium (NE) for the stage game encountered at every time step, and the second term refers to applying a RL technique (such as Q-learning) to model the temporal structure in the sequential decision-making process.

Both [17] and [18] discuss several challenges of MARL, including the combinatorial complexity, the multi-dimensional learning objectives and the issue of non-stationarity. The scalability issue when there are more than two players is extensively discussed, as the majority of MARL algorithms is only capable of solving games with only two players, in particular, two-player zero-sum games [17]. Solving general-sum games entails an entirely different level of difficulty than solving team games (cooperative games) or zero-sum games. Even in the simplest case of a two-player general sum normal-form game, finding a Nash equilibrium is PPAD-complete [5]. Thus, designing learning algorithms in a multi agent system with more than two players is a challenging task. One major reason is that the solution concept, such as Nash equilibrium, is difficult to compute in general due to the curse of dimensionality of the multi-agent problem itself.

3.4 Our contribution

The problem that we study can be classified as a multi-agent general-sum stage game, which is very difficult to solve due to the combinatorial complexity of multi-agent problems. To avoid these curses of dimensionality, we restrict ourselves to two-player cases, for which we propose a model in which the collective problem is split into individual problems. The individual policies are calculated based on a *belief* on the policies of the other player. This belief is represented by a success probability for trading. In this way, the problems are relatively small and the policies can be calculated in parallel using a direct method (backward induction), so that we are assured of finding optimal policies for both players. The resulting collective behaviour follows from simulations, where individual departments act according to their optimal policy given a predetermined trading probability. We study the cost of differences between the predetermined probability and the actual success probability. For this purpose, we introduce the concept of a trading equilibrium and provide an algorithm that is able to find trading equilibria for the radiology capacity allocation problem that we study.

Our contribution is twofold. To the best of our knowledge, we are the first to study the constrained resource problem at a radiology department for multiple departments from a decentralized perspective. We contribute to a solution for the practical problem by proposing a new system of dynamic capacity budgeting and presenting a mathematical framework in which we look for a well-functioning scan budget budget system. Our model can be used to evaluate the costs and benefits if a scan budget system is introduced in practice. Hence, we provide a theoretical foundation for the practical effects on acceptance ratios of introducing a scan budget system at the radiology department. Next to the practical use of the model, our work contributes to literature on two-player multi-agent models by studying the problem from a decentralized perspective, using single-player Markov Decision Processes that can be easily solved with direct methods. In this way problems that arise in classical MARL settings, such as computational intractability and nonconvergence of solutions obtained with approximation methods, are avoided. Therefore, our model allows practical multi-agent problems to be solved very fast. This also enables the use of direct solutions methods that provide exact solutions. For the purpose of evaluating an external mechanism (such as a scan budget system) using policies, exact solutions are more convenient than solutions obtained with approximation methods. With solutions obtained by approximation methods, distinguishing differences in results caused by the external mechanism and differences in results caused by the approximation can be a challenging task. Using exact solutions, i.e., optimal policies, ensures that experiments are repeatable and yield the same results for the same parameter settings. Therefore, differences in results can be completely attributed to the external mechanism when exact solutions are used. Thus, our work provides an efficient and noise-free framework to evaluate external mechanisms.

Chapter 4

Methods

In this chapter, we present the methods that we use to answer our research questions. In Section 4.1, we start by describing the problem that we study in more detail. In Section 4.2, we describe the solution approach and give an overview of the different models that we use. In Section 4.3, we provide the mathematical formulation of the problem. In Section 4.4, the first ingredient of the model, the allocation rule, is discussed. The second ingredient, the individual Markov Decision Process for a department, is described in Section 4.5. In Section 4.6, the trading probabilities are studied. We define definitions and provide algorithms to find trading equilibria. Finally, the analysis of the collective case is discussed in Section 4.7.

4.1 Problem description

We consider a scan budget system as a solution to the workload problem at the radiology department. The largest disadvantage of such a system is the inefficient use of capacity, due to inflexible individual scan budgets that cannot adapt to fluctuations of the demand. This can result in large departmental differences in the acceptance ratios of scans. Our main research question therefore is

"What are the effects of a scan budget system with trading on the type 1 and type 2 acceptance ratios?"

To answer this question for a our two-department setting, we need to be able to calculate acceptance ratios for a scan budget system with a given allocation mechanism. The proportion of scans that is accepted by a department is given by the department's *acceptance ratio*. The definition of the department i's annual acceptance ratio for type j is as follows:

$$A_i^j = \frac{\#\text{accepted scan candidates type } j \text{ department } i}{\#\text{ arrived scan candidates type } j \text{ department } i}$$

The relevance of these acceptance ratios is illustrated using Figure 4.1. In this figure, the scan budgets C_i and the demands D_i of two departments are depicted.

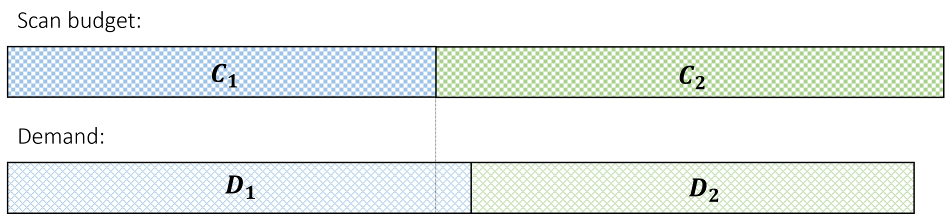

Figure 4.1: Scan budgets and demands of two departments.

The demand of department 1, D_1, exceeds the scan budget of department 1, C_1. Therefore, part of the scan candidates that arrive at department 1 must be rejected. On the other hand, department 2 has not used its full scan budget at the end of the year and the remaining scan budget of department 2 would be enough for player 1's excess. In other words, the total scan budget is large enough for both department 1 and department 2's demand, but part of department 1's demand must be rejected because the individual scan budget is not sufficient. This phenomenon is referred to as the *the cost of inflexibility*. Under centralized admission control, this would not happen, because with centralized admission control the capacity is shared over all departments and therefore very flexible. As a solution to the cost of inflexibility, we explore the possibility of trading among departments, as trading (mutual exchange of capacity) could increase the acceptance ratios. As departments are self-interested, a central mechanism is necessary to stimulate trading.

We consider a decentralized admission control setting with scan budgets and a central allocation mechanism for two departments in which the departments are modeled as self-interested agents, which means that they may behave strategically in order to maximize their benefits. Both departments are allocated a scan budget that can be used to request scans. We consider two types of scan requests: medically necessary scans according to protocol, and additional scans that are not according to a protocol. We refer to the first type as *type 1* scans and to the second as *type 2* scans. We assume that a type 1 scan has a higher priority than a type 2 scan to both departments. We make a distinction between the number of scan requests that arrive at a department, hereafter referred to as *candidates*, and the number of scans that are actually requested from the radiology department by departments. In this setting, we analyze an allocation mechanism by calculating the acceptance ratios for both departments. To find these acceptance ratios, we need to model both the individual behaviour of departments and the collective behaviour of departments under the scan budget system with trading.

4.2 Solution approach

To find the acceptance ratios under a central allocation mechanism, we need to include several factors in our model. We consider the two main research topics that we discussed in Section 2.4, individual acceptance and trading policy and the collective behaviour. We also consider multiple priorities and the capacity allocation mechanism. In this section we explain how these ingredients are modeled and we discuss the relation between the topics.

To account for multiple priorities, we consider two types of scans: type 1 and type 2. Type 1 scans have priority over type 2 scans, as type 1 scans are medically necessary. In general, a department's scan budget should be sufficient to accept all type 1 scan candidates, but it might not be sufficient for all type 1 and type 2 scan candidates. Thus, when a scan budget system is introduced, a department might not always want to accept type 2 scan candidates, because this could lead to forced rejection of type 1 scan candidates. Therefore, departments will design a policy that prescribes under which circumstances type 2 scans are accepted.

Next to the decision whether to accept a type 2 scan candidate, there is also the possibility to trade budget with the other department. Therefore, a department's policy also includes under which circumstances a department will buy or sell scan budget. We model and analyze these individual policies using a Markov Decision Process (MDP), where we assume that departments are self-interested and want to maximize their own rewards. They obtain a reward for accepting type 2 scans and they receive a penalty for every type 1 scan they reject. The solution to this MDP is a policy that prescribes under which circumstances a department should accept a type 2 scan and under which circumstances a department should trade to maximize the department's total reward.

The capacity allocation mechanism describes the process of allocating scan budgets to departments. The radiology department is responsible for this mechanism. The capacity allocation mechanism can be used to stimulate trading. Although the words suggest otherwise, there is neither money nor other means of payment involved in the trading process. 'Trading' is a possible solution to the problem of fluctuations in the demand and the inflexibility of the individual scan budgets. The idea is that departments can resolve these fluctuations among themselves by exchanging capacity: a department that has more capacity left during that year can give it away to a department that needs it. We assume that each department's individual goal is to maximize its own number of scans, first the type 1 scans and then the type 2 scans. For this reason, helping other departments by giving away capacity does not seem obvious. Therefore, with the annual capacity allocation process, radiology takes into account not only the number of type 1 and 2 scans requested, but also the capacity a department has given away and received from other departments. We express this as a balance. This means that departments can be rewarded or penalized by means of increasing or decreasing next year's scan budget. The purpose of this is to give departments an incentive to help other departments, if necessary, to increase the overall acceptance ratios and thus decrease the cost of inflexibility. That brings us back to the individual policy. A department's policy should not only take the current year into account, but also next year, because the actions of department during the current year influence the scan budget that is allocated to the department next year. That means that the individual MDPs are influenced by the allocation mechanism.

In this research, we study the cost of inflexibility for different allocation mechanisms. We assume that departments behave rationally, so we model the behavior of departments using Markov Decision Processes (MDPs) and evaluate the policies to see how the allocation rule affects the number of scans that is performed. To analyze the behavior of departments under the allocation rule, we model the situation for each individual department as an MDP, solve this MDP and assume that departments act according to the optimal policy. However, as departments are allowed to trade capacity, the individual MDPs are not independent, but have some interaction with each other. This is modeled using trading probabilities that are part of the transition probabilities in the MDP. How the trading probabilities affect the optimal policy and how the policy of one department influences the trading probabilities of other departments, is studied extensively in this chapter. An overview of the models is given in the following figure.

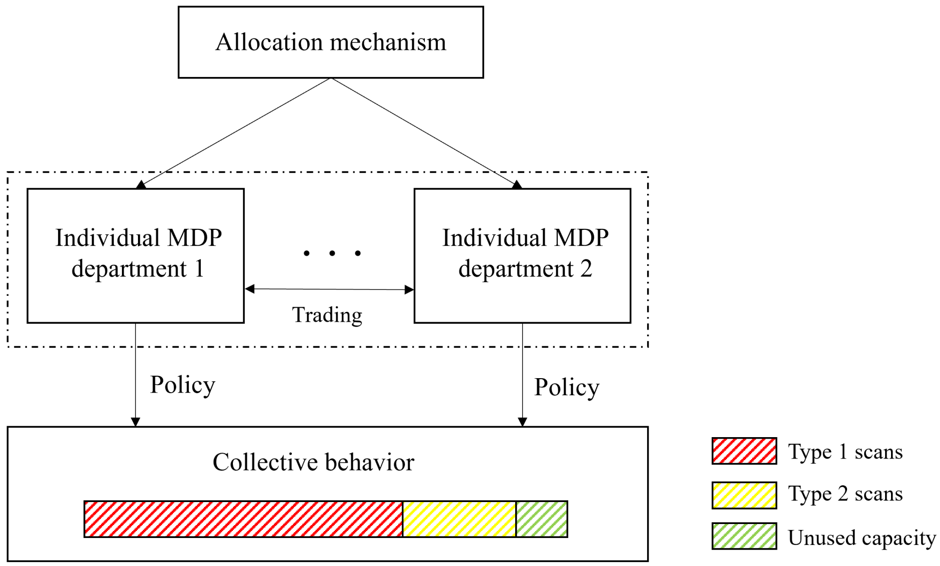

Figure 4.2: Model overview

The starting point is an allocation rule, described in Section 4.4. The allocation rule gives rise to two MDPs, modeling the situations the two departments of the hospital. These individual MDPs are described in detail in Section 4.5. In Section 4.6, we discuss the trading probabilities and describe algorithms to find trading equilibria. After solving the MDPs, we obtain two policies. Using these policies, we evaluate the allocation ratios. This process is described in Section 4.7.

4.3 Model formulation

In this section, we present the model that we use to answer our research question. Given an allocation rule h over a period of H years, we evaluate the set of acceptance ratios

$$\mathcal{A}(h) = \{A_i^1, A_i^2 : i = 1, 2\}$$

that follow from the solution to the set of MDPs

$$\mathcal{M}(h) = \{\mathcal{M}(c_0^{i,y}) = \{\mathcal{S}_i, \mathcal{A}_i, q_i, \mathcal{R}_i\} : i = 1, 2, y = 1, \ldots, H\},$$

where $\mathcal{M}(c_0^{i,y})$ is the Markov decision process that describes the decision making process of department i in year y with scan budget $c_0^{i,y}$. This process includes deciding on the acceptance of scans and trading. The scan budget $c_0^{i,y}$ is given by the allocation rule h. The allocation rule is defined in Chapter 4.4 and the MDPs are defined in Section 4.5. Before we explicitly describe the allocation rule and the MDPs, we start in this subsection with defining the setting that we model.

General setting

We investigate the consequences of a scan budget system on the acceptance ratios on a radiology department in a tertiary care center with two departments over multiple years. Under a scan budget system, every year each department ($i = 1, 2$) gets assigned a budget thay can be used to request that year, called their *scan budget*. Here, no distinction is made between different scans, for example MRI and CT, which means that we assume that there is one capacity for all type of scans, regardless of their workload and level of complexity. The cost of a scan is 1.

We model the situation over $y = 1, \dots, H$ years. The scan budget of department i in year y is denoted by $c_0^{i,y}$. A department's *remaining budget* is their scan budget minus the number of scans that have been requested by the department during the year. Hence, for every scan that is requested by a department, the department's remaining budget decreases by one. When a department's remaining budget is zero, the department cannot request any more scans from the radiology department during the current year, unless they would buy additional budget.

Trading

In our model we allow mutual exchange of capacity. The term *trading* is used to describe a department receiving (*buy*) capacity from and donating (*sell*) capacity to another department. There is no money involved in this trading process. In this model, departments can only trade during so-called *trading epochs* and the amount of capacity that can be traded is fixed and denoted by ζ. *Buying* implies that a department places an order for ζ units of capacity and *selling* means that a department offers ζ units of capacity. Departments are not obliged to sell or buy during a trading epoch, they can also decide to do nothing.

Scan types and demand

The number of type 1 scan candidates that arrive at department i during the year is a random variable D_i. Besides that, departments also request type 2 scans. The number of this type of scan candidates that arrive at department i is denoted by E_i. We assume that D_i and E_i are Poisson distributed.

Allocation rule

The scan budget that is allocated to department i at the beginning of year y is denoted by $c_0^{i,y}$. At the beginning of the time horizon, all departments receive an initial scan budget of

$$c_0^{i,1} = (1 + b)\,\mathbb{E}[D_i],$$

where $b \geq 0$ is the *buffer constant*. A department can either use its scan budget to request its own scan candidates, or they can trade it with other departments. After the first year, the allocation of capacity is not only based on the expected demand, but also on the department's behaviour. This is done according to an *allocation rule* that is clear to all departments, such that departments can take the consequences of their actions on their capacity for next year into account. The allocation rule is a function $h : \mathbb{Z} \to \mathbb{N}$ that maps characteristics of the behavior of a department during the year to a number. This number is the new scan budget that is allocated to the department at the start of the next year.

