APPROXIMATELY-OPTIMAL MECHANISMS IN AUCTION DESIGN, SEARCH THEORY, AND MATCHING MARKETS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Hedyeh Beyhaghi
December 2019

APPROXIMATELY-OPTIMAL MECHANISMS IN AUCTION DESIGN, SEARCH THEORY, AND MATCHING MARKETS

Hedyeh Beyhaghi, Ph.D.
Cornell University 2019

Algorithmic mechanism design is an interdisciplinary field, concerned with the design of algorithms that are used by strategic agents. This field has applications in many real-world settings, such as auction design, search problems, and matching markets. In this thesis we study the mechanisms in these areas through lenses of simplicity and practicality. We pursue two main directions: study the strength of simple mechanisms, and improve the efficiency of practical mechanisms. In auction design, we consider a setting where a seller wants to sell many items to many buyers, and establish a tight gap between the efficiency of a simple and commonly-used auction with the complicated revenue-optimal auction. In search theory context, we introduce *Pandora's problem with alternative inspections* and provide the first approximately-optimal mechanism for this problem. In Pandora's problem with alternative inspections, a searcher wants to select one out of n elements whose values are unknown ahead of time. The searcher evaluates the elements one by one and can choose among different costly ways to evaluate each element, the order to evaluate the elements, and how long to continue the search, in order to maximize her utility. In matching markets, we propose theoretical models that closely capture the participant behaviors in the real world, and provide methods to optimize the already implemented mechanisms.

BIOGRAPHICAL SKETCH

Hedyeh Beyhaghi is pursuing a Ph.D. in Computer Science at Cornell University. Her research focuses on Algorithm Design with an emphasis on Algorithmic Game Theory and Mechanism Design. During her PhD studies, she was a long-term visitor at the Simons Institute for theory of computing at Berkeley, an intern at Research and Ad Exchange groups at Google, and an Ivy-Plus Exchange Scholar in the Computer Science Department at Princeton University. Before joining Cornell, Hedyeh received her B.S. degree in Computer Engineering from Sharif University of Technology.

I dedicate this dissertation to my family.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Éva Tardos for guiding me through my PhD program. Her ability to clearly convey the merit of research directions and identify those worth pursuing has been very influential to me. I learned a lot by observing the way that she approaches research problems, interprets the results, and finds connections between different concepts. I would like to thank her for the countless research discussions and inspiring ideas.

Many thanks to Matt Weinberg for hosting me at Princeton University. Working with Matt shaped my understanding of many problems in mechanism design and taught me many mathematical tools that I did not possess before. I would like to thank him for always being available to provide feedback and advice. I am also grateful to Bobby Kleinberg for our collaboration and also for being a great mentor to me. His advice has been very influential at different points in my academic life and has helped me make better career decisions. Thank you as well to Joe Halpern for serving on my dissertation committee and for the many great things he taught me about writing and presentations.

I would also like to thank everyone whom I have had the pleasure of collaborating with: Bobby, Éva, Matt, Maryam Bahrani, Nishanth Dikkala, Negin Golrezaei, Brendan Lucier, Renato Paes Leme, Martin Pál, Daniela Saban, Ariel Schvartzman, Sahil Singla, and Balu Sivan. Special thanks to Renato and Alex Van Esbroeck for hosting me during my summer internship at Google. Thank you as well to the Simons Institute for the Theory of Computing at UC Berkeley for an amazing semester on Economics and Computation.

Thank you to my friends in the theory groups at Cornell and Princeton who made work more enjoyable: Meryem Essaidi, Matheus Ferreira, Sumegha Garg, Pooya Jalaly, Thodoris Lykouris, Divya Mohan, Rad Niazadeh, Vidhya Ra-

maswamy, and Rahmtin Rotabi.

Finally, I would like to thank my family for their support and encouragement over the years. I am especially grateful to Mom, memory of Dad, Mehdi, and Hadi for everything that they've done for me. I wish to dedicate this dissertation to them.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Algorithmic mechanism design is an interdisciplinary field, lying at the intersection of economic game theory, optimization, and computer science. The goal of this field is to design algorithms that perform well when used by *selfish* agents. These agents choose their actions strategically to maximize their gains.

In algorithm design, finding approximately optimal algorithms is mainly motivated for the class of NP-hard problems. However, since strategic settings are much more complex than non-strategic ones, in algorithmic mechanism design, it is often the goal to find *simple* mechanisms that closely approximate the optimal solutions. The difference between complexity of algorithms and mechanisms arises even in the simplest algorithmic problems. For example, the simple algorithmic problem of finding the maximum value becomes complicated when the inputs are provided by selfish agents who may have incentives to misreport the values to achieve a better outcome. In contrast to complex optimal mechanism, simple mechanisms are easier to describe and require less knowledge about the environment. The tradeoff between simplicity and optimality triggers a fundamental question: *How powerful are simple mechanisms; i.e., how close to optimal are their solutions?*

On the other hand, since mechanism design questions stem from real-world problems there is an implementation step after designing the mechanism. If the mechanism designed does not take all the practical limitations into account, the implemented mechanism will not be exactly as designed. Therefore, the result is a mechanism that although practical, does not meet the same performance guarantees and can still be improved. This mismatch between the implementation and theory motivates studying the following question: *Can we improve widely used* *mechanisms without hurting their practicality?*

In this thesis we study the above questions. For the rest of this chapter we provide more details of the context and our approach regarding each direction.

1.1 Simple vs Optimal Mechanisms

1.1.1 Simplicity Measures

Because of the complexity of optimal mechanisms, the study of simple mechanisms are often well-motivated. However, depending on the context of the problem, simplicity has different interpretations. We discuss two notions of simplicity used in this thesis.

Often times optimal mechanisms need to have detailed knowledge about the environment. Simple mechanisms, on the other hand, do not depend on the specific knowledge about the environment and perform reasonably well irrespective of the underlying setting. This notion of simplicity has applications in many environments, e.g., when inputs of the problem are assumed to be drawn from a distribution. A mechanism that does not need to know the distribution is called *prior-independent*. Compared to prior-dependent optimal mechanisms, prior-independent mechanisms are more robust; i.e., although small perturbations in the environment require a change in the optimal mechanisms, the same prior-independent mechanisms still work in the new setting.

As an example consider a seller who wants to sell an item to one of many potential buyers. The seller wants to run a mechanism that will maximize her revenue. The optimal mechanism in this setting needs to know the distribution that buyer values come from. However, there are prior-independent mechanisms that approximate the optimal revenue. Meaning, for any distribution, their performance is within a reasonable factor of the optimal mechanism.

Another notion of simplicity, specifically in online mechanisms, is independence of the past history. An online mechanism consists of multiple rounds and at each round the mechanism needs to make a decision. The optimal mechanism may need to make a decision that depends on the outcome of previous rounds. However, there may be approximately-optimal mechanisms that commit to decisions ahead of time and don't rely on the past history. These simple mechanisms are called *non-adaptive.*

For example, consider the following maximization problem. There are n elements whose values are random variables $r_1, \dots, r_n$ drawn from distributions $F_1, \dots, F_n$, respectively. A searcher knows the distributions in advance and can only realize the value of an element by probing it. The searcher is allowed to probe the elements one by one and can probe at most k elements. The objective is to maximize the highest probed value. The searcher decides at each point what element to probe next until she probes k elements. The optimal policy for this problem is adaptive, i.e., the searcher needs to select the next element to probe based on previously realized values. However, there exist approximately optimal non-adaptive policies that select the k elements to probe ahead of time and do not depend on already realized values.

1.1.2 Optimality Measures

Similar to simplicity, there are different optimality measures; i.e., there are different ways to measure the performance of simple mechanisms compared to the optimal ones.

In this thesis we consider two notions for measuring optimality. One of the most standard optimality measures stems from the field of approximation algorithms. This notion measures the efficiency of the mechanism as a multiplicative guarantee; i.e., the optimal solution is always guaranteed to be within a (predetermined) multiplicative factor of the returned solution. However, there are other approaches to measure the performance of simple mechanisms. For example in *resource augmentation* paradigm, the performance of the mechanism corresponds to the amount of extra resources the mechanism needs, in order to perform as well as the optimal mechanism (without the extra resources). The less resources the mechanism needs, the more efficient it is considered. The main idea behind this measure is that instead of designing a superior mechanism, the mechanism designer can invest in acquiring more resources to improve the performance. An application of this measure is presented in the next subsection.

1.1.3 Competition Complexity in Auction Design

Auction Design is one of the major areas in algorithmic mechanism design. In the auction design problems studied in this thesis, a seller wants to sell a set of items to one or many potential buyers. The seller's goal is to maximize the revenue of selling the items. Each buyer has a private value for each set of items. The value for each set of items is how much the buyer is willing to pay to receive that set. The buyers' values are drawn from a distribution that corresponds to the population that the buyers come from. The utility of each buyer upon receiving a set of items is her value for that set minus the price she pays. The buyers are strategic and their goal is to maximize their utility.

Consider selling m items to n buyers. Suppose the buyers are additive, i.e., their value for a set of items sums their values for items in that set. Furthermore, buyers' values for individual items are independent across items and buyers. How many buyers do we need to add such that the revenue of a simple prior-independent auction (with additional buyers) exceeds the optimal revenue with n buyers? This question is well-motivated for the following reasons. The optimal auction is unknown or complicated even for simple multi-item settings. For example, the optimal mechanism for selling two i.i.d. items to a single buyer may need randomization and can have uncountably many options for the buyer.[1] The known approximately-optimal auctions only recover a small fraction of the optimal revenue. The optimal and many approximately-optimal auctions are not robust and need to have full knowledge of the distributions. A small error or perturbation in distributions may make the previously optimal auction suboptimal. Lastly, instead of investing resources to learn the distributions and implement the optimal auction the seller can invest in increasing the number of buyers; for example by advertising, and recover the optimal revenue of the original setting. The necessary number of additional buyers for a specific mechanism in order to achieve the optimal revenue is called the *competition complexity* of that mechanism.

This problem has a clear answer in the single-item setting. The elegant result of Bulow and Klemperer [12] shows that when buyers are i.i.d. and the distributions

---

[1] Each option for a buyer is a pair of allocation, which specifies the probability of receiving each item, and a payment.

are regular[2], only adding one buyer is sufficient to achieve full revenue with the second price auction. One reason that makes the analysis more complicated for the multi-item setting is that the optimal auction and hence the optimal revenue are unknown. In Chapter 2, we prove that for n buyers with additive valuations over m independent items, the number of additional buyers for the mechanism that sells the items separately via the second price auction, in order to achieve the optimal revenue, is at most $n(ln(1 + m/n) + 2)$, and also at most $9\sqrt{nm}$. When $n \leq m$ the first bound is tight (up to constant factors). When $n \geq m$, the second bound is tight (up to constant factors) for any argument that starts from the benchmark introduced in [24]. Our results strictly improve the results of prior works including [24, 26] regarding selling separately to additive buyers with independent items.

1.1.4 Adaptivity Gap in Online Stochastic Optimization

Online stochastic optimization involves solving optimization problems when the inputs are random variables whose values are revealed one-by-one in an online sequence. Although solving the problem in the full information setting (where we know all the values) might be an easy task, the randomness and online arrival of inputs may require a complicated solution.

Consider the following stochastic optimization problem that we study in this thesis. Suppose a firm wants to hire an employee from a list of candidates. Each candidate has a score, unknown to the firm. The firm only knows the score distribution for each candidate. The evaluation cost depends on the candidate and the evaluation method used. Weitzman [59] studies the problem when a candidate cannot be selected unless the firm undertakes a unique evaluation process to assess

---

[2] A distribution with CDF F is regular if $x - \frac{1-F(x)}{f(x)}$ is monotone non-decreasing.

her. He surprisingly shows that the optimal solution has a simple structure; it assigns an index to each candidate at the beginning and evaluates the candidates in decreasing order of indices until a realized score exceeds the next index.

The assumption of unique evaluation methods for candidates limits the applicability of the model. In the real world the firm may be able to choose among distinct possible methods to evaluate a candidate where each of these methods comes at a distinct cost and reduces the uncertainty of the reward by a different amount. For example, the firm can decide to assess a candidate through a lengthy internship process, evaluate her through a shorter interview process, or skip any such costly evaluations and hire directly. Doval [22] shows that this well-motivated model with alternative inspections cannot be solved optimally by any simple ranking-based policies; a special case of the model was studied in a wireless networking context by Guha et al. [33] who also showed the suboptimality of simple ranking-based policies. It is also unknown whether there exists any polynomial-time algorithm to compute the optimal policy. This motivates the study of approximately optimal policies that are simple and computationally efficient.

In Chapter 3, we provide the first non-trivial approximation guarantees for this problem. Our technical contribution is introducing *committing policies* and finding their approximation guarantees. A committing policy is one that, before it evaluates any candidates, must pre-commit to a specific method of evaluating each candidate; in addition, it pre-commits to an order in which the candidates will be evaluated. The committing policies correspond to non-adaptive solutions in stochastic optimization. We show that the optimal committing policy attains at least $1 - \frac{1}{e} \approx 0.63$ fraction of the optimal policy's value. We further show that committing policies cannot achieve more than 4/5 of the optimum; implying a 4/5 adaptivity gap. This matches the lower bound of 4/5 proven by Guha et al. [33] in a special case of our model, thereby showing that their analysis of the adaptivity gap is tight for that special case.

1.2 Practical Limitations

Practical limitations add another layer to the difficulty of designing mechanisms. Many times the mechanisms designed theoretically, do not consider all these limitations. Therefore the implemented mechanisms are not exactly as designed, and result in different outcomes compared to what were predicted theoretically. To overcome this mismatch, an alternative approach is to understand the implemented mechanisms better, and find ways to optimize them while keeping them practical.

1.2.1 Practical Limitations in Matching Markets

In many centralized markets like the National Residency Matching Program (NRMP), the matching between residency applicants and positions is found via the *deferred acceptance* algorithm. While the deferred acceptance algorithm assumes providing full preference list by participants, in practice, reporting all desired choices is infeasible for participants in many circumstances. For instance, in NRMP an interview process occurs before the matching process, and applicants are only allowed to list the hospitals they interviewed with. Therefore, among nearly 5000 residency programs, applicants have to substantially limit their choices. Thus, even if it is in applicants' best interest to provide a preference list including all their acceptable positions[3], residents must be strategic about the hospital programs they choose to interview.

The question of how to select a short list of positions becomes especially interesting when the participants have incomplete information about other participants' preferences and preferences are somewhat aligned, i.e, some hospitals and applicants are more popular than the others. If there is no alignment in preferences, the positions look the same in terms of popularity and it makes sense for the applicants to apply to their top choices. In the case of full information setting, each applicant knows exactly what position they get matched to and they only need to apply to those programs. In Chapter 4, we propose a simple model that captures both imperfect information and a degree of alignment in preferences. We study three questions: (1) how strategic applicants decide where to apply; (2) how these decisions are affected by the market primitives; and (3) what the *social welfare*[4] of the outcome is in these markets. We find that the equilibrium of our model captures the main behavior observed in practice where applicants apply to a mix of positions consisting mostly of places where they are reasonably likely to get accepted, as well as a few "reach" applications to positions of very high quality, and a few "safe" applications to positions of lower than their expected level. Our most interesting finding in the analysis of applicants' behavior with respect to market primitives is when we study the behavior as a function of number of applications. We find that when the applicants are allowed to apply to *more* positions, sometimes they prefer to *decrease* the number of their safe choices. Finally, we show that the social welfare produced by the equilibrium is reasonably close to

---

[3] in deferred acceptance algorithm, applicants receive a better (or at least an equally good) outcome by reporting their preferences truthfully

[4] The social welfare is defined as the sum of values that participants receive from the match, where the value of being matched to more preferred hospitals/doctors is higher than the less preferred.

the optimal welfare. This result is surprising because in order to receive a higher quality position applicants may apply to many reach positions and (due to the limited number of applications) deprive some of the lower quality positions from any applicants, which potentially can do significant damage to the overall welfare.

Similar to applicants, hospitals have to limit their choices. Among hundreds of applications that a hospital receives, they can only conduct interviews with a limited number of prospective residents. Studying a market where both sides are constrained is an interesting problem even for the simple case of *symmetric* markets i.e., the case in which all doctors and hospitals are equally popular.

In Chapter 5, we study the effect of limiting number of applications sent by applicants on the efficiency of the matching when the number of interviews conducted by hospitals is small. The notion of efficiency we use is the size of outcome matching. Similar to Chapter 4 the mechanism used is the deferred acceptance algorithm but unlike the previously mentioned model with a degree of alignment in preferences, in this chapter, we consider a symmetric model. Although it seems reasonable for the matching size to be increasing in the number of applications, our first main result shows that when the number of interviews is limited, a market with a limited number of applications achieves higher efficiency compared to a market with no limits. Note that this property is specific to the limited-interview model; with no limit on the number of interviews, matching size is increasing in the number of applications. Our second main result is that a system of treating all applicants equally (where everyone has the same limit on the allowed number of applications) is more efficient than allowing a small set to apply to one more or less position. Hence, when the expected size of matching is drawn as a function of the expected number of applications the resulting figure is scallop-shaped, with peaks occurring at integral expected number of applications. This property is also specific to the limited-interview model.

1.3 Roadmap

Optimal mechanisms are complex and impractical in many settings. We overcome these undesired properties by providing simple and practical mechanisms that are approximately optimal. We study prior-independent and non-adaptive mechanisms. We measure their performance either by directly comparing their solution to the optimum, or quantifying the amount of resources they need to achieve the optimum. Furthermore, we study the mechanisms used in practice, and find ways to make them more efficient while keeping them practical.

In Chapter 2, we find the optimal competition complexity of the Vickrey-Clarke-Groves mechanism for selling m independent items to n additive buyers. In Chapter 3, we introduce Pandora's problem with alternative inspections and provide the first non-trivial mechanism that approximates the optimum. Chapter 4 and 5 are devoted to matching markets. In Chapter 4, we propose a theoretical model explaining the applicants' behaviors in matching markets. In Chapter 5, we provide methods to optimize the already implemented matching mechanisms.

Chapter 2 and Appendix A are based on joint work with S. Matthew Weinberg [10]. Chapter 3 is based on joint work with Robert Kleinberg [7]. Chapter 4 and Appendix B are based on joint work with Daniela Saban and Éva Tardos [8]. Chapter 5 and Appendix C are based on joint work with Éva Tardos [9].

CHAPTER 2

OPTIMAL (AND BENCHMARK-OPTIMAL) COMPETITION COMPLEXITY FOR ADDITIVE BUYERS OVER INDEPENDENT ITEMS

2.1 Introduction

In the past decade, the TCS community has made radical progress developing the theory of multi-dimensional mechanism design. In particular, it was previously well-understood the optimal multi-item auctions are prohibitively complex, even with just $m = 2$ items, and even subject to fairly restricted instances [11, 37, 39, 58, 50, 20]. Yet, starting from seminal work of Chawla, Hartline, and Kleinberg [16], a large body of work now proves that simple auctions are in fact approximately optimal in quite general settings [17, 18, 38, 47, 6, 60, 55, 19, 14, 25], helping to explain the prevalence of simple auctions in practice. Still, it would be a reach to claim that this agenda is convincingly resolved.

In particular, the thought of settling for 50% (or even 90%) of the optimal achievable revenue may be a non-starter for high-stakes auctions. Indeed, there are no hard constraints forcing the auctioneer to use a simple auction. Still, *Prior-independent* auctions are desirable since they don't require the auctioneer to understand the population from which consumers are drawn. *Deterministic* and *Dominant Strategy Truthful* auctions are desirable because consumers' strategic behavior is easier to predict. *Computationally tractable* auctions are desirable because they can be efficiently found. On the other hand, it is hard to imagine that auctioneers stand a hard line on simplicity if additional market research or outsourcing computation would increase revenues, even modestly.

The resource augmentation paradigm takes a different view: spend effort recruiting additional bidders rather than carefully designing a superior auction. We are therefore interested in answering the following question: for a given auction setting, *how many additional bidders* are necessary for a simple auction (with additional bidders) to guarantee greater expected revenue than the optimal (without)? Eden et al. term the answer to this question the *competition complexity* [24].

This question was first studied in seminal work by Bulow and Klemperer in the context of single-item auctions [12]. Remarkably, they show that just a single additional bidder suffices for the second-price auction to guarantee greater expected revenue than Myerson's optimal auction [49] (without the additional bidder), subject to a technical condition on the population called regularity. For multi-item auctions, similar results have even more bite, as the optimal multi-item auction is considerably more complex than Myerson's (which is deterministic, dominant strategy truthful, and computationally tractable, but not prior-independent). Our main result is optimal bounds on the competition complexity for the core setting of additive bidders with independent items. Specifically,

Main Result: The competition complexity of n bidders with additive values over m independent items is at most $n(2 + \ln(1 + m/n))$, and also $9\sqrt{nm}$. When $n \leq m$ the first bound is tight (up to constant factors). When $n \geq m$, the second bound is tight (up to constant factors) for any argument that uses the benchmark introduced in [24].

2.1.1 Brief Technical Overview

Formally, we consider n bidders drawn independently from a distribution D. We study the now-standard setting of additive bidders over m independent items: each bidder's value v_j for item j is drawn independently from some single-variate distribution D_j, and her value for a set S of items is $\sum_{j \in S} v_j$. The simple mechanism we study is to sell the items separately, either via the second-price auction in the case of regular distributions, or Myerson's optimal single-item auction in the general case.[1] Observe that, since the bidders are additive and values are independent, selling the items separately is really just m separate single-item problems. We are interested in understanding the minimum $c(n, m)$ such that selling separately to $n+c(n, m)$ bidders drawn from D yields greater expected revenue than the optimal mechanism with n bidders drawn from D for *any* $D = \times_{j=1}^{m} D_j$.

Our approach starts from the benchmark proposed in [24]. That is, Eden et al. propose an upper bound on the optimal achievable revenue with n bidders drawn from D via the duality framework of [13].[2] We defer a definition of this benchmark to Section 2.2.2: it defines a *Virtual Value* $\Phi_j(\vec{v}_i)$ of a bidder with values $\vec{v}_i$ for item j, and upper bounds the optimal expected revenue with $\mathbb{E}[\sum_j \max_{i \in [n]} \{\Phi_j(\vec{v}_i)\}]$. We defer most details to the technical sections, but briefly note that at this point, our analysis diverges from prior work. Eden et al. use an elegant coupling argument to connect this benchmark to the expected revenue of selling separately with additional bidders [24]. The high-level distinction in our approach is a significantly more in-depth analysis of this benchmark. In particular, our analysis makes more

---

[1] For irregular distributions, it is known that no guarantees are possible with prior-independence, even for a single item. The example to have in mind is a distribution with a point mass at p with probability $1/p$ and 0 otherwise: as $p \to \infty$, any auction that achieves revenue close to optimum must sell the item to the bidder with value p for price close to p whenever there is exactly one. It is impossible to have a single auction that does this for all p.

[2] We note that this upper bound can also be derived without duality using techniques of [18].

extensive use of Myersonian virtual value theory (Sections 2.3 and 2.4), which reduces the problem to questions purely regarding whether various methods of drawing correlated values from $[0,1]$ stochastically dominate one another (Section 2.5).

2.1.2 Connection to Related Works

The two works most directly related to ours are [24] and [26]. The one-sentence distinction between our results and these is that we strictly improve their main results regarding selling separately to additive bidders with independent items, but do not address alternative settings. For example, this text contains no results beyond additive bidders (considered in [24]), or results for mechanisms aside from selling separately (considered in [26]).

"Little n Regime": For n additive bidders with $m = \Omega(n)$ independent items, Eden et al. [24] prove an upper bound of $n+2(m-1)$ on the competition complexity. Feldman et al. [26] prove that selling separately to $O(n\ln(m/n)/\varepsilon)$ additional buyers exceeds a $(1-\varepsilon)$ fraction of the optimal revenue (without the additional buyers). Our main result essentially achieves the greatly improved bound of [26] (and improves it further) without losing *any* revenue: we prove a competition complexity bound of $n(2+\ln(1+m/n))$. This guarantee is tight up to constant factors (and remains tight even if one is willing to lose an ε fraction), due to a lower bound of [26].

"Big n Regime": For n additive bidders with $m = o(n)$ independent items, Eden et al. [24] prove a competition complexity bound of $n+2(m-1)$. Feldman et al. [26] prove that for any ε, there exists a constant $\delta(\varepsilon)$ such that if $n \geq m/\delta(\varepsilon)$, selling separately (without any additional bidders) achieves a $(1-\varepsilon)$ fraction of the optimal revenue. Our main result improves the guarantee of [24] to $9\sqrt{nm}$ and also implies the result of [26] (with $\delta(\varepsilon) = \varepsilon^2/81$). Note in particular that any sublinear competition complexity bound implies the [26] result for a different $\delta(\cdot)$, but that linear bounds do not. So our improvement from linear to sublinear is significant in this regard. Moreover, we show in Section A.1 that this is tight (up to constant factors) for any approach starting from the benchmark proposed in [24]. We further show (also in Section A.1) that there does not exist any function only of m upper bounding the competition complexity: as $n \to \infty$ the competition complexity approaches ∞ as well (at a rate of at least $\Omega(\ln n)$).

Other works that study the competition complexity of auctions include seminal work of Bulow and Klemperer, who study the $m = 1$ case, work of Liu and Psomas (who study the competition complexity of dynamic auctions) and Roughgarden et al. (who study the unit-demand setting) [12, 48, 54]. These works are thematically related, but both the results and techniques have little overlap.

Some of the aforementioned works which prove approximation guarantees for simple mechanisms use similar techniques to derive a tractable benchmark that upper bound on the achievable revenue [16, 17, 18, 38, 47, 6, 60, 55, 19, 14, 25]. However, it is worth noting that all of these works proceed by immediately splitting the benchmark into multiple simpler terms and finding approximately optimal mechanisms to cover each term separately. The best of those mechanism guarantees approximate optimality to revenue. This greatly simplifies analysis, at the cost of an additional constant factor. Because competition complexity results target the full original revenue, losing this initial constant factor can make future analysis impossible. As a result, while benchmarks may be shared by these lines of work, analysis of the benchmarks is often quite different.

Finally, it is worth noting that recent work follows two approaches to derive revenue upper bounds in these works. Some (including this text) use virtual value theory [16, 17, 54, 18, 13, 14, 24, 25, 48, 27]. Others use a more direct probabilitistic approach [38, 47, 6, 60, 55, 19, 5, 26]. For the most part, similar approximation guarantees are achievable through both approaches. With respect to these lines of work, our results (which yield exact competition complexity bounds) in comparison to those of [26] (which lose an arbitrarily small ε) suggest the virtual value approach may be desirable if one cares about small losses.

2.1.3 Roadmap

Our main result tightly characterizes the competition complexity in the litte n regime, and tightly characterizes the competition complexity in the big n regime among proofs which use the same benchmark as [24].

In Section 3.2, we provide the necessary preliminaries surrounding the benchmark of [24] and virtual value theory. In Section 2.3 we provide a near-complete proof of our results when $n = 1$ as a warm-up. In Section 2.4, we analyze the benchmark and reduce the analysis to proving stochastic dominance of certain correlated random variables drawn from $[0, 1]$. In Section 2.5 we prove the required claims regarding stochastic dominance (which at this point are purely mathematical claims and no longer reference auctions). In Appendix A.1 we: (a) recap the lower bound of [26] witnessing that our results are tight in the little n regime, (b) provide a lower bound witnessing that our results are tight in the big n regime (among proofs which use the same benchmark as [24]), and (c) prove that the competition complexity of n bidders with additive valuations over m independent items approaches ∞ as $n \to \infty$.

