

PRICING AND EQUILIBRIUM ANALYSIS OF NETWORK MARKET SYSTEMS

A Dissertation

by

VAMSEEDHAR REDDY REDDYVARI RAJA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Srinivas Shakkottai
Co-Chair of Committee,	P. R. Kumar
Committee Members,	Jean-Francois Chamberland
	Natarajan Gautam
Head of Department,	Miroslav M. Begovic

May 2018

Major Subject: Electrical Engineering

Copyright 2018 Vamseedhar Reddy Reddyvari Raja

ABSTRACT

Markets have been the most successful method of identifying value of goods and services. Both large and small scale markets have gradually been moving into the Internet domain, with increasingly large numbers of diverse participants. In this dissertation, we consider several problems pertaining to equilibria in networked marketplaces under different application scenarios and market sizes. We approach the question of pricing and market design from two perspectives. On the one hand, we desire to understand how self-interested market participants would set prices and respond to prices resulting in certain allocations. On the other hand, we wish to evaluate how best to allocate resources so as to attain efficient equilibria. There might be a gap between these viewpoints, and characterizing this gap is desirable.

Our technical approaches follow the number of market participants, and the nature of trades happening in the market. In our first problem, we consider a market of providing communication services at the level of providing Internet transit. Here, the transit Internet Service Provider (ISP) must determine billing volumes and set prices for its customers who are firms that are content providers, sinks, or subsidiary ISPs. Demand from these customers is variable, and they have different impacts on the resources that the transit ISP needs to provision. Using measured data from several networks, we design a fair and flexible billing scheme that correctly identifies the impact of each customer on the amount of provisioning needed.

While the customer set in the first problem is finite, many marketplaces deal with a very large number of agents that each have ephemeral lifetimes. Here, agents arrive, participate in the market for some time, and then vanish. We consider two such markets in such a regime. The first is one of apps on mobile devices that compete against each other for cellular data service, while the second is on service marketplaces wherein many providers compete with each other for jobs that consider both prices and provider reputations while making choices

between them. Our goal is to show that a Mean Field Game can be used to accurately approximate these systems, determine how prices are set, and characterize the nature of equilibria in such markets.

Finally, we consider efficiency metrics in large scale resource sharing networks in which bilateral exchange of resources is the norm. In particular, we consider peer-to-peer (P2P) file sharing under which peers obtain chunks of a file from each other. Here, contrary to the intuition that chunks must be shared whenever one peer has one of value to another, we show that a measure of suppression is needed to utilize resources efficiently. In particular, we propose a simple and stable algorithm entitled Mode suppression that attains near optimal file sharing times by disallowing the sharing of the most frequent chunks in the system.

DEDICATION

To my wife and parents

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Dr. Srinivas Shakkottai for his continuous guidance and motivation throughout my PhD studies. It would not have been possible to complete this work without his help. I would like to thank my main collaborators, Dr. Amogh Dhamdhere, Dr. Vijay Subramanian and Dr. Parimal Parag for their invaluable guidance and feedback. Their ideas gave me direction whenever I was lost. Finally, I would like to thank my committee members for their constructive feedback and the many concepts that I learned from their courses.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Srinivas Shakkottai, P. R. Kumar and Jean-Francois Chamberland of the Department of Electrical Engineering and Professor Natarajan Gautam of the Department of Industrial Engineering.

The analyses depicted in Chapter 3 was conducted in part by Vinod Ramaswamy and Mayank Manjrekar of the Department of Electrical Engineering.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a Research Assistantship from Texas A&M University.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
CONTRIBUTORS AND FUNDING SOURCES	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	xi
LIST OF TABLES	xiii
1. INTRODUCTION	1
2. ISP TRANSIT BILLING	4
2.1 Introduction	4
2.2 Datasets	6
2.2.1 SWITCH dataset	6
2.2.2 IXP dataset	6
2.3 Longitudinal Study of 95 th Percentile billing	7
2.3.1 Calculation of 95th percentile	7
2.3.2 Classification of networks	7
2.3.3 95 th percentile to average ratio	8
2.3.4 Skewness of the traffic distribution	9
2.4 Fairness of 95 th Percentile Billing	13
2.4.1 Shapley Value Percentile Billing	13
2.4.2 Overlap rank	15
2.4.3 Provision ratio	16
2.4.4 Towards a new billing mechanism	17
2.5 Measuring Billing Volumes	18
2.6 Optimization Framework	19
2.7 Conclusions	26
3. A MEAN FIELD GAME APPROACH TO QOE-AWARE CELLULAR SCHEDULING	28
3.1 Introduction	28

3.1.1	Mean Field Games.....	30
3.1.2	Overview of Chapter	31
3.2	Mean Field Game Model.....	32
3.2.1	Auction System.....	33
3.2.2	Queueing System	34
3.2.3	Mean Field Equilibrium.....	34
3.2.4	Agent's decision problem.....	35
3.2.5	Invariant distribution.....	37
3.2.6	Mean field equilibrium	37
3.3	Properties of Optimal Bid Function.....	38
3.4	Existence of MFE.....	42
3.5	MFE Existence: Proof.....	43
3.5.1	Continuity of the map \mathcal{F}	43
3.5.2	$\mathcal{F}(\mathcal{P})$ contained in a compact subset of \mathcal{P}	48
3.6	Approximation Results: PBE and MFE	50
3.7	Simulation Results	52
3.8	Conclusions.....	54
4.	MEAN FIELD EQUILIBRIA OF PRICING AND WORK-QUALITY SELECTION GAMES IN INTERNET MARKETPLACES.....	56
4.1	Introduction.....	56
4.1.1	Mean Field Games.....	59
4.1.2	Main Contributions.....	60
4.2	Mean Field Model	62
4.2.1	State Space of Agent	62
4.2.2	State Transitions	64
4.2.3	Bid, Payment and Cost	66
4.2.4	Effective Bid and Customer Choice	66
4.3	State Dynamics and MFE.....	66
4.3.1	Agent's Decision Problem.....	67
4.3.2	Invariant Distribution	69
4.3.3	Mean Field Equilibrium.....	69
4.4	Existence of Optimal Policy	70
4.5	Properties of Optimal Policy.....	72
4.6	Existence of Mean Field.....	75
4.6.1	General Auction	75
4.6.2	Continuity of the map \mathcal{F}	77
4.6.3	Continuity of \mathcal{F}_A	79
4.6.4	Compactness of Range Space \mathcal{F}_S	80
4.7	Non-Uniform Prior.....	81
4.8	Numerical Experiments	83
4.8.1	Determining z from Data	83
4.8.2	Large Particle Simulations	85
4.8.3	Strategic η_c	88

4.9	Conclusion.....	88
5.	PROVABLE STABLE POLICIES IN P2P FILE SHARING.....	91
5.1	Introduction.....	91
5.1.1	Related Work on Stable Algorithms.....	93
5.1.2	Main Results.....	94
5.2	System Model.....	96
5.2.1	Contact Processes.....	97
5.2.2	State space.....	97
5.2.3	State transitions.....	98
5.3	Mode Suppression policy.....	99
5.3.1	Stability Region of Mode Suppression Policy.....	100
5.4	Threshold Mode Suppression.....	104
5.4.1	Rate Transition Matrix (Q) of TMS.....	105
5.4.2	Stability of TMS.....	106
5.5	Simulation Results.....	109
5.5.1	Stability of Mode Suppression Policy.....	110
5.5.2	Missing Piece Syndrome in Random Chunk Selection.....	110
5.5.3	Chunk Frequency Evolution.....	111
5.5.4	Sojourn times.....	112
5.6	Conclusions.....	113
6.	CONCLUSIONS.....	116
	BIBLIOGRAPHY.....	118
	APPENDIX A. PROOFS FROM CHAPTER 3.....	124
A.1	Proof of Lemma 2.....	124
A.2	Proofs from Section 3.5.....	126
A.2.1	Proofs Pertaining to Section 3.5-A: Step 1.....	126
A.2.2	Proofs Pertaining to Section 3.5-A: Step 2.....	127
A.2.3	Proofs Pertaining to Section 3.5-A: Step 3.....	128
A.2.4	Proofs Pertaining to Section 3.5-B.....	130
	APPENDIX B. PROOFS FROM CHAPTER 4.....	131
B.1	Proofs from Section 4.4.....	131
B.1.1	Details of proof of Lemma 10.....	131
B.1.2	Existence of optimal policy.....	132
B.1.3	Details of Proof of Theorem 12.....	134
B.2	Proofs from Section 4.5.....	136
B.2.1	Details of Proof of Lemma 11.....	137
B.2.2	Details of Proof of Lemma 12.....	138
B.2.3	Details of Proof of Lemma 13.....	140

B.2.4	Details of Proof of Lemma 15	142
B.3	Proofs from Section 4.6.....	145
B.3.1	Details of Proof of Lemma 18	145
B.3.2	Details of Proof of Lemma 19	146
B.3.3	Details of Proof of Lemma 21	147
B.3.4	Details of Proof of Lemma 22	148
B.3.5	Details of Proof of Lemma 23	149
B.3.6	Details of Proof of Lemma 26	150
B.4	Proofs from Section 4.7.....	152
B.4.1	Details of Proof of Lemma 29	152
APPENDIX C. PROOFS FROM CHAPTER 5.....		153
C.1	Auxiliary Results	153
C.2	Proof of Theorem 17.....	155
C.2.1	Choosing Parameters:	158

LIST OF FIGURES

FIGURE	Page
2.1 Mean 95 th Percentile to average ratio for different network types in the SWITCH dataset.	9
2.2 Mean 95 th Percentile to average ratio for IXPs, using different network classifications.	10
2.3 Mean skewness for different network types in the SWITCH dataset. Heavy-outbound networks have a higher skewness, especially in the last 4 years.	11
2.4 Mean skewness for networks in the IXP dataset.	12
2.5 Shapley value percentiles: SWITCH dataset (Mar 2012) and IXP dataset (SIX, Aug 2013).	14
2.6 Overlap rank vs 95 th Percentile rank for IXP dataset (Aug 2013) and one month of SWITCH dataset (Jan 2012). A large fraction of networks lie far from the diagonal, meaning they have a large billing volume but little overlap with the provider’s peaks, or vice versa.	16
2.7 Inverse Provision Ratio order vs Average traffic order of SVPs and OWPs for February 2012 SWITCH data.	23
2.8 Distribution of percentage of orders preserved and Box and Whisker plot of difference in ranks for four years.	24
2.9 Percentage of amount of billing volume reduction with $\mathcal{E} = 5\%$ elasticity. The x-axis denote the index of the customers and y-axis denote the percentage reduction achieved by moving 5% of traffic to off-peak periods.	25
2.10 Percentage of amount of billing volume reduction with $\mathcal{E} = 15\%$ elasticity. The x-axis denote the index of the customers and y-axis denote the percentage reduction achieved by moving 5% of traffic to off-peak periods.	26
3.1 The game consists of an auction part (blue/dark tiles) and a queue dynamics part (beige/light tiles). The system is at MFE if the distribution of the bid X is equal to the assumed bid distribution ρ	33
3.2 Convergence to MFE bid distribution.	52
3.3 MFE queue length distribution.	52

3.4	MFE optimal bid function.....	52
4.1	The effect of reputation on sale price. Reputation is classified as “high” (=1) or “low” (=0). We consider four conditions of phones: fair, good, mint and new. Comparing phones of the same condition (shown in the same color), we observe a markup of about \$50 for a high reputation.	57
4.2	Mean Field Model: Parameters related to the mechanism are shown in blue, while the state transition parameters are shown in red.	63
4.3	Value Function variation with queue size and reputation.	85
4.4	Variation of type with queue size, reputation and quality.....	86
4.5	Optimal effective bid variation with queue size and reputation.	87
4.6	η_c chosen at different states.	88
4.7	η_c distribution with time.	89
5.1	Number of peers in the system when $m = 5$, $U = 1$ and $\mu = 1$. Random becomes unstable in some cases, whereas MS and DMS are always stable.	109
5.2	Evolution of peers and chunk frequencies under the random chunk selection policy. One of the chunks becomes a “missing chunk” (red/starred line).....	111
5.3	Chunk frequency evolution in a system with $m = 5$ chunks under different policies when starting from the state of a “missing-chunk” (whose frequency is indicated by a red/dashed line). Rarest-first is clearly unstable, since it cannot recover, whereas the other protocols manage to bring the chunk back into peer circulation and stabilize the system.....	112
5.4	Mean mixing sojourn times of policies for different values of m . Distributed mode suppression has the best performance in all cases.	113
5.5	Mean stationary sojourn times of policies for different values of m . Distributed mode suppression has the best performance in all cases.	114

LIST OF TABLES

TABLE	Page
4.1 Value of \bar{z}_κ for different products.	84
5.1 Comparison of global chunk distribution based policies.	96

1. INTRODUCTION

In recent years, there has been a proliferation of different kinds of markets operating over a range of timescales and catering to a variety of resources and services. The efficient functioning of these markets depends on the properties of the mechanisms that determine how users interact, and what kind of equilibria are attained. However, analyzing the impact of policies and equilibrium achieved in these markets is a challenging task, especially when the number of users or agents are large. In this thesis, we analyze mechanism design and equilibrium properties of four different markets, each with an increasing numbers of users.

The first problem we look at is the Internet transit service provider market, in which large transit providers such as AT&T and Level 3, provide transit services to level 2 customers. The 95th percentile method for calculating a customer's billable transit volume has been the industry standard used by such transit providers for over a decade due to its simplicity. Here, the average amount of bandwidth used is computed for every 5 minutes for the entire month and then the 95th Percentile of all these entries is considered as the bandwidth utilization of that ISP. In this chapter, we evaluate the 95th Percentile pricing mechanism from the perspective of transit providers, using a decade of traffic statistics from a European transit provider, SWITCH, and more recent traffic statistics from 3 Internet Exchange Points (IXPs). We find that over time, heavy-inbound and heavy-hitter networks are able to achieve a lower 95th-to-average ratio than heavy-inbound and moderate-hitter networks, possibly due to their ability to better manage their traffic profile. We also show that 95th Percentile does not reflect a customer's contribution to the provider's peak load. We discuss how a fair allocation of costs could potentially be achieved based on determining each network's Shapley value – a concept from cooperative game theory. However, computing the Shapley value has exponential complexity. We then propose a new transit billing optimization framework that is fair, flexible and computationally inexpensive. Our approach is based on the *Provision Ratio*, a metric that estimates the contribution of a customer to the

provider's peak traffic and acts as a proxy for the Shapley value. The proposed mechanism can be framed as convex optimization problem, and has fairness properties similar to the optimal (in terms of fairness) Shapley value allocation, with a much smaller computational complexity.

The second problem we deal with is the auction-based packet-scheduling in cellular networks. We study auction-theoretic scheduling in this setting using the idea of a mean field equilibrium (MFE). Here, agents model their opponents through an assumed distribution over their action spaces, and play the best response action against this distribution. We say that the system is at MFE if this best response action turns out to be a sample drawn from the assumed distribution. In our setting, the agents are smart phone apps that generate packets that require, have costs associated with packet queueing delay, and bid against each other for service from base stations. The users of the apps spend a geometrically distributed amount of time on each app, and then move on to another. We show that in a system in which we conduct a second-price auction at each base station and schedule the winner at each time, there exists an MFE that will schedule the user with highest value at each time. We further show that the scheme can be interpreted as a longest-queue-first type policy. The result suggests that auctions can implicitly attain the same desirable results as queue-length based scheduling. We also present results on the convergence between a system with a finite number of agents to a mean field case as the number of agents become large. Finally, we show simulation results illustrating the simplicity of computation of the MFE in our setting.

Our next problem deals with price setting in Internet marketplaces, which have become a popular way of selling goods and services. Examples include current e-commerce systems such as Amazon Mechanical Turk, Hackers list, and Swappa, which specialize in different niches. We model such an Internet marketplace using a set of servers that choose prices for performing jobs. Each server has a queue of unfinished jobs, and is penalized for delay by the market maker via a holding cost. Each server is allowed to choose the work-quality when performing a job, with a job done at a higher work-quality than its inherent value

incurring a cost and jobs truthfully report the “quality” with which they were completed. The best estimate of quality based on these reports is the “reputation” of the server. A server bases its pricing decision on the distribution of its competitors prices and reputations. An entering job chooses the best server based on a combination of price and reputation. We seek to understand how prices would be determined in such a marketplace using the idea of the mean field equilibrium. We show the existence of an MFE and characterize the impact of reputation in allowing servers to declare higher prices than their competitors. We illustrate our results by a numerical study of the system via simulation with parameters chosen from data gathered from existing Internet marketplaces.

Our final problem is on analyzing Peer-to-Peer (P2P) networks, whose ability to scale throughput up in proportion to the arrival rate of peers has recently been shown to be crucially dependent on the chunk sharing policy employed. Some policies can result in low frequencies of a particular chunk, known as the missing chunk syndrome, which can dramatically reduce throughput and lead to instability of the system. For instance, commonly used policies that nominally “boost” the sharing of infrequent chunks such as the well-known rarest-first algorithm have been shown to be unstable. Recent efforts have largely focused on the careful design of boosting policies to mitigate this issue. We take a complementary viewpoint, and instead consider a policy that simply prevents the sharing of the most frequent chunk(s). Following terminology from statistics wherein the most frequent value in a data set is called the mode, we refer to this policy as mode suppression. We prove the stability of this algorithm using Lyapunov techniques. We also design a distributed version that suppresses the mode via an estimate obtained by sampling three randomly selected peers. We show numerically that both algorithms perform well at minimizing total download times, with distributed mode suppression outperforming all others that we tested against.

2. ISP TRANSIT BILLING

2.1 Introduction

Transit providers are an important piece of the Internet ecosystem, providing customers with access to the rest of the Internet. But the future role of transit providers is uncertain, given continuously falling transit prices and increased propensity for networks to interconnect directly (peering) [42, 14], essentially routing around traditional transit providers. These business risks increase the pressure on transit providers to optimize their transit billing schemes to remain competitive.

There are two components to today’s Internet transit billing scheme: the volume of traffic for which a customer network is billed (the *billing volume*), and a function that computes price based on this volume. The industry standard for determining the billing volume is the *95th percentile* method [44, 15]: a transit provider measures the utilization of a customer link in 5-minute bins throughout a month, and then computes the *95th Percentile* of these values as the billing volume. The *95th Percentile* method has three attractive properties: it is simple to implement; it uses data that the provider typically already collects; and it approximates the load a customer imposes on the provider’s network while forgiving a few anomalous traffic bursts. An important aspect of the second billing component (the pricing function) is that providers generally offer volume discounts, such that the per-bit price decreases as billing volume increases [42].

In this chapter, we first visit the *95th Percentile* billing scheme from the perspective of a provider, to investigate whether this scheme approximately achieves its intended objective of providing an easy-to-compute approximation of a customer’s traffic load to the provider. We first use 10 years of historical data from SWITCH, a Swiss research/academic network, and more recent data from 3 Internet Exchange Points (IXPs) to investigate how the *95th Percentile* of a customer’s traffic relates to: (1) its total traffic volume, (2) its nature as a

predominantly inbound/outbound customer, and (3) its behavior as a heavy vs. moderate hitter. Second, we study the *fairness* of the 95th percentile scheme, and define a new metric called the *provision ratio* to investigate the relationship between the 95th Percentile of customer and the contribution of that customer to the provider’s traffic load.

Analysis of these data sets reveals evidence that over the years the customers with a predominantly outbound traffic profile are able to maintain a lower 95th-to-average ratio than predominantly inbound customers, meaning that they have a lower billing volume for the same amount of traffic sent. Furthermore, the 95th-percentile pricing mechanism is *unfair*, because for many customers the 95th Percentile may not reflect their cost burden to the provider, as there is little overlap between the customer’s peak and the overall (provider) peak traffic. Our results motivate the need to look for alternatives to the 95th Percentile billing method that can better approximate a customer’s cost burden to the provider without adding too much additional measurement or computational overhead.

We then present a framework for determining the billing volume for each customer in a manner that is fair, computationally inexpensive, and flexibly allows the provider to provide incentives (discounts) to certain customers. While solutions such as the Shapley value method exist to assign billing volumes to customers in a fair manner, they are computationally too expensive to implement at scale without approximations. Further, those methods are not flexible enough to accommodate all the constraints of transit providers, e.g., restricting billing percentiles to a certain range or offering incentives to certain classes of customers. Our billing framework is based on a new metric called the *Provision Ratio*, which reflects a customer’s contribution to the provider’s peak traffic load. By assigning billing volumes per-customer, providers can exercise fine-grained control over their billing and provide discounts to customers that contribute minimally to the provider’s peak traffic. The transit provider can use such incentives as a means for attracting new customers.

2.2 Datasets

2.2.1 SWITCH dataset

Our first dataset comes from SWITCH, a Swiss Research/Academic network which provides Internet connectivity to major universities and organizations in Switzerland. Currently, SWITCH connects about 50 research and education sites, acting as a transit provider for traffic that originates or is destined to those networks. SWITCH also provides connectivity to the public Internet via commercial providers, and hosts content caches of two large content providers. For traffic billing, SWITCH measures the utilization of each border router interface in both inbound and outbound directions in 5-minute intervals. To present a longitudinal analysis, we use historical datasets from SWITCH from January 2003 to December 2012.

2.2.2 IXP dataset

The second dataset consists of traffic statistics published by 2 Internet Exchange Points (IXPs) – Budapest Internet Exchange (BIX), and Interlan Internet Exchange (ILAN) These IXPs publish MRTG graphs with 5-minute utilization (inbound and outbound) for each network connected to the public peering fabric of the IXP. We collected these graphs every day for 8 months from August 2013 to March 2014 for BIX and 5 months for ILAN and used Optical Character Recognition tools [10] to parse them. BIX had 62 networks connected to its public peering fabric, while SIX and ILAN had 48 and 55 networks, respectively. Networks connect to IXPs to create (settlement-free) peering connections with other participating networks, and so the traffic statistics we see at an IXP are for a connected network’s peering traffic¹. Castro et al. [10] showed that transit traffic and peering traffic have similar diurnal patterns and peak-to-valley ratios; in fact, the transit traffic for a network can be well-approximated as a multiplicative factor of the peering traffic. In our analysis we consider the IXP as proxy for a transit provider, and the networks connected to it as its customers.

¹While not explicitly disallowed, transit sale over the shared IXP fabric is rare [43]

2.3 Longitudinal Study of 95th Percentile billing

We first describe two common methods of computing the 95th Percentile traffic volume, and how the two methods can treat customers differently. We then classify networks based on two criteria: (i) major direction of traffic (inbound, outbound, and balanced); and (ii) volume of traffic (heavy-hitter and moderate-hitter), and present a longitudinal view of the traffic properties of these network types.

2.3.1 Calculation of 95th percentile

Although 95th Percentile billing is the industry standard, there are two common implementations and several possible variations. The first method measures the inbound and outbound traffic in every 5 minutes over the month, calculates the 95th percentile for each direction, and uses the maximum of these two values. Most transit provider references to computing the 95th Percentile use this method, e.g., [18, 4], so we use it in our subsequent analysis. The second method records the maximum of inbound and outbound traffic in each five minute interval, and calculates the 95th Percentile value from the resulting data set. This second method seems to be less common although we found a few transit providers that bill using this method [1, 12]. The second method will yield a value greater than or equal to the first method, and the results will differ significantly for customers with balanced traffic profiles, but with inbound peaks occurring at different times from outbound peaks. We computed the 95th Percentile for each network in the SWITCH dataset over 10 years. We found that the median ratio of the 95th Percentile value for each network, computed using these two methods is close to 1, but the widest difference induces a 20% higher transit bill using the second method.

2.3.2 Classification of networks

Direction of Traffic: We divide networks into three categories based on the dominant direction of traffic. For each network, we measure the traffic that terminates within that network (inbound) and traffic that originates from that network (outbound). If the inbound

traffic of the network is more than twice the outbound traffic we classify it as *heavy-inbound*, and if the outbound traffic is more than twice the inbound traffic we classify the network as *heavy-outbound*. Networks that do not satisfy either condition are classified as *balanced*. Typically, content providers are heavy-outbound, while eyeball providers are heavy-inbound.

Volume of Traffic: We next classify networks based on the volume of traffic they generate/consume over a month into *heavy-hitter* and *moderate-hitter* networks. To define the two classes we evaluated the traffic contribution by the top 20% of networks in each month of the SWITCH and IXP datasets. The top 20% of networks consistently contributed between 80 and 90% of total traffic in the SWITCH dataset, and 75% of total traffic in the IXP dataset. Based on this observation, we classify the top 20% of networks in each month as *heavy-hitter networks* and the rest as *moderate-hitter networks*.

2.3.3 95th percentile to average ratio

For each customer network, we first evaluate the 95th Percentile to average traffic ratio; the average reflects the total volume of traffic, whereas the 95th Percentile value gives an idea of the peak, and is also the traffic volume for which the customer is billed. If the two significantly differ, it suggests that the customer is paying primarily for its burstiness. Figure 2.1 shows the mean of the 95th Percentile to average traffic ratio over time for networks in the SWITCH dataset classified by traffic direction and traffic volume.

First, we observe that the 95th Percentile to average ratio has been fairly stable over the years for each type of network, despite the dramatic changes in overall inter-domain traffic patterns that have occurred during the same time. In the last 4 years, the mean ratio for heavy-outbound networks is between 2 and 3, while the mean for heavy-inbound networks is between 3.25 and 4. For balanced networks, the ratio is less than 3.25. Hence, heavy-inbound networks in general have higher 95th Percentile traffic compared to heavy-outbound or balanced networks for the same average traffic. Consequently, heavy-inbound networks have a higher billing volume than heavy-outbound networks for the same amount of total traffic sent. We observe that the mean ratio is between 2.25 and 3 for heavy-hitter networks,

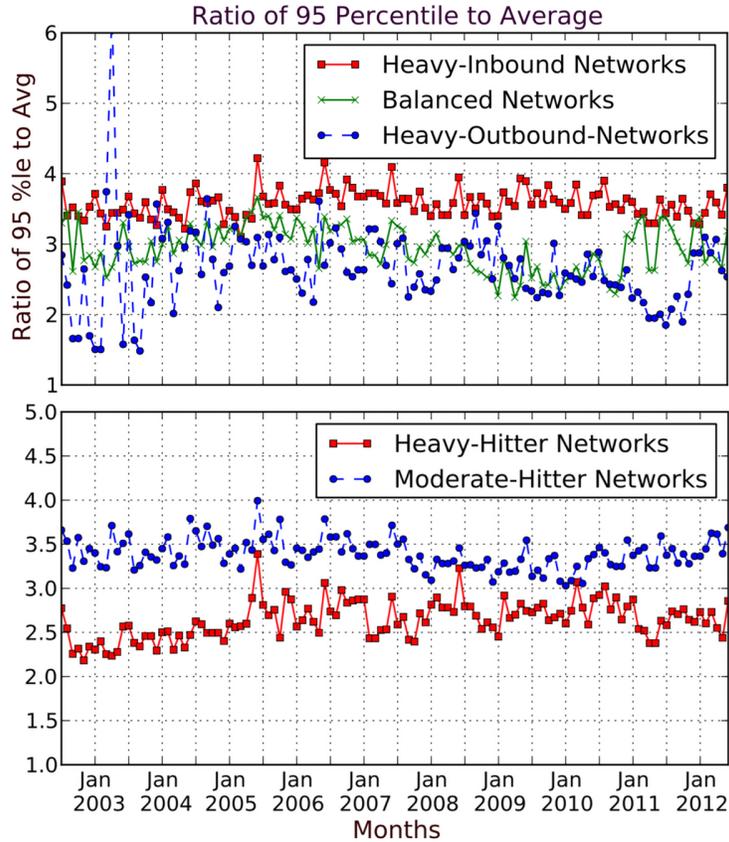


Figure 2.1: Mean 95th Percentile to average ratio for different network types in the SWITCH dataset.

especially in the last 4 years. However, the mean ratio always exceeds 3 for moderate-hitter networks in those 4 years.

Figure 2.2 shows the mean 95th Percentile to average ratio for different classes of networks in the IXP dataset. We observe that the mean ratio is higher for heavy-inbound networks than for heavy-outbound networks, consistent with our analysis of the SWITCH dataset. With the exception of BIX, the mean 95th Percentile to average ratio for networks at the other two IXPs is larger for moderate-hitter networks than for heavy-hitter networks, meaning that moderate-hitter networks have a burstier traffic profile than heavy-hitter networks.

2.3.4 Skewness of the traffic distribution

The above analysis shows that heavy-inbound and moderate-hitter networks have a higher

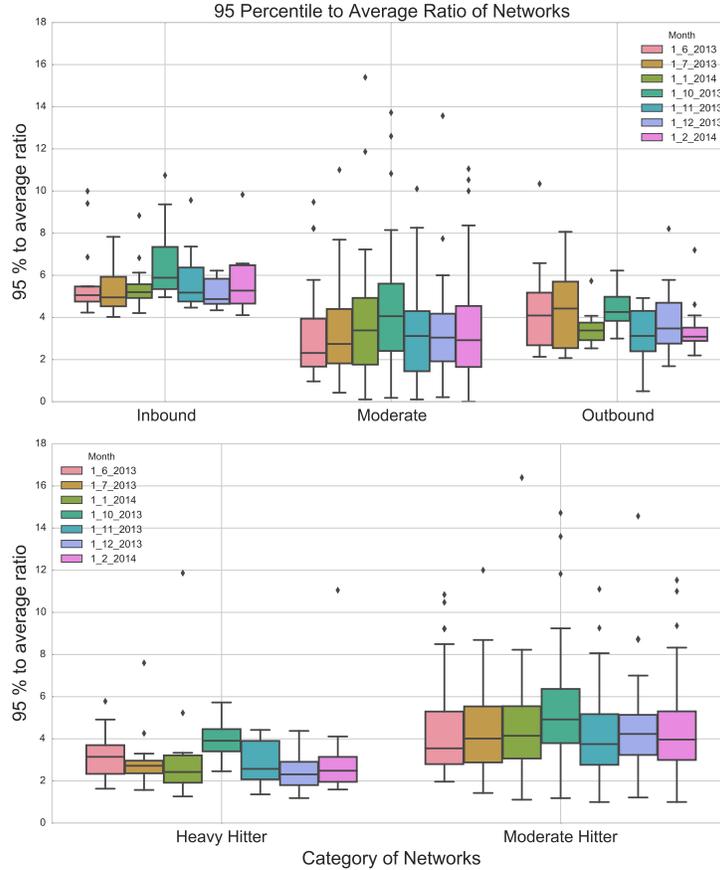


Figure 2.2: Mean 95th Percentile to average ratio for IXPs, using different network classifications.

95th-to-average ratio as compared to other networks, meaning that their traffic profile is likely to be burstier. Figure 2.3 illustrates the difference by plotting the mean skewness of the traffic distribution for each network type.

Skewness reveals how much the traffic distribution leans to one side of the mean; for a random variable X : $\text{Skewness} = E[(X - \mu)^3] / (E[(X - \mu)^2])^{3/2}$, where μ is the mean. If a probability distribution function is unimodal, then higher positive skew implies few values higher than the mean, i.e., the 95th Percentile value would be closer to the average. The empirical probability mass function for the traffic of each network is unimodal for our data sets. Heavy-outbound networks have high positive skew (the mean is between 5 and 25),

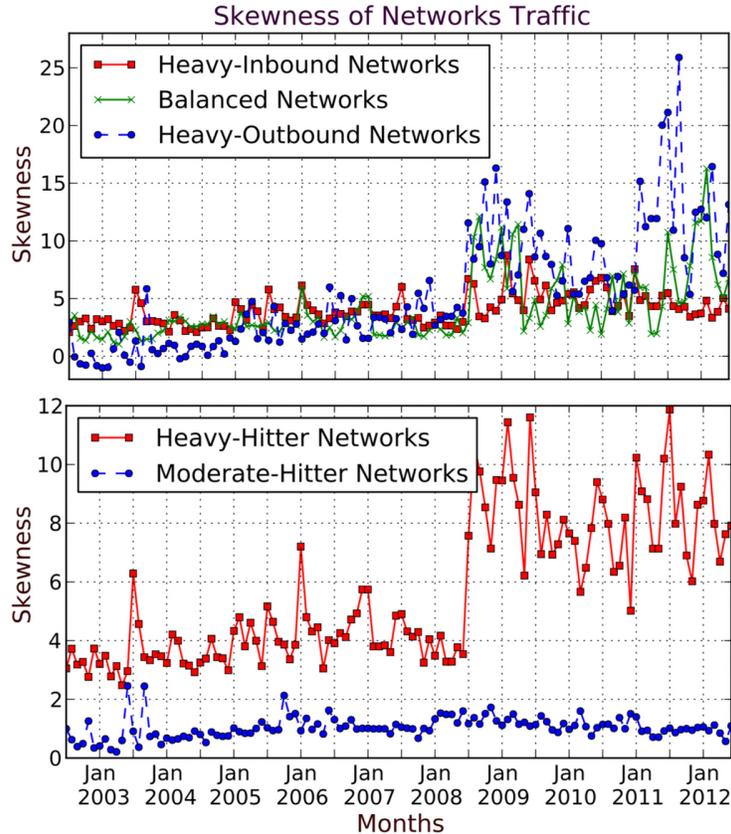


Figure 2.3: Mean skewness for different network types in the SWITCH dataset. Heavy-outbound networks have a higher skewness, especially in the last 4 years.

especially in the last 4 years², compared to heavy-inbound networks or balanced networks, whose mean skewness is between 0 and 12 and 5 and 15, respectively. Similarly, heavy-hitter networks have higher positive skew than moderate-hitter networks. Figure 2.4 shows the mean skew of traffic for networks at each IXP, classified according to dominant traffic direction and traffic volume. As in the SWITCH dataset, heavy-outbound and heavy-hitter networks generally have a larger skewness than heavy-inbound and moderate-hitter networks.

In summary, the 95th-to-average ratio has been stable for various classes of networks in our dataset over the last decade, indicating that a high-percentile billing scheme is still useful. Certain networks (particularly heavy-outbound and heavy-hitter networks) are able to achieve a lower 95th Percentile to average ratio (perhaps using intelligent means of traffic

²The level shifts around 2009 coincide with SWITCH connecting to AMS-IX, acquiring hundreds of new peers, though the set of customers over which we compute statistics is unchanged.

shaping), and hence a lower billing volume for the same total amount of transit traffic. Traffic smoothing may allow networks to achieve a lower transit bill, but this says little about the contribution of those networks to the provider's peak traffic. The 95th Percentile of a network does not account for *when* the peaks occur, and so it is unclear whether it is fair to charge each customer using the same percentile.

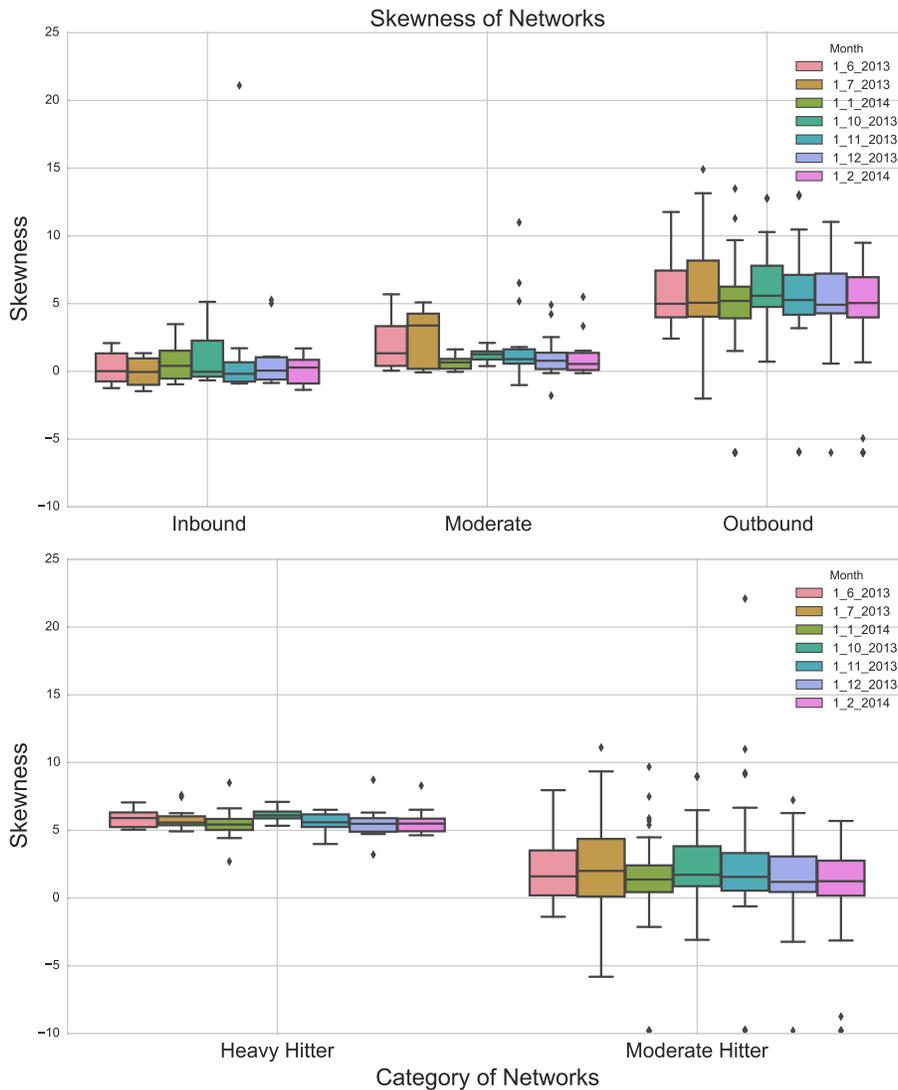


Figure 2.4: Mean skewness for networks in the IXP dataset.

2.4 Fairness of 95th Percentile Billing

Motivated by the preceding discussion, we now focus on the fairness of the 95th Percentile billing mechanism. We consider a billing mechanism fair if the amount of resources used by a network is reflected in the amount it is charged. An appealing idea in this context is the Shapley value, which assigns costs to the members in a cooperative game [49]. It possesses many attractive properties – it is *efficient*, i.e., the sum of costs assigned to each member is the total cost to the system, and it is *symmetric*, i.e., two members that have the same contribution will be assigned the same cost.

2.4.1 Shapley Value Percentile Billing

Stanojevic et al. [52] presented a model of the ISP cost allocation problem as a cooperative game. The cost function of a group is the 95th Percentile of the total traffic obtained by adding the traffic of all members in that group. This cost estimate is consistent with the idea that the transit provider must provision for peak traffic, and is itself billed by its provider based on this value. The Shapley value (ϕ_i) of network i is then uniquely defined by $\phi_i = \frac{1}{\mathcal{N}!} \sum_{\pi \in \Pi} (\mathcal{V}(S(\pi, i)) - \mathcal{V}(S(\pi, i) \setminus i))$ where \mathcal{V} is the cost function, Π is the set of all possible permutations of players \mathcal{N} and $S(\pi, i)$ is the set of all players in ordering π before i and including i .

Once we determine the Shapley value of each network, we need to map it to a billing percentile. Let the volume corresponding to the 95th Percentile value of the total traffic be \mathcal{V} . Then (by efficiency) the Shapley values of the customer networks will satisfy $\mathcal{V} = \sum_i \phi_i$. Let the volume corresponding to the 95th Percentile of network i be x_i . Then the total volume billed by the transit provider under the 95th Percentile billing scheme is $\sum_i x_i$, which we define as \mathcal{X} . Trivially, $\mathcal{X} \geq \mathcal{V}$. For an apples-to-apples comparison between the two billing schemes, we define the normalized Shapley value of network i as $s_i = \phi_i \mathcal{X} / \mathcal{V}$, so that the total billing volume in both cases is \mathcal{X} . Then each network can be charged based on a percentile that yields the traffic volume closest its normalized Shapley value, which is the

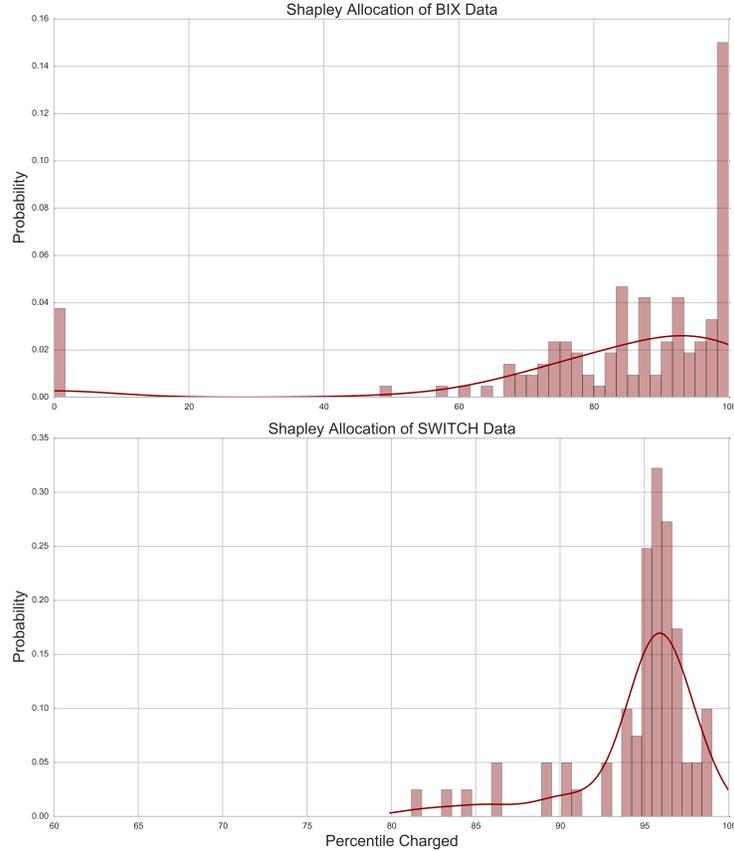


Figure 2.5: Shapley value percentiles: SWITCH dataset (Mar 2012) and IXP dataset (SIX, Aug 2013).

“Shapley value percentile” of that network.

Computation of the Shapley value is quite complex—with N users, it has complexity order of $\mathcal{O}(N!)$. Even for a moderate size ISP, which has around 50 users, the complexity is of the order of 10^{64} . Stanojevic et al. [52] used a Monte Carlo approximation, which achieves a good trade-off between accuracy and complexity. We used this approximation to find the Shapley value percentile for the SWITCH dataset (month of March 2012) and the SIX IXP (August 2013). The results are shown in Figure 2.5. Clearly, the Shapley value percentiles are widely different from the 95th Percentile .

In addition to computational complexity, the Shapley value percentile can be anywhere between 0 and 100. This approach lacks the ability of restricting the charging percentiles

to a fixed range. The handicaps of directly using the Shapley value motivate a need for a simple proxy that captures its essence. A key observation is that a traffic profile has greater Shapley value when it is concentrated during the peak periods when demand is highest. Thus, Shapley value percentile billing would charge users with high peak traffic higher than users with off peak traffic.

2.4.2 Overlap rank

Building on the intuition developed in the last section that it is fair to charge more to networks with traffic during peak periods than off-peak periods, we will show how the current 95th Percentile billing mechanism can lead to unfairness as it does not consider peak and off-peak periods. We define the peak periods of a transit provider as those in which the total traffic carried by the transit provider exceeds the 95th Percentile of the provider’s total traffic. We similarly define the peak slots for customer networks. Based on the number of peak slots of networks that overlap with peak slots of the total traffic, we rank the networks from highest to lowest and call it the *overlap rank*. Thus, a network with rank 0 has the maximum number of peak slots that occur during the same time intervals as the peak slots of the transit provider. We also rank networks based on their 95th Percentile and call it the 95th Percentile rank.

Figure 2.6 plots overlap rank vs. percentile rank (normalized to 100) for the IXP dataset (first 3 plots) and one month (January 2012) from the SWITCH dataset (far right). If networks with high 95th Percentile rank also had high overlap rank, most points would appear on the diagonal, and imply that 95th Percentile billing is charging the contributors who necessitate the provisioning of large transit links. Figure 2.6 tells a different story. The points below the diagonal, especially those in the red shaded area (16% of networks for SWITCH) have a high 95th Percentile rank but a low overlap rank, which means that their peaks are mostly in the peak period, but their billing volume is relatively lower. Analogously, the points above the diagonal line, especially in the gray region (15% of networks for SWITCH) correspond to low 95th Percentile rank and high overlap rank. Their contribution to the peak

period is low but they have a relatively high billing volume. Similar observations can also be made from the IXP graphs in Figure 2.6.