4.4 Allocation rule

In this section, we discuss the capacity allocation mechanism. This mechanism is characterized by the *allocation rule*, a function that determines the scan budget that a department receives at the beginning of a new year. The allocated scan budgets depend on the total radiology budget. The total radiology budget is denoted by B. It is given by

$$B = (1 + b) \sum_i \mathbb{E}[D_i].$$

The parameter $b \geq 0$ is a buffer. In the choice of b, the variance of $\sum_i \mathbb{E}[D_i]$ is taken into account. The underlying idea of this is that the total budget is always sufficient to serve all type 1 scans. In the allocation process in the first year, we follow the proportional division rule by Thomson [14]: $c_i = \lambda\,\mathbb{E}[D_i]$, where λ is chosen so that $\sum_i \lambda c_i = B$. Thus in the first year, department i receives an amount proportional to their expected type 1 demand. Therefore the initial scan budget of department i in year 1 is given by $c_0^i = (1+b) \cdot \mathbb{E}[D_i]$. For next years, the allocation rule h is followed, which is an adaptation to the proportional rule. However, it always satisfies the property that department i receives a minimum scan budget of $\mathbb{E}[D_i]$.

Next to providing enough budget such that the expected type 1 scans demand can be met, another goal of the allocation rule is to stimulate departments to sell capacity to departments that need it for their type 1 scans. There is a natural incentive for departments to place orders and buy capacity, as there is no money involved so buying capacity is actually free and having more capacity leads to a higher number of scans that can be requested, and thus to a higher direct reward. However, the process of ordering and offering only works if there is also an incentive for departments to sell their capacity. For that reason, we design an allocation rule that is based on a department's *trading balance*: the total capacity that has been sold minus the capacity that has been bought by the department. In this way, departments can be rewarded or penalized for their trading behaviour.

The allocation rule that is used in year $y = 2, \ldots, H$ is the following. If department i's trading balance is positive, it receives an amount of

$$c_i = (1+rb)\,\mathbb{E}[D_i]$$

the next year, where we have $r \geq 1$. Notice that a positive trading balance implies that a department has sold more capacity than it has bought. With two players, this means that the other department's trading balance must be negative. To ensure that the total scan budget is equal to $\sum_i (1+b)\,\mathbb{E}[D_i]$, department j (which has a negative trading balance) receives an amount of

$$C_j = (1+b)\,\mathbb{E}[D_j] + b(1-r)\,\mathbb{E}[D_i].$$

We require

$$1 \leq r \leq \frac{\min_i D_i}{\max_i D_i} + 1 \tag{4.1}$$

to ensure that each department receives at least $\mathbb{E}[D_i]$ units of capacity. Summarizing, the allocation rule is given by

$$h_r(f_i) = \begin{cases} (1+rb)\,\mathbb{E}[D_i] & \text{if } f_i > 0, \\ (1+b)\,\mathbb{E}[D_i] + b(1-r)\,\mathbb{E}[D_j] & \text{if } f_i < 0, \\ (1+b)\,\mathbb{E}[D_i] & \text{else.} \end{cases}$$

4.5 Individual department MDP

We model the situation of a single department during one year as a Markov Decision Process (MDP), where the states capture the available capacity and, during regular epochs, the presence of a type 2 scan candidate for the department. During regular epochs, the decision must be made whether or not to accept a scan if there is a scan candidate of type 2. There is no decision for type 1 scans, because we assume that these are always accepted, unless the remaining budget of the department is zero. The presence of type 1 scan candidates is incorporated in the model using the transition probabilities. It is not captured by the state. However, there is a counter in the state that keeps track of the number of type 1 scans that is declined during the year. During trading epochs, the department can decide to buy or sell capacity or to do nothing. A department aims to maximize its total reward, that is, maximize the total number of scans that is requested. However, when a type 1 scan is declined, the department receives a negative reward (penalty) at the end of the year. Thus, in some cases, it might be beneficial to decline type 2 scans to ensure that there is enough capacity left for type 1 candidates in the future. For every type 2 scan that is accepted, the department receives a reward. The final reward depends on the capacity that is allocated to the department for next year and the number of type 1 scans that was declined during the year. The transition probabilities are determined by the demand, that is, by the distribution of D_i and E_i and the behavior of other departments (trading probabilities). In this section, we describe the Markov Decision Process in detail.

4.5.1 Time horizon and decision epochs

The time horizon of an individual MDP is one year. In our model, we consider two types of decision epochs. The first one is the *regular epoch*. In these epochs, departments decide if they request scan candidates if they are available. The set of all regular decision epochs is denoted by $\mathcal{T}^R$. The second type is the *trading epoch*, denoted by $\mathcal{T}^T$. During these epochs, departments can decide to buy or sell capacity. We assume that scan candidates do not arrive during trading epochs, so the only decision during trading epochs is buy/sell/do nothing.

The majority of decision epochs are regular epochs. We define the number of regular epochs between trading epochs including one trading epoch as δ. The number of trading epochs each year is denoted by σ and the time horizon is given by $T = \delta(\sigma + 1) - 1$. We define the set of trading epochs and allocation epochs as follows:

$$\mathcal{T}^T = \{\delta k : k = 1, \dots, \sigma\}.$$

The set that is formed by the remaining decision epochs, the regular decision epochs, is

$$\mathcal{T}^R = \{t : t = 1, \dots, T, t \notin \mathcal{T}^T\}.$$

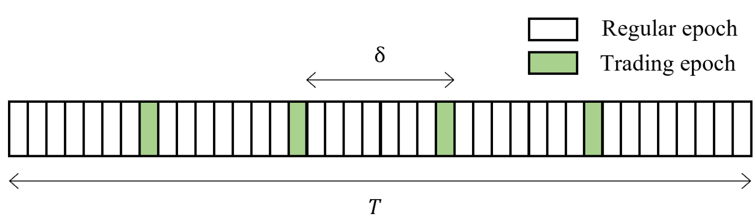

Figure 4.3: Representation of the time horizon T and decision epochs.

4.5.2 States

For $t \in \mathcal{T}^R, t \neq T$, a typical state takes the form

$$s_t = (c, d, f, g),$$

where c is the remaining budget for the current year for the department and $d \in \{0, 1\}$ is a variable that indicates the presence of a type 2 scan candidate. If $d = 0$, there are no candidates and if $d = 1$, there is a scan candidate of type 2. The trading balance is denoted by f, where we define the trading balance as follows:

$$f = \text{sold budget - bought budget}.$$

Finally, g represents the number of type 1 scan candidates that have been rejected during the year.

In the trading decision epochs $t \in \mathcal{T}^T$ and in the final state $t = T$, we have $d = 0$, as we assume that no scan candidates arrive during trading epochs:

$$s_t = (c, 0, f, g) \quad \text{for } t \in \mathcal{T}^T \cup T.$$

4.5.3 Actions

In the regular decision epochs $t \in \mathcal{T}^R$, an action takes the simple form

$$a_t = a, \quad a \in \{0, 1\}$$

where a is the decision for a waiting scan. If $a = 1$, the scan is accepted. Otherwise, the scan is declined. Evidently, a_t should satisfy $a_t \leq d$ (department can only accept when there is a scan candidate and $a_t \leq c$ (departments can only accept a scan when their remaining budget is positive).

During trading epochs, $t \in \mathcal{T}^T$, the action space is given by

$$\mathcal{A}_t = \{-1, 0, 1\},$$

where -1 stands for buying ζ units of capacity, 0 stands for doing nothing and 1 stands for selling ζ units of capacity.

4.5.4 Rewards

The reward of state-action pair (s, a) is given by

$$r(s_t, a) = \begin{cases} a_t, & \text{if } t \in \mathcal{T}^R, t < T \\ 0 & \text{if } t \in \mathcal{T}^T, \\ V_i^{y+1}(h(s_T)) - \alpha g_T & \text{if } t = T, y < H. \\ -\alpha g_T & \text{if } t = T, y = H. \end{cases} \tag{4.2}$$

In regular decision epochs, the reward of accepting a scan ($a = 1$) is 1, and the reward of rejecting a scan or doing nothing in case there is no scan candidate ($a = 0$) is 0. In trading epochs, the reward is always 0, regardless of the action. The final reward in year $y < H$ is $V_i^{y+1}(h(s_T))$, the value of the starting state of the MDP $\mathcal{M}(c_0^{i,y+1})$, the MDP for department i for next year with scan budget $c_0^{i,y+1} = h(s_T)$. This value is calculated according to Equation 4.3. It is subtracted by αg_T, the penalty for declining g_T type 1 scans. Finally, the final reward in the final year H is determined by the penalty for the number of type 1 scans that were rejected during year H.

4.5.5 Transition probabilities

Let the probability that a type i scan candidate arrives be denoted by π_i. We assume that at most one scan candidate can arrive per decision epoch. Then the transition probabilities $w(s_{t+1}|s_t, a_t)$ for $t, t+1 \in \mathcal{T}^R$ are given by

$$w(s_{t+1}|s_t, a_t) = \begin{cases} \pi_1(1-\pi_2) & \text{if } d_t = 0, c_{t+1} = c_t - \mathbb{1}_{c_t>0}, d_{t+1} = 0, g_{t+1} = g_t + \mathbb{1}_{c_t=0}, \\ \pi_1\pi_2 & \text{if } d_t = 0, c_{t+1} = c_t - \mathbb{1}_{c_t>0}, d_{t+1} = 1, g_{t+1} = g_t + \mathbb{1}_{c_t=0}, \\ (1-\pi_1)(1-\pi_2) & \text{if } d_t = 0, c_{t+1} = c_t, d_{t+1} = 0, \\ (1-\pi_1)\pi_2 & \text{if } d_t = 0, c_{t+1} = c_t, d_{t+1} = 0, \\ 1-\pi_2 & \text{if } d_t = 1, c_{t+1} = c_t - a_t, d_{t+1} = 0, \\ \pi_2 & \text{if } d_t = 1, c_{t+1} = c_t - a_t, d_{t+1} = 1. \end{cases}$$

These probabilities can be explained as follows. In the first line $d_t = 0$, which means that there is no type 2 scan candidate at time t. There is a type 1 scan candidate, which is only possible if $d_t = 0$ because at most one candidate can arrive during a decision epoch. The type 1 scan is either accepted (in case $c_t > 0$) or rejected (in case $c_t = 0$). In the first case, the remaining budget c_t decreases by 1. In the latter case, the reject counter g_t increases by 1. There is no type 2 request at $t + 1$, which occurs with probability $(1 - \pi_2)$. The probability of a type 1 arrival is π_1, so the total probability of this case is $\pi_1(1-\pi_2)$. The second line is almost identical, except for the value of d_{t+1}. In the second line $d_{t+1} = 1$ so this line represents the case where a type 1 scan request arrives at time t and a type 2 scan requests arrives at time $t+1$. This happens with probability $\pi_1\pi_2$. The third line represents the case when there is no arrival at time t and no type 2 arrival at time t, whereas the fourth line represents the case when there is no arrival at time t and a type 2 arrival at time t. The fifth and sixth line represent the cases when there is a type 2 scan at time t. If this scan is accepted, $a_t = 1$ and $c_{t+1} = c_t - 1$. If this scan is rejected, $a_t = 0$ and $c_{t+1} = c_t$. With probability π_2, a type 2 scan arrives at time $t+1$ (sixth case), and with probability $1 - \pi_2$, no type 2 scan arrives at time $t+1$ (fifth case).

For $t \in \mathfrak{T}^R, t+1 \in \mathfrak{T}^T$, we have

$$w(s_{t+1}|s_t, a_t) = \begin{cases} \pi_1 & \text{if } d_t = 0, c_{t+1} = c_t - \mathbb{1}_{c_t>0}, g_{t+1} = g_t + \mathbb{1}_{c_t=0}, \\ (1-\pi_1) & \text{if } d_t = 0, c_{t+1} = c_t, \\ 1 & \text{if } d_t = 1, c_{t+1} = c_t - a_t. \end{cases}$$

In the first two lines, there is no type 2 scan candidate at time t. No scan candidates arrive during a trading epoch. The first line represents the case when there is a type 1 scan at time t. This happens with probability π_1. With probability $1 - \pi_1$, there is no type 1 scan. This is represented by the second line. The third line represents the presence of a type 2 scan candidate at time t. In this case, the scan is either accepted or rejected and the transition is deterministic, because the probability of a scan arrival during a trading epoch is zero.

For $t \in \mathfrak{T}^T, t+1 \in \mathfrak{T}^R$, we have

$$w(s_{t+1}|s_t, a_t) = \begin{cases} (1-\pi_2)p_i(t) & \text{if } a_t = -1, c_{t+1} = c_t + \zeta, d_{t+1} = 0, f_{t+1} = f_t - 1, \\ \pi_2 p_i(t) & \text{if } a_t = -1, c_{t+1} = c_t + \zeta, d_{t+1} = 1, f_{t+1} = f_t - 1, \\ (1-\pi_2)(1-p_i(t)) & \text{if } a_t = -1, c_{t+1} = c_t, d_{t+1} = 0, f_{t+1} = f_t, \\ \pi_2(1-p_i(t)) & \text{if } a_t = -1, c_{t+1} = c_t, d_{t+1} = 1, f_{t+1} = f_t, \\ (1-\pi_2)q_i(t) & \text{if } a_t = 1, c_{t+1} = c_t - \zeta, d_{t+1} = 0, f_{t+1} = f_t + 1, \\ \pi_2 q_i(t) & \text{if } a_t = 1, c_{t+1} = c_t - \zeta, d_{t+1} = 1, f_{t+1} = f_t + 1, \\ (1-\pi_2)(1-q_i(t)) & \text{if } a_t = 1, c_{t+1} = c_t, d_{t+1} = 0, f_{t+1} = f_t, \\ \pi_2(1-q_i(t)) & \text{if } a_t = 1, c_{t+1} = c_t, d_{t+1} = 1, f_{t+1} = f_t, \\ (1-\pi_2) & \text{if } a_t = 0, c_{t+1} = c_t, d_{t+1} = 0, f_{t+1} = f_t, \\ \pi_2 & \text{if } a_t = 0, c_{t+1} = c_t, d_{t+1} = 1, f_{t+1} = f_t. \end{cases}$$

In the first four lines, the action at time t is buy. With probability $p_i(t)$, the trade succeeds. In that case, the remaining budget is increased by the trading amount ζ and the trading balance decreases by 1. With probability $1 - p_i(t)$, the trade does not succeed. In this case, the remaining budget and the trading balance do not change. With probability π_2, a type 2 scan request arrives at time $t+1$. In lines 5-8, the action at time t is sell. With probability $q_i(t)$, the trade succeeds. In that case, the remaining budget decreases by the trading amount ζ and the trading balance increases by 1. With probability $1 - q_i(t)$, the trade does not succeed. In this case, the remaining budget and the trading balance do not change. Again, a type 2 scan request arrives at time $t+1$ with probability π_2. In the final two lines, the action during the trading epoch is do nothing. In that case, the probability is determined by the arrival probability of a type 2 candidate π_2.

4.5.6 Optimality equations

The optimality equations are given by

$$u_t^*(s_t) = \max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{s_{t+1} \in S} w(s_{t+1}, a) u_{t+1}^*(s_{t+1}) \right\} \tag{4.3}$$

for $t = 1, \dots, T-1$. For $t = T$, we add the boundary condition

$$u_T^*(s_T) = r_T(s_T).$$

The maximum action in state s_t is given by

$$A_{s_t,t}^* = \arg\max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{j \in S} w(j|s_t, a) u_{t+1}^*(j) \right\}.$$

Any policy that selects an action in $A_{s_t,t}^*$ for each $s_t \in S$ for all $t = 1, \dots, T$ is optimal.

4.6 Trading probabilities

In this section, we consider the trading probabilities. The trading probabilities connect the individual policy to the policy of the other department. The trading probabilities that are used in an individual MDP are the estimated success probabilities of trading. The actual success probabilities for a player depend on the policy of the other player and might not always match the estimated probabilities that generated the policy for the player. In this section, we introduce theoretical concepts that help us understand the effects of the estimated trading probabilities, the true trading probabilities and the difference between these probabilities on the optimal policies and the acceptance ratios. In Section 4.6.1, we define estimated and true trading probabilities. Moreover, we relate the optimal policies obtained by (4.3) to the concept of a best response and we introduce the definition of a trading equilibrium. In Section 4.6.2, we present the Subdomain backward induction algorithm, that gives an overview of the influence of the esimated trading probabilities on the optimal policy. Finally, in Section 4.6.3, we present an algorithm that finds all trading equilibria.