2.2 Notation and Preliminaries

We consider a setting with n i.i.d. bidders with additive valuations over m independent items. That is, there are single-variate distributions D_j for all $j \in [m]$, and bidder i's value v_j for item j is drawn independently from D_j. Bidder i's value for the bundle S is just $\sum_{j \in S} v_j$. We will use the following notation:

- $\textsc{Rev}_n(D)$ to denote the revenue of the optimal (possibly randomized) Bayesian Incentive Compatible[3] mechanism when played by n bidders whose values for m items are drawn from D. In our setting, we will always have $D = \times_j D_j$ for some single-variate distributions D_j.
- $\text{VCG}_n(D)$ to denote the revenue achieved by the VCG mechanism when played by n bidders whose values for m items are drawn from D. In our setting, the VCG mechanism simply runs a second-price auction on each item separately with no reserve.
- $\text{S}\textsc{Rev}_n(D)$ to denote the revenue achieved by Myerson's mechanism run separately on each item, when played by n bidders whose values for m items are drawn from D. Note that for all n and distributions D over additive valuations, $\text{S}\textsc{Rev}_n(D) \geq \text{VCG}_n(D)$.

---

[3] A mechanism is Bayesian Incentive Compatible if it is in every bidder's interest to bid truthfully, conditioned on all other bidders bidding truthfully as well. That is, assuming that all other bidders submit bids drawn from D_{-i}, bidder i best responds by bidding their true values.

2.2.1 Myerson's Lemma, Bulow-Klemperer, and Virtual Values

Here, we briefly recap basic facts about the theory of virtual values due to Myerson [49]. We include some proofs and sketches in Appendix A.3, and refer the reader to [40] (Definition 3.11) for a deeper treatment of these concepts (or [13], Definition 8 for discrete distributions). Note that much of the theory extends to independent (but non-i.i.d.) bidders with slightly more complex statements. As we only consider i.i.d. bidders, we omit the extra notation. Below, when we write X^+ for a random variable X, we mean $\max\{X, 0\}$.

Definition 1 (Virtual Values and Ironing [49]). *For a continuous single-variate distribution with CDF $F(\cdot)$ and PDF $f(\cdot)$, the virtual valuation function $\varphi_F(\cdot)$ satisfies $\varphi_F(v) = v - \frac{1-F(v)}{f(v)}$. If $\varphi_F(\cdot)$ is monotone non-decreasing, F is said to be* regular. *If not, $\overline{\varphi}_F(\cdot)$ is the* ironed virtual value *function (see [40] for a formal definition), and is monotone non-deceasing. When F is regular, $\overline{\varphi}_F(\cdot) = \varphi_F(\cdot)$.*

Theorem 1 ([49]). *Let D be any single-variate distribution. Then for all n:*

$$\mathrm{SR{\small EV}}_n(D) = \mathrm{R{\small EV}}_n(D) = \mathbb{E}_{\vec{v} \leftarrow D^n}\left[\left(\overline{\varphi}_D\left(\max_{i\in[n]}\{v_i\}\right)\right)^+\right],$$

$$\mathrm{VCG}_n(D) = \mathbb{E}_{\vec{v} \leftarrow D^n}\left[\left(\varphi_D\left(\max_{i\in[n]}\{v_i\}\right)\right)\right].$$

Fact 1. *For any single-variate distribution D, and any value v, let $D_{\geq v}$ denote the distribution D conditioned on exceeding v. Then $v = \mathbb{E}_{w \leftarrow D_{\geq v}}[\varphi(w)] \leq \mathbb{E}_{w \leftarrow D_{\geq v}}[\overline{\varphi}(w)]$.*

Finally, we recall the seminal result of Bulow and Klemperer [12]:

Theorem 2 ([12]). *For any regular single-variate distribution D, $\mathrm{VCG}_{n+1}(D) \geq \mathrm{R{\small EV}}_n(D)$.*

2.2.2 Duality Benchmarks

Here we state an upper bound on $\textsc{Rev}_n(D)$ when D is additive over independent items. The bound is derived using the duality framework of Cai et al. [13], and first used by Eden et al. [24] (it is also possible to derive this particular bound without duality [18]). When referring to this benchmark in text, we call it the EFFTW[4] benchmark. Parsing the benchmark requires additional notation:

- v_{ij} denotes the value of bidder i for item j.
- D_j denotes the marginal distribution of item j. We use $\overline{\varphi}_j(\cdot)$ to denote $\overline{\varphi}_{D_j}(\cdot)$.
- For a variable X, if X has no point-masses, then we simply define $F(x) = \Pr[X < x] = \Pr[X \leq x]$. If $X = x$ with strictly positive probability, then we define $F(x)$ to be a random variable drawn uniformly from the interval $[\Pr[X < x], \Pr[X \leq x]]$. Importantly, note that the random variable $F(X)$ is drawn uniformly from $[0, 1]$ for any random variable X.
- For a distribution $D := \times_j D_j$, we partition the space $\mathbb{R}_+^m$ into m disjoint *regions*. For each $j \in [m]$, we define $R_j := \{\vec{v}_i \in \mathbb{R}_+^m \mid F_j(v_{ij}) > F_k(v_{ik})\ \forall k \neq j\}$. That is, $\vec{v}_i$ is in region R_j if item j has the highest *quantile*. Observe that his partition may be randomized if D has point masses (and is deterministic with probability 1 if D has no point masses).

Theorem 3 ([13, 24]). *Let D be additive over m independent items. Then:*

$$\textsc{Rev}_n(D) \leq \sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow D^n}\left[\max_{i \in [n]} \left\{\overline{\varphi}_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) + v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\right\}\right].$$

---

[4] The acronym stands for the authors' last names in [24].

If we think of the *Virtual Value* of bidder i for item j as equal to Myerson's ironed virtual value, $\overline{\varphi}_j(v_{ij})^+$, if item j has the highest quantile in $\vec{v}_i$, and equal to the value, v_{ij}, if not, then Theorem 3 claims that the expected revenue of the optimal mechanism does not exceed the sum over all items of the expected maximum virtual value for that item. Theorem 3 is an application of Corollary 28 in [13], together with the observation that our defined regions are upwards-closed.

2.3 Warm-Up: Single Bidder

In this section, we illustrate one portion of our improved anlaysis via the single bidder setting. This will also help identify one significant point of departure from [24]. Observe that the EFFTW benchmark simplifies significantly for a single bidder, as there is only one element of $[n]$, and the benchmark simply sums the virtual value of the item with the highest quantile plus the values of all other items.

2.3.1 Brief Recap of [24]

The main idea in the single-bidder approach of [24] is to couple draws of m bidders for item j with draws of a single bidder for m items via their quantiles. Specifically, they observe the following: consider fixed quantiles $q_1, \dots, q_m$ drawn independently. and uniformly from $[0, 1]$.

- **Benchmark Analysis:** Use the quantiles drawn to determine values for each of m items. If q_j is the largest quantile drawn, then item j contributes $\overline{\varphi}_j(F_j^{-1}(q_j))^+$ to the benchmark. If q_j is not the largest quantile drawn, then item j contributes $F_j^{-1}(q_j)$ to the benchmark.

- **VCG Analysis:** Use the quantiles drawn to determine values of each of m bidders for item j. If q_j is the largest quantile drawn, then bidder j contributes $\overline{\varphi}_j(F_j^{-1}(q_j))$ to the virtual surplus of VCG. If q_j is not the largest quantile drawn, then some other bidder wins the item and pays at least $F^{-1}(q_j)$, so at least $F^{-1}(q_j)$ is contributed by some bidder $\neq j$ to the revenue.

The above reasoning is not far from a formal proof that $\text{SR}\textsc{ev}_m(D) \geq \text{R}\textsc{ev}_1(D)$. Some care is required to make sure Theorem 1 is applied correctly (since we wish to count bidder j's contribution to the revenue of VCG using her ironed virtual value but the other bidders' contributions directly via payments), but the above reasoning is the key step. The main idea is that if we couple the quantiles drawn for the benchmark with quantiles drawn for selling just item j, then the revenue achieved from selling just item j to m bidders drawn from D_j exceeds the contribution of item j to the benchmark *for all quantiles drawn*.

2.3.2 Our Analysis

The main challenge that the previous analysis overcomes is the following: the contribution of item j to the benchmark is sometimes in the form of a virtual value, and sometimes in the form of a value. There is no "natural" random variable that takes exactly this form, and it is tricky to analyze directly. So the previous analysis finds a clever way to "recreate" it using this coupling argument. Unfortunately though, direct coupling arguments like this should not hope to prove a competition complexity better than $m-1$, as there are m random variables that need to be coupled.

Our approach instead is to reason about the contribution of item j to the benchmark *exclusively in terms of virtual values*, using Fact 1. Specifically, consider the following proposition, which rewrites the contribution of item j to the benchmark. Below, $X_L(1,m)$ denotes the following random variable: first, draw one quantile $X_{1,1}$ uniformly at random from $[0,1]$. Then, draw $m-1$ quantiles uniformly at random from $[0,1]$ and label them $Y_{1,m-1}$ thru $Y_{m-1,m-1}$. If $X_{1,1} > Y_{\ell,m-1}$ for all ℓ, then set $X_L(1,m) = X_{1,1}$. Otherwise, let ℓ^* denote a uniformly random element from $\{\ell \mid Y_{\ell,m-1} > X_{1,1}\}$ and set $X_L(1,m) = Y_{\ell^*,m-1}$.

Proposition 1. *For all $D = \times_j D_j$ and all items j,*

$$\mathbb{E}_{\vec{v}\leftarrow D}\left[\overline{\varphi}_j(v_j)^+ \cdot \mathbb{I}(\vec{v} \in R_j) + v_j \cdot \mathbb{I}(\vec{v} \notin R_j)\right] \leq \mathbb{E}[\overline{\varphi}_j(F_j^{-1}(X_L(1,m)))^+].$$

Proof. The main idea is to get a lot of mileage from Fact 1: ideally, any time $\vec{v} \notin R_j$, rather than contribute v_j to the benchmark, we will contribute the virtual value of a random draw from D_j conditioned on exceeding v_j. To begin, let's couple quantiles drawn for the benchmark with quantiles drawn for the experiment defining $X_L(1,m)$ so that $X_{1,1} = F_j(v_j)$ and $Y_{\ell,m-1} = F_\ell(v_\ell)$ for $\ell < m, \ell \neq j$, and $Y_{j,m-1} = F_m(v_m)$ (if $j \neq m$, otherwise there is no $Y_{m,m-1}$ to define). Observe that indeed the quantiles are all drawn independently and uniformly from $[0,1]$. Moreover, we have:

- Whenever $\vec{v} \in R_j$, $X_L(1,m) = X_{1,1} = F_j(v_j)$. Therefore, we conclude that:

$$\mathbb{E}_{\vec{v}\leftarrow D}\left[\overline{\varphi}_j(v_j)^+ \cdot \mathbb{I}(\vec{v} \in R_j)\right] = \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_L(1,m))^+ \cdot \mathbb{I}(X_L(1,m) = X_{1,1})\right]. \tag{2.1}$$

- Conditioned on $\vec{v} \notin R_j$, $X_L(1,m)$ is a uniformly random sample from $[X_{1,1}, 1]$. This is because there is some strictly positive number of ℓs such that $Y_{\ell,m-1} > X_{1,1}$. Conditioned on being $> X_{1,1}$, each such value is drawn uniformly from

$[X_{1,1}, 1]$. And then $X_L(1,m)$ picks one of them uniformly at random. Using Fact 1, we therefore conclude that:

$$\begin{aligned}\mathbb{E}_{\vec{v}\leftarrow D}\left[v_j \cdot \mathbb{I}(\vec{v} \notin R_j)\right] &\leq \mathbb{E}_{\vec{v}\leftarrow D}\left[\mathbb{E}_{x\leftarrow D_{j,\geq v_j}}\left[\overline{\varphi}_j(x)\right] \cdot \mathbb{I}(\vec{v} \notin R_j)\right] \\ &\leq \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_L(1,m))) \cdot \mathbb{I}(X_L(1,m) \neq X_{1,1})\right]. \quad (2.2)\end{aligned}$$

It is now easy to see that the left-hand sides of the two equations sum together to yield item j's contribution to the benchmark, while the two right-hand sides sum together to yield (at most) $\mathbb{E}[\overline{\varphi}_j(F_j^{-1}(X_L(1,m)))^+]$, proving the proposition. □

Proposition 1 gives an upper bound on the contribution of item j to the benchmark written as the expectation of a virtual value of *some* distribution $(F^{-1}(X_L(1,m)))$. This is convenient because we can write the revenue achieved by using Myerson's optimal auction for selling item j to $1+c$ bidders as the expectation of a virtual value of another distribution (the maximum of $1+c$ draws from D_j). Therefore, if we can relate these two distributions (for instance, by proving that one stochastically dominates the other), we can relate these two expectations. Below, let $X_S(1,c)$ denote the maximum of $1+c$ i.i.d. draws from the uniform distribution on $[0,1]$.

Corollary 1. *If $X_S(1,c)$ stochastically dominates $X_L(1,m)$, then for all D that are additive over m independent items,* $\text{SRev}_{1+c}(D) \geq \text{Rev}_1(D)$.

Proof. Observe first that by Theorem 1 we have:

$$\text{SRev}_{1+c}(D) = \sum_j \mathbb{E}_{\vec{x}\leftarrow D_j^{1+c}}\left[\overline{\varphi}_j\left(\max_{i\in[1+c]}\{x_i\}\right)^+\right] = \sum_j \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_S(1,c))^+\right].$$

By Proposition 1 (and Theorem 3), we have:

$$\begin{aligned}\textsc{Rev}_1(D) &\leq \sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow D}\left[\overline{\varphi}_j(v_j)^+ \cdot \mathbb{I}(\vec{v} \in R_j) + v_j \cdot \mathbb{I}(\vec{v} \notin R_j)\right] \\ &\leq \sum_{j=1}^{m} \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_L(1,m)))^+\right].\end{aligned}$$

Observe that $\overline{\varphi}_j(\cdot)$ is a monotone non-decreasing function, and F_j^{-1} is also monotone non-decreasing. As such, if $X_S(1,c)$ stochastically dominates $X_L(1,m)$, $\overline{\varphi}_j(F_j^{-1}(X_S(1,c)))$ stochastically dominates $\overline{\varphi}_j(F_j^{-1}(X_L(1,m)))$, which allows us to conclude that $\sum_j \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_S(1,c))^+\right] \geq \sum_{j=1}^{m} \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_L(1,m)))^+\right]$. Therefore, we may conclude that if $X_S(1,c)$ stochastically dominates $X_L(1,m)$, $\textsc{SRev}_{1+c}(D) \geq \textsc{Rev}_1(D)$. $\square$

At this point, we've reduced the problem of deriving competition complexity upper bounds to a purely mathematical problem relating stochastic dominance of $X_S(1,c)$ and $X_L(1,m)$. The proof of this claim for $n=1$ is not an especially instructive special case, so we defer the final step to Section 2.5. So we wrap up our warm-up by citing Theorem 8:

Corollary 2 (of Theorem 8)**.** *When $c \geq 2+\ln(m+1)$, $X_S(1,c)$ stochastically dominates $X_L(1,m)$.*

Theorem 4. *Let D be a distribution that is additive over m independent items. Then $\textsc{SRev}_{2+\ln(m+1)}(D) \geq \textsc{Rev}_1(D)$. If each D_j is regular, then also $\mathrm{VCG}_{3+\ln(m+1)}(D) \geq \textsc{Rev}_1(D)$.*

Proof. Theorem 3 upper bounds $\textsc{Rev}_1(D)$ with the EFFTW benchmark. Proposition 1 further upper bounds the EFFTW benchmark with $\sum_j \mathbb{E}[\overline{\varphi}_j(F_j^{-1}(X_L(1,m)))^+]$, which is the sum over all items of the expected virtual value of a quantile drawn from $X_L(1,m)$. Corollary 1 argues that if $X_L(1,m)$ is stochastically dominated by $X_S(1,c)$ (the maximum of $c+1$ i.i.d. draws uniformly from $[0,1]$), then we may replace $X_L(1,m)$ with $X_S(1,c)$ in the upper bound, which is exactly $\textsc{SRev}_{1+c}(D)$. Finally, Corollary 2 claims that indeed $X_S(1,c)$ stochastically dominates $X_L(1,m)$ when $c \geq 2 + \ln(m+1)$ (and the final $+1$ when each D_j is regular comes from going from SRev to VCG using Bulow-Klemperer). □

This concludes our exposition for a single bidder. Above we introduced one new idea which departs from prior work: instead of directly treating the benchmark which involves both values and virtual values, rewrite the benchmark to involve only virtual values and reduce the problem to purely mathematical questions about stochastic dominance of $X_L(1,m)$ and $X_S(1,c)$.

2.4 Multiple Bidders

In this section, we overview our approach for the general case. The key simplifying feature of the single-bidder case that allowed us to isolate one key idea is that for each item j, that item has the highest quantile or it doesn't. In the multi-bidder case, there are multiple bidders, some of whom will have their highest quantile for item j, some of whom won't. So we must actually engage with the "$\max_{i\in[n]}$" in the benchmark. Our approach will be different depending on whether n is big or little relative to m. We begin with the little n case as it is more similar to the single-bidder case.

2.4.1 Part One: When $n \leq m$

Our key step is conceptually similar to Proposition 1, but the random variables involved are necessarily more complex. We first make the following observation (also made in [24]). Below, $v_{(\ell)j}$ denotes the ℓ^{th} highest value for item j (among all bidders). All omitted proofs can be found in Appendix A.4.

Observation 1.

$$\begin{aligned}&\mathbb{E}_{\vec{v}\leftarrow D^n}\left[\max_{i\in[n]}\left\{\overline{\varphi}_j(v_{ij})^+\cdot\mathbb{I}(\vec{v}_i\in R_j)+v_{ij}\cdot\mathbb{I}(\vec{v}_i\notin R_j)\right\}\right]\\&\leq\mathbb{E}_{\vec{v}\leftarrow D^n}\left[\max\left\{v_{(1)j}\cdot\mathbb{I}(\vec{v}_{(1)}\notin R_j),\overline{\varphi}_j(v_{(1)j}),v_{(2)j}\right\}\right].\end{aligned}$$

Next, we want to rewrite the right-hand side above using random variables simliar to $X_L(1,m)$. This time, let $X'_L(n,m)$ denote the following random variable: first, draw n quantiles $X_{1,n},\dots,X_{n,n}$ independently and uniformly at random from $[0,1]$. Relabel them so that $X_{(1),n}\geq\dots\geq X_{(n),n}$. Then, draw $m-1$ quantiles uniformly at random from $[0,1]$ and label them $Y_{1,m-1}$ thru $Y_{m-1,m-1}$. If $X_{(1),n}>Y_{\ell,m-1}$ for all ℓ, then set $X'_L(n,m)=X_{(1),n}$. Otherwise, let ℓ^* be a uniformly random element from $\{\ell|Y_{\ell,m-1}>X_{(1),n}\}$ and set $X'_L(n,m)=Y_{\ell^*,m-1}$.

Proposition 2. *For all $D=\times_j D_j$, and all items j:*

$$\begin{aligned}&\mathbb{E}_{\vec{v}\leftarrow D^n}\left[\max\left\{v_{(1)j}\cdot\mathbb{I}(\vec{v}_{(1)}\notin R_j),\overline{\varphi}_j(v_{(1)j}),v_{(2)j}\right\}\right]\\\leq&\mathbb{E}\left[\max\left\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))),F_j^{-1}(X_{(2),n})\right\}\right].\end{aligned}$$

Proposition 2 helps us replace any instances of $v_{(1)n}$ in the benchmark with a randomly drawn virtual value, but we still need to do the same for $v_{(2)n}$ (which so far has essentially just been rewritten as $F_j^{-1}(F_j(v_{(2)n}))$). Now, let $W_{2,n}$ be a uniformly random draw from $[X_{(2),n},1]$, and define $X_L(n,m)=\max\{X'_L(n,m),W_{2,n}\}$. By making use of Fact 1, we can conclude:

Corollary 3.

$$\mathbb{E}\left[\max\left\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), F_j^{-1}(X_{(2),n})\right\}\right] \leq \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_L(n,m)))\right].$$

Now, we are nearly ready to wrap up the $n \leq m$ case. Similarly to the single-bidder case, define $X_S(n,c)$ to be the maximum of $n+c$ i.i.d. draws uniformly form $[0,1]$.

Corollary 4. *If $X_S(n,c)$ stochastically dominates $X_L(n,m)$, then* $\text{SRev}_{n+c}(D) \geq \text{Rev}_n(D)$*. If each D_j is regular, then* $\text{VCG}_{n+c}(D) \geq \text{Rev}_n(D)$*.*

Finally, Theorem 8 claims that when $c \geq n \cdot (2 + \ln(1 + m/n))$, $X_S(n,c)$ indeed stochastically dominates $X_L(n,m)$. Combining Corollary 4 with Theorem 8 therefore concludes:

Theorem 5. *For all D that are additive over m independent items,* $\text{SRev}_{n+n\cdot(2+\ln(1+m/n))}(D) \geq \text{Rev}_n(D)$*. If each marginal of D_j is regular, then* $\text{VCG}_{n+n\cdot(2+\ln(1+m/n))}(D) \geq \text{Rev}_n(D)$*.*

When $n \leq m$, this is tight up to constant factors, due to a lower bound of [26] (see Appendix A.1 for the construction). But when $n \geq m$, this is still linear in n. We therefore provide an alternative argument in the following section which achieves the optimal (up to constant factors) competition complexity *that is achievable starting from the EFFTW benchmark* of $\Theta(\sqrt{nm})$.

2.4.2 Part Two: When $n \geq m$

At a high level, the main difference between how we should analyze the $n \leq m$ case and the $n \geq m$ is as follows: Observation 1 immediately upper bounds the

EFFTW by upper bounding $\overline{\varphi}_j(v_{(2)j})^+ \cdot \mathbb{I}(\vec{v}_{(2)} \in R_j) + v_{(2)j} \cdot \mathbb{I}(\vec{v}_{(2)} \notin R_j)$ with $v_{(2)j}$. When $n \leq m$, this upper bound is unlikely to be much of a relaxation, because it's likely that $v_{(1)j} \notin R_j$ anyway. But when $n \gg m$, we're extremely unlikely to have $v_{(1)j} \notin R_j$, and this upper bound is wasteful. Indeed, this step is what limits the analysis in [24] to $\Omega(n)$. The first step for the $n \geq m$ case is to address this.

Proposition 3. *For all items j, all $\ell \in [n]$, and all distributions D that are additive over independent items:*

$$\begin{aligned}&\mathbb{E}_{\vec{v} \leftarrow D^n}\left[\max_{i \in [n]} \left\{\overline{\varphi}_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) + v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\right\}\right] \\ \leq &\mathbb{E}_{\vec{w} \leftarrow D_j^{n+(m-1)(\ell-1)}}\left[\max\left\{\overline{\varphi}_j(w_{(1)}), w_{(\ell)}\right\}\right].\end{aligned}$$

We now want to take a simliar step to the previous case and replace $w_{(\ell)}$ with a randomly drawn virtual value using Fact 1. Here, define the random variable $X_B(n,\ell)$ as follows. First, draw $X_{1,n}, \ldots, X_{n,n}$ independently and uniformly at random from $[0,1]$. Then, randomly draw $W_{\ell,n}$ uniformly from $[X_{(\ell),n}, 1]$, and set $X_B(n,\ell) := \max\{X_{(1),n}, W_{\ell,n}\}$.

Lemma 1. *For any single-dimensional distribution D, and any n':*

$$\mathbb{E}_{\vec{w} \leftarrow D^{n'}}\left[\max\left\{\overline{\varphi}_j(w_{(1)}), w_{(\ell)}\right\}\right] \leq \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_B(n',\ell)))\right].$$

Corollary 5 below follows from Proposition 3 and Lemma 1 with $n' := n+(m-1)(\ell-1)$.

Corollary 5. *If $X_S(n,c)$ stochastically dominates $X_B(n+(\ell-1)(m-1),\ell)$, then for any D that is additive over m independent items,* $\text{SRev}_{n+c}(D) \geq \text{Rev}_n(D)$. *If each marginal D_j is regular, then* $\text{VCG}_{n+c}(D) \geq \text{Rev}_n(D)$.

Finally, Theorem 7 states that $X_S(n,c) = X_S(n+(\ell-1)(m-1), c-(\ell-1)(m-1))$

stochastically dominates $X_B(n+(\ell-1)(m-1),\ell)$ whenever $c-(\ell-1)(m-1) \geq \frac{4n+4(\ell-1)(m-1)}{\ell-1}$. Setting $\ell := \sqrt{n}m+1$, we get $c \geq \sqrt{nm}+4\sqrt{nm}+4(m-1)$.

Theorem 6. *For all D that are additive over m independent items, $\text{SRev}_{n+5\sqrt{nm}+4(m-1)}(D) \geq \text{Rev}_n(D)$. If each marginal D_j is regular, then $\text{VCG}_{n+5\sqrt{nm}+4(m-1)}(D) \geq \text{Rev}_n(D)$. In particular, if $n \geq m$, $5\sqrt{nm}+4(m-1) \leq 9\sqrt{nm}$.*

2.5 Stochastic Dominance via Additional Samples

In this section, we consider purely questions about whether one distribution stochastically dominates another (Sections 2.3 and 2.4 outline the connection between these problems and our main result). Recall the following ingredients in our experiments:

- $X_{1,n},\dots,X_{n,n}$ are n i.i.d. draws from the uniform distribution on $[0,1]$, and then relabeled so that $X_{(1),n} \geq \dots \geq X_{(n),n}$.
- $Y_{1,m-1},\dots,Y_{m-1,m-1}$ are $m-1$ i.i.d. draws from the uniform distribution on $[0,1]$, and then relabeled so that $Y_{(1),m-1} \geq \dots \geq Y_{(m-1),m-1}$.
- $Z_{1,c},\dots,Z_{c,c}$ are c i.i.d. draws from the uniform distribution on $[0,1]$, and then relabeled so that $Z_{(1),c} \geq \dots \geq Z_{(c),c}$.
- $W_{\ell,n}$ is a random draw from the uniform distribution on $[X_{(\ell),n},1]$.

SRev Experiment(n,c): Output $X_S(n,c) := \max\{X_{(1),n}, Z_{(1),c}\}$.

Big n Benchmark Experiment(n, ℓ): Output $X_B(n,\ell) := \max\{X_{(1),n}, W_{\ell,n}\}$.

Little n Benchmark Experiment(n, m): Let j^* be the largest index such that $Y_{(j^*),m-1} > X_{(1),n}$ (if such a j^* exists). If no such j^* exists, output $X_L(n,m) := \max\{X_{(1),n}, W_{2,n}\}$. Otherwise, pick an index j uniformly at random from $\{1,\ldots,j^*\}$ and output $\max\{Y_{(j),m-1}, W_{2,n}\}$.