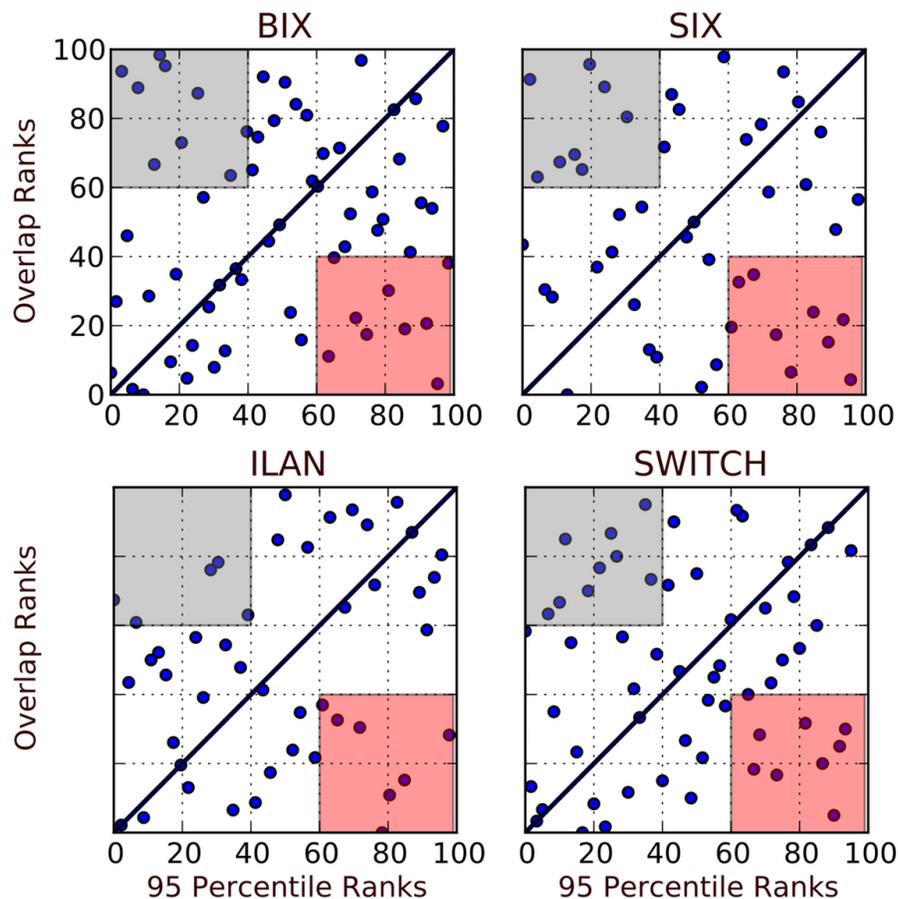


Figure 2.6: Overlap rank vs 95th Percentile rank for IXP dataset (Aug 2013) and one month of SWITCH dataset (Jan 2012). A large fraction of networks lie far from the diagonal, meaning they have a large billing volume but little overlap with the provider’s peaks, or vice versa.

2.4.3 Provision ratio

The overlap rank considers only the cardinality of overlap slots, without accounting for diverse traffic volumes. A good proxy for the Shapley value should capture the volume during peak slots, appropriately normalized with the amount of traffic generated by the network. We define the **provision ratio** (PR) of a network as the ratio of the average traffic during

the peak slots of total traffic to the 95th Percentile of that network’s traffic.

$$PR_i = \frac{\text{Total traffic of } i \text{ during peak slots} / \# \text{ of peak slots}}{95^{\text{th}} \text{ Percentile of } i\text{'s traffic}}$$

The PR is essentially the ratio of traffic contributed by the network during the peak time slots (or average capacity provided to that network during these peaks) to the peak traffic of that network (excluding the top 5% of bursts); It can be viewed as the fraction of a network’s peak traffic that occurs during the provider’s peak periods. We propose that the PR can be an important component of a billing mechanism, because it captures the contribution of a network’s traffic to the provider’s peak. The PR is also robust to the exact thresholds used to compute it – we found that in our datasets, the provision ratio is robust to the exact threshold for defining a peak slot, e.g., if we change the 95th Percentile to 85th percentile, the provision ratio does not change significantly.

The provision ratio is not equal to the Shapley value percentile in an absolute sense, but in a relative sense it appears to have the right characteristics. To quantify the similarity between the two, we find the percentage of orders preserved between all possible pairs of networks in both datasets. A transit provider with N customers will have ${}^N C_2$ customer pairs. For each pair, order is preserved if the network that is charged a higher Shapley percentile also has a higher provision ratio. We find that for the SWITCH dataset, the provision ratio preserves between 76% and 82% of orders in the SWITCH dataset (each month of 2012) and 89%, 75%, and 82% for the SIX, BIX, and ILAN IXPs, respectively (August 2013). The strong similarity of orders indicates that provision ratio is indeed order preserving.

2.4.4 Towards a new billing mechanism

One could argue that the 95th Percentile billing scheme is an approximation, aiming for simplicity and predictability over fairness. At the other extreme is Shapley value pricing, which charges each user differently based on their actual contribution to the provider’s

costs. An open challenge is how to achieve both objectives – fairness and low computational complexity. We are currently exploring the use of the provision ratio in a scheme that determines the optimal percentile to charge a given customer. The objective of this scheme would be to vary the billing percentile per customer, and to use the provision ratio as a measure of the contribution of a customer to the provider’s peak traffic. This pricing scheme would automatically assign lower billing percentiles (i.e., give discounts) to customers whose peak traffic does not contribute significantly to the provider’s peak, and higher percentiles to customers that contribute most to the provider’s peak. An important criterion for such a scheme is that the provider should be able to communicate information about its peak and off-peak periods to customers, without having to make its traffic profile available publicly. For this purpose, the provider could design a tool that accepts a customer’s traffic profile and analyzes it in relation to its own traffic to determine the percentile at which it would charge the customer. Such a scheme would retain the attractive properties of burstable billing (because it is still based on a billing percentile), while better accounting for a network’s contribution to total provider costs. Our initial investigation indicates that this problem can be formulated as a convex optimization, and hence solved efficiently.

2.5 Measuring Billing Volumes

We present a framework for percentile-based measurement of billing volumes. Consider a transit provider with N customers indexed by i , $i \in \{1, 2, \dots, N\}$. Each month, the transit provider must determine the billing volumes of each customer.

The relationship between billing volume and billing percentile can be expressed using the cumulative distribution function (CDF) of the customer network’s traffic. First, both the inbound and outbound traffic volumes are measured in 5-minute intervals, and are used to calculate the average transmission rates during each interval. Denote the empirical CDFs of customer network i ’s transmission rates by $\mathcal{F}_{i(in)}(\cdot)$ and $\mathcal{F}_{i(out)}(\cdot)$, for inbound and outbound directions, respectively. Also, denote the inverse cumulative distribution functions by $\mathcal{F}_{i(in)}^{-1}(\cdot)$ and $\mathcal{F}_{i(out)}^{-1}(\cdot)$. If the CDF function is not one-to-one, the inverse will be an interval (due to

monotonicity of the CDF function). If this is the case, we take the supremum of the interval to be the value of the inverse, *i.e.*,

$$\mathcal{F}_{i(in)}^{-1}(y) = \sup \{x | \mathcal{F}_{i(in)}(x) = y\}, \quad (2.1)$$

and similarly for outbound traffic. We then decide on whether the traffic is inbound or outbound dominated by comparing the 95th Percentile volumes of the two, *i.e.*, we compare $V_{i(in)}(0.95) = \mathcal{F}_{i(in)}^{-1}(0.95)$ with $V_{i(out)}(0.95) = \mathcal{F}_{i(out)}^{-1}(0.95)$. We choose the overall CDF of the customer i 's traffic to be the one with the larger 95th Percentile. Thus, if $V_{i(in)}(0.95) > V_{i(out)}(0.95)$, then we set $\mathcal{F}_i(x) = \mathcal{F}_{i(in)}(x)$. Correspondingly, the volume billed by the 95th Percentile scheme is $V_i(0.95) = \mathcal{F}_i^{-1}(0.95)$, and the sum total volume of billed traffic is $V_{95} = \sum_{i=1}^N V_i(0.95)$. As described in Section 4.1, the 95th Percentile method is unfair because it does not account for the fact that the temporal traffic profile of customers might impose very different loads on the transit provider. For instance, a customer whose traffic is concentrated in the peak periods of overall traffic would require the transit provider to provision more capacity than one whose traffic is in the off-peak periods. A fair scheme should ensure that the amount of resources used by a customer should be reflected in its corresponding billing volume.

While both the Shapley value and Provision Ratio can be translated into percentile-based volume measurement methods, they do not directly allow us to restrict the range of acceptable billing percentiles. As our objective is to incentivize customers to occupy off-peak periods, while not excessively dis-incentivizing those who do not, we desire a framework that incorporates both fairness as well as flexibility in choosing billing percentiles.

2.6 Optimization Framework

We seek a scheme whereby customers occupying off-peak periods are given rebates, while those that do not are charged extra. However, we also wish to ensure that the billing percentiles are not overly large or small. Finally, this must be done at no loss of net revenue

to the transit provider. How can we achieve these goals?

Suppose that the transit provider uses a price function $\mathcal{B}(\cdot)$ to translate traffic volumes into dollar charges. Often, this function is (approximately) concave and increasing [42] to ensure discounts for large volume customers. We do not propose to alter the billing function, but instead use $\mathcal{B}(\cdot)$ as is. Let the revenue obtained through 95th Percentile based volume measurement be M_{95} . Then the solution to the following optimization problem attains our goals:

$$\max_{\{p_i\}} \sum_{i=1}^N (0.95 - p_i) \omega_i - \gamma \left(\sum_{i=1}^N (0.95 - p_i) \right)^2 \quad (2.2)$$

$$s.t. \quad L \leq p_i \leq H, \quad \forall i \in \{1, 2, \dots, N\}, \quad (2.3)$$

$$\sum_{i=1}^N \mathcal{B} \left(\tilde{\mathcal{F}}_i^{-1}(p_i) \right) \geq M_{95}. \quad (2.4)$$

Here, the objective (2.2) is to ensure that the billing volume percentile is reduced below 0.95 as much as possible, i.e. provide the maximum possible incentives to customers. To provide incentives for off-peak customers, we set the weight $w_i = (1/\rho_i)^\alpha$, where $\alpha \geq 1$. Since the weight varies inversely with the normalized Provision Ratio, maximizing the objective would assign larger p_i values to customers with smaller weights i.e., high occupancy during peak times. The second term in the objective is to smooth it, as otherwise the solution would be to set p_i to extreme high or low values. Parameter γ is used to decide the desired smoothing.

We next have a (convex) constraint (2.3) that ensures that the percentiles output by the optimization lie in an acceptable interval between $[L, F]$. Constraint (2.4) ensures that the transit provider does not suffer any loss of revenue (as compared to 95th Percentile based volume measurements). As defined above, $\mathcal{B}(\cdot)$ is a concave billing function. Now, since the inverse CDF of traffic, \mathcal{F}_i^{-1} , is empirical, it might not have any particular form. Hence, we approximate it using a concave function $\tilde{\mathcal{F}}_i^{-1}$ in the range $[L, F]$. In practice, we employed

an approximation of the form $\tilde{\mathcal{F}}^{-1}(x) = a + bx + c\sqrt{x}$ with $c \geq 0$. Notice that the concave approximation immediately implies that the constraint becomes convex.

Our problem formulation is in the form of convex optimization, and hence the solution can be easily computed using convex solvers. If we ensure that $0.95 \in [L, H]$, then 0.95 satisfies the constraints. Then setting $p_i = 0.95$ for all i would result in an objective value of zero. Maximization of the objective can only increase the value, which means that the optimal value should be non-negative. We denote the set of percentile values that solve the optimization problem (2.2)–(2.4) by $\{\hat{p}_i\}$, and refer to them as the *optimal weighted percentiles* (OWPs). In the next section we calculate the OWPs for customer networks using multiple data traces, and compare the values with the equivalent Shapley value percentiles (SVPs), in order to gauge the fairness achieved by this method. In this section first we compare the fairness achieved by SVP versus OWP, using data sets of traffic seen by real transit providers and then study the advantages of new billing mechanism if the traffic is elastic. Our first data set (the “SWITCH” data set) is from SWITCH, a European transit provider that serves educational institutions and some commercial organizations. The second data set (the “IXP” data set) is parsed from MRTG graphs published by three European Internet exchange points (IXPs): SIX, BIX and ILAN. From both data sets we extract traffic rates of each customer network at 5-minute intervals.

Our comparison of SVP and OWP proceeds as follows. For each customer i , we first calculate the 95th Percentile billing volume $V_i(0.95)$, and use a billing function $\mathcal{B}(x) = 50x^{0.7}$ to translate these volumes into dollar charges. This form of the billing function is based on real-world transit prices [42] and has been used in prior work [11, 14]. We refer to the sum total revenue obtained over all customers as R_{95} , and use it as the minimum target revenue that both the SVP and OWP schemes should assure to the transit provider. We then compare which customers are targeted for higher/lower percentile billing in each method to check if both methods are aligned in their conception of fairness.

To find the Shapley value percentiles (SVPs) corresponding to the above revenue tar-

get, we use the same formulation as Section 2.4.1. Since calculating the Shapley value is computationally intensive, we used a Monte Carlo approximation [52] with 10000 iterations. Here, the idea is to pick random subsets of customers in Shapley value evaluation equation, and average the value over such subsets. Then, for each customer i , we set $S_i = \sigma_i R_{95}$, and determine the set of SVPs $\{p_i^S\}$ using $p_i^S = \mathcal{F}_i(S_i)$. Note, that for accurate results even this process is computationally expensive.

To find the optimal weighted percentiles (OWPs), we limit the allowable percentiles to 3 units above and below 95%, that is $L = 92\%$ and $H = 98\%$ in (2.3). We set $\alpha = 5$ when selecting the weights, and a smoothing parameter $\gamma = 200$. We used the Levenberg Marquardt algorithm [30] for approximating the inverse CDF function with a concave function, and found that the normalized least squares errors are less than 10^{-2} . The result of our optimization is a set of percentiles $\{\hat{p}_i\}$. Note that the complexity of these calculations is small as compared to determining the SVP.

We computed the SVPs and OWPs for four years of SWITCH data and 3 months of IXP data. When we plotted the distribution of these percentiles, the support of SVPs varied widely. For example the support of SVPs for February 2012 SWITCH data is $[0.83, 0.99]$. However, by design, the support of all OWP distributions is $[0.92, 0.98]$. Also, as desired, the OWP scheme reduced the billing percentiles of many customers, while increasing that of only a few. Since our conception of fairness is that of the Shapley value, we consider the OWP method fair if the *same* customers are targeted for high/low percentile billing as in the SVP method. We now show this kind of order preservation is largely maintained between SVP and OWP. We first visualize percentile information for a month in the SWITCH data set in Figure 2.7. We place the individual customers in increasing order of the reciprocal of their Provision Ratios on the x-axis, and their average traffic on the y-axis. For example, a customer with a large reciprocal of Provision Ratio (*i.e.*, it occupies off-peak periods) and small average traffic would appear in the bottom right of the plot. Each circle or square represents a customer network, while the size and intensity of fill color is proportional to the

relative percentile used to bill them.

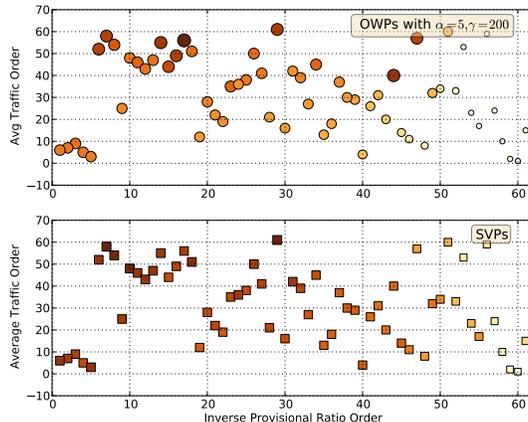


Figure 2.7: Inverse Provision Ratio order vs Average traffic order of SVPs and OWPs for February 2012 SWITCH data.

We observe that in the SVP scheme, there is a gradual increase in billing percentiles from the bottom right to the top left (with a few exceptions). The same trend is observed in OWP. Although the actual percentiles used to bill are different, we see that by-and-large the same customers are targeted in both schemes.

While the visualization indicates the validity of the OWP scheme in preserving fairness, we would prefer to use numerical metrics. We define two such metrics, and show that SVP and OWP are well aligned on both metrics. Our first metric is that of order preservation. We say that order is preserved between two customers i and j if $p_i^s > p_j^s$ implies that $\hat{p}_i > \hat{p}_j$. We compute the percentage of orders preserved in each month of our data sets. For the SWITCH data set, this gives 48 samples over four years from 2009 to 2012. We plot the distribution of percentage of orders preserved in Figure 2.8. Here, the x-axis is the percentage of orders preserved in that sample, while the y-axis is the number of samples that had that value. We see that all values are above 70% and many values are around 80%, indicating strong order preservation between SVP and OWP. We observed similar results for IXP data sets:

the percentage of orders preserved is above 78%. The second metric that we consider is the

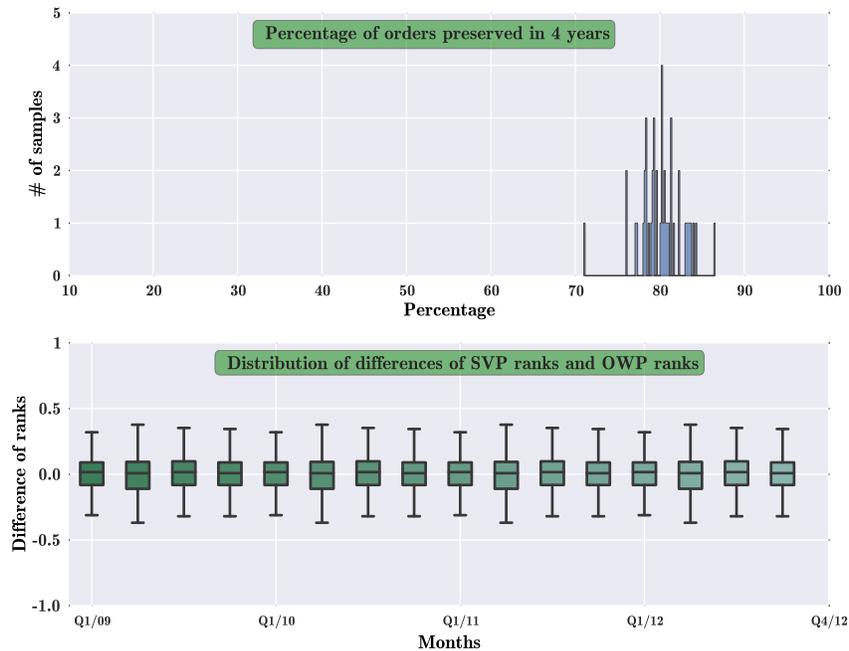


Figure 2.8: Distribution of percentage of orders preserved and Box and Whisker plot of difference in ranks for four years.

difference in ranks of billing percentiles. Consider the two sets of billing percentiles $\{p_i^s\}$ and $\{\hat{p}_i\}$, corresponding to SVP and OWP methods, respectively. We can arrange the percentile values in ascending order in each set. Let r_i^s and \hat{r}_i refer to the order in which $\{p_i^s\}$ and $\{\hat{p}_i\}$, respectively appear in the ordered sets. We call r_i^s and \hat{r}_i as the *ranks* of customer i according to the two schemes, and consider the normalized rank difference $(r_i^s - \hat{r}_i)/N$. The difference must lie in $[-1, 1]$, and a large difference would mean that the ranks are very different, while a small one indicates that they are close to each other. We group the data into three-month intervals (quarters) and present a box-whisker plot of the distributions of the normalized absolute differences over each quarter. Here, the bottom and top of each “box” represents 1st and 3rd quartiles of the distribution for that quarter (*i.e.*, 50% of the samples are contained in both boxes together), while the bottom and top “whiskers” are equal to 1.5

times the 1st and 3rd quartiles of the distribution. We observe that the distributions tightly concentrate around 0, indicating strong preservation of ranks between the SVP and OWP schemes.

Now we will see how this new billing mechanism could benefit the transit provider when the traffic is elastic. Let us say that every customer tries to minimize the cost incurred by moving the traffic from peak slots to non peak slots of previous month. But, moving all the traffic is not feasible so we assume a bound on elasticity \mathcal{E} that is the maximum amount of traffic that can be moved. Let us assume that they employ min max strategy for removing traffic from peak slots and max min strategy for filling the non-peak slots. The percentage difference in the traffic reduction for different elasticities is shown in the figures below 2.9 and 2.10. The x-axis denote the index of users or customers and y-axis denote the percentage reduction in billing volume achieved by moving traffic.

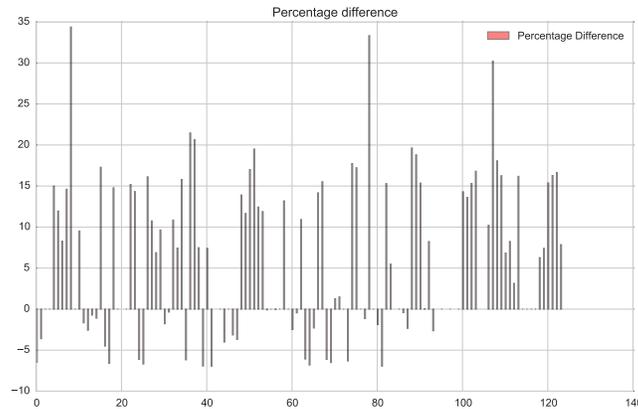


Figure 2.9: Percentage of amount of billing volume reduction with $\mathcal{E} = 5\%$ elasticity. The x-axis denote the index of the customers and y-axis denote the percentage reduction achieved by moving 5% of traffic to off-peak periods.

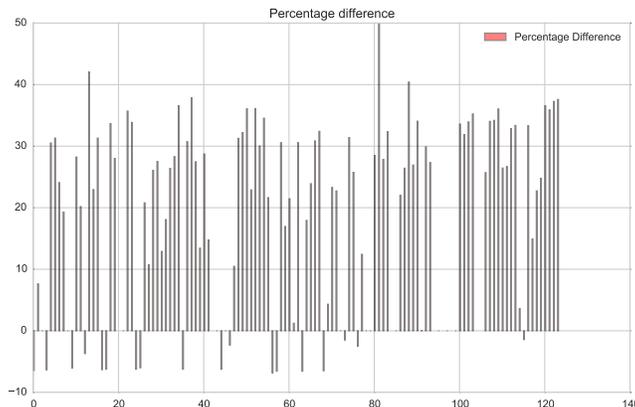


Figure 2.10: Percentage of amount of billing volume reduction with $\mathcal{E} = 15\%$ elasticity. The x-axis denote the index of the customers and y-axis denote the percentage reduction achieved by moving 5% of traffic to off-peak periods.

2.7 Conclusions

In this chapter, our goal was to empirically examine the effectiveness and fairness of the 95th percentile pricing scheme, using a decade of historical traffic data and provide an alternate schemes. Our analysis shows that over the years, certain networks have lower 95th-to-average ratio than others – for the datasets we studied, networks with predominantly inbound traffic have higher 95th-to-average ratios, and would incur a higher billing volume than those with predominantly outbound traffic (for the same amount of total traffic), and similarly for moderate hitters vs. heavy hitters. Furthermore, we find that the 95th percentile pricing scheme can be unfair, as the 95th Percentile traffic of a network is often unrelated to the amount of time that network’s peak traffic overlaps that of its provider, nor does it accurately represent the contribution of that network to the provider’s peak traffic.

We defined a new metric, the Provision Ratio (PR) for a network, which is easy to compute and is able to capture the contribution of a customer traffic to the provider’s peak. By using the Provision Ratio as a weight factor in the optimization scheme, the ISP is able to assign lower percentiles to users that have a low contribution to the provider’s peak periods. This scheme achieves a notion of fairness similar to the Shapley value, is efficient,

and can provide rebates to customers that contribute less to the ISP's peak traffic. We also investigated how much the customers can save by moving their traffic into off peak periods.

The setup that we considered has one seller (the transit provider) and a finite number of customers. In large Internet marketplaces such as Amazon, eBay etc., there are a large number of sellers and customers. The techniques we used to analyze finite user markets are intractable in this setting. In the next two chapters, we will study two such markets, each with a large number of agents. The first market that we consider is resource allocation in cellular network, which we discuss in the next chapter.

3. A MEAN FIELD GAME APPROACH TO QOE-AWARE CELLULAR SCHEDULING

3.1 Introduction

There has recently been a rapid increase in the use of smart hand-held devices for Internet access. These devices are supported by cellular data networks, which carry the packets generated by apps running on these smart devices. These apps can be modeled as queues that arrive when the user starts the app, and depart when the user terminates that app. Packets generated by an app are buffered in a queue corresponding to that app. Queueing delays impact the quality of user experience (QoE) based on the app being used. For example, the QoE of a streaming application might be more sensitive to queueing delays than that of a file download. Users move around cells that each has a base station, and scheduling a particular user provides service to the queue that represents his/her currently running app. User interest could shift from app to app, regardless of whether or not there are buffered packets. Hence, a queue might terminate and be replaced by a new one even if there are jobs waiting for service.

An important problem in cellular data networks is that for scheduling, *i.e.*, determining which queues receive service at each time instant. Most work on scheduling has focused on the case of a finite number of infinitely long lived flows, with the objective being to maximize the total throughput. A seminal piece of work under this paradigm introduced the so-called *max-weight algorithm* [54]. Here, the drift of a quadratic Lyapunov function is minimized by maximizing the queue-length weighted sum of acceptable schedules. Later work (*e.g.*, [33, 16, 40]) has used a similar approach in a variety of network scenarios. If queues arrive and depart, then a natural scheduling policy in the single server case is a *Longest-Queue-First (LQF)* scheme, in which each server serves the longest of the queues requesting service from it. LQF has many attractive properties, such as minimizing the expected value of the

longest queue in the system.

The above approach assumes that the queue length values are given to the scheduler, and if we have heterogeneous apps, the cost on QoE of queueing for each can be determined from the queue lengths. However, while the downlink queue lengths would naturally be available at a cellular base station, the only way to obtain uplink queue information is to ask the users themselves. However, a larger value of queue length results in a higher probability of being scheduled under all the above policies, implying an incentive to lie. Also, such a framework does not allow for the human end user to decide his or her own priorities for the QoEs for different apps. For instance, one user might place a high importance on the QoE of Skype and Facebook, whereas another might have high importance for YouTube and Pandora. Thus, traditional approaches do not account for the human in the loop that is the ultimate consumer of the service.

An appealing idea is to use a pricing scheme to inform scheduling decisions for cellular data access. These prices could be in the form of tokens issued by the cellular service provider that are used as currency in the service market. An example of a pricing approach is presented in [22], which describes an experimental trial of a system in which day-ahead prices are announced to users, who then decide on whether or not to use their 3G service based on the price at that time. However, these prices have to be determined through trial and error. Can we determine prices by using an auction?

Our key objective is to design an incentive compatible scheduling scheme that behaves in a (weighted) LQF-like fashion. We consider a system in which each app bids for service from the base station that it is currently connected to. The auction is conducted in a second-price fashion, with the winner being the app that bids highest, and the charge being the value of the second highest bid. It is well known that such an auction promotes truth-telling [28]. Would the scheduling decisions resulting from such auctions resemble that of LQF? Would conducting such an auction repeatedly over time with queues arriving and departing result in some form of equilibrium?

3.1.1 Mean Field Games

We investigate the existence of an equilibrium using the theory of Mean Field Games (MFG) [29, 55, 3, 9, 27, 56]. In MFG, the players assume that each opponent would play an action drawn *independently* from a fixed distribution over its action space. The player then chooses an action that is the best response against actions drawn in this fashion. We say that the system is at Mean Field Equilibrium (MFE) if this best response action is itself a sample drawn from the assumed distribution.

The MFG framework offers a structure to approximate so-called Perfect Bayesian Equilibrium (PBE) in dynamic games. PBE requires each player to keep track of their beliefs on the future plays of every other opponent in the system, and play the best response to that belief. This makes the computation of PBE intractable when the number of players is large. The MFG framework simplifies computation of the best response, and often turns out to be asymptotically optimal.

Work on MFGs has mostly focused on showing the existence, accuracy and stability of MFE [29, 55, 3, 9]. In the space of queueing systems, some recent work considers the game of sampling a number of queues and joining one [56]. However, ours is a scheduling problem in queueing system interacting with an auction system, which we believe is unique.

In the space of applications, Iyer *et al.* [26] study advertisers competing via a second price auction for spots on a webpage. The bid must lie in a finite real interval, and the winner can place an ad on the webpage. With time, the advertisers learn about the value of winning (probability of making a sale). Li *et al.* [31] consider the problem of mechanism design for truthful state revelation (number of packets at each station) in a wireless D2D streaming system. Their main result is to generalize the Groves mechanism to the mean field regime. Both use some version of fixed point theorem to show existence of the MFE.

Neither of the above consider the use of auctions for service scheduling in queueing systems, which is the basis of our problem. In our setup, the state is the queue with arrivals and departures, and we allow bids to lie in the full positive real line. These considerations

result in significant technical work to show existence and characterize the MFE.

3.1.2 Overview of Chapter

We introduce the Mean Field Game (MFG) in Section 3.2. Here, a selected agent (app) has a belief ρ about the bid distribution of the other agents, and assumes that their bids will be drawn independently from this distribution. The state of the agent is its current queue length, and it faces a per-time step cost that is a function of its queue length, which models the impact on QoE of queueing delay. The agent must place a bid based on the belief and its current state and belief about other agents.

We consider the problem of determining the cost minimizing bid function and the corresponding value function as a Markov Decision Process in Section 3.3. We show that the Bellman operator corresponding to the MDP is a contraction mapping with a unique minimum, implying that value iteration would converge to the best response bid. Further, we show that the best response bid is monotone increasing in queue length. We call the bid distribution across agents that results from playing the best responses as γ .

We next prove the existence of the MFE in Sections 3.4–3.5 by verifying the conditions of the Schauder Fixed Point Theorem. We need to show that the mapping between the assumed bid distribution ρ to the resultant bid distribution γ is continuous, and that the space in which γ lies is contained in a compact subset of the space from which ρ is drawn. In order to do this, we first show that the mapping between ρ and best response bid function is continuous, and then show that the map between ρ and the state (queue length) distribution is continuous. Putting these together yields the required continuity conditions. We then verify the conditions of the Arzelà-Ascoli Theorem for showing compactness.

We show in Section 3.6 that the MFE in our case is an asymptotically accurate approximation of a PBE. The result follows from the fact that any finite subset of agents is unlikely to have interacted with any of the others as the number of agents becomes large. Finally, we present simulation results in Section 3.7, showing that MFE computation is straightforward.

Discussion: In the case of a single cost function (homogeneous QoE for all apps) the best

response bid function is monotone increasing in the queue length regardless of ρ . This implies that the service regime corresponding to MFE is identical to LQF (or a weighted version if we have different QoE classes for apps). Further, our simulations suggest that if the base stations were to compute the empirical bid distribution and return it to the users, the eventual bid distribution would be the MFE. Thus, the desirable properties of LQF are a natural result of auction-based scheduling, while the queue length distribution would be that generated by LQF.

3.2 Mean Field Game Model

We consider a system consisting of N cells and NM agents (apps), which are randomly permuted in these cells at each time instant, with each cell having exactly M agents¹. The model is consistent on the likely evolution of the 5G cellular system, wherein we expect a large number of small, dense cells and much user mobility across different cells. The mobility model is identical to the basic framework used in work on mobile wireless networks [21]. Each cell contains a base station, which conducts a second price auction to choose which agent to serve. Each agent must choose its bid in response to its state and its belief over the bids of its competitors.

Figure 3.1 illustrates the MFG approximation, which is accurate in the limit as N becomes large. An MFG is described from the perspective of a single candidate agent i , which assumes that the actions of all its competitors are drawn independently from some distribution. The asymptotic accuracy of the independence assumption follows from a standard argument on the *propagation of chaos* whose details are provided in Section 3.6. In Figure 3.1, the auction (shown as blue/dark tiles) and the queue dynamics (shown as beige/light tiles). Since we focus on a single agent, we do not need to specify its identity explicitly, unless we wish to compare its actions with those of other agents. Hence, we will drop the index i where possible for ease of notation.

¹Note that our results are essentially unchanged if M is a random variable with finite support.

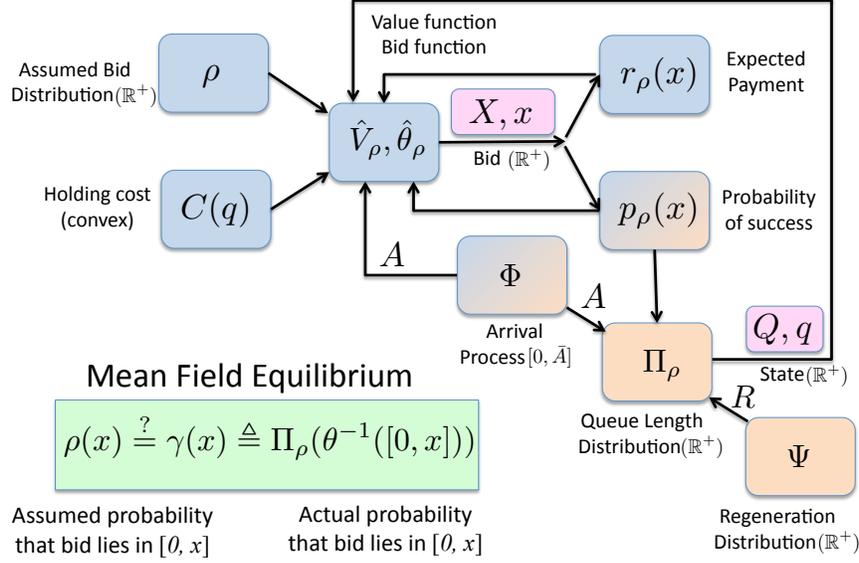


Figure 3.1: The game consists of an auction part (blue/dark tiles) and a queue dynamics part (beige/light tiles). The system is at MFE if the distribution of the bid X is equal to the assumed bid distribution ρ .

3.2.1 Auction System

At each time step k , the agent of interest competes in a second price auction against $M - 1$ other agents, whose bids are assumed to be independently drawn from a continuous, finite mean (cumulative) bid distribution ρ , with support \mathbb{R}^+ . The state of the agent is its current queue length q (the random variable is represented by Q). The queue length induces a holding cost $C(q)$, where $C(\cdot)$ is a strictly convex and increasing function. The cost function could be one of a finite set of cost functions, modeling the impact on QoE of queuing for the currently running app for that agent. However, since the analysis is identical for one or a finite set of cost functions, we focus on the single cost function in the analysis below.

Suppose that the agent bids an amount $x \in \mathbb{R}^+$. The outcomes of the second price auction are that the agent would obtain one unit of service with probability $p_\rho(x)$ and would have to pay an expected amount of $r_\rho(x)$ when all the other bids are drawn independently from ρ . Further, the queue has future job arrivals according to distribution Φ , with the random job

size being denoted by A . Finally, the app can terminate at any time instant with probability $1 - \beta$. Based on these inputs, the agent needs to determine the value of its current state $\hat{V}_\rho(q)$, and the best response bid to make $x = \hat{\theta}_\rho(q)$. The assumption that only a single unit of service is provided at each base station is for simplicity of notation, and our results are unchanged if we are allowed to choose some $\tilde{M} < M$ agents as winners in each auction. The mechanism followed would be a $\tilde{M} + 1^{\text{th}}$ price auction in that case.

3.2.2 Queueing System

The queueing dynamics are driven by the arrival process Φ and the probability of obtaining a unit of service being $p_\rho(x)$ as described above. When the user terminates an app, he/she immediately starts a fresh app, *i.e.*, a new queue takes the place of a departing queue. The initial condition of this new queue R is drawn from a regeneration distribution Ψ , whose support is \mathbb{R}^+ . The invariant distribution associated with this queueing system (if it exists) is denoted Π_ρ .

We make the following assumptions.

Assumption 1. *At each time k , the arrivals $\{A_k\}$ are i.i.d random variables distributed according to Φ . We assume that $A_k \in [0, \bar{A}]$, where \bar{A} is finite. Also, these random variables have a bounded density function, ϕ (*i.e.*, $\|\phi\| < c_\phi$, where $\|\cdot\|$ is the sup norm).*

Assumption 2. *The regeneration values $\{R_k\}$ are i.i.d random variables distributed according to Ψ , and they have a bounded density ψ (*i.e.*, $\|\psi\| < c_\psi$, where $\|\cdot\|$ is the sup norm).*

Assumption 3. *The holding cost function $C : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is continuous, increasing and strictly convex. We also assume that C is $O(q^m)$ for some integer m .*

The polynomial form above is for technical reasons, but is not very restrictive since many convex functions can be approximated quite well.

3.2.3 Mean Field Equilibrium

The probability that the agent's bid lies in the interval $[0, x]$ is equal to the probability that the agent's queue length lies in some set whose best response is to bid between $[0, x]$.

Thus, the probability of the bid lying in the interval $[0, x]$ is $\Pi_\rho(\hat{\theta}_\rho^{-1}([0, x]))$, which we define as $\gamma(x)$. According to the assumed (cumulative) bid distribution, the probability of the same event is $\rho(x)$. If $\rho(x) = \gamma(x)$, it means that the assumed bid distribution is consistent with the best response bid distribution, and we have an MFE.

3.2.4 Agent's decision problem

Let the candidate be agent i . Suppose that the belief over the bid of a random agent has cumulative distribution ρ . We assume that $\rho \in \mathcal{P}$ where,

$$\mathcal{P} = \{G | G \text{ is a continuous c.d.f, } \int (1 - G(x))dx < E\},$$

for some $E < \infty$, to be defined later. Besides, its current state, the information available with the agent about the market at any time prior to the auction only includes the following:

1. The bids it made in each previous auction from last regeneration.
2. The auctions that it won.
3. The payments made for the auctions won.

Let $H_{i,k}$ be the history vector containing the above information available to agent i at time k . Suppose that the random variable representing the bid made by agent i at time k is denoted by $X_{i,k}$, with the realized value being $x_{i,k}$. Also, let $\bar{X}_{-i,k} = \max_{j \in M_{i,k}} X_{j,k}$, represent the maximum value of $M - 1$ draws from the distribution ρ . Thus, $\bar{X}_{-i,k}$ is the value of the highest opposing bid. The agent's decision problem is to choose a *bid function* θ_i , which maps its available information to a bid at each time $x_{i,k}$.

Since the time of regeneration T_i^k is a geometric random variable, the expected cost of agent i can be written as

$$V_{i,\rho}(H_{i,k}; \theta_i) = \mathbb{E} \left[\sum_{t=k}^{\infty} \beta^t [C(Q_{i,t}) + r_\rho(X_{i,t})] \right], \quad (3.1)$$

where the expectation is over future state evolutions. By replacing the belief with ρ , we have made an agent's decision problem independent of other agents' strategies. Hence, we represent the cost by $V_{i,\rho}(H_{i,k};\theta_i)$. Also, $r_\rho(x) = \mathbb{E}[\bar{X}_{-i,k}\mathbf{I}\{\bar{X}_{-i,k} \leq x\}]$ is the expected payment when i bids x under the assumption that the bids of other agents are distributed according to ρ . Hence, given ρ , the win probability in the auction is

$$p_\rho(x) = \mathbb{P}(\bar{X}_{-i,k} \leq x) = \rho(x)^{M-1}. \quad (3.2)$$

The expected payment when bidding x is,

$$r_\rho(x) = \mathbb{E}[\bar{X}_{-i,k}\mathbf{I}\{\bar{X}_{-i,k} \leq x\}] = xp_\rho(x) - \int_0^x p_\rho(u)du. \quad (3.3)$$

The state process $Q_{i,k}$ is Markov and has a transition kernel

$$\begin{aligned} \mathbb{P}(Q_{i,k+1} \in B | Q_{i,k} = q, X_{i,k} = x) = \\ \beta p_\rho(x)\mathbb{P}((q-1)^+ + A_k \in B) + \beta(1-p_\rho(x))\mathbb{P}(q + A_k \in B) + (1-\beta)\Psi(B), \end{aligned} \quad (3.4)$$

where $B \subseteq R^+$ is a Borel set and $x^+ \triangleq \max(x, 0)$. Recall that $A_k \sim \Phi$ is the arrival between $(k)^{th}$ and $(k+1)^{th}$ auction and Ψ is density function of the regeneration process. In the above expression, the first term corresponds to the event that agent wins the auction at time k , while the second corresponds to the event that it does not. The last term captures the event that the agent regenerates after auction k . The agent's decision problem can be modeled as an infinite horizon discounted cost MDP. It can be shown that there exists an optimal Markov deterministic policy for our MDP [53]. Then, from (3.1), the optimal value function of the agent is

$$\hat{V}_{i,\rho}(q) = \inf_{\theta_i \in \Theta} \mathbb{E} \left[\sum_{t=1}^{\infty} \beta^t [C(Q_{i,t}) + r_\rho(X_{i,t})] | Q_{i,0} = q \right], \quad (3.5)$$

where Θ is the space of Markov deterministic policies.

Note that user index is redundant in the above expression as we are concerned with a single agent's decision problem. In future notations, we will omit the user subscript i .

3.2.5 Invariant distribution

Given cumulative bid distribution ρ and a Markov policy $\theta \in \Theta$, the transition kernel given by (3.4) can be re-written as,

$$\begin{aligned} \mathbb{P}(Q_{k+1} \in B | Q_k = q) &= \beta p_\rho(\theta(q)) \mathbb{P}((q-1)^+ + A_k \in B) \\ &\quad + \beta(1 - p_\rho(\theta(q))) \mathbb{P}(q + A_k \in B) + (1 - \beta) \Psi(B). \end{aligned} \quad (3.6)$$

Then, we have an important result in the following lemma:

Lemma 1. *The Markov chain described by the transition probabilities in (3.6) is positive Harris recurrent and has a unique invariant distribution.*

Proof. From (3.6) we note that, $\mathbb{P}(Q_{k+1} \in B | Q_k = q) \geq (1 - \beta) \Psi(B)$, where $0 < \beta < 1$ and Ψ is a probability measure. The result then follows from results in Chapter 12, Meyn and Tweedie [38]. \square

We denote the unique invariant distribution by $\Pi_{\rho, \theta}$.

3.2.6 Mean field equilibrium

As described in the Introduction, the mean field equilibrium requires the consistency check that the bid distribution γ induced by the invariant distribution Π_{ρ, θ_ρ} should be equal to the bid distribution conjectured by the agent, i.e., ρ . Thus, we have the following definition of MFE:

Definition 1 (Mean field equilibrium). *Let ρ be a bid distribution and θ_ρ be a stationary policy for an agent. Then, we say that (ρ, θ_ρ) constitutes a mean field equilibrium if*

1. θ_ρ is an optimal policy of the decision problem in (3.5), given bid distribution ρ ; and

2. $\rho(x) = \gamma(x) \triangleq \Pi_\rho(\theta_\rho^{-1}([0, x])), \forall x \in \mathbb{R}^+, \text{ where } \Pi_\rho = \Pi_{\rho, \theta_\rho}.$

Note that the game theoretic definition of the MFE is considers the existence of an invariant distribution at a fixed time as the number of agents becomes asymptotically large. In keeping with extending the ideas of a Bayesian Nash Equilibrium to the system with a large number of agents, the definition does not require the occupancy distribution to converge to the invariant distribution from an arbitrary initial condition as time becomes large [26, 31, 3]. The approach may be contrasted with the stochastic systems literature that often studies the mean field in the case where both the number of controllers and time go to infinity (in some order), but the control policy is fixed (not a best response), and the objective is to show that interchanging limits of time and number of particles produces the same steady state distribution [5].

3.3 Properties of Optimal Bid Function

The decision problem given by (3.5) is an infinite horizon, discounted Markov decision problem. The optimality equation or Bellman equation corresponding to the decision problem is

$$\begin{aligned} \hat{V}_\rho(q) = & C(q) + \beta \mathbb{E}_A(\hat{V}_\rho(q + A)) \\ & + \inf_{x \in \mathbb{R}^+} \left[r_\rho(x) - p_\rho(x) \beta \mathbb{E}_A \left(\hat{V}_\rho(q + A) - \hat{V}_\rho((q - 1)^+ + A) \right) \right], \end{aligned} \quad (3.7)$$

where $A \sim \Phi$, and we use the notation $\max(0, z) = z^+$. Note that the decision problem above is independent of the regeneration distribution Ψ , since the game simply ends at any time with probability $1 - \beta$ from the agent's perspective. However, from a system perspective, the Markov chain describing the state transition is correctly represented by (3.6).