4.6.1 Estimated and true trading probabilities

In the MDP that is used to describe the decision process of department i, the probability that department i succeeds in buying or selling ζ units of capacity during a trading epoch t is denoted by $p_i(t)$ and $q_i(t)$, respectively. These variables are the *estimated trading probabilities* and connect the strategy of the individual player to the strategy of the other player. The solution of the MDP with trading probabilities $p_i(t)$ and $q_i(t)$ can be interpreted as the optimal policy under the belief of player i that the success probabilities of buying and selling are $p_i(t)$ and $q_i(t)$. This policy is denoted by $\Pi^*(p_i(t), q_i(t))$.

The success probabilities for player i depend on the behaviour of player j and vice versa. Thus, $p_i(t)$ and $q_i(t)$ are a reflection of the belief of player i on the strategy of player j. These do not have to match the *true trading probabilities*. We denote the true trading probabilities for player i by $\tilde{p}_i(t)$ and $\tilde{q}_i(t)$. These can be defined as follows. If player i wants to buy capacity at time t, the probability that this trade succeeds, $\tilde{p}_i(t)$, is equal to the probability that player j is willing to sell capacity at time t. So we define the true trading probabilities for player i in the following way:

$$\tilde{p}_i(t) = \text{probability that player } j \text{ is in a state at time } t \text{ where he is willing to sell} \tag{4.4}$$

and similarly

$$\tilde{q}_i(t) = \text{probability that player } j \text{ is in a state at time } t \text{ where he is willing to buy.} \tag{4.5}$$

These probabilities follow from the policy $\Pi^*(p_j(t), q_j(t))$, the policy that is optimal when the estimated trading probabilities for player j are $p_j(t)$ and $q_j(t)$. Let $S_j^s \subseteq S_t$ be the set of states where the optimal action for player j is 'sell', and let $S_j^b \subseteq S_t$ be the set of states where the optimal action for player j is 'buy'. Notice that these sets depend on the policy that player j is using, which again depends on the estimated trading probabilities of player j, $p_j(t)$ and $q_j(t)$. Therefore, we let $\Pi^*(p_j(t), q_j(t))$ denote the optimal policy for player j when the estimated trading probabilities are $p_j(t)$ and $q_j(t)$. Then (4.4) and (4.5) can be written as follows:

$$\tilde{p_i}(t) = \sum_{s \in S_j^s} \mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s | s_0) \tag{4.6}$$

and

$$\tilde{q_i}(t) = \sum_{s \in S_j^b} \mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s | s_0), \tag{4.7}$$

where $\mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s | s_0)$ is the probability that the state of player j at time t under policy $\Pi^*(p_j(t), q_j(t))$ is s_j, given that the starting state is s_0.

With these definitions, we can introduce the definition of a *best response*.

Definition 4.6.1 (Best response)**.** Policy Π_i is a best response to policy Π_j with

$$\sum_{s \in S_j^s} \mathbb{P}^{\Pi_j}(s_j = s | s_0) = \tilde{p_i}(t)$$

$$\sum_{s \in S_j^b} \mathbb{P}^{\Pi_j}(s_j = s | s_0) = \tilde{q_i}(t)$$

if Π_i satisfies

$$\Pi_i = \Pi^*(\tilde{p_i}(t), \tilde{q_i}(t)).$$

To be able to model the strategies of the departments correctly, we need to understand the dynamics behind the estimated trading probabilities and the true trading probabilities. In this chapter, we focus on the question whether it is possible for all players to estimate the trading probabilities correctly. That is, we investigate if there exist estimated trading probabilities $p_i(t) = \tilde{p_i}(t), q_i(t) = \tilde{q_i}(t)$ for $i = 1, 2$. To answer this question, we introduce the concept of an equilibrium of trading probabilities.

Definition 4.6.2 (Trading equilibrium)**.** A 4-tuple of trading probabilities $(p_i(t), q_i(t), p_j(t), q_j(t))$ is a *trading equilibrium* if it satisfies the following equations:

$$p_i(t) = \sum_{s \in S_j^s} \mathbb{P}^{\Pi_j^*(p_j(t), q_j(t))}(s_j = s | s_0) \quad \text{for } t \in \mathcal{T}^T, \tag{4.8}$$

$$q_i(t) = \sum_{s \in S_j^b} \mathbb{P}^{\Pi_j^*(p_j(t), q_j(t))}(s_j = s | s_0) \quad \text{for } t \in \mathcal{T}^T, \tag{4.9}$$

$$p_j(t) = \sum_{s \in S_i^s} \mathbb{P}^{\Pi_i^*(p_i(t), q_i(t))}(s_i = s | s_0) \quad \text{for } t \in \mathcal{T}^T, \tag{4.10}$$

$$q_j(t) = \sum_{s \in S_i^b} \mathbb{P}^{\Pi_i^*(p_i(t), q_i(t))}(s_i = s | s_0) \quad \text{for } t \in \mathcal{T}^T, \tag{4.11}$$

$$0 \leq p_i(t), p_j(t), \leq 1 \quad \text{for all } t \in \mathcal{T}^T. \tag{4.12}$$

That is, if the estimated trading probabilities of players i and j equal the true trading probabilities of the players, respectively.

If the true trading probabilities for player i that follow from the optimal strategy Π_j^* of player j induced by the equilibrium probabilities are used as estimated trading probabilities to calculate the optimal policy Π_i^* for player i, the true trading probabilities for player j that follow from this policy yield the strategy Π_j^*. Moreover, if both players use the strategies induced by estimated probabilities equal to the equilibrium probabilities $p_i(t), q_i(t), p_j(t)$ and $q_j(t)$, the true trading probabilities are given by $p_i(t), q_i(t), p_j(t)$ and $q_j(t)$. These properties are summarized in the following theorem.

Theorem 4.6.1. *Let $\Pi_i^* = \Pi^*(p_i(t), q_i(t))$ and $\Pi_j^* = \Pi^*(p_j(t), q_j(t))$ denote the optimal policies of player i and j induced by the estimated trading probabilities $p_i(t), q_i(t)$ and $p_j(t), q_j(t)$, respectively. If $(p_i(t), q_i(t), p_j(t), q_j(t))$ is a trading equilibrium, these policies satisfy the following properties:*

(i) The true trading probabilities when player i and player j act according to Π_i^* and Π_j^*, respectively, are given by $p_i(t), q_i(t), p_j(t)$ and $q_j(t)$.

(ii) Policy Π_i^* is the best response to policy Π_j^* and vice versa.

Proof. (i) The true trading probabilities are given by equations (4.6.1) and (4.7). By the definition of a trading equilibrium, we have

$$\tilde{p_i}(t) = \sum_{s \in S_j^s} \mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s|s_0) = p_i(t),$$

$$\tilde{q_i}(t) = \sum_{s \in S_j^b} \mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s|s_0) = q_i(t),$$

$$\tilde{p_j}(t) = \sum_{s \in S_i^s} \mathbb{P}^{\Pi^*(p_i(t), q_i(t))}(s_j = s|s_0) = p_j(t),$$

$$\tilde{q_j}(t) = \sum_{s \in S_j^b} \mathbb{P}^{\Pi^*(p_i(t), q_i(t))}(s_j = s|s_0) = q_j(t).$$

(ii) If player j plays policy Π_j^*, by (i) the true trading probabilities for player i under this policy are

$$\tilde{p_i}(t) = \sum_{s \in S_j^s} \mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s|s_0) = p_i(t),$$

$$\tilde{q_i}(t) = \sum_{s \in S_j^b} \mathbb{P}^{\Pi^*(p_j(t), q_j(t))}(s_j = s|s_0) = q_i(t).$$

The optimal policy for player i, given that the true probabilities are given by $p_i(t)$ and $q_i(t)$, is

$$\Pi^*(p_i(t), q_i(t)) = \Pi_i^*,$$

which is a best response by definition. The converse also holds by symmetry.

□

It is unclear whether a trading equilibrium always exists. In the next sections, we provide an algorithm that identifies policies $\Pi^*(p(t), q(t))$ for all $0 \leq p(t), q(t) \leq 0$ and an algorithm that finds all trading equilibria, in case they exist.

4.6.2 Subdomain backward induction algorithm

From the system in equations (4.8)-(4.12) we can see that the trading probabilities and the optimal policy influence each other, and hence solving the system is not trivial. In this section, we provide an algorithm that calculates all best response policies $\Pi^*(p(\tilde{t}), q(\tilde{t}))$ for $0 \leq p(\tilde{t}), q(\tilde{t}) \leq 1$ for MDPs with one trading epoch, denoted by $\tilde{t}$. In this algorithm maxima are obtained in (4.3) so that we are assured of obtaining an optimal Markovian deterministic policy for one player as a function of the trading probabilities. We follow the principle of the Backward Induction Algorithm from the book by Puterman [12] for finite-horizon MDPs. We adapt this algorithm such that it calculates a department's optimal policy as a function of the department's estimated trading probabilities, $0 \leq p, q \leq 1$. It is based on the fact that the value functions $u_t^*(s_t)$ are functions of p and q and can be discontinuous:

$$u_t^*(s_t) = \begin{cases} f_1(p,q) & \text{if } p, q \in D_{s_t}^1 \\ \vdots & \\ f_k(p,q) & \text{if } p, q \in D_{s_t}^k \end{cases}$$

We have $D_{s_t}^1 \cup \cdots \cup D_{s_t}^k = D^{\mathbb{1}}$, where $D^{\mathbb{1}}$ is the region defined by $0 \leq p, q \leq 1$. For $s_t \in S$ we denote the collection of these sets by D_{s_t}:

$$D_{s_t} = \{D_{s_t}^1, \ldots, D_{s_t}^k\}. \tag{4.13}$$

To introduce the algorithm we first consider the backward induction algorithm for finite-horizon MDPs by Puterman [12]. Our algorithm is an adapted version of this algorithm. This algorithm is presented below.

Algorithm 1: Backward induction algorithm [12]

Result: $u_t^*(s_t), A_{s_t,t}^*$: value and optimal action for all $s_t \in S$.
Initialization: initialize $u_T^*(s_T) = r_T(s_T)$ for all $s_T \in S$.
for $t = T-1, \ldots, 1$ **do**
  for $s_t \in S$ **do**
    $u_t^*(s_t) = \max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{j \in S} w(j|s_t, a) u_{t+1}^*(j) \right\}$.
    $A_{s_t,t}^* = \arg\max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{j \in S} w(j|s_t, a) u_{t+1}^*(j) \right\}$.
  end
end

We perform this algorithm for the MDP as defined in Section 4.5. The algorithm starts at $t = T$. For $t = T, \ldots, \tilde{t}+1$, the functions $u_t^*(s_t)$ do not depend on the trading probabilities. Hence the value function $u_t^*(s_t)$ is not a function of p and q for all $t > \tilde{t}$. At the trading epoch $\tilde{t}$, we calculate $u_t^*(s_t)$ according to

$$u_t^*(s_t) = \max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{j \in S} w(j|s_t, a) u_{t+1}^*(j) \right\} \tag{4.14}$$

In trading epochs, the direct reward $r(s_t, a)$ is zero, so $u_t^*(s_t)$ is calculated by calculating the maximum of the functions $\sum_{j \in S} w(j|s_t, a) u_{t+1}^*(j)$ for $a \in \{-1, 0, 1\}$.

For $a = -1$, i.e., 'buy', this function is given by

$$\begin{aligned} u_{t+1}^*(s, a) = p((1-\pi_2)u_{t+1}^*(c_t + \zeta, 0, f_t - \zeta, g_t) + \pi_2 u_{t+1}^*(c_t + \zeta, 1, f_t - \zeta, g_t)) \\ +(1-p)((1-\pi_2)u_{t+1}^*(c_t, 0, f_t, g_t) + \pi_2 u_{t+1}^*(c_t, 1, f_t, g_t)) \\ = pB(s_t) + (1-p)N(s_t). \end{aligned}$$

For $a = 0$, i.e., 'do nothing', this function is given by

$$u_{t+1}^*(s, a) = (1-\pi_2)u_{t+1}^*(c_t, 0, f_t, g_t) + \pi_2 u_{t+1}^*(c_t, 1, f_t, g_t) = N(s_t).$$

For $a = 1$, i.e., 'sell', this function is given by

$$\begin{aligned} u_{t+1}^*(s, a) = q((1-\pi_2)u_{t+1}^*(c_t - \zeta, 0, f_t + \zeta, g_t) + \pi_2 u_{t+1}^*(c_t - \zeta, 1, f_t + \zeta, g_t)) \\ +(1-q)((1-\pi_2)u_{t+1}^*(c_t, 0, f_t, g_t) + \pi_2 u_{t+1}^*(c_t, 1, f_t, g_t)) \\ = qS(s_t) + (1-q)N(s_t). \end{aligned}$$

where we have

$$\begin{aligned}
B(s_t) &= (1-\pi_2)u^*_{t+1}(c_t+\zeta, 0, f_t-\zeta, g_t) + \pi_2 u^*_{t+1}(c_t+\zeta, 1, f_t-\zeta, g_t),\\
N(s_t) &= (1-\pi_2)u^*_{t+1}(c_t, 0, f_t, g_t) + \pi_2 u^*_{t+1}(c_t, 1, f_t, g_t),\\
S(s_t) &= (1-\pi_2)u^*_{t+1}(c_t-\zeta, 0, f_t+\zeta, g_t) + \pi_2 u^*_{t+1}(c_t-\zeta, 1, f_t+\zeta, g_t).
\end{aligned}$$

$B(s_t), N(s_t)$ and $S(s_t)$ can be interpreted as the expected value of the MDP after a successful buy, after an unsuccessful trade or doing nothing, and after a successful sell, respectively. Using these expressions, we can rewrite (4.14):

$$u^*_t(s_t) = \max\{pB(s_t) + (1-p)N(s_t), N(s_t), qS(s_t) + (1-q)N(s_t)\} \tag{4.15}$$

The solution of this equation is given by the following theorem.

Theorem 4.6.2. *If $s_t \in S_{\tilde{t}}$, the value functions $u^*_t(s_t)$ and optimal actions $A^*_{s_t,t}(s_t)$ are given by*

$$u^*_t(s_t) = \begin{cases} pB(s_t) + (1-p)N(s_t) & \text{if } B(s_t) > N(s_t) \text{ and } p > \beta(s_t)q,\\ N(s_t) & \text{if } N(s_t) \geq B(s_t) \text{ and } N(s_t) \geq S(s_t),\\ qS(s_t) + (1-q)N(s_t) & \text{if } S(s_t) > N(s_t) \text{ and } p \leq \beta(s_t)q, \end{cases}$$

and

$$A^*_{s_t,t}(s_t) = \begin{cases} -1 & \text{if } B(s_t) > N(s_t) \text{ and } p > \beta(s_t)q,\\ 0 & \text{if } N(s_t) \geq B(s_t) \text{ and } N(s_t) \geq S(s_t),\\ 1 & \text{if } S(s_t) > N(s_t) \text{ and } p \leq \beta(s_t)q, \end{cases}$$

where

$$\beta(s_t) = \frac{S(s_t) - N(s_t)}{B(s_t) - N(s_t)}.$$

Proof. The value function $u^*_t(s_t)$ is calculated by solving

$$\max\{pB(s_t) + (1-p)N(s_t), N(s_t), qS(s_t) + (1-q)N(s_t)\}.$$

The maximum value is $pB(s_t) + (1-p)N(s_t)$ if

$$pB(s_t) + (1-p)N(s_t) > N(s_t) \implies pB(s_t) - pN(s_t) > 0 \implies B(s_t) > N(s_t)$$

and

$$\begin{aligned}
& pB(s_t) + (1-p)N(s_t) > qS(s_t) + (1-q)N(s_t)\\
\implies & p(B(s_t) - N(s_t)) + N(s_t) > q(S(s_t) - N(s_t)) + N(s_t)\\
& \implies p > q\frac{S(s_t) - N(s_t)}{B(s_t) - N(s_t)} = q\beta s_t.
\end{aligned}$$

The action corresponding to this maximum is 'buy' so the optimal action is given by $A^*_{s_t,t}(s_t) = -1$.