The main results of this section are as follows:

Theorem 7. *When $c \geq 4n/(\ell-1)$, $X_S(n,c)$ stochastically dominates $X_B(n,\ell)$.*

Theorem 8. *When $c \geq n \cdot (2+\ln(1+m/n))$, $X_S(n,c)$ stochastically dominates $X_L(n,m)$.*

Intuitively, we might expect $X_S(n,c)$ to stochastically dominate $X_B(n,\ell)$ right around $c = 2n/\ell$. This is because $\mathbb{E}[Z_{(1),c}] = 1-1/(c+1)$, and $\mathbb{E}[W_{\ell,n}] = 1-\frac{\ell}{2(n+1)}$. Of course, this observation doesn't come close to proving stochastic dominance, especially because $X_{(1),n}$ and $W_{\ell,n}$ aren't independent. But it does give us an idea of the right ballpark to shoot for. The following proposition will be used in the proof of both theorems.

Proposition 4. *Let $c \geq 4n/(\ell-1)$. Then for all p, $\Pr[Z_{(1),c} > p | X_{(1),n} < p] \geq \Pr[W_{\ell,n} > p | X_{(1),n} < p]$. When $\ell = 2$, this can be improved to $c \geq n$.*

Before getting into the proof, let's unpack the role of conditioning on $X_{(1),n}$. $Z_{(1),c}$ and $X_{(1),n}$ are independent, so $\Pr[Z_{(1),c} > p | X_{(1),n} < p] = \Pr[Z_{(1),c} > p]$. On the other hand, $W_{\ell,n}$ and $X_{(1),n}$ are *positively* correlated: the lower bound on the range from which $W_{\ell,n}$ is drawn is $X_{(\ell),n}$, which is positively correlated with $X_{(1),n}$.

So certainly if we could prove the lemma without the conditioning on $X_{(1),n} < p$, the desired proposition would hold. This approach works for $\ell = 2$ (and indeed shows up in our proof as a base case), but without conditioning the conclusion is otherwise false for larger ℓ.

Proof. The proof will proceed by induction on n, ℓ. We begin with the base case, $\ell = 2$. $Z_{(1),c}$ is easy to reason about: $Z_{(1),c}$ is just the maximum of c i.i.d. draws uniformly from $[0, 1]$. So:

$$\Pr[Z_{(1),c} > p | X_{(1),n} < p] = \Pr[Z_{(1),c} > p] = 1 - p^c. \tag{2.3}$$

Now we turn to $W_{2,n}$. As referenced in the foreword to the proof, for this case the proposition statement holds even without conditioning on $X_{(1),n} < p$. Indeed, observe that without conditioning on $X_{(1),n} < p$, $X_{(2),n}$ is just the second-highest of n i.i.d. draws uniformly from $[0, 1]$, and $W_{2,n}$ is drawn uniformly from $[0, 1]$, but conditioned on exceeding $X_{(2),n}$. That is, $W_{2,n}$ is actually identically distributed to $X_{(1),n}$, and is distributed according to the maximum of n i.i.d. draws uniformly from $[0, 1]$. Therefore, when $c = n$, $W_{2,n}$ is identically distributed to $Z_{(1),c}$, and the conclusion holds. That is:

$$\Pr[W_{2,n} > p | X_{(1),n} < p] \leq \Pr[W_{2,n} > p] = 1 - p^n. \tag{2.4}$$

As such, we have proved the base case (in fact, a slightly stronger claim): for all n, and $\ell = 2$ when $c \geq n = n/(\ell - 1)$, $Z_{(1),c}$ stochastically dominates $W_{2,n}$. Now we turn to the inductive step, which is significantly more involved. As referenced in the foreword, we must take a different approach for larger ℓ, as the desired claim is false if we remove conditioning on $X_{(1),n} < p$.

To this end, we'll first observe that when $p = 1$, $\Pr[Z_{(1),c} > 1] = \Pr[W_{\ell,n} >$

$1|X_{(1),n} < 1] = 0$, and when $p \to 0$, $\Pr[Z_{(1),c} > p] = \Pr[W_{\ell,n} > p|X_{(1),n} < p] = 1$. So the desired inequalities hold at both endpoints of $[0,1]$, and we'd like to reason about $p \in (0,1)$. To accomplish this, it will actually be easier to compare $\Pr[Z_{(1),c} > p] \cdot \Pr[X_{(1),n} < p]$ to $\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]$ (observe that this simply multiplies both conditional probabilities in our original comparison by $\Pr[X_{(1),n} < p]$), and consider the derivative with respect to p.

So let $f_{1,n}(\cdot)$ denote the density of $X_{(1),n}$. Then $\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p] = \int_0^p f_{1,n}(q) \cdot \Pr[W_{\ell,n} > p|X_{(1),n} = q]dq$. By Leibniz' rule, the derivative of this with respect to p is:

$$\begin{aligned}
&\frac{\partial \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{\partial p} \\
=&\frac{\partial \int_0^p f_{1,n}(q) \cdot \Pr[W_{\ell,n} > p|X_{(1),n} = q]dq}{\partial p} \\
=&f_{1,n}(p) \cdot \Pr[W_{\ell,n} > p|X_{(1),n} = p] + \int_0^p f_{1,n}(q) \cdot \frac{\partial \Pr[W_{\ell,n} > p|X_{(1),n} = q]}{\partial p} dq.
\end{aligned}$$

Let's first unpack the left-most term with the following lemma.

Lemma 2. *For all $\ell, n > 1$, $\Pr[W_{\ell,n} > p|X_{(1),n} = p] = \Pr[W_{\ell-1,n-1} > p|X_{(1),n-1} < p]$.*

Proof. Observe that, conditioned on $X_{(1),n} = p$, $X_{(2),n}, \ldots, X_{(n),n}$ are $n-1$ (sorted) i.i.d. draws uniformly at random from $[0,p]$, and $X_{(\ell),n}$ is the $(\ell-1)^{th}$ highest of them. Put another way, $X_{(\ell),n}$ conditioned on $X_{(1),n} = p$ is *identically distributed* to $X_{(\ell-1),n-1}$ conditioned on $X_{(1),n-1} < p$. This therefore implies that $W_{\ell,n}$ conditioned on $X_{(1),n} = p$ and $W_{\ell-1,n-1}$ conditioned on $X_{(1),n-1} < p$ are identically distributed as well. □

Now we turn to the right-most term.

Lemma 3. *For all* p, ℓ, n, q, $\frac{\partial \Pr[W_{\ell,n} > p | X_{(1),n} = q]}{\partial p} = -\Pr[W_{\ell,n} > p | X_{(1),n} = q]/(1-p)$.

Proof. Let's first expand $\Pr[W_{\ell,n} > p | X_{(1),n} = q]$ by letting $f^q_{\ell,n}(\cdot)$ denote the density of $X_{(\ell),n}$ conditioned on $X_{(1),n} = q$.

$$\Pr[W_{\ell,n} > p | X_{(1),n} = q] = \int_0^q f^q_{\ell,n}(r) \cdot \frac{1-p}{1-r} dr.$$

This is simply because, conditioned on $X_{(\ell),n} = r$, the probability that $W_{\ell,n}$ (a uniformly random draw from $[r, 1]$) exceeds p is exactly $\frac{1-p}{1-r}$. Taking now the derivative with respect to p (again by Leibniz' rule), we see that:

$$\frac{\partial \Pr[W_{\ell,n} > p | X_{(1),n} = q]}{\partial p} = -\int_0^q f^q_{\ell,n}(r)/(1-r) dr = -\Pr[W_{\ell,n} > p | X_{(1),n} = q]/(1-p).$$

□

Using Lemma 3, we can now rewrite:

$$\begin{aligned}\int_0^p f_{1,n}(q) \cdot \frac{\partial \Pr[W_{\ell,n} > p | X_{(1),n} = q]}{\partial p} dq &= \frac{-1}{1-p} \int_0^p f_{1,n}(q) \cdot \Pr[W_{\ell,n} > p | X_{(1),n} = q] dq \\ &= \frac{-\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{1-p}\end{aligned}$$

And using both Lemmas 2 and 3 we can now simplify (the second equality follows by recalling that $f_{1,n}(\cdot)$ is the density of $X_{(1),n}$):

$$\frac{\partial \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{\partial p} \tag{2.5}$$

$$= f_{1,n}(p) \cdot \Pr[W_{\ell-1,n-1} > p | X_{(1),n-1} < p] - \frac{\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{1-p} \tag{2.6}$$

$$= np^{n-1} \cdot \Pr[W_{\ell-1,n-1} > p | X_{(1),n-1} < p] - \frac{\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{1-p} \tag{2.7}$$

From here we'll show that whenever $\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p] \geq \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p]$ (i.e. whenever what we're trying to prove at p is violated), then the derivative trends towards satisfying our desired claim.

Lemma 4. *If* $\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p] \geq \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p]$*, then:*

$$\frac{\partial \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{\partial p} \geq 0.$$

Proof. Observe first that $\Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] = p^n \cdot (1 - p^c)$. As such, we also have $\frac{\partial \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p]}{\partial p} = np^{n-1}(1-p^c) - cp^{n+c-1}$. So if the hypotheses of the lemma are satisfied, then by Equation (2.5) we can write:

$$\begin{aligned}
&\frac{\partial \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{\partial p} \\
=& np^{n-1}(1-p^c) - cp^{n+c-1} - np^{n-1} \cdot \Pr[W_{\ell-1,n-1} > p | X_{(1),n-1} < p] \\
&+ \frac{\Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{1-p} \\
\geq& np^{n-1}(1-p^c) - cp^{n+c-1} - np^{n-1} \cdot (1 - p^{4(n-1)/(\ell-2)}) + \frac{p^n(1-p^c)}{1-p}.
\end{aligned}$$

In the inequality, we have used two facts. First, we have used the inductive hypothesis, which claims that $\Pr[W_{\ell-1,n-1} > p | X_{(1),n-1} < p] \leq \Pr[Z_{(1),4(n-1)/(\ell-2)} > p | X_{(1),n-1} < p] = 1 - p^{4(n-1)/(\ell-2)}$. Second, we have used the hypothesis of the lemma statement. Next, we can substitute $(1-p^c) = (1 + p + \ldots + p^{c-1}) \cdot (1-p)$ (and make some other algebraic simplifications) to get:

$$\begin{aligned}
&\frac{\partial \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{\partial p} \\
&\geq -(n+c)p^{n+c-1} + np^{n-1+4(n-1)/(\ell-2)} + p^n \cdot \sum_{j=0}^{c-1} p^j.
\end{aligned}$$

Recall again that we are hoping to prove that the above term is ≥ 0. As $p \geq 0$, the above term is ≥ 0 if and only if (dividing all terms by p^{n-1} and rearranging):

$$np^{4(n-1)/(\ell-2)} + \sum_{j=1}^{c} p^j \geq (n+c)p^c. \tag{2.8}$$

To conclude that Equation (2.8) indeed holds for all $p \in [0,1]$, consider the convex function $f(x) := p^x$, and the random variable A where $A = j$ with probability $1/(n+c)$ for all $j \in \{1,\dots,c\}$, and $A = 4(n-1)/(\ell-2)$ with probability $n/(n+c)$. Then the left-hand side of the equation above is exactly $(n+c)\cdot\mathbb{E}[f(A)]$. Therefore, we may conclude by Jensen's inequality that the left-hand side exceeds $(n+c)\cdot f(\mathbb{E}[A]) = (n+c)\cdot p^{\frac{4(n-1)n/(\ell-2)+c(c+1)/2}{n+c}}$. As $p \in [0,1]$, $(n+c)\cdot p^{\frac{4(n-1)n/(\ell-2)+c(c+1)/2}{n+c}} \geq (n+c)p^c$ if and only if $\frac{4(n-1)n/(\ell-2)+c(c+1)/2}{n+c} \leq c$. So finally, our only remaining job is to see what values of c satisfy:

$$\frac{4(n-1)n}{\ell-2} + \frac{c(c+1)}{2} \leq nc + c^2. \tag{2.9}$$

Indeed, observe that when $c = 4n/(\ell-1)$, we get:

$$\begin{gathered}
\frac{4(n-1)n}{\ell-2} + \frac{c(c+1)}{2} \overset{?}{\leq} nc + c^2 \\
\frac{4}{\ell-2} - \frac{4}{n(\ell-2)} + \frac{4^2}{2(\ell-1)^2} + \frac{4}{2n(\ell-1)} \overset{?}{\leq} \frac{4}{\ell-1} + \frac{4^2}{(\ell-1)^2} \\
2(\ell-1)^2 - 2(\ell-1)^2/n + 4(\ell-2) + (\ell-2)(\ell-1)/n \overset{?}{\leq} 2(\ell-2)(\ell-1) + 2\cdot 4\cdot(\ell-2) \\
2(\ell-1) - \ell(\ell-1)/n \overset{?}{\leq} 4\cdot(\ell-2) \\
2\frac{\ell-1}{\ell-2} - \frac{\ell(\ell-1)}{(\ell-2)n} \overset{?}{\leq} 4
\end{gathered}$$

The last inequality indeed holds as $\ell \geq 3$.

$\square$

Now we are ready to wrap up the proof of the proposition. We have just shown (Lemma 4) that for all p, either $\Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p] \geq 0$, or $\frac{\partial \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p]}{\partial p} - \frac{\partial \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]}{\partial p} \geq 0$. We now are ready to claim that this implies that $\Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p] \geq 0$ for all $p \in [0, 1]$.

Indeed, define $G(p) := \Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p]$. Then we have shown that for all p, either $G(p) \geq 0$ or $G'(p) \geq 0$. Moreover, we know that $G(0) = 0 - 0 = 0$. So assume for contradiction that there exists some p with $G(p) < 0$. Then because $G(\cdot)$ is continuous (and $G(0) = 0$), there exists some open interval (q, p) such that $G(x) < 0$ on (q, p), while $G(q) = 0$. But now we have a contradiction: By Lemma 4, $G'(x) \geq 0$ on (q, p), and $G(q) = 0$. Therefore, we must also have $G(x) \geq 0$ on (q, p), contradicting our initial assumption.

Therefore, we cannot have $G(x) < 0$ anywhere on $[0, 1]$, meaning that $\Pr[Z_{(1),c} > p \wedge X_{(1),n} < p] - \Pr[W_{\ell,n} > p \wedge X_{(1),n} < p] \geq 0$. This is identical to the claim that $\Pr[Z_{(1),c} > p | X_{(1),n} < p] \geq \Pr[W_{\ell,n} > p | X_{(1),n} < p]$. This completes the proof of the proposition.

□

Theorem 7 now follows nearly directly from Proposition 4.

Proof of Theorem 7. We can directly compute $\Pr[X_S(n, c) > p] = \Pr[X_{(1),n} > p] + \Pr[Z_{(1),c} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p]$. Similarly, $\Pr[X_B(n, \ell) > p] = \Pr[X_{(1),n} > p] + \Pr[W_{\ell,n} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p]$. By Proposition 4, when $c \geq 4n/(\ell - 1)$, we get:

$$\Pr[Z_{(1),c} > p | X_{(1),n} < p] \geq \Pr[W_{\ell,n} > p | X_{(1),n} < p].$$

Therefore,

$$\begin{aligned}[X_{(1),n} > p] + \Pr[Z_{(1),c} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p] \\ \geq \Pr[X_{(1),n} > p] + \Pr[W_{\ell,n} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p].\end{aligned}$$

This implies

$$\Pr[X_S(n,c) > p] \geq \Pr[W_{\ell,n} > p].$$

As the above holds for all $p \in [0,1]$, this proves that $X_S(n,c)$ stochastically dominates $X_B(n,\ell)$. □

Theorem 8 will require one more similar proposition.

Proposition 5. *Let* $c \geq n \cdot (1 + \ln(1 + m/n))$, *then for all* p, $\Pr[Z_{(1),c} > p | X_{(1),n} < p] \geq \Pr[Y^*_{n,m-1} > p | X_{(1),n} < p]$, *where* $Y^*_{n,m-1}$ *is the random variable equal to* 0 *if* $Y_{1,m-1} < X_{(1),n}$, *and is otherwise equal to* $Y_{j,m-1}$ *for a uniformly random* $j \in \{j | Y_{j,m-1} > X_{(1),n}\}$.

Proof. The proof of Proposition 5 is more direct than that of Proposition 4. This time, we can just directly compute $\Pr[Y^*_{n,m-1} > p | X_{(1),n} < p]$. We again begin by observing that $\Pr[Z_{(1),c} > p | X_{(1),n} < p] = \Pr[Z_{(1),c} > p] = 1 - p^c$.

We now turn to $Y^*_{n,m-1}$. Observe first that $Y^*_{n,m-1} = 0$, conditioned on $X_{(1),n} = q$, with probability exactly q^{m-1}. This is because $Y^*_{n,m-1}$ is 0 whenever each of $m-1$ i.i.d. draws uniformly from $[0,1]$ are all $< q$. Now, conditioned on $Y^*_{n,m-1} > 0$ (and also $X_{(1),n} = q$), observe that $Y^*_{n,m-1}$ is just a random draw from the uniform distribution on $[X_{(1),n}, 1]$. This is because, conditioned on $Y^*_{n,m-1} > 0$ and $X_{(1),n} = q$, $Y^*_{n,m-1}$ simply picks uniformly at random among some number of i.i.d. random variables drawn uniformly from $[X_{(1),n}, 1]$. Therefore, conditioned on $Y^*_{n,m-1} > 0$, and $X_{(1),n} = q$, $Y^*_{n,m-1}$ exceeds p with probability exactly $\frac{1-p}{1-q}$. Therefore, we can compute (below, let $f_1(\cdot)$ denote the density of $X_{(1),n}$, and $F_1(\cdot)$ denote the CDF):

$$\begin{aligned}
&\Pr[Y^*_{n,m-1} > p | X_{(1),n} < p] \\
=&\int_0^p \frac{f_1(q)}{F_1(p)} \cdot \Pr[Y^*_{n,m-1} > 0 | X_{(1),n} = q] \cdot \Pr[Y^*_{n,m-1} > p | Y^*_{n,m-1} > 0 \wedge X_{(1),n} = q] dq \\
=&\int_0^p \frac{nq^{n-1}}{p^n} \cdot (1 - q^{m-1}) \cdot \frac{1-p}{1-q} dq \\
=&\frac{n(1-p)}{p^n} \cdot \int_0^p q^{n-1} \cdot \frac{1-q^{m-1}}{1-q} dq \\
=&\frac{n(1-p)}{p^n} \cdot \int_0^p q^{n-1} \cdot (\sum_{i=0}^{m-2} q^i) dq \\
=&\frac{n(1-p)}{p^n} \cdot \left[\sum_{i=0}^{m-2} q^{i+n}/(n+i) \right]_0^p \\
=&\frac{n(1-p)}{p^n} \cdot \sum_{i=0}^{m-2} p^{i+n}/(n+i) \\
=&n(1-p) \cdot \sum_{i=0}^{m-2} p^i/(n+i) \\
=&1 - np^{m-1}/(n+m-2) + n \cdot \sum_{i=1}^{m-2} p^i (1/(n+i) - 1/(n+i-1)) \\
=&1 - np^{m-1}/(n+m-2) - \sum_{i=1}^{m-2} np^i/(n+i)(n+i-1)
\end{aligned}$$

Before proceeding, we quickly observe that the sums of the coefficients of non-zero powers of p is -1 (that is, $n/(n+m-2) + \sum_{i=1}^{m-2} n/(n+i)(n+i-1) = 1$). This is because the third-from-the-bottom equality is clearly equal to 0 when $p = 1$, and so the bottom equality must be equal to 0 when $p = 1$ as well.

$$\begin{aligned}\Pr[Z_{(1),c} > p | X_{(1),n} < p] &\overset{?}{\geq} \Pr[Y^*_{n,m-1} > p | X_{(1),n} < p] \\ 1 - p^c &\overset{?}{\geq} 1 - np^{m-1}/(n+m-2) - \sum_{i=1}^{m-2} np^i/(n+i)(n+i-1) \\ p^c &\overset{?}{\leq} np^{m-1}/(n+m-2) + \sum_{i=1}^{m-2} np^i/(n+i)(n+i-1)\end{aligned}$$

From here, we again apply Jensen's inequality. Let $f(x) := p^x$ (which is convex), and let A denote the random variable which is equal to $m-1$ with probability $n/(n+m-2)$ and equal to i with probability $n/(n+i)(n+i-1)$ for all $i \in \{1, \ldots, m-2\}$. By reasoning in the previous paragraph (that the coefficients of non-zero powers of p sum to -1), this is indeed a distribution. Then Jensen's inequality asserts that $\mathbb{E}[f(A)] \geq f(\mathbb{E}[A])$. Moreover, the right-hand side above is exactly $\mathbb{E}[f(A)]$. As such, we get that:

$$np^{m-1}/(n+m-2) + \sum_{i=1}^{m-2} np^i/(n+i)(n+i-1) \geq p^{n(m-1)/(n+m-2)+\sum_{i=1}^{m-2} ni/(n+i)(n+i-1)}.$$

As $p \in [0,1]$, this means that our desired inequality is satisfied as long as $c \geq n(m-1)/(n+m-2) + \sum_{i=1}^{m-2} ni/(n+i)(n+i-1)$. But now observe that:

$$\begin{aligned}&n(m-1)/(n+m-2) + \sum_{i=1}^{m-2} ni/(n+i)(n+i-1) \\ &= n\left(\frac{m-1}{n+m-2} + \sum_{i=1}^{m-2} \frac{i}{(n+i)(n+i-1)}\right) \\ &\leq n \cdot \left(1 + \sum_{i=1}^{m-2} 1/(n+i)\right) \\ &\leq n \cdot \left(1 + \ln(\frac{n+m-2}{n})\right) \\ &\leq n \cdot \left(1 + \ln(1 + \frac{m}{n})\right).\end{aligned}$$

□

And now we can prove Theorem 8, which essentially combines Proposition 4 and Proposition 5 (with some extra work).

Proof of Theorem 8. We again directly compute $\Pr[X_S(n,c) > p] = \Pr[X_{(1),n} > p] + \Pr[Z_{(1),c} > p | X_{(1),n} < p]$. Similarly, $\Pr[X_L(n,m) > p] = \Pr[X_{(1),n} > p] + \Pr[W_{2,n} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p] + \Pr[Y^*_{n,m-1} > p | X_{(1),n} < p \wedge W_{2,n} < p] \cdot \Pr[X_{(1),n} < p \wedge W_{2,n} < p]$, where again $Y^*_{n,m-1}$ is defined to be 0 if $Y_{1,m-1} < X_{(1),n}$, and otherwise equal to $Y_{j,m-1}$ for a uniformly random $j \in \{j | Y_{j,m-1} > X_{(1),n}\}$.

By Proposition 4 we have that $\Pr[W_{2,n} > p | X_{(1),n} < p] \leq \Pr[Z_{(1),n} > p]$. Now we need to reason about $Y^*_{n,m-1}$ conditioned on $X_{(1),n} < p$ and $W_{2,n} < p$, which is not directly related to any previous propositions. However, we claim that $Y^*_{n,m-1}$ and $W_{2,n}$ are positively correlated, conditioned on $X_{(1),n} < p$.

Lemma 5. $\Pr[Y^*_{n,m-1} > p | X_{(1),n} < p \wedge W_{2,n} < p] \leq \Pr[Y^*_{n,m-1} > p | X_{(1),n} < p]$.

Proof. For this proof, for random variables A, B, C, when we say A and B are conditionally independent, conditioned on C we mean that for all c, the random variables A and B are conditionally independent, conditioned on the event $C = c$. We use this shorthand to avoid cumbersome notation.

The proof consists of three steps: (a) we first show $Y^*_{n,m-1}$ and $W_{2,n}$ are conditionally independent, conditioned on $X_{(1),n}$, (b) we show that, conditioned on $X_{(1),n} < p$, $Y^*_{n,m-1}$ is positively correlated with $X_{(1),n}$, and (c), $W_{2,n}$ is positively correlated with $X_{(1),n}$. Together, this essentially lets us claim that additionally conditioning on $W_{2,n} < p$ only serves to lower $X_{(1),n}$, which lowers the probability that $Y^*_{n,m-1} > p$.

Observe first that conditioned on $X_{(1),n}, \ldots, X_{(n),n}$, $Y^*_{n,m-1}$ and $W_{2,n}$ are independent (this is just by definition: they are drawn independently, but the distributions from which they are drawn depend on $X_{(1),n}, \ldots, X_{(n),n}$). However, observe that the distribution from which $Y^*_{n,m-1}$ is independently drawn can be defined as a function only of $X_{(1),n}$. This means that, conditioned on $X_{(1),n}$, $X_{(2),n}$ and $Y^*_{n,m-1}$ are conditionally independent. Similarly, the distribution from which $W_{2,n}$ is independently drawn from can be described as a function only of $X_{(2),n}$, which we just claimed is conditionally independent of $Y^*_{n,m-1}$, conditioned on $X_{(1),n}$. Therefore, $Y^*_{n,m-1}$ and $W_{2,n}$ are conditionally independent, conditioned on $X_{(1),n}$.[5]

Next, we want to claim that, for all $r \leq q \leq p$, $\Pr[Y^*_{n,m-1} > p | X_{(1),n} = q] \geq \Pr[Y^*_{n,m-1} > p | X_{(1),n} = r]$. To see this, observe the following equivalent method for drawing $Y^*_{n,m-1}$: First, draw $Y_{1,m-1}, \ldots, Y_{m-1,m-1}$ i.i.d. from the uniform distribution on $[0, 1]$. Permute them into random order.[6] Then, let j be the smallest index such that $Y_{j,m-1} > X_{(1),n}$. If no such j exists, set $Y^*_{n,m-1} = 0$. Otherwise, set $Y^*_{n,m-1} = Y_{j,m-1}$. Now, let's couple draws for $Y^*_{n,m-1}$ conditioned on $X_{(1),n} = q$ and $X_{(1),n} = r$ by fixing the values $Y_{1,m-1}, \ldots, Y_{m-1,m-1}$ and the random permutation. Then think of $Y^*_{n,m-1}$, conditioned on $X_{(1),n} = r$ (respectively, $X_{(1),n} = q$) as scanning the values sequentially until it hits one whose value exceeds r (respectively, q).

- If the (permuted) sequence $Y_{1,m-1}, \ldots, Y_{m-1,m-1}$ has no values $> p$, we have $Y^*_{n,m-1} < p$ in both cases.
- If the sequence has values $> p$, but a value $\in (q, p)$ precedes all values $> p$, then again $Y^*_{n,m-1} < p$ in both cases. This is because both scans stop at the value $\in (q, p)$ which is not $> p$.

---

[5] Note that $Y^*_{n,m-1}$ and $W_{2,n}$ are *not* conditionally independent, conditioned on $X_{(1),n} < p$. They are only conditionally independent, conditioned on $X_{(1),n} = q$ (for some q).

[6] Actually, this step is not necessary, but it helps the analogy to state it.

- If the sequence has value $> p$, and the first one is *not* preceded by any value $\in (r, p)$, then $Y^*_{n,m-1} > p$ in both cases. This is because both scans stop at a value $> p$ and output it.
- If the sequence has a value $> p$, but a value $\in (r, q)$ precedes all values $> p$ but *no* value $\in (q, p)$ precedes the first value $> p$: then $Y^*_{n,m-1} > p$ when conditioned on $X_{(1),n} = q$, but $Y^*_{n,m-1} < p$ when conditioned on $X_{(1),n} = r$. This is because the $X_{(1),n} = q$ scan skips over the value $\in (r, q)$, and stops at the value $> p$, whereas the $X_{(1),n} = r$ scan stops at the value $\in (r, q)$.

This covers all cases, and proves that for all $r \leq q \leq p$, $\Pr[Y^*_{n,m-1} > p | X_{(1),n} = q] \geq \Pr[Y^*_{n,m-1} > p | X_{(1),n} = r]$.