Define the set of functions

$$\mathcal{V} = \left\{ f : \mathbb{R}^+ \mapsto \mathbb{R}^+ : \sup_{q \in \mathbb{R}^+} \left| \frac{f(q)}{w(q)} \right| < \infty \right\}, \quad (3.8)$$

where $w(q) = \max\{C(q), 1\}$. Note that \mathcal{V} is a Banach space with w -norm,

$$\|f\|_w = \sup_{q \in \mathbb{R}^+} \left| \frac{f(q)}{w(q)} \right| < \infty.$$

Also, define the operator T_ρ as

$$(T_\rho f)(q) = C(q) + \beta \mathbb{E}_A f(q + A) + \inf_{x \in \mathbb{R}^+} [r_\rho(x) - p_\rho(x) \beta (\mathbb{E}_A (f(q + A) - f((q - 1)^+ + A)))] , \quad (3.9)$$

where $f \in \mathcal{V}$. It is straightforward to show that the infimum in the above operator occurs at

$$\beta \Delta f(q)^+, \quad (3.10)$$

where $\Delta f(q) = \mathbb{E}_A (f(q + A) - f((q - 1)^+ + A))$. Then, substituting from (3.2), (3.3) and (3.10), (3.9) can be rewritten as

$$(T_\rho f)(q) = C(q) + \beta \mathbb{E}_A f(q + A) - \int_0^{\beta \Delta f(q)^+} p_\rho(u) du. \quad (3.11)$$

The following lemma characterizes the optimal solution.

Lemma 2. *Given a cumulative bid distribution ρ ,*

1. *There exists a $j \in \mathbb{N}$ such that $T_\rho^j : \mathcal{V} \rightarrow \mathcal{V}$ is a contraction mapping. Hence, there exists a unique $f_\rho^* \in \mathcal{V}$ such that $T_\rho f_\rho^* = f_\rho^*$, and for any $f \in \mathcal{V}$, $T_\rho^n f \rightarrow f_\rho^*$ as $n \rightarrow \infty$.*
2. *The fixed point f_ρ^* of operator T_ρ is the unique solution to the optimality equation (3.7), i.e., $f_\rho^* = \hat{V}_\rho$.*
3. *Let, $\hat{\theta}_\rho(q) = \beta \Delta \hat{V}_\rho(q)^+$. Then, $\hat{\theta}_\rho$ is an optimal policy.*

Proof. The proof is similar to Theorem 8.3.6 in [25]. An exception to be noted here is that the action space in our case is not a compact set which violates Assumption 8.3.1(a) in

[25]. However, this assumption can be overridden if the statement of Lemma 8.3.8(a) in [25] holds true. This applies to our case since, as derived in (3.10), a minimizer exists for the infimum operator in (3.9) for every q . Further, Theorem 8.3.6 is specified with $j = 1$ (a one-step contraction). Hence, we replace Assumption 8.3.2(b) with an appropriate condition to obtain a j -step contraction. Please refer to Appendix A.1 for details. \square

Corollary 1. *An optimal policy of the agent's decision problem (3.5) is given by*

$$\hat{\theta}_\rho(q) = \beta \mathbb{E}_A \left[\hat{V}_\rho(q + A) - \hat{V}_\rho((q - 1)^+ + A) \right].$$

We now establish that \hat{V}_ρ and $\hat{\theta}_\rho$ are continuous and increasing functions.

Lemma 3. *Given a cumulative bid distribution function ρ*

1. \hat{V}_ρ is a continuous increasing function.
2. $\hat{\theta}_\rho$ is a continuous strictly increasing function.

Proof. Let $f \in \mathcal{V}$. Suppose f is a continuous monotone increasing function. We first prove that $T_\rho f$ is also continuous monotone increasing function. Since, $T_\rho^n f \rightarrow \hat{V}_\rho$ according to Statement 2 of Lemma 2, we conclude that \hat{V}_ρ also has the same property.

Let $q > q'$. Then,

$$\begin{aligned} T_\rho f(q) - T_\rho f(q') &= C(q) - C(q') + \beta \mathbb{E}_A(f(q + A) - f(q' + A)) \\ &\quad + \inf_x [r_\rho(x) - \beta p_\rho(x) \mathbb{E}_A(f(q + A) - f((q - 1)^+ + A))] \\ &\quad - \inf_x [r_\rho(x) - \beta p_\rho(x) \mathbb{E}_A(f(q' + A) - f((q' - 1)^+ + A))] \\ &\stackrel{(a)}{\geq} \beta \mathbb{E}_A(f(q + A) - f(q' + A)) + \beta \inf_x [p_\rho(x) \mathbb{E}_A(f(q' + A) \\ &\quad - f((q' - 1)^+ + A) - f(q + A) + f((q - 1)^+ + A))] \\ &\geq \beta \min \{ \mathbb{E}_A(f(q + A) - f(q' + A)) + \\ &\quad \mathbb{E}_A(f((q - 1)^+ + A) - f((q' - 1)^+ + A)) \} \stackrel{(b)}{\geq} 0, \end{aligned}$$

where (a) follows from the assumption that $C(\cdot)$ is an increasing function, and (b) follows from the assumption that $f(\cdot)$ is an increasing function.

To prove that $T_\rho f$ is continuous consider a sequence $\{q_n\}$ such that $q_n \rightarrow q$. Since f is a continuous function, $f(q_n + a) \rightarrow f(q + a)$. Then, by using dominated convergence theorem, we have $\mathbb{E}_A f(q_n + A) \rightarrow \mathbb{E}_A f(q + A)$ and $\mathbb{E}_A f((q_n - 1)^+ + A) \rightarrow \mathbb{E}_A f((q - 1)^+ + A)$. Also, $\Delta f(q_n) \geq 0$ as f is an increasing function. Then, from (3.11), we get that

$$\begin{aligned} T_\rho f(q_n) &= C(q_n) + \beta \mathbb{E}_A f(q_n + A) - \int_0^{\beta \Delta f(q_n)} p_\rho(u) du \\ &\rightarrow C(q) + \beta \mathbb{E}_A f(q + A) - \int_0^{\beta \Delta f(q)} p_\rho(u) du = T_\rho f(q). \end{aligned}$$

Hence, $T_\rho f$ is a continuous function. This yields Statement 1 in the lemma.

Now, to prove the second part, assume that Δf is an increasing function. First, we show that $\Delta T_\rho f$ is an increasing function. Let $q > q'$. From (3.11), for any $a < \bar{A}$ we can write

$$\begin{aligned} &(T_\rho f)(q + a) - (T_\rho f)((q - 1)^+ + a) - (T_\rho f)(q' + a) + (T_\rho f)((q' - 1)^+ + a) \\ &= C(q + a) - C((q - 1)^+ + a) - C(q' + a) + C((q' - 1)^+ + a) \\ &\quad + \beta \mathbb{E}_A f(q + a + A) - \beta \mathbb{E}_A f((q - 1)^+ + a + A) \\ &\quad - \beta \mathbb{E}_A f(q' + a + A) + \beta \mathbb{E}_A f((q' - 1)^+ + a + A) \\ &\quad - \int_{\beta \Delta f(q' + a)}^{\beta \Delta f(q + a)} p_\rho(u) du + \int_{\beta \Delta f((q' - 1)^+ + a)}^{\beta \Delta f((q - 1)^+ + a)} p_\rho(u) du \\ &= C(q + a) - C((q - 1)^+ + a) - C(q' + a) + C((q' - 1)^+ + a) \\ &\quad + \beta \mathbb{E}_A f((q + a - 1)^+ + A) - \beta \mathbb{E}_A f((q - 1)^+ + a + A) \\ &\quad - \beta \mathbb{E}_A f((q' + a - 1)^+ + A) + \beta \mathbb{E}_A f((q' - 1)^+ + a + A) \\ &\quad + \int_{\beta \Delta f(q' + a)}^{\beta \Delta f(q + a)} 1 - p_\rho(u) du + \int_{\beta \Delta f((q' - 1)^+ + a)}^{\beta \Delta f((q - 1)^+ + a)} p_\rho(u) du \end{aligned}$$

It can be easily verified that $\mathbb{E}_A(f(q + a - 1)^+ + A) - \mathbb{E}_A(f(q - 1)^+ + a + A) - \mathbb{E}_A(f(q' + a - 1)^+ + A) + \mathbb{E}_A(f(q' - 1)^+ + a + A) \geq 0$ as f is increasing (due to Statement 1 of this lemma).

From the assumption that Δf is increasing, the last two terms in the above expression are also non-negative. Now, taking expectation on both sides, we obtain $\Delta T_\rho f(q) - \Delta T_\rho f(q') \geq \Delta C(q) - \Delta C(q') > 0$. Therefore, from Statements 2 and 3 of Lemma 2, we have

$$\hat{\theta}_\rho(q) - \hat{\theta}_\rho(q') = \Delta \hat{V}_\rho(q) - \Delta \hat{V}_\rho(q') \geq \Delta C(q) - \Delta C(q') > 0.$$

Here, the last inequality holds since C is a strictly convex increasing function. \square

3.4 Existence of MFE

The main result showing the existence of MFE is as follows.

Theorem 4. *There exists an MFE $(\rho, \hat{\theta}_\rho)$ such that*

$$\rho(x) = \gamma(x) \triangleq \Pi_\rho \left(\hat{\theta}_\rho^{-1}[0, x] \right), \forall x \in \mathbb{R}^+.$$

We first introduce some useful notation. Let $\Theta = \{\theta : \mathbb{R} \mapsto \mathbb{R}, \sup_{q \in \mathbb{R}^+} \left| \frac{\theta(q)}{w(q)} \right| < \infty\}$. Note that Θ is a normed space with w -norm. Also, let Ω be the space of absolutely continuous probability measures on \mathbb{R}^+ . We endow this probability space with the topology of weak convergence. Note that this is same as the topology of point-wise convergence of continuous cumulative distribution functions.

We define $\theta^* : \mathcal{P} \mapsto \Theta$ as $(\theta^*(\rho))(q) = \hat{\theta}_\rho(q)$, where $\hat{\theta}_\rho(q)$ is the optimal bid given by Corollary 1. It can easily verified that $\hat{\theta}_\rho \in \Theta$. Also, define the mapping Π^* that takes a bid distribution ρ to the invariant workload distribution $\Pi_\rho(\cdot)$. Later, using Lemma 4 we will show that $\Pi_\rho(\cdot) \in \Omega$. Therefore, $\Pi^* : \mathcal{P} \rightarrow \Omega$. Finally, define \mathcal{F} as $(\mathcal{F}(\rho))(x) = \gamma(x) = \Pi_\rho(\hat{\theta}_\rho^{-1}([0, x]))$. Lemma 6 will show that \mathcal{F} maps \mathcal{P} into itself.

Now to prove the above theorem we need to show that \mathcal{F} has a fixed point, i.e $\mathcal{F}(\rho) = \rho$.

Theorem 5 (Schauder Fixed Point Theorem). *Suppose $\mathcal{F}(\mathcal{P}) \subset \mathcal{P}$. If $\mathcal{F}(\cdot)$ is continuous, and $\mathcal{F}(\mathcal{P})$ is contained in a convex and compact subset of \mathcal{P} , then $\mathcal{F}(\cdot)$ has a fixed point.*

In next section, we show that the mapping \mathcal{F} satisfies the conditions of the above theorem, and hence it has a fixed point. Note that \mathcal{P} is a convex set. Therefore, we just need to show that the other two conditions are satisfied.

3.5 MFE Existence: Proof

3.5.1 Continuity of the map \mathcal{F}

To prove the continuity of mapping \mathcal{F} , we first show that θ^* and Π^* are continuous mappings. To that end, we will show that for any sequence $\rho_n \rightarrow \rho$ in uniform norm, we have $\theta^*(\rho_n) \rightarrow \theta^*(\rho)$ in w -norm and $\Pi^*(\rho_n) \Rightarrow \Pi^*(\rho)$ (where \Rightarrow denotes weak convergence). Finally, we use the continuity of θ^* and Π^* to prove that $\mathcal{F}(\rho_n) \rightarrow \mathcal{F}(\rho)$.

Step 1: Continuity of θ^* : First we will look at continuity of θ^* .

Theorem 6. *The map θ^* is continuous.*

Proof. Define the map $V^* : \mathcal{P} \mapsto \mathcal{V}$ that takes ρ to $\hat{V}_\rho(\cdot)$. We begin by showing that $\|\hat{\theta}_{\rho_1} - \hat{\theta}_{\rho_2}\|_w \leq K\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w$, which means that the continuity of the map V^* implies the continuity of the map θ^* .

Then we show two simple properties of the Bellman operator. The first is that for any $\rho \in \mathcal{P}$ and $f_1, f_2 \in \mathcal{V}$,

$$\|T_\rho f_1 - T_\rho f_2\|_w \leq \hat{K}\|f_1 - f_2\|_w \quad (3.12)$$

for some large \hat{K} , independent of ρ . This result is available in Lemma 32 in Appendix A.2.1.

Second, let T_{ρ_1} and T_{ρ_2} be the Bellman operators corresponding to $\rho_1, \rho_2 \in \mathcal{P}$ and let $f \in \mathcal{V}$. We show that

$$\|T_{\rho_1} f - T_{\rho_2} f\|_w \leq 2(M - 1)K_1\|f\|_w\|\rho_1 - \rho_2\|. \quad (3.13)$$

This result is available in Lemma 33 in Appendix A.2.1.

We then have

$$\begin{aligned}
\|T_{\rho_1}^j \hat{V}_{\rho_2} - T_{\rho_2}^j \hat{V}_{\rho_2}\|_w &\leq \|T_{\rho_1}^j \hat{V}_{\rho_2} - T_{\rho_1}^{j-1} T_{\rho_2} \hat{V}_{\rho_2}\|_w \\
&\quad + \|T_{\rho_1}^{j-1} T_{\rho_2} \hat{V}_{\rho_2} - T_{\rho_1}^{j-2} T_{\rho_2}^2 \hat{V}_{\rho_2}\|_w + \dots + \|T_{\rho_1} T_{\rho_2}^{j-1} \hat{V}_{\rho_2} - T_{\rho_2}^j \hat{V}_{\rho_2}\|_w \\
&\leq \hat{K}^{j-1} \|T_{\rho_1} \hat{V}_{\rho_2} - T_{\rho_2} \hat{V}_{\rho_2}\|_w + \dots + \|T_{\rho_1} T_{\rho_2}^{j-1} \hat{V}_{\rho_2} - T_{\rho_2}^j \hat{V}_{\rho_2}\|_w \quad (3.14) \\
&\leq (\hat{K}^{j-1} + \dots + 1) \|T_{\rho_1} \hat{V}_{\rho_2} - T_{\rho_2} \hat{V}_{\rho_2}\|_w \\
&\leq 2(M-1)K \|\rho_1 - \rho_2\| (\hat{K}^{j-1} + \dots + 1) \|\hat{V}_{\rho_2}\|_w \quad (3.15)
\end{aligned}$$

Here, (3.14) and (3.15) follow from (3.12) and (3.13), respectively.

Now, let j be such that $T_{\rho_1}^j$ is an α -contraction, which is guaranteed to exist by lemma 2. Note that Statement 1 of Lemma 2 implies that such a $j < \infty$ exists. Then we have

$$\begin{aligned}
\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w &= \|T_{\rho_1}^j \hat{V}_{\rho_1} - T_{\rho_2}^j \hat{V}_{\rho_2}\|_w \\
&\leq \|T_{\rho_1}^j \hat{V}_{\rho_1} - T_{\rho_1}^j \hat{V}_{\rho_2}\|_w + \|T_{\rho_1}^j \hat{V}_{\rho_2} - T_{\rho_2}^j \hat{V}_{\rho_2}\|_w \\
&\implies (1 - \alpha) \|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w \leq \|T_{\rho_1}^j \hat{V}_{\rho_2} - T_{\rho_2}^j \hat{V}_{\rho_2}\|_w \quad (3.16)
\end{aligned}$$

Finally, from (3.15) and (3.16), we get

$$\begin{aligned}
\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w &\leq \frac{2(m-1)K(\hat{K}^{j-1} + \dots + 1) \|\rho_1 - \rho_2\|}{1 - \alpha} \|\hat{V}_{\rho_2}\|_w \\
&\leq \frac{2(m-1)K(\hat{K}^{j-1} + \dots + 1) \|\rho_1 - \rho_2\|}{1 - \alpha} \\
&\quad \times (\|\hat{V}_{\rho_1}\|_w + \|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w).
\end{aligned}$$

Therefore, if $\frac{2(m-1)K(\hat{K}^{j-1} + \dots + 1)}{1 - \alpha} \|\rho_1 - \rho_2\| < \frac{1}{2}$, then

$$\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w \leq \frac{4(m-1)K(\hat{K}^{j-1} + \dots + 1)}{1 - \alpha} \|\hat{V}_{\rho_1}\|_w \|\rho_1 - \rho_2\|$$

Hence, the maps V^* and θ^* are continuous. □

Step 2: Continuity of the map Π^* : Let $\Pi_{\rho,\theta}(\cdot)$ be the invariant distribution generated by any θ . Recall that Π^* takes $\rho \in \mathcal{P}$ to probability measure $\Pi_\rho(\cdot) = \Pi_{\rho,\hat{\theta}_\rho}(\cdot)$. First, we show that $\Pi_{\rho,\theta}(\cdot) \in \Omega$, where Ω is the space of absolutely continuous measures (with respect to Lebesgue measure) on \mathbb{R}^+ .

Lemma 4. *For any $\rho \in \mathcal{P}$ and any $\theta \in \Theta$, $\Pi_{\rho,\theta}(\cdot)$ is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^+ .*

Proof. $\Pi_{\rho,\theta}(\cdot)$ can be expressed as the invariant queue-length distribution of the dynamics

$$q \rightarrow \begin{cases} Q' + A & \text{with probability } \beta \\ R & \text{with probability } (1 - \beta), \end{cases}$$

where $A \sim \Phi$ and $R \sim \Psi$, and Q' is a random variable with distribution generated by the conditional probabilities

$$\mathbb{P}(Q' = q|q) = 1 - p_\rho(\hat{\theta}(q))$$

$$\mathbb{P}(Q' = (q - 1)^+|q) = p_\rho(\hat{\theta}(q))$$

Let Π' be the distribution of Q' . Then for any Borel set B , Π can be expressed using the convolution of Π' and Φ :

$$\Pi_{\rho,\theta}(B) = \beta \int_{-\infty}^{\infty} \Phi(B - y) d\Pi'(y) + (1 - \beta)\Psi(B). \quad (3.17)$$

If B is a Lebesgue null-set, then so is $B - y \ \forall y$. So, $\Phi(B - y) = 0$ and $\Psi(B) = 0$ and therefore $\Pi_\rho(B) = 0$. □

We now develop a useful characterization of $\Pi_{\rho,\theta}$. Let

$$\Upsilon_{\rho,\theta}^{(k)}(B|q) = \mathbb{P}(Q_k \in B | \text{no regeneration}, Q_0 = q)$$

be the distribution of queue length Q_k at time k induced by the transition probabilities (3.6) conditioned on the event that $Q_0 = q$ and that there are no regenerations until time k . We can now express the invariant distribution $\Pi_{\rho,\theta}(\cdot)$ in terms of $\Upsilon_{\rho,\theta}^{(k)}(\cdot|q)$ as in the following lemma.

Lemma 5. *For any bid distribution $\rho \in \mathcal{P}$ and for any stationary policy $\theta \in \Theta$, the Markov chain described by the transition probabilities in (3.6) has a unique invariant distribution $\Pi_{\rho,\theta}(\cdot)$. Also $\Pi_{\rho,\theta}$ and $\Upsilon_{\rho,\theta}^{(k)}$ are related as follows:*

$$\Pi_{\rho,\theta}(B) = \sum_{k \geq 0} (1 - \beta) \beta^k \mathbb{E}_{\Psi}(\Upsilon_{\rho,\theta}^{(k)}(B|Q)), \quad (3.18)$$

where $\mathbb{E}_{\Psi}(\Upsilon_{\rho,\theta}^{(k)}(B|Q)) = \int \Upsilon_{\rho,\theta}^{(k)}(B|q) d\Psi(q)$.

Proof. $\Upsilon_{\rho,\theta}^{(k)}(B|q)$ is the queue length distribution assuming no regeneration has happened yet, and the regeneration event occurs with probability β independently of the rest of the system. It is then easy to find $\Pi_{\rho,\theta}(B)$ in terms of $\Upsilon_{\rho,\theta}^{(k)}(B|q)$ by simply using the properties of the conditional expectation, and the theorem follows. Note that in $\mathbb{E}_{\Psi}(\Upsilon_{\rho,\theta}^{(k)}(B|Q))$, the random variable is the initial condition of the queue, as generated by Ψ . Full details are available in Appendix A.2.2. \square

We shall now prove the continuity of Π^* in ρ .

Theorem 7. *The mapping $\Pi^* : \mathcal{P} \mapsto \Omega$ is continuous.*

Proof. By Portmanteau theorem [8], we only need to show that for any sequence $\rho_n \rightarrow \rho$ in w -norm and any open set B , $\liminf_{n \rightarrow \infty} \Pi_{\rho_n}(B) \geq \Pi_{\rho}(B)$. By Fatou's lemma,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \Pi_{\rho_n}(B) &= \liminf_{n \rightarrow \infty} \sum_{k=0}^{\infty} (1 - \beta) \beta^k \mathbb{E}_{\Psi_R}[\Upsilon_{\rho_n}^{(k)}(B|Q)] \\ &\geq \sum_{k=0}^{\infty} (1 - \beta) \beta^k \mathbb{E}_{\Psi_R}[\liminf_{n \rightarrow \infty} \Upsilon_{\rho_n}^{(k)}(B|Q)] \end{aligned} \quad (3.19)$$

where $Q \sim \Psi_R$. Let $\Upsilon_\rho^{(k)} = \Upsilon_{\rho, \hat{\theta}_\rho}^{(k)}$. We prove in Lemma 34 (see Appendix A.2.2) that $\liminf_{n \rightarrow \infty} \Upsilon_{\rho_n}^{(k)}(B|q) \geq \Upsilon_\rho^{(k)}(B|q)$ for every $q \in \mathbb{R}^+$, and the proof follows. \square

Step 3: Continuity of the mapping \mathcal{F} : Now, using the results from Step 1 and Step 2, we establish continuity of the mapping \mathcal{F} . First, we show that $\mathcal{F}(\rho) \in \mathcal{P}$.

Lemma 6. *For any $\rho \in \mathcal{P}$, let $\gamma(x) = (\mathcal{F}(\rho))(x) = \Pi_\rho(\hat{\theta}_\rho^{-1}([0, x]))$, $x \in \mathbb{R}^+$. Then, $\gamma \in \mathcal{P}$.*

Proof. From the definition of Π_ρ , it is easy to see that γ is a distribution function. Since $\hat{\theta}_\rho$ is continuous and strictly increasing function as shown in Lemma 3, $\hat{\theta}_\rho^{-1}(\{x\})$ is either empty or a singleton. Then, from Lemma 4, we get that $\Pi_\rho(\hat{\theta}_\rho^{-1}(\{x\})) = 0$. Together, we get that $\gamma(x)$ has no jumps at any x and hence it is continuous.

To complete the proof, we need to show that the expected bid under $\gamma(\cdot)$ is finite. In order to do this, we construct a new random process \tilde{Q}_k that is identical to the original queue length dynamics Q_k , except that it never receives any service. We show that this process stochastically dominates the original, and use this property to bound the mean of the original process by a finite quantity independent of ρ . Full details are presented in Appendix A.2.3. \square

We now have the main theorem.

Theorem 8. *The mapping $\mathcal{F} : \mathcal{P} \mapsto \mathcal{P}$ given by $(\mathcal{F}(\rho))(x) = \Pi_\rho(\hat{\theta}_\rho^{-1}([0, x]))$ is continuous.*

Proof. Let $\rho_n \rightarrow \rho$ in uniform norm. From previous steps, we have $\hat{\theta}_{\rho_n} \rightarrow \hat{\theta}_\rho$ in w -norm and $\Pi_{\rho_n} \Rightarrow \Pi_\rho$. Then, using Theorem 5.5 of Billingsley [8], one can show that the push-forwards also converge:

$$\Pi_{\rho_n}(\hat{\theta}_{\rho_n}^{-1}(\cdot)) \Rightarrow \Pi_\rho(\hat{\theta}_\rho^{-1}(\cdot)).$$

Then, $\mathcal{F}(\rho_n)$ converges point-wise to $\mathcal{F}(\rho)$ as it is continuous at every x , i.e., $(\mathcal{F}(\rho_n))(x) \rightarrow (\mathcal{F}(\rho))(x)$ for all $x \in \mathbb{R}^+$.

It is easy to show that in the norm space \mathcal{P} , point-wise convergence implies convergence in uniform norm. This result is proved in Lemma 35 in Appendix A.2.3. This completes the proof. \square

3.5.2 $\mathcal{F}(\mathcal{P})$ contained in a compact subset of \mathcal{P}

We show that the closure of the image of the mapping \mathcal{F} , denoted by $\overline{\mathcal{F}(\mathcal{P})}$, is compact in \mathcal{P} . As \mathcal{P} is a normed space, sequential compactness of any subset of \mathcal{P} implies that the subset is compact. Hence, we just need to show that $\overline{\mathcal{F}(\mathcal{P})}$ is sequentially compact. Sequential compactness of a set $\overline{\mathcal{F}(\mathcal{P})}$ means the following: if $\{\rho_n\} \in \overline{\mathcal{F}(\mathcal{P})}$ is a sequence, then there exists a subsequence $\{\rho_{n_j}\}$ and $\rho \in \overline{\mathcal{F}(\mathcal{P})}$ such that $\rho_{n_j} \rightarrow \rho$. We use Arzelà-Ascoli theorem and uniform tightness of the measures in $\mathcal{F}(\mathcal{P})$ to show the sequential compactness. The version that we will use is stated below:

Theorem 9 (Arzelà-Ascoli Theorem). *Let X be a σ -compact metric space. Let \mathcal{G} be a family of continuous real valued functions on X . Then the following two statements are equivalent:*

1. *For every sequence $\{g_n\} \subset \mathcal{G}$ there exists a subsequence g_{n_j} which converges uniformly on every compact subset of X .*
2. *The family \mathcal{G} is equicontinuous on every compact subset of X and for any $x \in X$, there is a constant C_x such that $|g(x)| < C_x$ for all $g \in \mathcal{G}$.*

Suppose a family of functions $\mathcal{D} \subseteq \mathcal{P}$ satisfies the equivalent conditions of the Arzelà-Ascoli theorem and in addition satisfy the uniform tightness property, i.e., $\forall \epsilon > 0$ there exists x_ϵ such that for all $f \in \mathcal{D}$ $1 \geq f(x_\epsilon) \geq 1 - \epsilon$. Then, for any sequence $\{\rho_n\} \subset \mathcal{D}$, there exists a subsequence $\{\rho_{n_j}\}$ that converges uniformly on every compact set to a continuous increasing function ρ on \mathbb{R}^+ . As \mathcal{D} is uniformly tight it can be shown that ρ_{n_j} converges uniformly to ρ and that $\rho \in \mathcal{P}$. Therefore, $\overline{\mathcal{D}}$ is sequentially compact in the topology of uniform norm.

In the following, we show that $\mathcal{F}(\mathcal{P})$ satisfies uniform tightness property and condition 2 in Arzelà-Ascoli theorem. First verifying the conditions of Arzelà-Ascoli theorem, note that

the functions in consideration are uniformly bounded by 1. To prove equicontinuity, consider a $\gamma = \mathcal{F}(\rho)$ and let $x > y$.

$$\begin{aligned}\gamma(x) - \gamma(y) &= \Pi_\rho(\theta_\rho(q) \leq x) - \Pi_\rho(\theta_\rho(q) \leq y) \\ &= \Pi_\rho(y < \theta_\rho(q) \leq x)\end{aligned}\tag{3.20}$$

Lemma 7. *For any interval $[a, b]$, $\Pi_\rho([a, b]) < c \cdot (b - a)$, for some large enough c .*

Proof. The proof follows easily from our characterization of Π_ρ in terms of $\Upsilon_\rho^{(k)}$. \square

The above lemma and equation (3.20) imply that $\gamma(x) - \gamma(y) \leq c(\theta_\rho^{-1}(x) - \theta_\rho^{-1}(y))$. To show equicontinuity, it is enough to show that $\limsup_{y \uparrow x} \frac{\gamma(x) - \gamma(y)}{x - y} \leq K(x)$ for some K independent of ρ , which we will show now.

$$\begin{aligned}\limsup_{y \uparrow x} \frac{\gamma(x) - \gamma(y)}{x - y} &= \limsup_{y \uparrow x} \frac{\Pi_\rho(y < \theta_\rho(q) \leq x)}{x - y} \\ &= \limsup_{y \uparrow x} \frac{\Pi_\rho([\theta_\rho^{-1}(y), \theta_\rho^{-1}(x)])}{x - y} \\ &\leq c \limsup_{y \uparrow x} \frac{\theta_\rho^{-1}(x) - \theta_\rho^{-1}(y)}{x - y} \\ &= c \limsup_{y \uparrow x} \frac{\theta_\rho^{-1}(x) - \theta_\rho^{-1}(y)}{\theta_\rho \theta_\rho^{-1}(x) - \theta_\rho \theta_\rho^{-1}(y)}\end{aligned}$$

Let $x' = \theta_\rho^{-1}(x)$ and $y' = \theta_\rho^{-1}(y)$. Now,

$$\begin{aligned}\limsup_{y \uparrow x} \frac{\gamma(x) - \gamma(y)}{x - y} &\leq c \limsup_{y' \rightarrow x'} \frac{x' - y'}{\theta_\rho(x') - \theta_\rho(y')} \\ &= c \limsup_{y' \rightarrow x'} \frac{x' - y'}{\beta \Delta V(x') - \beta \Delta V(y')} \\ &\leq c \limsup_{y' \rightarrow x'} \frac{x' - y'}{\beta (\Delta C(x') - \Delta C(y'))} \\ &\leq c \frac{1}{H(x')}\end{aligned}$$

Where,

$$0 < H(x') = \begin{cases} \mathbb{E}_A[C'(x' + A) - C'(\overline{x-1} + A)] & x' > 1 \\ \mathbb{E}_A[C'(x' + A)] & x' \leq 1 \end{cases}$$

and $C'(x) = \frac{dC(x)}{dx}$.

Finally, we have the following lemma showing that $\mathcal{F}(\mathcal{P})$ is uniformly tight.

Lemma 8. $\mathcal{F}(\mathcal{P})$ is uniformly tight, i.e., for any $\epsilon > 0$ and any $f \in \mathcal{F}(\mathcal{P})$, there exists an $x_\epsilon \in \mathbb{R}$ such that $1 - \epsilon \leq f(x_\epsilon) \leq 1$.

Proof. From Lemma 6, we have $\mathcal{F}(\mathcal{P}) \subseteq \mathcal{P}$. Hence, the expectation of the bid distributions in $\mathcal{F}(\mathcal{P})$ is bounded uniformly. An application of Markov inequality will give uniform tightness. \square

3.6 Approximation Results: PBE and MFE

In this section we prove that the mean field policy is an ϵ -Nash equilibrium. We have the following theorem:

Theorem 10. Let $(\rho, \hat{\theta}_\rho)$ constitute an MFE. Suppose at time 0 the queue length of the users is set independently across users according to Π_ρ ; and that their initial belief is also consistent. Also, suppose that all queues except queue 1 play the MFE policy $\hat{\theta}_\rho$. Then, for any policy θ^N of queue 1 that may be history dependent and any $q \in \mathbb{R}^+$, we have

$$\limsup_{N \rightarrow \infty} V_{1, \mu_{1,0}}^N(q; \hat{\theta}_\rho, (\hat{\theta}_\rho)_{-1}) - V_{1, \mu_{1,0}}^N(q; \theta^N, (\hat{\theta}_\rho)_{-1}) \leq 0,$$

where $\mu_{1,0} = \Pi_\rho$ and the superscript N has been added to explicitly indicate the dependence on the number of cells.

The main idea behind the proof is a result called *propagation of chaos*, and it identifies conditions under which any finite subset of the state variables are independent from each

other. We state this result now in our context. We only provide brief sketches of proofs in this section, since the methodology is much the same as [20] and space constraints do not allow us to present the full version of the proofs here.

Lemma 9 (Propagation of chaos). *For any fixed indices i_1, \dots, i_k , let $\mathcal{L}(Q_{i_1}^N(t), \dots, Q_{i_k}^N(t))$ denote the probability law of the k -tuple of corresponding queues in the MN -queue system, at time t . Suppose that $\mathcal{L}((Q_{i_1}^N(0), \dots, Q_{i_k,0}^N(0))) = \otimes^k \Pi_\rho$, where (ρ, θ_ρ) is the solution to the MFE equation. Also, suppose that all queues are following mean field equilibrium strategy. Then for any $T > 0$, we have*

$$\mathcal{L}(Q_{i_1}^N(T), \dots, Q_{i_k}^N(T)) \Rightarrow \otimes^k \Pi_\rho,$$

as $N \rightarrow \infty$.

Proof. We shall only consider the case $k = 2$; the proof of the general case is similar. We can follow the proof of Theorems 4.1 and 5.1 in Graham and Meleard [20]. The proof is divided into two parts; the first part proves that for any two agents i and j ,

$$\|\mathcal{L}(Q_i^N, Q_j^N) - \mathcal{L}(Q_i^N) \otimes \mathcal{L}(Q_j^N)\|_D \rightarrow 0,$$

where the subscript D refers to the total variation norm. In the second, we show that $\mathcal{L}(Q_i^N) \Rightarrow \Pi_\rho$. Both parts rely on studying interaction graphs, defined in [20], which characterize the amount of interactions that any finite subset of agents may have had in the past. □

The proof of Theorem 10 is as follows. Suppose we start at time $t = 0$ with queue length of agent 1 being $Q_1(0)$. We can choose a time T large enough so that the value added by auctions occurring after time T is less than ϵ , due to discounting. Thus, the difference in value contributed by these auctions, when using policy θ^N and $\hat{\theta}_\rho$ can be bounded by 2ϵ , and we can restrict attention to the first T auctions.

Using ideas similar to Lemma 9, we show that probability of the event E^N that other agents that interact with agent 1 at time t have never been influenced by agent 1 goes to 1 as N becomes large. Thus, the belief of distribution of queue lengths of other agents encountered converges to Π_ρ according to Lemma 9. Then we can show that for any $\epsilon > 0$ and N ,

$$V_{1,\mu_{1,0}}^N(q; \theta_\rho, (\theta_\rho)_{-1}) - V_{1,\mu_{1,0}}^N(q; \theta^N, (\theta_\rho)_{-1}) \leq \epsilon,$$

which yields the desired result.

3.7 Simulation Results

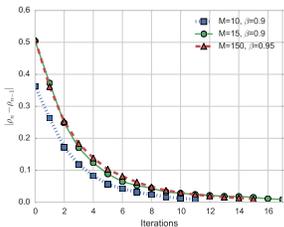


Figure 3.2: Convergence to MFE bid distribution.

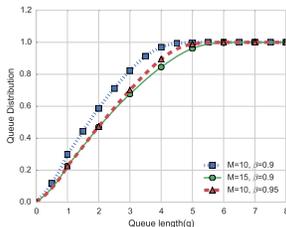


Figure 3.3: MFE queue length distribution.

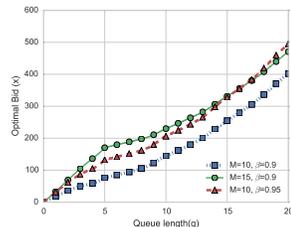


Figure 3.4: MFE optimal bid function.

We now turn to computing the MFE distribution. We simulate a large system with 100,000 users distributed among 10,000 cells with 10 users per cell. For simplicity of simulation, we truncate and discretize both state and bid spaces. The truncated state space is $\mathcal{S} = \{0.01m, 0 \leq m \leq 2000\}$, while the bid space is $\mathcal{X} = \{0.15m, 0 \leq m \leq 3000\}$. The job arrival and regeneration distributions are both chosen to be uniform over interval $[0, 1]$. The service rate of each base station is assumed to be 5 units per time slot. Finally, the holding cost function is chosen as $C(q) = q^2$.

Our simulation simply follows the choices made by each agent and calculating the empirical distribution that would result at each time step. Let $\rho_0(x) = \min\{0.001x, 1\}$, $x \text{ in } \mathcal{X}$

and $\Pi_0 = \Psi$. For every positive integer n , do the following:

1. Compute the optimal value function, \hat{V}_n , which is the unique solution to the following equation,

$$\hat{V}_n(q) = C(q) + \beta \mathbb{E}_A[\hat{V}_n(q + A)] - \sum_{x \leq \beta \Delta \hat{V}_n(q), x \in \mathcal{X}} p_{\rho_n}(x),$$

where $\Delta \hat{V}_n(q) = \mathbb{E}_A[\hat{V}_n(q + A)] - \mathbb{E}_A[\hat{V}_n((q - 1)^+ + A)]$. We apply value iteration (Section 6.10 [46]) to compute an approximate solution to the above equation.

2. Compute the optimal bid function, $\hat{\theta}_n$ as

$$\hat{\theta}_n(q) = \beta \mathbb{E}_A[\hat{V}_n(q + A) - \hat{V}_n((q - 1)^+ + A)].$$

3. Next, all agents employ the optimal bid policy from Step 2 and update their states. Then we compute the candidate steady state distribution by evaluating empirical distribution of state.
4. Finally, compute the empirical bid distribution, $\rho_{n+1}(x) = \Pi_n[\hat{\theta}_n^{-1}([0, x])]$. If $\|\rho_n - \rho_{n+1}\| < \epsilon$, then $\rho = \rho_{n+1}$ and exit. Otherwise, set $n = n + 1$ and go to Step 1.

If the algorithm converges, then its output distribution ρ , is an approximation of the MFE bid distribution.

We simulated the algorithm for three set of parameters: 1. ($\beta = 0.9, M = 10$), 2. ($\beta = 0.95, M = 10$) and 3. ($\beta = 0.9, M = 15$). Also, we chose the accuracy parameter $\epsilon = 0.008$. Figure 3.2 shows that the algorithm converges in less than 20 iterations in all three cases. In each iteration, Step 1 (value iteration) is the most computationally intensive. It converges in 80 recursions, with each recursion having to update $|\mathcal{S}|$ number of variables, and with each variable update requiring at most $|\mathcal{X}|$ number of arithmetic operations. All together, the computational complexity of each iteration is in the order of $80 \times |\mathcal{S}| \times |\mathcal{X}|$ arithmetic operations.

The queue length distributions at MFE are shown in Figure 3.3. We observe that the distribution curves exhibit a rightward shift with increase in β or M . Note that larger β makes queues live longer without regeneration, while higher M reduces each individual's average service rate. Hence, the queues get longer on average. We show the optimal bid functions at MFE in Figure 3.4. As expected from our analysis, the bid functions are monotonically increasing in queue length.

3.8 Conclusions

Our algorithm for computing the MFE immediately suggests a simple implementation scheme. Suppose each mobile device has a network interface manager on which the human user sets up priorities for different apps. The interface manager also should be aware of the cost functions corresponding to the QoE of different apps. Now, suppose that the base stations were to calculate the empirical bid distribution at each time instant, and return it to the interface manager. The interface manager plays its best response to this bid distribution. Value iteration could be done either on each device or at a data center and provided as a look up table to the interface manager. The base stations combine all the bids to create a new empirical bid distribution. Such a proceeding is essentially identical to the algorithm that we employed above, and would converge in a similar fashion.

We had assumed a single cost function for the agent (app), but as long as there are a finite set of cost functions (corresponding to a finite number of app types) we can incorporate the cost function as part of the state of the agent, with the cost function being chosen according to some distribution at each regeneration (corresponding to choosing to start a new app with some probability). Such a modification causes no changes to any of our analysis. Then the mean field bid distribution accounts for the distribution of cost functions (app popularities), and the agent takes a decision based on this distribution as before. The only difference to the achieved equilibrium is that it now follows a weighted version of LQF, with the weights corresponding to the cost on QoE of the competing apps.

To summarize our work, we explored the question of whether it is possible to design an

scheduling policy that allows for declaration of value by humans in the loop and effects on QoE, and which has the attractive properties of a (weighted) LQF service regime. We used a mean field framework to show that as the number of agents in the system becomes large, this objective can indeed be fulfilled using a second price auction at each server. Our design appears to lend itself well to implementation and this will be our future goal.

In this chapter, we studied a large scale market of agents (queues) competing for service, and established the existence of an efficient Mean Field Equilibrium. In the next chapter, we will consider a complementary problem of competition across many small scale providers that participate in an Internet marketplace.

4. MEAN FIELD EQUILIBRIA OF PRICING AND WORK-QUALITY SELECTION GAMES IN INTERNET MARKETPLACES

4.1 Introduction

The world has recently seen an explosion in the growth of Internet marketplaces. While large scale online goods retailers are well established, recently developed market places have allowed small scale players to participate, as well as for the proliferation of the sale of services as well as goods. For example, Amazon Mechanical TurkTM and Hacker’s List¹ are marketplaces to match tasks whose completion requires human intelligence, with appropriate individuals that can perform them. Marketplaces that enable small scale goods sellers to reach their target audience range from broad ones such as Amazon and Ebay, or more specialized ones such as Swappa² that focuses on used smart phones.

The heterogeneous mix of agents who have different lifetimes and differing abilities to complete jobs has meant that establishing credentials in a marketplace is usually by means of a reputation scheme, wherein users that have interacted with these agents leave feedback based on their experience. It has been empirically observed that having a higher reputation carries a sale price premium to the tune of about 5 – 10% [19]. To observe this phenomenon in practice, we collected the data from Swappa.com, which is an online market place for the sale of used smart phones. Here, a seller is allowed to indicate the condition of the phone (fair, good, mint or new), and customers provide star ratings of the sellers. Each successful transaction is posted under the heading of “recent sales” with the make and model of the phone, the sale price and the rating of the seller. We collected data for 2 months and obtained 4233 unique data points. We focussed on the two major service providers in the US (AT&T and Verizon), so as to avoid effects of lower demand associated with smaller carriers.

We found that reputation on Swappa tends to pool around 1 star or 5 stars, with relatively

¹<http://hackerslist.com>

²<http://swappa.com>

few points in-between. Hence, we divide our data set into two classes of reputation 0 and 1. The distribution of the sale price of phones at the two different reputations for different conditions of phones is plotted in the Figure 4.1. The “violin plot” shown is a box plot along with a depiction of the empirical distribution of the data around it. We note that the effect of reputation on sale prices appears to be uniform across the data, and a high reputation allows for a markup of about \$50 on average. For example, a phone declared as “fair” by a seller with a low reputation sells at a price that is less by \$50 (on average) than a phone declared to be in the same condition by one with a high reputation.

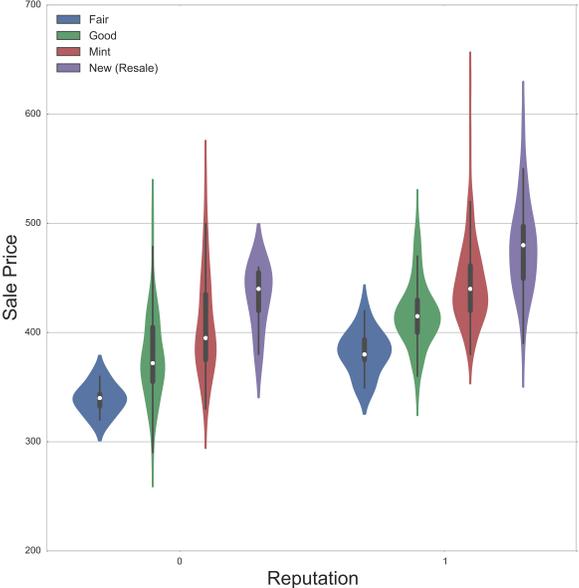


Figure 4.1: The effect of reputation on sale price. Reputation is classified as “high” (=1) or “low” (=0). We consider four conditions of phones: fair, good, mint and new. Comparing phones of the same condition (shown in the same color), we observe a markup of about \$50 for a high reputation.

Our objective in this chapter is to study the dynamics of competition across agents providing a nominally identical service or good in an Internet marketplace. Here, the agents (sellers) offering the service have an inherent competence or quality of service that is unknown

to potential customers. This “work-quality” parametrizes characteristics such as how well the agent is capable of performing the job (eg. on Hacker’s List), the customer service that he provides (eg. on Amazon), whether he truthfully reveals the condition of the good (eg. on Swappa) and so on. Customers do not see the work-quality directly, but are only aware of the posted reputation. As such, they take a decision on which agent to choose for service based on a combination of the offered price and reputation. The customer updates the reputation once the job is completed.