Similarly, the value $qS(s_t) + (1-q)N(s_t)$ with corresponding action 'sell' is the maximum if

$$qS(s_t) + (1-q)N(s_t) > N(s_t) \implies qS(s_t) - qN(s_t) > 0 \implies S(s_t) > N(s_t)$$

and

$$\begin{aligned}
& qS(s_t) + (1-q)N(s_t) > pB(s_t) + (1-p)N(s_t)\\
\implies & q(S(s_t) - N(s_t)) + N(s_t) > p(B(s_t) - N(s_t)) + N(s_t)\\
& \implies p \leq q\frac{S(s_t) - N(s_t)}{B(s_t) - N(s_t)} = q\beta s_t.
\end{aligned}$$

Finally, $N(s_t)$ is maximal when

$$N(s_t) > pB(s_t) + (1-p)N(s_t) \implies N(s_t) > B(s_t)$$

and

$$N(s_t) > qS(s_t) + (1-q)N(s_t) \implies N(s_t) > S(s_t).$$

□

It can be seen that the value and the optimal action may depend on the estimated trading probabilities. If $N(s_t)$ is strictly smaller than either $B(s_t)$ or $S(s_t)$, the optimal action depends on p and q via the inequalities $p > \beta(s_t)q$ and $p < \beta(s_t)q$. These inequalities do not have a solution for $B(s_t) > N(s_t) > S(s_t)$ and $S(s_t) > N(s_t) > B(s_t)$, respectively, as $\beta < 0$ in these cases and p, q are both non-negative. In the other cases, we can represent the inequality by a line in the (p, q)-plane (with $0 \le p, q \le 1$) that splits the plane in two parts: both parts represent the region for which an action is optimal. This is illustrated in Figure 4.4.

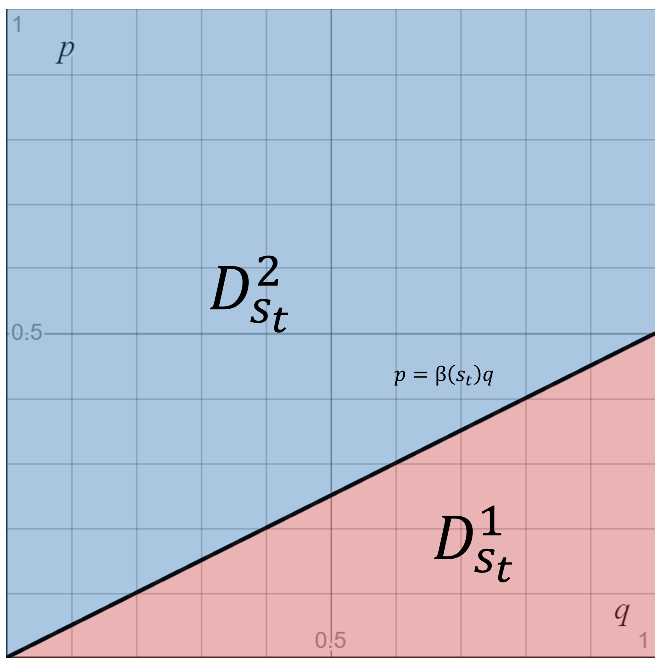

Figure 4.4: Subdomains induced by the subdomain backward induction algorithm. In $D^2_{s_t}$, $p > \beta(s_t)q$. In $D^1_{s_t}$, $p < \beta(s_t)q$.

Hence, in the case that $B(s_t), S(s_t) > N(s_t)$ for $s_t \in \mathfrak{T}^T$, the value $u^*_t(s_t)$ and optimal action $A_{s_t,t}(s_t)$ are given by

$$u^*_t(s_t) = \begin{cases} pB(s_t) + (1-p)N(s_t) & \text{if } p, q \in D^2_{s_t} \\ qS(s_t) + (1-q)N(s_t) & \text{if } p, q \in D^1_{s_t} \end{cases}$$

$$A^*_{s_t,t}(s_t) = \begin{cases} -1 & \text{if } p, q \in D^2_{s_t} \\ 1 & \text{if } p, q \in D^1_{s_t}, \end{cases}$$

where

$$D^1_{s_t} = \{p, q : 0 \le p, q \le 1, p \le \beta(s_t)q\}$$
$$D^2_{s_t} = \{p, q : 0 \le p, q \le 1, p > \beta(s_t)q\}.$$

In the other cases, that is the cases that do not have a solution for $0 \leq p, q \leq 1$ and the case $N(s_t) > B(s_t), S(s_t)$, the optimal action does not depend on p and q.

Now for $t < \tilde{t}$, $u_t^*(s_t)$ is calculated according to Equation 4.14, but as Theorem 4.2 shows, the functions $u_{t+1}^*(j)$ can be functions of p and q and need not be continuous. Therefore we need to specify the domain (i.e., the values of p and q) for which $u_t^*(s_t)$ is calculated. To do that, we identify M_{s_t}, the set of states $j \in S_{t+1}$ that are reachable from s_t, i.e., for which $w(j|s_t, a) > 0$ for some $a \in A_{s_t}$. Then we take the $|M_{s_t}|$*-fold Cartesian product* of the subdomains of the functions $u_{t+1}^*(j)$ for $j \in M_{s_t}$. This is defined as follows.

Definition 4.6.3 (n-fold Cartesian product)**.** The n-fold Cartestian product over the sets $D_1, \ldots, D_n$ can be defined as the set

$$D_1 \times \cdots \times D_n = \{D_1^j \cap \cdots \cap D_n^j | D_i^j \in D_i \text{ for every } i \in \{1, \ldots, n\}\},$$

where D_i is defined as in Equation 4.13.

This is illustrated using Figures 7 and 8.

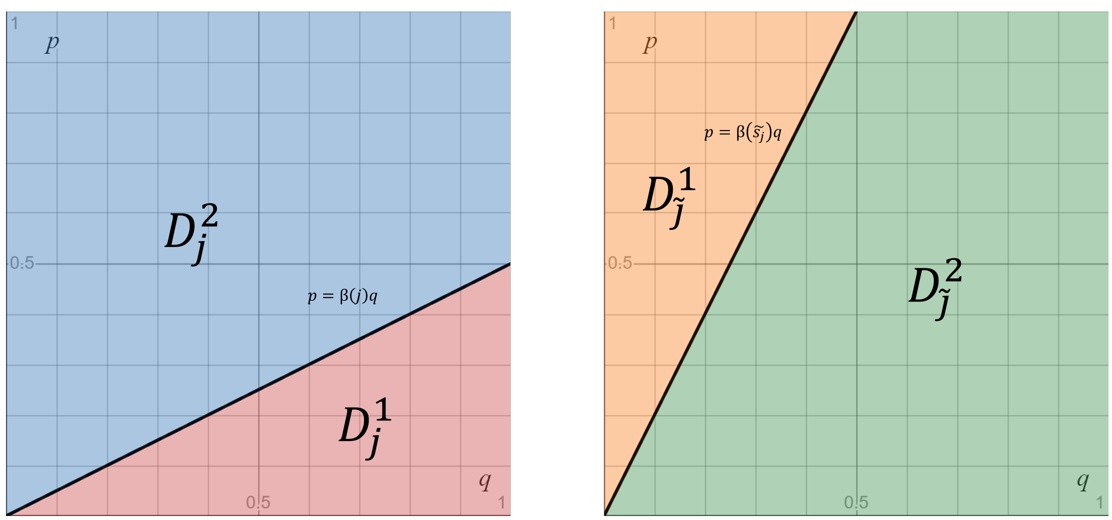

Figure 4.5: $M_{s_t} = \{j, \tilde{j}\}$. The domain is split in two subdomains for both $u_{t+1}^*(j)$ and $u_{t+1}^*(\tilde{j})$.

We consider state s_t. Suppose that the set of states that are reachable from s_t is given by $M_{s_t} = \{j, \tilde{j}\}$ and that the value functions of these states are given by

$$u_{t+1}^*(j) = \begin{cases} f_1^j(p,q) & \text{if } p, q \in D_j^1 \\ f_2^j(p,q) & \text{if } p, q \in D_j^2 \end{cases}$$

and

$$u_{t+1}^*(j) = \begin{cases} f_1^{\tilde{j}(p,q)} & \text{if } p, q \in D_{\tilde{j}}^1 \\ f_2^{\tilde{j}(p,q)} & \text{if } p, q \in D_{\tilde{j}}^2. \end{cases}$$

Hence, $D_j = \{D_j^1, D_j^2\}$ and $D_{\tilde{j}} = \{D_{\tilde{j}}^1, D_{\tilde{j}}^2\}$ (Figure 4.5). The 2-fold Cartesian product of the sets of regions D_j and $D_{\tilde{j}}$ is defined as

$$M' = \{(D_j^1 \cap D_{\tilde{j}}^1), (D_j^1 \cap D_{\tilde{j}}^2), (D_j^2 \cap D_{\tilde{j}}^1), (D_j^2 \cap D_{\tilde{j}}^2)\}.$$

We have

$$D_j^1 \cap D_{\tilde{j}}^1 = \emptyset$$

$$D_j^1 \cap D_{\tilde{j}}^2 = \{p, q : 0 \leq p, q \leq 1, p \leq \beta(j)q\} = D_{s_t}^3$$

$$D_j^2 \cap D_{\tilde{j}}^1 = \{p, q : 0 \leq p, q \leq 1, p > \beta(\tilde{j})q\} = D_{s_t}^1$$

$$D_j^2 \cap D_{\tilde{j}}^2 = \{p, q : 0 \leq p, q \leq 1, \beta(j)q < p \leq \beta(\tilde{j})q\} = D_{s_t}^2,$$

so $D_{s_t} = \{D_{s_t}^1, D_{s_t}^2, D_{s_t}^3\}$. These sets are depicted in Figure 4.6. Using these sets, the maximization problem in Equation 4.14 can be solved for all subdomains $D \in D_{s_t}$. In this way, $u_t^*(s_t)$ can be calculated. This procedure is formalized in Algorithm 2, the Subdomain backward induction algorithm.

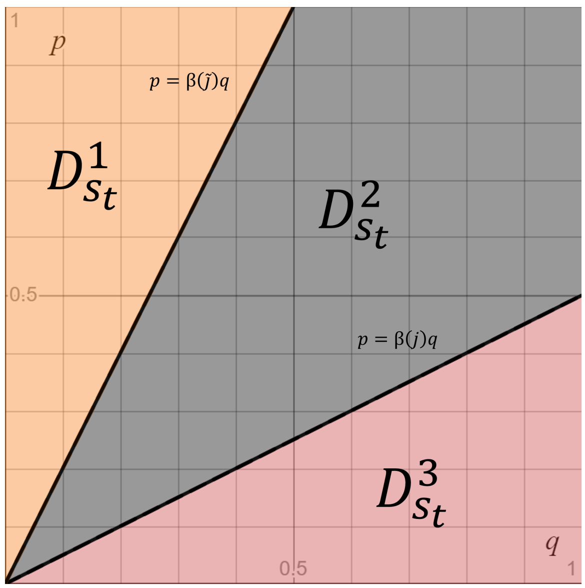

Figure 4.6: Cartesian product of D_j and $D_{\tilde{j}}$.

Algorithm 2: Subdomain backward induction algorithm

Result: $u_t^*(s_t), A_{s_t,t}^*$: value and optimal action for all $s_t \in S$ as function of p and q.
Initialization: initialize $u_T^*(s_T) = r_T(s_T)$ for all $s_T \in S$.
for $t = T-1, \ldots, 1$ **do**
    for $s_t \in S$ **do**
        $M = \emptyset$
        for j *s.t.* $w(j|s_t, a) > 0$ *for some* $a \in A_{s_t}$ **do**
            $M = M + j$
        end
        $M' = |M|-$fold Cartesian product of the sets $D_j, j \in M$.
        for $D \in M'$ **do**
            if $D \neq \emptyset$ **then**
                $u_t^*(s_t) = \max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{j \in S} w(j, s_t, a) u_{t+1}^*(j) \right\}$ for $p, q \in D$.
                $A_{s_t,t}^* = \arg\max_{a \in A_{s_t}} \left\{ r(s_t, a) + \sum_{j \in S} w(j, s_t, a) u_{t+1}^*(j) \right\}$ for $p, q \in D$.
            end
        end
    end
end

By adjusting the backward induction algorithm in this way, we can calculate values and optimal actions for all states for discontinuous value functions. Moreover, the algorithm creates a clear overview of the different policies that are induced by the trading probabilities. When the algorithm terminates, the output is the value and optimal action for all $s_t \in S$ as function of p and q, defined on the set of subdomains $D_{s_t} = \{D^1_{s_t}, \ldots, D^k_{s_t}\}$. The set of subdomains of the starting state, M_{s_0}, consists of elements D_i that all induce a policy that is optimal for the values of p and q in the subdomain. Therefore, the number of different optimal policies is given by the cardinality of this set, $|M_{s_0}|$.

4.6.3 Trading equilibrium algorithm

Let the starting states of the system for player i, j be $s_{1,i}, s_{1,j}$, respectively. The set of subdomains for this state is denoted by

$$D_{s_1,i} = \{D^1_{s_1,i}, \ldots, D^k_{s_1,i}\}.$$

The subdomain backward induction algorithm calculates the value and optimal action for all $s_t \in S$ as function of p and q. Each subdomain D_i corresponds to a policy that is optimal for the player when the estimated trading probabilities lie in D_i. So for a given p and q, we are able to calculate the optimal policy for a player using the subdomain backward induction algorithm, using

$$\Pi^*(p, q) = \Pi^*_{D_i} \quad \text{for all } p, q \in D_i. \tag{4.16}$$

We now present an algorithm that uses these subdomain-policies to find the trading equilibria from Definition 4.2.

Algorithm 3: Trading equilibrium algorithm

Input: optimal policies $\Pi^*_{D_i}, \Pi^*_{D_j}$ for all $D_i \in D_{s_1,i}, D_j \in D_{s_1,j}$.
Result: $l(D_i), l(D_j)$ for all subdomains $D_i \in D_{s_1,i}, D_j \in D_{s_1,j}$: true trading probabilities under optimal policies $\Pi^*_{D_i}, \Pi^*_{D_j}$, set of trading equilibria TE.
Initialization: initialize $l(D_i), l(D_j) = -\infty$ for all $D_i \in D_{s_1,i},, D_j \in D_{s_1,j}$ and $TE = \emptyset$.
for $D_i \in D_{s_1,i}$ **do**
  $p_j(D_i) = \sum_{s \in S^s_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{D_i}}(s_{i\tilde{t}} = s | s^i_0)$
  $q_j(D_i) = \sum_{s \in S^b_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{D_i}}(s_{i\tilde{t}} = s | s^i_0)$
  Set $l(D_i) = (p_j(D_i), q_j(D_i))$
  Take D_j s.t. $l(D_i) \in D_j$
  if $l(D_j) \neq -\infty$ **then**
    if $l(D_j) \in D_i$ **then**
      $TE = TE + (p_i(D_j), q_i(D_j), p_j(D_i), q_j(D_i))$
    end
  end
  else
    $p_i(D_j) = \sum_{s \in S^s_{j\tilde{t}}} \mathbb{P}^{\Pi^*_{D_j}}(s_{j\tilde{t}} = s | s^j_1)$
    $q_i(D_j) = \sum_{s \in S^b_{j\tilde{t}}} \mathbb{P}^{\Pi^*_{D_j}}(s_{j\tilde{t}} = s | s^j_1)$
    Set $l(D_j) = (p_i(D_j), q_i(D_j))$
    if $l(D_j) \in D_i$ **then**
      $TE = TE + ((p_i(D_j), q_i(D_j), p_j(D_i), q_j(D_i))$
    end
  end
end

Theorem 4.6.3. *Algorithm 3 finds all solutions for the trading equilibrium equations (2)-(6). If there is no solution, the algorithm will terminate with $TE = \emptyset$.*

Proof. To prove the statement, we first prove that all trading equilibria are found by the algorithm. To that end, suppose that $(\bar{p_i}, \bar{q_i}, \bar{p_j}, \bar{q_j})$ is a trading equilibrium and let $\bar{p_i}, \bar{q_i} \in \bar{D_i}$ w.l.o.g. For

$\bar{D}_i$, the algorithm calculates

$$p_j(\bar{D}_i) = \sum_{s \in S^s_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{\bar{D}_i}}(s_{i\tilde{t}} = s | s^i_0) = \sum_{s \in S^s_{i\tilde{t}}} \mathbb{P}^{\Pi^*(\bar{p}_i, \bar{q}_i)}(s_{i\tilde{t}} = s | s^i_0) = \bar{p}_j$$

and

$$q_j(\bar{D}_i) = \sum_{s \in S^b_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{\bar{D}_i}}(s_{i\tilde{t}} = s | s^i_0) = \sum_{s \in S^b_{i\tilde{t}}} \mathbb{P}^{\Pi^*(\bar{p}_i, \bar{q}_i)}(s_{i\tilde{t}} = s | s^i_0) = \bar{q}_j.$$

The first equality follows from Equation 4.16 and the second equality follows from the definition of a trading equilibrium. Then

$$l(\bar{D}_i) = (\bar{p}_j, \bar{q}_j) \in D_j.$$

For the subdomain D_j, the algorithm calculates

$$p_i(\bar{D}_j) = \sum_{s \in S^s_{j\tilde{t}}} \mathbb{P}^{\Pi^*_{\bar{D}_j}}(s_{j\tilde{t}} = s | s^j_1) = \sum_{s \in S^s_{j\tilde{t}}} \mathbb{P}^{\Pi^*(\bar{p}_j, \bar{q}_j)}(s_{j\tilde{t}} = s | s^i_1) = \bar{p}_i$$

and

$$q_i(\bar{D}_j) = \sum_{s \in S^b_{j\tilde{t}}} \mathbb{P}^{\Pi^*_{\bar{D}_j}}(s_{j\tilde{t}} = s | s^j_1) = \sum_{s \in S^b_{j\tilde{t}}} \mathbb{P}^{\Pi^*(\bar{p}_j, \bar{q}_j)}(s_{j\tilde{t}} = s | s^j_1) = \bar{q}_i.$$

Since $l(\bar{D}_j) = (\bar{p}_j, \bar{q}_j) \in \bar{D}_i$, we have

$$(\bar{p}_i, \bar{q}_i, \bar{p}_j, \bar{q}_j) \in TE.$$

Hence, the trading equilibrium is found by the algorithm. This holds for any trading equilibrium, because the algorithm visits all $D_i \in D_{s_1,i}$ and

$$\bigcup_{D_i \in D_{s_1,i}} D_i = \{p, q : 0 \leq p, q \leq 1\}.$$

Moreover, the set $D_{s_1,i}$ is finite, so the algorithm will always terminate.