Finally, we make the same claim for $W_{2,\ell}$: for all $r \leq q$, $\Pr[W_{2,\ell} > p | X_{(1),n} = q] \geq \Pr[W_{2,\ell} > p | X_{(1),n} = r]$. This claim is more straight-forward: $W_{2,\ell}$ is drawn from a uniform distribution on $[X_{(2),\ell}, 1]$. So clearly, $\Pr[W_{2,\ell} > p | X_{(2),n} = q] \geq \Pr[W_{2,\ell} > p | X_{(2),n} = r]$ whenever $r \leq q$. Moreover, $X_{(2),n}$ is distributed according to the maximum of $n-1$ i.i.d. uniform draws from $[0, X_{(1),n}]$, so the distribution of $X_{(2),n}$ conditioned on $X_{(1),n} = q$ stochastically dominates that of $X_{(2),n}$ conditioned on $X_{(1),n} = r$ whenever $q \geq r$. Both observations together allow us to conclude that for all $r \leq q$, $\Pr[W_{2,\ell} > p | X_{(1),n} = q] \geq \Pr[W_{2,\ell} > p | X_{(1),n} = r]$.

Now we may put all three claims together to prove the lemma. $\square$

Now with Lemma 5, we can wrap up the proof. We now know that:

$$
\begin{aligned}
&\Pr[X_L(n,m) > p] \\
&= \Pr[X_{(1),n} > p] \\
&+ \Pr[W_{2,n} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p] \\
&+ \Pr[Y^*_{n,m-1} > p | X_{(1),n} < p \wedge W_{2,n} < p] \cdot \Pr[X_{(1),n} < p \wedge W_{2,n} < p] \\
&\leq \Pr[X_{(1),n} > p] \\
&+ \Pr[W_{2,n} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p] \\
&+ \Pr[Y^*_{n,m-1} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p \wedge W_{2,n} < p] \\
&\leq \Pr[X_{(1),n} > p] \\
&+ \Pr[W_{2,n} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p] \\
&+ \Pr[Z_{(1),n(1+\ln(1+m/n))} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p \wedge W_{2,n} < p] \\
&\leq \Pr[X_{(1),n} > p] + \Pr[W_{2,n} > p \vee Z_{(1),n(1+\ln(1+m/n))} > p | X_{(1),n} < p] \cdot \Pr[X_{(1),n} < p]
\end{aligned}
$$

At this point, observe that the random variables $W_{2,n}$ and $Z_{(1),n(1+\ln(1+m/n))}$ are independent (and also conditionally independent, conditioned on $X_{(1),n} < p$). Therefore:

$$
\begin{aligned}
&\Pr[W_{2,n} < p \wedge Z_{(1),n(1+\ln(1+m/n))} < p | X_{(1),n} < p] \\
=&\Pr[W_{2,n} < p | X_{(1),n} < p] \cdot \Pr[Z_{(1),n(1+\ln(1+m/n))} < p | X_{(1),n} < p]
\end{aligned}
$$

By Proposition 4, we know that $\Pr[W_{2,n} < p | X_{(1),n} < p] \geq \Pr[Z_{(1),n} <$

$p|X_{(1),n} < p]$. Therefore, we get that:

$$\begin{aligned}
&\Pr[W_{2,n} < p \wedge Z_{(1),n(1+\ln(1+m/n))} < p|X_{(1),n} < p] \\
\geq &\Pr[Z_{(1),n} < p|X_{(1),n} < p] \cdot \Pr[Z_{(1),n(1+\ln(1+m/n))} < p|X_{(1),n} < p] \\
= &\Pr[Z_{(1),n} < p] \cdot \Pr[Z_{(1),n(1+\ln(1+m/n))} < p] \\
= &\Pr[Z_{(1),n(2+\ln(1+m/n))} < p].
\end{aligned}$$

Therefore,

$$\Pr[Z_{(1),n(2+\ln(1+m/n))} > p] \geq \Pr[W_{2,n} > p \vee Z_{(1),n(1+\ln(1+m/n))} > p|X_{(1),n} < p].$$

So substituting all the way back, we get that:

$$\begin{aligned}
[X_L(n,m) > p] &\leq \Pr[X_{(1),n} > p] + \Pr[Z_{(1),n(2+\ln(1+m/n))} > p] \cdot \Pr[X_{(1),n} < p] \\
&= \Pr[X_S(n, n(2 + \ln(1 + m/n))) > p].
\end{aligned}$$

This completes the proof. □

CHAPTER 3

PANDORA'S PROBLEM WITH ALTERNATIVE INSPECTIONS

3.1 Introduction

Search theory, which concerns the ways in which costs of obtaining information affect the structure and outcome of optimization procedures, was born in 1961 when the economist George Stigler [57] sought to understand the phenomenon of price dispersion. When sellers charge different prices for identical goods, why do consumers ever choose the higher-priced seller? Stigler realized that this counter-intuitive behavior could be explained by *search frictions* whereby consumers must expend costly effort to find and/or evaluate sellers.

The insight that optimization has qualitatively different outcomes under search frictions resounded beyond economics, and particularly within computer science. Models of costly information acquisition have been incorporated into information retrieval, robotics, database theory, distributed systems, and of course also into sub-areas of CS such as algorithmic pricing and mechanism design that explicitly relate to economics.

From a mathematical standpoint, the most foundational model of optimal search was articulated by Martin Weitzman [59] under the name *Pandora's problem*. The basic elements of the problem are as follows. A searcher is allowed to select a prize from one of n closed boxes. The values of the prizes inside the boxes are independent random variables, sampled from (not necessarily identical) distributions that are known to the searcher. The searcher chooses a sequence of operations, each of which is either *opening* a box or *selecting* the box. Opening box

i has an associated cost c_i and results in learning the value v_i of the prize contained inside. Selecting box i results in a payoff of v_i and immediately ends the search process; this operation can only be performed after box i has been opened. The searcher's goal is to design an adaptive policy (i.e., a choice of which operation to perform next, for every possible past history of operations and their outcomes) to maximize the expectation of the prize selected, minus the sum of the inspection costs accrued while opening boxes.

A priori, it would appear that the solution to Pandora's problem may be horribly complex. An optimal adaptive policy must specify the next operation to be performed given any past history. If each v_i is drawn from a distribution with support size s, the number of possible histories is s^n, so adaptive policies in general have exponential description size. It is easy to see that the optimal policy can be implemented in space $\text{poly}(n, s)$, but there is no obvious reason why the complexity of Pandora's problem should lie anywhere below PSPACE.

Surprisingly, though, the solution to Pandora's problem is not complex at all. Weitzman proved that the optimal policy has a beautifully simple structure: one computes a *reservation value* for each box, sorts them in decreasing order of reservation value, and opens them in this order, stopping and selecting the first open box whose prize value exceeds the reservation value of every remaining closed box. This entire process can be implemented to run in time $O(ns \log(ns))$.

A key assumption in Pandora's problem is that before selecting a box, the searcher must undertake a predetermined process to open the box. This assumption limits the applicability of the model. For example, when using Pandora's problem to model a firm searching for an employee to hire, boxes represent job candidates. Before hiring a candidate, the firm can decide to assess her through a lengthy internship process, evaluate her through a short interview process, or skip such costly evaluations and hire directly. This motivates a version of Pandora's problem in which a searcher can select one of multiple ways to open a box where a different value distribution and cost is associated with each method of inspection. We call this variant Pandora's problem with alternate inspection.

As a special case of the above problem suppose that before selecting a box, the searcher can either pay the cost to open the box and realize the value inside it; or select the box without opening it. We call this, Pandora's problem with nonobligatory inspection. The optimal solution to this close variant is rather complex and lacks the nice structure of the optimal solution in the original problem. For example, Doval [22] presents an example of a problem instance (Problem 3 in [22]) with three boxes — A, B, and C — such that the optimal policy first opens box C, but the question of whether it subsequently opens box A before B or vice-versa depends on the value of the prize discovered inside box C. As before, one can easily show that this variant of Pandora's problem belongs to PSPACE, but unlike Weitzman's version of Pandora's problem, there is no evidence that this version is easier than PSPACE-complete.

These considerations motivate the study of approximately optimal policies that are computationally efficient, structurally simple, or both. Our work follows this direction.

3.1.1 Our results and techniques

To put our results in context, we begin this section with an easy observation showing that simple, computationally efficient policies can attain at least a $\frac{1}{2}$-

approximation to the optimal policy of Pandora's problem with nonobligatory inspection. Consider the following two policies.

1. [**Policy A**] Run Weitzman's optimal policy, ignoring the fact that the searcher has the option to select boxes without opening them.
2. [**Policy B**] Leave every box closed, and select the one with the highest expected value.

Among all adaptive policies, Policy A is the one that maximizes the expected net contribution (i.e., the value if selected, minus the inspection cost) of open boxes, whereas Policy B maximizes the expected net contribution of closed boxes. Hence the combined value of Policies A and B bounds from above the combined value that the optimal policy obtains from both open and closed boxes. The better of A and B must consequently attain at least half the value of the optimal policy.

For any specified $\varepsilon > 0$, it is not hard to construct a problem instance such that neither Policy A nor Policy B attains more than $\frac{1}{2} + \varepsilon$ of the value of the optimal policy. To achieve a better approximation factor, we focus on a broader class of simple policies that includes both of the aforementioned ones.

In the more general problem of Pandora's box with alternative inspections, let us define a committing policy to be one that, before it inspects any boxes, must pre-commit to exactly one alternative of opening each box; in addition it pre-commits to an order in which the boxes will be opened. Such a policy is almost non-adaptive; the only way in which it may adjust its behavior in response to information revealed during the search process is that it may terminate the search early. In this sense, questions about the ability of committing policies to approximate the optimal adaptive policy are akin to questions about *adaptivity* *gaps* in stochastic optimization [2, 21, 34, 35].

The foregoing discussion inspires two interrelated questions.

Question 1 *For which values of α is there a simple policy that α-approximates the optimal adaptive policy?*

Question 2 *What is the worst-case ratio between the value of the optimal committing policy and that of the optimal adaptive policy?*

We show that for the general case of Pandora's problem with alternative inspections, the optimal committing policy always attains at least $1 - \frac{1}{e} = 0.63\ldots$ fraction of the optimal policy's value. This furnishes a non-trivial lower bound on the answers to Questions 1 and 2 above. We also provide an example where the optimal committing policy does not achieve more than $\frac{4}{5}$-approximation to the optimal adaptive policy. The main question left open by our work is whether the factor of $1 - \frac{1}{e}$ for the general case of n boxes can be improved. For the special case of n boxes with nonobligatory inspection, the factor of 4/5 is known to be tight due to the work of Guha et al. [33].

In the remainder of this section, we briefly discuss the techniques used to achieve these results. Our proof that committing policies attain a $(1 - \frac{1}{e})$-approximation to the value of the optimal adaptive policy starts with a crucial observation: Pandora's problem with alternative inspections can be recast as an equivalent problem in which each box in the original problem corresponds to a tuple of boxes, where each box in the tuple has its own value distribution and inspection cost. To make the problem with tupled boxes equivalent to the original problem instance, we must impose an additional constraint that search policies for the tupled-box problem may open at most one of the multiple boxes in each tuple. This reduction, which appears simple and natural in hindsight, is crucial because it enables the application of two powerful tools. The first is a lemma of Kleinberg, Waggoner, and Weyl [43] that reduces the analysis of policies for Pandora's problem and generalizations to the analysis of algorithms for the same optimization problem when the values of items are revealed for free, but are sampled from modified distributions. In Section 3.2.3 we generalize the lemma to account for policies that can choose among different ways to open a box, a generalization which is vital for our application. The second tool is a theorem of Asadpour and Nazerzadeh about the adaptivity gap of stochastic submodular function maximization problems. Once Pandora's problem with alternative inspections has been transformed into a form where these two ingredients apply, the derivation of the $(1 - \frac{1}{e})$-approximation result becomes nearly automatic. The combination of the two ingredients — the Kleinberg-Waggoner-Weyl amortization lemma from [43] together with adaptivity gaps for stochastic probing — was pioneered by Singla [56] to solve a problem he refers to as *constrained utility maximization in the price of information model*, which generalizes our tupled-box problem with probing constraints.

3.1.2 Related Work

We have already discussed the foundational work on optimal search theory in economics, particularly Weitzman's paper [59] that introduced Pandora's problem and derived its solution. The optimality of Weitzman's procedure turns out to be a special case of the Gittins Index Theorem [30, 31], which ironically was proven earlier although Weitzman obtained his results independently and the connection between these two theorems was only realized afterward.

To the best of our knowledge we are the first to study Pandora's problem with alternative inspections. However, the specific case of nonobligatory inspection has been studied before. In the search literature Doval [22] is known to be the first to introduce Pandora's problem with nonobligatory inspection. However, the same problem was studied sooner by Guha et al. [33] in the context of wireless networks. The main contribution of [33] is a 4/5 approximation to the Pandora's problem with non-obligatory inspection when the support of each box value distribution is discrete and finite. Although they also use the notion of committing policies, their techniques for proving the approximation guarantees extensively use the properties of nonobligatory inspection setting and do not extend to the case of alternative inspections.

Doval [22] was the first to address Pandora's problem with nonobligatory inspection as an extension to Weitzman's model, though special cases were anticipated in earlier unpublished work by Postl [51]. In addition to examples illustrating that optimal policies in general need to be adaptive (as described above), Doval's main results identify sufficient conditions for the optimal policy to have a simple structure. In particular, Theorem 1 in [22] identifes a sufficient condition under which the optimal policy is a committing policy. The sufficient condition is quite technical, but one corollary is that a committing policy is optimal whenever boxes have equal inspection costs and are totally ordered by the "mean-preserving spread" relation. Doval also provides a complete solution for the case when the boxes have equal costs, the value of each is sampled from a distribution with two-point support, and the lower support point is the same for all boxes. Attias et al. [3] study specific cases of the non-obligatory problem and provide approximately optimal algorithms.

The blending of Pandora's problem with ideas from combinatorial optimization and algorithmic game theory was initiated by Kleinberg, Waggoner, and Weyl [43]. Their paper introduced a novel method for analyzing optimal and approximately-optimal policies for Pandora's problem and generalizations, by relating the expected utility of the policy to expected values of related quantities in a simpler environment without inspection costs. The paper primarily applies this method to analyze the price of anarchy of a descending price auction when bidders face a cost to inspect their own value, but it also analyzes various extensions including one in which inspection is optional; the price of anarchy of the descending auction in this setting is shown to be no worse than $\frac{1}{2} - \frac{1}{2e} \approx 0.316$. Singla [56] applied the analysis technique introduced in [43] to a much broader family of combinatorial optimization problems, providing a general transformation to convert *frugal algorithms* (a type of greedy algorithm) for combinatorial optimization problems into policies for solving combinatorial counterparts to Pandora's problem, i.e., generalizations in which the searcher still must pay a cost to open each box, but may be allowed to select multiple boxes, subject to feasibility constraints on the set of selected boxes. As noted earlier, among the problems solved in [56] is a constrained utility maximization problem featuring probing constraints that generalize the probing constraint in our tupled-box problem.

Adaptivity gaps have been studied for various stochastic optimization problems. Any such problem consists of a set of elements whose values are independent random variables. The algorithm knows the distributions of these variables, but not the actual realizations. The only way to learn the actual realizations is to probe these elements. If the value of the optimal adaptive probing policy can always be approximated, to within a factor of α, by the value of a simple policy that performs probes in a fixed, predetermined order until a stopping time is reached, then we say the problem has an adaptivity gap of α. In one of the earliest papers on adaptivity gaps in stochastic optimization, Dean, Goemans, and Vondrak [21] studied a stochastic variant of the 0/1 knapsack problem, where items have deterministic values but their sizes are independent random variables and the act of placing an item in the knapsack reveals its size. They showed that adaptivity gap is constant and provided constant factor non-adaptive approximations.

The proof of our main result makes use of adaptivity gaps for stochastic submodular optimization with constraints on probing. Asadpour and Nazerzadeh [2] bound the adaptivity gap to $1 - \frac{1}{e}$ for maximizing stochastic monotone submodular functions when elements to probe should satisfy matroid feasibility constraints. Adaptivity gaps for much more general families of constraints were subsequently proven by Gupta, Nagarajan, and Singla [34, 35]. In addition to feasibility constraints over sets of elements to probe, there may also be constraints on the ordering of the probes. Gupta, Nagarajan, and Singla showed a constant adaptivity gap for submodular functions under arbitrary prefix-closed constraints on the sequence of elements probed [35].

3.2 Preliminaries

In this section, we formally define our model and discuss three related problems:*Pandora's problem with nonobligatory inspection*, *Pandora's problem with required inspection* and *maximizing a stochastic monotone submodular function*. Then we introduce a class of search procedures called *committing policies* and explain why the optimal committing policy has a simple structure.

3.2.1 Model

An agent has a set of n boxes. Box i, $1 \leq i \leq n$, contains a prize, v_i, distributed according to distribution $F_i(v_i)$ with expected value $\mathbb{E}v_i$. Prizes inside boxes are independently distributed.

The agent sequentially inspects boxes, and search is with recall. For each box the searcher has the option to choose at most one inspection method from $1, 2, \ldots, m$. Each inspection method decreases the uncertainty about the prize by a different amount. Below we explain how we can consider a different distribution of prizes for each pair of box and inspection method.

For an easier illustration, first consider Pandora's problem with nonobligatory inspection. In this problem there are two inspection methods: Method $j = 1$ fully inspects and realizes the value of the prize inside the box. Method $j = 2$ does not inspect at all. Suppose the prize inside box i is v. If the searcher chooses method 1, she learns v. If she chooses method 2, she does not learn v, but may still select this box. If she inspects with method 2 and selects the box, her decision to select the box is independent of v (she has not seen the value so her decision cannot possibly be dependent), and in expectation she gets $\mathbb{E}v_i$. Therefore, equivalently we can assume that while the realized value of box i with inspection method 1 is a random draw from distribution F_i, the realized value with inspection method 2 is a random draw from the deterministic distribution with value $\mathbb{E}v_i$.

This notion of considering a different distribution for each inspection method generalizes beyond the extreme cases of full inspection and no inspection. For example, we can think of each inspection method as only identifying a group of values that the prize belongs to. This concept describes inspection methods in the spectrum of full inspection to no inspection. In this case, the expected value of the prize upon selecting this box, is the expected value of the group that the prize belongs to. The distribution corresponding to this inspection method of the box, is defined as a distribution over the expected values of the groups that the values belong to.

Therefore, we assume for inspection method j, $1 \leq j \leq m$, the prize inside box i, v_{ij}, is distributed according to distribution $F_{ij}(v_{ij})$. The inspection cost for box i associated with method j is c_{ij}. While F_{ij} and c_{ij} are known; v_{ij} is not. Although in the above illustration, for each box i, the expected value of distributions F_{ij} is equal to $\mathbb{E}v_i$ and more costly inspection methods corresponds to more granular distributions, we do not need any such assumptions for our techniques to work.

Given a set of uninspected boxes, $\mathcal{U}$, and a vector of realized sampled prizes, v, the agent decides whether to stop or to continue search; if she decides to continue search she decides which box in $\mathcal{U}$ to inspect next and what method to use for inspection. If she decides to inspect box i with method j, she pays cost c_{ij} to instantaneously learn her value v_{ij}. If she decides to stop the search, she selects the maximum value realized so far.

We use $\mathbb{I}_{ij}$ as an indicator for box i being inspected by method j and $\mathbb{A}_{ij}$ as an indicator for the agent obtaining box i after inspecting it with method j. Since one box can be obtained, $\sum_{i,j} \mathbb{A}_{ij} \leq 1$. The agent is an expected utility maximizer, where utility, u, is defined as the value of the box selected minus the sum of inspection costs paid. Given v, the vector of realized sampled prizes, and the two vectors of indicator variables, $\mathbb{A}$ and $\mathbb{I}$, respectively indicating which boxes were selected and inspected, we have:

$$u(v, \mathbb{A}, \mathbb{I}) = \sum_{i,j} (\mathbb{A}_{ij} v_{ij} - \mathbb{I}_{ij} c_{ij}).$$

3.2.2 Nonobligatory Inspection

The nonobligatory inspection is a special case of the model introduced above. Box i, $1 \leq i \leq n$, contains a prize, v_i, distributed according to distribution $F_i(v_i)$ with expected value $\mathbb{E}v_i$. For each box i the agent has the options to pay inspection cost c_i and realize value v_i, or select it without inspection and in expectation achieve $\mathbb{E}v_i$. To observe this is a special case, note that for each box we can define two methods of inspection: one with distribution $F_i(v_i)$ and c_i, and the other with deterministic distribution with value $\mathbb{E}v_i$ and cost 0.

3.2.3 Required Inspection

Consider imposing the additional constraint that there is a single way to open a box. We drop the subscript j for this subsection. For this constraint we require $\mathbb{A}_i \leq \mathbb{I}_i$ for each i.

Weitzman [59] finds the optimal procedure to maximize expected utility when inspection is required. The optimal solution is an index-based policy, in which the agent inspects boxes in decreasing order of their indices, σ_i, where σ_i is the unique solution to

$$\mathbb{E}_{v_i \sim F_i}\left[(v_i - \sigma_i)^+\right] = c_i$$

and is also known as the reservation value of box i. The search stops either when one of the realized values is above the reservation value of every remaining uninspected box, or when the agent has inspected all of the boxes.

Kleinberg et al. [43] develop a new interpretation of Weitzman's characterization. They introduce an important property of policies that we will call "non-exposure", defined as follows.

Definition 2 (Non-exposed Policy)**.** *A policy is non-exposed if it is guaranteed to select any inspected box whose value is found to satisfy* $v_i > \sigma_i$*. In other words, a policy is non-exposed if the event* $(\mathbb{I}_i - \mathbb{A}_i)(v_i - \sigma_i) > 0$ *has probability zero, for every box* i*.*

The key to the analysis of Weitzman's optimal policy in [43] is a family of random variables $\kappa_i \triangleq \min\{v_i, \sigma_i\}$ defined for each box i. Kleinberg et al. prove that for any policy that satisfies the required-inspection constraint $\mathbb{A}_i \leq \mathbb{I}_i$, the net contribution of box i to the expected value of the policy is bounded above by $\mathbb{E}[\mathbb{A}_i \kappa_i]$, with equality if and only if the policy is non-exposed.

Lemma 6. *[43] Given any* F_i *and any policy that satisfies* $\mathbb{A}_i \leq \mathbb{I}_i$ *pointwise,*

$$\mathbb{E}\left[\mathbb{A}_i v_i - \mathbb{I}_i c_i\right] \leq \mathbb{E}[\mathbb{A}_i \kappa_i]. \tag{3.1}$$

Furthermore, this holds with equality for every box i *if and only if the policy is non-exposed.*

Lemma 6 can be interpreted as providing an accounting scheme that amortizes a policy's expected inspection costs by deducting them from the expected value of the box it eventually selects. This accounting scheme exactly characterizes the value of non-exposed policies, and furnishes an upper bound on the value of every other policy. The benefit of the amortization is that it reduces the problem of analyzing policies for Pandora's problem to the (generally simpler) problem of analyzing rules for selecting boxes in an environment where the value of box i is κ_i, and this value can be queried at no cost. A first application of this technique is the following characterization of the optimal policy with required inspection, and its expected utility.

Corollary 6. *[43] Weitzman's policy on boxes $1 \leq i \leq n$ with distributions F_i and inspection costs c_i, achieves expected utility $\mathbb{E}[\max_i \kappa_i]$; the expected utility of any other policy cannot exceed this bound.*

Since Pandora's problem with alternative inspections allows for a choice among multiple ways to open a box we will need a generalization of Lemma 6 that pertains to such policies.

Lemma 7. *Given any policy for Pandora's problem with alternative inspections, and any i and j, The inequality*

$$\mathbb{E}\left[\mathbb{A}_{ij} v_{ij} - \mathbb{I}_{ij} c_{ij}\right] \leq \mathbb{E}[\mathbb{A}_{ij} \kappa_{ij}] \tag{3.2}$$

is always satisfied, and the two sides are equal for all i and j if and only if the policy is non-exposed.

Proof. The proof is similar to the proof of Lemma 6 in [43]. Using the definition of σ_{ij}, and independence of v_{ij} and $\mathbb{I}_{ij}$:

$$\mathbb{E}[\mathbb{A}_{ij} v_{ij} - \mathbb{I}_{ij} c_{ij}] = \mathbb{E}[\mathbb{A}_{ij} v_{ij} - \mathbb{I}_{ij}(v_{ij} - \sigma_{ij})^+] \leq \mathbb{E}[\mathbb{A}_{ij}(v_{ij} - (v_{ij} - \sigma_{ij})^+)] = \mathbb{E}[\mathbb{A}_{ij} \kappa_{ij}].$$

The inequality follows because $\mathbb{A}_{ij} \leq \mathbb{I}_{ij}$. Furthermore, if the policy is non-exposed, then $\mathbb{A}_{ij} < \mathbb{I}_{ij}$ implies $(v_{ij} - \sigma_{ij})^+ = 0$, so the inequality holds with equality. □

3.2.4 Stochastic Submodular Maximization

Consider the problem of maximizing a stochastic monotone submodular function f with respect to a matroid constraint $\mathcal{M}$. Suppose $f : \mathbb{R}_+^n \to \mathbb{R}_+$ is a function of n random variables, namely, $\mathcal{A} = \{X_1, X_2, \cdots, X_n\}$. Assume f is submodular, meaning

$$\forall x, y \in \mathbb{R}_+^n \quad f(x) + f(y) \geq f(x \wedge y) + f(x \vee y) \tag{3.3}$$

where $x \wedge y$ and $x \vee y$ respectively denote the coordinate-wise minimum and maximum of vectors x and y.

A policy π picks the elements to inspect one by one (perhaps, based on the realized value of the previous elements) until it stops. Once π stops, the current state is a random vector $\Theta^\pi = (\theta_1, \theta_2, \cdots, \theta_n)$, where θ_j denotes the realization of X_i, if i is inspected by the policy, and is equal to 0 otherwise. The objective of stochastic submodular maximization is to optimize the expected value of a policy, i.e., $\underset{\pi}{\text{Maximize}}\ \mathbb{E}[f(\Theta^\pi)]$, subject to feasibility. The feasibility constraint is modeled using a matroid. For a given matroid $\mathcal{M}$ defined on the ground set of the aforementioned random variable set $\mathcal{A}$, a policy π is called feasible if the subset of random variables it inspects is always an independent set of $\mathcal{M}$.

Asadpour and Nazerzadeh [2] compare the performance of the best **adaptive** and **non-adaptive** policies. In adaptive policies, at each point in time all the information regarding the previous inspections of the policy is known. In other words, the policy has access to the actual realized value of all the elements it has inspected so far. In contrast, non-adaptive policies do not have access to such information and should make their decisions (about which random variables to inspect) before observing the outcome of any of them. They show that there exists a non-adaptive policy that achieves at least a $1 - \frac{1}{e} \approx 0.63$ fraction of the value of the optimal adaptive policy.

Lemma 8. *[2] There exists a non-adaptive policy that achieves* $1 - \frac{1}{e} \approx 0.63$ *fraction of the optimal policy in maximizing a stochastic monotone submodular function with respect to matroid feasibility.*

We now use the multilinear relaxation of f to define the value of *fractional non-adaptive policies* [2]. A fractional non-adaptive policy is determined by a vector $y \in [0,1]^n$. This policy inspects elements in the (random) set Y, a set that is defined to include each $X_i \in \mathcal{A}$ with probability y_i, independently for each i.