It is observed in [19] that delays in processing jobs are considered quite unfavorably, and indeed marketplaces such as Mechanical Turk allow the imposition of deadline by which jobs must be completed, while sites such as Amazon allow the customer to cancel the transaction without loss until the item is shipped. Consistent with this idea, we assume that the marketplace is interested in ensuring high throughput and low delays, and hence imposes a holding cost on each seller based on its number of pending jobs. The agent providing service has a random lifetime, after which he leaves the system. Such behavior is typical of small scale service providers and retailers.

How should an agent with a certain reputation and a certain number of pending jobs to be completed price his good or service against competing offers? The challenge faced by an agent is to tradeoff the need to set a low price to incentivize customers to explore its service (and potentially enhance its reputation), while at the same time trying to maximize its payoff at the current reputation by setting higher prices. While doing so, the agent needs to model the likely prices set by the competing providers.

Modeling the actions of each individual competitor in a marketplace where the number of agents is very large is prohibitively complex for any particular agent. Indeed, given that a potential customer only views the offered prices and reputations of a finite subset³ of agents, a particular agent is unlikely to have gone head-to-head with any of the competitors it faces for some particular job if the agent lifetimes are all randomly drawn. Thus, it might in fact

³This subset will be platform-specific but typically a small and random set of agents, even with specific search terms included.

be unnecessary for any agent to compute the behavior of each potential competitor in order to determine the optimal price. This setting of infinite agents, each with ephemeral lifetimes and with random finite subsets interacting with each other at each step is the typical setting of a mean field game.

4.1.1 Mean Field Games

The problem of an agent estimating the probable actions of a randomly drawn subset of competitors at each time, taking a best response action, and updating the belief using Bayes' Rule after observing the realized actions of the competitors is the scenario in a repeated Bayesian game. A Mean Field Game (MFG) is the limiting case of a Bayesian game under which the number of agents becomes large, and the interactions between any finite subset of agents becomes infrequent. When these conditions are met, it is an accurate approximation to assume that the states and hence the actions of any finite subset of competitors are independent of each other, which is the so-called chaos hypothesis [20]. This hypothesis provides the basis for the MFG [29] under which we consider the system from the viewpoint of a single agent that has a belief that his opponents actions would be drawn independently of each other from some candidate probability distribution, and takes a best response action based on his current state and this belief. The system is said to be at a Mean Field Equilibrium (MFE) if this action is itself a sample consistent with the candidate distribution.

The MFG approach has recently been used in several different problems on games with a large number of agents, each subset of which meet infrequently [55, 2, 9, 56, 26, 34, 31, 32]. Most pertinent to our work are [56, 26, 34]. In [56] jobs arrive into a system of many queues and must decide how many queues to sample, with the objective of joining the least loaded queue. There is a cost to sampling each queue and hence each job must take a decision based on its belief about how many queues are being sampled by the other jobs. The problem of auctioning spots on a webpage to competing advertisers is considered in [26], where the set of competing advertisers at each auction are drawn randomly from a large pool. The

advertisers must learn the true value of winning, i.e, the probability of selling an item via the advertisement over time. A problem of multiple queues competing using an auction to obtain service from a server (a cellular network) is studied in [34]. Here, the objective is to show that the MFE that results has small service delays.

Our problem can be seen as a system of many servers competing for jobs that pick the one that has the best combination of price and reputation; any inference on load or innate ability has to be inferred from agents having similar reputation. Thus, the jobs estimate the work-quality provided by the servers via their reputation, and the estimate becomes more accurate as more jobs are served by any particular server. Our objective is to characterize the policy employed by the servers to attract jobs to themselves. Our problem is distinct from those considered in the literature, and is methodologically similar only in that we utilize the general framework of showing existence of MFE by verifying the conditions of a fixed point theorem. We will point out the difference and our innovations in the text.

4.1.2 Main Contributions

Our analytical model takes the form of a system in which jobs (customers) enter at a constant rate, and a subset of agents (servers) declare prices in order to obtain these jobs. Each server has an inherent work-quality of job completion (a parameter that indicates the probability of providing “good” or “bad” service), which known to itself but is unknown to jobs. However, each job leaves a (potentially noisy) “good” or “bad” review of the server that it utilized, and the server’s reputation (the best estimate of quality thus far) is visible to all jobs. A key novelty of our model is that the customer selects a server based on a linear combination of both price and the server’s reputation. The challenge faced by the server is to determine the optimal price, based on the fact that reputation is only updated when a customer accepts its price. Thus, a low price would result in a high probability of being selected, and a possibility of enhancing one’s reputation (particularly if it is less than the true work-quality). However, such an action would also mean that the revenue obtained through completing the job is low. Furthermore, servers are charged a holding cost by the

marketplace so as to incentivize short delays in completion.

Our first contribution is in establishing the existence of an optimal bidding policy and characterizing policy under the mean field regime. Our main insight that enables us to obtain a closed form solution is the observation that our marketplace can be viewed as an auction mechanism in which the servers bid their “effective price,” which is combination of both their declared price and reputation, but actually receive a reward (if selected) of the declared price. Posed in this fashion, we have a repeated first price reverse auction (lowest bid wins) wherein the agents estimate the future while bidding. We also determine an equivalent of the agent’s “type” (the terminology used in mechanism design literature to compare the fitness to win of different agents) that encapsulates the residual work, the reputation, the work-quality and future rewards. Since we show that the system operates in a manner similar to a standard auction, the agent with the highest type wins the auction. We further show that the type is decreasing in residual work and increasing in reputation.

Our second contribution is methodological. We show the existence of the Mean Field Equilibrium under our marketplace game across servers. Our proof technique requires the establishment of a continuous map from the assumed bid space (the belief distribution) to the resultant bid distribution of the agents bidding under that consistent belief. The procedure involves following the map from belief \rightarrow bid \rightarrow state transition kernel \rightarrow invariant state distribution \rightarrow resultant bid distribution. Finally, we show that the map is between convex compact spaces, establishing the result. Unique to our current work is a state space that is countable and uncountable variables that requires the use of results of [24] to show existence of a best response and [6] in showing continuity of the map.

Our final contribution is on data collection and numerical studies. Although there has been prior work on calculating the price markup engendered by a good reputation, none that we are aware of takes into account the impact on future reputation while choosing the current price. In order to answer this question numerically, we use the data from the used smart phone market, Swappa, to determine the right parameters for the value of reputation

in our model. We show numerically system with a large number of agents that the MFE is attained within a small number of iterations. We then provide comparative estimates of how each parameter (residual work, good reviews, total reviews and true quality) impact the “type” of the agent and his bid. Apart from verifying the analytical results, we illustrate the tradeoff between trying to incentivize customers to explore by setting low prices when reputation is less than the true quality, and to exploit one’s reputation when the reputation exceeds the true quality. We also provide time series analysis of high and low quality agents and show how the former play a long range strategy of building up reputation and then exploiting it, while the latter play the short range strategy of making as much as possible before the system learns their true nature.

4.2 Mean Field Model

We consider a system with NM servers that act as strategic agents in our system. These agents are randomly permuted and divided into N clusters with M agents in each cluster. At each discrete time instant $k = 1, 2, 3, \dots$, M jobs arrive into the system, and each job is assigned to a different cluster. Each agent in a cluster specifies a price (that we refer to as a bid) for completion of the job, and if its bid is accepted by the job, the agent is awarded the payment stated in its bid. We study the system in the mean field regime in which $N \rightarrow \infty$, and the agents assume independence across their competitors’ actions. Our setup is typical and the asymptotic accuracy of the approximation follows immediately from [20]. A proof of the asymptotic convergence of the finite system to the mean field system using the results of [20] in a setup that that is similar at a high level to ours is available in [26, 34]. Since the proof is similar in our current context. Figure 4.2 gives an overview of our model and the details are given below.

4.2.1 State Space of Agent

In the mean field regime, we consider a generic agent (server) who competes against $M - 1$ other agents, whose bids are believed to be drawn independently from some distribution.

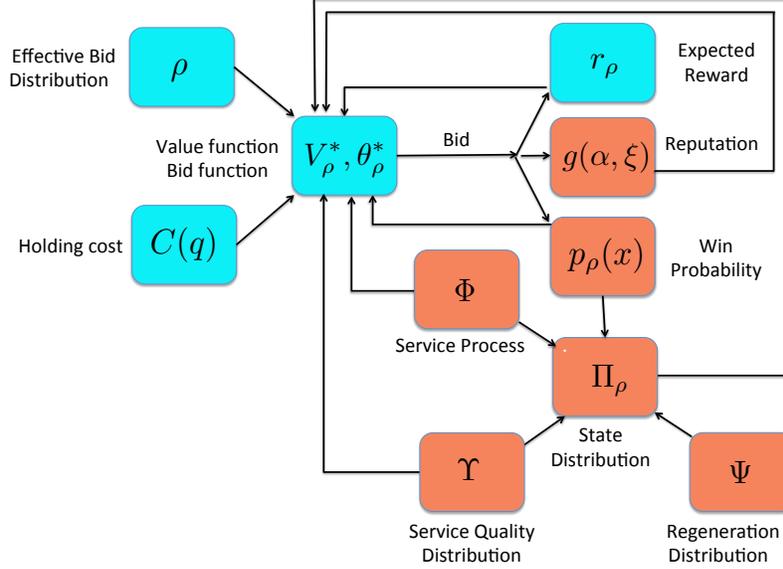


Figure 4.2: Mean Field Model: Parameters related to the mechanism are shown in blue, while the state transition parameters are shown in red.

The state of this agent consists of the queue length, the service quality and its reputation. We denote the residual work load in the queue of the agent just before time instant k by $q_k \in [0, \infty]$ (the random variable is represented by Q_k). An agent can complete each job with either a “high” quality w.p. η or “low” quality w.p. $1 - \eta$. The quality parameter $\eta \in (0, 1]$ is the realization of the random variable H that follows a distribution Υ , and is assumed to be fixed for the lifetime of the agent. An agent can perform a good job with quality he chooses η_c . If he chooses a quality $\eta_c > \eta$, he incurs a cost of $K(\eta, \eta_c)$ and if he chooses $\eta_c \leq \eta$, the cost is 0.

The marketplace tries to estimate the agents’s quality via a reputation scheme in which customers provide feedback on whether they received high or low quality service. The reputation of an agent is estimated via a β -distribution $\beta(\alpha_k, \beta_k)$, where α_k is the number of positive ratings (random variable A_k) and β_k is the number of negative ones (random variable B_k). Hence, the reputation function $g(\alpha_k, \xi_k) = \alpha_k / (\alpha_k + \beta_k)$. Finally, let $\xi_k = \alpha_k + \beta_k$ be the total number of ratings received up to time k , with the equivalent random variable Ξ_k . Then the realized state of the agent at time k is the tuple $(q_k, \alpha_k, \xi_k, \eta)$. For simplicity,

we assume that the reputation is updated immediately after the agent’s bid is accepted by the customer, and that the job reports the realized quality truthfully. In reality, there might be a random interval of time for the job to be completed and the customer to rate the service quality. Thus, the reputational updates might be bursty, which would complicate the model. However, our assumption implies that the agents would be conservative in their behavior and would do no worse in a delayed or bursty rating system. There might also be noise in reporting the realized quality with which the job was performed, i.e., the customer may provide inaccurate feedback. This effect is easy to incorporate into our model, since it simply changes the probability that good quality job completion would result in an increased number of positive ratings.

4.2.2 State Transitions

The state of the system could change due to the following events.

Departures:At each time slot, the server completes d_k units of work. Here, $d_k \in (0, \infty)$ is the realization of the random variable D_k drawn from the distribution Φ , which is assumed to have a finite mean. When the d_k units of work are completed, the queue state changes from q_k to $(q_k - d_k)^+$.

Arrivals:We assume a new job of unit size arrives in the cluster of the selected agent at each time slot. Let $x_k \in [0, \infty]$ be the bid submitted by the server in slot k (the equivalent random variable is denoted X_k). This bid is the output of the policy employed by the agent $\theta_\rho(q_k, \alpha_k, \xi_k, \eta_k)$, where ρ is the belief of the agent about its competitors actions, and will be defined precisely at the end of this section. If the server wins the auction, the queue state q_k changes to $q_k + 1$. Let W_k denote a unit random variable, whose realization w_k takes a value 1 if the server i wins the action in slot k , and else 0. As indicated above, with each arrival, the reputation changes to $(\alpha_k + 1, \xi_k + 1)$ with probability η_c and $(\alpha_k, \xi_k + 1)$ with probability $1 - \eta_c$.

Regeneration:Each server has a geometrically distributed lifetime and can quit in any time slot w.p. $1 - \delta$. This assumption is consistent with the fact that the human agents (often

students) at small scale service sites such as Mechanical Turk or even retail sites like Ebay are active for a while, and then move on to professional pursuits. We assume that a departing agent stops bidding for new jobs, but would complete the residual work in his queue. The agent is immediately replaced with a newly arriving agent. We refer to such a departure and replacement as a regeneration event. Thus, from the system perspective, the agent can regenerate at any time with probability $1 - \delta$. When regeneration occurs the agent's initial state is $(R_q, R_\alpha, R_\xi, R_\eta)$ drawn from the distribution Ψ .

In summary, from the perspective of the system, the state of the agent at time $k + 1$ is

$$q_{k+1} = \begin{cases} R_q & \text{if regenerates at } k \\ (q_k - d_k + w_k)^+ & \text{otherwise.} \end{cases} \quad (4.1)$$

$$(\alpha_{k+1}, \xi_{k+1}) = \begin{cases} (R_\alpha, R_\xi) & \text{if regenerates at } k \\ (\alpha_k, \xi_k) & \text{if } w_k = 0 \\ (\alpha_k + 1, \xi_k + 1) & \text{w.p. } \eta_c \text{ if } w_k = 1 \\ (\alpha_k, \xi_k + 1) & \text{w.p. } (1 - \eta_c) \text{ if } w_k = 1. \end{cases} \quad (4.2)$$

$$\eta_{k+1} = \begin{cases} R_\eta & \text{if regenerates at } k \\ \eta_k & \text{otherwise.} \end{cases} \quad (4.3)$$

Note that from the perspective of the agent, there is actually no regeneration, and it simply departs the system at some time. Thus, the evolution of state from an agent's view point (and hence its optimal action) would be different from the high-level view point of the system as a whole. We take heed of this point while developing the best response policy in Section 4.3.

4.2.3 Bid, Payment and Cost

The marketplace charges a holding cost $C(q_k)$ to the sever at each time slot k . We assume that $C(q_k)$ is a polynomial of degree $p \geq 1$ and convex increasing in queue size. Also, at each time slot if the server's bid is accepted, it obtains a payment equal to its bid y_k (random variable Y_k). The net payoff for the agent in time slot k is $y_k w_k - C(q_k)$. Note that we assume that the agent does not experience a cost for performing work, since we can subtract out this constant cost per unit time from the model.

4.2.4 Effective Bid and Customer Choice

Customers (jobs) choose the agents (servers) based on both the bid amount as well as reputation. We model the *effective bid* of the server as the linear combination $x_k \triangleq y_k - z g(\alpha_k, \xi_k)$, where $g(\alpha, \xi) = \alpha/\xi$ and z is a parameter that trades off bid and reputation. X_k is used to denote the random variable corresponding to effective bid. Then the assumption is that customers choose the agent that has the lowest effective bid. Customers are assumed to be short-lived, quasi-linear utility and myopic. Thus, they use the information they have when they appear to make a best guess of reward, hence the particular form of the effective bid.

Such an effective bid is consistent with the idea of reputation as enabling premium prices [19]. Clearly, the function $g(\alpha, \xi) = \alpha/\xi$ is increasing in α and decreasing in ξ , in tune with empirical observations [19]. Note that the smallest effective bid is $-z$, corresponding to an actual bid of 0, and a value of $g(\alpha, \xi) = 1$. More general models of reputation can also be represented by $g(\cdot)$. Finally, we denote the (cumulative) distribution of effective bids by ρ .

4.3 State Dynamics and MFE

As described earlier, under the mean field approximation we consider the bid selection problem faced by a generic agent of interest, who assumes that the effective bids of all other agents are drawn independently from a candidate (cumulative) distribution ρ and calculates a best response based on his current state. Note that for simplicity, we will analyze our

system using the effective bid as the parameter chosen by the agent; the translation back to the real bid simply involves an additive term. While performing the calculation of effective bid, the agent assumes that ρ is fixed for all time, which, as we will see in the definition presented at the end of this section, would indeed be the case at MFE.

We denote the space of candidate effective bid distributions by \mathcal{P} . The elements of \mathcal{P} are taken as cumulative probability distributions with support as a finite interval $[-z, \mathcal{T}]$ in \mathbb{R}^+ , and which are absolutely continuous w.r.t the Lebesgue measure. Furthermore, we assume that each member of \mathcal{P} is zero until a threshold, and then strictly increases over a finite interval to reach 1. The reason for this choice stems from the intuition that even an agent with a zero queue length would not bid 0, since it stands to get a higher payoff by bidding a non-zero value. Hence, there should exist a non-zero “reserve bid” below which no bids will occur. We will show the existence of a mean field equilibrium in this space in Section 4.6.

4.3.1 Agent’s Decision Problem

As the agent of interest maintains a belief ρ about the effective bid distribution of all its competitors, the agent’s objective function is

$$V_{\rho, \eta}(\theta) = \mathbb{E}_{\theta} \left[\sum_{t=k}^{\infty} \delta^t (-C(Q_t) + r_{\rho}(X_t, A_t, \Xi_t, H_t)) \right], \quad (4.4)$$

where expectation is over future state evolutions under policy θ , and $r_{\rho}(\cdot)$ is the payoff (or reward) function that we define next. Note that we have explicitly shown the quality parameter η , which remains fixed over the lifetime of the agent.

Expected Reward: Since the agent (server) receives the full value of the bid if selected by the customer, the expected reward if the agent’s effective bid is x is simply $r_{\rho}(x, \alpha, \xi) = p_{\rho}(x)(x + z g(\alpha, \xi))$. Here, $p_{\rho}(x)$ is the probability of the effective bid being accepted (a “win”). Thus, the mechanism is essentially a sealed bid reverse first price auction using effective bids, followed by payment with an additive markup. As described in Section 4.2, the customer decides which server (agent) to choose by comparing their effective bids and

choosing the one that is the smallest. Let $\mathcal{C}_1, \mathcal{C}_2, \dots$ represent a set of $M - 1$ I.I.D random variables, each with (cumulative) distribution ρ . The probability of winning for a server that has an effective bid of x is then given by

$$p_\rho(x) = \mathbb{P}\left(x \leq \min_j \mathcal{C}_j\right) \quad (4.5)$$

$$= \left(\mathbb{P}\left(x \leq \mathcal{C}_1\right)\right)^{(M-1)} = \left(1 - \rho(x)\right)^{(M-1)} \quad (4.6)$$

Hence, the reward

$$r_\rho(x, \alpha, \xi) = (x + z g(\alpha, \xi)) \left(1 - \rho(x)\right)^{(M-1)}. \quad (4.7)$$

Agent State Evolution Process: We can now write down the evolution of state from the perspective of the generic agent of interest. The state process (Q, A, Ξ, H) for an agent with quality parameter η is

$$\begin{aligned} & \mathbb{P}(Q_{k+1} \in B, A_{k+1} = \alpha, \Xi_{k+1} = \xi, H_{k+1} = \eta | Q_k = q_k, A_k = \alpha_k, \Xi_k = \xi_k, H_k = \eta_k, X_k = x, Y_k = \eta_c) \\ &= \delta p_\rho(x) \left(\mathbb{P}((Q_k + 1 - D_k)^+ \in B) (\mathbf{1}_{\alpha_{k+1}, \xi_{k+1}}(\alpha, \xi)) \eta_c \mathbf{1}_{\eta_k}(\eta) \right. \\ &+ \mathbb{P}((Q_k + 1 - D_k)^+ \in B) (\mathbf{1}_{\alpha_k, \xi_{k+1}}(\alpha, \xi)) (1 - \eta_c) \left. \mathbf{1}_{\eta_k}(\eta) \right) \\ &+ \delta (1 - p_\rho(x)) \left(\mathbb{P}((Q_k - D_k)^+ \in B) (\mathbf{1}_{\alpha_k, \xi_k}(\alpha, \xi)) \right) \mathbf{1}_{\eta_k}(\eta). \end{aligned}$$

We can now formulate the agent's bid selection problem as a Markov Decision Problem (MDP). The optimal value function of the agent is

$$V_{\rho, \eta} = \sup_{\theta \in \Theta} \mathbb{E}_\theta \left[\sum_{t=k}^{\infty} \delta^t (-C(Q_t) + r_\rho(X_t, A_t, \Xi_t)) \right], \quad (4.8)$$

where Θ is the space of Markov deterministic policies.

4.3.2 Invariant Distribution

The state process from the system-wide perspective is different from that of a single agent, since it also involves regenerations that are irrelevant to any particular agent. However, the state transitions occur under the Markov policy θ chosen by the agent in (4.8). In other words, the policy employed by the agent is optimal from its own perspective, and generates a state transition kernel for the system as a whole. Thus, the state transition kernel \mathcal{Q} that describes the evolution of (Q, A, Ξ, H) under policy θ is as given below.

$$\begin{aligned}
& \mathbb{P}(Q_{k+1} \in B, A_{k+1} = \alpha, \Xi_{k+1} = \xi, H_{k+1} \in B' | Q_k = q_k, A_k = \alpha_k, \Xi_k = \xi_k, H_k = \eta_k) \\
&= \delta p_\rho(\theta(q_k, \alpha_k, \xi_k, \eta_k)) \left(\mathbb{P}((Q_k + 1 - D_k)^+ \in B)(\mathbf{1}_{\alpha_{k+1}, \xi_{k+1}}(\alpha, \xi)) \eta_k \mathbf{1}_{B'}(\eta_k) \right) \\
&+ \mathbb{P}((Q_k + 1 - D_k)^+ \in B)(\mathbf{1}_{\alpha_k, \xi_{k+1}}(\alpha, \xi))(1 - \eta_k) \mathbf{1}_{B'}(\eta_k) + \delta(1 - p_\rho(\theta(q, \alpha_k, \xi_k, \eta_k))) \\
&\left(\mathbb{P}((Q_k - D_k)^+ \in B)(\mathbf{1}_{\alpha_k, \xi_k}(\alpha, \xi)) \mathbf{1}_{B'}(\eta_k) \right) + (1 - \delta) \Psi(B, \alpha_{k+1}, \xi_{k+1}, B') \tag{4.9}
\end{aligned}$$

We denote the invariant distribution of the above transition kernel by Π_ρ . Under the mean field approximation, the invariant distribution Π_ρ corresponds to the empirical state distribution the original (infinite agent) system. Thus, if some randomly chosen agent in the infinite agent system were sampled, its state would be drawn from Π_ρ .

4.3.3 Mean Field Equilibrium

Definition 2 (Mean field equilibrium). *Let ρ be a candidate effective bid distribution and θ_ρ be a stationary policy for an agent. Then, we say that (ρ, θ_ρ) constitutes a mean field equilibrium if*

1. θ_ρ is an optimal policy of the decision problem in (4.8), given effective bid distribution ρ ; and
2. $\rho(x) = \gamma(x) \triangleq \Pi_\rho\left(\{(q, \alpha, \xi, \eta) : (q, \alpha, \xi, \eta) \in \theta_\rho^{-1}([-z, x])\}\right), \forall x \in \mathbb{R}^+$.

4.4 Existence of Optimal Policy

In this section, we first show that optimal work-quality decision can be taken independent of bid decision. We then show that there indeed exists an optimal bid policy θ_ρ^* that would maximize the objective function shown in (4.8). Our argument follows a procedure outlined in Hernández-Lerma et al. [24]. We then obtain a closed form expression for the optimal bid function by viewing the server (agent) selection scheme as a version of a first price reverse auction.

The Bellman equation corresponding to our problem is

$$V_\rho(q, \alpha, \xi, \eta) = \sup_{x \in [-z, \infty]} \max_{\eta_c \in [0, 1]} \left\{ r_\rho(x, \alpha, \xi) - C(q) + \delta \mathbb{E}_D \left[p_\rho(x) V_\rho((q+1-D)^+, \alpha+1, \xi+1, \eta) \eta_c + p_\rho(x) (V_\rho((q+1-D)^+, \alpha, \xi+1, \eta) (1-\eta_c) + (1-p_\rho(x)) (V_\rho((q-D)^+, \alpha, \xi, \eta))) \right] \right\} \quad (4.10)$$

Lemma 10. *Optimal work-quality decision is independent of the optimal bid decision and the bellman equation for the optimal bid is given by*

$$V_\rho(q, \alpha, \xi, \eta) = \sup_{x \in [-z, \infty]} \left\{ p_\rho(x) \left(x + z g(\alpha, \xi) + \Delta_q V_\rho(q, \alpha, \xi, \eta) + K^*(\Delta_\alpha V_\rho(q, \alpha, \xi, \eta)) \right) \right\}, \quad (4.11)$$

where $\Delta_\alpha V_\rho(q, \alpha, \xi, \eta) = \delta \mathbb{E}_D [(V_\rho((q+1-D)^+, \alpha+1, \xi+1, \eta) - V_\rho((q+1-D)^+, \alpha, \xi+1, \eta))]$, and $\Delta_q V_\rho(q, \alpha, \xi, \eta) = \delta \mathbb{E}_D [V_\rho((q+1-D)^+, \alpha, \xi+1, \eta) - V_\rho((q-D)^+, \alpha, \xi, \eta)]$ and K^* is the convex conjugate of K .

Proof. The proof is given in the Appendix B.1.1. □

Define the space of possible value functions \mathcal{V} as

$$\mathcal{V} = \left\{ f : (\mathbb{R}^+, \mathbb{N}, \mathbb{N}, [0, 1]) \rightarrow \mathbb{R}^+ : \sup_{q \in \mathbb{R}, \xi \in \mathbb{N}, \alpha \leq \xi} |f(q, \alpha, \xi, \eta)| < \infty \right\}. \quad (4.12)$$

Next, define the Bellman operator T_ρ as follows:

$$(T_\rho f)(q, \alpha, \xi) = \delta \mathbb{E}_D \left[f((q - D)^+, \alpha, \xi, \eta) \right] - C(q) + \sup_{x \in [-z, \infty]} \left\{ p_\rho(x) \left(x + z g(\alpha, \xi) + \Delta_q f(q, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta)) \right) \right\}.$$

where, $\Delta_\alpha f(q, \alpha, \xi, \eta) = \delta \mathbb{E}_D [f((q + 1 - D)^+, \alpha + 1, \xi + 1, \eta) - f((q + 1 - D)^+, \alpha, \xi + 1, \eta)]$, and $\Delta_q f(q, \alpha, \xi, \eta) = \delta \mathbb{E}_D [f((q + 1 - D)^+, \alpha, \xi + 1, \eta) - f((q - D)^+, \alpha, \xi, \eta)]$.

We show the existence of the optimal value function (and hence effective bid function) by verifying certain properties of our MDP, and then applying results from [24].

Theorem 11. *For the MDP problem,*

1. *The Bellman equation (4.11) has a unique solution V_ρ^**
2. *There exists a deterministic stationary policy θ_ρ^* , which achieves this V_ρ^**
3. *The operator T_ρ is a contraction mapping that is $\|T_\rho u - T_\rho u'\|_w \leq \gamma' \|u - u'\|_w, \forall u, u' \in \mathcal{V}$ with $\gamma' < 1$.*

Proof. First, we need three auxiliary lemmas 36, 37, 38, which we prove in Appendix B.1.1. These lemmas are equivalent to the Assumptions 8.3.1, 8.3.2 and 8.3.3 in [24]. Parts 1 and 2 above follow from Theorem 8.3.6 of [24], while Part 3 follows from the proposition 8.3.9 of [24]. □

Now that we have shown the existence of the optimal value and effective bid functions, we characterize the effective bid (and hence the actual bid) in terms of the value function. We achieve this characterization by casting our agent selection mechanism as a first price reverse auction with a total of M agents and then using the Envelope Theorem [39] to derive the optimal bid. In our auction, each agent assumes that the states of the other $M - 1$ agents are drawn in an I.I.D fashion from a distribution Π . Our objective is to identify the bid of a particular player as a function of his realized state. We have the following result.

Theorem 12. *The optimal effective bid is given by the following conditional expectation which has a similar form to a general first price reverse auction.*

$$\theta_\rho^*(q, \alpha, \xi, \eta) = -\mathbb{E}[\tilde{\nu} \mid \tilde{\nu} \leq \nu(q, \alpha, \xi, \eta)] \quad (4.13)$$

$$\text{where, } \nu(q, \alpha, \xi, \eta) = zg(\alpha, \xi) + \Delta_q V_\rho(q, \alpha, \xi, \eta) + K^*(\Delta_\alpha V_\rho(q, \alpha, \xi, \eta)), \quad (4.14)$$

$$\text{and } \tilde{\nu} \sim \max_{j=1,2,\dots,M-1} \nu(Q_j, A_j, \Xi_j, H_j) \text{ with } (Q_j, A_j, \Xi_j, H_j) \sim \Pi \quad \forall j \quad (I.I.D). \quad (4.15)$$

Proof. According to our model, a job chooses the agent that submits the lowest effective bid. Thus, we can think of the mechanism as a standard reverse auction in which each agent submits an effective bid, the agent with lowest bid wins and is paid some premium in addition to the effective bid that depends on reputation. We use the envelope theorem to derive the closed form expression of the bid, and in doing so we come up with the notion of the “type” of the agent as function of its state, which characterizes its probability of winning the auction. The details are given in Appendix B.1.3 □

The idea of the agent type defined in (4.14) allows us to rewrite the dynamic auction in the same fashion as a static one-shot auction. Essentially, the value of winning for the agent both currently and in the future is captured by his type. We will study the properties of the type, both analytically and numerically, to obtain insights on how it impacts the bid of the agent under different state realizations.

4.5 Properties of Optimal Policy

In this section, we derive several useful properties of the optimal bid function, which, apart from providing insight into how the state of an agent impacts his bid, also will be useful in establishing the existence of the MFE in Section 4.6. Most of these results are intuitively clear. Our first result shows the decreasing nature of the value function in the queue length.

Lemma 11. *Given a cumulative bid distribution function ρ , V_ρ^* is a continuous strictly decreasing function of q .*

Proof. Our approach is as follows. We assume that $f \in \mathcal{V}$ is a strictly decreasing function of q . We prove that under this assumption, $T_\rho f$ is also a strictly decreasing function of q . Since $T_\rho^n(f) \rightarrow V_\rho^*$, V_ρ^* will also be a strictly decreasing function of q . The details are given in Appendix B.2 \square

Our next result follows a similar approach to show the increasing nature of effective bid with α .

Lemma 12. *For a given effective bid distribution ρ , V_ρ^* strictly increases with α .*

Proof. We assume that f strictly increases with α and we prove that $T_\rho f$ also strictly increases with α . Since $T_\rho^n f \rightarrow V_\rho^*$, V_ρ^* will increase with α . The details are given in Appendix B.2.2 \square

We next prove a useful property that the value of the transition upon being selected by the job, ΔV_ρ^* , is decreasing in queue length.

Lemma 13. *Let $\nu_f(q, \alpha, \xi, \eta) = zg(\alpha, \xi) + \Delta_q f(q, \alpha, \xi, \eta) + \sup_{\eta_c} \{\Delta_\alpha f(q, \alpha, \xi, \eta)\}$, then*

1. $\nu_{V_\rho^*}(q, \alpha, \xi, \eta)$ strictly decreases with q .
2. $\nu_{V_\rho^*}(q_H, \alpha, \xi, \eta) - \nu_{V_\rho^*}(q_L, \alpha, \xi, \eta) \leq \Delta C(q_L) - \Delta C(q_H)$, where $\Delta C(q) = \delta \mathbb{E}_D [C((q+1-D)^+) - C((q-D)^+)]$.

Proof. Again we assume that both properties hold for a function $f()$ and prove that they are true for $Tf()$. Full details are given in Appendix B.2.3 \square

The result indicates that the agent has a reducing incentive to accept more jobs as the queue increases, and would hence choose larger effective bids. We now prove this result.

Lemma 14. *Given a cumulative distribution function ρ , the optimal policy $X^*(q, \alpha, \xi)$ strictly increases with queue length.*

Proof. From Theorem 12, the optimal policy can be written as

$$X^*(q, \alpha, \xi, \eta) = -\mathbb{E}[\tilde{\nu} \mid \tilde{\nu} \leq \nu(q, \alpha, \xi, \eta)].$$

We see that $X^*(q, \alpha, \xi, \eta)$ decreases with ν as the expectation value term increases with ν . We know that $\nu(q, \alpha, \xi, \eta) = zg(\alpha, \xi) + \delta\Delta V(q, \alpha, \xi)$. From Lemma 13, $\Delta V(q, \alpha, \xi)$ is strictly decreasing with q , which means that $\nu(q, \alpha, \xi)$ strictly decreases with q . Hence, the optimal policy strictly increases with queue length. \square

Corollary 2. *If all agents have the same fixed reputation (i.e., we have a pure price competition), then by Lemma 14, the optimal policy is equivalent to each job joining the shortest queue in the cluster.*

The last result is interesting, since it suggests that if the market place enforces a holding cost, and all servers are equally good, the customers are automatically incentivized to join the queue promises the lowest delay. Essentially, the market mechanism effectively forces agents to reveal (a function of) their true queue lengths in this case.

Lemma 15. *Given a cumulative distribution function ρ , $\nu(q, \alpha, \xi, \eta)$ increases with α for a given q, ξ, η .*

Proof. We will again show the preservation of the property through the Bellman function. We assume $\nu_f(q, \alpha, \xi, \eta) = zg(\alpha, \xi) + \Delta_q f(q, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))$ is increasing with α and then show $\nu_{T_\rho f}(q, \alpha, \xi, \eta)$ also increases with α . Details are given in the Appendix B.2.4. \square

This result shows that the type is increasing with reputation, and hence if all else is held constant an agent with a higher reputation is more likely to be selected by a job. In Section 4.8, we will numerically study impact of the relationships that we have derived in this section, and [provide insights on how their variation impacts the choices made by the agent.

4.6 Existence of Mean Field

In this section, we prove the existence of the mean field equilibrium in our system. Formally, the main result of this section is as follows.

Theorem 13. *There exists an MFE (ρ, θ^*) such that $\rho(x) = \gamma(x) \triangleq \Pi_\rho(\theta_{\rho,z}^{*-1}[-z, x])$, where $\theta_{\rho,z}^*(q, \alpha, \xi, \eta) = \theta_\rho^*(q, \alpha, \xi, \eta) - zg(\alpha, \xi)$.*

We first need the following notation. We define the space of effective bid policies $\Theta = \left\{ \theta : (\mathbb{R}^+, \mathbb{N}, \mathbb{N}, [0, 1]) \rightarrow [-z, \infty) \mid \|\theta\|_w < \infty \right\}$. We also define $\Theta^* : \mathcal{P} \rightarrow \Theta$ such that $(\Theta^*(\rho))(q, \alpha, \xi, \eta) = \theta_{\rho,z}^*(q, \alpha, \xi, \eta) \triangleq \theta_\rho^*(q, \alpha, \xi, \eta) - zg(\alpha, \xi)$. Also, define the mapping $\Pi^* : \mathcal{P} \rightarrow \Omega$ that takes the bid distribution to the invariant state distribution Π_ρ . Finally, we define \mathcal{F} as $(\mathcal{F}(\rho))(x) = \gamma(x) = \Pi_\rho(\theta_{\rho,z}^*[-z, x])$. We will show that \mathcal{F} maps \mathcal{P} into \mathcal{P} itself.

We will establish the existence of the MFE by verifying the conditions of the following fixed point theorem.

Theorem 14. *(Schauder Fixed Point Theorem). Suppose $\mathcal{F}(\mathcal{P}) \subset \mathcal{P}$. If $\mathcal{F}(\mathcal{P})$ is contained in a convex and compact subset of \mathcal{P} , then $\mathcal{F}(\cdot)$ has a fixed point.*

In our model all the users employ first price auction and the function \mathcal{F} depends on this auction format. Instead of proving the compactness for this \mathcal{F} , we will prove a much general result which proves existence of MFE for any standard auction. But, first we will find the relations between different auction formats. Then we will prove continuity and compactness results. The key idea here is to look from the perspective of second price auction.

4.6.1 General Auction

Theorem 15. (Payoff Equivalence Theorem) *Let h be the value distribution, and let A denote a standard reverse auction and $\mathcal{N}_{h,A}$ denote Nash-equilibrium of auction A when the value distribution is h . Then*

1. *Equilibrium bid distribution in A is $\rho_A(x) = h(\mathcal{N}_{A,h}^{-1}(x))$ and in Second price auction S is h .*

2. Reward, $r_{\rho_A, A}(\mathcal{N}_{h, A}(\nu)) = r_{h, S}(\nu)$.

3. Probability of winning in equilibrium, $Pr(\mathcal{N}_{h, A}(\nu) | \rho_A) = Pr(\nu | h)$.

Let LM_S denote the market with second price reverse auction and LM_A denote a market with standard reverse auction.

Theorem 16. *If (ρ_S, θ_S) denote Mean Field Equilibrium of LM_S and let A be any standard auction with symmetric increasing Nash equilibrium $\mathcal{N}_{A, h}$ when the value is distributed according to h , then LM_A has an MFE (ρ_A, θ_A) , such that*

1. $\rho_A(x) = \rho_S(\mathcal{N}_{A, \rho_S}^{-1}(x))$

2. $\theta_A(q, \alpha, \xi, \eta) = \mathcal{N}_{A, \rho}(\theta_S(q, \alpha, \xi, \eta))$

Lemma 16. *The value function of second price auction under assumption ρ_S is equal to the value function of auction A with the assumption ρ_A . That is $V_{\rho_S, S}(q, \alpha, \xi, \eta) = V_{\rho_A, A}(q, \alpha, \xi, \eta)$.*

Proof. We show that for any policy X under S , there exists a policy X_A , such that if $X(q, \alpha, \xi, \eta) = b$, $X_A(q, \alpha, \xi, \eta) = \mathcal{N}_{A, \rho_A}(b)$ which will have same value function.

$$V_{\rho_S, S}^X(q_0, \alpha_0, \xi_0, \eta_0) = \mathbb{E}_X \left[\sum_k \delta^k (r_{\rho_S, S}(b_k) - C(q_k)) \right] \quad (4.16)$$

$$\begin{aligned} V_{\rho_A, A}^{X_A}(q_0, \alpha_0, \xi_0, \eta_0) &= \mathbb{E}_{X_A} \left[\sum_k \delta^k (r_{\rho_A, A}(b'_k) + zg(\alpha_k, \xi_k) - C(q_k)) \right] \\ &= \mathbb{E}_{X_A} \left[\sum_k \delta^k (r_{\rho_A, A}(\mathcal{N}_{A, \rho_S}(b_k)) + zg(\alpha_k, \xi_k) - C(q_k)) \right] \\ &= \mathbb{E}_{X_A} \left[\sum_k \delta^k (r_{\rho_S, S}(b_k) + zg(\alpha_k, \xi_k) - C(q_k)) \right] \\ &= \mathbb{E}_X \left[\sum_k \delta^k (r_{\rho_S, S}(b_k) + zg(\alpha_k, \xi_k) - C(q_k)) \right] \\ &= V_{\rho_S, S}^X(q_0, \alpha_0, \xi_0, \eta_0) \end{aligned}$$

Since \mathcal{N}_A is strictly increasing, for any policy X_A we can find a policy X for which the value functions agree. Therefore the optimal value functions should also agree and the optimal policy $\theta_A(q, \alpha, \xi, \eta) = \mathcal{N}_{A, \rho_S}(\theta_S(q, \alpha, \xi, \eta))$.

□

Lemma 17. *The state distribution remains same $\Pi_{\rho_A, A}(q, \alpha, \xi, \eta) = \Pi_{\rho_S, S}(q, \alpha, \xi, \eta)$*

Proof. The transition kernel only depends on $p_\rho(\theta^*(q, \alpha, \xi, \eta)$ and the following holds

$$p_{\rho_S}(\theta_S^*(q, \alpha, \xi, \eta)) = p_{\rho_A}(\theta_A^*(q, \alpha, \xi, \eta)), \quad (4.17)$$

the steady state distribution will be same for any standard auction.

□

4.6.2 Continuity of the map \mathcal{F}

Now we will show the continuity of the map of first price auction \mathcal{F} . To prove this, first we will prove the continuity of Θ^* and Π^* with ρ .

4.6.2.1 Continuity of Θ^*

Lemma 18. *$\Theta^* : \mathcal{P} \rightarrow \Theta$ is a continuous function of ρ .*

Proof. We have, $\Theta^*(\rho)(q, \alpha, \xi, \eta) = \theta_\rho^*(q, \alpha, \xi, \eta)$ and

$$\theta_\rho^*(q, \alpha, \xi) = \theta_{\rho, z}^*(q, \alpha, \xi) + zg(\alpha, \xi) \quad (4.18)$$

$$= \arg \max_{x \in [-z, \infty)} \left\{ p_\rho(x)(x + zg(\alpha, \xi, \eta) + \Delta_q V_\rho(q, \alpha, \xi, \eta) + K^*(\Delta_\alpha V_\rho(q, \alpha, \xi, \eta))) \right\} + zg(\alpha, \xi). \quad (4.19)$$

Let $h(\rho, x) = p_\rho(x, \alpha, \xi)(x - \delta \Delta V_\rho(q, \alpha, \xi))$, and $\mathcal{R}(\rho) = [0, \infty)$. We use Berge's maximum theorem [6] to establish the continuity of θ_ρ . The details are given in Appendix B.3.1 □

4.6.2.2 Continuity of Π^*

Lemma 19. *For any $\rho \in \mathcal{P}$ and any $\theta \in \Theta$, $\Pi_{\rho, \theta}(\cdot)$ is absolutely continuous with respect to the product measure of Lebesgue measure on \mathcal{R}^+ , the counting measure on \mathcal{N}^+ and counting*

measure on \mathcal{N}^+ .

Proof. The proof is given in Appendix B.3.2 □

We develop a useful alternate characterization of Π_ρ . Define $\Gamma_\rho^{(k)}(B, \alpha, \xi, B'|q_0, \alpha_0, \xi_0, \eta_0)$ as the probability that the state at time slot k is in the set (B, α, ξ, B') given that the initial state is $(q_0, \alpha_0, \xi_0, \eta_0)$ and there is no regeneration until time k .

$$\Gamma_\rho^{(k)}(B, \alpha, \xi, B'|q_0, \alpha_0, \xi_0, \eta_0) = \mathbb{P}((Q^k, \alpha^k, \xi^k, \eta^k) \in (B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0), \text{no regen})$$

Lemma 20. $\Pi_\rho(B, \alpha, \xi, B') = \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_\Psi \left[\Gamma_\rho^k((B, \alpha, \xi, B') | (State)) \right]$ where,
 $\mathbb{E}_\Psi \left[\Gamma_\rho^k((B, \alpha, \xi, B') | (State)) \right] = \int \Gamma_\rho^k((B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0)) d\Psi(q_0, \alpha_0, \xi_0, \eta_0)$.

Proof.

$$\begin{aligned} \Pi_{\rho, \theta}(B, \alpha, \xi, B') &= \lim_{m \rightarrow \infty} \Pi_{\rho, \theta}^m(B, \alpha, \xi, \eta) \\ &= \lim_{m \rightarrow \infty} \sum_{k=0}^m \delta^{m-k} (1 - \delta) P(Q^{m-k} \in B, \alpha^{m-k} = \alpha, \xi^{m-k} = \xi, \eta^{m-k} \in B' | \text{regen at } k) \\ &= \lim_{m \rightarrow \infty} \sum_{k=0}^m \left[\delta^{m-k} (1 - \delta) \int \Gamma_{\rho, \theta}^{m-k}((B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0)) d\Psi(q_0, \alpha_0, \xi_0, \eta_0) \right] \\ &= \lim_{m \rightarrow \infty} \sum_{k=0}^m \delta^{m-k} (1 - \delta) \mathbb{E}_\Psi \left[\Gamma_{\rho, \theta}^{m-k}(B, \alpha, \xi, B' | (State)) \right] \\ &= \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_\Psi \left[\Gamma_{\rho, \theta}^k((B, \alpha, \xi, B') | (State)) \right] \\ \text{where, } \mathbb{E}_\Psi \left[\Gamma_{\rho, \theta}^k((B, \alpha, \xi, B') | (State)) \right] &= \int \Gamma_{\rho, \theta}^k((B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0)) d\Psi(q_0, \alpha_0, \xi_0, \eta_0) \end{aligned}$$

□

Now we will use the previous characterization to derive the density of the distribution Π_ρ .