To prove that TE exclusively consists of trading equilibria and $TE = \emptyset$ when there are no trading equilibria, suppose to the contrary that there are no trading equilibria, but TE is nonempty. That is, suppose that $(\tilde{p}_i, \tilde{q}_i, \tilde{p}_j, \tilde{q, j}) \in TE$, where $(\tilde{p}_i, \tilde{q}_i, \tilde{p}_j, \tilde{q, j})$ is not a trading equilibrium. By construction of TE, we have

$$\tilde{p}_j = p_j(\tilde{D}_i) = \sum_{s \in S^s_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{\tilde{D}_i}}(s_{i\tilde{t}} = s | s^i_1) \tag{4.17}$$

and

$$\tilde{q}_j = q_j(\tilde{D}_i) = \sum_{s \in S^b_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{\tilde{D}_i}}(s_{i\tilde{t}} = s | s^i_1) \tag{4.18}$$

for some $D_i \in D_{s_1,i}$. Moreover, $(\tilde{p}_j, \tilde{q}_j) \in D_j$ and

$$\tilde{p}_i = p_i(\tilde{D}_j) = \sum_{s \in S^s_{j\tilde{t}}} \mathbb{P}^{\Pi^*_{\tilde{D}_j}}(s_{j\tilde{t}} = s | s^j_1) = \sum_{s \in S^s_{j\tilde{t}}} \mathbb{P}^{\Pi^*(\tilde{p}_j, \tilde{q}_j)}(s_{j\tilde{t}} = s | s^j_1) \tag{4.19}$$

$$\tilde{q}_i = q_i(\tilde{D}_j) = \sum_{s \in S^b_{j\tilde{t}}} \mathbb{P}^{\Pi^*_{\tilde{D}_j}}(s_{j\tilde{t}} = s | s^j_1) = \sum_{s \in S^b_{j\tilde{t}}} \mathbb{P}^{\Pi^*(\tilde{p}_j, \tilde{q}_j)}(s_{j\tilde{t}} = s | s^j_1). \tag{4.20}$$

Since $(\tilde{p}_i, \tilde{q}_i, \tilde{p}_j, \tilde{q, j}) \in TE$, we have

$$l(\tilde{D}_j) = (\tilde{p}_i, \tilde{q}_i) \in \tilde{D}_i.$$

Therefore, Equations 4.17 and 4.18 can be rewritten:

$$\tilde{p_j} = p_j(\tilde{D_i}) = \sum_{s \in S^s_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{\tilde{D}_i}}(s_{i\tilde{t}} = s | s^i_1) = \sum_{s \in S^s_{i\tilde{t}}} \mathbb{P}^{\Pi^*(\tilde{p_i}, \tilde{q_i})}(s_{i\tilde{t}} = s | s^i_1) \tag{4.21}$$

and

$$\tilde{q_j} = q_j(\tilde{D_i}) = \sum_{s \in S^b_{i\tilde{t}}} \mathbb{P}^{\Pi^*_{\tilde{D}_i}}(s_{i\tilde{t}} = s | s^i_1) = \sum_{s \in S^b_{i\tilde{t}}} \mathbb{P}^{\Pi^*(\tilde{p_i}, \tilde{q_i})}(s_{i\tilde{t}} = s | s^i_1). \tag{4.22}$$

Now Equations 4.19-4.22 satisfy the trading probability equations. Thus, $(\tilde{p_i}, \tilde{q_i}, \tilde{p_j}, \tilde{q,j})$ is a trading equilibrium and we have a contradiction. We conclude that $TE = \emptyset$. □

4.7 Analysis of collective behaviour

In this section, we describe how we calculate the acceptance ratios under a given allocation rule as formulated in Section 4.3. We first obtain individual policies, and after that, we perform simulations to obtain the collective behaviour. Under this collective behaviour, we calculate the acceptance ratios for both departments.

4.7.1 Finding individual policies

We start with finding the individual policies for all departments. These policies are the solutions to the set of MDPs $\mathcal{M}(h)$ as defined in Section 4.3. Under all allocation rules, the minimum amount of capacity that a department receives is $\mathbb{E}[D_i]$, and the maximum amount of capacity that a department receives is $(1 + rb)\,\mathbb{E}[D_i]$. Therefore, we need to solve the MDPs $\mathcal{M}(c^{i,y}_0)$ with $c^{i,1}_0 = (1 + b)\,\mathbb{E}[D_i]$ and $c^{i,y}_0 = \mathbb{E}[D_i], \ldots, (1 + rb)\,\mathbb{E}[D_i]$ for $y = 2, \ldots, H$. The state space, action space, transition probabilities and reward function are defined according to the model that we described in Section 4.5. The MDPs are solved using Algorithm 4.

Algorithm 4: Master algorithm individual policies

Result: Optimal policies $\Pi(c^{i,y}_0)$ and values $V(c^{i,y}_0)$ for $\mathcal{M}(c^{i,y}_0)$ for $i = 1, \ldots, N, y = 1, \ldots, H$ and $c^{i,y}_0 = \mathbb{E}[D_i], \ldots, (1 + rb)\,\mathbb{E}[D_i]$.
Input: $b_i, \alpha_i, r, \zeta, \sigma, p_i, q_i$.
for $i = 1, \ldots, N$: **do**
  for $y = H, \ldots, 1$: **do**
    for $c = \mathbb{E}[D_i], \ldots, (1 + rb)\,\mathbb{E}[D_i]$ **do**
      Calculate $u^*_t(s_t)$ for all $s_t \in S$ using backward induction algorithm for $\mathcal{M}(c^{i,y}_0)$.
      Set $V(c^{i,y}_0) = \sum_{s \in S_1} p(s)u^*_1(s)$, where $p(s)$ is initial state distribution.
    end
  end
end

The algorithm can be explained using Figure 4.7. Notice that the MDPs are connected by the final rewards (Equation 4.2). At the final year of the horizon, $y = H$, the final rewards only depend on the state, or more specifically, on the number of rejected type 1 scans and the parameter α, so they can be calculcated directly. Therefore, the algorithm starts with solving the MDPs in year $y = H$. For $y < H$, the final rewards depend on the value of starting in the next year with some allocated capacity c. Thus $V(c^{i,y+1}_0)$ serves as input to calculate $V(c^{i,y}_0)$. The output of the algorithm is the set of policies $\Pi(c^{i,y}_0)$for $\mathcal{M}(c^{i,y}_0)$ for $i = 1, \ldots, N, y = 1, \ldots, H$ and $c^{i,y}_0 = \mathbb{E}[D_i], \ldots, (1 + rb)\,\mathbb{E}[D_i]$.

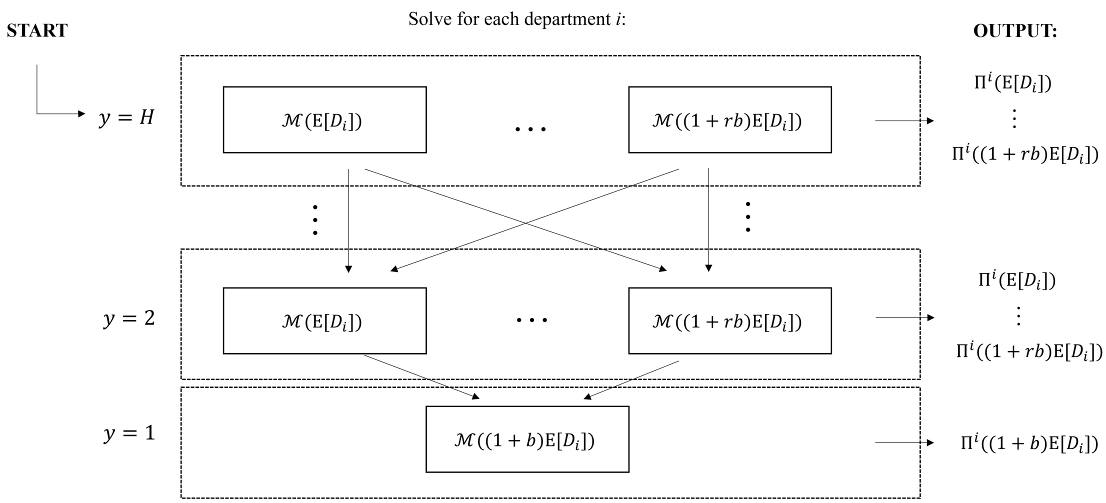

Figure 4.7: Visualization of Algorithm 4.

4.7.2 Simulation & performance analysis

To measure the effect of the allocation rule, we perform Algorithm 4 to find optimal policies for the players. After that, we perform simulations to find the proportion of type 1 and type 2 scans that are accepted when the players act according to the optimal policy. This process is depicted in Figure 4.8.

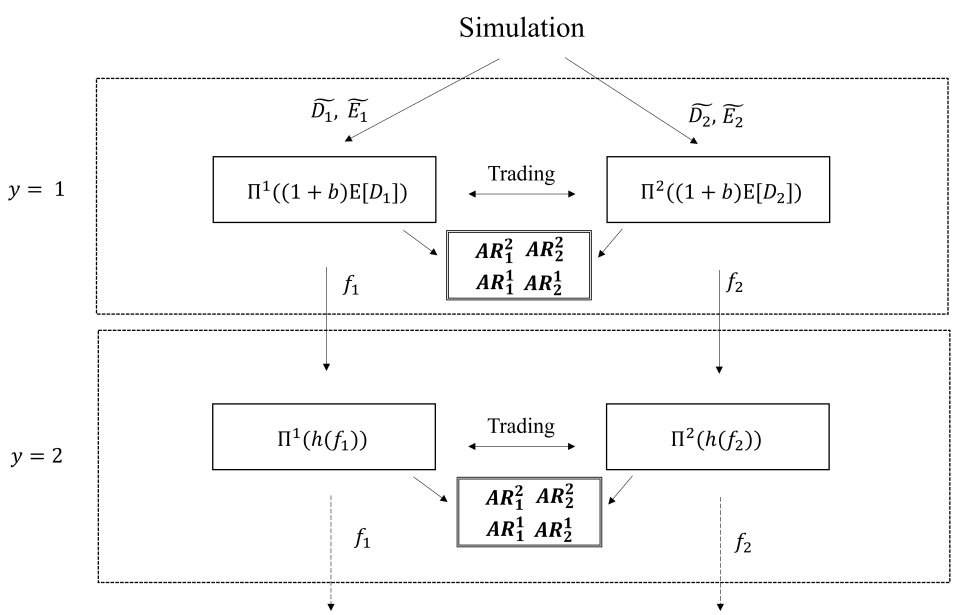

Figure 4.8: Visualization of the simulation to to obtain acceptance ratios under collective behaviour.

Each year, the demands D_i and E_i are simulated. Both departments act according to the optimal policy for scan budget $c_0^{i,y}$. In the first year, this scan budget is given by $(1+b)\,\mathbb{E}[D_i]$ for both departments. In the years after the first year, the scan budget is given by $c_0^{i,y} = h(f_i^{y-1})$, the trading balance of the previous year. This trading balance is determined by the amount of budget that was actually sold or bought by a department. Trading only succeeds when one department performs action 'sell', and the other performs action 'buy' during a trading epoch. In each year, we track the allocation ratios for both scan types for both departments.

Chapter 5

Numerical results

In this section, we present results from numerical experiments. We perform these experiments to answer our research questions. These were given by:

1. How can we model the acceptance and trading policy of a single department under a budget system with trading, and what does this policy look like?
2. What is the effect of a reward and penalty mechanism carried out by radiology for trading on a single department's willingness to trade?
3. How does a department's belief on other departments' willingness to trade influence the individual acceptance and trading policy?
4. What are the effects of the individual policies on the collective behaviour?
5. What are the acceptance ratios of type 1 and type 2 under the collective behaviour?

We answer these questions step-by-step: we first evaluate the individual policy for a given set of parameters. The parameters that are used in the experiments are given in Table 5.1.

Parameter	**Description**
$\mathbb{E}[D_i]$	Expected number of type 1 scan candidates for department i.
$\mathbb{E}[E_i]$	Expected number of type 2 scan candidates for department i.
b	Buffer.
α	Penalty for rejecting type 1 scan.
r	Allocation rule parameter.
ζ	Trading amount.
σ	Number of trading epochs.
p_i	Success probability buying for player i.
q_i	Success probability selling for player i.

Table 5.1: Parameters

These parameters include the estimated success probabilities for buying and selling as well as the allocation rule parameter. Thus sub-question 2 and sub-question 3 can be answered by evaluating the individual policies under different values of r and p_i, q_i. Then, we answer the fourth and fifth question by performing simulations of the collective behaviour, where the departments act according to their individual policies. With these simulations, we evaluate the trading that takes place under the collective behaviour and the total acceptance ratios of departments as well as the type-specific acceptance ratios. Finally, we perform an extensive sensitivity analysis for all parameters in Table 5.1 to see how the parameters influence the acceptance ratios. We only show results for symmetric cases, as results for asymmetric cases do not fundamentally differ from symmetric cases, i.e., have similar acceptance ratios.

5.1 Effect of trading

In this section, we discuss numerical results from simulations. To evaluate the effect of trading on the acceptance ratios, we compare three cases:

1. **No trading (NT):** the case where players have an annual trading budget $c_0^{i,y}$ (where $\sum_i c_0^{i,y} = \sum_i (1+b)\,\mathbb{E}[D_i]$ for all y) without trading possibilies and act according to the optimal policies calculated by Algorithm 4.
2. **Trading (T)**: the case where players have an annual trading budget $c_0^{i,y}$ (where $\sum_i c_0^{i,y} = \sum_i (1+b)\,\mathbb{E}[D_i]$ for all y) and the possibility to buy or sell ζ units of capacity during trading epochs. Players act based on the optimal policies calculated by Algorithm 4.
3. **Centralized admission control (CA)**: in this case, all scan candidates are collected by a central agent, who decides based on the remaining capacity if a scan can be accepted. We calculate the policy of the central agent by Algorithm 4 with $H = 1$ and $c_0 = (1+b)(\mathbb{E}[D_1] + \mathbb{E}[D_2])$.