We use $F(y)$ to denote the expected value obtained by the fractional non-adaptive policy associated with y. Using the notation Θ^Y to denote the random vector Θ^π when π is the non-adaptive policy associated with set Y, we have

$$F(y) := \sum_{Y \subseteq \{0,1\}^n} \left[\left(\prod_{i \in Y} y_i \prod_{i \notin Y} (1 - y_i) \right) \mathbb{E} f(\Theta^Y) \right]. \tag{3.4}$$

Lemma 9. *[2] For any monotone submodular function with matroid* $\mathcal{M}$ *feasibility constraint, for any* y *in the base polytope of* $\mathcal{M}$*, there exists an integral (deterministic) non-adaptive policy with expected value greater than or equal to* $F(y)$*.*

3.2.5 Committing Policies with Alternative Inspections

Consider the problem of maximizing expected utility for the box problem with alternative inspections (as discussed in Section 3.2.1). A class of policies that will be central to our analysis are the *committing policies*, which were discussed in Section 3.1 and are defined formally as follows.

Definition 3 (Committing Policy)**.** *A policy is called* committing *if there exists* *an index j_i assigned to each box i as the method of inspecting the box, and a total ordering of over the boxes, denoted by $\prec$, such that the following properties hold.*

1. *The policy never inspects box i using any method other than j: $\forall i, j \neq j_i \;\; \mathbb{I}_{ij} = 0$.*
2. *If i, k and $i \prec k$ then the policy never inspects i before it has inspected k.*

3.3 $1 - \frac{1}{e}$ Approximation

In this section we analyze the worst-case ratio between the value of the optimal committing policy and that of the optimal policy. We also show there is a gap between these two policies.

Theorem 9. *There exists a committing policy that achieves at least $1 - \frac{1}{e} \approx 0.63$ of the optimal utility for the box problem with nonobligatory inspection.*

We establish a correspondence between the box problem and stochastic submodular optimization. Recall from Section 3.2.4 that an instance of stochastic submodular optimization is specified by a set of random variables $\mathcal{A} = \{X_1, X_2, \cdots, X_m\}$, a submodular function $f : \mathbb{R}_+^m \to \mathbb{R}_+$, and a matroid $\mathcal{M}$ with ground set $\mathcal{A}$. A policy π chooses (either adaptively or non-adaptively) a subset $I \subseteq \mathcal{A}$ of random variables whose values it probes, subject to the constraint that I must be an independent set in $\mathcal{M}$. The value obtained when running policy π is the random variable $f(\Theta^\pi)$, where Θ^π denotes the random vector $(\theta_1, \ldots, \theta_m)$ specified by setting $\theta_i = X_i$ if $i \in I$ and $\theta_i = 0$ otherwise.

Definition 4. *[Associated Stochastic Optimization Problem] Given an instance of Pandora's problem with alternative inspections, the* associated stochastic optimization problem *has a random variable for each pair of (box i, method to open box i) denoted by*

$$\mathcal{A} = \{X_{i=1,\ldots,n,\, j=1,\ldots,m}\},$$

submodular objective function

$$f(\theta_{i=1,\ldots,n,\, j=1,\ldots,m}) = \max_{i,j} \theta_{i,j},$$

and matroid constraint $\mathcal{M}$ defined by the partition matroid whose independent sets are all the subsets of $\mathcal{A}$ that contain at most one element of each tuple $\{X_{i,j}\}_{j=1}^{m}$. The distributions of the random variables are defined as follows: $X_{i,j}$ is drawn from the same distribution as κ_{ij}.

Probing the element $X_{i,j}$ of the tuple i in the associated stochastic optimization problem corresponds to inspecting box i with method j in the box problem. This correspondence is formalized by the following pair of policy transformations.

Definition 5. *Let $\mathcal{I}$ denote an instance of Pandora's problem with alternative inspections, and let $\mathcal{J}$ denote the associated stochastic optimization problem.*

If π is any (possibly adaptive) policy for Pandora's problem $\mathcal{I}$, let $\Phi(\pi)$ denote the adaptive policy for $\mathcal{J}$ that simulates π running in $\mathcal{I}$ and performs the following sequence of probes: whenever π inspects box i with method j, $\Phi(\pi)$ probes $X_{i,j}$.

If ρ is a non-adaptive policy for stochastic optimization problem $\mathcal{J}$ and $B(\rho) \subset \mathcal{A}$ is the set of random variables that ρ probes, let $S(\rho)$ denote the pairs of (box, method) $\{i, j \mid X_{i,j} \in B(\rho)\}$ and let $\Psi(\rho)$ denote the committing policy $\mathcal{P}^{S(\rho)}$ for Pandora's problem $\mathcal{I}$.

In the following lemmas, as in the preceding definition, $\mathcal{I}$ denotes an instance of Pandora's problem with alternative inspections and $\mathcal{J}$ denotes its associated stochastic optimization problem. If π is a policy for either problem $\mathcal{I}$ or $\mathcal{J}$, we will use the notation $u(\pi)$ to denote the expected utility of running policy π. In the case of Pandora's problem this means $u(\pi) = \mathbb{E}\left[\sum_{i,j}(\mathbb{A}_{ij}v_{ij} - \mathbb{I}_{ij}c_{ij})\right]$. In the case of the associated stochastic optimization problem it means $u(\pi) = \mathbb{E}\left[f(\Theta^{\pi})\right]$.

Lemma 10. *If ρ is a non-adaptive policy for $\mathcal{J}$ and $\Psi(\rho)$ is the corresponding committing policy for $\mathcal{I}$, then*

$$u(\Psi(\rho)) \;\geq\; u(\rho). \tag{3.5}$$

Proof. Couple the probability spaces of the two optimization problems such that when the prize inside box i is v_{ij}, the value of random variable $X_{i,j}$ equals $\kappa_{ij} = \min\{v_{ij}, \sigma_{ij}\}$. Note that such a coupling exists, because the random variables $\{X_{i,j}\}_{i=1,\ldots,n,\, j=1,\ldots,m}$ are mutually independent and $X_{i,j}$ has the same marginal distribution as κ_{ij} by construction.

By construction, policy $\Psi(\rho) = \mathcal{P}^{S(\rho)}$ is non-exposed. According to Lemma 6, then,

$$u(\Psi(\rho)) = \mathbb{E}[\max_i \tilde{\kappa}_i], \tag{3.6}$$

where $\tilde{\kappa}_i = \kappa_{ij}$ if $(i,j) \in S(\rho)$. As for $u(\rho) = \mathbb{E}[f(\Theta^{\rho})]$, by the definition of f and of Θ^{ρ} we have

$$u(\rho) = \mathbb{E}[\max_i \tilde{\theta}_i] \tag{3.7}$$

where $\tilde{\theta}_i = X_{i,j} = \kappa_{ij}$ if $X_{i,j} \in B(\rho)$, and $\tilde{\theta}_{i,j} = 0$ otherwise. In the former case $\tilde{\kappa}_{ij} = \tilde{\theta}_{ij}$ whereas in the latter case $\tilde{\kappa}_{ij} \geq 0 = \tilde{\theta}_{ij}$. Hence $\tilde{\kappa}_{ij} \geq \tilde{\theta}_{ij}$ pointwise. Combining this inequality with (3.6)-(3.7) and using the fact that the random variables $\{\tilde{\kappa}_{ij}\}$ are mutually independent, as are $\{\tilde{\theta}_{ij}\}$, the inequality $u(\Psi(\rho)) \geq u(\rho)$ follows. □

Lemma 11. *If π is any (possibly adaptive) policy for Pandora's problem $\mathcal{I}$, and $\Phi(\pi)$ is the corresponding policy for the associated stochastic optimization problem, then $u(\Phi(\pi)) \geq u(\pi)$.*

Proof. As in the proof of Lemma 10, couple the probability spaces of the two optimization problems such that the value of the random variable $X_{i,j}$ equals $\kappa_{ij} = \min\{v_{ij}, \sigma_{ij}\}$. By construction of policy $\Phi(\pi)$, the set of random variables it probes is $\{X_{i,j} \mid \mathbb{I}_{ij} = 1\}$. Hence, if we define $\tilde{\kappa}_{ij} = \kappa_{ij}$ when $\mathbb{I}_{ij} = 1$ and 0 otherwise, then we have

$$u(\Phi(\pi)) = \mathbb{E}[\max_{i,j} \tilde{\kappa}_{i,j}] \geq \sum_{i,j} \mathbb{E}[\mathbb{A}_{ij}\tilde{\kappa}_{ij}]. \tag{3.8}$$

Lemma 7 implies the following upper bound on $u(\pi)$.

$$u(\pi) = \mathbb{E}\left[\sum_{i,j}(\mathbb{A}_{ij}v_{ij} - \mathbb{I}_{ij}c_{ij})\right] \leq \mathbb{E}\left[\sum_{i,j}\mathbb{A}_{ij}\tilde{\kappa}_{ij}\right] \tag{3.9}$$

Combining this relation with inequality (3.8) completes the proof. ☐

Proof of Theorem 9. If π denotes the optimal policy for an instance $\mathcal{I}$ of Pandora's problem with nonobligatory inspection, and $\mathcal{J}$ denotes the associated stochastic optimization problem, let ρ denote an optimal non-adaptive policy for $\mathcal{J}$. We have the chain of inequalities

$$u(\rho) \geq \left(1 - \tfrac{1}{e}\right) \cdot u(\Phi(\pi)) \geq \left(1 - \tfrac{1}{e}\right) u(\pi)$$

where the first inequality is Lemma 10, the second is Lemma 8, and the third is Lemma 11. ☐

Lower Bound on Adaptivity Gap

The following example shows a 4/5 gap between committing policies and the optimal policy. More specifically the example considers adaptivity gap in Pandora's problem with nonobligatory inspection, implying a lower bound on the gap in the more general problem of alternative inspections.

Example 1. *Consider boxes A and B. Suppose box A has value* 0 *with probability* $\frac{1}{2}$ *and value* 1 *with probability* $\frac{1}{2}$*, and its inspection cost is* 0*. Box B has value* 0 *with probability* $1 - \frac{1}{N}$ *and value* N *with probability* $\frac{1}{N}$*; and its inspection cost is* $\frac{N-1}{2N}$.

The optimal policy starts with inspecting box A, and if the value is 0, selects uninspected box B. If the value of box A is 1, the optimal policy inspects box B and takes the maximum value of the two boxes. The expected utility of this policy is

$$\begin{aligned}&\frac{1}{2}\cdot\frac{1}{N}\cdot N+\frac{1}{2}\left(-\frac{N-1}{2N}+\frac{1}{N}\cdot N+\left(1-\frac{1}{N}\right)\cdot 1\right)\\=&\frac{1}{2}+\frac{1}{2}\left(\frac{3}{2}-\frac{1}{2N}\right)\end{aligned}$$

which approaches $\frac{5}{4}$ *as* N *goes to infinity.*

Policies $\mathcal{P}^1$*,* $\mathcal{P}^2$ *and* $\mathcal{W}$ *each achieve utility* 1*: Policy* $\mathcal{W}$*, inspects both boxes and obtains the maximum value. The expected utility in this case is* $-\frac{N-1}{2N}+\frac{1}{N}\cdot N+(1-\frac{1}{N})\cdot\frac{1}{2}$*. Policy* $\mathcal{P}^1$ *starts by inspecting box B. If box B has value* N*, it selects it. Otherwise it selects uninspected box A. Therefore it has utility* $-\frac{N-1}{2N}+\frac{1}{N}\cdot N+(1-\frac{1}{N})\cdot\frac{1}{2}$*. Policy* $\mathcal{P}^2$ *inspects box A. If the value is 0, it selects uninspected box B. If the value of box A is 1, it is indifferent between selecting box A and uninspected box B. The expected utility in this case is* $\frac{1}{2}\cdot\frac{1}{N}\cdot N+\frac{1}{2}\cdot 1$.

CHAPTER 4

TWO-SIDED MATCHING WITH LIMITED NUMBER OF APPLICATIONS

4.1 Introduction

We consider a two-sided matching market where individuals apply to positions and monetary transfers are not permitted. The main goal in such markets is typically to find a stable matching, i.e. a matching in which no pair of agents would rather be matched with each other than with their current partners. Given the preference lists of the applicants and the positions, a stable matching can easily be found using the much celebrated Gale-Shapley mechanism [28], which constitutes the basis for many real-world centralized matching mechanisms. Examples include the National Residency Matching Program (NRMP) used to match medical students with residency programs, public schools assignments in the US, and the assignment of seats in public universities in other countries (see e.g. [53]).

An important feature of the markets described above, which is usually not modeled by the classical work studying the Gale-Shapley mechanism, is that the number of applications sent by an individual is in general much lower than the number of open positions. For example, in school and university assignment systems there is often a constraint on the number of applications that can be submitted. Even if no such formal constraint is imposed, applications still might require a fair amount of work (e.g university applications typically require a separate essay for each school) or might impose a time constraint (e.g. in the NRMP, interviews must be conducted before the application process) which then naturally limits the number of applications submitted per individual.

The goal of this chapter is to study the incentive properties of the the Gale-Shapley mechanism in systems with limited number of applications. It is well known that in the Gale-Shapley stable matching mechanism with men proposing to women, it is dominant strategy for the men to report their preferences truthfully [52], while women can benefit from strategic behavior [29]. However, the truthful reporting for men is relying on the fact that they can report their complete preference list. When the number of applications is limited, the mechanism is no longer truthful and deciding where to apply to becomes non-trivial.

To illustrate, consider the college application process where options on either side of the market are not ex-ante identical: preferences of applicants and schools are highly correlated —not surprisingly, applicants generally prefer good schools, and all schools want to fill their positions with the best applicants. If the student has a low SAT score, he will very likely not get into a top-tier school. However, while college application tools and high school councilors try to help students understand the level of school they should expect to be admitted at, many low performing students still choose to apply to one (or a couple) of their dream schools. Is this behavior rational? Similarly, top-students with almost perfect SATs tend to add a couple of safe choices to their application lists. In any case, applicants usually do not (and should not) simply truncate their true preference lists, and have to employ more complex strategies instead. The question then becomes: how should applicants decide where to apply in matching markets with short lists, when agents are not ex-ante identical and preferences are correlated?

The objective of this chapter is three-fold: understand how strategic applicants decide where to apply, how these decisions are affected by the market primitives and, what is the "quality" of the outcome in these markets.

Our Results. We consider a simple model that allows us to study the questions proposed above. We assume that both sides of the market (referred to as doctors and hospitals hereafter) are divided into tiers. For simplicity, we assume two tiers on each side, high and low, and the number of doctors and hospitals to be equal in each tier, although some of our results can be extended to more tiers and unequal sides.[1] To model correlations in preferences, we assume that all agents agree that the high-tier doctors are better than the low-tier ones, and top tier hospitals are better than lower tier ones, but have heterogeneous within-tier preferences which we assume to be drawn uniformly at random. With our assumptions, the Gale-Shapley algorithm with full preference lists would match high doctors to high hospitals and low doctors to low hospitals, so we think of this assignment as the reasonable expectation of the doctors.

We study the mechanism in which each doctor submits a list of K applications, for some parameter K of the problem. Each hospital ranks all doctors, and we run the doctor-proposing Gale-Shapley algorithm. The utility that an agent (hospital or resident) obtains from a match is driven primarily by the tier of his partner. With preferences within each tier uniformly random, applicants simply need to decide how many hospitals in each tier to list.[2]

Characterization of Nash Equilibrium. To understand how applications are decided in this model, we use the large market approximation introduced in [1]. Our first main result is Theorem 10, showing that there exists a unique symmetric Nash equilibrium of this game, which has a nice structure: either there is a

---

[1] In particular, our equilibrium characterization is also valid if the number of agents on each side is not equal.

[2] It is easy to see that within each tier he will list his top choices. We defer the discussion to the Section 5.2, once the model is formally introduced.

pure-strategy equilibrium, or an equilibrium in mixed strategies in which doctors randomize between "close" configurations (there exists k_h, k_l in $\{0, \ldots, K-1\}$ such that doctors in tier $i \in \{h, l\}$ randomize between listing k_i or $k_i + 1$ high hospitals and the rest low hospitals.) This equilibrium captures the main behavior observed in practice: while most applications are devoted to the most reasonable alternatives, top doctors still list some safe hospitals, while bottom doctors list a couple of "reach" hospitals.

Worst case welfare properties. We measure welfare of an assignment of doctors to hospitals by considering the value both for doctors as well as for hospitals. Unassigned doctors and unfilled hospital positions derive no value. We assume that the value for a doctor of being assigned to a low hospital is 1, and the value for being assigned to a high hospital is v. While doctors have strict preferences among individual hospitals within a tier, we assume that these differences in value are small relative to the overall quality of the hospital; i.e. what tier they belong to. Similarly, we use 1 as the value for a hospital for being assigned a low doctor and v for the value for a high doctor.

Our second main result, Theorem 11, shows that the social welfare produced by the Nash equilibrium is always within a factor of 2 of the optimal welfare. This low price of anarchy is surprising, as selfish behavior of the doctors (of applying to too many "reach" hospitals), is depriving some of the lower hospitals from any doctors, and hence can do significant damage to welfare. In fact, if we allow for the number of high hospitals to be (much) smaller than the number of top quality applicants, the Nash equilibrium produces significantly lower welfare than that of the social optimum; in such case, the welfare loss is no longer bounded by a constant and rather becomes a function of the ratio of applicants to hospitals. To understand the effect of applications to “safe” and “reach” schools, we also compare the welfare of the Nash equilibrium against that produced by a simple market design, in which doctors can only apply to hospitals in their own tier. It is not hard to see that the welfare of this simple market design is within a $e/(e-1)$ factor of the optimum. We show that the Nash outcome can be either worse or better than this simple design in terms of social welfare, it is at most a $e/2(e-1) \approx 0.79$ factor worse than this simple restricted solution.

Sensitivity to the market design. Through a combination of theory results and simulations, we study how the primitives of the market impact the strategies used by doctors, and the total welfare that is produced. First, we show that the number of high applications k_h submitted by high doctors is monotone in K, but the number of applications sent to “safe” hospitals can change in surprising ways, e.g., can decrease when increasing number of overall applications. We give an example, where with $K = 5$ applications high doctors apply each to 1 high hospital and 4 low hospitals, while with $K = 6$ applications they apply to 3 hospitals of each type. We find empirically that high doctors typically apply to at most 1-2 low hospitals (“safe” options), while low doctors can be applying to many more high hospitals. We also study the effect of increasing K on welfare empirically for various values for high match, and ratios of low and high hospitals and applicants.

Related Literature. To the best of our knowledge, the first paper that studied incentives in Gale Shapley mechanism with short preference lists is Immorlica and Mahdian [41]. They show that, when doctors have short lists, with high probability the stable matching is unique and truth-telling is the best response for hospitals. Arnosti [1] studies the quality of the outcomes in terms of the welfare under different (hospital) preference structures. However, both papers assume that the short preference lists of applicants describe their true full preferences, thus ignoring the strategic aspects on how doctors choose what hospitals to list.

The literature of college admissions also considers the issue of students being able to apply to a select set of schools. [36] and [15], study the problem with two colleges (similar to our model of two tiers of hospitals) and heterogeneous students who must decide which of the two schools they apply to. Avery and Levin [4] study a game theoretic model of early decision systems used in US universities, where students can apply early to a single school, either just signaling their preference, or committing to the school if admitted. These papers consider the decentralized problem where not only doctors are strategic on where to apply, but also hospitals are strategic on setting their admissions standards or early admission policies. In contrast, we consider a centralized system but each tier has multiple hospitals, and we allow doctors to submit a list containing more than one (or two) hospitals. While this adds realism to the model and the insights obtained, it also adds significant complications to the analysis.

In the papers mentioned so far, as well as in our work, the applicants know their own preferences and decide to apply to hospitals or schools aiming to maximize the quality of the school they get accepted at. Kadam [42], Drummond et al [23], and Kleinberg et al [44] study models where costly exploration is necessary to discover the applicant's ranking. Due to the discovery cost, these models also lead to exploring (and then ranking) only a limited set of options. Drummond et al [23] and Kadam [42], study a two-stage model, in which doctors choose a short application set and interview at this set of hospitals. In both cases, they consider markets where both doctors and hospitals are (almost perfectly) vertically differentiated.

On the other hand, we consider a model where many ex-ante identical students (those in the same tier) might be competing for the same school. Also, we provide bounds on the overall welfare implication of the system design.

4.2 Model

We consider a two-sided market consisting of doctors and open positions in hospitals (hospitals for short). For simplicity, we assume there are N doctors and N positions, and that each hospital has only one open position, so we focus on one-to-one matching. Each agent has a full ordered list of preferences for agents on the other side of the market. Doctors, however, can only submit to the system a preference list of length K, for some parameter K of the system. Hospitals, on the other hand, submit a list ranking all doctors.

Once doctors and hospitals submit their preferences, a doctor-proposing deferred acceptance algorithm is run to determine the final assignment of doctors to hospitals. The algorithm starts with all doctors unmatched. In each step, all unmatched doctors, who have not yet exhausted their options, apply to their most preferred hospital among those to which he/she has not yet applied. Now each hospital tentatively accepts their most preferred doctor from the doctors now applying and the one who has been tentatively assigned to the selected hospital, and rejects all other applicants. The procedure is repeated until all unmatched doctors have been rejected from every hospital in their lists. Naturally, the algorithm uses the lists submitted by the doctors (not their true preferences) and the hospitals' true preferences over doctors.

To model the preferences of the agents, we assume that each side is divided into two tiers, high and low. These tiers allow us to divide the agents into types, which we use to capture a vertical component in preferences by assuming that all doctors prefer any high tier hospital to every low tier one, and similarly all hospitals prefer any high quality doctor to any low quality ones. Under this assumption, agents' true preferences rank first all agents in the top tier, then all agents in the low tier. Within each tier, we assume that the preferences are drawn independently and uniformly at random. Having only two tiers on each side allows us to simplify the analysis and the exposition, while clearly allowing us to distinguish between doctors applying to "reach", in level, and "safe" hospitals, a key feature of our model.[3] However, the arguments used to characterize the equilibrium can be to extended to multiple tiers. For most of the chapter, we will assume that there are n high quality doctors and rn low quality doctors, and n high hospitals and rn low hospitals[4], with $N = n + rn$. Under this assumption, the Gale-Shapley matching algorithm with full preference lists will match high doctors to high hospitals, and low doctors to low hospitals.[5]

We consider the game when doctors are allowed to list only up to K hospitals, for parameters $K << n$. As argued in the introduction, in general doctors will not want to simply list their top K choices of hospitals; high doctors might prefer to list a safe option to avoid being unmatched, while low doctors who also prefer the high hospitals, should know that they are unlikely to obtain such a match. A natural strategy for each doctor would be to apply to the top K hospitals within its own tier: top K high hospitals for high quality doctors, and top K high hospitals for low quality doctors. We call this the *simple application system.* Since doctors

---

[3] When a low quality doctor chooses to apply to a high hospital, we view this as a *"reach"* application, and when a high quality doctor applies to a low hospital, we view this as a *"safe"* application.

[4] When rn is not a integer number, we just approximate it to the closest integer

[5] This assumption is not used in the equilibrium characterization in Section 4.3.

are making offers, and hence the algorithm is truthful for them, doctors should order the same-tier hospitals using their true preference, and the only strategic decision they need to make is selecting the hospitals to apply to.[6]

The question we want to study is how many reach or safe hospitals do doctors apply to, and what is the effect of this selfish choice on the quality of Nash equilibria in terms of social welfare. To study these questions, we must assign values to the agent's matches. We want to focus on the low and high distinction, and hence will use the following very simple approximation for values. The value that a doctor (from either tier) derives from getting matched to a low hospital is 1, and the value of assigned to a high tier hospitals is $v > 1$. For simplicity, we have assumed that the value is given by the tier and is independent of the actual rank of the partner in his preference list, that is, we assumed that the difference in value within each tier is negligible compared to the difference in the value of the tiers. Similarly, the value for a hospital (from either tier) to be assigned a low quality doctor is 1, while the value for a hospital to be assigned a high quality doctor is v, again independent of the actual rank on the hospital's preference list. We define the *social welfare* to be the sum of values obtained by all agents in the market. Under these assumptions, it is easy to see that the maximum achievable social welfare is $2vn + 2rn$, $2v$ for matching a high doctor to a high hospital and 2 for matching a low doctor to a low hospital. Despite its simplicity, this model allows us to capture correlation in preferences through the tier structure and, unlike previous work, we also allow for horizontal differentiation through the within-tier heterogeneity.

---

[6] It is not hard to extend the result of Immorlica and Mahdian [41] to our tiered model to show that with large n and constant K algorithm is also strategy proof for hospitals with high probability, so thus we assume that hospitals report their preferences truthfully.

Equilibrium analysis. We study a “large market approximation” of this model, as introduced by Arnosti [1]. In particular, we study the outcomes in the limiting case where n grows (i.e. the numbers of doctors and hospitals grow), while the ratio of doctors to hospitals, the proportion of doctors and hospitals in each tier, the lengths of doctors' lists, and the valuations for each hospital/doctor are held fixed. For example, when $K = 1$ and we consider the simple application system of all doctors applying to the top hospital in their own tier, we obtain that an $(1 - 1/e) \approx 0.63$ fraction of the doctors will be matched to hospitals.[7]

When $K > 1$ —doctors apply to more than one hospital— as n grows we can approximate the probability that a doctor is accepted at a hospital as fixed probability, independent of previous applications. Given the tiered structure of the hospital's preferences, we can think about the final allocation resulting from two separate steps, where first the allocation for high-type doctors is finalized and, once this is fixed, the low-type doctors are allocated. Then, focusing on the high tier of doctors, sending some k applications to high tier hospitals, the above approximations the probability p that one single application is accepted satisfies the following fixed point equation.

Proposition 6. *[1] Using the above large market approximation with each high doctor applying to k high hospitals, the probability of a single application getting accepted is defined by equation:*

$$(1-p)^k = e^{-(1-(1-p)^k)/p} \tag{4.1}$$

Proof. Since each application is accepted independently with probability p, the

---

[7] To see why, note that a high hospital will hire its best applicant, so the only hospitals unmatched are those with no application; this occurs with probability $(1 - 1/n)^n$ using the fact that the choice of all applicants is a random hospital in their tier. The approximation then follows as $(1 - 1/n)^n \to 1/e$ as $n \to +\infty$.

probability that a doctor with k applications eventually gets matched is $(1-(1-p)^k)$, and so the expected expected number of hospitals a single doctor makes an offer to throughout the Gale-Shapley process to is

$$1+(1-p)+(1-p)^2+...+(1-p)^{k-1}=\frac{1}{p}(1-(1-p)^k)$$

Each hospital remains matched if it receives any applications. Using the approximation that all these offers are independent, the probability that a hospital didn't get an offer is then

$$(1-1/n)^{\frac{n}{p}(1-(1-p)^k)}\approx e^{-(1-(1-p)^k)/p}$$

and so the expected number of matched hospitals is then $n(1-e^{-(1-(1-p)^k)/p})$. The equation claimed by the lemma then follows, as the number of matched hospitals is the same as the number of matched doctors. □

Analogous formulas also apply when the number of doctors and hospitals is not the same on the two sides of the market. Furthermore, these formulas can also be extended to the case where different doctors apply to a different number of hospitals; this becomes useful when, for example, doctors use a mixed-strategy. In addition, similar formulas allow us to derive the probability that an application from a low-type doctor is accepted. In that case, we need to take into account that, every time a low doctor applies to a hospital, there exist some probability that the hospital is already taken by a high-doctor, and thus his application will be automatically rejected due to the tiered structure in preferences. We denote the probability that a high and low hospital is already taken by $(1-\alpha_H)$ and $(1-\alpha_L)$ respectively. We make use of these large market approximations when deriving the results in this chapter.