Lemma 21. *The density of the Π_ρ exists and is given by,*

$$\begin{aligned} & \Pi_\rho(q, \alpha, \xi, \eta) \\ &= \sum_{k \geq 1} \delta^k (1 - \delta) \sum_{a \in L_k} \psi(G_k((q, \alpha, \xi), a), \eta) [\Gamma^k(a_k = a | (G_k((q, \alpha, \xi), a), \eta))] J_k((q, \alpha, \xi), a) \end{aligned} \quad (4.20)$$

Proof. The proof is given in the Appendix B.3.3. □

Lemma 22. *The mapping $\Pi^* : \mathcal{P} \rightarrow \Omega$ is continuous in $\rho \in \mathcal{P}$.*

Proof. We need to prove that as ρ_n converges to ρ in ω -norm, $\Pi_{\rho_n, \theta_{\rho_n}}$ converges to Π_{ρ, θ_ρ} in distribution. We use the Portmanteau Theorem to prove this. The details are given in Appendix B.3.4 □

Lemma 23. *For any ρ , $\gamma(x) = \Pi_\rho(\theta_{\rho, z}^{*-1}[-z, x]) \in \mathcal{P}$.*

Proof. \mathcal{P} is the space of continuous and finite mean distribution functions. So, we need to prove that $\gamma(x)$ is continuous and has finite mean. The details are given in Appendix B.3.5 □

We are now ready to show the continuity of the mapping F , between ρ to γ .

Lemma 24. *$\mathcal{F} : \mathcal{P} \rightarrow \mathcal{P}$ is continuous with respect to ρ .*

Proof. We have $(\mathcal{F}(\rho))(x) = \Pi_\rho(\theta_{\rho, z}^{*-1}[-z, x])$, and Π_ρ and θ_ρ^* are continuous in ρ . Let $\rho_n \rightarrow \rho$ in uniform norm, then from 22 $\Pi_{\rho_n} \rightarrow \Pi_\rho$, and from 18, $\theta_{\rho_n}^* \rightarrow \theta_\rho^*$. From Theorem 5.5 of [8], the push forward is also continuous, so, $\Pi_{\rho_n}(\theta_{\rho_n}^{*-1}(\cdot)) \rightarrow \Pi_\rho(\theta_\rho^{*-1}(\cdot))$. Therefore $\mathcal{F}_{\rho_n}(x) \rightarrow \mathcal{F}_\rho(x)$. □

4.6.3 Continuity of \mathcal{F}_A

We have proved the continuity of the map in the first price auction format. Now we will show that this implies map corresponding to any standard auction format is continuous.

Lemma 25. *If \mathcal{F} is continuous then a second price reverse auction mapping \mathcal{F}_S is also continuous.*

Proof. If $\rho_{S,n} \rightarrow \rho_S$, we need to prove that,

$$\mathcal{F}_S(\rho_{S,n}) \rightarrow \mathcal{F}_S(\rho_S) \quad \text{where, } \mathcal{F}_S(\rho_S) = \Pi_{\rho_S, S} \left(\theta_{\rho_S, S}^{-1}([-z, x]) \right). \quad (4.21)$$

Let $\rho_n = \rho_{S,n}(\mathcal{N}^{-1}(x))$.

$$\begin{aligned} \mathcal{F}_S(\rho_{S,n}) &= \Pi_{\rho_{S,n}, S} \left(\theta_{\rho_{S,n}, S}^{-1}([-z, x]) \right) \\ &= \Pi_{\rho_n} \left(\theta_{\rho_{S,n}, S}^{-1}([-z, x]) \right) \\ &= \Pi_{\rho_n} \left(\theta_{\rho_n}^{-1}(\mathcal{N}([-z, x])) \right) \\ &= \mathcal{F}(\rho_n)(\mathcal{N}(x)) \\ &\rightarrow \mathcal{F}(\rho)(\mathcal{N}(x)) \\ &= \Pi_{\rho} \left(\theta_{\rho}^{-1}(\mathcal{N}([-z, x])) \right) \\ &= \Pi_{\rho_S, S} \left(\theta_{\rho_S, S}^{-1}([-z, x]) \right) = \mathcal{F}_S(\rho_S) \end{aligned}$$

There fore, \mathcal{F}_S is continuous. □

Now we will show the compactness of the range space. It is easy to show this in second price auction. So, we will prove the compactness for second price auction and established the MFE existence in second price auction which establishes the existence of MFE in any standard auction.

4.6.4 Compactness of Range Space \mathcal{F}_S

The final condition that we need to verify is the compactness of the range space of \mathcal{F} . To prove this first we will prove the equi-continuity of $\mathcal{F}(\mathcal{P})$.

Lemma 26. *$\mathcal{F}_S(\mathcal{P})$ is equi-continuous.*

Proof. Let $\gamma \in \mathcal{F}(\mathcal{P})$. To prove the set of $\gamma(x)$ are equi-continuous we will show that $\gamma'_+(x) = \limsup_{y \rightarrow x} \frac{\gamma(y) - \gamma(x)}{y - x}$ is bounded. The details are given in the Appendix B.3.6. \square

Since \mathcal{F}_S is equi-continuous by Arzela-Ascoli lemma \mathcal{F}_S is compact. By Schauder Fixed point theorem there exists a fixed point γ_S in second price auction. From Lemma 27 ρ_A is the MFE in the first price auction

Lemma 27. $\rho_A(x) = \rho_S(\mathcal{N}_{A,\rho_S}^{-1}(x))$ is the MFE in the first price auction, that is $\mathcal{F}_A(\rho_A) = \Pi_{\rho_A,A}(\theta_{\rho_A,A}^{-1}([-z, x]))$.

Proof.

$$\begin{aligned}
\gamma_{A,\rho_A}(x) &= \mathcal{F}_A(\rho_A) = \Pi_{\rho_A,A}(\theta_{\rho_A,A}^{-1}([-z, x])) \\
&= \Pi_{\rho_S,S}(\theta_{\rho_A,A}^{-1}([-z, x])) \\
&= \Pi_{\rho_S,S}(\theta_{\rho_S,S}^{-1}(\mathcal{N}_{A,\rho_S}^{-1}([-z, x]))) \\
&= \rho_S(\mathcal{N}_{A,\rho_S}^{-1}([-z, x])) \\
&= \rho_A(x)
\end{aligned}$$

\square

4.7 Non-Uniform Prior

We assumed that the Bernoulli parameter of the quality distribution of agents is drawn from the Uniform distribution, but this may not be true in all situations. If the prior distribution from which the Bernoulli parameter is drawn is known to the users then the optimal quality estimator will be a Bayesian estimator. In this section we will prove that any such Bayesian estimator will satisfy certain properties and establish the existence of MFE for any prior distribution.

We proved the existence of MFE for any $g(\alpha, \xi)$ function as long as $g(\cdot)$ is increasing in alpha and decreasing in ξ . Now we prove that any Bayesian estimator will satisfy these

properties. First we will find the expression for the Bayesian estimator given a prior distribution.

Lemma 28. *If the prior distribution of η is μ , then the Bayesian estimator is given by,*

$$g(\alpha, \xi, \mu) = \frac{\int_{\eta} \eta \text{Binom}(\alpha, \xi - \alpha, \eta) d\mu(\eta)}{\int_{\eta} \text{Binom}(\alpha, \xi - \alpha, \eta) d\mu(\eta)}.$$

Proof. We observe (α, ξ) and know the distribution μ of η . Then we have

$$\begin{aligned} \mathbb{P}((\alpha, \xi) | \eta = \eta^*) &= \text{Binom}(\alpha, \xi - \alpha, \eta^*) \\ \implies \frac{d\mathbb{P}(\eta = \eta^* | (\alpha, \xi))}{d\eta} &= \frac{\frac{d\mathbb{P}(\eta = \eta^*, (\alpha, \xi))}{d\eta}}{\mathbb{P}((\alpha, \xi))} \\ &= \frac{\frac{d\mu(\eta^*)}{d\eta} \mathbb{P}((\alpha, \xi) | \eta = \eta^*)}{\int_{\eta} \mathbb{P}((\alpha, \xi) | \eta' = \eta) d\mu(\eta)} \\ &= \frac{\frac{d\mu(\eta^*)}{d\eta} \text{Binom}(\alpha, \xi - \alpha, \eta^*)}{\int_{\eta} \text{Binom}(\alpha, \xi - \alpha, \eta) d\mu(\eta)}. \end{aligned}$$

Best estimator is the mean under the *aposteriori* distribution, therefore,

$$g(\alpha, \xi, \mu) = \frac{\int_{\eta} \eta \text{Binom}(\alpha, \xi - \alpha, \eta) d\mu(\eta)}{\int_{\eta} \text{Binom}(\alpha, \xi - \alpha, \eta) d\mu(\eta)}.$$

□

Lemma 29. *The following properties holds true,*

1. $g(\alpha + 1, \alpha + 1 + \beta, \mu) - g(\alpha, \alpha + \beta, \mu) \geq 0$
2. $g(\alpha, \alpha + \beta + 1, \mu) - g(\alpha, \beta + \alpha, \mu) \leq 0$

Proof. The details are given in the Appendix B.4.1. □

Corollary 3. *Bayesian estimator $g()$ has the following properties,*

1. $g(\alpha, \xi, \mu) \leq g(\alpha + 1, \xi, \mu)$

$$2. g(\alpha, \xi, \mu) \geq g(\alpha, \xi + 1, \mu)$$

Proof. Let $\beta = \xi - \alpha$ and $\hat{g}(\alpha, \beta, \mu) = g(\alpha, \alpha + \beta, \mu)$ then using lemma 29,

$$\begin{aligned} & g(\alpha + 1, \xi, \mu) - g(\alpha, \xi, \mu) \\ &= g(\alpha + 1, \alpha + \beta, \mu) - g(\alpha, \beta + \alpha, \mu) \\ &= g(\alpha + 1, \alpha + \beta, \mu) - g(\alpha + 1, \alpha + \beta + 1, \mu) + g(\alpha + 1, \alpha + \beta + 1, \mu) - g(\alpha, \beta + \alpha, \mu) \\ &= \left(g(\alpha + 1, \alpha + \beta, \mu) - g(\alpha + 1, \alpha + \beta + 1, \mu) \right) + \left(g(\alpha + 1, \beta + \alpha + 1, \mu) - g(\alpha, \alpha + \beta, \mu) \right) \\ &\geq 0 \end{aligned}$$

The second part can be proved in a similar way. □

4.8 Numerical Experiments

In this section, we present the insights gained from numerically evaluating the optimal policy and the MFE in a numerical fashion. Our first objective is to obtain a sense of what parameters would be accurate in real Internet marketplaces today.

4.8.1 Determining z from Data

As mentioned in the introduction, we gathered data from Swappa.com in order to determine the price premium that customers would potentially be willing to pay in order to receive service from an agent with a higher reputation. We first divide the phones into classes such that all devices in a class have the identical make, model and condition. Hence, the effective sale price should be same for any particular class, with the markup only depending on the reputation of the seller. To determine the effect of reputation, we minimize the square of the difference of the effective bid between all members of a class. Hence, the trade off parameter for a product κ is,

$$z_\kappa = \arg \min_z \sum_i \sum_j |(m_i - zr_i) - (m_j - zr_j)|^2,$$

where i and j denote the indices of the data points corresponding to class κ . Since this expression is quadratic we can find the minimizer by differentiating and finding the critical point. Thus,

$$z_\kappa = \frac{\sum_i \sum_j (m_i - m_j)(r_i - r_j)}{\sum_i \sum_j (r_i - r_j)^2}.$$

We use the lowest observed sale price of a class as the base price of that class, and denote it as \underline{m}_κ . Since the phones have different base prices, we desire to normalize the z_κ values to obtain an estimate of the parameter for a canonical product with a base price of \$100. In order to do this, we simply set $\bar{z}_\kappa = z_\kappa \times 100/\underline{m}_\kappa$. We evaluated the value of \bar{z}_κ for different classes of phones and the results are as shown in Table 1. The average of \bar{z}_κ appropriately weighted by the number of sample points in each class is 9.5%. In other words, the markup of a perfect reputation for a canonical product with a base price of 100 is of the order of \$9.50. We then assume that since every agent bears this base price of \$100 to provide

Table 4.1: Value of \bar{z}_κ for different products.

Phone	New	Mint	Good	Fair
iphone6 ATT-16G	5.45	7.75	6.5	-
iphone6 VRZ-16G	10.9	6.0	6.5	1.58
iphone6 VRZ-64G	11.32	11.46	6.4	-
iphone6 Plus ATT-16G	16	7.67	4.7	-
iphone6 Plus ATT-64G	11.07	9.09	7.33	9.72
Samsung NT4 ATT-32G	18.03	13.57	12.21	-
Samsung S6 VRZ-32G	12.62	11.297	11.7	2.98
Samsung S6 VRZ-64G	7.36	11.98	2.23	-
Samsung S6 ATT-32G	12.94	8.13	10.27	5.64
Samsung S6 ATT-64G	19.24	10.57	15.98	-
Samsung NT4 VRZ-32G	6.71	13.0	15.63	10.32

service, we can equivalently set the service cost equal to zero, and only consider the markup due to reputation.

4.8.2 Large Particle Simulations

We next determine the MFE distribution. Since our state space is $\mathbb{R} \times \mathbb{N} \times \mathbb{N}$, we discretize the state space to a resolution of 0.01 for queue lengths, and truncate all at 20 units. Similarly, we also discretize the bid space. We assume that the amount of service at any time step is uniformly distributed in $[0, 0.25]$, and that there are $M = 10$ agents per cluster. We set the parameter $z = 10$. Further, we choose the holding cost as $C(q) = q^2$.

In our simulations, we consider 1,000,000 agents (sometimes called “particles” in the literature) divided into 1000 clusters. At each time step, we calculate the empirical distribution of states (corresponds to Π ,) and using value iteration over the Bellman equation (4.11), we calculate the best response bid for each state. Each agent then places a bid based on its state, the lowest effective bid wins and we update the states of all agents. We then calculate the empirical state and bid distributions. We assume that the system has converged to the MFE, when the maximum point-wise difference between the effective bid distributions from one iteration to the next are less than 0.001. Essentially, we try to approximate uniform convergence. We found that convergence occurred quickly within 6 iterations of state.

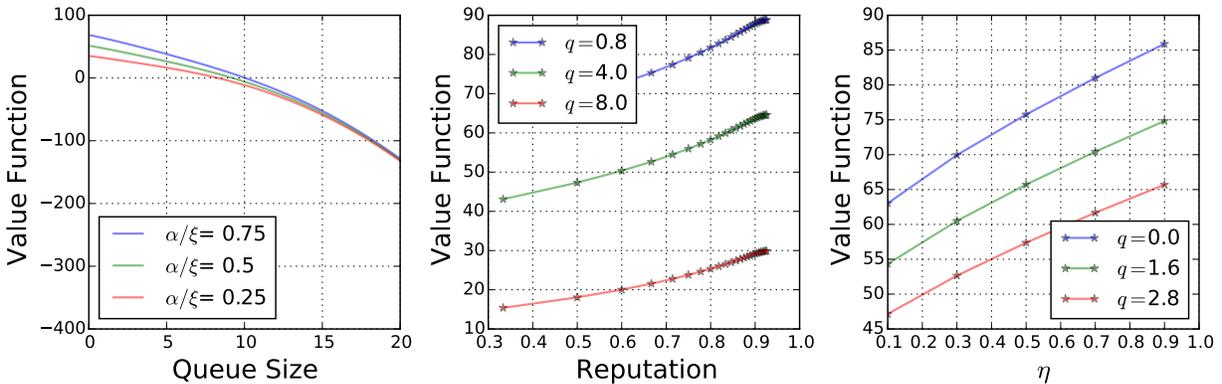


Figure 4.3: Value Function variation with queue size and reputation.

The dependence of the optimal value function with queue length, reputation and true

quality at MFE is plotted in Figure 4.3. We observe that the value function strictly decreases with queue size and increases with reputation and quality as expected.

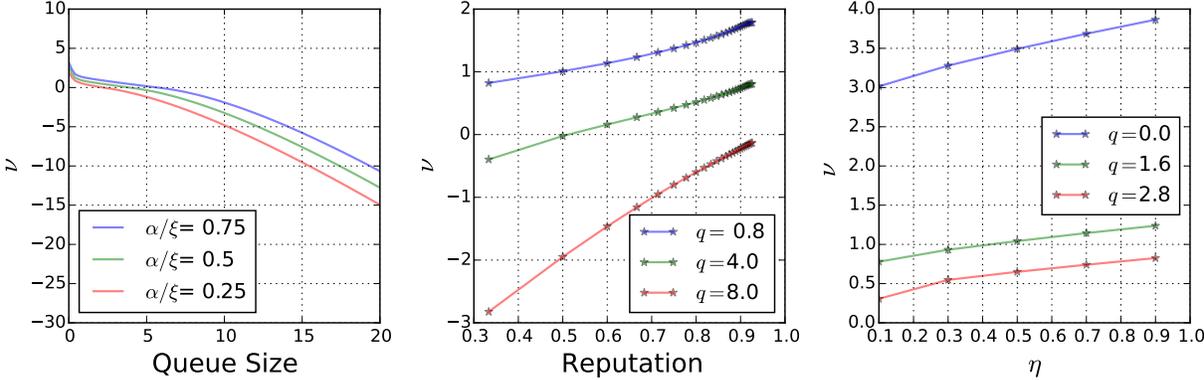


Figure 4.4: Variation of type with queue size, reputation and quality.

The variation of the “type” $\nu(q, \alpha, \xi, \eta)$ with its arguments is shown in Figure 4.4. Recall, that the type maps the state of the agent to its likelihood of obtaining the job, with a larger type yielding higher probability of winning. Figure 4.4 A shows that the type is decreasing in queue length, with the decrease becoming more pronounced as reputation falls. Hence, as desired, the agent is less likely to win as its residual work builds up. Figure 4.4 B shows the variation of type for increasing reputation. As expected, the an enhanced reputation results in a larger type. What is particularly interesting is the relationship between the true quality of the agent (0.6 in this experiment) and the shape of the type as the reputation grows. We see that curve has a point of inflexion and is concave on one side and convex on the other. This is essentially the tradeoff between incentivizing exploration and exploiting reputation. When reputation is below the true quality the marginal increase in type for increasing reputation is less that when the reputation is higher than the true quality. In other words, an agent with an abnormally large reputation for his given quality has a very high probability of winning. Figure 4.4 C shows that the type is increasing with quality as expected.

In Figure 4.5 A and B, we show the dependence of optimal effective bid and the corresponding real bid with queue size. The bids are increasing as the queue length increases, which is consistent with the observation made in Figure 4.4 A that the type decreases with queue length. Essentially, the agent bids high so as to avoid winning the job (or if he wins, the payoff is greater than the holding cost imposed). Figure 4.5 C and D illustrate the same effect of reputation versus true quality observed in Figure 4.4 B. We notice that the effective bid has a point of inflection, and the effective bid drops faster when the reputation is larger than the true quality, meaning that the agent is more likely to win.

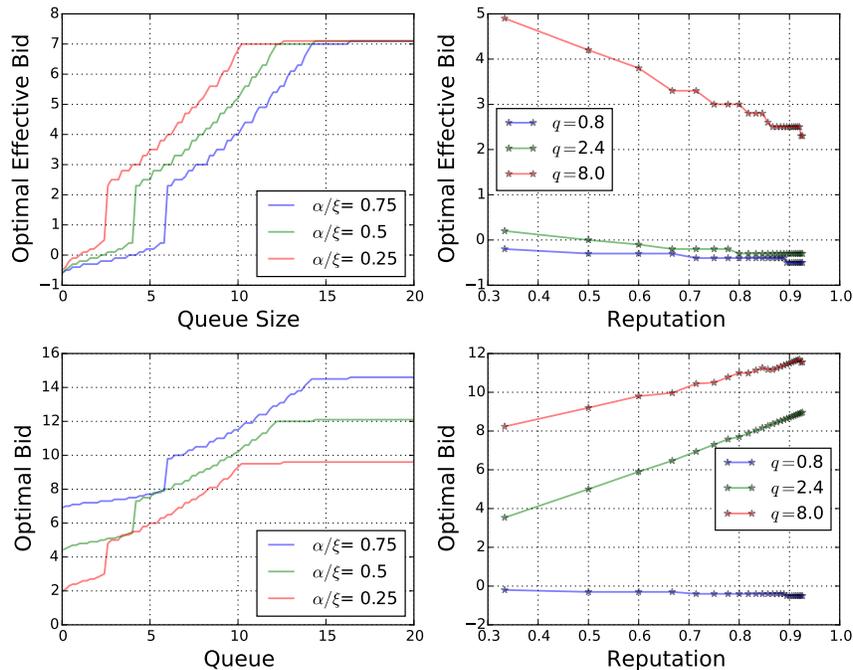


Figure 4.5: Optimal effective bid variation with queue size and reputation.

We next study the temporal evolution of average reputation and bids of agents that have a particular quality. In Figure 4.5 E and F we group all agents that have a quality of 0.8 and plot their average reputation with time and compare it with the average over all the agents that have a quality of 0.4. One would expect that the reputation should gradually

converge to the true quality as time proceeds and this is indeed observed. More interesting is the evolution of average bids of these agents shown in the Figure 4.6. We see that the high quality agents ($\eta = 0.8$) bid low initially to enhance their reputation, and then gradually raise their bids as the reputation increases. However, the low quality agents try to exploit their initial reputation as much as possible by bidding high, and then their bids fall and flatten out quickly as their quality becomes apparent in the reputation.

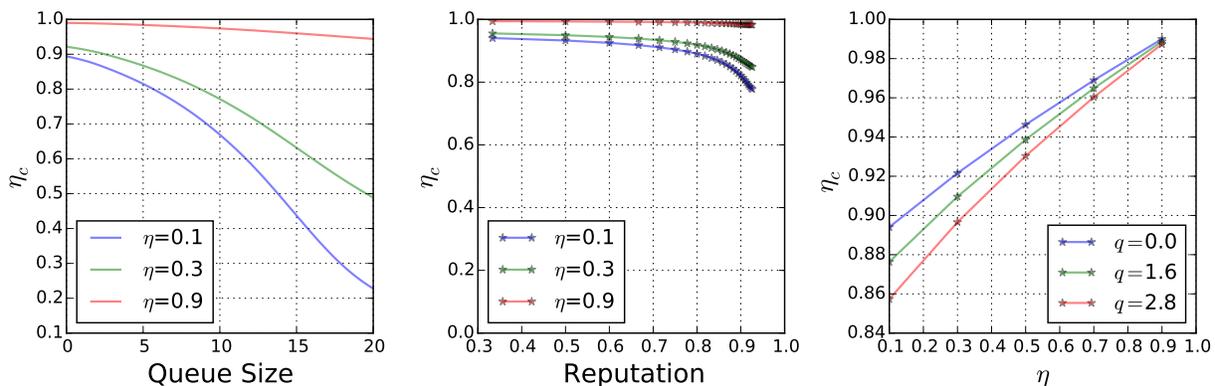


Figure 4.6: η_c chosen at different states.

4.8.3 Strategic η_c

To observe the effects of strategic aspects of quality we simulated the system with 0 queue cost. First we plotted the distribution of η_c chosen by the agents. Figure 4.7 shows the distribution of η_c both in strategic case and non strategic case. We can see that in non strategic case since the quality parameter is drawn from uniform distribution the η_c distribution is uniform. But in case of strategic case there is more weight at higher reputations. So the agents are performing the jobs at higher qualities to get better reputation scores.

4.9 Conclusion

In this chapter we considered the problem of determining the optimal price (bid) selection policy in a competitive marketplace that has many service providers that are relatively

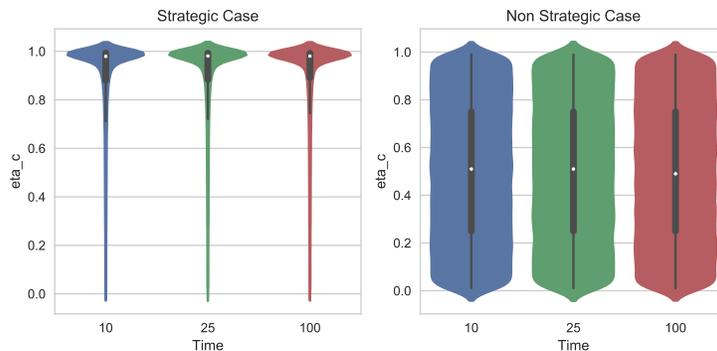


Figure 4.7: η_c distribution with time.

short lived. We setup a model in which jobs come at a steady rate into the marketplace and use both reputation and price jointly as a means of selecting a service provider. We showed that our marketplace can be thought of a version of a first price reverse auction, and characterized the best response bidding policy under the mean field assumption. We also showed that the assumption is valid by proving the existence of a MFE. We gathered data to find an accurate parameter to represent the tradeoff between reputation and offered price, and then numerically studied the system to understand the interplay between the different state variables of an agent.

Our findings suggested that a service provider that has a high inherent quality should try to establish its reputation by bidding low initially and securing many jobs, and then exploit that reputation by bidding high later on. However, a service provider with a low quality would seek to exploit what reputation he has by bidding high, since the value of each customer secured is in the payment it obtains, and not in any enhancement of reputation (indeed, the reputation would likely decrease until it hits the true quality).

In this chapter, our goal was to understand how competing sellers or service providers set prices for their goods while accounting for their reputation, and established the existence of Mean Field equilibrium. The market that we considered was conventional in the sense that providers and consumers were well differentiated. However, with the rise of the sharing economy, markets are moving in the direction of producers and consumers frequently switching

roles. A sharing system of the kind is a P2P network in which the resource being traded is link capacity. In the next chapter, we will study how to allocate resources efficiently in such P2P file sharing networks.

5. PROVABLE STABLE POLICIES IN P2P FILE SHARING

5.1 Introduction

Peer-to-Peer (P2P) file sharing networks such as BitTorrent [13] have been studied intensely in recent years, using analytical models, simulation studies, and large scale field experiments. This interest partly stems from the dominance of P2P as a source of Internet traffic in past years. Even today, although the traffic fraction has reduced to around 3-4% in North America, P2P sharing still occupies a significant fraction of about 30% of traffic in the Asia-Pacific region [50]. Interest also stems from a desire to understand the thought-provoking phenomenon of apparent scaling up of the throughput of a P2P network as the number of peers grows, which enables them to effectively distribute content with low file-download times during high demand situations called *flash-crowds*.

In a P2P network, a file is divided into fixed-size chunks, and a peer possessing a set of chunks can upload those chunks to other peers that need them. Once a peer has downloaded all chunks, it could continue to serve other peers or leave the system. A so-called *seed server* that possesses all chunks and never leaves is often used to ensure that no particular chunk ever goes missing. It is the feature of integrating the upload capacity of each peer into the system that is supposed to enable system-wide throughput scaling up with the number of peers. However, since peers can only share chunks that they possess, it is crucial to ensure the wide availability of all chunks to enable maximum usage of available upload capacity with each peer.

The problem of ensuring that all chunks are easily obtainable—ideally by engendering equal numbers of copies of each chunk over the network—was considered by the original designers of P2P networks. For example, BitTorrent, which is the most popular P2P network protocol, uses an algorithm called *rarest-first* (RF) to try to achieve this goal [13]. Here, the idea is to keep a running estimate of the frequency of all chunks in the system. When a peer

has a chance to download a chunk, it chooses the least frequent (i.e., the “rarest”) among all the chunks that it needs. In practice, peers keep track of the frequency of chunks in local subsets. Intuition suggests that such “boosting” of rare chunks might ensure a near-uniform empirical distribution of chunks.

Recent work has postulated that under some conditions, the rarest-first policy used by BitTorrent actually does not achieve its goal, and can actually be harmful to system performance. In particular, [23] studied a chunk-level model of P2P sharing under which new peers that do not possess any chunks arrive into the system at some rate, contacts between peers happen at random, and at each contact a chunk is transferred to a requesting peer under a given policy. Peers depart immediately after completing the file download. The objective was to determine if the system is *stable* under a given policy, i.e., at any time is the number of peers that have not yet received the whole file finite or is it exploding to infinity? The result was that under several policies including rarest-first and random chunk selection, a particular chunk can become very rare across the network—a phenomenon referred to as the *missing chunk syndrome*. This causes the creation of a large set of peers that are missing only that one chunk, referred to as the *one club*. In turn, the seed server must serve the missing chunk to almost all peers (which then depart), which means that the system is unstable unless the upload capacity of the seed server is of the order of the arrival rate of peers into the system. Thus, the phenomenon largely negates the value of the P2P system.

More recently, experimental studies have revealed that the missing piece syndrome is an observable phenomenon occurring in BitTorrent networks [36]. The results show that when the seed server has low or intermittent upload capacity, the throughput of the system saturates as the number of peers grows. In turn, this causes lengthened stay of peers in the system between arrival and completion, where an increasingly large number of peers are waiting to obtain the final chunk before leaving. In other words, designing policies that can ensure stability of a P2P network under a fixed seed server capacity for all peer arrival rates is practically relevant.

5.1.1 Related Work on Stable Algorithms

There has been extensive work on P2P networks, and we refer here only to those directly relating to the scaling properties of a single swarm. A large system assumption was made in [47, 57, 51], and the evolution of peers and seeds is described using a system of differential equations. While [47, 57] study the stationary regime and indicate the stability of BitTorrent-like systems for all arrival rates, [51] considers the transient regime and studies how much seed server capacity is needed to attain a target sojourn time (the time between the arrival of a peer and its completing the file download). Results on stability and scaling here require that at least a fixed fraction of the peers' upload capacity can always be utilized—an implicit assumption of chunk availability. As shown in [23], this assumption need not hold for all chunk selection policies, and a chunk-level model is needed for accurate analysis.

Chunk-level models have considered the missing chunk problem from two angles. The first method is to explicitly insist that peers that have completed the download should stay in the system as servers for some period of time. For example, [17] presents results on fairness vs. system performance based on how long peers stay after completion. In a more recent work [58], it was analytically shown that the system is stable as long as peers stay long enough to serve of the order of one additional chunk after completion. Indeed, in the original BitTorrent implementation this often happened naturally, since most users manually stopped participation at some point after download was completed. However, current implementations allow for the peer to depart immediately after completion, which can lead to the instability observed in [36].

The second method is to assume that peers would leave immediately after completion, and to design the chunk sharing policy such that the missing chunk syndrome is avoided. Some algorithms of this nature are “boosting” policies that can be thought of as modified versions of rarest-first. For example, the *rare chunk* (RC) algorithm studied in [48, 41, 45] picks three peers at random and chooses a chunk that is available with exactly one of the selected peers (called a “rare” chunk). Also studied in [45] is a variant of this algorithm

called the *common chunk* (CC) algorithm, which proceeds as in the RC algorithm when the peer has no chunks, then follows a policy of sampling a single peer with random selection among its required chunks until it only needs one more chunk, and then proceeds by sampling three peers and only downloading a chunk if every chunk with it appears at least twice with the sampled peers. However, although stable, these algorithms appear to have long sojourn times in some settings [7].

More recent work on chunk sharing policies [7] describes an algorithm called *group suppression* (GS), which is based on observations made in [23]. The policy is based on computing the empirical distribution of the states in the system, where a state of a peer is the set of chunks available with that peer. Peers that belong to the state with highest frequency are not allowed to upload chunks to peers that have fewer chunks than themselves, thus suppressing entry into the highest frequency group. Although this policy appears to have low sojourn times in simulations, it is somewhat complex since it requires the knowledge of the entire empirical state distribution. Furthermore, the authors are only able to prove stability in a P2P network with exactly 2 chunks, while the stability of the general case is left as a conjecture.

A different model is presented in [35], wherein peers arrive into the system already possessing one randomly selected chunk. This system is stable for many policies (including random chunk selection), but is constrained by the fact that the initial chunk has to be provided by the seed server. Thus, in this case too the seed’s capacity must scale with the arrival rate of peers, and the system might be unstable otherwise.

5.1.2 Main Results

The nominal objective of the rarest-first policy is to ensure a uniform chunk distribution across the network, which it actually does not achieve in all cases, causing instability as shown in [23]. Our intuition is that rather than following a policy of boosting low-frequency chunks as rarest-first does, simply preventing the most frequent chunk(s) from being shared would allow less frequent chunks to catch up, and drive the empirical distribution of chunks

towards the desired uniform distribution. Implicitly, this would also remove a small fraction of the upload capacity, keeping peers in the system a little longer, and hence enabling them to share more copies of rare chunks.

Following this intuition, we propose a policy that we refer to as *mode suppression* (MS), which is based on terminology used in statistics in which the *mode* is the most frequent value(s) in a data set. Thus, we keep track of the frequency of chunks in the system, and when a peer contacts another peer, it is allowed to download any chunk except the one(s) belonging to the mode. Any chunk may be downloaded if all chunks are equally frequent (i.e., if all chunks belong to the mode). The policy is simple to implement, and since all that is needed is the frequency of chunks; which is already a part of BitTorrent, the information needed for decision making is low.

We consider a continuous time model in the manner of [23, 45, 7] in which peers that have no chunks, enter the system according to a Poisson process with a certain arrival rate. There is a seed server that has an independent Poisson clock of a fixed rate, and at each clock tick, it contacts a single peer and uploads a chunk to it following a given policy. Each peer also has an independent Poisson clock of a fixed rate, and at each clock tick, the peer contacts a randomly selected peer and uploads a packet to it following the same policy. Our main analytical result is that under this model, mode suppression is *stable under all peer arrival rates* in a system in which the *file is divided into any number of chunks*.

The result follows from stability analysis using an intuitive Lyapunov function. Although the proof is somewhat involved, we show that mode suppression indeed is able to ensure that the drift is negative in all relevant cases, yielding stability.

We observe through simulations that mode suppression actually does come very close to attaining a uniform distribution of chunks in the system. In particular, we start the system in a corner case where one of the chunks is available only at the seed server, and observe the evolution of the system afterwards. We also compare with random chunk selection, rarest-first, and group suppression. A comparison of results on these algorithms is presented in

Table 5.1, where m is the number of chunks that the file is divided into. Furthermore, in simulations to compare sojourn time, we found that GS and MS have similar performance, but neither is better than the other in all cases.

Table 5.1: Comparison of global chunk distribution based policies.

Algorithm	$m = 2$	$m > 2$	Information
Random	Unstable	Unstable	None
Rarest-First	Unstable	Unstable	Chunk Frequency
Group Suppression	Stable	Unknown	Complete Distribution
Mode Suppression	Stable	Stable	Chunk Frequency

Since all the above algorithms require some kind of global chunk frequency or chunk set frequency for decision making, we then design a version of mode suppression that only depends on localized information. Under this algorithm, which we refer to as *distributed mode suppression* (DMS), a peer contacts three other peers at random, and among the chunks available with more than one peer, we define the *local mode* to be the chunk(s) with greatest frequency. The peer is allowed to download any chunk that is not part of the local mode. Any chunk may be downloaded if all chunks are equally frequent. We compare the sojourn time of this algorithm numerically with the other localized decision making algorithms discussed in Section 5.1.1. All the distributed algorithms do about as well or significantly better than the global-chunk-frequency-based algorithms, with DMS having the *lowest sojourn time of all algorithms in all cases*.

5.2 System Model

We consider a P2P file sharing system for a single file divided into m chunks. This file sharing system has a unique seed that has all m chunks, and the seed stays in the system indefinitely. Peers arrive according to a Poisson process with rate λ . Each incoming peer arrives without any chunks and stays in the system till it obtains all m chunks of the file.

In this model, a peer leaves as soon as it has all m chunks of the file. The peers can receive the chunks in two ways, either directly from the seed or from other peers.

Whenever the seed or a peer contact another peer, it is deemed as a contact. Therefore, each peer and the seed have individual contact processes corresponding to the sequence of contact instants. Upon contact the seed or the peer transfer a missing chunk to the contacted peer, according to a *chunk selection policy*. When chunk selection policy depends solely on the current state of the system, it is called a Markov chunk selection policy.

5.2.1 Contact Processes

The time interval between two contacts are assumed to be random, independent, and identically exponentially distributed, i.e. all contact processes are assumed to be independent and Poisson. The Poisson contact rate for the seed is assumed to be U , and each peer is assumed to have a common contact rate of μ .

5.2.2 State space

At any time t , the number of peers in the system with a proper subset of chunks $S \subset [m]$ is denoted by $X_S(t) \in \mathbb{N}_0 \triangleq \{0, 1, \dots\}$. The system at time t can be represented by the state

$$X(t) = (X_S(t) : S \subset [m]).$$

The total number of peers at any time t is denoted by

$$|X(t)| = \sum_{S \subset [m]} X_S(t).$$

For any Markov chunk selection policy, the continuous time process $\{X(t), t \geq 0\}$ is Markov with countable state space $\mathcal{X} \triangleq \mathbb{N}_0^{\mathcal{P}([m]) \setminus [m]}$. The *stability region* is defined as the set of arrival rates λ , for which the continuous time Markov chain $X(t)$ is positive recurrent.

5.2.3 State transitions

The generator matrix for the process $X(t)$ is denoted by Q . For this continuous time Markov chain, there can only be a single transition in an infinitesimal time. We denote the system state as $x \in \mathcal{X}$ just before any transition, and let e_S be the unit vector in the dimension corresponding to a proper subset $S \subset [m]$.

There are three types of possible transitions. First type of state transition is the arrival of a new peer, that leads to an increase in the number of peers with no chunks. The corresponding transition rate is denoted by

$$Q(x, x + e_\emptyset) = \lambda.$$

Second and third type of transitions occur, when a peer with $S \subset [m]$ chunks receives a chunk $j \notin S$ from the contacting seed/peer. In both these cases, the next state is denoted by $\mathcal{T}_{S,j}(x)$. Second type of state transition occurs when the reception of new chunks doesn't lead to a departure. This transition is denoted by

$$\mathcal{T}_{S,j}(x) \triangleq x - e_S + e_{S \cup \{j\}}, \quad x_S > 0, |S| < m - 1.$$

Third type of state transition occurs for a peer with $m - 1$ chunks, which departs the system after getting the last chunk upon contact. This transition is denoted by

$$\mathcal{T}_{S,j}(x) \triangleq x - e_S, \quad x_S > 0, |S| = m - 1.$$

At a system state x , if the contacting source has T chunks and the contacted receiving peer has S chunks, then the set of available chunks that can be transferred is $T \setminus S$. Selection of which chunk to transfer is called the *chunk selection policy*, that governs the evolution of the process $X(t)$. In particular, the last two transition rates $Q(x, \mathcal{T}_{S,j}(x))$ can only be computed for a specific Markov chunk selection policy. We describe the proposed chunk

selection policy and the corresponding transition rates in the following section.

5.3 Mode Suppression policy

In this section, we describe the mode suppression policy and provide its rate transition matrix. First, let us establish some notation. The set of allowable transfers from a peer with set of chunks T to a peer with set of chunks S , is denoted by $A(x, T, S) \subseteq T \setminus S$, and the cardinality of this set is denoted by $h(x, T, S)$, that takes integral values between 0 and m . Recall that the seed has all the chunks, and hence the set of allowable chunk transfers by the seed is $A(x, [m], S)$. Below, we describe the specifics of selecting the set of allowable transfers.

If there are no peers in the system, there is no need for chunk transfer. Hence without any loss of generality, we consider the mode suppression policy when there exist peers in the system, or $|x| > 0$. Here, we assume that each peer has the knowledge of all chunk frequencies in the system. Frequency of the j th chunk is

$$\pi_j(x) \triangleq \frac{\sum_{S \in \mathcal{S}} x_S}{|x|}.$$

The chunks that attain the highest frequency $\arg \max\{\pi_j(x) : j \in [m]\}$ are called the modes of the chunk frequencies. The set of modes is defined as

$$I(x) \triangleq \{i \in [m] : \pi_i(x) \geq \pi_j(x), \forall j \neq i\}.$$

The mode suppression policy restricts transmission of chunks that belong to the set of modes. Specifically, when the index set $I(x)$ is a strict subset of all chunks, the contacting source excludes the most popular chunk(s) (i.e., the modes) from the set of allowable transfers. Otherwise, when all chunks are equally popular, the source allows all possible transfers.

Mathematically, one can write the allowable transfer set for mode suppression policy as

$$A(x, T, S) = \begin{cases} T \setminus (S \cup I(x)), & I(x) \subset [m], \\ T \setminus S, & I(x) = [m]. \end{cases}$$

From the superposition of independent Poisson contact processes, the rate at which either the seed or one of the peers with chunk j contact any peer is also Poisson with the aggregate rate

$$R_j(x) \triangleq U + \mu \sum_{T:j \in T} x_T = U + \mu|x|\pi_j(x).$$

The probability of the source contact process contacting a peer with chunk subset S is $\frac{x_S}{|x|}$. If the contacting source has T chunks, then it can transfer one out of $h(x, T, S)$ available chunks to the contacted peer with S chunks. The transition of type $\mathcal{T}_{S,j}$ occurs when either the seed or one of the peers with chunk $j \notin S$ contact a peer with chunks S , and transfer chunk j among all the possible choices. From the thinning and superposition of independent Poisson processes, we can write for $j \notin S$ and $x_S > 0$

$$Q(x, \mathcal{T}_{S,j}(x)) = \frac{x_S}{|x|} \left(\frac{U}{h(x, [m], S)} + \mu \sum_{T:j \in T} \frac{x_T}{h(x, T, S)} \right).$$

5.3.1 Stability Region of Mode Suppression Policy

In this section we characterize the stability region of mode suppression policy.

Theorem 17. *The stability region of the mode suppression policy is $\lambda \geq 0$ for file-sharing systems with at least two chunks, and positive contact rates U, μ .*

Proof. To prove the positive recurrence of the continuous time Markov chain $X(t)$, we employ

Foster-Lyapunov criteria [37]. We consider the following Lyapunov function,

$$V(x) = \sum_{i=1}^m \left((\bar{\pi} - \pi_i) |x| \right)^2 + C_1 ((1 - \bar{\pi}) |x|) + C_2 \left(M - \sum_{i=1}^m \pi_i |x| \right)^+, \quad (5.1)$$

where, C_1, C_2 and M are positive constants that depend on m, λ, U, μ , and $\bar{\pi} = \max_i \pi_i$. Note that the explicit dependency of $\pi(x)$ on x is not shown for simplicity.

The intuition behind this Lyapunov function is as follows. Since the nominal objective is to attain a uniform distribution, we should expect that the policy should promote negative Lyapunov drift whenever the current state differs from uniformity. Hence, our Lyapunov function is designed to penalize for the cases where chunks have differing frequency, where some might have zero frequency, and where all have zero frequency.

The expected rate of change of potential function for a Markov process $X(t)$ from state x is called the mean drift from this state, and is given by

$$\sum_y Q(x, y)(V(y) - V(x)) = QV(x).$$

Mean drift from a state x for the Markov process $X(t)$ in terms of its generator matrix Q can be written as

$$QV(x) = Q(x, x + \emptyset)(V(x + \emptyset) - V(x)) + \sum_{j \in [m]} Q(x, \mathcal{T}_{S,j}(x))(V(\mathcal{T}_{S,j}(x)) - V(x)).$$

First, we compute the mean drift corresponding to a new peer arrival. The arrival of a new peer does not change the number of peers with chunk $j \in [m]$. However, it does lead to a unit increase in the number of peers in the system. That is,

$$Q(x, x + e_\emptyset)(V(x + e_\emptyset) - V(x)) = \lambda C_1.$$

The rest of the proof proceeds as follows. We divide the states into two cases when the

chunk frequency is (i) non-uniform and (ii) uniform, and in each case we show that the drift is negative.