We start by evaluating the individual policy for a department with expected demand sizes $\mathbb{E}[D] = 10$ and $\mathbb{E}[E] = 3$. We set the buffer $b = 0.1$, the value of a type 1 scan $\alpha = 2$, the trading amount $\zeta = 1$ and the number of trading epochs $\sigma = 1$. The estimated success probabilities of trading are both set at 1, so we calculate the policy for a department that has full confidence in trading. The allocation rule parameter is first set to 1, meaning that there is no reward or penalty for trading. We perform Algorithm 4, that outputs policies for H years with different initial budgets. The possible initial budgets are given by $\mathbb{E}[D] = 10, \ldots, (1+br)\,\mathbb{E}[D] = 11$ and the number of years is set to $H = 10$. The policy for year 1 with initial budget 11 is depicted in Figure 5.1.

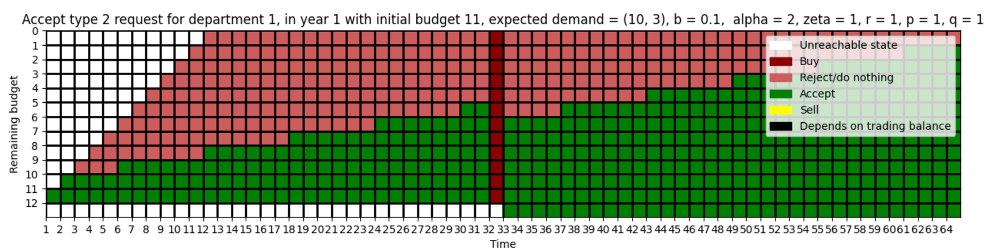

Figure 5.1: Individual policy with $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\alpha = 2$, $\zeta = 1$, $r = 1$, $p = 1$, $q = 1$.

On the y-axis, the remaining budget is given. The colors represent the optimal action for a given remaining budget and time. Pink and green correspond to the decisions 'accept' and 'reject', respectively. So for example, at time $t = 8$ a type 2 request is accepted if the remaining budget is 9 or higher. At $t = 32$, there is a trading epoch. In this trading epoch, the optimal action is always buy, regardless of the remaining budget. This can be explained by the estimated trading probabilities and the allocation rule parameter: there is no penalty or reward for trading and the estimated success probability for buying is 1. Thus, by performing the action 'buy' the department is guaranteed to increase its remaining budget by 1 without being penalized for it. We find that the optimal policy does not change over the years and does not depend on the initial budget.

Increasing the allocation parameter r to 1.5 yields the policy in Figure 5.2.

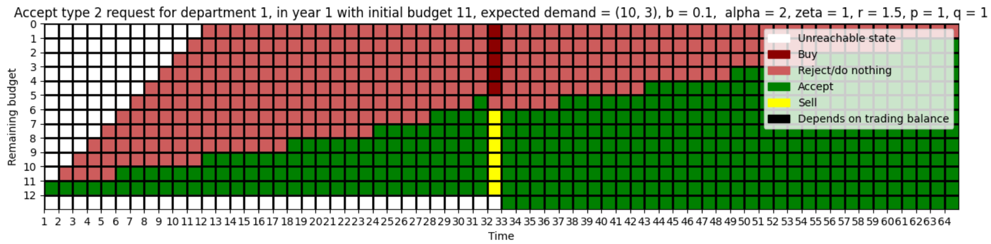

Figure 5.2: Individual policy with $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\alpha = 2$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$.

In this case, the optimal action in the trading epoch depends on the remaining budget. If the remaining budget is less than 5, the department will (try to) buy budget. If the capacity is 5, the department will neither buy nor sell. In case the remaining budget is greater than 5, the optimal action is sell. These actions can be related to the expected size of the type 1 demand: the expected number of type 1 scan candidates that arrive during the half year after the trading epoch is 5.

Next, we consider the collective behaviour of two departments. We start by performing 1000 simulations for $H = 10$ years with $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$, $b = 0.1$. $\alpha = 2$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$. We consider the Case N, where the players have an annual scan budget, without trading possibilities. The resulting acceptance ratios (ARs) are given in Figure 5.3. The average acceptance ratios (ARs) for both types over the years are very stable, there are no large annual differences. For that reason, we do not show separate results for all years. Instead, we show the average numbers over 1000 simulations with a time horizon of $H = 10$ years, where the standard deviation of the ARs is also given.

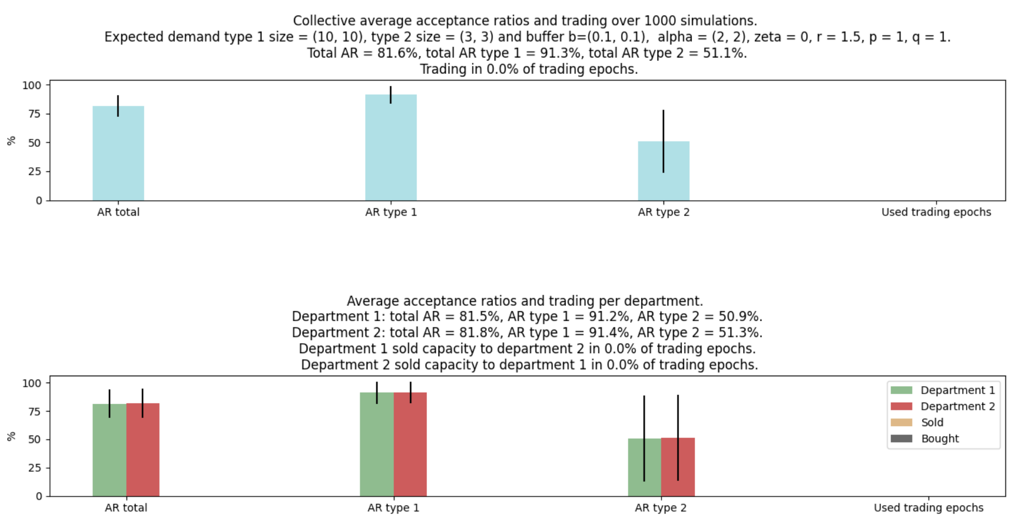

Figure 5.3: (NT) simulation for two departments with $\mathbb{E}[D_i] = 10$, $\mathbb{E}[E_i] = 3$, $b = 0.1$ and $\alpha = 2$ Moreover, $\zeta = 0$, so there is no trading.

In the upper graph, the total acceptance ratios and the used total capacity is shown. It can be seen that the average acceptance ratio for type 1 and type 2 scans together over all departments is 81.6% and that the total ARs for type 1 and type 2 are 91.3% and 51.1%, respectively. There is a clear difference in type 1 and type 2 scans: only half of the type 2 scans are accepted, while the acceptance ratio for type 1 scans is 91.3%. In the lower graph, these numbers are split out for the departments. It can be seen that there are hardly any differences between the departments, which is as expected for a symmetric situation without interaction between the departments. The standard deviations are also given in both graphs. The standard deviations for the type 1 acceptance ratios are small, and there is only a small difference between the standard deviation in the collective case and the standard devations of the type 1 ARs of the departments. The standard deviation of the AR of type 2 scans is significantly larger than the other standard deviations. This can be explained by two factors. First, the numbers of the type 2 demand are small, which results in large differences in ARs, because the acceptance of a single scan has a relatively large influence on the acceptance ratio. Second, the number of type 2 scans that are accepted depends on the number of type 1 scans that arrive, and is therefore more variable than the number of type 1 scans that is accepted. Finally, we also look at the average absolute annual difference (AAAD) between the ARs for type 1 of both departments. This number gives an indication of the 'fairness' of the allocation mechanism. For this case, the AAAD is 9.6 percentage point (pp).

Next we consider T, the case where we allow trading. We perform another 1000 simulations, but now the departments can trade capacity. We set $\zeta = 1$ and $r = 1.5$. The averages over the 1000 simulations can be seen in Figure 5.4.

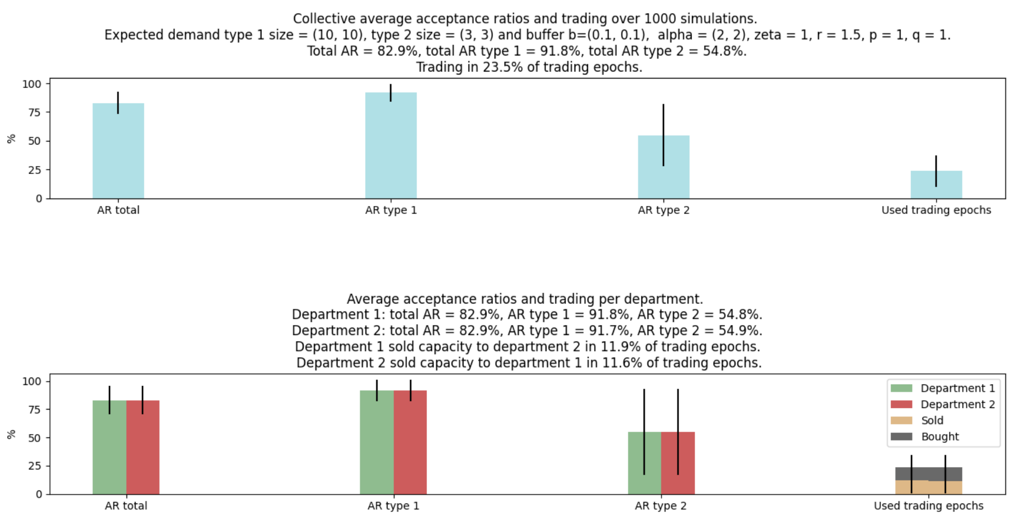

Figure 5.4: Case T simulation for two departments with $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$, $b = 0.1$, $\alpha_1 = 2$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$.

It can be seen that both acceptance ratios (slightly) increase now that there is an opportunity to trade, so the total number of rejected scans decreases. The total acceptance ratios are both 82.9%, similar to the total acceptance ratios in Case 1, and higher than the total acceptance ratios in Case 2. Thus, trading seems to reduce the waste of capacity. Trading actually happens: it can be seen that trading took place in 23.5% of the trading epochs. Again, the performance measures are symmetric: there are no clear structural differences between the departments. The AAAD is 8.2 pp, which means that a trading possibility also reduces the annual departmental difference.

To investigate the dynamics between departments, we zoom in on a single simulation. This simulation is given in Figure 5.5.

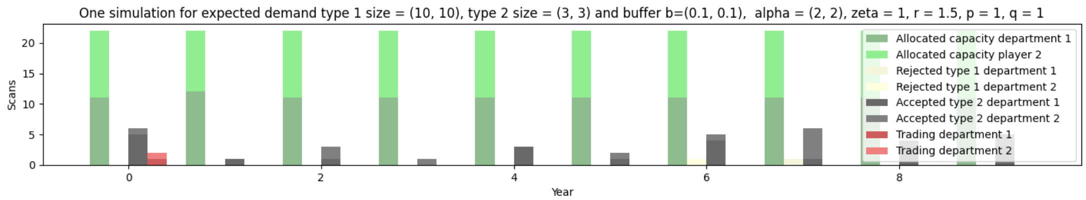

Figure 5.5: Case T simulation for two departments with $\mathbb{E}[D_i] = 10$, $\mathbb{E}[E_i] = 3$, $b = 0.1$ and $\alpha_1 = 2$, $\zeta = 1$, $r = 1.5$.

In this figure, we can see that trading takes place in the first year. In this year, department 1 sells and therefore receives more capacity in year 3. Department 2 buys capacity from department 1 and therefore receives less capacity in year 3. In year 2, there is no trading, so after this year the allocated capacity is set to the standard amount ($(1+b)\,\mathbb{E}[D]$) again for both departments.

Finally, we perform 1000 simulations with centralized admission control. We set the expected total type 1 demand $\mathbb{E}[D] = 20$, the total type 2 demand $\mathbb{E}[E]$=6 and $\alpha = 2$.

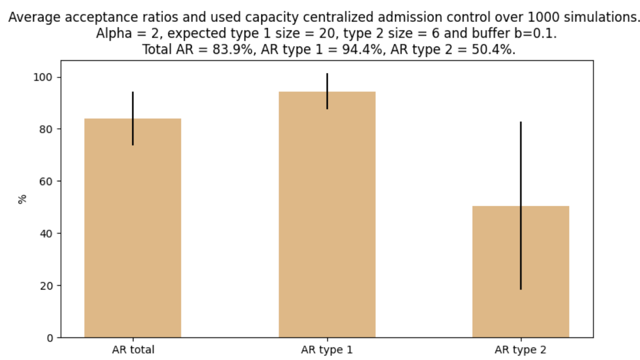

Figure 5.6: Case CA simulation (centralized admission control) with $\mathbb{E}[D] = 20$, $\mathbb{E}[E] = 6$, $b = 0.1$ and $\alpha = 2$.

In this case, the total acceptance ratio increases to 84.0%. The acceptance ratio for type 1 scans is 94.5%, which is higher than the acceptance ratio for type 1 scans in Case T (91.8%). The acceptance ratio for type 2 scans is 50.6%, which is smaller than the acceptance ratio in Case T (54.8%). The standard deviation of the AR of type 2 is still large. We conclude that under centralized admission control, the acceptance ratio for type 1 increases whereas the acceptance ratio for type 2 decreases, leading to a positive net result as the total acceptance ratio increases compared to Case T. The acceptance ratios under centralized admission control serve as an upper bound for the acceptance ratios for the decentralized cases.

5.2 Sensitivity analysis of parameters

In this section, we study the influence of the parameters on the acceptance ratios. For all experiments, we set $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$ for both departments.

Buffer

We start with a sensitivity analysis of the buffer parameter b. In the following table, we present the total acceptance ratios for different values of b. The first value is the total acceptance ratio for both types, the second value represents the total acceptance ratio for type 1 scans and the third number represents the total acceptance ratio for type 2 scans.

$b\downarrow$	**Case NT**	**Case T**	**Case CA**
0.1	81.6 / 91.3 / 51.1 / 9.6	82.9 / 91.8 / 54.8 / 8.2	84.0 / 94.5 / 50.6
0.2	87.3 / 93.5 / 66.7 / 7.3	88.5 / 93.9 / 70.3 / 6.7	90.1 / 95.9 / 70.8
0.3	91.6 / 95.4 / 78.5 / 5.4	92.6 / 95.6 / 82.6 / 4.9	94.4 / 97.7 / 83.3
0.4	94.7 / 96.9 / 87.2 / 3.7	95.0 / 97.3 / 87.5 / 3.5	97.1 / 98.4 / 92.6
0.5	96.7 / 98.0 / 92.3 / 2.2	97.5 / 98.4 / 94.5 / 2.2	98.7 / 99.2 / 97.1

Table 5.2: Acceptance ratios for $E[D] = 10$, $E[E] = 3$, $\alpha = 2$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$, $\sigma = 4$ and different values of b. The first number is the total acceptance ratio, the second number is the acceptance ratio for type 1 scans, the third number is the acceptance ratio for type 2 scans and the fourth number is the AAAD.

In Table 5.2 there are a few patterns visible. The first observation is that when the buffer increases, the total number of rejected scans decreases for all cases. This is a logical result of the relationship between the buffer and the total capacity: increasing the buffer means that the total capacity will also increase, leading to more scans that can be accepted. Comparing case NT with case T, we observe that the algorithm with one trading epoch performs better in every way: the requested number of both type 1 and type 2 scans for case T is larger. From this, we can conclude that the possibility of trading reduces the waste of capacity. This can be caused by the trading process itself, but it could also be due to the belief of players that they will be able to buy capacity when they need it. However, the latter would only increase the number of requested type 2 scan candidates, and from the results in Table 5.2 it is clear that also the type 1 ratio improves in case T. Therefore, we can conclude that not only the possibility of trading improves the performance, but also that trading actually takes place and that it leads to less rejected scans, reducing the waste of capacity. The effect of trading is also visible in the AAADs: trading possibilities reduce the annual departmental difference. The effect of trading decreases when the buffer increases. This makes sense: when the buffer is increases, the probability that a department runs out of budget decreases. Thus, departments will buy less often. They might want to sell more, but trading only happens when there is another player to sell to. This happens less when the buffer is larger.