4.3 Existence and Structure of Equilibrium

We call an equilibrium *symmetric* if all high type, as well as all low type doctors use the same strategy, while different types are allowed to have different strategies. As discussed above, the strategic decision of a doctor can be thought of as an ordered pair $(k, K-k)$, where the first (resp. second) component indicates how many high-type (resp. low type) hospitals he/she is ranking.

Theorem 10. *In a market with two tiers, a symmetric equilibrium always exists and is unique, and for each type of agents* $i \in \{L, H\}$*, the equilibrium strategy is either pure, or is a randomization between consecutive strategies* $(k_i, K-k_i)$ *and* $(k_i+1, K-k_i-1)$ *for some* $k_i \in \{0, \ldots, K\}$*.*

We provide a proof sketch here for high doctors only, which one can think of separately due to the tiered structure of the hospital's preferences. The proof for the low doctors is the same, except that it must account that some hospitals are already taken by a high doctor (in that case, the low doctor gets automatically rejected). We divide the proof into two lemmas, first about the structure, and then stating the existence of equilibria. We give a short sketch of their proofs. We include full details of the proof in Appendix B.1.

Lemma 12. *If a symmetric equilibrium exists for high doctors, it is either an equilibrium in pure strategies, or it is a mixed strategy equilibrium where agents randomize between consecutive strategies* $(k, K-k)$ *and* $(k+1, K-k-1)$*.*

Sketch. Consider a symmetric equilibrium. Under the large market approximation discussed in the Section 5.2, each application of a doctor to a high hospital is accepted with some probability p, and an application to a low hospital is accepted with a different probability p'. Given the values p and p' we can express the utility of a high doctor as a function $f(k)$ of the number of applications to high hospitals. For example $f(K) = v(1-(1-p)^K)$ as $(1-p)^K$ is the probability under this model that none of his K applications are accepted.

Given p and p', the best response of a single agent is to maximize $f(k)$ over integers. (By the large market assumption, p and p' remain unchanged regardless of the agent's action.) The main observation is that the function $f(k)$, when viewed as a function of a real variable k, is strictly concave (see the Appendix for details). This implies that the maximum over integers is either a single integer or two neighboring integer values. $\square$

To simplify notation, for a symmetric strategy of high doctors let X be the expected number of high application of a doctor, and $k = \lfloor X \rfloor$. Let $p(X)$ (resp. $p'(X)$) be the probability that a single application of a doctor is accepted by an individual high (resp. low) hospital he applies to, when the strategy is given by X. These probabilities are implicitly defined by equations analogous to Equation (C.1). Despite this implicit definition, one can prove that $p(X)$ is continuous and monotone decreasing, and $p'(X)$ is continuous and monotone increasing. Intuitively, the more slots used to list high hospitals, the smaller the probability that a single application succeeds in getting accepted, and this relation is continuous.

Lemma 13. *A symmetric equilibrium for high doctors must exist.*

Sketch. The idea of the proof is to express the equilibrium condition using $p(X)$ and $p'(X)$ defined above. If X is an equilibrium, and $k = \lfloor X \rfloor \neq X$, then doctors should be neutral between applying to either k or $k+1$ high-type hospitals. Consider the possible $k+1$st application sent to high hospitals. At this point, the choice of sending one more applications to high hospitals, versus sending all remaining $K-k$ applications to low hospitals must have equal value.

$$p(X)v + (1-p(X))(1-(1-p'(X))^{K-k-1}) = (1-(1-p'(X))^{K-k}) \tag{4.2}$$

It is not hard to see, similar to the proof of Lemma 12, that if the above equation holds, then X defines an equilibrium.

The idea of the proof is to consider the function $g(X)$, which is the difference of the two sides of the above equation. We have that $g(0) > 0$ and $g(K) < 0$, g is continuous at all non-integer points, and is strictly decreasing. So there is either an X with $g(X) = 0$, or an integer point k where $g(X)$ changes sign. We show that in the latter case, k defines a pure strategy equilibrium. ☐

4.4 Equilibrium Efficiency

To study the efficiency of equilibrium, we can consider two different benchmarks:

- Full list optimum, or OPTIMUM: the matching resulting from running the Gale-Shapley algorithm when doctors have a full list. In this case, everybody is matched to an agent in his own tier. Noting that the contribution to social welfare of a match on top is $2v$ and of a match in the low tier is 2, the total welfare of this benchmark is $2(v+r)n$, independent of K.[8]
- SIMPLE: the matching resulting of every doctor sending all applications to hospitals in their own tier. We focus on the even simpler benchmark of simple

---

[8] This benchmark is overestimating the actual welfare, as it assumes that doctors can rank all hospitals. While it will be fairer to compare the Nash welfare to that achieved by the optimum under short lists, we use full-list optimum as it is easier to compute.

with $K = 1$,which makes the computations smoother, referred hereafter as SIMPLE.

In the SIMPLE benchmark (with $K = 1$), the probability that a hospital is open in high tier is $(1 - 1/n)^n \approx 1/e$, similarly also $1/e$ in low tier. Therefore with large number of doctors, there are $n\frac{e-1}{e}$ matchings in the high tier and $rn\frac{e-1}{e}$ in the low tier with the total welfare $2(v + r)n\frac{e-1}{e}$. For simplicity of computation, we assume $v \geq e/(e - 1) \approx 1.58$; this assumption is somewhat reasonable as the tiers represent vertically differentiation and hence there should be a non-negligible difference between high and low tiers.

We first consider the case where $K = 1$, as it allows for closed-form computations as opposed to relaying on the fixed-point equations. We then generalize our bound for $K > 1$.

Equilibrium when $K = 1$. We now provide a closed-form characterization of the equilibrium when $K = 1$ and $v \geq e/(e - 1)$. First, high doctors apply to a high hospital with probability 1, as even when all high doctors apply to high hospitals, the probability that an application sent by a doctor is accepted is $(e - 1)/e$ which produces an expected value of $v(e - 1)/e > 1$. Now assume low tier doctors send an application to a high tier hospital with probability x, then the expected number of applications to top is xrn (with a very small error term for large values of n). Now the probability that a high hospital accepts a low doctor comes out to be $(1 - e^{-xr})n/(exrn) = (1 - e^{-xr})/(erx)$. Also, all rn low hospitals are available and so the probability of an application to such a hospital being accepted is $(1 - e^{-(1-x)})/(1 - x)$. This is a Nash equilibrium if either (1) $x = 1$ and $v(1 - e^{-r})/(er) \geq 1$, or (2) $x = 0$ and $v/e \leq (1 - 1/e)$, or (3) $0 < x < 1$

and

$$\frac{v(1-e^{xr})}{exr} = \frac{1-e^{-(1-x)}}{1-x}. \tag{4.3}$$

The social welfare is $2vn(1-1/e)+n(v+1)(1-e^{-xr})/e+2rn(1-e^{-(1-x)})$.

Note that even the special case with $K=1$ already exhibits some of the issues we discussed in the introduction. While high doctors apply to hospitals in their own tier, low type doctors already have incentives to apply to a “reach” (high) hospital with some probability if $v > e-1 \approx 1.72$.

Bounds on the efficiency of the equilibrium. It is natural to conjecture that welfare of the Nash equilibrium should be at least as good as that of SIMPLE, as allowing low quality doctors to target high tier hospitals has significant potential to improve welfare , even though it can also hurt welfare by leaving some low hospitals unoccupied. However, we found that this is not necessarily the case even for large v.[9] Two forces can contribute to this phenomena. Doctors applying to top-tier hospitals create a negative externality for each-other by increasing the congestion, and thus decreasing the probability that other applicants get accepted. Second, the low-tier doctors trade-off only the value they expect to obtain by getting an appointment in hospitals of different tiers (v versus 1). However, the trade-off for social welfare is $v+1$ versus 2, as the welfare of the hospitals also needs to be accounted for. Although the gap can exist, in the examples we found the decrease in welfare to be minimal. Next we show that this is also true in the worst case.

Proposition 7. *The social welfare in Nash equilibrium in the case when* $v \geq e/(e-1)$ *and* $K=1$ *is at least a* $\frac{e}{2(e-1)} > 0.79$ *fraction of the welfare of* SIMPLE.

---

[9] Examples are provided in Table B.3.2 in the appendix.

Sketch. By our previous discussion, when $v > e/(e-1)$ every high doctor applies to high hospital in equilibrium; thus, the welfare produced by the high doctors in the equilibrium and SIMPLE agrees. Next, using equation 4.3 we can find some v such that strategy x is an equilibrium for low doctors. Our goal is to find the equilibrium with highest loss of welfare. We show this happens when x converges to 1 and the efficiency in this case is more than $\frac{e}{2(e-1)}$ fraction of SIMPLE. The intuition behind is that being strategic makes agents to apply more to top than what would be optimal in terms of welfare: by matching to a high hospital they obtain value v versus 1 for matching low, while the social welfare the contribution of those matches is $v+1$ and 2 respectively. □

We next leverage the previous result to obtain a bound on the welfare of the Nash equilibrium for the general case with $K \geq 1$ applications.

Theorem 11. *When $v \geq e/(e-1)$, the loss of efficiency of equilibrium when doctors have list of size $K \geq 1$ compared to* OPTIMUM *is at most a factor of* 2*.*

Remark. Note that compared to the optimal welfare with full preference lists, the equilibrium welfare can be as low as only a $(1-1/e) \approx 0.63$ fraction: this is the case with a single tier when $K = 1$, as we expect random applications to miss a $1/e$ fraction of hospitals.

sketch of Theorem 11. By Proposition 7, when $K = 1$ SIMPLE achieves $\frac{e-1}{e}$ fraction of OPTIMUM social welfare. Therefore the welfare in the equilibrium for $K = 1$ is at least $\frac{e}{2(e-1)} \frac{e-1}{e} = \frac{1}{2}$ fraction of OPTIMUM welfare. As OPTIMUM is by definition independent of K, it suffices to show that the welfare of the Nash equilibrium cannot decrease as K increases. We show this in two steps. Full proof is in

Appendix B.2. As a first step, we compare the welfare of SIMPLE with the welfare achieved in the equilibrium of a game in which high doctors can submit a list of length $K > 0$ and low ones submit a list of length 1. To see why this also satisfies the bound of Proposition 7, note that as K increases, we will show in Proposition 8 that high tier doctors have more applications on top. This will certainly increase the welfare of the high doctors. As shown in Lemma 26, this increase will at least compensate for the loss incurred by low tier doctors.

Next, we show in Lemmas 27 and 28 that social welfare of this compared to SIMPLE is further improved when low doctors also have a list of size K. The two lemmas break the proof into two cases, if all doctors send at least one application to the low tier, then welfare is at least as high as the SIMPLE benchmark. In the case, when doctors send less than one application to the low tier, we use the equilibrium properties to show that welfare at this equilibrium is at least as high as the welfare when only the high tier could send $K > 1$ applications. □

Remark. It is important to note that the result in Theorem 11 is indeed dependent on our assumption that the number of high tier hospitals is the same as high tier doctors. The quality of Nash equilibria can be a lot lower than the optimum when there is a shortage of high tier hospitals. For an example, assume v is huge, we have n top tier doctors, and only αn top hospitals $\alpha << 1$, $(1 - \alpha)n$ low tier hospitals, and $r = 0$. Now all top tier doctors applicants apply only to high tier hospitals if say $v > 1/\alpha$ achieving overall welfare less than $2v\alpha n$. On the other hand, the optimal social welfare is $2v\alpha n + (1 - \alpha)(v + 1) \approx v(1 + \alpha)n$, so the ratio is $2\alpha/(1 + \alpha) \to 0$ as $\alpha \to 0$.

4.5 Properties of Equilibria

In this section we study how the equilibrium strategies and the welfare achieved at the equilibrium are affected by the primitives of the market. While a few properties we can prove formally, for many we rely on simulations. The basic parameters used in our simulations are

- The length of the preference lists K varies between 1 and 10.
- The number of doctors and hospitals of the same tier agree. However, we vary the proportion of high to low agents, by considering the proportion of high doctors, $1/(r+1)$ to be equal to $0.1, 0.2, 0.3, 0.4, 0.5$.
- As usual, the value of the low type is assumed to be equal to 1. For the value of the high type, we consider $v = 1.001, 1.01, 1.05, 1.1, 1.25, 1.5$, and all the way up to $v = 10$ using .5 increments.

The simulation results can be found in Appendix B.3.2, but we describe our main findings next.

Effect on Equilibrium Strategies. First we consider the equilibrium strategy of high doctors as a function of K, the number of applications, and a function of r, the ratio of low and high quality doctors and hospitals. We defer the proof to Appendix B.3.1.

Proposition 8. *Let $X^*_H(K, v, r)$ denote the equilibrium strategy for the high type doctors when the length of lists is K, the ratio of low to high doctors is r, and the ratio of values of high and low match is v, as we defined in Section 4.3. Then, $X^*_H(K, v, r)$ is a monotone increasing function of K, v, and decreasing in r, e.g.,* *the number of applications from the high doctors to high hospitals strictly increases with the length of the list.*

Sketch. Let the equilibrium strategy be $X^* = X^*(K)$ with the upper integer part $k = \lceil X^* \rceil$. Consider the solution where with $K+1$ applications, doctors still send X^* applications to high hospitals. We show that doctors sending k applications on top are better off than doctors sending $k-1$ applications. This shows that in equilibrium there must be more applications on top implying $X^*(K+1) \geq X^*(K)$. Using similar arguments, one can show the dependence on both v and r. □

One would expect the equilibrium strategy to increase "smoothly" in K: if one extra slot is available (K increases by one), the number of high hospitals ranked will increase by at most one. Surprisingly, this is not true in general. As an example, consider an instance with the same proportion of high and low doctors and high and low hospitals and $v = 1.001$. Then, when $K = 5$, the high-type doctors play the pure strategy $(1, 4)$, and when $K = 6$ they play $(3, 3)$. This discontinuity can be understood as follows: as X_H^* is increasing in K, by adding an extra slot to the list more high-type doctors will match with high-type hospitals (on average). This implies that the availability of low type hospitals increased as less high-type doctors will attempt to match to them; thus, a doctor needs to list less low-type hospitals to obtain the same probability of being matched to one of them, potentially devoting more space to listing high-type hospitals. Although not true in general, we tested it computationally and it appears that if v is "big enough" then $X_H^*(K) < X_H^*(K+1) \leq X_H^*(K) + 1$.

On the other hand, the effect that a change in primitives has in the strategy for low doctors X_L^* is more involved. For instance, one would like to establish a monotonicity result in K for the strategy of the low-type. However, both situations might arise and the strategy of the low-type doctors can either increase or decrease. This is maybe not so surprising: as X_H^* is increasing in K, so more high-type doctors will match with high-type hospitals on expectation and thus less hospitals will be free. Hence, even though low doctors might use the extra space in their list to try to reach out to high hospitals, they are less likely to be accepted and thus the value of listing a high-hospital also decreases. In general, we found that when v is relatively high, $X_L^*(K)$ is weakly increasing; even though as K increases it is less likely that an application will be accepted, the fact that v is high makes it attractive to spend some slots reaching to high hospitals. (This is not a surprise; if v is sufficiently high, all high doctors apply high and thus they don't have a direct externality on the low doctors - low hospitals matches.) In addition, when v is close to one, $X_L^*(K)$ is decreasing. This can be understood by the fact that, when K is low, high doctors are very likely to be matched to low-type hospitals and thus low hospitals become less attractive to low doctors as they might already be taken. However, as K increases, X_H^* also increases, making high hospitals less attractive (more of them are taken) and low hospitals potentially more attractive as less high doctors might be matching low. Finally, for intermediate values of v (e.g, $v = 1.5$) we observe that X_L^* might decrease and then increase or the other way around, depending on r. See Figure 4.5.

Regarding the dependence on v, when K is low, X_L^* might either increase or decrease as v increases. On the other hand, it always increases for high K. The intuition is similar that of the dependence on K; when K is low, an increase in v can have a huge impact on X_H^*, but when K relatively high, most applications of high doctors will be to high hospitals anyway, so the impact in X_H^* is less. Finally, X_L^* always increases as r decreases. This is because of two effects, both due to

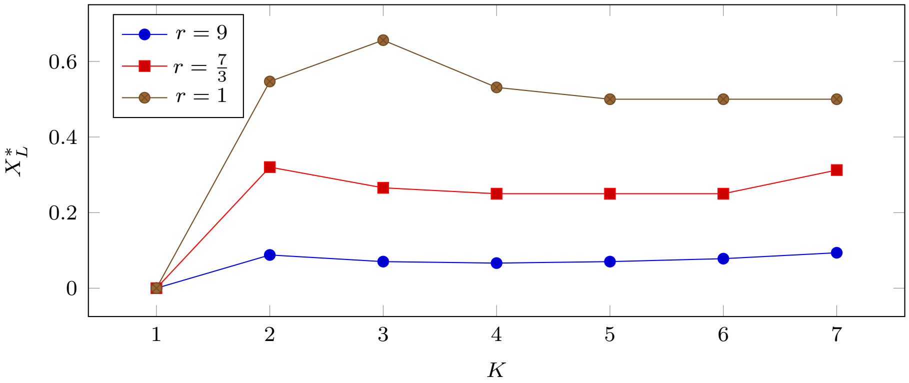

Figure 4.1: Equilibrium strategy for the low type (X_L^*) as a function of the length of the list K for different market compositions. In all cases, $v = 1.5$.

X_H^* decreasing. First, more high hospitals remain available to low doctors, which makes them more attractive. Second, there are relatively more high doctors and less low hospitals, so the number of low hospitals available to low doctors decreases.

Effect on Welfare. Regarding the social welfare produced by the Nash equilibrium solution we note that it is always increasing in K. In general, the number of within-tier matches increases as K increases, which agrees with the intuition that if the doctors could submit the full-preference list, the assortative matching would arise. The welfare is also increasing in v.

Surprisingly, the welfare is not always increasing in the ratio of high to low doctors, even though the percentage of high agents for a fixed size population increased. In particular, when v is close to one, the welfare decreases as a higher proportion of doctors and hospitals is of high type. This is due to the fact that high doctors split their applications between high and low hospitals; hence, it is more likely that low doctors will apply to a hospital that is already taken by a high doctor, thus wasting the applications and reducing the total welfare.

CHAPTER 5

TWO-SIDED MATCHING WITH LIMITED NUMBER OF INTERVIEWS

5.1 Introduction

We continue studying two sided matching markets where applicants apply to positions. As discussed in Chapter 4, an important aspect that is not modeled under the classical deferred acceptance algorithm is the fact that individuals can only consider a limited set of options. In the previous chapter, we focused on the limit on the number of applications that applicants can send; however, typically both sides of the market are limited in the number of options they can consider.

Consider the national residency matching program (NRMP) as a prominent example of matching under the deferred acceptance algorithm. In NRMP, there are nearly 5000 residency programs and applicants have to limit their choices when applying. Similarly, interviewing consumes significant time of the hospitals; therefore, the hospitals can only grant interviews to a limited set of doctors. An alternative view of the limitations of the employer side is a limit on the offers for one position: based on interviewing all or none of the applicants, they may rank only a limited number of their favorite applicants. Such selective offer making is observed in the academic job markets, e.g., when some high ranked departments prefer to wait till next year, rather than make offers to applicants they liked less.

The goal of this chapter is to study the effect of the number of applications and the number of interviews on the resulting matching. We think of our model as a simplified version of the matching process of the national residency matching program. In this program, the matching mechanism has two stages. In the first stage, doctors express their interest to a set of hospitals, and hospitals choose to interview a subset of the applicants. In the second stage, both doctors and hospitals submit their ordered preference list over the set they interviewed with. The centralized matching market uses these preference lists to perform the deferred acceptance algorithm and outputs the final matching.

We will primarily consider a symmetric market, and as a primary metric of the social welfare of the system, we will mainly consider the size of the matching found. In asymmetric markets, where some applicants and some jobs are a-priory better, and preferred by most applicants, the ability to consider only limited set of choices gives rise to strategic issues (with top candidates getting too many offers, and a number of jobs remaining unfilled due to this). Our focus here is large symmetric markets, where it will be dominant strategy for both sides to report their true preferences. We find a number of surprising effects of this limited choice, resulting in two policy recommendations for the system designers.

- With limited capacity for granting interviews, social welfare may be maximized by severely limiting the number of applications an applicant can send.
- The best social welfare is achieved by a fair system, where all applicants can send the same number of applications. Allowing a small subset to send even just one additional application decreases the overall welfare of the system.

The size of the resulting matching is a classical measure of quality of the matching system. For example, the CRA report [32] from 2010 recommending changes to the computer science job-market reported unfilled jobs as a measure of need for improvement in the mechanism. If there were no constraints on the number of interviews for applicants or programs, and both parties were willing to list everybody on the other side, the outcome of deferred acceptance algorithm would be a matching of maximum size (fully matching the smaller side of the market). With limited number of applications and interviews, the resulting matching can be much smaller. There are a number of different sources for this loss in the size of the matching. When doctors apply to a limited set of hospitals, there will be hospitals that have received less applications than their available positions. Even if hospitals receive many more applications than their positions, they might end up with free slots in case the doctors they interview choose other hospitals. Limiting the number of interviews by hospitals further reduces the number of potential matches, e.g., doctors may end up with too few or no interviews, and even with interviews, they may not get an offer.

Our Results. We study the effect of limitation on the number of applications and interviews in matching markets and show surprising effects on the matching outcomes.

- We show that in markets where jobs can only consider a limited number of candidates for interview (or only willing to make offers to a few of their favorite applicants) it is best to also have a low limit on the number of application a candidate can send. Note that we show this effect in Section 5.3 in symmetric markets, where a-priory all jobs and all candidates are equally likely to be great, and the judgments of different employers and different candidates are independent.

  Although it might seem intuitive that allowing applicants to apply to more positions increases the size of the resulting matching, we show that this is not true by computing the expected size of the matching formally.

When ranking of applications is correlated, such selective offer making is well-known to lead to many departments unable to fill their positions, as top applicants get many offers, and others don't get any. Notice that we observe the same phenomenon in a model where rankings of different hospitals are independent, and drawn from the uniform distribution.

- We show that to achieve the highest social welfare it is best to have a system that is totally fair: all applicants can send exactly the same number of applications. Allowing a small subset to send even just one additional application decreases the overall welfare of the system.

  To show this, we allow a small subset of applicants to apply to one extra position. Surprisingly, we find that the increase in applications in this way decreases the size of the matching. On the other hand, if starting from the case when all doctors apply to the same number of hospitals, we restrict a small subset of doctors to one fewer application, this too decreases the size of the matching.

  The illustration of this result is a scallop-shape function of the size of the marching as a function of the number of applications per candidate as shown in Figure 5.4. This illustration shows the role of fairness in maximizing matching size. While the scallop shape is more pronounced in the outcome with a very small number of interviews, the same effect remains true for any limit, as shown by our formula for the expected size of the resulting matching.

- In Section 5.4 we show that the same phenomena exists even in unbalanced markets, where the number of applicants and positions is different. In unbalanced markets, the optimum number of applications for social welfare (the size of the resulting matching) is different compared to the balanced case, and depends on the ratio between the two sides of the market. However, in this more general model, the effect of not treating everybody equally still can hurt the efficiency.

- In section 5.5 we study an extension when the hospitals and doctors' preferences are a-priori uniform, but become correlated after interviewing, and find that the same phenomena remains.
- In Section 5.3 we also study the effect of different methods for distributing the applications, when the total number of allowed applications is fixed. We find the optimal method of distributing the applications when each hospital selects only one doctor. We show that in this case, the maximum matching is achieved when doctors either apply to k or $k+1$ positions; this allocation dominates any other way of distributing the applications, showing to maximize the size of the resulting matching, it is best to treat all applicants as evenly as possible.
- Finally, in Section 5.6 we study the same phenomenon in a two-tier model: both doctors and hospitals are divided into high-tier and low-tier, and all parties prefer a match to a high-tier partner to a low-tier one. To evaluate social welfare in this model, we need to model this difference in value. We use a parameter v to model the extent to which both sides value high-tier matches compared to low-tier. In this model, both applicants and hospitals need to make a strategic choice about how to divide their applications and interviews between the two tiers. In the symmetric markets of Sections 5.3 and 5.4 increasing the number of interviews can only increase the size of the resulting matching. In the two-tier system modeling different qualities of applicants and positions in Section 5.6 we find that due to the strategic behavior of doctors (aiming to get employed by better hospitals) the social

welfare can sometimes be improved with limiting the number of interviews.

Roadmap. In Section 5.2, we formally reintroduce the large market model of [1] that we use throughout the chapter. In Section 5.3 we analytically compute the expected size of the matching with limited number of interviews and applications. We start by considering the simple special case with just one interview in Sections 5.3.1 and 5.3.2, which is then extended to the more complex case of multiple interviews using the deferred acceptance algorithm, in subsection Section 5.3.3. In Section 5.4 we show that the same phenomena remain true even in unbalanced markets (when one side is significantly different size than the other).The results of the last two sections are obtained via simulation. In Section 5.5 we study a model where the hospitals and doctor's preferences are a-priori unform, but become correlated after interviewing. In Section 5.6 we consider a two-tiered market when a subset of doctors as well as a subset of hospitals (or jobs) are viewed as better by all, and study the Nash equilibrium of the resulting system.

Related Literature. Papers studying matching with short lists have different viewpoints on how the short lists are selected. For example, Immorlica and Mahdian [41], Kojima and Pathak [45], and Arnosti [1] study matchings when preference lists are inherently short, i.e. the participants prefer to stay unmatched rather than being matched outside their preference lists. In contrast, Avery and Levin [4], Kadam [42], Drummond et al. [23], and Beyhaghi et al. [8] study how applicants make the strategic choice to select a limited number of positions for their short preference lists. In this paper, because of the uniform preference of participants we do not focus on their strategic behavior.

Arnosti [1], Lee and Schwarz [46], and Kadam [42] study efficiency of matching either as matching size or social welfare in presence of short lists. Arnosti evaluates social welfare under different preference models. However unlike our model, in [1] the preference lists are limited only for one side of the market. Similar to us, Lee and Schwarz study matching size in a setting with an interview stage, where the ex-ante preferences are i.i.d. However, they solve the problem of some doctors not receiving any offer while others receive many, by coordinating the set of doctors that each hospital interviews. In contrast, we assume that the interviews are selected in a decentralized way without imposing coordination; and our solution is to limit the number of applications and let the applicants have the same number of applications. Similar to our work, Kadam studies a model with limit on both sides of the market and shows that limiting the length of lists can have positive effects of the size of matching. However, in [42] these effects are due to almost common preferences: When some doctors are more preferred, with a stricter limit on the interviews for doctors, the more preferred doctors do not accept interviews with less preferred hospitals. We show that limiting the length can be helpful in the exactly opposite case where all participants are identical in terms of popularity.

5.2 Model

We consider a two-sided market consisting of doctors and open positions in hospitals (hospitals for short). We assume there are n doctors and rn positions, and that each hospital has only one open position, so we focus on one-to-one matching.