Case 1: $I(x) \subsetneq [m]$: In this case, no popular content is allowed to be transferred. Hence, any transition of type $\mathcal{T}_{S,j}(x)$ occurs only for $j \notin I(x)$. When $S \cup \{j\} \subsetneq [m]$, this transition leads to unit increase in the number of peers with chunk j , and no change in the number of peers with other chunks. The corresponding change in potential function for $S \cup \{j\} \subsetneq [m]$ and $M \in \mathbb{Z}_+$ equals

$$V(\mathcal{T}_{S,j}(x)) - V(x) = 1 - 2(\bar{\pi} - \pi_j)|x| - C_2 1_{\{M > \sum_i \pi_i |x|\}}.$$

Since the number of popular chunks has to be at least unity, this difference is strictly negative for all non-zero states x . For this transition, we can trivially bound the cardinality of the allowable transfers by $\sup_T h(x, T, S) \leq |S^c|$. This provides a lower bound on the transition rate

$$Q(x, \mathcal{T}_{S,j}(x)) \geq \frac{x_S}{|S^c||x|} R_j.$$

When $S = \{j\}^c$, it is clear that the set of allowable transfer is $\{j\}$ for the contacting sources. Hence, $h(x, T, S) = |S^c| = 1$ and the transition rate is

$$Q(x, \mathcal{T}_{S,j}(x)) = \frac{x_S}{|x|} R_j.$$

Further, the transition $\mathcal{T}_{S,j}(x)$ leads to a departure from the system of peer with $S = \{j\}^c$ chunks. That is, this transition leads to a unit decrease in number of peers with chunks other than j . The change in potential function $V(\mathcal{T}_{S,j}(x)) - V(x)$ for the transition from state x to state $\mathcal{T}_{S,j}(x)$, for $S \cup \{j\} = [m]$ and $M \in \mathbb{Z}_+$, is upper bounded by

$$1 - 2(\bar{\pi} - \pi_j)|x| + C_2(m - 1)1_{\{M + m - 1 > \sum_i \pi_i |x|\}}.$$

The fraction of users that have all the pieces except j th piece is denoted by $\gamma_j(x) \triangleq \frac{x_{\{j\}^c}}{|x|}$, and the aggregate number of chunks in the system at all peers is denoted by $r \triangleq \sum_{S \subseteq [m]} |S| x_S = \sum_{i \in [m]} \pi_i |x|$. Aggregating all the above results and notations, and observing that $|S^c| \leq m$, we can find an upper bound on the mean drift from state x as

$$C_1 \lambda - \sum_{j \notin I(x)} \frac{R_j}{m} \left[(2(\bar{\pi} - \pi_j)|x| - 1)(1 - \pi_j) + C_2(1 - \pi_j - \gamma_j)1_{\{M > r\}} - \gamma_j C_2 m(m-1)1_{\{M+m-1 > r\}} \right]. \quad (5.2)$$

We will divide the state space in to three regions and show that in each region the drift is negative. The details are given in the Appendix C. Now, we consider the uniform chunk frequency case.

Case 2: $I(x) = [m]$: In this case, the chunk frequencies are identical, that is $\bar{\pi} = \pi_i$ for each chunk $i \in [m]$, and any chunk j can be transferred. This also implies that the contact rate $R_j = U + \mu\pi_j|x|$ is uniform for all chunks j , and can be denoted by $R = U + \mu\bar{\pi}|x|$. For $S \subsetneq \{j\}^c$, a transition of type $\mathcal{T}_{S,j}(x)$ doesn't lead to any departure from the system. The number of peers with chunk j has a unit increase by one, and chunk j becomes the popular chunk. There is no change in the number of peers for other chunks. Hence the potential change, due to this transition, is

$$V(\mathcal{T}_{S,j}(x)) - V(x) = m - 1 - C_1 - C_2 1_{\{M > r\}}.$$

For $S = \{j\}^c$, a transition of type $\mathcal{T}_{S,j}(x)$ leads to the departure of the receiving peer from the system. In this case, the number of peers with chunk j remains same, the number of peers having other chunks has a unit decrease. The potential change due to this transition, is

$$V(\mathcal{T}_{S,j}(x)) - V(x) \leq m - 1 - C_1 - C_2(m-1)1_{\{M+m-1 > r\}}.$$

Using the same techniques as in Case 1, we can upper bound the drift of state x by,

$$C_1\lambda - R\left[(C_1 - m + 1)(1 - \bar{\pi}) + C_2(1 - \bar{\pi} - \gamma_j)1_{\{M>r\}} - \gamma_j C_2 m(m - 1)1_{\{M+m-1>r\}}\right]. \quad (5.3)$$

Similar to Case 1, we will divide the state space into three regions and show that in each region the drift is negative. The details are given in the Appendix C. \square

5.4 Threshold Mode Suppression

In the previous section we proposed a new chunk selection policy called Mode Suppression and proved that the stability region is $\lambda > 0$. In this section we generalize the Mode Suppression and propose a new policy called Thresholded Mode Suppression and discuss its stability.

In Mode suppression the chunks in the mode are suppressed even if their frequency count (number of peers that have this chunk) is marginally higher than the others. Consider the situation in which the frequency count of all the chunks is same except for one chunk which has a count one less than the others. In this case no other chunk will be allowed to be transferred except for the one chunk. So, this results in frequent suppression and hence higher sojourn times.

To address this problem we can relax the constraint for suppression. We will define a threshold parameter denoted by $T \in \mathbb{N}$ and suppress the chunks in the mode only if the frequency count of the mode is more than the least frequency item by at least T units.

Recall that $\pi_j(x) = \sum_{S:j \in S} \frac{x_S}{|x|}$, $\bar{\pi}(x) = \max_j \pi_j(x)$, $\underline{\pi}(x) = \min_j \pi_j(x)$ and $\mathcal{M}(x)$ is the set of mode indices, $\mathcal{M}(x) = \{k | \pi_k(x) = \bar{\pi}(x)\}$. In **Threshold Mode Suppression** (TMS) the set of suppressed chunks D_T in state x for some threshold $T \in \mathbb{N}$ is defined as,

$$D_T(x) = \{k | \pi_k(x) \in \mathcal{M}(x) \text{ and } \bar{\pi}(x)|x| \geq \underline{\pi}(x)|x| + T\}. \quad (5.4)$$

When $T = 1$, the Threshold Mode Suppression will become vanilla Mode Suppression and

when $T = \infty$, this will be equivalent to Random Chunk selection policy as there will not be any suppression.

5.4.1 Rate Transition Matrix (Q) of TMS

The set of chunks a peer B can transfer to S is denoted by $A(x, B, S)$ and the cardinality is denoted by $h(x, B, S) = |A(x, B, S)|$. In TMS as we suppress the chunks in the set D_T , the set of allowed packets $A(x, B, S)$ is given by,

$$A(x, B, S) = B \setminus (S \cup D_T). \quad (5.5)$$

Now we will compute the components of the rate matrix. When an arrival happens the state of the system changes from x to $x + e_\phi$. As the arrival process is a poisson process with rate λ ,

$$Q(x, x + e_\phi) = \lambda \quad (5.6)$$

Let us denote the state a peers goes to from S if it receives a chunk $j \notin S$ by $\mathcal{T}_{S,j}$. In our model as the peer leaves the system as soon as it receives all the chunks, $\mathcal{T}_{S,j}$ is equal to,

$$\mathcal{T}_{S,j}(x) = \begin{cases} x - e_S + e_{S \cup \{j\}} & \text{if } S \cup \{j\} \subset [m] \\ x - e_S & \text{if } S \cup \{j\} = [m] \end{cases} \quad (5.7)$$

The transition rate from S to $\mathcal{T}(S, j)$, $\forall S : j \notin S$ in TMS is given by,

$$Q(x, \mathcal{T}_{S,j}(x)) = \begin{cases} \frac{x_S}{|x|} \left(\frac{U}{h(x, [m], S)} + \mu \sum_{B:j \in B} \frac{x_S}{h(x, B, S)} \right) & \text{if } j \notin D_T(x) \\ 0 & \text{if } j \in D_T(x). \end{cases} \quad (5.8)$$

All other entries in the rate transition matrix other than the diagonal entries are 0 and the diagonal entries are equal to the negative sum of all the entries in that row.

5.4.2 Stability of TMS

First we prove some auxiliary results before proving the main result.

Lemma 30. *In TMS, $\underline{\pi}$ can at most be $\frac{m-1}{m}$.*

Proof.

$$\sum_i \pi_i |x| \leq (m-1)|x| \quad (5.9)$$

$$\implies m\underline{\pi}|x| \leq (m-1)|x| \quad (5.10)$$

$$\implies \underline{\pi} \leq \left(1 - \frac{1}{m}\right) \quad (5.11)$$

$$\implies (1 - \underline{\pi}) \geq \frac{1}{m} \quad (5.12)$$

□

Lemma 31. *In TMS, during not-suppression case, $1 - \bar{\pi} \geq \frac{1}{2m}$ when $|x| \geq 2mT$.*

Proof. In TMS we have,

$$\bar{\pi}|x| \leq \underline{\pi}|x| + T \quad (5.13)$$

$$\implies \bar{\pi} \leq \underline{\pi} + \frac{T}{|x|} \quad (5.14)$$

$$\implies \bar{\pi} \leq \frac{m-1}{m} + \frac{T}{|x|} \quad (5.15)$$

$$(5.16)$$

When $|x| > 2mT$, $\bar{\pi} \leq 1 - \frac{1}{m} + \frac{1}{2m} \implies 1 - \bar{\pi} \geq \frac{1}{2m}$ □

Theorem 18. *The stability region of Threshold Mode Suppression (TMS) is $\lambda > 0$ for any finite threshold $T < \infty$, if $m \geq 2$, $\mu > 0$ and $U > 0$.*

Proof. The proof is similar to the proof of stability of Mode Suppression and the same Lyapunov function with a different constant C_1 works in this case. We will mainly highlight

the differences here. The Lyapunov function is,

$$V(x) = \sum_{i=1}^m ((\bar{\pi} - \pi_i)|x|)^2 + C_1((1 - \bar{\pi}))|x| + C_2\left(M - \sum_i \pi_i|x|\right)^+ \quad (5.17)$$

We will divide the state space into two regions - suppressing region and not-suppressing region. Suppressing region is when $D_T(x)$ is not null set and not-suppressing region is when $D_T(x)$ is null set. First we will start with suppressing region and show that the drift is negative except in a finite set.

Region 1: Suppressing Region: $D_T(x) \neq \phi$

We further divide this region into three regions,

$$\mathcal{R}_{11}: \{x|D_T(x) \neq \phi\} \cap \{x|\bar{\pi}(x) \geq \delta\}:$$

$$\begin{aligned} QV(x) &\leq C_1\lambda + \sum_{j \notin D_T(x)} \frac{1}{m}(1 - \pi_j)(U + \mu\pi_j|x|)(-2|x|(\bar{\pi} - \pi_j) + 1) \\ &\leq C_1\lambda - \bar{\pi}|x|\left(\frac{2U}{m^2} \wedge \mu\right) + 1 \quad \text{if } |x| > M + m - 1 \\ &\leq C_1\lambda - \bar{\pi}|x|g(\mu, U) + 1 \quad \text{if } |x| > M + m - 1 \end{aligned} \quad (5.18)$$

$$\leq -\epsilon \quad \text{for a large } x \text{ where, } g(\mu, U) = \min\left\{\frac{2U}{m^2}, \mu\right\}. \quad (5.19)$$

$$\mathcal{R}_{12}: \{x|D_T(x) \neq \phi\} \cap \{x|\bar{\pi}(x) < \delta, \bar{\pi}(x)|x| \geq \frac{M}{m}\}:$$

$$\begin{aligned} QV(x) &\leq C_1\lambda + \sum_{j \notin D_T(x)} \frac{1}{m}(1 - \pi_j)(U + \mu\pi_j|x|)(-2|x|(\bar{\pi} - \pi_j) + \delta C_2 m^2(m - 1)) \\ &\leq C_1\lambda - \bar{\pi}|x|\left(\frac{2U}{m^2} \wedge \mu(2 - \delta C_2 m^2(m - 1))\right) + K' \\ &\leq C_1\lambda - \bar{\pi}|x|\left(\frac{2U}{m^2} \wedge \mu\right) + K' \end{aligned}$$

$$\leq C_1\lambda - \bar{\pi}|x|g(\mu, U) + K' \quad (5.20)$$

$$\leq -\epsilon \quad \text{for a large } x. \quad (5.21)$$

\mathcal{R}_{13} : $\{x|D_T(x) \neq \phi\} \cap \{x|\bar{\pi}(x) < \delta, \bar{\pi}(x)|x| < \frac{M}{m}\}$:

$$\begin{aligned} QV(x) &\leq C_1\lambda + \sum_{j \notin D_T(x)} \frac{1}{m} (1 - \pi_j)(U + \mu\pi_j|x|) (-2|x|(\bar{\pi} - \pi_j) + 1) \\ &\quad - C_2 \frac{U}{m} \left(\frac{1}{m} - \delta(m(m-1) + 1) \right) \\ &\leq C_1\lambda - \bar{\pi}|x|g(\mu, U) + 1 - C_2 \frac{U}{m} \left(\frac{1}{m} - \delta(m(m-1) + 1) \right) \end{aligned} \quad (5.22)$$

$$\leq -\epsilon \quad \text{for } C_2 \text{ sufficiently large.} \quad (5.23)$$

Region 2: Not-Suppressing Case: $D_{TMS} = \phi$

In this case we have, $(\bar{\pi}(x) - \underline{\pi}(x))|x| < T$.

\mathcal{R}_{21} : $\{x|D_T(x) \neq \phi\} \cap \{x|\bar{\pi}(x) \geq \delta\}$:

$$\begin{aligned} QV(x) &\leq C_1\lambda - (U + \mu\bar{\pi}|x|) \left[\frac{(1 - \bar{\pi})}{m} \left(C_1 - (2T + 1)(m - 1) \right) \right] \\ &\leq C_1\lambda - \mu\bar{\pi}|x| \left[\frac{(1 - \bar{\pi})}{m} \left(C_1 - (2T + 1)(m - 1) \right) \right] \\ &\leq C_1\lambda - \frac{\mu}{2m^2} (C_1 - (2T + 1)(m - 1))\bar{\pi}|x| \end{aligned} \quad (5.24)$$

$$\leq -\epsilon \quad \text{for a large } x. \quad (5.25)$$

\mathcal{R}_{22} : $\{x|D_T(x) \neq \phi\} \cap \{x|\bar{\pi}(x) < \delta, \bar{\pi}(x)|x| \geq \frac{M}{m}\}$:

$$QV(x) \leq C_1\lambda - \frac{\bar{\pi}|x|\mu}{2m^2} \left(C_1 - (2T + 1)(m - 1) - \delta C_2 m^2 (m - 1) \right) \quad (5.26)$$

$$\leq -\epsilon \quad \text{for a large } x. \quad (5.27)$$

$$\mathcal{R}_{23}: \{x | D_T(x) \neq \phi\} \cap \{x | \bar{\pi}(x) < \delta, \bar{\pi}(x)|x| < \frac{M}{m}\}:$$

$$QV(x) \leq C_1 \lambda - \frac{\bar{\pi}|x|\mu}{2m^2} \left(C_1 - (2T + 1)(m - 1) \right) - C_2 \frac{U}{m} \left(\frac{1}{2m} - \delta(m(m - 1) + 1) \right) \quad (5.28)$$

$$\leq -\epsilon \quad \text{for } C_2 \text{ large enough.} \quad (5.29)$$

□

Though the TMS policy is stable for any finite threshold T for any $\lambda > 0$, the sojourn time of TMS depends on the threshold(T) we choose. We found empirically that choosing the threshold twice the number of chunks, that is $T = 2m$ gives the best sojourn times. More details are given in the simulations section.

5.5 Simulation Results

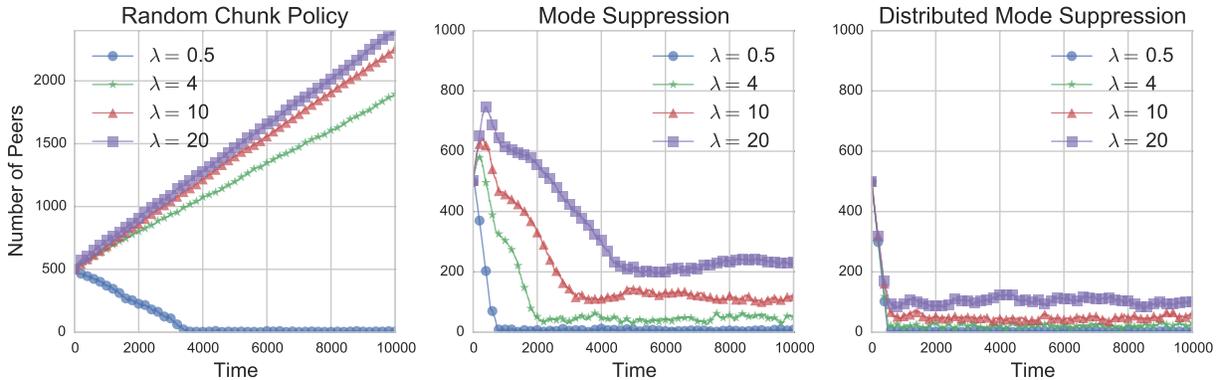


Figure 5.1: Number of peers in the system when $m = 5$, $U = 1$ and $\mu = 1$. Random becomes unstable in some cases, whereas MS and DMS are always stable.

In this section, we show the results from numerical simulations that illustrate the performance of different chunk selection policies. Recall that our candidate policies are (i) random chunk selection, (ii) rarest-first, (iii) rare chunk, (iv) common chunk, (v) group suppression, (vi) mode suppression, and (vii) distributed mode suppression. A description of these poli-

cies can be found in Section 5.1. For all the simulations, we kept the peer contact rate and seed contact rate to 1. To simulate a Poisson process, we make use of the fact that inter arrival times of a Poisson process follow an exponential distribution. Each peer in the system, including the seed, generates an exponential random variable with mean $\frac{1}{\mu} = \frac{1}{U} = 1$, and the peer or the seed with the smallest value gets a chance to contact another peer. After the contact, a chunk transfer takes place instantaneously according to the chosen chunk selection policy.

5.5.1 Stability of Mode Suppression Policy

We begin the simulation with 500 empty peers. Whenever a peer receives all the chunks, it immediately leaves the system. In Figure 5.1, we plot the number of peers in the system as time progresses for three different policies, namely (i) random chunk selection, (ii) mode suppression, and (iii) distributed mode suppression. The purpose of simulating the random chunk selection policy, which is known to be unstable, is to provide a visual representation of what an unstable regime appears like, in order to compare with stable policies. In this simulation, the number of chunks is taken as 5, and the peer arrival rate (λ) is varied. We observe that when the peer arrival rate is less than seed rate ($\lambda = 0.5 < 1 = U$), the random chunk selection policy is stable and in all other cases $\lambda > U$, the number of peers grows large and the system is unstable. However, in case of mode suppression and distributed mode suppression, the system is stable for all arrival rates.

5.5.2 Missing Piece Syndrome in Random Chunk Selection

We observed in Figure 5.1 that the random chunk selection policy is not stable when $\lambda > U$. We determine the reason for this instability by observing the evolution of the chunk frequency. In Figure 5.2, we plot the time evolution of the number of peers and the fraction of peers having different chunks in the system, for the random chunk selection policy with $m = 5$ and $\lambda = 4$. We see that when number of peers becomes large, one chunk remains rare. As time progresses, the chunk represented by the red/starred line becomes rare and

remains rare forever. However, all other chunks are available with most of the peers. This is precisely the formation of the *one-club* caused by the *missing piece syndrome*.

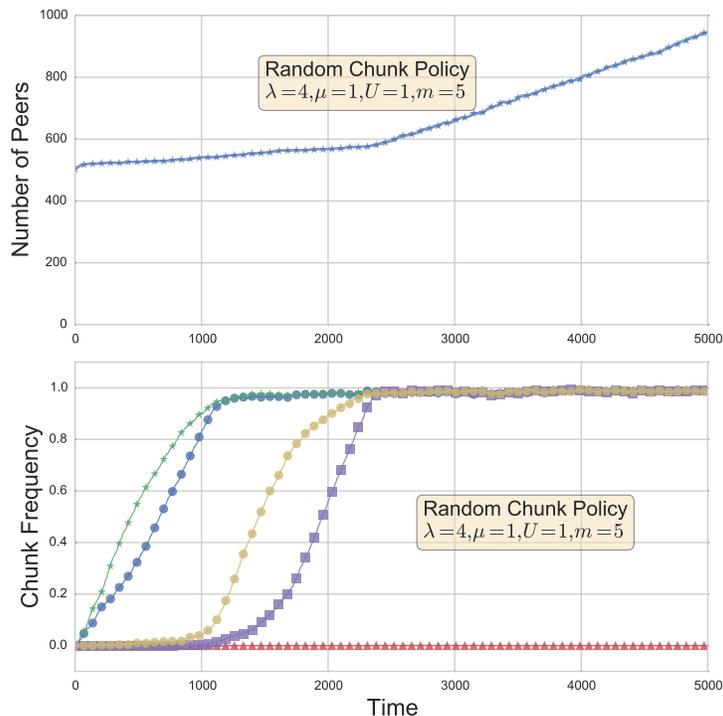


Figure 5.2: Evolution of peers and chunk frequencies under the random chunk selection policy. One of the chunks becomes a “missing chunk” (red/starred line).

5.5.3 Chunk Frequency Evolution

A stable chunk selection policy has to be robust to the one-club state. In other words, a stable policy should be able to boost the frequency of a rare chunk. To see how different policies handle the one-club situation, we start the system with 500 peers that have all the chunks except first chunk (i.e., all peers are part of the one-club). In Figure 5.3, we plot the evolution of the chunk frequency for different policies under this initial condition. We see that when using the rarest-first policy, the rare chunk remains rare and abundant chunks remain abundant—a clear sign of instability. In all stabilizing policies, the rare

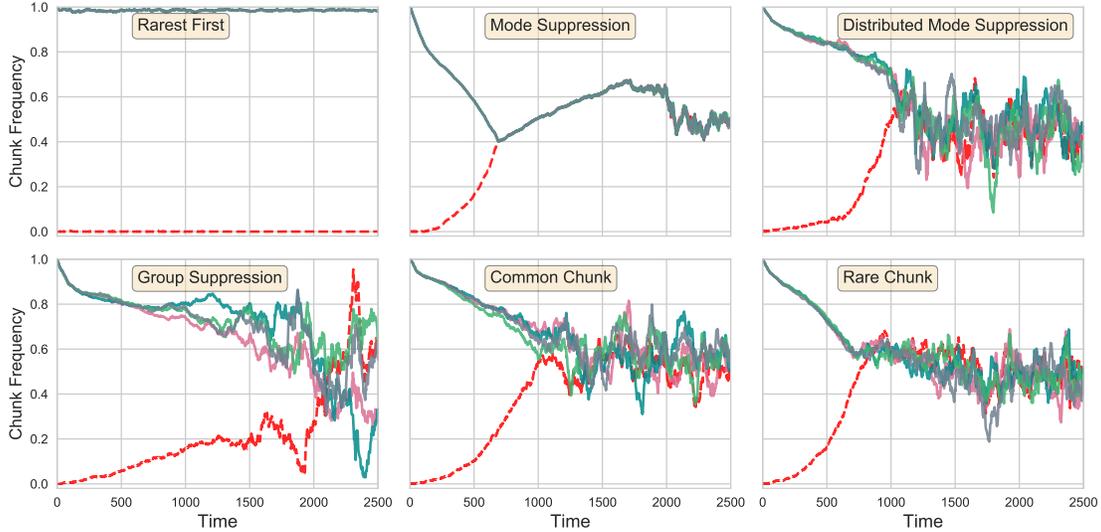


Figure 5.3: Chunk frequency evolution in a system with $m = 5$ chunks under different policies when starting from the state of a “missing-chunk” (whose frequency is indicated by a red/dashed line). Rarest-first is clearly unstable, since it cannot recover, whereas the other protocols manage to bring the chunk back into peer circulation and stabilize the system.

chunk is made available by giving priority to that chunk in some way. For instance, in case of mode suppression, no other chunk will be transmitted until the frequency of the rare chunk is equal to the frequency of all other chunks. Once this happens, the frequencies of the different chunks remain almost same, and hence we only see a thin spread across the frequencies. Other policies also manage to bring the rare chunk back into circulation and the corresponding statistics become similar to all other chunks. We also observe that the *stabilization time* to increase the frequency of rare chunk to the same level as that of other chunk frequencies, is much shorter for MS and DMS when compared to other algorithms.

5.5.4 Sojourn times

In addition to stability, an important performance metric is the sojourn time of a peer, which is defined as the amount of time a peer spends in the system collecting all chunks before leaving. In Figure 5.4 and Figure 5.5, we plot the average sojourn times of the peers for different policies, for the arrival rate $\lambda = 4$ and different numbers of chunks that the

file is divided into (m). In Figure 5.4 we only included the first 1000 peers that left the system and in the Figure 5.5 we included rest of the peers. We can observe that the TMS has best sojourn times in both cases. Though the Group Suppression has good stationary sojourn times the mixing times are large and have very high variance. In $m = 3$, the average stationary sojourn times of TMS, DMS and group suppression are very close to m . Since the rate of peer contact is 1, this fact indicates that these algorithms attains the best possible trade-off between suppression (to keep peers in the system) and sharing (to enable peers to gather chunks).

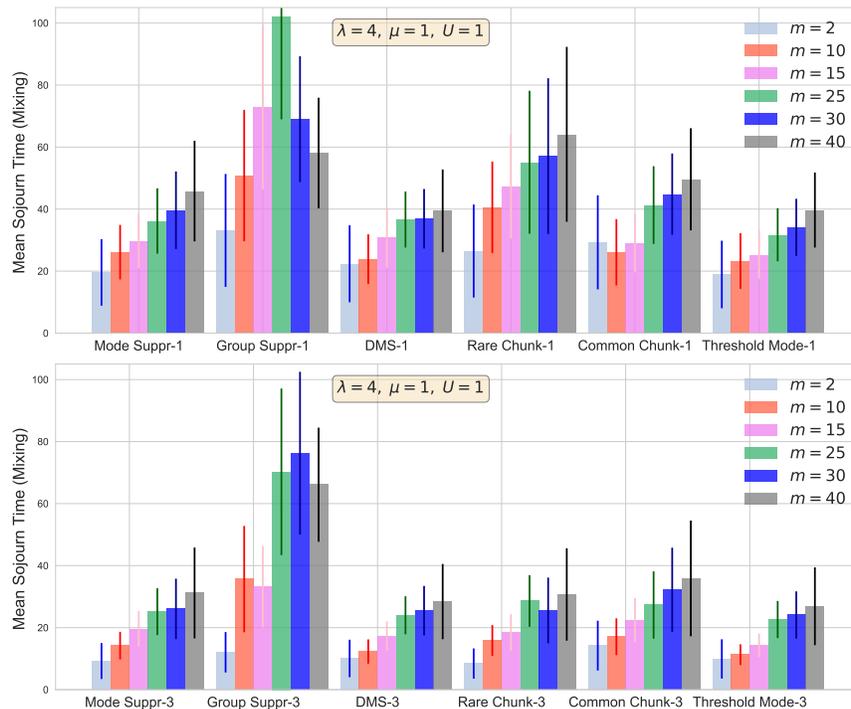


Figure 5.4: Mean mixing sojourn times of policies for different values of m . Distributed mode suppression has the best performance in all cases.

5.6 Conclusions

In this chapter, we analyzed the scaling behavior of a P2P swarm with reference to its stability when subjected to an arbitrary arrival rate of peers. It has been shown earlier

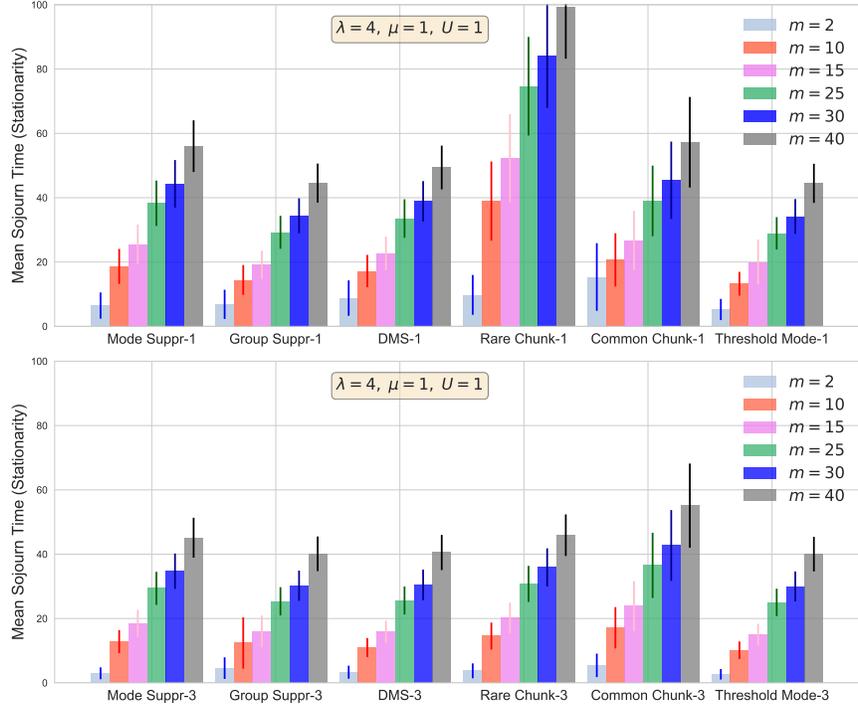


Figure 5.5: Mean stationary sojourn times of policies for different values of m . Distributed mode suppression has the best performance in all cases.

that not all chunk sharing policies are stable in such a regime, and our goal was to design a simple and stable policy that yields low sojourn times. Our main observation was that, contrary to the traditional approach of boosting the availability of rare chunks, preventing the spread of the most frequent chunk(s) yields a simple and stable policy that we entitled mode suppression (MS), and a generalized version that we refer to as Threshold Mode Suppression (TMS). We analytically proved their stability, and showed using numerical studies that Threshold Mode Suppression achieves near optimal sojourn times.

Our results indicate that there is a delicate trade-off between sharing (i.e., uploading a useful chunk if at all possible) and suppression (i.e., trying to reduce chunk transfers to keep peers in the system so that they can help others). The chunk selection policy has a fundamental impact on this trade-off. On one hand, by suppressing some chunk sharing (as in the GS, MS or DMS algorithms), we can ensure peers stay longer at the expense

of increasing sojourn time, with too much suppression leading to instability. On the other hand, trying too hard to be work conserving (maximizing sharing as in random or RF) with the idea of reducing sojourn times can lead to instability due to chunk starvation.

6. CONCLUSIONS

In this thesis, we studied the equilibrium and resource allocation efficiency of different network market systems. In chapter 2, we studied the Internet Service Provider (ISP) transit billing market. We showed that the current 95th Percentile percentile billing mechanism is unfair through a study of measured transit data volumes. We then considered the design of a fair cost allocation scheme using the concept of Shapley value percentile (SVP) from the framework of cooperative game theory. To overcome the complexity of SVP, we proposed a new billing scheme using a convex optimization framework, and a new metric called Provision Ratio. The new billing mechanism is flexible enough for the transit provider to incorporate billing constraints and close to SVP in terms of fairness.

After analyzing market with finite agents then we examined market systems with large number of agents in chapters 3 and 4. First, in chapter 3 we modeled auction-based scheduling in cellular system as a Bayesian game. As the number of agents becomes asymptotically large, it is difficult to find a Bayesian Nash equilibrium. We then used the notion of Mean Field Equilibrium (MFE) and established that this system has a unique MFE. We also showed that the best response policy in this system takes the form of longest queue first policy. Then in chapter 4, we considered the problem of determining the optimal price selection policy in a competitive marketplace that has many service providers that are relatively short lived. We showed that our marketplace can be thought of a version of a first price reverse auction, and characterized the best response bidding policy under the mean field assumption. We again established the existence of a Mean Field Equilibrium.

Finally in chapter 5, we analyzed a P2P network file sharing system in which there is a single seed that has all chunks, and non-altruistic peers that leave the system as soon as they obtain all the chunks. It is established in the literature that rarest-first chunk selection policy, and in general any work conserving policy is unstable if the arrival rate is large. Although counter intuitive, some amount of suppression of sharing chunks is needed

to promote system stability. In this spirit, we proposed a simple chunk selection policy called Mode Suppression in which we suppress the sharing of chunks that are widely available in the system. We showed that this policy is easy to implement as it depends only on the marginal chunk frequencies, and is stable for any peer arrival rate. We also developed a variant called Threshold Mode Suppression in which we suppress the most popular chunks only if they are significantly more abundant than the least frequency chunks. We showed through simulations that TMS has the best sojourn times among all the policies.

BIBLIOGRAPHY

- [1] AboveNet. Monitoring 95-percentile. <http://john.de-graaff.net/wiki/doku.php/links/95-percentile>.
- [2] S. Adlakha and R. Johari. Mean field equilibrium in dynamic games with complementarities. In *IEEE Conference on Decision and Control (CDC)*, pages 6633–6638. IEEE, 2010.
- [3] Sachin Adlakha, Ramesh Johari, and Gabriel Y Weintraub. Equilibria of dynamic games with many players: Existence, approximation, and market structure. *Journal of Economic Theory*, 156:269–316, 2015.
- [4] Axis Internet. 95th Percentile. <http://www.axint.net/95th>.
- [5] M. Benaim and J-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11):823–838, 2008.
- [6] C. Berge. *Topological Spaces: Including a treatment of multi-valued functions, vector spaces, and convexity*. Courier Corporation, 1963.
- [7] Omer Bilgen and Aaron Wagner. A new stable peer-to-peer protocol with non-persistent peers. In *Proceedings of INFOCOM*, pages 1783–1790, 2017.
- [8] Patrick Billingsley. *Convergence of probability measures*, volume 493. Wiley-Interscience, 2009.
- [9] V. Borkar and R. Sundaresan. Asymptotics of the invariant measure in mean field models with jumps. In *49th Annual Allerton Conference on Communication, Control, and Computing*, pages 1258–1263, 2011.
- [10] Ignacio Castro, Rade Stanojevic, and Sergey Gorinsky. Using Tuangou to reduce IP transit costs. *IEEE/ACM Transactions on Networking*, 2013.

- [11] H. Chang, S. Jamin, and W. Willinger. To peer or not to peer: Modeling the evolution of the internet's AS-level topology. In *Proc. IEEE INFOCOM*, 2006.
- [12] Cline Communications. 95th Percentile billing. <http://clinecommunications.net/?ID\=33>.
- [13] Bram Cohen. Incentives build robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer systems*, volume 6, pages 68–72, 2003.
- [14] A. Dhamdhere and C. Dovrolis. The Internet is Flat: Modeling the Transition from a Transit Hierarchy to a Peering Mesh. In *Proceedings of CoNEXT*, 2010.
- [15] X. Dimitropoulos, P. Hurley, A. Kind, and M. Stoecklin. On the 95-Percentile Billing Method. In *Proc. of PAM*, 2009.
- [16] A. Eryilmaz and R. Srikant. Joint Congestion Control, Routing and MAC for Stability and Fairness in Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 24(8):1514–1524, August 2006.
- [17] Bin Fan, John Lui, and Dah-Ming Chiu. The design trade-offs of BitTorrent-like file sharing protocols. *IEEE/ACM Transactions on Networking (TON)*, 17(2):365–376, 2009.
- [18] G4 Communications. 95th Percentile Usage Billing Policy. http://www.g4communications.com/docs/G4_95th_Percentile_Usage.pdf.
- [19] A. Ghose, P. G Ipeirotis, and A. Sundararajan. Opinion mining using econometrics: A case study on reputation systems. In *Annual Meeting-Association for Computational Linguistics*, volume 45, page 416, 2007.
- [20] Carl Graham and Sylvie Méléard. Chaos hypothesis for a system interacting through shared resources. *Probability Theory and Related Fields*, 100(2):157–174, 1994.
- [21] Matthias Grossglauser and David N. C. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Trans. Netw.*, 10(4), August 2002.

- [22] Sangtae Ha, Soumya Sen, Carlee Joe-Wong, Youngbin Im, and Mung Chiang. Tube: time-dependent pricing for mobile data. In *SIGCOMM*, pages 247–258, 2012.
- [23] Bruce Hajek and Ji Zhu. The missing piece syndrome in peer-to-peer communication. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 1748–1752. IEEE, 2010.
- [24] Hernandez-Lerma and Jean Lasserre. *Further Topics on Discrete-Time Markov Control Processes*. Springer, Newyork, 1999.
- [25] O. Onesimo Hernandez-Lerma and Jean-Bernard Lasserre. *Further topics on discrete-time Markov control processes*. Applications of mathematics. Springer, 1999.
- [26] K. Iyer, R. Johari, and M. Sundararajan. Mean field equilibria of dynamic auctions with learning. *Management Science*, 60(12):2949–2970, 2014.
- [27] Krishnamurthy Iyer, Ramesh Johari, and Mukund Sundararajan. Mean field equilibria of dynamic auctions with learning. *ACM SIGecom Exchanges*, 10(3):10–14, 2011.
- [28] Vijay Krishna. *Auction theory*. Academic press, 2009.
- [29] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japan Journal of Mathematics*, 2007.
- [30] K. Levenberg. A method for the solution of certain problems in least squares. *Quarterly of applied mathematics*, 2:164–168, 1944.
- [31] J. Li, R. Bhattacharyya, S. Paul, S. Shakkottai, and V. Subramanian. Incentivizing sharing in realtime D2D streaming networks: A mean field game perspective. In *Proceedings of IEEE Infocom*, Hong Kong, 2015.
- [32] J. Li, B. Xia, X. Geng, H. Ming, S. Shakkottai, V. Subramanian, and L. Xie. Energy coupon: A mean field game perspective on demand response in smart grids. In *Proceedings of ACM Sigmetrics*, Portland, OR, 2015.
- [33] Xiaojun Lin and Ness B. Shroff. Joint rate control and scheduling in multihop wireless

- networks. In *in Proceedings of IEEE Conference on Decision and Control*, pages 1484–1489, 2004.
- [34] Mayank Manjrekar, Vinod Ramaswamy, and Srinivas Shakkottai. A mean field game approach to scheduling in cellular systems. In *Proceedings of INFOCOM*, pages 1554–1562, 2014.
- [35] Laurent Massoulié and Milan Vojnovic. Coupon replication systems. *IEEE/ACM Transactions on Networking (TON)*, 16(3):603–616, 2008.
- [36] Diego Ximenes Mendes, Edmundo de Souza e Silva, Daniel Menasche, Rosa Leao, and Don Towsley. An experimental reality check on the scaling laws of swarming systems. In *Proceedings of INFOCOM*, pages 1647–1655, 2017.
- [37] Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.
- [38] Sean P Meyn, Richard L Tweedie, and Peter W Glynn. *Markov chains and stochastic stability*, volume 2. Cambridge University Press, 2009.
- [39] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, pages 583–601, 2002.
- [40] Michael J. Neely, Eytan Modiano, and Chih ping Li. Fairness and optimal stochastic control for heterogeneous networks. In *Proceedings of IEEE INFOCOM*, pages 1723–1734, 2005.
- [41] Ilkka Norros, Hannu Reittu, and Timo Eirola. On the stability of two-chunk file-sharing systems. *Queueing Systems*, 67(3):183–206, 2011.
- [42] William B Norton. Internet Transit Prices - Historical and Projected. <http://drpeering.net/AskDrPeering/blog>, 2014.
- [43] William B Norton. Transit Traffic at Internet Exchange Points? *Drpeering.net blog*: <http://drpeering.net/AskDrPeering/blog>, 2014.

- [44] A. Odlyzko. Internet Pricing and the History of Communications. *COMPUTER NETWORKS*, 36:493–517, 2001.
- [45] Barlas Oguz, Venkat Anantharam, and Ilkka Norros. Stable distributed p2p protocols based on random peer sampling. *IEEE/ACM Transactions on Networking (TON)*, 23(5):1444–1456, 2015.
- [46] Martin L Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, Inc., 1994.
- [47] Dongyu Qiu and Rayadurgam Srikant. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In *ACM SIGCOMM computer communication review*, volume 34, pages 367–378. ACM, 2004.
- [48] Hannu Reittu. A stable random-contact algorithm for peer-to-peer file sharing. In *IWSOS*, pages 185–192. Springer, 2009.
- [49] Alvin E. Roth. *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. Cambridge University Press, 1988.
- [50] SANDVINE. Global Internet phenomena report. URL: <https://www.sandvine.com/trends/global-internet-phenomena/>, 2016.
- [51] S. Shakkottai and R. Johari. Demand Aware Content Distribution on the Internet. *IEEE/ACM Transactions on Networking*, 18(2), April 2010.
- [52] R. Stanojevic, N. Laoutaris, and P. Rodriguez. On Economic Heavy Hitters: Shapley Value Analysis of 95th-percentile Pricing. In *Proc. of ACM SIGCOMM IMC*, 2010.
- [53] Ralph E Strauch. Negative dynamic programming. *The Annals of Mathematical Statistics*, 37(4):871–890, 1966.
- [54] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. In *IEEE Transactions on Automatic Control*, volume 37, pages 1936–1948, 1992.

- [55] H. Tembine, J.Y. Le Boudec, R. El-Azouzi, and E. Altman. Mean field asymptotics of Markov decision evolutionary games and teams. In *International Conference on Game Theory for Networks, (GameNets)*, pages 140–150, 2009.
- [56] Jiaming Xu and Bruce Hajek. The supermarket game. In *ISIT*, 2012.
- [57] Xiangying Yang and Gustavo De Veciana. Performance of peer-to-peer networks: Service capacity and role of resource sharing policies. *Performance evaluation*, 63(3):175–194, 2006.
- [58] Ji Zhu and Bruce Hajek. Stability of a peer-to-peer communication system. *IEEE Transactions on Information Theory*, 58(7):4693–4713, 2012.

APPENDIX A

PROOFS FROM CHAPTER 3

A.1 Proof of Lemma 2

We may rewrite the the definition of T_ρ in (3.9) as

$$(T_\rho f)(q) = \inf_{x \in \mathbb{R}^+} S_f(q, x) \tag{A.1}$$

where $S_f(q, x) = C(q) + r_\rho(x) + \beta \mathbb{E}_{Q_1}[f(Q_1)|q, x]$. Given the current state and action pair, (q, x) , the first two terms in $S_f(q, x)$ constitute the current cost, while the last term is the future expected cost, where Q_1 is one-step future state variable. Further, from (3.9), we have

$$\begin{aligned} \mathbb{E}_{Q_1}[f(Q_1)|q, x] &= (1 - p_\rho(x)) \mathbb{E}_A[f(q + A)] \\ &\quad + p_\rho(x) \mathbb{E}_A[f((q - 1)^+ + A)]. \end{aligned}$$

The proof proceeds through a verification of the assumptions of Theorem 8.3.6 in [25]. An exception is that action space in our case is not a compact set which violates Assumption 8.3.1(a) in [25]. However, this assumption can be overridden if the statement of Lemma 8.3.8(a) in [25], equivalently Condition (3) below, holds true. Further, we desire to show the existence of a $j \in \mathbb{N}$ such that T_ρ^j is a contraction mapping. Since Theorem 8.3.6 is derived for $j = 1$, we replace Assumption 8.3.2(b) with Condition (5) given below.

Now, we prove the following statements.

1. $C(q) + r_\rho(x)$ is a continuous function in $x \in \mathbb{R}^+$.
2. $\mathbb{E}_{Q_1}[f(Q_1)|q, x]$ is continuous in $x \in \mathbb{R}^+$ for every $f \in \mathcal{V}$.
3. For any $f \in \mathcal{V}$, there exists a measurable function $\theta_f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that $\theta_f(q)$ attains minimum in (A.1). Further, $S_f(q, \theta_f(q))$ is a measurable function for any $f \in \mathcal{V}$.

4. There exists a nonnegative constant c_1 such that $\sup_x |C(q) + r_\rho(x)| \leq c_1 w(q)$ where $w(q) = \max\{C(q), 1\}$.
5. There exists $j \in \mathbb{N}$ and c_2 with $c_2 < 1$ such that $\beta^j \sup_{\vec{x}} \mathbb{E}_{Q_j}[w(Q_j)|q, \vec{x}] \leq c_2 w(q) + c_3$ where Q_j is j -step future state variable and \vec{x} is a j -length sequence of actions.
6. The function $\mathbb{E}[w(Q_1)|q, x]$ is continuous in $x \in \mathbb{R}^+$

Conditions (1) and (2) are obvious from the continuity of $r_\rho(x)$ and $p_\rho(x)$. Further, as derived in (3.10), $\theta_f(q) = \Delta f(q)^+$ where $\Delta f(q) = \mathbb{E}_A(f(q + A) - f((q - 1)^+ + A))$. The measurability of functions $\theta_f(q)$ and $S_f(q, \theta(q))$ are evident from their definitions. Condition (4) holds true from the definition of $w(q)$ and from the fact that

$$r_\rho(x) \leq \lim_{y \rightarrow \infty} r_\rho(y) < (M - 1) \int_0^\infty (1 - \rho(x)) dx < (M - 1)E$$

where the last inequality follows as $\rho \in \mathcal{P}$. Condition (5) follows from the fact that,

$$\begin{aligned} \beta^j \mathbb{E}_{Q_j}[w(Q_j)|q, \vec{x}] &\leq \beta^j \max\{1, C(q + j\bar{A})\} \\ &\leq \beta^j (k_1 w(q) + k_2), \end{aligned}$$

where \bar{A} , as defined in Assumption 1, is the maximum arrival possible between any two adjacent auctions and $k_1 > 0, k_2$ are some constants independent of j . The above results follows from (A.2) and the definition of $w(q)$. Then, there exists a j such that $\beta^j k_1 = c_2 < 1$ and hence (5) holds. Finally, the last condition follows from Condition (2) as $w(q) \in \mathcal{V}$.