Value of type 1 scan

The next parameter that we consider is α, the value of a type 1 scan relative to the value of a type 2 scan. Until now $\alpha = 2$, but for our next experiment we set $\alpha = 5$. The individual policy for this new setting of parameters is the following:

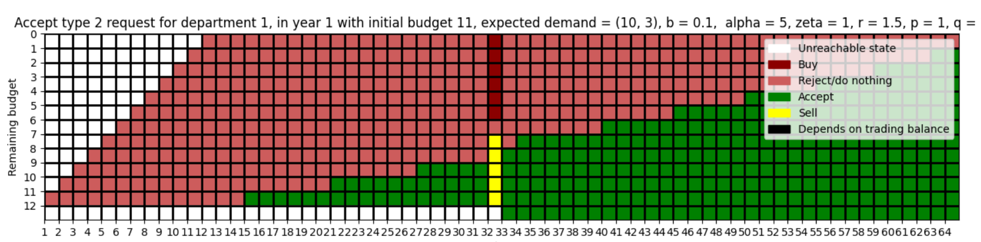

Figure 5.7: Individual policy with $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$, $b = 0.1$, $\alpha = 5$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$.

Compared with the policy with $\alpha = 2$ (Figure 5.2), the structure of the policies are similar, but the threshold for accepting a type 2 scan is much higher for $\alpha = 5$. Just before the trading epoch, type 2 scan candidates are accepted only when the remaining budget is greater than 8. For the policy with $\alpha = 2$, this value is 5. Also the trading policy is different: when the remaining budget is greater than 6, the optimal action is sell, and when the remaining budget is smaller than 6, the optimal action is buy. If the remaining budget is exactly 6, the optimal action is to do nothing. Thus, if α increases, the willingness of departments to buy increases. This can be explained by the risk of running out of budget. This risk increases if a department sells capacity. For increasing α, the penalty for rejecting type 1 scans also increases. Therefore, departments want to decrease this risk for larger values of α. We evaluate the collective behaviour for two symmetric departments with $\alpha = 5$. The resulting acceptance ratios are given in Table 5.3.

$\alpha \downarrow$	**Case NT**	**Case T**	**Case CA**
2	81.6 / 91.3 / 51.1 / 9.6	82.9 / 91.8 / 54.8 / 8.2	84.0 / 94.5 / 50.6
5	77.9 / 94.0 / 26.9 / 6.6	78.6 / 94.6 / 27.5 / 5.7	80.9 / 96.7 / 30.7
20	75.3 / 94.3 / 15.1 / 6.2	76.2 / 95.2 / 15.3 / 5.1	78.2 / 97.0 / 18.5
200	73.0 / 94.5 / 8.0 / 6.3	74.7 / 95.4 / 8.9 / 5.0	76.6 / 97.2 / 11.0

Table 5.3: Acceptance ratios for $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$ and different values of α. The first number is the total acceptance ratio, the second number is the acceptance ratio for type 1 scans, the third number is the acceptance ratio for type 2 scans and the fourth number is the AAAD.

This table shows results for the different cases that are similar to results in Table 5.2. Also for larger values of α, trading improves both the type 1 AR and the type 2 AR and the annual absolute difference between departments is smaller for case T. From this table, we can conclude that increasing α leads to a slightly larger AR for type 1, a smaller AR for type 2 scans and also a smaller general AR. Thus, increasing α mainly leads to a policy that rejects more type 2 scans, and does not guarantee that all type 1 scans are accepted. This can be explained by the sizes of $\mathbb{E}[D], \mathbb{E}[E]$ and the buffer b: even if a department's policy is to reject all type 2 scans regardless of the budget, the department may still run out of capacity because of type 1 scans.

Allocation rule parameter

In this section we consider the allocation rule parameter r. The penalty and reward for buying and selling are determined by the value of r. We consider a situation with buffer $b = 0.3$, because this gives more possibilities to tune r, because of the integer restriction of the allocated budget. We only consider Case T, as r is only relevant when departments have the opportunity to trade, which is not the case for Case NT and Case CA. Besides the acceptance ratios and $AAAD$, we also consider an additional performance measure: the number of trading epochs in which trading takes place. The performance measures are given in Table 5.4.

$r \downarrow$	**AR total / type 1 / type 2 / AAAD / used trading epochs (%)**
1	92.2 / 95.2 / 82.7 / 5.2 / 0.0
1.3	92.5 / 95.9 / 81.5 / 4.7 / 20.8
1.6	91.6 / 95.9 / 77.5 / 5.0 / 0.5
1.9	91.6 / 96.1 / 77.2 / 5.0 / 0.0

Table 5.4: Acceptance ratios and the used trading epochs for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $p = 1$, $q = 1$, $\sigma = 4$ and different values of ζ.

This table shows interesting results. Trading practically only happens when $r = 1.3$. In this case, the reward and penalty are 1 and -1, respectively. For $r = 1.6$ and $r = 1.9$, the rewards are 2 and 3 and the penalties are -2 and -3, respectively. For $r = 1$, there is no reward or penalty. From the table it is clear that trading does not happen in these cases. We zoom in on the individual policy for $r = 1.6$ in Figure 5.8.

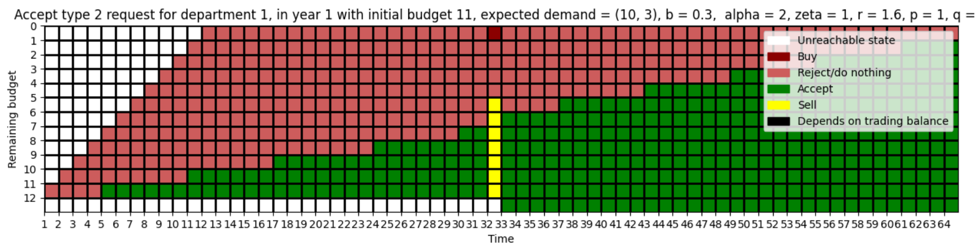

Figure 5.8: Individual policy with $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$, $b = 0.3$, $\alpha = 2$, $\zeta = 1$, $r = 1.6$, $p = 1$, $q = 1$.

This picture explains why trading almost never happens: a department only sells when it has ran out of budget at the trading epoch. Apparently, the penalty for buying (-2) is too high and is not worth the additional unit of budget that a department would receive with buying. This also explains the ARs in Table 5.4. In this table it can be seen that the type 1 acceptance ratios are as least as high and even higher for $r = 1.9$, although there is no trading for this value. This can be explained by the policy. Because the department does not buy at the trading, it cannot anticipate on receiving an extra unit of budget during the trading epoch as in the case with $r = 1.3$ (Figure 5.9). Therefore, the threshold for accepting type 2 scans is higher, which results in a lower AR for type 2 and a higher AR for type 1 in the end.

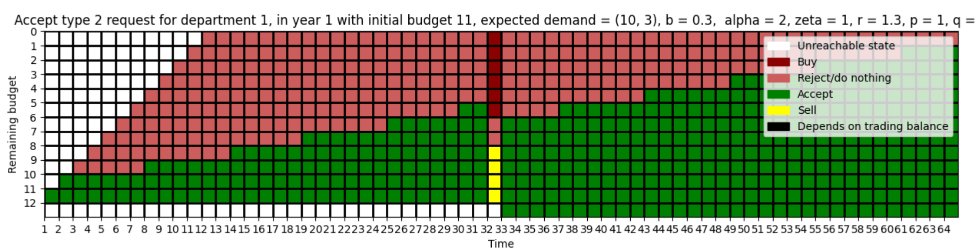

Figure 5.9: Individual policy with $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$, $b = 0.3$, $\alpha = 2$, $\zeta = 1$, $r = 1.3$, $p = 1$, $q = 1$.

Trading amount

In this section we consider the trading amount ζ. We only consider Case T, as ζ is only relevant when departments have the opportunity to trade. Besides the acceptance ratios, we also consider an additional performance measure: the number of trading epochs in which trading takes place. The performance measures are given in Table 5.5.

$\zeta \downarrow$	**AR total / type 1 / type 2 / AAAD / used trading epochs (%)**
1	82.9 / 91.8 / 54.8 / 8.4 / 23.5
2	83.8 / 89.3 / 65.8 / 9.5 / 5.8
3	83.7 / 88.5 / 68.4 / 9.7 / 0.2

Table 5.5: Acceptance ratios and the used trading epochs for $\mathbb{E}[D] = 10$, $\mathbb{E}[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $p = 1$, $q = 1$, $\sigma = 4$ and different values of ζ.

We see that for $\zeta > 1$, the number of used trading epochs decreases, and thus that less trading takes place. However, the AR for type 2 increases. To understand this, we look at the individual policy in Figure 5.10.

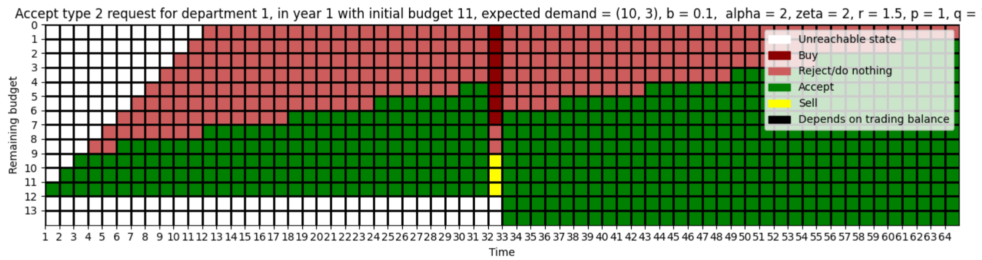

Figure 5.10: Individual policy with $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\alpha = 2$, $\zeta = 2$, $r = 1.5$, $p = 1$, $q = 1$.

With $\zeta = 2$, the department buys whenever its remaining budget is less than 7. Thus, in comparison to the policy for $\zeta = 1$ (Figure 5.2) where the department buys when the remaining budget is less than 5, the probability that a department buys budget is higher. In addition, the trading amount is also higher, and therefore the department anticipates on having a much higher budget after the trading epoch. Therefore the threshold for accepting type 2 scans is lower, which results in a higher AR for type 2. However, as trading rarely happens contrary to the department's belief, the AR for type 1 scans is lower.

Number of trading epochs

We consider the number of trading epochs σ. We set $\zeta = 1$ again and calculate the policies for $\sigma = 4$. The individual policy for this case is depicted in Figure 5.11.

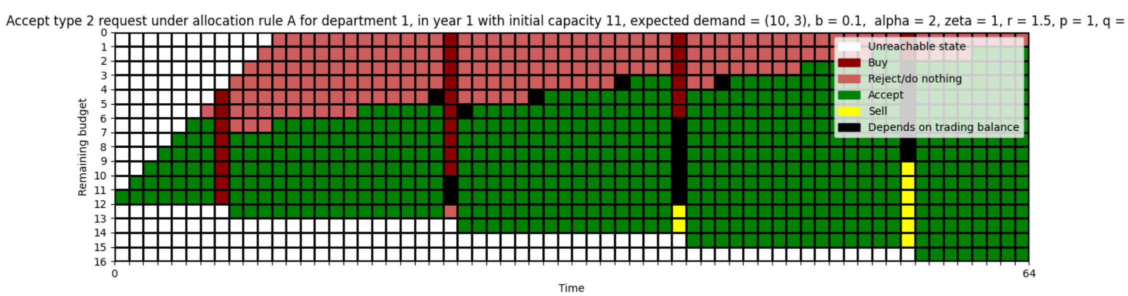

Figure 5.11: Individual policy with $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\alpha = 2$, $\zeta = 1$, $r = 1.5$, $p = 1$, $q = 1$, $\sigma = 4$.

This policy is different from the policies with one trading epoch in multiple aspects. The first thing that we observe is the acceptance of type 2 scans: the threshold is much lower than the threshold in the policy in Figure 5.2. This can be explained by the larger number of trading epochs and by the estimated trading probabilities. The probabilities are 1, which implies that departments receive an additional unit of scan budget when they buy during a trading epoch. With $\sigma = 4$, this means that departments 'believe' they are assured of receiving 4 extra units of budget per year. We also observe that there are a few states where the optimal action depends on the trading balance, especially in the third and fourth trading epoch. This can be explained by the simple nature of the allocation rule: departments receive a reward whenever the trading balance is greater than zero, and receive a penalty whenever the trading balance is smaller than zero. With four trading epochs, the trading balance can be any integer value between -4 and 4. So selling (buying) might be optimal when the trading balance is 0, but whenever the trading balance is smaller (greater) than zero, there is no additional reward (penalty) for selling (buying). Finally, we observe that selling only takes place in the third and fourth trading epochs. This is a logical result from the simple allocation rule: the exact amount that has been sold does not matter, as long as it is greater than zero. Therefore, selling during one epoch is enough to receive the reward.

Departments prefer the later trading epochs, because in those epochs the risk of having a larger demand and therefore losing more budget than expected is lower, as the number of remaining decision epochs is smaller for the later epochs.

We also investigate the influence of multiple trading epochs on the collective behaviour. The performance measures are given in Table 5.6.

$\sigma \downarrow$	**AR total / type 1 / type 2 / AAAD / used trading epochs (%)**
1	82.9 / 91.8 / 54.8 / 8.2 / 23.5
2	85.1 / 91.0 / 65.9 / 9.1 / 14.6
3	86.0 / 90.2 / 72.7 / 9.6 / 9.3
4	87.2 / 90.1 / 77.8 / 10.3 / 4.8

Table 5.6: Acceptance ratios and the used trading epochs for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $\zeta = 1$, $p = 1$, $q = 1$ and different values of σ. The first number is the total acceptance ratio, the second number is the acceptance ratio for type 1 scans, the third number is the acceptance ratio for type 2 scans and the fourth number is the percentage of trading epochs in which trading takes place.

For $\sigma > 1$, the total acceptance ratio is higher than for only one trading epoch. However, a remarkable fact is that this increase is not caused by a higher acceptance ratio for type 1 scans, because this acceptance ratio decreases in comparison to the case with one trading epoch. The increase is caused by the larger acceptance ratio for type 2 scans. There is an outstanding difference between the acceptance ratios for type 2 scans for the cases. This difference can be explained by the thresholds for type 2 acceptance decisions in Figure 5.11, which are caused by the high values of the estimated trading probabilities. With $p = q = 1$, departments have full confidence in trading, which results in less risk-averse behaviour when it comes to accepting type 2 scans. However, the true buying probability under the collective behaviour is actually be much smaller, especially in the early trading epochs. In a symmetric situation, both players will decide to buy in the first and second trading epoch, and therefore the true buying probabilities are zero at these epochs. However, both players anticipate on a successful trade, because their policies are based on the belief that both trading probabilities are one. Thus, their 'optimal' policies might not be optimal, for example because they accept too many type 2 requests and are forced to reject type 1 requests later in the year, because they did not get the additional capacity they anticipated on. This also explains the lower acceptance ratio for type 1 scans.

Trading probabilities

If we calculate the policy for $\sigma = 4$ and the true trading trading policies under the policy in Figure 5.11 $p = 0.15, q = 0.85$, the performance and policy is follows (Table 5.7, Figure 5.12):

$p, q \downarrow$	**AR total / type 1 / type 2 / AAAD / used trading epochs (%)**
1, 1	87.2 / 90.1 / 77.8 / 10.3 / 4.8
0.15, 0.85	85.8 / 93.7 / 61.0 / 6.5 / 8.1

Table 5.7

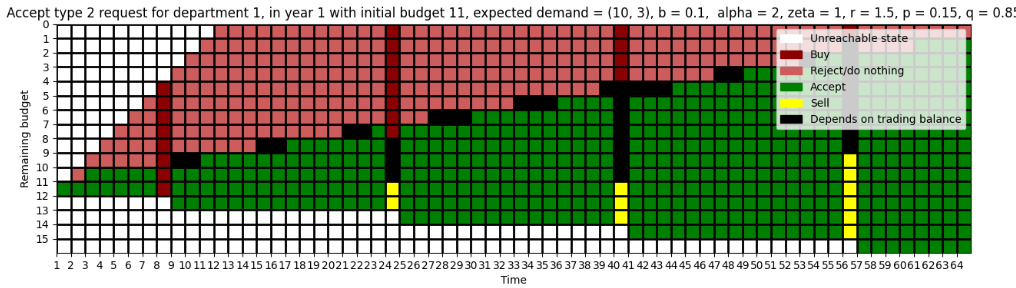

Figure 5.12: Individual policy with $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\alpha = 2$, $\zeta = 1$, $r = 1.5$, $p = 0.15$, $q = 0.35$, $\sigma = 4$.