We start with a model where there is a universal limit on the number of applications by doctors and the number of interviews by hospitals. The fixed limit on the number of applications by a doctor is inspired by NRMP that grants all doctors 10 initial applications. We relax this assumption later and compare the effect of allocating different number of applications to different doctors. We also have a fixed limit on the number of interview conducted by hospitals, modeling that hospitals spend more or less the same time and budget on interviews.

We consider a two-stage matching mechanism. In the first stage, doctors apply to a number of hospitals and ask for interview. Hospitals then choose a subset from their applications to conduct interviews. In the interview process, doctors and hospitals gather information and learn their preferences. In the second stage, at the end of interviews, both doctors and hospitals order the list that they interviewed with, based on their preferences. They submit their ordered preferences to the system which performs a doctor-proposing deferred acceptance algorithm to determine the final assignment of doctors to hospitals. The algorithm starts with all doctors unmatched. In each step, the algorithm simulates all unmatched doctors, who have not yet exhausted their options, propose to their most preferred hospital among those to which he/she has not yet proposed. Now the algorithm simulates that each hospital tentatively accepts their most preferred doctor from the doctors now proposing and the one who has been tentatively assigned to this hospital, and rejects all other doctors. The procedure is repeated until all unmatched doctors have been rejected from every hospital in their lists.

We use the following terminology throughout the whole chapter.

Application: The procedure in the first stage where the doctors submit an application and ask for interview.

Grant interview or reject for interview: The procedure in the first stage where the hospitals grant interviews to a number of doctors that have applied for the position and reject the rest.

Valid application: Any application that is granted an interview is a valid application. Other applications are *invalid.*

Proposal: The procedure in the second stage that simulates doctors proposing to hospitals as steps of the deferred acceptance algorithm.

Offer: In the special case where the hospitals limit their number of interviews to one, conducting interviews is of no use because the hospital offers the position to the selected doctor in any case. Therefore we use offer instead of interview in this case.

Modified implementation: For the sake of analysis, it is useful to define a *modified implementation* of the matching procedure. Similar to the original implementation, in the first stage doctors apply to hospitals using all their applications and hospitals conduct interviews with a subset of their applicants. The difference between the two implementations arises in the second stage. In the modified implementation doctors submit their ordered preferences including both the hospitals they interviewed and the hospitals that rejected them for interview. Therefore when the system simulates doctor-proposing deferred acceptance algorithm, doctors are allowed to propose using any of their applications including invalid ones. Since the invalid applications do not exist in hospitals' lists they get rejected immediately and the outcomes of both implementations are the same. In contrast to the modified implementation, we refer to the main procedure previously defined as the *original* implementation.

Semi-proposal: The procedure in the second stage of the modified implementation that simulates doctors proposing to hospitals as steps of the deferred acceptance algorithm.

In the beginning of the procedure, everybody has prior preference over the other side. For Sections 5.3 and 5.4 we assume that everybody's preference comes independently from uniform distribution over the other side. Both doctors and hospitals refine their preference order for the list they interviewed. However, we assume that the overall distribution of doctors preferences and hospitals preferences can be treated as outcome of new independent draws from uniform distribution. Refinement in the way mentioned, does not have an effect on our analysis. For analytic purposes we can assume that both the initial distribution, as well as the ranking after the interviews, are uniformly random.

Lemma 14. *Suppose prior to interviews preferences are drawn independently from uniform distributions, and after interviews, the overall distribution of preferences are uniform and independent. In this case, assuming that the preferences remained unchanged leads to the same analytic results including the same matching size.*

We assume that throughout the whole process, there is no coordination between doctors or hospitals. If hospitals knew which applicants had more/less interviews, they may prefer to grant interviews to doctors with less interviews, as these are more likely to be matched with. Our model assumes no such strategic behavior.

For most of the chapter we focus on the size of the matching as our notion of efficiency. We define social welfare as the ratio of size of the matching outcome to the size of the maximum matching. Since we assume all doctors/hospitals prefer to be matched to any hospital/doctor rather than being unmatched, in a maximum-size matching everybody on the less populated side of the market is matched.

Definition 6 (Social Welfare)**.** *We define social welfare of the matching outcome as the ratio of the size of the matching outcome compared to the size of the maximum-size matching. In a two sided market, maximum-size matching will be the size of the smaller side of the market.*

Throughout this chapter we use the "large market approximation" of the setting as introduced by [1]. This approximation assumes that the properties of the market, such as the limit on the number of applications, the limit on the number of interviews and the ratio between the size of the two sides of the market, is fixed while the size of the market grows to ∞, and studies the properties of the matching in the limiting case.

We claim that in this model, neither doctors or hospitals will be strategic: When doctors apply to hospitals for interview, they pick their favorite hospitals, with respect to their prior preference. Also hospitals select their favorite doctors to interview based on their prior. After interviews doctors and hospitals refine their preference order based on what they learned in the interviews stage. They are truthful with the orders they use for doctors/hospitals they interviewed. To see why, note that in the first stage due to the symmetry and independence in the market, no doctor or hospital is considered more popular than others. Therefore neither doctors or hospitals can benefit from applying or granting interviews to somebody other than their favorite selection. After the interviews, again neither hospitals or doctors can benefit by submitting their lists in a non-truthful order. It is not beneficial for the doctors (the proposing side) as shown in [28]. Also as shown in [41], the probability that it is beneficial for hospitals vanishes as the size of the market grows to ∞, which is the setting we study in this chapter. We summarize the statement in the following lemma.

Lemma 15. *Both doctors and hospitals are truthful during the matching procedure: In the first stage, doctors apply to their favorite hospitals; and hospitals grant interviews to their favorite applicants.*
In the second stage, both doctors and hospitals submit the list that they interviewd, ordered from the most preferred to the least preferred.

5.3 Effect of Number of Interviews

In this section we study the effect of number of interviews/offers with the same number of doctors and hospitals in the market. In 5.3.1 and 5.3.2 we find the efficiency of matching as a function of the expected number of applications by doctors, when hospitals only select one applicant in the first stage of the mechanism. We begin with the case where all doctors apply to the same number of hospitals and then we move to the case where doctors send different number of applications. We find allocating the number of applications equally results in the most efficient matching and allowing a small subset to apply to an extra position or restricting a small set to apply to one less position, results in a smaller matching. In Section 5.3.3 we study the matching outcome when hospitals grant multiple interviews/offers. We find that treating applicants equally is most efficient even in this more general model.

5.3.1 Same Number of Applications and Making One Offer

In this part we assume that hospitals only select one of their applicants in the first stage of the matching. As shown in Lemma 15, they choose their favorite applicant. Since interviews are conducted to compare the selected doctors, in this case that only one doctor is selected, interviews are of no use from the hospitals' point of view. Therefore the hospital immediately offers the position to the selected doctor. On the other hand, the outcome of the deferred acceptance algorithm can be easily found in this case; each doctor will be matched to his/her favorite hospital among those who gave an offer. Therefore without going through the complication of the deferred acceptance algorithm, we assume that each doctor accepts the most preferred offer. This discussion implies that the size of matching in this case, is the number of doctors who receive an offer.

We study the social welfare of the matching with respect to the number of applications by doctors. First, consider the simple case where doctors are allowed to send one application. Since in this case each doctor receives at most one offer, all the offers are accepted and the size of matching is equal to the number of offers which is the same as the number of hospitals who received an application.

Example 2. *With one application for each doctor and hospitals accepting one interview, the social welfare of the matching approaches* $(1 - 1/e) \approx 0.63$ *in a large market. In a matching between hospitals and doctors the size of matching equals the number of matched hospitals. Each hospital that receives an application is matched. Because when a hospital accepts a doctor, that doctor does not have any other offers and accepts the matchThe probability of a hospital receiving at least one application is* $1 - (1 - 1/n)^n$ *which tends to* $1 - 1/e$ *with* n *approaching* ∞*.*

Now to observe the effect of having more applications, consider the extreme case where doctors have no limit and are allowed to apply to all hospitals. If the size of matching is monotone increasing in the number of applications, this case creates the largest matching. Since hospital preferences are uniformly random, each hospital selects an applicant to make an offer to, uniformly at random. Therefore the number of doctors who receive an offer is:

$$lim_{n\to\infty} 1 - (1 - \frac{1}{n})^n = 1 - \frac{1}{e}$$

which shows that although the number of applications increased from one to n, there is no increase in the size of matching. Therefore either $1 - \frac{1}{e}$ is the largest matching possible or applying to $1 < k < n$ positions, achieves a larger matching which, if true means that matching size is not increasing in the number of applications.

Definition 7 (covered hospital)**.** *A hospital is covered if it receives at least one application.*

Suppose that all doctors are allowed to apply to k positions, therefore the number of total applications is nk. We show the number of covered hospitals is $\approx n(1 - e^{-k})$.

Proposition 9. *When all doctors apply to exactly k uniformly random hospitals, with large market approximation the expected fraction of covered hospitals is* $(1 - e^{-k})$.

proof sketch. Similar to example 2, the probability of a hospital receiving any application is $\approx 1 - (1 - 1/n)^{nk}$, which tends to $(1 - e^{-k})$ in the limit. □

Next we find the probability of a random application turning into an offer. In Lemma 15 we showed that doctors apply to their favorite positions. Since doctor preferences are independent, applications of different doctors are also independent. Also since with large n and small number of applications, the probability that a fresh truly random choice of a doctor is identical to a previous choice approaches to 0, with large market assumptions, we assume that different applications of a doctor are independent. Therefore we can assume that the destination hospitals of all applications are selected uniformly and independently at random. On the other hand doctors have no knowledge about the number of doctors applying to each hospital and all hospitals look symmetric to them. Thus from every doctors perspective, each hospital that they applied to, makes an offer to them with some probability p, and this probability is independent among the applications.

Proposition 10. *When each doctor applies to k random positions and each hospital offers its position randomly to one of its applicants, the expected social welfare of matching equals*

$$1 - (1 - \frac{(1 - e^{-k})}{k})^k.$$

Proof. With the above arguments we know from the doctors perspective, each hospital that they applied to, offers them a position independently with probability p. Therefore the expected number of total offers is the product of p and the total number of applications. As discussed previously the total number of offers equals the number of covered hospitals and is equal to $n(1 - e^{-k})$. Also since each doctor applies to k positions, the total number of applications is nk. Therefore $p = \frac{1-e^{-k}}{k}$ and the probability for a doctor to receive an offer is $(1 - (1 - \frac{(1-e^{-k})}{k})^k)$. The expected social welfare of matching is equal to the probability of a doctor receiving an offer, which implies the conclusion. □

Figure 5.1 shows the size of matching with respect to the number of applications.

The nonmonotonicity with respect to number of applications is a property of limited number of interviews. As proved in Proposition 20 and shown in Figure C.1, the size of matching is increasing in a setting with unlimited interviews.

As seen in Figure 5.1 the maximum size of matching occurs when all doctors apply to 3 hospitals. Therefore from a market design point of view, allowing three applications in this case is optimal in terms of social welfare.

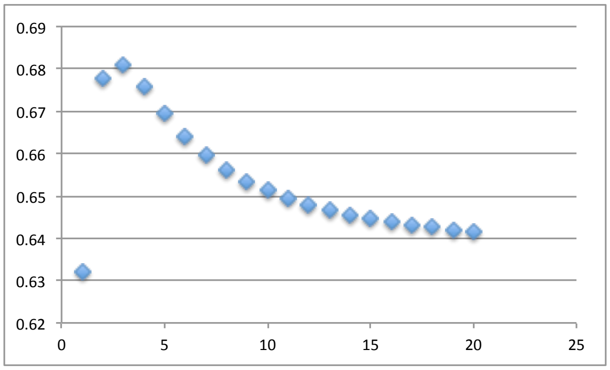

Figure 5.1: Size of matching with respect to the number of applications, when hospitals make only one offer.

Proposition 11. *In markets with the same number of applicants as positions where participants' preferences are drawn independently and uniformly at random, if all doctors are allowed to apply to k positions and hospitals grant a single offer, allowing doctors to apply to* 3 *hospitals results in the largest matching outcome when size of the market grows to ∞.*

As shown in Figure 5.1, having more applications can have both positive and negative effects: As a positive effect, with more applications the number of hospitals who receive an application increases. This can potentially increase the size of matching. On the other hand as the number of applications increases, the hospitals become more congested. Since he total number of offers is limited by the number of hospitals (each hospital makes at most one offer), this causes an increase in the probability of rejection of an application. This may increase the number of doctors with no offer, leading to a negative effect on the size of matching. As seen in the picture, The biggest jump in the size of matching occurs when moving from one application to two applications. When increasing the number of applications from one to two, the fraction of covered hospitals changes from $1 - 1/e \approx 0.63$ to $1 - 1/e^2 \approx 0.86$, which is the highest increase when adding more applications.

This causes the highest increase in the matching size. The increase stops with only three number of applications, with which the fraction of covered hospitals is $1 - 1/e^3 \approx 0.95$. From this point forward the negative effect takes over and because of the random allocation of hospitals some doctors receive multiple offers while others receive none.

5.3.2 Fractional Expected Number of Applications and Making One Offer: scallop-shape

Unlike Section 5.3.1, where there was a universal limit on the number of applications by doctors, in this section we study the social welfare when doctors are allowed to send different number of applications. First we show that for any expected number of applications x, the optimal social welfare occurs when doctors send either $\lceil x \rceil$ or $\lfloor x \rfloor$ applications. Then we study the social welfare as a function of the expected number of applications. Surprisingly, we observe that granting extra applications to a small set of doctors and also retracting an application from a small set both hurt the market; suggesting that unfair treatment is not efficient in terms of social welfare.

The following proposition shows that for any expected number of applications x, the optimal social welfare occurs when doctors send either $\lceil x \rceil$ or $\lfloor x \rfloor$ applications.

Proposition 12. *With any expected number of applications, x, the distribution of number of applications that achieves the highest social welfare, is one that allocates $\lfloor x \rfloor$ applications to some doctors and $\lceil x \rceil$ to the others, such that the expected number of applications equals x.*

Proof. Suppose the applications are distributed in a different way. Therefore there are two doctors with l and k applications such that $k - l \geq 2$. We show that the size of matching is improved if we allocate $f = \lfloor (k+l)/2 \rfloor$ and $c = \lceil (k+l)/2 \rceil$ applications to those doctors. This alteration does not change the probability of receiving offers by other doctors as the applications are independent from doctors perspective. So the only difference is the probability of receiving offers by these two doctors. Let p be the probability of a random application to lead to an offer. We claim

$$1 - (1-p)^k + 1 - (1-p)^l \leq 1 - (1-p)^c + 1 - (1-p)^f.$$

Since $(1-p)^c$ and $(1-p)^f$ have the same product as $(1-p)^k$ and $(1-p)^l$, their sum is larger when the two factors are far apart, therefore:

$$(1-p)^c + (1-p)^f \leq (1-p)^k + (1-p)^l$$

which implies the conclusion. □

This argument shows that in order to find the optimal social welfare for different expected number of applications, we only need to study the case where doctors apply to the same number of hospitals –as studied in 5.3.1– or two consecutive numbers. To find the size of matching, we first find the number of hospitals who receive an application.

Proposition 13. *The fraction of covered hospitals, when the expected number of applications is x, equals to $(1 - e^{-x})$.*

Proof. Similar to Proposition 9. □

Using the proposition above, we are ready to find the social welfare with different expected number of applications.

Theorem 12. *The social welfare of the matching with expected number of applications $k < x < k+1$, when applicants either apply to k or $k+1$ positions is:*

$$(\lceil x \rceil - x)(1 - (1 - \frac{(1-e^{-x})}{x})^k) + (x - \lfloor x \rfloor)(1 - (1 - \frac{(1-e^{-x})}{x})^{k+1})$$

this function is illustrated in Figure 5.2.

Proof. As shown in the proof of Proposition 10, from the doctors perspective, the probability of a random application, leading to an offer is $\frac{\text{number of covered hospitals}}{\text{number of applications}} = \frac{1-e^{-x}}{x}$. Therefore the expected social welfare of the matching which is the same as the probability of a random doctor receiving an offer is:

$$(\lceil x \rceil - x)(1 - (1 - \frac{(1-e^{-x})}{x})^k) + (x - \lfloor x \rfloor)(1 - (1 - \frac{(1-e^{-x})}{x})^{k+1}).$$

$\square$

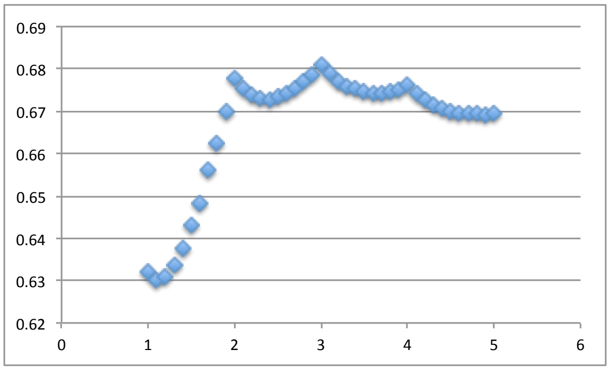

Figure 5.2: Size of matching with respect to expected number of applications, when hospitals make only one offer.

The scallop-shape Figure 5.2 illustrates the size of matching as a function of expected number of applications. The unusual behavior of the function at integer points shows that allowing a small group to apply to one more position, or limiting the number of applications of a small group to one less application, has a negative effect on size of matching.

This property is also a consequence of limited number of interviews. As proved in Proposition 20 and shown in Figure C.1, the size of matching is increasing even between two consecutive integers in a setting with unlimited interviews.

Also, if we compare this diagram with Figure C.1 as both use the same method for distributing applications; i.e. same number of doctors applying to k or $k + 1$ positions, we observe that this unusual shape belongs to the case with limited number of interviews, and if there is no limit on the number of interviews the size of the matching is monotone increasing.

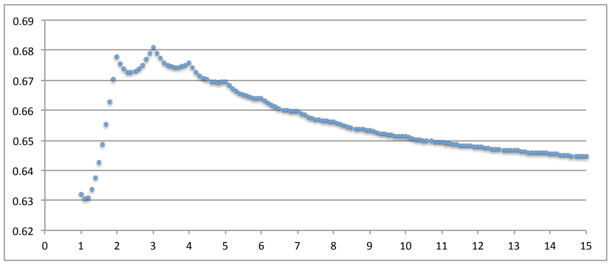

Figure 5.3: Size of matching with respect to expected number of applications, when hospitals make only one offer. (Extended version of Figure 5.2.)

As seen in Figure 5.3 the maximum size of matching occurs when all doctors apply to 3 hospitals, when doctors either apply to the same number of hospitals or two consecutive numbers. Also as shown in Proposition 12, for any expected number of application this is the optimal distribution of applications in terms of social welfare. This implies the following proposition which is a generalization of proposition 11: Applying to 3 hospitals is not only optimal when all doctors apply to the same number of positions, but it is also optimal for any arbitrary allocation of applications to doctors.

Corollary 7. *In markets with the same number of applicants as positions where participants' preferences are drawn independently and uniformly at random, and hospitals make one offer, allowing doctors to apply to 3 hospitals results in the* *largest matching outcome when size of the market grows to ∞.*

5.3.3 Granting multiple interviews

In this subsection we focus on the more realistic case where hospitals grant multiple interviews and we observe that the size of matching as a function of expected number of applications has a similar structure. Unlike the previous subsections where each hospital immediately made an offer to their top applicant, in this case they grant multiple interviews and have backup options in case they do not match their top applicant.

With multiple interviews, finding the final matching requires actually running the usual procedure of the deferred acceptance algorithm and finding the size of matching is more complicated. In the previous case where hospitals made immediate offers to their top applicant, each doctor was matched to their most preferred hospital that they received an offer from. However, when hospitals interview multiple doctors this is not clear anymore. A doctor can be rejected by a hospital in the first stage before the interview, if the hospital does not grant an interview to the doctor. Also a doctor can be rejected in the second stage as a step of the deferred acceptance algorithm. With these different potential outcomes of an application, the probability of an application leading to an offer is more complicated than the cases previously studied.

Theorem 13. *Suppose each doctor has k applications and each hospital grants interviews to at most k' doctors. The social welfare of the matching in this case is $1-(1-p)^k$ where $0<p<1$ is the solution to*

$$1-(1-p)^k = \sum_{i=1}^{k'-1}\left[(1-(1-\frac{1-(1-p)^k}{pk})^i)\frac{k^i e^{-k}}{i!}\right] + (1-(1-\frac{1-(1-p)^k}{pk})^{k'})(1-e^{-k}\sum_{i=0}^{k'-1}\frac{k^i}{i!}).$$

Also when the expected number of applications is $k \leq x < k+1$, the social welfare is upperbounded by $(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})$, where $0 < p < 1$ is the solution to

$$\begin{aligned}
&(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1}) \\
=&\sum_{i=1}^{k'-1}\left[(1-(1-\frac{(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})}{px})^i)\frac{x^i e^{-x}}{i!}\right] \\
&+(1-(1-\frac{(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})}{px})^{k'})(1-e^{-x}\sum_{i=0}^{k'-1}\frac{x^i}{i!}).
\end{aligned}$$

The upperbound is tight when all doctors apply to either k or $k+1$ positions.

Proof. We first invoke Lemma 16 to show that we can treat a doctor's proposals independently.

Lemma 16. *[1] Suppose p is the probability of a single proposal resulting in a permanent match in the deferred acceptance procedure. By large market assumptions, rejection of the first proposals of a doctor does not affect the probability of acceptance of the later proposals.*

Proof. A more complete statement of the lemma and its proof appear in Proposition 19 in the appendix. □

By Lemma 16 from a doctor's perspective each hospital they propose to is available to them with equal probability, and their availabilities are independent.Consider the modified implementation defined in Section 5.2. In the second stage of the game doctors send semi-proposals to hospitals. Let p be the probability of a random semi-proposal becoming a permanent match. By independence the probability that a doctor becomes matched is $1-(1-p)^k$. Also the expected number of semi-proposals made by a doctor is $1+(1-p)+\ldots+(1-p)^{k-1} = \frac{1-(1-p)^k}{p}$.

Lemma 17. *The probability of an application turning into a semi-proposal is independent of the application being valid.*

Proof. An application is turned into a semi-proposal if the previous semi-proposals are rejected. Same holds for a valid application: a valid application is turned into a proposal if the previous semi-proposals are rejected. Since preferences are distributed uniformly at random for both doctors and hospitals, the fact that an application is valid or invalid does not have any effect on its rank in the preference list. Therefore the probability of an application turning into a semi-proposal conditioned on the application being valid is the same as the unconditional probability. □

Now we find the total number of proposals. The probability of a random valid application turning into a proposal is the ratio of total number of proposals to valid applications. Similarly the probability of a random application turning into a semi-proposal is the ratio of total number of proposals to applications. By Lemma 17, these two ratios are equal. Therefore the total number of proposals can be computed based on total number of applications, valid applications and semi-proposals. We already know the total number of applications and semi-proposals, and only need to find the number of valid applications.

Corollary 8. *The expected number of proposals that a hospital receives is*

$$\frac{\textit{expected number of semi-proposals} \times \textit{expected number of valid applications}}{\textit{expected number of applications}}.$$

Because of the independent and uniform preferences of doctors, the number of applications that a hospital receives in the first stage comes from a Poisson distribution with mean k. The hospitals accepts all of them if the number of received applications, i, is less than k'. Otherwise the hospital accepts k' of them. Therefore the probability that a hospital has i valid applications is $\frac{k^i e^{-k}}{i!}$ for $i < k'$ and $1-e^{-k}\sum_{j=0}^{k'-1}\frac{k^j}{j!}$ for $i = k'$. Therefore the expected number of valid applications that a hospital receives is $\sum_{i=1}^{k'-1} i \cdot \frac{k^i e^{-k}}{i!} + k'(1-e^{-k}\sum_{j=0}^{k'-1}\frac{k^j}{j!})$. By corollary 8, the expected number of proposals a hospital receives is

$$\frac{1-(1-p)^k}{p} \cdot \frac{\sum_{i=1}^{k'-1} i \cdot \frac{k^i e^{-k}}{i!} + k'(1-e^{-k}\sum_{i=0}^{k'-1}\frac{k^i}{i!})}{k}.$$

From the hospitals' point of view, each of their valid applications has the same probability of becoming a proposal. Let q be the probability of a random valid application turning into a proposal. In a doctors-proposing deferred-acceptance algorithm, if a hospital is once tentatively matched it will remain matched forever. A hospital becomes tentatively matched if it receives a proposal, in other words if at least one of its valid applications becomes a proposal. The probability that a random hospital with j valid applications receives a proposal is $1-(1-q)^j$. Therefore the probability that a random hospital is matched equals

$$\sum_{i=1}^{k'-1}\left[(1-(1-q)^i)\frac{k^i e^{-k}}{i!}\right] + (1-(1-q)^{k'})(1-e^{-k}\sum_{i=0}^{k'-1}\frac{k^i}{i!}).$$

The probability of a random valid application turning into a proposal is equal to the ratio of the total number of proposals to the total number of valid applications. By Lemma 17, this ratio is equal to the ratio of the total number of semi-proposals to the total number of applications. Therefore, $q = \frac{1-(1-p)^k}{pk}$.

In a matching of two equal sides, the probability of a doctor being matched is equal to the probability of a hospital being matched. Therefore,

$$\begin{aligned} 1-(1-p)^k &= \sum_{i=1}^{k'-1}\left[(1-(1-q)^i)\frac{k^i e^{-k}}{i!}\right] + (1-(1-q)^{k'})(1-e^{-k}\sum_{i=0}^{k'-1}\frac{k^i}{i!}) \\ &= \sum_{i=1}^{k'-1}\left[(1-(1-\frac{1-(1-p)^k}{pk})^i)\frac{k^i e^{-k}}{i!}\right] + (1-(1-\frac{1-(1-p)^k}{pk})^{k'})(1-e^{-k}\sum_{i=0}^{k'-1}\frac{k^i}{i!}) \end{aligned}$$

where the second equality is derived by substituting q. By solving the equation for $0 \leq p \leq 1$ the size of the matching, $1-(1-p)^k$ is found.

Similar reasoning holds when the expected number of applications is an arbitrary real number x such that $k \leq x < k+1$ and each doctor has either k or $k+1$ applications. The size of the matching in this case is $(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})$, where p is computed using the equation below.

$$\begin{aligned}
&(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1}) \\
&=\sum_{i=1}^{k'-1}\left[(1-(1-q)^i)\frac{x^ie^{-x}}{i!}\right]+(1-(1-q)^{k'})(1-e^{-x}\sum_{i=0}^{k'-1}\frac{x^i}{i!}) \\
&=\sum_{i=1}^{k'-1}\left[(1-(1-\frac{(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})}{px})^i)\frac{x^ie^{-x}}{i!}\right] \\
&\quad+(1-(1-\frac{(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})}{px})^{k'})(1-e^{-x}\sum_{i=0}^{k'-1}\frac{x^i}{i!})
\end{aligned}$$

The last equality is derived by using $q=\frac{(1-(x-\lfloor x \rfloor))(1-(1-p)^k)+(x-\lfloor x \rfloor)(1-(1-p)^{k+1})}{px}$.