Given that the above conditions are met, we can prove the first statement of Lemma 2. The proof is essentially identical to that of Theorem 8.3.6 in [25]. The second statement of the lemma can be obtained by comparing (3.9) and (3.7). The last part of the lemma follows from (3.10).

A.2 Proofs from Section 3.5

In this section, we present details of proofs that were omitted from Section 3.5. We divide this section into parts based on the development of that section.

A.2.1 Proofs Pertaining to Section 3.5-A: Step 1

Lemma 32. *Suppose $\rho_1, \rho_2 \in \mathcal{P}$. Then, $\|\hat{\theta}_{\rho_1} - \hat{\theta}_{\rho_2}\|_w \leq K\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w$*

Proof. For any $q \in \mathbb{R}^+$, by the definition of $\hat{\theta}_\rho$ we have,

$$\begin{aligned}
& |\hat{\theta}_{\rho_1}(q) - \hat{\theta}_{\rho_2}(q)| \\
&= |\beta[\mathbb{E}_A[\hat{V}_{\rho_1}(q+A) - \hat{V}_{\rho_1}((q-1)^+ + A) - \hat{V}_{\rho_2}(q+A) + \hat{V}_{\rho_2}((q-1)^+ + A)]]| \\
&\leq \beta\mathbb{E}_A|\hat{V}_{\rho_1}(q+A) - \hat{V}_{\rho_2}(q+A)| + \beta\mathbb{E}_A|\hat{V}_{\rho_1}((q-1)^+ + A) - \hat{V}_{\rho_2}((q-1)^+ + A)| \\
&\leq \beta\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w\mathbb{E}_A(w(q+A) + w((q-1)^+ + A)) \leq K\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w w(q)
\end{aligned}$$

□

Lemma 33. *Let $\rho \in \mathcal{P}$ and $f_1, f_2 \in \mathcal{V}$. Then,*

$$\|T_\rho f_1 - T_\rho f_2\|_w \leq \hat{K}\|f_1 - f_2\|_w \quad (\text{A.2})$$

Proof. Using the characterization of T_ρ from eq. (3.11), we have that, for any $q \in \mathbb{R}^+$

$$\begin{aligned}
|T_\rho f_1(q) - T_\rho f_2(q)| &\leq \beta\|f_1 - f_2\|K_1 w(q) + \left| \int_{\beta\Delta f_2(q)}^{\beta\Delta f_1(q)} |\rho^{M-1}(u)| du \right| \\
&\leq \beta\|f_1 - f_2\|K_1 w(q) + \beta|\Delta f_1(q) - \Delta f_2(q)| \\
&\leq \beta(K_1 + K_2)\|f_1 - f_2\|w(q)
\end{aligned}$$

□

A.2.2 Proofs Pertaining to Section 3.5-A: Step 2

Proof of Lemma 5. For brevity, denote $\Pi_{\rho,\theta}(\cdot)$ be $\Pi(\cdot)$ and $\Upsilon_{\rho,\theta}^{(k)} = \Upsilon^{(k)}$. Let $-\tau$ be the last time before 0 the chain regenerated. We have

$$\Pi(B) = \sum_{k=0}^{\infty} \mathbb{P}(B, \tau = k) \quad (\text{A.3})$$

$$= \sum_{k=0}^{\infty} \mathbb{P}(\tau = k) \mathbb{P}(B | \tau = k) \quad (\text{A.4})$$

Since the regeneration events are independent of the queue-length and occur geometrically with probability $(1 - \beta)$, $\mathbb{P}(\tau = k) = (1 - \beta)\beta^k$. Hence,

$$\Pi(B) = \sum_{k=0}^{\infty} (1 - \beta)\beta^k \mathbb{P}(Q_0 \in B | \tau = k) \quad (\text{A.5})$$

$$= \sum_{k=0}^{\infty} (1 - \beta)\beta^k \mathbb{E}(\mathbb{E}(\mathbf{1}_{Q_0 \in B} | \tau = k, Q_{-k} = Q) | \tau = k) \quad (\text{A.6})$$

$$= \sum_{k=0}^{\infty} (1 - \beta)\beta^k \mathbb{E}(\Upsilon^{(k)}(B|Q) | \tau = k) \quad (\text{A.7})$$

$$= \sum_{k=0}^{\infty} (1 - \beta)\beta^k \mathbb{E}_{\Psi_R}(\Upsilon^{(k)}(B|Q)). \quad (\text{A.8})$$

since $Q_{-k} \sim \Psi_R$ given $\tau = k$. □

Lemma 34. $\liminf_{n \rightarrow \infty} \Upsilon_{\rho_n}^{(k)}(B|q) \geq \Upsilon_{\rho}^{(k)}(B|q)$

Proof. The proof proceeds through mathematical induction on k . For $k = 0$, we have $\Upsilon_{\rho_n}^{(0)}(B|q) = \mathbf{1}_{(q \in B)}$ and hence the hypothesis holds true. Suppose that the hypothesis is true till $k = m - 1$. To prove the lemma, we just need to verify that the hypothesis holds for $k = m$. Let $\mathbb{P}_{q,\rho}(\cdot)$ be the one step transition kernel of the queue dynamics conditioned on the following facts: the initial state is q , the bids are generated according to the optimal policy given by Corollary 1 and no regeneration. Verify that $\mathbb{P}_{q,\rho_n}(\cdot) \implies \mathbb{P}_{q,\rho}(\cdot)$ by considering the integrals of a bounded continuous function. Then, by Skorokhod representation theorem, there exists X_n and X on common probability space such that $X_n \sim \mathbb{P}_{q,\rho_n}$, $X \sim \mathbb{P}_{q,\rho}$ and

$X_n \rightarrow X$ a.s. We have,

$$\liminf \Upsilon_{\rho_n}^{(m)}(B|q_n) = \liminf \mathbb{E}(\Upsilon_{\rho_n}^{(m-1)}(B|X_n)) \quad (\text{A.9})$$

$$\geq \mathbb{E}(\liminf \Upsilon_{\rho_n}^{(m-1)}(B|X_n)) \quad (\text{A.10})$$

$$\geq \mathbb{E}(\Upsilon_{\rho}^{(m-1)}(B|X)) \quad (\text{A.11})$$

$$= \Upsilon_{\rho}^{(m)}(B|q), \quad (\text{A.12})$$

where eq. (A.10) follows from Fatou's lemma, and eq. (A.11) follows from the induction hypothesis. \square

A.2.3 Proofs Pertaining to Section 3.5-A: Step 3

Details of proof of Lemma 6. To complete the proof, we need to show that the expected bid under the cumulative distribution function $\hat{\rho}$ is bounded from above by a constant that is independent of $\hat{\rho}$. To that end, define a new Markov random process \tilde{Q}_k with the probability transition matrix

$$\mathbb{P}(\tilde{Q}_{k+1} \in B | \tilde{Q}_k = q) = \beta \mathbf{1}_{(q+\bar{A} \in B)} + (1 - \beta) \Psi_R(B) \quad (\text{A.13})$$

where \bar{A} is the maximum possible arrival between any two consecutive auction instants. The process \tilde{Q}_k has an invariant distribution which is given by,

$$\tilde{\Pi}(B) = \sum_{k=0}^{\infty} (1 - \beta) \beta^k \mathbb{E}_{\Psi_R}(\mathbf{1}_{(q+k\hat{A}) \in B}). \quad (\text{A.14})$$

The proof of the above result is identical to that of Lemma 5. For any q given, the above probability measure (A.13) stochastically bounds the probability measure in eq. (3.6). Therefore, it can be shown that $\tilde{\Pi}$ stochastically dominates Π_{ρ} for all $\rho \in \mathcal{P}$, i.e, $\Pi_{\rho} \preceq \tilde{\Pi}$.

Now, the expected value of the optimal bid function $\hat{\theta}_\rho(q)$ under Π_ρ satisfies,

$$\mathbb{E}_{\Pi_\rho}[\hat{\theta}_\rho(q)] \leq \mathbb{E}_{\tilde{\Pi}}[\hat{\theta}_\rho(q)] \leq \mathbb{E}_{\tilde{\Pi}}[\hat{V}_\rho(q + \bar{A})] \quad (\text{A.15})$$

$$\leq \sum_{k=0}^{\infty} (1 - \beta) \beta^k \mathbb{E}_{\Psi_R}(\hat{V}_\rho(q + (k + 1)\bar{A})) \quad (\text{A.16})$$

Above, the first inequality follows from stochastic dominance of $\tilde{\Pi}$ and the second inequality is due to the definition of optimal bid function. From (3.7), we can observe that for any ρ , $\hat{V}_\rho(q) \leq \sum_{k=0}^{\infty} \beta^k C(q + k\bar{A})$ independent of ρ . Since $C(q) \in O(q^m)$ for some m , we have $\hat{V}_\rho(q) \in O(q^m)$. Then, $\mathbb{E}_{\Psi_R}(\hat{V}_\rho(q + (k + 1)\bar{A})) \in O(k^m)$ as the moments of Ψ_R are bounded. This directly gives that $\mathbb{E}_{\Pi_\rho}[\hat{\theta}_\rho(q)]$ is bounded by the some constant that is independent of ρ and, hence independent of $\hat{\rho}$. \square

Lemma 35. *In \mathcal{P} , pointwise convergence implies uniform convergence.*

Proof. Let $\rho_n, \rho \in \mathcal{P}$ and $\rho_n \rightarrow \rho$ point-wise. Given $\epsilon > 0$, choose L large enough so that $\rho(L) > 1 - \epsilon$. Since ρ is continuous function by definition, it is uniformly continuous on the compact set $[0, L]$. Therefore, we can construct a sequence $0 = x_1 < x_2 < \dots < x_k = L$ such that and $|\rho(x_{i+1}) - \rho(x_i)| < \epsilon$. Let J be large enough so that for all $n > J$, $|\rho(x_i) - \rho_n(x_i)| < \epsilon$ for all i . For any y such that $x_i < y < x_{i+1}$,

$$\begin{aligned} |\rho(y) - \rho_n(y)| &< |\rho(y) - \rho(x_i)| + |\rho(x_i) - \rho_n(x_i)| + |\rho_n(y) - \rho_n(x_i)| \\ &< |\rho(x_{i+1}) - \rho(x_i)| + |\rho(x_i) - \rho_n(x_i)| + |\rho_n(x_{i+1}) - \rho_n(x_i)| \\ &< 2|\rho(x_{i+1}) - \rho(x_i)| + |\rho(x_i) - \rho_n(x_i)| + 2\epsilon < 5\epsilon \end{aligned} \quad (\text{A.17})$$

While if $L < y$, then

$$|\rho(y) - \rho_n(y)| < |\rho(y) - \rho_n(L)| + |\rho_n(L) - \rho(L)| + |\rho(y) - \rho(L)| \quad (\text{A.18})$$

$$< 1 - \rho(L) + \epsilon + \epsilon + 1 - \rho(L) < 4\epsilon. \quad (\text{A.19})$$

Therefore, $|\rho(y) - \rho_n(y)| < 5\epsilon$ for all $n > J$ and hence ρ_n converges to ρ uniformly. \square

A.2.4 Proofs Pertaining to Section 3.5-B

Proof of Lemma 7. We know that $\Pi([a, b]|\rho, \theta) = \sum_{k \geq 0} (1 - \beta)\beta^k \mathbb{E}_{\Psi_R}(\Upsilon_\rho^{(k)}([a, b]|Q_0))$. Let A_k be the net arrivals and D_k be the net departures till time k . Then,

$$\Upsilon_\rho^{(k)}([a, b]|Q_0) = \mathbb{E}(\mathbf{1}_{(Q_0 + A_k - D_k \in [a, b])} | Q_0) \tag{A.20}$$

$$= \mathbb{E}(\mathbb{E}(\mathbf{1}_{(Q_0 + A_k - D_k \in [a, b])} | D_k, Q_0) | Q_0) \tag{A.21}$$

$$= \mathbb{E}(\mathbb{E}(\mathbf{1}_{(A_k \in [a - Q_0 + D_k, b - Q_0 + D_k])} | Q_0, D_k) | Q_0) \tag{A.22}$$

$$\leq c_1 \cdot (b - a). \tag{A.23}$$

The above results hold since the random variable A_k is independent of Q_0 and D_k for any k and it has a bounded density function. Therefore, $\mathbb{E}_{\Psi_R}(\Upsilon_\rho^{(k)}([a, b]|Q_0)) \leq c \cdot (b - a)$ for all $k > 0$. For $k = 0$, we know that Ψ_R has a bounded density which implies $\Psi_R([a, b]) \leq c_1 \psi \cdot (b - a)$. These two results prove that there is a large enough c such that $\Pi_\rho([a, b]) < c \cdot (b - a)$. \square

APPENDIX B

PROOFS FROM CHAPTER 4

B.1 Proofs from Section 4.4

B.1.1 Details of proof of Lemma 10

By grouping the terms containing x in equation (4.10), we get

$$V_\rho(q, \alpha, \xi, \eta) \tag{B.1}$$

$$= \delta \mathbb{E}_D \left[V_\rho((q - D)^+, \alpha, \xi, \eta) \right] - C(q) +$$

$$\sup_{x \in [-z, \infty]} \max_{\eta_c \in [0, 1]} \left\{ p_\rho(x) (-K(\eta_c, \eta) + x + z g(\alpha, \xi) + \delta \Delta V_\rho(q, \alpha, \xi, \eta)) \right\},$$

$$= \delta \mathbb{E}_D \left[V_\rho((q - D)^+, \alpha, \xi, \eta) \right] - C(q) +$$

$$\sup_{x \in [-z, \infty]} \left\{ p_\rho(x) \left(x + z g(\alpha, \xi) + \max_{\eta_c \in [0, 1]} \left\{ -K(\eta_c, \eta) + \delta \Delta V_\rho(q, \alpha, \xi, \eta) \right\} \right) \right\},$$

$$= \delta \mathbb{E}_D \left[V_\rho((q - D)^+, \alpha, \xi, \eta) \right] - C(q) +$$

$$\sup_{x \in [-z, \infty]} \left\{ p_\rho(x) \left(x + z g(\alpha, \xi) + \mathbb{E}_D \left[\delta V_\rho((q + 1 - D)^+, \alpha, \xi + 1, \eta) \right] \right) \right\} \tag{B.2}$$

$$- \delta V_\rho((q - D)^+, \alpha, \xi, \eta) + \max_{\eta_c \in [0, 1]} \left\{ -K(\eta_c, \eta) \right.$$

$$\left. + \eta_c \delta \mathbb{E}_D \left[(V_\rho((q + 1 - D)^+, \alpha + 1, \xi + 1, \eta) - V_\rho((q + 1 - D)^+, \alpha, \xi + 1, \eta)) \right] \right\} \Big\}$$

$$= \delta \mathbb{E}_D \left[V_\rho((q - D)^+, \alpha, \xi, \eta) \right] - C(q) + \sup_{x \in [-z, \infty]} \left\{ p_\rho(x) \left(x + z g(\alpha, \xi) + \Delta_q V_\rho(q, \alpha, \xi, \eta) \right) \right.$$

$$\left. + \max_{\eta_c \in [0, 1]} \left\{ -K(\eta_c, \eta) + \eta_c \Delta_\alpha V_\rho(q, \alpha, \xi, \eta) \right\} \right\}, \tag{B.3}$$

Let us assume $K(x)$ is ∞ when $x \notin [0, 1]$ and let K^* denote the convex conjugate of K , then equation B.3 can be simplified as,

$$V_\rho(q, \alpha, \xi, \eta) = \delta \mathbb{E}_D \left[V_\rho((q - D)^+, \alpha, \xi, \eta) \right] - C(q) +$$

$$\sup_{x \in [-z, \infty]} \left\{ p_\rho(x) \left(x + z g(\alpha, \xi) + \Delta_q V_\rho(q, \alpha, \xi, \eta) + K^* (\Delta_\alpha V_\rho(q, \alpha, \xi, \eta)) \right) \right\}. \quad (\text{B.4})$$

B.1.2 Existence of optimal policy

We first have a result indicating that the per stage cost and transitions are bounded. For consistency with [24], we define the per-state cost function

$$c((q, \alpha, \xi, \eta), x) = C(q) - p_\rho(x)(x + z g(\alpha, \xi)),$$

We then have the following result.

Lemma 36. *There exist nonnegative constants c_1 and c_2 , with $1 \leq c_2 < 1/\delta$, and a weight function $\omega \geq 1$ such that for every state (q, α, ξ)*

1. $\sup_{x \in [-z, \infty]} |c((q, \alpha, \xi), x)| \leq \omega(q, \alpha, \xi)$ and
2. $\sup_{x \in [-z, \infty]} \int \omega(s) \mathcal{Q}(ds|q, \alpha, \xi, x) \leq c_2 \omega(s) + c_3$,

where \mathcal{Q} is the transition kernel defined in (4.9), $x \in [-z, \infty]$ is any effective bid, and s represents a three tuple corresponding to state.

Proof. Consider a candidate weight function of the form $\omega(q, \alpha, \xi) = \max\{1, C(q)\}$. Part 1 follows from the following argument.

$$\begin{aligned} c((q, \alpha, \xi), y) &= C(q) - p_\rho(x)(x + z g(\alpha, \xi)) \\ &\leq C(q) \leq \omega(q, \alpha, \xi). \end{aligned}$$

Now, consider $\int \omega(s) \mathcal{Q}(ds|q, \alpha, \xi, x)$. As $\omega(s)$ is increasing in q and the maximum value that q can take for any x is $q + 1$, we have

$$\sup_{x \in [-z, \infty]} \int \omega(s) \mathcal{Q}(ds|q, \alpha, \xi, x) \leq \omega(q + 1, \alpha, \xi)$$

Recall that $C(q)$ is a polynomial function of q of degree p . Then we have

$$\begin{aligned} \frac{\omega(q+1, \alpha, \xi)}{\omega(q, \alpha, \xi)} &= \frac{a_1(q+1)^p + a_2(q+1)^{(p-1)} + \dots}{a_1(q)^p + a_2q^{(p-1)} + \dots} \\ &= \frac{a_1(1+1/q)^p + \frac{a_2}{q}(1+1/q)^{(p-1)} + \dots}{a_1 + \frac{a_2}{q} + \dots} \\ &< \frac{1}{\delta} \text{ for } q \text{ greater than some } q_0. \end{aligned}$$

Let $c_2 = \max_{q>q_0} \frac{\omega(q+1, \alpha, \xi)}{\omega(q, \alpha, \xi)}$ and $c_3 = \max_{q \leq q_0} w(q+1)$. Then we have

$$\sup_{x \in [-z, \infty]} \int w(s) \mathcal{Q}(ds|q, \alpha, \xi, x) \leq c_2 w(s) + c_3$$

with $c_2 < 1/\delta$. □

The next result shows the continuity of the expected weight of transition in the effective bid x .

Lemma 37. *For every state (q, α, ξ) , the function $\omega'(q, \alpha, \xi, x) \triangleq \int \omega(s) \mathcal{Q}(ds|q, \alpha, \xi, x)$ is continuous in $x \in [-z, \infty]$.*

Proof. Using the same candidate weight function of the form $w(q, \alpha, \xi) = \max\{1, C(q)\}$, we have

$$\begin{aligned} \omega'(q, \alpha, \xi, x) &:= \int \omega(s) \mathcal{Q}(ds|q, \alpha, \xi, x) \\ &= \mathbb{E} \left[C((q+1-D)^+) p_\rho(x) + C((q-D)^+) (1 - p_\rho(x)) \right] \end{aligned}$$

Therefore, ω' is continuous with x if p_ρ is continuous with x , which in turn depends on the continuity of $\rho(x)$. As we assumed that $\rho(\cdot)$ is continuous, ω' is continuous with respect to x . □

We need one final result on compactness before we can show the existence of the optimal bid function.

Lemma 38. For every state (q, α, ξ) ,

1. The control-constraint set is compact;
2. The cost per stage is lower semi-continuous in action
3. The function $u'(q, \alpha, \xi, x) := \int u(s) \mathcal{Q}(ds|q, \alpha, \xi, x)$ is continuous in x for every function $u \in \mathcal{V}$.

Proof. We assumed that the support of the mean field distribution of effective bid is a compact set $[z, \mathcal{T}]$ in \mathbb{R}^+ . This means that no agent will choose an effective bid above \mathcal{T} , as the probability of winning is 0. Hence, the action will lie in a compact set $[-z, \mathcal{T}]$. Now, to prove the second part of the theorem, consider the cost in each stage $C(q) - p_\rho(x)(x + z g(\alpha, \xi))$. Expanding the probability of winning, we get $C(q) - (x + z g(\alpha, \xi)) \left(1 - \rho(x)\right)^{M-1}$. Since, we assumed that ρ is continuous, the cost per each stage is continuous with respect to x .

As above, the function depends on x through $p_\rho(x)$ in the third part of the theorem. Hence, continuity follows from the continuity of ρ . \square

B.1.3 Details of Proof of Theorem 12

Since we know that the optimal bid exists, we replace sup with max in (4.11), and the Bellman equation to calculate x becomes

$$\begin{aligned}
& V_{\rho, \eta}(q, \alpha, \xi) \\
&= \delta \mathbb{E}_D \left[V_{\rho, \eta}((q - D)^+, \alpha, \xi) \right] - C(q) + \max_{x \in [-z, \infty)} \left\{ p_\rho(x)(x + z g(\alpha, \xi) + \delta \Delta V_{\rho, \eta}(q, \alpha, \xi)) \right\} \\
&= \delta \mathbb{E}_D \left[V_{\rho, \eta}((q - D)^+, \alpha, \xi) \right] - C(q) + \max_{x \in [-z, \infty)} \left\{ p_\rho(x)(x + \nu(q, \alpha, \xi, \eta)) \right\}, \text{ where,}
\end{aligned}$$

$$\nu(q, \alpha, \xi, \eta) = z g(\alpha, \xi) + \delta \Delta V_{\rho, \eta}(q, \alpha, \xi). \quad (\text{B.5})$$

Note that $\nu(q, \alpha, \xi, \eta)$ essentially determines the “type” of the agent in the mechanism design terminology. Since we have a standard auction format, we expect the agent with the largest value of ν (smallest effective bid) to win the auction. Let

$$U(\nu(q, \alpha, \xi, \eta)) = \max_{x \in [-z, \infty)} \left\{ p_\rho(x)(x + \nu(q, \alpha, \xi, \eta)) \right\} \quad (\text{B.6})$$

With a slight abuse of notation, let $X^*(\nu)$ be the optimal bid function, assumed to be decreasing in ν ; essentially the users generate their new types ν and then participate in the reverse auction. We will see that this assumption indeed holds true once we determine the function. We assume symmetry in that all agents use the same bid function. Now, if x_{opt} is the optimal bid then

$$U(\nu(q, \alpha, \xi, \eta)) = \left\{ (x_{opt} + \nu(q, \alpha, \xi, \eta)) \mathbb{P}^{(M-1)}[X^{*-1}(x_{opt})] \right\} \quad (\text{B.7})$$

$$= \left\{ (x_{opt} + \nu(q, \alpha, \xi, \eta)) \times \mathbb{P}^{(M-1)}[(\nu(Q, A, \Xi, H)) \leq \nu((q, \alpha, \xi, \eta))] \right\}, \quad (\text{B.8})$$

where the decreasing property of $X^*(\cdot)$ is used in the last step. Using the form in (B.7) and $X^*(\cdot)$ being invertible we can transform the bid determination problem by directly incorporating the bid function as

$$U(\nu(q, \alpha, \xi, \eta)) = \max_{x \in [-z, \infty)} \left\{ (x + \nu(q, \alpha, \xi, \eta)) \mathbb{P}^{(M-1)}[X^{*-1}(x)] \right\} \quad (\text{B.9})$$

In the current form we can apply the Envelope Theorem [39], to obtain

$$\begin{aligned} \frac{dU(\nu(q, \alpha, \xi, \eta))}{d(\nu(q, \alpha, \xi, \eta))} &= \mathbb{P}^{M-1}(X^{*-1}(x_{opt})) \\ &= \mathbb{P}^{M-1}(\nu(Q, A, \Xi, H)) \leq \nu((q, \alpha, \xi, \eta)) \end{aligned}$$

The minimum possible effective bid is $x = -z$, under which the probability of winning is 1

but payoff is 0. This can occur at a reputation $\alpha/\xi = 1$. Therefore,

$$U(\nu(q, \alpha, \xi, \eta)) = \int_{\Omega} \mathbb{P}^{M-1}[(\nu(Q, A, \Xi, H) \leq \nu(q', \alpha', \xi', \eta'))] dq' d\alpha' d\xi' d\eta' \quad (\text{B.10})$$

where $\Omega = \{(q', \alpha', \xi', \eta') : \nu(q', \alpha', \xi', \eta') \leq \nu(q, \alpha, \xi, \eta)\}$. Thus, we have

$$\begin{aligned} x_{opt} &= -\nu(q, \alpha, \xi, \eta) + \frac{\int \mathbb{P}^{M-1}[(\nu(Q, A, \Xi, H) \leq \nu(q', \alpha', \xi', \eta'))] dq' d\alpha' d\xi' d\eta'}{\mathbb{P}^{(M-1)}[(\nu(Q, A, \Xi, H) \leq \nu(q, \alpha, \xi, \eta))]} \\ &= -\mathbb{E}[\tilde{\nu} \mid \tilde{\nu} \leq \nu(q, \alpha, \xi, \eta)], \end{aligned}$$

where the last expression follows by a standard integration-by-parts argument when $\mathbb{P}[(\nu(Q, A, \Xi, H) \leq \nu(q, \alpha, \xi, \eta))]$ is absolutely continuous.

B.2 Proofs from Section 4.5

Lemma 39. *If there exists x_f^*, x_g^* such that $f(x_f^*) = \sup_x \{f(x)\}$ and $g(x_g^*) = \sup_x \{g(x)\}$, then the following inequalities holds true,*

1. $\sup\{f(x)\} - \sup\{g(x)\} \leq \sup\{f(x) - g(x)\}$
2. $\sup\{f(x)\} - \sup\{g(x)\} \geq -\sup\{g(x) - f(x)\} = \inf\{g(x) - f(x)\}$
3. $\sup\{f(x)\} - \sup\{g(x)\} \leq f(x_f^*) - g(x_f^*)$, where $x_f^* = \arg \max f(x)$
4. $\sup\{f(x)\} - \sup\{g(x)\} \geq f(x_g^*) - g(x_g^*)$, where $x_g^* = \arg \max g(x)$
5. $-\sup\{g(x) - f(x)\} \leq f(x_g^*) - g(x_g^*) \leq \sup\{f(x)\} - \sup\{g(x)\} \leq f(x_f^*) - g(x_f^*) \leq \sup\{f(x) - g(x)\}$

Proof. we have, $\sup\{f(x) + g(x)\} \leq \sup\{f(x)\} + \sup\{g(x)\}$. Substitue $f - g$ as f , we get

$$\begin{aligned} \sup\{f(x) - g(x) + g(x)\} &\leq \sup\{f(x) - g(x)\} + \sup\{g(x)\} \\ \implies \sup\{f(x)\} - \sup\{g(x)\} &\leq \sup\{f(x) - g(x)\} \end{aligned}$$

Now,

$$\begin{aligned}
& \sup\{f(x)\} - \sup\{g(x)\} \leq \sup\{f(x) - g(x)\} \\
& \implies -\sup\{f(x)\} + \sup\{g(x)\} \geq -\sup\{f(x) - g(x)\} \\
& \implies \sup\{f(x)\} - \sup\{g(x)\} \geq -\sup\{g(x) - f(x)\}
\end{aligned}$$

□

B.2.1 Details of Proof of Lemma 11

Let $q_L < q_H$, by assumption we have $f(q_L) > f(q_H)$. Now,

$$\begin{aligned}
& T_\rho f(q_L, \alpha, \xi, \eta) - T_\rho f(q_H, \alpha, \xi, \eta) = \\
& = \mathbb{E}_D[(f((q_L - D)^+, \alpha, \xi, \eta)\delta] - C(q_L) + \\
& \quad + \sup_x \left\{ p_\rho(x)(x + zg(\alpha, \xi) + \Delta_q f(q_L, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q_L, \alpha, \xi, \eta))) \right\} \\
& - \mathbb{E}_D[(f((q_H - D)^+, \alpha, \xi, \eta)\delta] + C(q_H) \\
& \quad - \sup_x \left\{ p_\rho(x)(x + zg(\alpha, \xi) + \Delta_q f(q_H, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q_H, \alpha, \xi, \eta))) \right\} \\
& \stackrel{a}{\geq} \delta \mathbb{E}_D[(f((q_L - D)^+, \alpha, \xi, \eta))] - \delta \mathbb{E}_D[(f((q_H - D)^+, \alpha, \xi, \eta))] \\
& \quad + \sup_x \left\{ p_\rho(x)(x + zg(\alpha, \xi) + \Delta_q f(q_L, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q_L, \alpha, \xi, \eta))) \right\} \\
& \quad - \sup_x \left\{ p_\rho(x)(x + zg(\alpha, \xi) + \Delta_q f(q_H, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q_H, \alpha, \xi, \eta))) \right\} \\
& \stackrel{b}{\geq} \delta \mathbb{E}_D[(f((q_L - D)^+, \alpha, \xi, \eta))] - \delta \mathbb{E}_D[(f((q_H - D)^+, \alpha, \xi, \eta))] \\
& \quad + p_\rho(x_b)(x_b + zg(\alpha, \xi, \eta) + \Delta_q f(q_L, \alpha, \xi, \eta) - K(\eta_b, \eta) + \eta_b(\Delta_\alpha f(q_H, \alpha, \xi, \eta))) \\
& \quad - p_\rho(x_b)(x_b + zg(\alpha, \xi, \eta) + \Delta_q f(q_H, \alpha, \xi, \eta) - K(\eta_b, \eta) + \eta_b(\Delta_\alpha f(q_H, \alpha, \xi, \eta))) \\
& \stackrel{c}{=} \delta \mathbb{E}_D[(f((q_L - D)^+, \alpha, \xi, \eta))] - \delta \mathbb{E}_D[(f((q_H - D)^+, \alpha, \xi, \eta))] \\
& \quad + p_\rho(x_b) \left(\Delta_q f(q_L, \alpha, \xi, \eta) - \Delta_q f(q_H, \alpha, \xi, \eta) \right. \\
& \quad \quad \left. + \eta_b(\Delta_\alpha f(q_H, \alpha, \xi, \eta)) - \eta_b(\Delta_\alpha f(q_H, \alpha, \xi, \eta)) \right)
\end{aligned}$$

$$\begin{aligned}
&\stackrel{d}{=} \delta \mathbb{E}_D [f((q_L - D)^+, \alpha, \xi, \eta)] - \delta \mathbb{E}_D [f((q_H - D)^+, \alpha, \xi, \eta)] \\
&\quad + p_\rho(x_b) \left(\delta \mathbb{E}_D [(1 - \eta_b) f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta)] \right. \\
&\quad \quad - (1 - \eta_b) f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta) + \\
&\quad \quad - (f((q_L - D)^+, \alpha, \xi, \eta) - f((q_H - D)^+, \alpha, \xi, \eta)) + \\
&\quad \quad \left. + \eta_b f((q_L + 1 - D)^+, \alpha + 1, \xi + 1, \eta) - \eta_b f((q_H + 1 - D)^+, \alpha + 1, \xi + 1, \eta) \right) \\
&\stackrel{e}{=} \delta (1 - p_\rho(x_b)) \mathbb{E}_D [f((q_L - D)^+, \alpha, \xi, \eta) - f((q_H - D)^+, \alpha, \xi, \eta)] \\
&\quad + \delta p_\rho(x_b) (1 - \eta_b) \mathbb{E}_D [f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta) \\
&\quad \quad - f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta)] \\
&\quad + \delta p_\rho(x_b) \eta_b \mathbb{E}_D [f((q_L + 1 - D)^+, \alpha + 1, \xi + 1, \eta) \\
&\quad \quad - f((q_H + 1 - D)^+, \alpha + 1, \xi + 1, \eta)] \\
&\stackrel{f}{>} 0,
\end{aligned}$$

where, (a) follows from the assumption that $C(q)$ is a strictly convex increasing function in q , (b) follows from Lemma 39.4. Equalities (c), (d) and (e) are just rearranging and expanding terms and (f) follows from our assumption that f is decreasing in q . Next, we prove the continuity. Let us take a sequence $q_n \rightarrow q$. By assumption, f is continuous, and therefore $f(q_n) \rightarrow f(q)$. As f is a decreasing function of queue, $f((q - D)^+, \alpha, \xi)$ is dominated by $f(0, \alpha, \xi)$, which is integrable. Hence by the Dominated Convergence theorem $\mathbb{E}_D(f((q_n - D)^+, \alpha, \xi)) \rightarrow \mathbb{E}_D(f((q - D)^+, \alpha, \xi))$. Therefore, $T_\rho^n f \rightarrow V_\rho^*$ is also continuous.

B.2.2 Details of Proof of Lemma 12

Let $\alpha_L < \alpha_H$, then $f(q, \alpha_H, \xi, \eta) > f(q, \alpha_L, \xi, \eta)$. We will prove that $T_\rho f(q, \alpha_H, \xi, \eta) > T_\rho f(q, \alpha_L, \xi, \eta)$.

$$\begin{aligned}
&T_\rho f(q, \alpha_H, \xi, \eta) - T_\rho f(q, \alpha_L, \xi, \eta) = \\
&= \mathbb{E}_D [f((q - D)^+, \alpha_H, \xi, \eta) \delta] - C(q) + \sup_x \left\{ p_\rho(x) (x + z g(\alpha_H, \xi)) \right.
\end{aligned}$$

$$\begin{aligned}
& + \Delta_q f(q, \alpha_H, \xi, \eta) + K^*(\Delta_\alpha f(q_L, \alpha_H, \xi, \eta)) \Big\} \\
& - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)\delta] + C(q) - \sup_x \left\{ p_\rho(x)(x + zg(\alpha_L, \xi) \right. \\
& \quad \left. + \Delta_q f(q, \alpha_L, \xi, \eta) + K^*(\Delta_\alpha f(q, \alpha_L, \xi, \eta))) \right\} \\
& = \mathbb{E}_D [(f((q-D)^+, \alpha_H, \xi, \eta)\delta] - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)\delta] \\
& \quad + \sup_x \left\{ p_\rho(x)(x + zg(\alpha_H, \xi) + \Delta_q f(q, \alpha_H, \xi, \eta) + K^*(\Delta_\alpha f(q_L, \alpha_H, \xi, \eta))) \right\} \\
& \quad - \sup_x \left\{ p_\rho(x)(x + zg(\alpha_L, \xi) + \Delta_q f(q, \alpha_L, \xi, \eta) + K^*(\Delta_\alpha f(q, \alpha_L, \xi, \eta))) \right\} \\
& \stackrel{a}{\geq} \mathbb{E}_D [(f((q-D)^+, \alpha_H, \xi, \eta)\delta] - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)\delta] \\
& \quad + \left\{ p_\rho(x_L^*)(x_L^* + zg(\alpha_H, \xi) + \Delta_q f(q, \alpha_H, \xi, \eta) + K^*(\Delta_\alpha f(q_L, \alpha_H, \xi, \eta))) \right\} \\
& \quad - \left\{ p_\rho(x_L^*)(x_L^* + zg(\alpha_L, \xi) + \Delta_q f(q, \alpha_L, \xi, \eta) + K^*(\Delta_\alpha f(q, \alpha_L, \xi, \eta))) \right\} \\
& = \mathbb{E}_D [(f((q-D)^+, \alpha_H, \xi, \eta)\delta] - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)\delta] \\
& \quad + p_\rho(x_L^*) [zg(\alpha_H, \xi) - zg(\alpha_L, \xi)] + p_\rho(x_L^*) (\Delta_q f(q, \alpha_H, \xi, \eta) \\
& \quad + K^*(\Delta_\alpha f(q_L, \alpha_H, \xi, \eta)) - \Delta_q f(q, \alpha_L, \xi, \eta) - K^*(\Delta_\alpha f(q, \alpha_L, \xi, \eta))) \\
& \stackrel{b}{\geq} \mathbb{E}_D [(f((q-D)^+, \alpha_H, \xi, \eta)\delta] - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)\delta] \\
& \quad + p_\rho(x_L^*) \left(\Delta_q f(q, \alpha_H, \xi, \eta) - \Delta_q f(q, \alpha_L, \xi, \eta) \right. \\
& \quad \quad \left. + \eta_L^* (\Delta_\alpha f(q_L, \alpha_H, \xi, \eta) - \Delta_\alpha f(q, \alpha_L, \xi, \eta)) \right) \\
& \stackrel{d}{=} \mathbb{E}_D [(f((q-D)^+, \alpha_H, \xi, \eta)\delta] - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)\delta] \\
& \quad + \delta p_\rho(x_L^*) \left(\eta_L^* (\mathbb{E}_D f((q+1-D)^+, \alpha_H+1, \xi+1) - \mathbb{E}_D f((q+1-D)^+, \alpha_L+1, \xi+1)) \right. \\
& \quad + (1 - \eta_L^*) (\mathbb{E}_D f((q+1-D)^+, \alpha_H, \xi+1) - \mathbb{E}_D f((q+1-D)^+, \alpha_L, \xi+1)) \\
& \quad \left. - (\mathbb{E}_D f((q-D)^+, \alpha_H, \xi) - \mathbb{E}_D f((q-D)^+, \alpha_L, \xi)) \right) \\
& \stackrel{e}{=} \delta(1 - p_\rho(x_L^*)) \mathbb{E}_D [(f((q-D)^+, \alpha_H, \xi, \eta)] - \mathbb{E}_D [(f((q-D)^+, \alpha_L, \xi, \eta)] \\
& \quad + \delta p_\rho(x_L^*) \left(\eta_L^* (\mathbb{E}_D f((q+1-D)^+, \alpha_H+1, \xi+1) - \mathbb{E}_D f((q+1-D)^+, \alpha_L+1, \xi+1)) \right. \\
& \quad \left. + (1 - \eta_L^*) (\mathbb{E}_D f((q+1-D)^+, \alpha_H, \xi+1) - \mathbb{E}_D f((q+1-D)^+, \alpha_L, \xi+1)) \right) \\
& \stackrel{f}{\geq} 0,
\end{aligned}$$

where, (a) and (b) follows from Lemma 39.4. Inequality (c) follows from monotnocity of $g(\alpha, \xi)$ with α . Equalities (d) and (e) are just rearranging and expanding terms and (f) follows from our assumption that f is increasing in α .

B.2.3 Details of Proof of Lemma 13

Let us assume if $q_L < q_H$ then, $\nu_f(q_H, \alpha, \xi, \eta) < \nu_f(q_L, \alpha, \xi, \eta)$. Now, consider

$$\begin{aligned}
& \nu_{T_\rho f}(q_H, \alpha, \xi, \eta) - \nu_{T_\rho f}(q_L, \alpha, \xi, \eta) \\
&= \Delta_q T_\rho f(q_H, \alpha, \xi, \eta) + \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \Delta_\alpha T_\rho f(q_H, \alpha, \xi, \eta) \} + \\
&\quad - \Delta_q T_\rho f(q_L, \alpha, \xi, \eta) - \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \Delta_\alpha T_\rho f(q_L, \alpha, \xi, \eta) \} \\
&= \delta \mathbb{E}_D [T_\rho f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta) - T_\rho f((q_H - D)^+, \alpha, \xi, \eta)] + \\
&\quad + \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D [T_\rho f((q_H + 1 - D)^+, \alpha + 1, \xi + 1, \eta) \\
&\quad\quad - T_\rho f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta)] \} \\
&\quad - \delta \mathbb{E}_D [T_\rho f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta) - T_\rho f((q_L - D)^+, \alpha, \xi, \eta)] \\
&\quad - \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D [T_\rho f((q_L + 1 - D)^+, \alpha + 1, \xi + 1, \eta) \\
&\quad\quad - T_\rho f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta)] \} \\
&\stackrel{a}{=} \delta \mathbb{E}_D [\Delta_q f((q_H - D)^+, \alpha, \xi, \eta) - \Delta_q f((q_L - D)^+, \alpha, \xi, \eta)] \\
&\quad + \delta \mathbb{E}_D [(C(q_L + 1 - D)^+ - C(q_L - D)^+) - (C(q_H + 1 - D)^+ - C(q_H - D)^+)] \\
&\quad + \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x) \left(x + \nu_f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta) \right) \right\} \right. \\
&\quad\quad \left. - \sup_x \left\{ p_\rho(x) \left(x + \nu_f((q_H - D)^+, \alpha, \xi, \eta) \right) \right\} \right] \\
&\quad - \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x) \left(x + \nu_f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta) \right) \right\} \right. \\
&\quad\quad \left. - \sup_x \left\{ p_\rho(x) \left(x + \nu_f((q_L - D)^+, \alpha, \xi, \eta) \right) \right\} \right] \\
&\quad + \quad \text{sup terms} \\
&\stackrel{b}{\leq} \delta \mathbb{E}_D [\Delta_q f((q_H - D)^+, \alpha, \xi, \eta) - \Delta_q f((q_L - D)^+, \alpha, \xi, \eta)] + \Delta C(q_L) - \Delta C(q_H)
\end{aligned}$$

$$\begin{aligned}
& + \delta \mathbb{E}_D \left[p_\rho(x_1^*) \left(\nu_f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta) - \nu_f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta) \right) \right] \\
& - \delta \mathbb{E}_D \left[p_\rho(x_1^*) \left(\nu_f((q_H - D)^+, \alpha, \xi, \eta) - \nu_f((q_L - D)^+, \alpha, \xi, \eta) \right) \right] \\
& + \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D \left[\Delta_\alpha f((q_H - D)^+, \alpha, \xi) \right. \right. \\
& \quad \left. \left. + \sup_x \left\{ p_\rho(x) (x + \nu_f((q_H + 1 - D)^+, \alpha + 1, \xi + 1, \eta)) \right\} \right. \right. \\
& \quad \left. \left. - \sup_x \left\{ p_\rho(x) (x + \nu_f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta)) \right\} \right] \right\} \\
& - \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D \left[\Delta_\alpha f((q_L - D)^+, \alpha, \xi) \right. \right. \\
& \quad \left. \left. + \sup_x \left\{ p_\rho(x) (x + \nu_f((q_L + 1 - D)^+, \alpha + 1, \xi + 1, \eta)) \right\} \right. \right. \\
& \quad \left. \left. - \sup_x \left\{ p_\rho(x) (x + \nu_f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta)) \right\} \right] \right\} \\
& \stackrel{c}{\leq} \delta \mathbb{E}_D \left[\Delta_q f((q_H - D)^+, \alpha, \xi, \eta) - \Delta_q f((q_L - D)^+, \alpha, \xi, \eta) \right] + \Delta C(q_L) - \Delta C(q_H) \\
& + \delta \mathbb{E}_D \left[p_\rho(x_1^*) \left(\nu_f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta) - \nu_f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta) \right) \right] \\
& - \delta \mathbb{E}_D \left[p_\rho(x_2^*) \left(\nu_f((q_H - D)^+, \alpha, \xi, \eta) - \nu_f((q_L - D)^+, \alpha, \xi, \eta) \right) \right] \\
& + \eta_c^* \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x) (x + \nu_f((q_H + 1 - D)^+, \alpha + 1, \xi + 1, \eta)) \right\} \right. \\
& \quad \left. - \sup_x \left\{ p_\rho(x) (x + \nu_f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta)) \right\} \right. \\
& \quad \left. - \sup_x \left\{ p_\rho(x) (x + \nu_f((q_L + 1 - D)^+, \alpha + 1, \xi + 1, \eta)) \right\} \right. \\
& \quad \left. + \sup_x \left\{ p_\rho(x) (x + \nu_f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta)) \right\} \right] \\
& + \delta \mathbb{E}_D \left[\eta_c^* \Delta_\alpha f((q_H - D)^+, \alpha, \xi) \right] - \eta_c^* \Delta_\alpha f((q_L - D)^+, \alpha, \xi) \\
& \stackrel{d}{\leq} \delta (1 - p_\rho(x_2^*)) \mathbb{E}_D \left[\nu_f((q_H - D)^+, \alpha, \xi, \eta) - \nu_f((q_L - D)^+, \alpha, \xi, \eta) \right] \\
& + \Delta C(q_L) - \Delta C(q_H) \\
& + \delta (1 - \eta_c^*) p_\rho(x_1^*) \mathbb{E}_D \left[\left(\nu_f((q_H + 1 - D)^+, \alpha, \xi + 1, \eta) \right. \right. \\
& \quad \left. \left. - \nu_f((q_L + 1 - D)^+, \alpha, \xi + 1, \eta) \right) \right]
\end{aligned}$$

$$\begin{aligned}
& + \eta_d^* p_\rho(x_3^*) \delta \mathbb{E}_D \left[\left(\nu_f((q_H + 1 - D)^+, \alpha + 1, \xi + 1, \eta) \right. \right. \\
& \qquad \qquad \qquad \left. \left. - \nu_f((q_L + 1 - D)^+, \alpha + 1, \xi + 1, \eta) \right) \right] \\
& \stackrel{e}{\leq} \Delta C(q_L) - \Delta C(q_H) \tag{B.11}
\end{aligned}$$

$$\stackrel{f}{\leq} 0, \tag{B.12}$$

where, (b), (c) and (d) follows from Lemma 39.4 and (e) follows from our assumption that ν_f is decreasing in q . Inequality (f) follows from the convexity of the cost function. As $T_\rho f \rightarrow V_\rho^*$, $\nu_{V_{\rho^*}}(q, \alpha, \xi)$ strictly decreases with queue.