For these estimated probabilities, the AR for type 1 is higher and trading takes place more often, resulting in a much smaller AAAD. This shows that accurate information about the true trading probabilities can significantly improve the general performance. In this subsection, we investigate the influence of the estimated trading probabilities on the individual policies and on the resulting collective behaviour and acceptance ratios. More specifically, we provide some insight in the 'cost' of estimation errors, that is, the difference between the acceptance ratios when the policies induced by the estimated probabilities are followed, and the acceptance ratios when the policies induced by the true trading probabilities are followed.

We start with studying the simple case we started with, with parameters $E[D] = 10$, $E[E] = 3$, $b = 0.1$, $\alpha = 2$, $\zeta = 1$, $r = 1.5$ and only one trading epoch for different values of p and q. The results are given in Table 5.8. In Table 5.9, the true trading probabilities for these cases are given. These can be interpreted as follows: if a department acts according to the policy calculated with estimated probabilities $p = 0.25$, $q = 0.25$, the probability that this department sells budget (i.e., the true buying probability for the other player) during any trading epoch is 0.58, and the probability that the department buys (i.e., the true selling probability for the other player) budget during any trading epoch is 0.24.

$p \downarrow q \rightarrow$	0.25	0.50	0.75
0.25	82.6 / 92.0 / 53.1 / 22.3	82.8 / 92.3 / 52.7 / 21.8	82.6 / 92.3 / 51.1 / 22.1
0.50	82.9 / 91.8 / 54.8 / 23.3	82.8 / 92.0 / 53.8 / 21.5	82.7 / 92.2 / 52.4 / 22.4
0.75	83.2 / 91.7 / 56.2 / 22.8	82.9 / 91.8 / 54.9 / 22.6	82.8 / 91.9 / 53.8 / 22.5

Table 5.8: Acceptance ratios and the used trading epochs for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $\zeta = 1$, $\sigma = 1$ and different values of p and q. The first number is the total acceptance ratio, the second number is the acceptance ratio for type 1 scans, the third number is the acceptance ratio for type 2 scans and the fourth number is the percentage of trading epochs in which trading takes place.

$p \downarrow q \rightarrow$	0.25	0.50	0.75
0.25	0.53, 0.24	0.54, 0.23	0.57, 0.21
0.50	0.51, 0.25	0.53, 0.24	0.55, 0.23
0.75	0.49, 0.26	0.51, 0.25	0.53, 0.24

Table 5.9: True trading probabilities for the other player (i.e., the first number refers to the probability that the other player can buy) for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $\zeta = 1$, $\sigma = 1$ and different values of p and q.

These tables show only small differences in the ARs and the number of used trading epochs. However, also in these small differences there are patterns visible: as p increases, the ARs for type 2 increase and the ARs for type 1 decrease. For q, the opposite holds. From the true trading probabilities in Table 5.9, we can conclude that the true trading probabilities are most likely close to $p = 0.5$ and $q = 0.25$. Thus, there is probably a trading equilibrium for $(0.5, 0.25, 0.5, 0.25)$. If we consider the performance measures for these equilibrium probabilities in Table 5.8, two things can be noticed. First, the trading is the highest for the trading equilibrium. Second, the AR for type 1 is not the highest for the trading equilibrium. This number is higher for larger estimated selling probabilities q and for smaller estimated buying probabilities p. We also consider the case with two trading epochs. The results for this case are given in Tables 5.10 and 5.11.

$p \downarrow q \rightarrow$	0.25	0.50	0.75
0.25	84.1 / 92.3 / 58.0 / 21.7	84.1 / 92.6 / 57.1 / 22.1	84.0 / 92.6 / 56.5 / 20.9
0.50	84.6 / 91.8 / 61.8 / 30.8	84.6 / 92.1 / 60.6 / 22.4	84.3 / 92.0 / 59.8 / 20.4
0.75	85.0 / 91.3 / 64.9 / 25.4	84.8 / 91.3 / 64.2 / 25.7	84.7 / 91.6 / 62.8 / 17.7

Table 5.10: Acceptance ratios and the used trading epochs for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $\zeta = 1$, $\sigma = 2$ and different values of p and q. The first number is the total acceptance ratio, the second number is the acceptance ratio for type 1 scans, the third number is the acceptance ratio for type 2 scans and the fourth number is the percentage of trading epochs in which trading takes place.

$p \downarrow q \rightarrow$	0.25	0.50	0.75
0.25	0.38, 0.34	0.26, 0.47	0.26, 0.47
0.50	0.36, 0.54	0.26, 0.51	0.25, 0.51
0.75	0.24, 0.71	0.23, 0.71	0.18, 0.71

Table 5.11: True trading probabilities for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $\zeta = 1$, $\sigma = 2$ and different values of p and q.

From these results, we can see that in a trading equilibrium, the values of p and q are probably close to $(0.25, 0.50$. In Table 5.11, we again observe a clear relation between p and q and the ARs for type 1 and type 2. Moreover, we also see that for the cases where the true trading probabilities are approximately $(0.25, 0.50)$, that is, policies with estimated probabilities $(0.25, 0.50), (0.25, 0.75), (0.50, 0.50), (0.50, 0.75)$, the policy with estimated trading probabilities $(0.25, 0.50)$ is indeed optimal in terms of the AR for type 1.

Finally, we perform a small iterative experiment for the last case. We start by calculating a policy with estimated trading probabilities $p = q = 1$. Then we evaluate the true trading probabilities under this policy, and use these true trading probabilities as the new estimated probabilities. We repeat this procedure until we arrive at an equilibrium.

$p, q \downarrow$	AR total / type 1 / type 2 / used trading epochs	$\tilde{p}, \tilde{q}$
1.0, 1.0	82.9 / 91.0 / 65.9 / 9.1 / 14.6	0.16, 0.73
0.16, 0.73	83.8 / 92.9 / 55.4 / 7.7 / 20.9	0.26, 0.46
0.26, 0.46	84.1 / 92.5 / 58.2 / 7.5 / 22.1	0.26, 0.48
0.26, 0.48	84.1 / 92.5 / 58.2 / 7.5 / 22.1	0.26, 0.48

Table 5.12: Acceptance ratios and the used trading epochs for $E[D] = 10$, $E[E] = 3$, $\beta = 0.1$, $\alpha = 2$, $r = 1.5$, $\zeta = 1$, $\sigma = 2$. The first number is the total acceptance ratio, the second number is the acceptance ratio for type 1 scans, the third number is the acceptance ratio for type 2 scans and the fourth number is the percentage of trading epochs in which trading takes place.

This iterative process shows several things. First, it shows that there is a trading equilibrium at $(0.26, 0.48, 0.26, 0.48)$ and that we are able to find this trading equilibrium in three iterations. We can of course not draw conclusions from this single case, but this might show that the number of subdomains for the cases with only two trading epochs is small, i.e., that there are only a few different policies that are optimal for different values of p and q. Tables 5.8 and 5.9 also show this for $\sigma = 1$. This would mean that finding trading equilibria using the Subdomain Backward Induction algorithm and Algorithm 3 can be done very efficiently for cases with small σ. As these algorithms were not implemented. this cannot be tested in this study, but this would be a promising result that deserves further research. It can also be observed that the policy with estimated trading probabilities $p = 0.16$ and $q = 0.73$ (second row) performs better in terms of the type 1 AR than the policy that should be optimal for these probabilities (last row). However, it can be seen that with the last policy trading happens more often, which also leads to a smaller AAAD. The AR for type 2 is also higher in the last row, which can be explained by the larger estimated success probability for buying compared to the case in the second row. Thus, we can conclude that having an 'incorrect' belief on the true trading probabilities does not necessarily lead to a decrease in the type 1 AR, as long as the department underestimates the success probability of buying. Underestimating the buying opportunities leads to more risk-averse behaviour, i.e., higher thresholds for the acceptance of type 2 scans. Overestimating the buying probability leads to an increase in the the type 2 AR and a decreasing type 1 AR, as shown in Figure 5.11 and Table 5.7.

Chapter 6

Conclusions and discussion

6.1 Conclusions

In this research, we have presented a model that can be used to analyze individual policies and collective behaviour under a scan budget system. We found that the acceptance and trading policy of an individual department is a threshold policy where optimal actions depend on the remaining budget and the time. Multiple parameters influence the threshold for accepting type 2 scans, but the most important ones are the value of a type 1 scan α and the estimates success probability of buying . Under the assumption that departments are self-interested and only motivated by budget (i.e., scans), a small reward and penalty in terms of budget is necessary to ensure that trading takes place. However, this reward and penalty should be close to the amount that is traded, because in case the reward and penalty are much higher, departments do not buy because the amount that is received by buying does not outweigh the penalty that is received for buying. Similarly, in case the penalty and reward are much lower than the trading amount, departments do not sell capacity, because the reward that is received for selling does not outweigh the amount of budget that is lost by selling. In both cases, trading does not happen. We have seen that a department's belief on the other department's willingness to trade affects the individual policy and therefore also the collective behaviour. If the success probability of buying is underestimated, this leads to more risk-averse behaviour, i.e., the type 1 AR increases, while the type 2 AR increases. If the success probability of buying is overestimated, the opposite happens. The influence of under- and overestimating of the selling probability on the acceptance is small. However, it can be seen that under- and overestimation of both trading probabilities affect the trading process and the average annual absolute departmental difference; these seem to be optimal for trading equilibria. In general, our experiments for the cases N, T and CA show that introducing the possibility of trading reduces the waste of capacity and improves the acceptance ratios for both type of scans. Moreover, trading leads to smaller average annual departmental differences, meaning that trading provides equity between departments. When multiple (> 3) trading epochs are used and estimated probabilities are close to the true trading probabilities, the acceptance ratios increase even more and approach the ratios under centralized admission (CA).

6.2 Discussion

6.2.1 Discussion of results

Because of the choice for a direct solution method, we can only calculate relatively small instances. The use of small instances can affect the stability of the results, because single events and decisions can have a large impact on the ARs. For example, if the total number of type 2 scans that arrive during a year is 3, the decision whether or not to accept a single scan makes a difference of 33 percentage point in the acceptance ratio. Therefore, including larger cases is desirable. Under the allocation rule that is used, there exist monotone policies that are optimal. This can be proved using Theorem 4.7.4. of [12]. This property can be exploited to solve larger instances.

Another disadvantage of small cases is that there are little possibilities for tuning the allocation rule. We require the allocated capacity to be at least $\mathbb{E}[D_i]$, at most $(1 + rb)\,\mathbb{E}[D_i]$ and integer-valued. So if $\mathbb{E}[D_i]$ is small, the difference between the minimum and maximum allocated amount is small, and different values of r could yield the same allocated amount, due to the integer restriction. Thus the allocation is very 'rough' for small cases; for larger cases, it is more sophisticated. However, this means that the performance of an allocation rule for small instances provides a lower bound for larger instances, because for larger cases the allocation rule can be (but does not need to be) refined. This also holds for the number of trading epochs. It is likely that larger numbers will improve the acceptance ratios.

In the current setup of the model and simulations, it is difficult to distinguish effects of the individual policy and effects of the collective behaviour on the acceptance ratios. For example, when the players are not in a trading equilibrium, it is unclear if smaller acceptance ratios are caused by having the 'wrong' belief, or that these are caused by the individual policy. That is, the current setting does not allow for evaluating the individual policies for different true trading probabilities. This could be studied by fixing a true trading probability and evaluate individual policies for this true trading probability, and vice versa.

6.2.2 Future research

To make the model useful in practice, a lot of things can be added and improved. First, a major assumption in the model is that the demand does not change over the years. This is not a realistic assumption, because in the real world, the demand is increasing. Therefore, the allocation mechanism in practice should not only be based on the trading balance, but also on the observed trends and on the expected demand for next year. Incorporating a prediction model in the current model would make the allocation rule more realistic. Moreover, different prices for different scans can be used to account for different levels of complexity of scans. The models of the individual policies can also be made more sophisticated by including for example seasonal influences on the demand, and imposing constraints on the differences between monthly acceptance ratios to avoid unfair treatment of different patients over time. If this is done, the individual policies could also be used in practice to advise departments on their policies under a scan budget.

The trading and collaboration can also be modeled differently. Currently, the assumption is that departments are only motivated by budget or capacity and therefore the budget allocation is the only tool that can be used to stimulate trading. This can be approached differently, for example if we assume that departments are also interested in maintaining good relationships with other departments, because they might benefit from these relationships by collaboration. This could be incorporated in the model by using a tie strength for the connection between two departments. This tie strength then influences the trading probabilities for these departments. This could be combined with the allocation rule that is used in our model.

Bibliography

[1] Zeinab Abtahi, Rashed Sahraeian, and Donya Rahmani. A stochastic model for prioritized outpatient scheduling in a radiology center. *International Journal of Engineering Transactions A: Basics*, 33(4), 2020.

[2] Anja Almén, Wolfram Leitz, and Sven Richter. National survey on justification of ct-examinations in sweden. 2009.

[3] Michael Bowling and Manuela Veloso. An analysis of stochastic game theory for multiagent reinforcement learning. Technical report, Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science, 2000.

[4] Lucian Buşoniu, Robert Babuška, and Bart De Schutter. Multi-agent reinforcement learning: An overview. *Innovations in multi-agent systems and applications-1*, pages 183–221, 2010.

[5] Xi Chen and Xiaotie Deng. Settling the complexity of two-player nash equilibrium. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 261–272. IEEE, 2006.

[6] Rodolfo Benedito Zattar da Silva, Flávio Sanson Fogliatto, André Krindges, and Moiseis dos Santos Cecconello. Dynamic capacity allocation in a radiology service considering different types of patients, individual no-show probabilities, and overbooking. *BMC health services research*, 21(1):1–24, 2021.

[7] Jie Gao, Terrence Wong, and Chun Wang. Coordinating patient preferences through automated negotiation: a multiagent systems model for diagnostic services scheduling. *Advanced Engineering Informatics*, 42:100934, 2019.

[8] Linda V Green, Sergei Savin, and Ben Wang. Managing patient service in a diagnostic medical facility. *Operations Research*, 54(1):11–25, 2006.

[9] Leon Hermans, Scott Cunningham, and Jill Slinger. The usefulness of game theory as a method for policy evaluation. *Evaluation*, 20(1):10–25, 2014.

[10] Heljä Oikarinen, Salme Meriläinen, Eija Pääkkö, Ari Karttunen, Miika T Nieminen, and Osmo Tervonen. Unjustified ct examinations in young patients. *European radiology*, 19(5):1161–1165, 2009.

[11] Jonathan Patrick, Martin L Puterman, and Maurice Queyranne. Dynamic multipriority patient scheduling for a diagnostic resource. *Operations research*, 56(6):1507–1525, 2008.

[12] Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons, 2014.

[13] James W Ryan, Aoife Hollywood, Aaron Stirling, Martina Glynn, Peter J MacMahon, and Ferdia Bolster. Evidenced-based radiology? a single-institution review of imaging referral appropriateness including monetary and dose estimates for inappropriate scans. *Irish Journal of Medical Science (1971-)*, 188(4):1385–1389, 2019.

[14] William Thomson. Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: a survey. *Mathematical social sciences*, 45(3):249–297, 2003.

[15] Ch Triantopoulou, I Tsalafoutas, P Maniatis, D Papavdis, G Raios, I Siafas, S Velonakis, and E Koulentianos. Analysis of radiological examination request forms in conjunction with justification of x-ray exposures. *European journal of radiology*, 53(2):306–311, 2005.

[16] HJ Valkenburg. *Exploratory research on the effects of strategic behaviour in long-term defence planning: a game-theoretic approach*. PhD thesis, Eindhoven University of Technology, 2022.

[17] Yaodong Yang and Jun Wang. An overview of multi-agent reinforcement learning from game theoretical perspective. *arXiv preprint arXiv:2011.00583*, 2020.

[18] Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of Reinforcement Learning and Control*, pages 321–384, 2021.

[19] Liping Zhou, Na Geng, Zhibin Jiang, and Xiuxian Wang. Public hospital inpatient room allocation and patient scheduling considering equity. *IEEE Transactions on Automation Science and Engineering*, 17(3):1124–1139, 2019.

[20] Liping Zhou, Zhibin Jiang, Na Geng, and Xiuxian Wang. Capacity allocation for balancing cost and fairness in a imaging facility setting. In *2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 721–725. IEEE, 2015.

[21] Maartje E Zonderland and Judith Timmer. Optimal allocation of mri scan capacity among competing hospital departments. *European journal of operational research*, 219(3):630–637, 2012.