Now the second part can be proved using same method. We know that $V_\rho^* = T_\rho V_\rho^*$, therefore

$$\nu_{V_\rho^*}(q_H, \alpha, \xi, \eta) - \nu_{V_\rho^*}(q_L, \alpha, \xi, \eta) \tag{B.13}$$

$$= \nu_{T_\rho V_\rho^*}(q_H, \alpha, \xi, \eta) - \nu_{T_\rho V_\rho^*}(q_L, \alpha, \xi, \eta) \tag{B.14}$$

$$\stackrel{a}{\leq} \Delta C(q_L) - \Delta C(q_H), \tag{B.15}$$

where, inequality (a) can be obtained from (B.11) replacing f with V_ρ^* .

B.2.4 Details of Proof of Lemma 15

We will again show the preservation of the property through the Bellman function. We assume $\nu_f(q, \alpha, \xi, \eta) = zg(\alpha, \xi) + \Delta_q f(q, \alpha, \xi, \eta) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))$ is increasing with α and then show $\nu_{T_\rho f}(q, \alpha, \xi, \eta)$ also increases with α . Let us assume if $\alpha_L < \alpha_H$ then, $\nu_f(q, \alpha_L, \xi, \eta) < \nu_f(q, \alpha_H, \xi, \eta)$. Now,

$$\begin{aligned}
& \nu_{T_\rho f}(q, \alpha_L, \xi, \eta) - \nu_{T_\rho f}(q, \alpha_H, \xi, \eta) \\
& = zg(\alpha_L, \xi) + \Delta_q T_\rho f(q, \alpha_L, \xi, \eta) + \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \Delta_\alpha T_\rho f(q, \alpha_L, \xi, \eta) \right\}
\end{aligned}$$

$$\begin{aligned}
& -zg(\alpha_H, \xi) - \Delta_q T_\rho f(q, \alpha_H, \xi, \eta) - \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \Delta_\alpha T_\rho f(q, \alpha_H, \xi, \eta) \} \\
= & zg(\alpha_L, \xi) - zg(\alpha_H, \xi) + \delta \mathbb{E}_D [T_\rho f((q+1-D)^+, \alpha_L, \xi+1, \eta) \\
& \quad - T_\rho f((q-D)^+, \alpha_L, \xi, \eta)] + \\
& + \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D [T_\rho f((q+1-D)^+, \alpha_L+1, \xi+1, \eta) \\
& \quad - T_\rho f((q+1-D)^+, \alpha_L, \xi+1, \eta)] \} \\
& - \delta \mathbb{E}_D [T_\rho f((q+1-D)^+, \alpha_H, \xi+1, \eta) - T_\rho f((q-D)^+, \alpha_H, \xi, \eta)] \\
& - \sup_{\eta_c} \{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D [T_\rho f((q+1-D)^+, \alpha_H+1, \xi+1, \eta) \\
& \quad - T_\rho f((q+1-D)^+, \alpha_H, \xi+1, \eta)] \} \\
\stackrel{a}{=} & zg(\alpha_L, \xi) - zg(\alpha_H, \xi) + \delta \mathbb{E}_D [\delta \mathbb{E}_{D'} f((q+1-D-D')^+, \alpha_L, \xi+1, \eta) \\
& \quad - \delta \mathbb{E}_{D'} f((q-D-D')^+, \alpha_L, \xi, \eta)] + \\
& - \delta \mathbb{E}_D [\delta \mathbb{E}_{D'} f((q+1-D-D')^+, \alpha_H, \xi+1, \eta) + \delta \mathbb{E}_{D'} f((q-D-D')^+, \alpha_H, \xi, \eta)] + \\
& + \delta \mathbb{E}_D [(C(q+1-D)^+ - C(q-D)^+) - (C(q+1-D)^+ - C(q-D)^+)] \\
& + \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x) (x + \nu_f((q+1-D)^+, \alpha_L, \xi+1, \eta)) \right. \right. \\
& \quad \left. \left. - \sup_x \left\{ p_\rho(x) (x + \nu_f((q-D)^+, \alpha_L, \xi, \eta)) \right\} \right\} \right] \\
& - \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x) (x + \nu_f((q+1-D)^+, \alpha_H, \xi+1, \eta)) \right. \right. \\
& \quad \left. \left. - \sup_x \left\{ p_\rho(x) (x + \nu_f((q-D)^+, \alpha_H, \xi, \eta)) \right\} \right\} \right] \\
& + \text{ other terms.} \\
\stackrel{b}{\leq} & zg(\alpha_L, \xi) - zg(\alpha_H, \xi) + \delta \mathbb{E}_D [\Delta_q f((q-D)^+, \alpha_L, \xi, \eta) - f((q-D)^+, \alpha_H, \xi, \eta)] + \\
& + \delta \mathbb{E}_D [p_\rho(x_1^*) (\nu((q+1-D)^+, \alpha_L, \xi+1, \eta)) - (\nu((q+1-D)^+, \alpha_H, \xi+1, \eta))] \\
& \quad - \delta \mathbb{E}_D [p_\rho(x_2^*) (\nu((q-D)^+, \alpha_L, \xi, \eta) - \nu((q-D)^+, \alpha_H, \xi, \eta))] \\
& + \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D \left[\delta \mathbb{E}_{D'} [f((q+1-D-D')^+, \alpha_L+1, \xi+1, \eta)] \right. \right. \\
& \quad \left. \left. - \delta \mathbb{E}_{D'} [f((q+1-D-D')^+, \alpha_L, \xi+1, \eta)] \right\}
\end{aligned}$$

$$\begin{aligned}
& + \sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_L + 1, \xi + 1, \eta)) \right\} \\
& \quad - \sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_L, \xi + 1, \eta)) \right\} \Bigg\} \\
& - \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D \left[\delta \mathbb{E}_{D'} [f((q+1-D-D')^+, \alpha_H + 1, \xi + 1, \eta)] \right. \right. \\
& \quad \left. \left. - \delta \mathbb{E}_{D'} [f((q+1-D-D')^+, \alpha_H, \xi + 1, \eta)] \right] \right. \\
& \quad \left. + \sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_H + 1, \xi + 1, \eta)) \right\} \right. \\
& \quad \left. - \sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_H, \xi + 1, \eta)) \right\} \right\} \Bigg\} \\
& \stackrel{c}{\leq} zg(\alpha_L, \xi) - zg(\alpha_H, \xi) + \delta \mathbb{E}_D [\Delta_q f((q-D)^+, \alpha_L, \xi, \eta) - f((q-D)^+, \alpha_H, \xi, \eta)] + \\
& \quad + \delta \mathbb{E}_D [p_\rho(x_1^*)(\nu((q+1-D)^+, \alpha_L, \xi + 1, \eta)) - (\nu((q+1-D)^+, \alpha_H, \xi + 1, \eta))] \\
& \quad - \delta \mathbb{E}_D [p_\rho(x_2^*)(\nu((q-D)^+, \alpha_L, \xi, \eta) - \nu((q-D)^+, \alpha_H, \xi, \eta))] \\
& \quad + \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D [\Delta_\alpha f((q-D)^+, \alpha_L, \xi, \eta)] \right\} \\
& \quad + \eta_c^* \left\{ \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_L + 1, \xi + 1, \eta)) \right\} \right. \right. \\
& \quad \quad \left. \left. - \sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_L, \xi + 1, \eta)) \right\} \right] \right\} \\
& \quad - \sup_{\eta_c} \left\{ -K(\eta_c) + \eta_c \delta \mathbb{E}_D [\Delta_\alpha f((q-D)^+, \alpha_H, \xi, \eta)] \right\} \\
& \quad - \eta_c^* \left\{ \delta \mathbb{E}_D \left[\sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_H + 1, \xi + 1, \eta)) \right\} \right. \right. \\
& \quad \quad \left. \left. - \sup_x \left\{ p_\rho(x)(x + \nu_f((q+1-D)^+, \alpha_H, \xi + 1, \eta)) \right\} \right] \right\} \\
& \stackrel{d}{\leq} \delta \mathbb{E}_D [\nu_f((q-D)^+, \alpha_L, \xi, \eta) - \nu_f((q-D)^+, \alpha_H, \xi, \eta)] + \\
& \quad + \delta \mathbb{E}_D [p_\rho(x_1^*)(1 - \eta_c^*)(\nu_f((q+1-D)^+, \alpha_L, \xi + 1, \eta)) - \nu_f((q-D)^+, \alpha_L, \xi, \eta)] \\
& \quad - \delta \mathbb{E}_D [p_\rho(x_2^*)(\nu_f((q-D)^+, \alpha_L, \xi, \eta) - \nu_f((q-D)^+, \alpha_H, \xi, \eta))] \\
& \quad + \delta \mathbb{E}_D [\eta_c^* p_\rho(x_3^*)(\nu_f((q+1-D)^+, \alpha_L + 1, \xi + 1, \eta) \\
& \quad \quad - \nu_f((q+1-D)^+, \alpha_H + 1, \xi + 1, \eta))] \\
& = \delta \mathbb{E}_D [(1 - p_\rho(x_2^*))\nu_f((q-D)^+, \alpha_L, \xi, \eta) - \nu_f((q-D)^+, \alpha_H, \xi, \eta)] + \\
& \quad + p_\rho(x_1^*)(1 - \eta_c^*)\delta \mathbb{E}_D [(\nu_f((q+1-D)^+, \alpha_L, \xi + 1, \eta))
\end{aligned}$$

$$\begin{aligned}
& - (\nu_f((q+1-D)^+, \alpha_H, \xi+1, \eta))] \\
& + \eta_c^* p_\rho(x_3^*) \delta \mathbb{E}_D [(\nu_f((q+1-D)^+, \alpha_L+1, \xi+1, \eta) \\
& \quad - \nu_f((q+1-D)^+, \alpha_H+1, \xi+1, \eta))] \\
& \leq 0
\end{aligned}$$

B.3 Proofs from Section 4.6

B.3.1 Details of Proof of Lemma 18

For the Berg's maximum theorem to hold we need the continuity of the objective function $h(\rho, x)$ and compactness of the correspondence $\mathcal{R}(\rho)$. If we prove V_ρ^* is continuous with ρ then joint continuity follows from the assumption that $\rho(x)$ is a continuous function of x . To prove continuity of V_ρ^* , consider

$$\begin{aligned}
\|V_{\rho_1}^* - V_{\rho_2}^*\|_w &= \|T_{\rho_1} V_{\rho_1}^* - T_{\rho_2} V_{\rho_2}^*\|_w \\
&= \|T_{\rho_1} V_{\rho_1}^* - T_{\rho_1} V_{\rho_2}^* + T_{\rho_1} V_{\rho_2}^* - T_{\rho_2} V_{\rho_2}^*\|_w \\
&\leq \|T_{\rho_1} V_{\rho_1}^* - T_{\rho_1} V_{\rho_2}^*\|_w + \|T_{\rho_1} V_{\rho_2}^* - T_{\rho_2} V_{\rho_2}^*\|_w \\
&\leq \alpha \|V_{\rho_1}^* - V_{\rho_2}^*\|_w + \|T_{\rho_1} V_{\rho_2}^* - T_{\rho_2} V_{\rho_2}^*\|_w \\
\implies (1 - \alpha) \|V_{\rho_1}^* - V_{\rho_2}^*\|_w &\leq \|T_{\rho_1} V_{\rho_2}^* - T_{\rho_2} V_{\rho_2}^*\|_w.
\end{aligned}$$

Now,

$$\begin{aligned}
& \|T_{\rho_1} f - T_{\rho_2} f\|_w \\
&= \left| \sup_{q, \alpha, \xi, \eta} \left\{ \delta \mathbb{E}_D(f(q-D)^+, \alpha, \xi, \eta) - C(q) + \sup_x \left\{ p_{\rho_1}(x)(x + zg(\alpha, \xi) \right. \right. \right. \\
& \quad \left. \left. \left. + \Delta_q f(q, \alpha, \xi) \right\} + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta)) \right. \right. \\
& \quad \left. \left. - \delta \mathbb{E}_D(f(q-D)^+, \alpha, \xi, \eta) + C(q) - \sup_x \left\{ p_{\rho_2}(x)(x + zg(\alpha, \xi) \right. \right. \right.
\end{aligned}$$

$$\begin{aligned}
& \left. + \Delta_q f(q, \alpha, \xi) \right\} + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta)) \Big\} \Big| \\
\leq & \left| \sup_{q, \alpha, \xi, \eta} \left\{ p_{\rho_1}(x^*)(x^* + zg(\alpha, \xi) + \Delta_q f(q, \alpha, \xi) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))) \right. \right. \\
& \left. \left. - p_{\rho_2}(x^*)(x^* + zg(\alpha, \xi) + \Delta_q f(q, \alpha, \xi) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))) \right\} \right| \\
= & \left| \sup_{q, \alpha, \xi, \eta} \left\{ (p_{\rho_1}(x^*) - p_{\rho_2}(x^*))(x^* + zg(\alpha, \xi) + \Delta_q f(q, \alpha, \xi) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))) \right\} \right| \\
= & \left| \sup_{q, \alpha, \xi, \eta} \left\{ ((1 - \rho_1(x^*))^{(M-1)} - (1 - \rho_2(x^*))^{(M-1)})(x^* + zg(\alpha, \xi) \right. \right. \\
& \left. \left. + \Delta_q f(q, \alpha, \xi) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))) \right\} \right| \\
\leq & \left| \sup_{q, \alpha, \xi, \eta} \left\{ \left((\rho_1(x^*)) - (\rho_2(x^*)) * (M - 1) \right) (x^* + zg(\alpha, \xi) \right. \right. \\
& \left. \left. + \Delta_q f(q, \alpha, \xi) + K^*(\Delta_\alpha f(q, \alpha, \xi, \eta))) \right\} \right| \\
\leq & K \|\rho_1 - \rho_2\| \\
\implies & \|V^*(\rho_1) - V^*(\rho_2)\|_w \leq K \|\rho_1 - \rho_2\|
\end{aligned}$$

Hence, V_ρ^* is continuous with ρ . $\mathcal{R}(\rho)$ as such is not compact, but, we will prove that we can restrict the correspondence to a compact set. The proof is exactly same as the proof of Lemma 38. $\mathcal{R}(\rho) = [0, \mathcal{M}_{\rho, q, \alpha, \xi}]$ is the compact correspondence. Since the conditions of Berge's Maximum theorem are satisfied θ_ρ is continuous in ρ .

B.3.2 Details of Proof of Lemma 19

The queue length dynamics are given by

$$q = \begin{cases} (Q' - D)^+ & \text{with probability } \delta \\ R & \text{with probability } (1 - \delta), \end{cases} \quad (\text{B.16})$$

where, $D \sim \Phi$ and $R \sim \Psi$ and

$$\mathbb{P}(Q' = q | q, \alpha, \xi) = 1 - p_\rho(\hat{\theta}(q, \alpha, \xi)) \quad (\text{B.17})$$

$$\mathbb{P}(Q' = (q + 1)|q, \alpha, \xi) = p_\rho(\hat{\theta}(q, \alpha, \xi)) \quad (\text{B.18})$$

Let Π' be the distribution of Q' . Then the distribution of queue length Π_q is the convolution of Π' and $\bar{\Phi}$, where $\bar{\Phi}$ is the distribution of $-D$.

$$\Pi_q(B, \alpha', \xi'|q, \alpha, \xi) = \delta f(\alpha', \xi') \int_{-\infty}^{\infty} \Phi(B - y) \Pi'(y|q, \alpha, \xi) + (1 - \delta) \Psi(B, \alpha', \xi')$$

If B is a Borel null set then so will $B - y$ and hence $\Pi_q(B) = 0$.

B.3.3 Details of Proof of Lemma 21

The density of invariant distribution Π_ρ is computed as follows,

$$\Pi_\rho(B_q, B_\alpha, B_\xi, B_\eta) = \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_\Psi \Gamma^k(B_q, B_\alpha, B_\xi, B_\eta | (q, \alpha, \xi, \eta)) \quad (\text{B.19})$$

$$= \sum_{k \geq 0} \delta^k (1 - \delta) \int_{\mathbb{R}^+, \mathbb{N}, \mathbb{N}, [0,1]} \Gamma^k(B_q, B_\alpha, B_\xi, B_\eta | (q, \alpha, \xi, \eta)) \psi(q, \alpha, \xi, \eta) dq d\alpha d\xi d\eta \quad (\text{B.20})$$

where, Γ^k represent the process without regeneration. Let $a \in L_k$ represents complete history till time k . This includes at each time slot, the amount of queue serviced, if the user had won the auction, if it won did he get a positive reward. Now,

$$\begin{aligned} & \mathbb{E}_\Psi [\Gamma^k(B_q, B_\alpha, B_\xi, B_\eta | (q, \alpha, \xi, \eta))] \\ &= \mathbb{E}_\Psi \left[\sum_{a \in L_k} \Gamma^k(a_k = a, \eta_k \in B_\eta | (q, \alpha, \xi, \eta)) \mathbf{I} \{F_k((q, \alpha, \xi), a) \in B_q \times B_\alpha \times B_\xi\} \right] \\ &= \sum_{a \in L_k} \mathbb{E}_\Psi [\Gamma^k(a_k = a, \eta_k \in B_\eta | (q, \alpha, \xi, \eta)) \mathbf{I} \{F_k((q, \alpha, \xi), a) \in B_q \times B_\alpha \times B_\xi\}] \\ &= \sum_{a \in L_k} \int_{\mathbb{R}^+, \mathbb{N}, \mathbb{N}, [0,1]} [\psi(q, \alpha, \xi, \eta) \Gamma^k(a_k = a, \eta_k \in B_\eta | (q, \alpha, \xi, \eta)) \\ & \quad \mathbf{I} \{F_k((q, \alpha, \xi), a) \in B_q \times B_\alpha \times B_\xi\}] dq d\alpha d\xi d\eta \end{aligned}$$

As η_k doesn't change before regeneration, $\eta \in B_\eta$. Let $u = F_k((q, \alpha, \xi), a)$ and $(q, \alpha, \xi) = G(u, a)$

$$\begin{aligned}
&= \sum_{a \in L_k} \int_{\mathbb{R}^+, \mathbb{N}, \mathbb{N}, [0,1]} \psi(G_k(u, a), \eta) [\Gamma^k(a_k = a | (G_k(u, a), \eta))] \\
&\quad \mathbf{I}\{\eta \in B_\eta; u \in B_q \times B_\alpha \times B_\xi\} J_k(u, a) d\eta du \\
&= \sum_{a \in L_k} \int_{B_q \times B_\alpha \times B_\xi \times B_\eta} \psi(G_k(u, a), \eta) [\Gamma^k(a_k = a | (G_k(u, a), \eta))] J_k(u, a) d\eta du \\
&= \int_{B_q \times B_\alpha \times B_\xi \times B_\eta} \sum_{a \in L_k} \psi(G_k(u, a), \eta) [\Gamma^k(a_k = a | (G_k(u, a), \eta))] J_k(u, a) d\eta du
\end{aligned}$$

Therefore the density is given by

$$\begin{aligned}
&\pi_\rho(q, \alpha, \xi, \eta) \\
&= \sum_{k \geq 1} \delta^k (1 - \delta) \sum_{a \in L_k} \psi(G_k((q, \alpha, \xi), a), \eta) [\Gamma^k(a_k = a | (G_k((q, \alpha, \xi), a), \eta))] J_k((q, \alpha, \xi), a)
\end{aligned}$$

B.3.4 Details of Proof of Lemma 22

By Portmanteau theorem, it is enough to show that for any open set (B, α, ξ, B') , $\liminf_{n \rightarrow \infty} \Pi_{\rho_n}(B, \alpha, \xi, B') \geq \Pi_\rho(B, \alpha, \xi)$. We have

$$\begin{aligned}
\liminf_{n \rightarrow \infty} \Pi_{\rho_n}(B, \alpha, \xi, B') &= \liminf_{n \rightarrow \infty} \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_\Psi \Gamma_{\rho_n}^k((B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0)) \\
&\geq \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_\Psi \left[\liminf_{n \rightarrow \infty} \Gamma_{\rho_n}^k((B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0)) \right] \\
&\geq \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_\Psi \left[\Gamma_\rho^k((B, \alpha, \xi, B') | (q_0, \alpha_0, \xi_0, \eta_0)) \right] \\
&= \Pi_\rho(B, \alpha, \xi)
\end{aligned}$$

B.3.5 Details of Proof of Lemma 23

First we will prove that $\gamma(x)$ is continuous. From 14 the optimal bid is strictly increasing in queue size for a given α and ψ . Therefore, $\theta_{\rho,z}^{-1}[-z, x]$ will at most have a countable number of states whose measure is 0. Hence, $\gamma(x)$ will not have any jumps in this region and is hence continuous.

Now we will show that γ has a finite mean. To prove this we need to prove that expected bid under ρ is bounded above by a constant that is independent of ρ . We define a new Markov random process \tilde{Q}_k with the probability transition matrix

$$\begin{aligned} \mathbb{P}(\tilde{Q}_{k+1} \in B, \alpha_{k+1} = \alpha, \xi_{k+1} = \xi, \eta_{k+1} \in B' | q_k = q_0, \alpha_k = \alpha_0, \xi_k = \xi_0, X_k = x, \eta_k = \eta_0) \\ = \delta \mathbb{I}_{(q_0+1) \in B, \alpha = \xi_0+1, \xi = \xi_0+1} + (1 - \delta) \Psi(B, \alpha, \xi, B'). \end{aligned}$$

This process will be equal to previous process if the user always wins, with out any departures and has the highest reputation. Let $\tilde{\Pi}$ be the steady state distribution. The steady state can be written as

$$\tilde{\Pi}(B, \alpha, \xi, \eta) = \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_{\Psi} \left[\Gamma_{\rho, \theta}^k ((B, \alpha, \xi, B') | (State)) \right] \quad (\text{B.21})$$

$$= \sum_{k \geq 0} \delta^k (1 - \delta) \mathbb{E}_{\Psi} \left[(\mathbb{I}_{(q+k) \in B, \alpha = \xi+k, \xi = \xi+k} | State) \right] \quad (\text{B.22})$$

Since this processes has always higher queue length and reputation values than the original process it can be shown that this stochastically dominates the original process.

Now consider

$$\mathbb{E}_{\Pi_{\rho}}(\hat{\theta}_{\rho}(q, \alpha, \xi)) \leq \mathbb{E}_{\tilde{\Pi}}(\hat{\theta}_{\rho}(q, \alpha, \xi)) \quad (\text{B.23})$$

$$\leq \sum_{k \geq 0} \delta^k (1 - \delta) \mathcal{T} \leq \mathcal{T} \quad (\text{B.24})$$

Therefore γ has a finite mean and hence $\gamma \in \mathcal{P}$.

B.3.6 Details of Proof of Lemma 26

Since $\gamma(x) \in \mathcal{F}(\mathcal{P})$, there exists a $\rho \in \mathcal{P}$ such that $\gamma(x) = (\mathcal{F}(\rho))(x)$. This implies

$$\gamma(x) = \Pi_\rho(\theta_{\rho,z}^{*-1}([-z, x])) \quad (\text{B.25})$$

$$= \int_{\eta=0}^1 \int_{\alpha=1}^\infty \int_{\xi=1}^\infty \int_{q=0}^\infty \pi_\rho(q, \alpha, \xi, \eta) \mathbf{I}[\theta_{\rho,z}^*(q, \alpha, \xi, \eta) \leq x] dq d\alpha d\xi d\eta \quad (\text{B.26})$$

Let

$$W(x, \alpha, \xi, \eta) = \int_{q=0}^\infty \pi_\rho(q, \alpha, \xi, \eta) \mathbf{I}[\theta_\rho^*(q, \alpha, \xi, \eta) \leq x] dq.$$

Then,

$$\gamma(x) = \int_{\eta=0}^1 \int_{\alpha=1}^\infty \int_{\xi=1}^\infty W(x, \alpha, \xi, \eta) d\alpha d\xi d\eta$$

By using reverse Fatou lemma we can show that,

$$\gamma'_+(x) \leq \int_{\eta=0}^1 \int_{\alpha=1}^\infty \int_{\xi=1}^\infty W'_+(x, \alpha, \xi, \eta) d\alpha d\xi d\eta$$

Where,

$$\begin{aligned} W'_+(x, \alpha, \xi, \eta) &= \limsup_{y \rightarrow x} \frac{W(y, \alpha, \xi, \eta) - W(x, \alpha, \xi, \eta)}{y - x} \\ &= \limsup_{y \rightarrow x} \frac{\int_{q=0}^\infty \pi_\rho(q, \alpha, \xi, \eta) \mathbf{I}[\theta_\rho^*(q, \alpha, \xi, \eta) \leq y] dq - \int_{q=0}^\infty \pi_\rho(q, \alpha, \xi, \eta) \mathbf{I}[\theta_\rho^*(q, \alpha, \xi, \eta) \leq x] dq}{y - x} \\ &= \limsup_{y \rightarrow x} \frac{\int_{q=m(y, \alpha, \xi, \eta)}^\infty \pi_\rho(q, \alpha, \xi, \eta) dq - \int_{q=m(x, \alpha, \xi, \eta)}^\infty \pi_\rho(q, \alpha, \xi, \eta) dq}{y - x} \\ &= \limsup_{y \rightarrow x} \frac{\int_{q=m(y, \alpha, \xi, \eta)}^{m(x, \alpha, \xi, \eta)} \pi_\rho(q, \alpha, \xi, \eta) dq}{y - x} \\ &\leq - \limsup_{y \rightarrow x} \frac{m(y, \alpha, \xi, \eta) - m(x, \alpha, \xi, \eta)}{y - x} \pi_\rho(m(y, \alpha, \xi, \eta), \alpha, \beta, \eta) \\ &= -m'_+(x, \alpha, \xi, \eta) \pi_\rho(m(x, \alpha, \xi, \eta), \alpha, \beta, \eta). \end{aligned}$$

Therefore,

$$\gamma'_+(x) \leq - \int_{\eta=0}^1 \int_{\alpha=1}^{\infty} \int_{\xi=1}^{\infty} m'_+(x, \alpha, \xi, \eta) \pi_\rho(m(x, \alpha, \xi, \eta), \alpha, \xi, \eta) d\alpha d\xi d\eta. \quad (\text{B.27})$$

Now we will bound $m'()$ by dividing into two cases $m > 1$ and $m \leq 1$.

Case 1: $m > 1$

$$\begin{aligned} -\frac{1}{m'_+(x, \alpha, \xi, \eta)} &= -\liminf_{y \rightarrow x} \frac{y - x}{m(y, \alpha, \xi, \eta) - m(x, \alpha, \xi, \eta)} \\ &\geq \delta \mathbb{E}_D \left[(C'(m(x, \alpha, \xi, \eta) + 1 - D)) - C'((m(x, \alpha, \xi, \eta) - D)^+) \right] \\ &= \delta \mathbb{E}_D \left[\frac{(C'(m(x, \alpha, \xi, \eta) + 1 - D)) - C'((m(x, \alpha, \xi, \eta) - D))}{(m + 1 - D) - (m - D)} \right] \\ &= \delta \mathbb{E}_D [C''(y)] \\ &\geq K\delta \end{aligned}$$

Case 2: $m \leq 1$

$$\begin{aligned} -\frac{1}{m'_+(x, \alpha, \xi, \eta)} &= -\liminf_{y \rightarrow x} \frac{y - x}{m(y, \alpha, \xi, \eta) - m(x, \alpha, \xi, \eta)} \\ &\geq \delta \mathbb{E}_D \left[(C'(m(x, \alpha, \xi, \eta) + 1 - D)) - C'((m(x, \alpha, \xi, \eta) - D)^+) \right] \\ &= \delta \int_0^m \left[\frac{(C'(m(\cdot) + 1 - d)) - C'((m(\cdot) - d))}{(m + 1 - D) - (m - D)} f_D(d) \right] \\ &\quad + \delta \int_m^1 (C'(m(\cdot) + 1 - d)) - C'(0) f_D(d) \\ &= \delta \int_0^m C''(y_1) f_D(d) + \delta \int_m^1 C''(y_2) (m + 1 - d) f_D(d) \\ &\geq \delta K F_D(m) + \delta \int_m^1 K (m + 1 - d) f_D(d) \\ &= \delta K \left(F_D(m) + (m + 1 - d) F_D(d)|_m^1 + \int_m^1 F_D(d) \right) \\ &= \delta K \left(F_D(m) + (m) - F_D(m) + \int_m^1 F_D(d) \right) \\ &= \delta K \left(m + \int_m^1 F_D(d) \right) \end{aligned}$$

$$\geq \delta K$$

There fore, $-m'(x, \alpha, \xi, \eta) \leq K = \frac{1}{\delta K}$.

B.4 Proofs from Section 4.7

B.4.1 Details of Proof of Lemma 29

$$g(\alpha, \alpha + \beta, \mu) = \frac{\int_{\eta} \eta \text{Binom}(\alpha, \beta, \eta) d\mu(\eta)}{\int_{\eta} \text{Binom}(\alpha, \beta, \eta) d\mu(\eta)}$$

Let $u(x)$ be the pdf of $\mu(x)$, then

$$\begin{aligned} g(\alpha, \alpha + \beta, \mu) &= \frac{\int_0^1 \binom{\alpha+\beta}{\alpha} \eta \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta}{\int_{\eta} \binom{\alpha+\beta}{\alpha} \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta} \\ &= \frac{\int_0^1 \eta \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta}{\int_{\eta} \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta} \end{aligned}$$

Now,

$$\frac{g(\alpha + 1, \alpha + 1 + \beta, \mu)}{g(\alpha, \alpha + \beta, \mu)} = \frac{\frac{\int_0^1 \eta^2 \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta}{\int_{\eta} \eta \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta}}{\frac{\int_0^1 \eta \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta}{\int_{\eta} \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta}} \quad (\text{B.28})$$

$$= \frac{\left(\int_0^1 \eta^2 \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta \right) \left(\int_0^1 \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta \right)}{\left(\int_0^1 \eta \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta \right)^2} \quad (\text{B.29})$$

Let $\rho(\eta) = \frac{\eta^{\alpha} (1-\eta)^{\beta} u(\eta)}{\int_0^1 \eta^{\alpha} (1-\eta)^{\beta} u(\eta) d\eta} \implies \frac{g(\alpha+1, \alpha+1+\beta, \mu)}{g(\alpha, \alpha+\beta, \mu)} = \frac{\mathbb{E}_{\rho}[\eta^2]}{\left(\mathbb{E}_{\rho}[\eta]\right)^2} \geq 1$. The inequality is due to applying Jensen's inequality applying to $f(x) = x^2$.

Therefore, $g(\alpha + 1, \alpha + 1 + \beta, \mu) \geq g(\alpha, \alpha + \beta, \mu)$. The second part can be proved by change of variables $1 - \eta \rightarrow \eta'$ and using the first property.

APPENDIX C

PROOFS FROM CHAPTER 5

C.1 Auxiliary Results

Lemma 40. *For each state x , the fraction of peers with least popular chunk is upper bounded by $\frac{m-1}{m}$.*

Proof. Any peer in the system can have at most $m - 1$ pieces, or else it would leave the system. The result follows from bounding the total number of pieces in the system as

$$m\underline{\pi}|x| \leq \sum_{i \in [m]} \pi_i|x| = \sum_{S \subsetneq [m]: S \neq \emptyset} |S|x_S \leq (m-1)|x|.$$

□

Recall $\bar{\pi}$ and $\underline{\pi}$ respectively denote the fraction of peers that have the most and least popular chunks. When all chunks are equally popular, then $\underline{\pi} = \bar{\pi} = \pi_j$ for each chunk j .

When the set of most popular chunks $I(x) \subsetneq [m]$, the least popular chunk is denoted by $\underline{j} \notin I(x)$, and $\underline{\pi} = \pi_{\underline{j}}$. In this case, the least popular chunks are possessed by at least one less peer than the corresponding number for other chunks. That is, when $\pi_i > \underline{\pi}$, we have $\pi_i|x| - \underline{\pi}|x| \geq 1$. Specifically, $2(\bar{\pi} - \underline{\pi})|x| - 1 \geq 1$.

Lemma 41. *Let $K_1 > 0, K_2 < 2$ be constants. For each $\epsilon > 0$ there exists an $N(K_1, K_2, \epsilon) \in \mathbb{R}^+$, such that if $\bar{\pi}|x| \geq N$, then for $I(x) \subsetneq [m]$, we have*

$$C_1\lambda - K_1 \sum_{j \notin I(x)} R_j(1 - \pi_j)(2(\bar{\pi} - \pi_j)|x| - K_2) < -\epsilon.$$

Proof. Lower bounding the summation over $[m] \setminus I(x)$ by a single term corresponding to the least popular chunk \underline{j} , and lower bounding $1 - \underline{\pi}$ by $\frac{1}{m}$ from Lemma 40, we can upper bound

the LHS of the above equation by

$$C_1\lambda - \frac{K_1}{m}R_j(2(\bar{\pi} - \pi_j)|x| - K_2).$$

To upper bound the above equation, we define η as the ratio of number of peers with the least and the most popular chunks. That is, $\underline{\pi} = \eta\bar{\pi}$ and $\eta \in \left[0, 1 - \frac{1}{\bar{\pi}|x|}\right]$, and we can write

$$\begin{aligned} R_j(2(\bar{\pi} - \pi_j)|x| - K_2) &= (U + \eta\bar{\pi}\mu|x|)(2\bar{\pi}(1 - \eta)|x| - K_2) \\ &= -K_2U + 2U\bar{\pi}|x|(1 - \eta) - K_2\eta\bar{\pi}|x|\mu + 2\bar{\pi}^2|x|^2\mu\eta(1 - \eta). \end{aligned}$$

Let us denote the above quadratic expression in η by $g(\eta)$. We can check that $g''(\eta) = -4\bar{\pi}^2|x|^2\mu < 0$. Hence, the function $g(\eta)$ is strictly concave and quadratic in η , with a unique maximum. This function attains minimum at the boundary values of η , and we can lower bound $g(\eta)$ as

$$\begin{aligned} g(\eta) &\geq \min\{g(\eta) : \eta \in [0, 1 - \frac{1}{\bar{\pi}|x|}]\} = g(0) \wedge g(1 - \frac{1}{\bar{\pi}|x|}) \\ &= \frac{K_1}{m} [U(2\bar{\pi}|x| - K_2) \wedge (2 - K_2)(U + \mu(\bar{\pi}|x| - 1))]. \end{aligned}$$

The result follows since $C_1\lambda - \frac{K_1}{m}g(\eta) < -\epsilon$ if $\bar{\pi}|x| > N$, where we can choose N to be

$$\max \left\{ \frac{1}{2} \left(\frac{C_1\lambda + \epsilon}{\frac{K_1}{m}U} + K_2 \right), \left(\frac{C_1\lambda + \epsilon}{\frac{K_1}{m}(2 - K_2)\mu} - \frac{U}{\mu} + 1 \right) \right\}.$$

□

Corollary 4. *Let $K_1 > 0, K_2 < 2$ be constants, $\bar{\pi}(x) \geq \delta$, and $I(x) \subsetneq [m]$. Then, for each $\epsilon > 0$, we can find an L such that when $|x| > L$,*

$$C_1\lambda - K_1 \sum_{j \notin I(x)} R_j(1 - \pi_j)(2(\bar{\pi} - \pi_j)|x| - K_2) < -\epsilon.$$

Proof. Fix $\epsilon > 0$, we choose the N from Lemma 41 and $L = \frac{N}{\delta}$. Then $\bar{\pi}|x| \geq N$, and the inequality holds. \square

C.2 Proof of Theorem 17

For any $\delta \in (0, 1)$, we can partition the state space into following three regions,

$$\mathcal{R}_1 = \{\bar{\pi} \geq \delta\}, \mathcal{R}_2 = \mathcal{R}_1^c \cap \{\bar{\pi}|x| \geq \frac{M}{m}\}, \text{ and}$$

$$\mathcal{R}_3 = \mathcal{R}_1^c \cap \{\bar{\pi}|x| < \frac{M}{m}\}.$$

For each $i \in [3]$, we can further subdivide each region R_i into

$$\mathcal{R}_{i1} = \{x \in \mathcal{R}_i, I(x) \subsetneq [m]\}, \quad \mathcal{R}_{i2} = \{x \in \mathcal{R}_i, I(x) = [m]\}.$$

All these regions have countable number of states. We will prove that in each region \mathcal{R}_{ij} where $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$, the mean drift $QV(x) < -\epsilon$ for all states $x \in \mathcal{R}_{ij} \setminus F_{ij}$ for some finite set F_{ij} dependent on ϵ . We fix $\epsilon > 0$, and choose $N(K_1, K_2, \epsilon)$ from Lemma 41.

Lemma 42. *For states in the region \mathcal{R}_1 , the total number of chunks is lower bounded by $\delta|x|$.*

Proof. The total number of chunks in the system r is lower bounded by the number of most popular chunk, i.e.

$$r = \sum_{i \in [m]} \pi_i |x| \geq \bar{\pi} |x|.$$

The result follows since $\bar{\pi} \geq \delta$ in the region \mathcal{R}_1 . \square

Lemma 43. *For states in the region $\mathcal{R}_2 \cup \mathcal{R}_3$, the fraction of peers γ_j with the set of chunks $\{j\}^c$ is upper bounded by δ .*

Proof. We can upper bound the number of peers with the set of chunks $\{j\}^c$ as

$$x_{\{j\}^c} \leq \sum_{S: i \in S, i \neq j} x_S = |x| \pi_i 1_{\{i \neq j\}} \leq |x| \bar{\pi}.$$

The result follows since $\bar{\pi} < \delta$ in region $\mathcal{R}_2 \cup \mathcal{R}_3$. \square

Region \mathcal{R}_{11} : For this region, we choose $N_{11} \triangleq N(\frac{1}{m}, 1, \epsilon)$ from Lemma 41, to define the finite set

$$F_{11} \triangleq \{\delta|x| \leq (M + m - 1) \vee N_{11}\}.$$

From Lemma 42, it follows that for the states $x \in \mathcal{R}_{11} \cap F_{11}^c$, the indicator corresponding to the event $\{M + m - 1 > r\}$ is zero in the mean drift of (5.2). Therefore, Corollary 4 implies that the mean drift in (5.2) is upper bounded by $-\epsilon$.

Region \mathcal{R}_{12} : For this region, we choose $C_1 > (m - 1)$ and

$$N_{12} \triangleq \frac{m(C_1 \lambda + \epsilon)}{\mu \delta (C_1 - m + 1)},$$

to define the finite set $F_{12} \triangleq \{\delta|x| \leq (M + m - 1) \vee N_{12}\}$. Let $x \in \mathcal{R}_{12} \cap F_{12}^c$. In this region, the indicator corresponding to the event $\{M + m - 1 > r\}$ is zero in the mean drift of (5.3). By choosing a lower bound on common contact rate $R \geq \mu \bar{\pi} |x|$, complement of the frequency $1 - \bar{\pi} \geq \frac{1}{m}$ from Lemma 40, and on the frequency $\bar{\pi} \geq \delta$ since $x \in \mathcal{R}_1$, we can bound the mean drift in equation (5.3) by $-\epsilon$.

Region \mathcal{R}_{21} : We can upper bound the fraction of peers $\gamma_j < \delta$ by Lemma 43, and upper bound $1 < m(1 - \pi_j)$ from Lemma 40. Thus, we can upper bound

$$\gamma_j C_2 m (m - 1) 1_{\{M + m - 1 > r\}} \leq \delta C_2 (1 - \pi_j) m^2 (m - 1).$$

Hence, we can upper bound the mean drift in (5.2) with

$$C_1\lambda - \sum_{j \notin I(x)} \frac{R_j(1 - \pi_j)}{m} (2(\bar{\pi} - \pi_j)|x| - 1 - \delta C_2 m^2(m - 1)).$$

When $\delta C_2 m^2(m - 1) < 1$, we can choose from Lemma 41

$$N_{21} \triangleq N \left(\frac{1}{m}, \delta C_2 m^2(m - 1) + 1, \epsilon \right).$$

For each $x \in \mathcal{R}_{21}$, $\bar{\pi}|x| \geq \frac{M}{m}$, and hence by selecting $\frac{M}{m} > N_{12}$, we ensure that the mean drift in (5.2) is bounded above by $-\epsilon$ in this region.

Region \mathcal{R}_{22} : In this region, $\underline{\pi}|x| \geq \frac{M}{m}$, and hence the total number of chunks $r = \sum_i \pi_i |x| \geq M$. We again choose $C_1 > m - 1$ and bound the common contact rate $R = U + \underline{\pi}\mu|x| \geq \frac{M}{m}\mu$. We also use the upper bound for fraction of peers $\gamma_j < \delta$ from Lemma 43, and the lower bound $1 - \underline{\pi} \geq \frac{1}{m}$ from Lemma 40, to upper bound the mean drift from equation (5.3) with

$$C_1\lambda - \frac{M}{m^2}\mu (C_1 - m + 1 - \delta C_2 m^2(m - 1)).$$

Hence, the mean drift for all states $x \in \mathcal{R}_{22}$ is bounded above by $-\epsilon$, if $\delta C_2 m^2(m - 1) < C_1 - m + 1$ and

$$\frac{M}{m} > N_{22} \triangleq \frac{(C_1\lambda + \epsilon)m}{\mu (C_1 - m + 1 - \delta C_2 m^2(m - 1))}.$$

Region \mathcal{R}_{31} : In this region, $m\bar{\pi}|x| < M$, and hence the total number of chunks $r = \sum_i \pi_i |x| \leq m\bar{\pi}|x| < M$. Therefore both the indicator functions associated with the events $\{M > r\}$ and $\{M + m - 1 > r\}$ equal unity in the equation (5.2). Recall that $\gamma_j < \delta$ by Lemma 43, $1 - \underline{\pi} > \frac{1}{m}$ by Lemma 40, $R_j \geq U$, and $2(\bar{\pi} - \underline{\pi})|x| - 1 > 0$ for state such that

$I(x) \subsetneq [m]$. Summarizing all these results, and lower bounding the summation over $I^c(x)$ by the least popular term, we can upper bound the mean drift with

$$C_1\lambda - C_2\frac{U}{m} \left(\frac{1}{m} - \delta(m(m-1) + 1) \right).$$

By choosing $\delta < \frac{1}{2m(m(m-1)+1)}$ and $C_2 = \frac{2m^2(C_1\lambda + \epsilon)}{U}$ we can bound the mean drift from state $x \in \mathcal{R}_{31}$ with $-\epsilon$.

Region \mathcal{R}_{32} : Similar to region \mathcal{R}_{31} both the indicator functions in equation (5.2) will be equal to unity. In this region, $\bar{\pi} = \underline{\pi}$. Using the bounds $\gamma_j < \delta$, $1 - \underline{\pi} > \frac{1}{m}$, $R \geq U$, and $C_1 > m - 1$, we can upper bound the mean drift in (5.3) for $x \in \mathcal{R}_{32}$ with

$$C_1\lambda - UC_2 \left(\frac{1}{m} - \delta(m(m-1) + 1) \right)$$

By choosing $\delta < \frac{2m-1}{2m^2(m(m-1)+1)}$ and $C_2 = \frac{2m^2(C_1\lambda + \epsilon)}{U}$ we can bound the drift with $-\epsilon$.

C.2.1 Choosing Parameters:

Following choice of C_1, C_2, M satisfy all the constraints,

$$C_1 > m - 1, \quad C_2 = \frac{2m^2(C_1\lambda + \epsilon)}{U}, \quad M > m \max \{N_{21}, N_{22}\},$$

where $m^2(m-1)\delta$ equals

$$\min \left\{ \frac{m - \frac{1}{2}}{m + \frac{1}{m-1}}, \frac{\frac{m}{2}}{m + \frac{1}{m-1}}, \frac{1}{C_2}, \frac{C_1 - m + 1}{C_2} \right\}.$$