By Proposition 12, we know that the size of the matching with expected $k \leq x < k+1$ applications is maximized when all doctors apply to either k or $k+1$ positions. This concludes the proof. □

5.4 Unbalanced Markets

In this section we show that the phenomenon of the positive effect of setting the same limit for all doctors is not limited to a balanced market – where the number of applicants and positions are the same – but extends to unbalanced markets with different number of applicants and positions. Although the same phenomenon exits

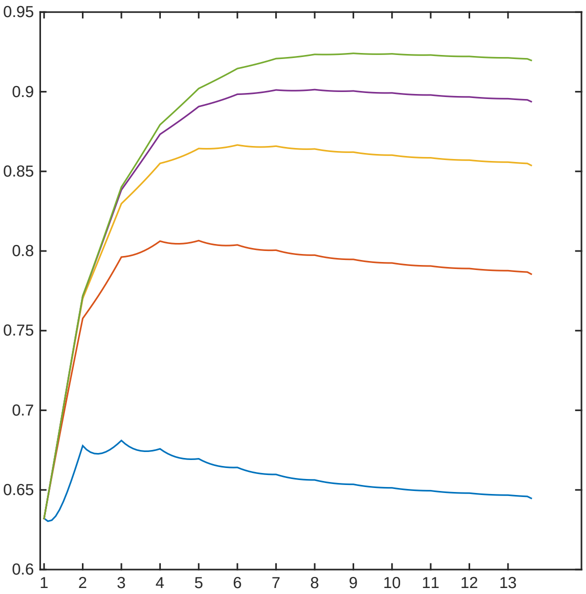

Figure 5.4: Size of matching with respect to expected number of application. The curves represents the settings where hospitals allow one to five interviews. The lowest curve belongs to one interview and the highest curve to five interviews.

more generally, the exact function is not preserved and the optimum number of applications depends on the ratio between the sizes of the two sides.

We study the case where hospitals make one offer and the ratio of the number of hospitals to the number of doctors is r. The results from Section 5.3.3 with multiple interviews also generalize to this setting. However, for simplicity and similarity of the outcomes, we just present the results for the model with one offer. Similar to the previous section, we can compute the size of the matching as a function of the expected number of applications, when doctors send k or $k+1$ applications.

Proposition 14. *The expected fraction of matched doctors with expected number of applications $k < x < k+1$, when applicants either apply to k or $k+1$ positions* *and when the ratio of number of hospitals to number of doctors is r is:*

$$(\lceil x\rceil - x)(1-(1-\frac{r(1-e^{-x/r})}{x})^k) + (x-\lfloor x\rfloor)(1-(1-\frac{r(1-e^{-x/r})}{x})^{k+1}).$$

Proof. As shown in the proof of Proposition 10, from the doctors perspective, the probability of a random application, leading to an offer is $\frac{\text{number of covered hospitals}}{\text{number of applications}}$. From the hospitals perspective, each application is equally likely to be sent to each hospital. Therefore the number of applications received is distributed as a Poisson distribution with $\lambda = \frac{x}{r}$. So the number of expected covered hospitals is $rn(1-e^{-x/r})$ and the probability of a random application, leading to an offer is $\frac{r(1-e^{-x/r})}{x}$. Therefore the expected fraction of doctors who are matched which is the same as the probability of a random doctor receiving an offer is:

$$(\lceil x\rceil - x)(1-(1-\frac{r(1-e^{-x/r})}{x})^k) + (x-\lfloor x\rfloor)(1-(1-\frac{r(1-e^{-x/r})}{x})^{k+1}).$$

□

Based on Definition 6, the social welfare of a matching is the ratio of the size of the matching to the size of maximum matching and in unbalanced markets, the maximum size is the size of the smaller side of the matching. For $r \geq 1$, the doctors make the smaller side therefore the social welfare is equal to expected fraction of doctors who are matched as computed in Proposition 14. If $r < 1$, the social welfare is the expected fraction of doctors who are matched as computed in Proposition 14 divided by r.

Figure 5.4 shows the social welfare of the matching as a function of expected number of applications for $r = \frac{1}{2}, 1, 2$. In the figure, the red function refers to $r = 2$, the blue function to $r = 1/2$ and the green function to $r = 1$.

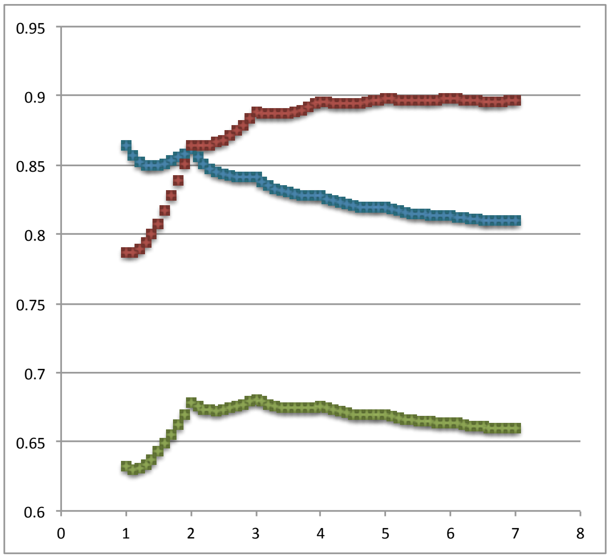

Figure 5.5: Social welfare of the matching with respect to the expected number of applications. The red curve refers to $r = 2$, blue curve to $r = 1/2$ and the green curve to $r = 1$, where r is the ratio of the number of hospitals to doctors.

As seen in the figure a similar structure holds for unbalanced networks, but the optimal number of applications depends on the factor of balancedness r. When the number of hospitals is half of the number of doctors, the market achieves its maximum size when each doctor applies to just one position. With more applications, hospitals become more congested, and the number of covered hospitals and therefore the number of total offers does not increase significantly. Therefore the rejection probability of applications increases and with higher probability a doctor remains unmatched. In contrast, when the number of hospitals is more than the number of doctors, allowing more applications has a positive effect. With more applications –while the number of applications is still small– more hospitals receive at least one application; therefore the number of offers increases considerably.

Also note that in the figure, the social welfare of the balanced market is generally lower than the social welfare of markets with $r = 1/2, 2$. This is not surprising since by definition the social welfare is the fraction of matched individuals on the smaller side of the market: doctors for $r \leq 1$, and hospitals for $r > 1$. For $r = 2$, with the same number of applications, more hospitals are covered that leads to more offers and a higher fraction of matched doctors. For $r = 1/2$, with the same number of applications a higher fraction of hospitals is covered which leads to a higher fraction of matched hospitals.

5.5 Beyond Uniform Preferences

In this section we go beyond the uniform and independent assumptions and use simulations to show that the properties of limited interviews exist more generally. The model studied in the previous sections assume that doctors and hospitals preferences remain uniform and independent after the interview stage. We relax this assumption. We study a model where the two sides have no information prior to the matching procedure. However, after the interview stage, all participants refine their lists with information learned in the interviews. Therefore preference lists may become correlated.

Without prior knowledge about the other side, the behaviors in the first stage is similar to what stated previously: doctors and hospitals pick the top of their lists. Therefore no strategic action is introduced in the first stage. In the second stage, doctors and hospitals are not anymore symmetric in terms of popularity. However as shown in [41] in a large market almost all participants prefer to be truthful.

To show how generally the properties hold, for simulations we study the other extreme in preferences: completely aligned preferences after interviews. We use a market of 10,000 doctors and hospitals. Similar to previous sections we start with uniform and independent preference in the first stage. For the second stage we simulate running the deferred acceptance algorithm when hospitals agree on a random preference over doctors. Figure 5.6 shows the size of matching as a function of number of applications when the the number of interviews for hospitals is limited to two interviews.

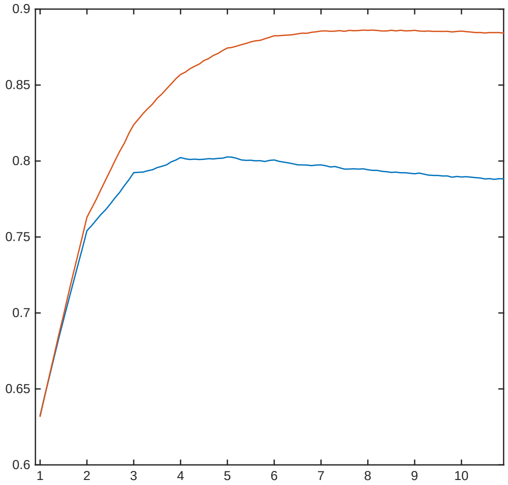

Figure 5.6: Size of matching with respect to expected number of application. In the first stage the preferences are uniform and independent. In the second stage hospitals preferences are aligned. The bottom curve belongs to two interviews and the top one to four interviews.

As observed in Figure 5.6 the properties of limited interviews are not restricted to uniform preferences. They hold even in the other extreme case where the preferences are completely correlated after interviews.

5.6 Improvement of social welfare with restricting interviews

In this section we study the effect of limiting the number of interviews in the more general model introduced in Chapter 4 and [8]. We first remind the readers of the model. Beyhaghi et al. [8] study the setting where the preference model is a combination of independence and correlation. They assume that both sides of the market are divided into tiers, where everybody prefers higher tiers to lower tiers (correlation) and inside a tier the preferences are uniformly random (independence). It is not hard to see that the model studied in the previous sections is a special case of the tiered model where the market only consists of one tier. In the model studied in the previous sections, limiting the number of interviews by hospitals had a negative effect on the social welfare as observed in Figure 5.4 which compares one and two interviews. This holds more generally since with a more restricted limit on the number of interviews by hospitals, more applications are rejected in the first stage of the process. Therefore doctors and hospitals submit shortened lists to be used by the deferred acceptance procedure, resulting in a matching with smaller size. Our main finding in this section is that a limit on interviews can improve the social welfare in the multi-tiered model.

In a multi-tiered market, matching to a higher tier doctor/hospital has a higher value for a participant compared to a lower tier, and the social welfare is the sum of the values of all participants (doctors and hospitals). For example in a two-tier market, matching to a high-tier doctor/hospital has value $v > 1$ for participants, while matching to a low-tier doctor/hospital has value 1.

Definition 8 (social welfare in a two-tier market [8]). *Social welfare is the sum* *of the value of matched doctors and hospitals, such that the value of matching to a high-tier doctor/hospital is v while the value of matching to a low-tier doctors/hospitals is* 1*.*

In a tiered market, the participants are strategic and their strategic action can hurt the social welfare [8]. Consider a two-tier market. Unlike the model in the previous section, since the market is not symmetric and high-tier hospitals are more popular, doctors are strategic. They need to decide how to select their limited number of hospitals both to increase the likelihood of being matched to a high-tier hospital, and to have backup options from the low tier in case they are not matched to the high tier. In the equilibrium of a two-tier market, doctors have a tendency to apply more to high-tier hospitals compared to a non-strategic setting when their goal is to maximize the social welfare [8]. This strategic behavior sometimes has a negative impact on social welfare of the matching. Applying to more high-tier hospitals makes high-tier hospitals congested and also reduces the number of matches of the high-tier doctors (more valuable doctors in terms of social welfare) since they do not use the benefit from potential matches to the low tier.

Although restricting the number of interviews hurts the social welfare in many cases, there are settings like Example 3 where it can improve the social welfare.

Example 3. *Consider a market with the same number of top-tier doctors, low-tier doctors, top-tier hospitals and low-tier hospitals. Suppose each doctor sends two applications and v, the value for the high tier, is ≈ 1.91. The social welfare of the matching when each hospital grants at most six interviews is higher compared to case where there is no limit on the number of interviews.*

The result in Example 3 is found via simulation. We observed $v \approx 1.91$ is the smallest value that when there is no limit on the number of interviews, in equilibrium all high-tier doctors send both of their applications to the high tier. When hospitals grant a limited number of interviews, the probability of an application leading an offer decreases. So high tier doctors gain less utility when sending both of their applications to the high tier and have more incentives to apply to the low tier, therefore they randomize between sending both applications to the high tier or one to the high and one to the low tier. By simulation we observed that when the number of interviews by each hospital is limited to six, this new flow of applications from high tier doctors to low tier hospitals increases the size of the matching and also improves the overall social welfare. With less number of interviews, many applications are rejected in the first stage and the social welfare decreases. Also with more number of interviews the flow of applications to the low tier is not significant to improve the social welfare. Therefore we observe that in this case a limit on the number of interviews by hospitals (six) results in a more efficient matching than unlimited number of interviews.

APPENDIX A

MISSING PROOFS OF CHAPTER 2

A.1 Lower Bound Examples

In this section, we analyze a new lowerbound in detail, and remind the reader of a lower bound from [26].

A.1.1 Theorem 5 is tight [26]

Here, we sketch the construction from [26]. There be n bidders and $m > c \cdot n$ items for some absolute constant c. Recall that $\mathcal{ER}^m$ denotes the distribution over values for m items for which each value is drawn i.i.d. from the equal revenue curve with revenue equal to 1. Then consider the posted-price mechanism which visits each buyer sequentially (in arbitrary order) and offers her the option to buy any set of $\frac{m}{4n}$ remaining items for price $\pi = \frac{m}{8} \cdot (\ln(m/n) + 1)$. Feldman et al. prove the following:

Proposition 15 ([26]). *The posted-price mechanism described above achieves expected revenue $\Omega(m \cdot n(1 + \ln(m/n)))$ for n buyers whose values for the m items are drawn i.i.d. from $\mathcal{ER}^m$.*

On the other hand, the revenue achieved by selling a single item to n buyers whose values are drawn i.i.d. from $\mathcal{ER}$ is well-understood. For the sake of completeness we repeat a proof below.

Proposition 16 (Folklore). $\text{Rev}_n(\mathcal{ER}) = n$.

Proof. It is clear that $\textsc{Rev}_1(\mathcal{ER}) = 1$: for any price p, the revenue achieved by setting price p is $p \cdot 1/p = 1$. This immediately implies that $\textsc{Rev}_n(\mathcal{ER}) \leq n$: even if the auctioneer had n copies of the item for sale, they would still not get revenue more than n.

Moreover, here is an auction that guarantees revenue approaching n as $p \to \infty$: post price p on the item, and sell to the lexicographically first bidder whose value exceeds p. Then the probability of sale is $1 - (1 - 1/p)^n \geq n/p - \binom{n}{2}/p^2$. So the revenue is at least $n - \binom{n}{2}/p$, which approaches n as $p \to \infty$. The second-price auction is optimal for n bidders drawn from $\mathcal{ER}$, and achieves revenue n (one could separately verify this by directly computing the expected second-highest value of n i.i.d. draws from $\mathcal{ER}$, if desired). $\square$

Together, these propositions claim that $\textsc{Rev}_n(\mathcal{ER}^m) = \Omega(m \cdot n(\ln(m/n) + 1))$, yet also $\text{VCG}_{n+c}(\mathcal{ER}^m) = m \cdot (n + c)$. Therefore, in order to possibly have $\text{VCG}_{n+c}(\mathcal{ER}^m) \geq \textsc{Rev}_n(\mathcal{ER}^m)$ we need to have $c = \Omega(n \ln(m/n))$. We therefore conclude:

Corollary 9 ([26]). *The competition complexity of n bidders with additive valuations over m i.i.d., regular items is at least $\Omega(n \ln(m/n))$.*

A.1.2 Theorem 6 is tight for the EFFTW benchmark

In this section, we prove that any analysis starting from the EFFTW benchmark can prove at best a competition complexity of $\sqrt{nm}$. Again, consider the distribution $\mathcal{ER}^m$. For the subsequent analysis, it will be helpful to instead think of replacing $\mathcal{ER}$ by a distribution with CDF $F(x) = 1 - 1/x$ for $x \in [1, p)$, and

$F(p) = 1$ (that is, an equal revenue curve truncated at p) for $p \to \infty$. We will not explicitly replace $\mathcal{ER}$ with this distribution, and we will always think of $p \to \infty$ in the subsequent analysis.

Our first lemma states that when considering $\mathcal{ER}^m$, max of sums in the EFFTW benchmark can be replaced by a sum of maxes. Observe that Lemma 18 does *not* generally hold for distributions other than $\mathcal{ER}^m$, and moreover that generally swapping the max and sum results in a benchmark that is unachievable with any finite competition complexity (just consider distributions that are point-masses at 1). But for equal revenue curves, this equation holds.

Lemma 18.

$$\begin{aligned}
&\sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n} \left[\max_{i \in [n]} \left\{ \varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) + v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j) \right\} \right] \\
=&\sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n} \left[\max_{i \in [n]} \left\{ \varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) \right\} + \max_{i \in [n]} \left\{ v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j) \right\} \right] \\
=&nm + \sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n} \left[\max_{i \in [n]} \left\{ v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j) \right\} \right]
\end{aligned}$$

Proof. The intuition is roughly as follows: for a single value drawn from the equal revenue curve truncated at p, we have $\varphi(v) = 0$ when $v < p$ and $\varphi(p) = p$. So the distribution of virtual values is 0 with probability $1 - 1/p$, and p with probability $1/p$. One can therefore informally think of the untruncated equal revenue curve as having virtual value distribution that is 0 with probability 1, and $+\infty$ with probability 0 (yielding an expected value of 1). Therefore, the argmaximum in the benchmark is some i with $\vec{v}_i \notin R_j$ with probability 1, and with probability 0 the argmaximum is some i with $\vec{v}_i \in R_j$ and $\varphi_j(v_{ij}) = +\infty$. As the φ term only interferes with probability 0, we can simply sum the terms instead. Of course, this is quite informal, but provides good intuition for where the proof is going (and why we need to consider $\mathcal{ER}$ truncated at p for $p \to \infty$ to be formal).

To begin the proof, observe first that whenever $\varphi_j(v_{ij})^+ > 0$, we have $\varphi_j(v_{ij})^+ = p \geq v_{i'j}$ for all i'. That is, whenever a virtual value is non-zero, it is at least as large as any value (because the virtual value is only non-zero when it is equal to p, the maximum possible value). Now, let $f^*(\cdot)$ denote the density of $\max_{i\in[n]} \{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}$, and let $P^*(x)$ denote the probability that $\max_{i\in[n]}\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\} = 0$, conditioned on $\max_{i\in[n]} \{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = x$. Then we can write:

$$\begin{aligned}&\sum_{j=1}^{m} \mathbb{E}_{\vec{v}\leftarrow(\mathcal{ER}^m)^n} \left[\max_{i\in[n]} \left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) + v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\right\}\right] \\ =&m \cdot \Bigg(p \cdot \Pr[\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = p] \cdot P^*(p) \\ &\quad + \int_0^p x \cdot f^*(x) \cdot P^*(x)dx \\ &\quad + p \cdot \Pr\left[\max_{i\in[n]} \left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\right\} > 0\right]\Bigg).\end{aligned}$$

The left two terms sum the expected contribution from bidders $\notin R_j$, and the last term covers the contribution from bidder $\in R_j$. Let's begin with the left-most term. We claim here that the left-most term approaches 0 as $p \to \infty$. To see this, observe that in order to possibly have $\max_{i\in[n]} \{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = p$, there must exist two pairs (i, j) for which $v_{ij} \geq p$. This is because in order to have $\vec{v}_i \notin R_j$, *and* $v_{ij} = p$, we must have $v_{ij'} \geq p$ for some $j' \neq j$. Observe that the probability that this happens is at most $\binom{nm}{2}/p^2$. This yields:

$$\lim_{p\to\infty} p \cdot \Pr[\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = p] \cdot P^*(p) \leq p \cdot \binom{nm}{2}/p^2 = 0.$$

Let's now analyze the integral. Here, we will simply claim that as

$p \to \infty$, $P^*(x) \to 1$ for all x. To see this, let's first understand how conditioning on $\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = x$ affects the distribution of $\max_{i\in[n]}\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\}$. Observe that conditioning on $\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = x$ may bias the distribution of the number of indices for which $\vec{v}_i \in R_j$. For example, if $\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = 0$, then we know that $\vec{v}_i \in R_j$ for all i. Similarly, if $\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}$ is large, then it is more likely that $\vec{v}_i \notin R_j$ for many i (because then more terms in the max are non-zero). But this conditioning does not bias the distribution of $\varphi_j(v_{ij})$ for those indices (conditioned on $\vec{v}_i \in R_j$). That is, once we condition in a set S of bidders with $\vec{v}_i \in R_j$ for all $i \in S$, the distribution of $\max_{i\in S}\{\varphi_j(v_{ij})\}$ is independent of $\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j\}$. So certainly $1 - P^*(x)$ is at most the probability that n independent draws from $\varphi_j(v_{ij})$, conditioned on $\vec{v}_i \in R_j$, are all less than p, as for all x there are at most n indicies for which $\vec{v}_i \in R_j$ (simply because there are n bidders). Observe further that the distribution of $\varphi_j(v_{ij})$, conditioned on $\vec{v}_i \in R_j$, is simply the maximum of m i.i.d. draws from $\mathcal{ER}$ (truncated at p). So the probability that a single one of these draws exceeds p is at most m/p by the union bound. Again taking a union bound over the n draws, the probability that any exceed p is at most mn/p. As $p \to \infty$, this approaches 0. Therefore, as $p \to \infty$, $P^*(x) \to 1$. Observe that we've now shown that:

$$\lim_{p\to\infty} m \cdot \left(p \cdot \Pr[\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\} = p] \cdot P^*(p) + \int_0^p x \cdot f^*(x) \cdot P^*(x)dx\right)$$
$$= 0 + m\int_0^\infty x \cdot f^*(x)dx$$
$$= \sum_{j=1}^m \mathbb{E}_{\vec{v}\leftarrow(\mathcal{ER}^m)^n}\left[\max_{i\in[n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}\right].$$

We now turn to the final term. First, consider all nm i.i.d. draws from $\mathcal{ER}$. There are two ways in which we can have $\max_{i\in[n]}\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\} > 0$. First, maybe the maximum of these nm draws is some bidder i's value for item j, and this $v_{ij} = p$. Or, maybe the maximum of these nm draws is some bidder i's value for some other item $j' \neq j$, but there is another bidder whose value for item j exceeds p (implying that there are at least two of the nm draws that exceed p).

For the first case, the probability that the maximum is some bidder i's value for item j is exactly $1/m$. Independently, the probability that this value exceeds p is $1-(1-1/p)^{nm} \in [nm/p - \binom{nm}{2}/p^2, nm/p]$. For the second case, the probability that at least two of the nm draws exceed p is at most $\binom{nm}{2}/p^2$. Therefore, we get that:

$$\Pr\left[\max_{i\in[n]}\left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\right\} > 0\right] \in [n/p - \binom{nm}{2}/(mp^2), n/p + \binom{nm}{2}/p^2].$$

And therefore

$$p \cdot \Pr\left[\max_{i\in[n]}\left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\right\} > 0\right] \in [n - \binom{nm}{2}/(mp), n + \binom{nm}{2}/p].$$

Implying that:

$$\lim_{p\to\infty} p \cdot \Pr\left[\max_{i\in[n]}\left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\right\} > 0\right] = n.$$

This completes the analysis of the right-most term, as now:

$$\begin{aligned}&\lim_{p\to\infty} m \cdot p \cdot \Pr\left[\max_{i\in[n]}\left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\right\} > 0\right]\\ &= \mathbb{E}_{\vec{v}\leftarrow(\mathcal{ER}^m)^n}\left[\max_{i\in[n]}\left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j)\right\}\right] = nm.\end{aligned}$$

Putting both parts together proves the lemma.

□

Now, we just need to analyze $\sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n} \left[\max_{i \in [n]} \{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}\right]$.

Proposition 17. *Let $n \geq 4m$. Then:*

$$\sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n} \left[\max_{i \in [n]} \{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}\right] =\geq \frac{m\sqrt{mn}}{14}.$$

Proof. We begin by considering the probability that the ℓ^{th} highest of nm draws from $\mathcal{ER}$ (or any distribution) is some bidder i's value for item j, *and* $\vec{v}_i \notin R_j$ *and* for all $\ell' < \ell$, either it is some bidder i''s value for an item $j' \neq j$, or $v_{i'} \in R_j$. That is, we are interested in computing the probability that the ℓ^{th} highest of the nm i.i.d. draws is the maximum value (times $\mathbb{I}(\vec{v}_i \notin R_j)$) for item j. Observe that this probability is well-defined: it does not depend on the particular values for the nm draws (and does not even depend on the distribution from which they are drawn). Indeed, whether the desired event occurs or not is only a function of how the draws are permuted among the nm values of bidders for items.

So denote by P_ℓ the probability that the ℓ^{th} highest of nm draws is some bidder i's value for item j *and* $\vec{v}_i \notin R_j$ *and* for all $\ell' < \ell$, either it is some bidder i''s value for an item $j' \neq j$, or $\vec{v}_{i'} \in R_j$. Let E_1 denote the event that the ℓ^{th} highest is some bidder's value for item j. Let E_2 denote the probability that $\vec{v}_{i(\ell)} \notin R_{j(\ell)}$, where the ℓ^{th} highest is assigned to bidder $i(\ell)$ and item $j(\ell)$. Let E_3 denote the event that for all $\ell' < \ell$, either $j(\ell') \neq j(\ell)$ or $\vec{v}_{i(\ell')} \in R_{j(\ell)}$. Then it is easy to see that $\Pr[E_1] = 1/m$. Moreover, events E_2 and E_3 only involve where the highest $\ell - 1$ draws go. In particular, observe that both E_2 and E_3 are independent of E_1. Moreover, observe that E_3 is only more likely to occur conditioned on E_1 and E_2: if one of the top $\ell - 1$ draws is for the same bidder as the ℓ^{th}, it is necessarily *not* for the same item j, and also necessarily *not* helping to put some other $\vec{v}_{i'} \notin R_j$ (by being larger than $v_{i'j}$) (hence making it more likely that all of the top $\ell - 1$

draws are permuted to some item other than j, and also that those values which are still permuted to item j are $\in R_j$). Therefore:

$$\begin{aligned} P_\ell &= \Pr[E_1] \cdot \Pr[E_2|E_1] \cdot \Pr[E_3|E_2 \wedge E_1] \\ &= \Pr[E_1] \cdot \Pr[E_2] \cdot \Pr[E_3|E_2 \wedge E_1] \\ &\geq \Pr[E_1] \cdot \Pr[E_2] \cdot \Pr[E_3] \\ &= \Pr[E_3 \wedge E_1] \cdot \Pr[E_2]. \end{aligned}$$

Let's first analyze the probability of E_2. Observe that event E_2 does *not* happen only if each of the top $\ell - 1$ draws are permuted to a different bidder than ℓ. So:

$$\begin{aligned} \Pr[E_2] = 1 - &\left(1 - \frac{m-1}{mn-1}\right) \cdot \left(1 - \frac{m-1}{mn-2}\right) \cdot \dots \cdot \left(1 - \frac{m-1}{mn-\ell+1}\right) \\ &\in [1 - (1 - \frac{m-1}{mn-1})^{\ell-1}, (\ell-1)/n]. \end{aligned}$$

To see the first equality, observe that $\frac{mn-1}{mn-\ell'}$ is the probability that the ℓ'^{th}-highest draw is permuted to a different bidder than the ℓ^{th}, conditioned on the first thru $(\ell'-1)^{th}$ draws all being permuted to a different bidder than the ℓ^{th}. This is because there are $m - 1$ items left for the same bidder, and $mn - \ell'$ remaining (bidder, item) pairs in total. To see the upper bound, observe that the probability that a single $\ell' < \ell$ is permuted to the same bidder as ℓ is exactly $(m-1)/(mn-1)$. So taking a union bound over all $\ell' < \ell$, we get an upper bound on the probability that *any* ℓ' is permuted to the same bidder as ℓ is at most $(\ell-1)(m-1)/(mn-1) \leq (\ell-1)m/(mn) = (\ell-1)/n$. The lower bound follows by just observing that there are $\ell - 1$ terms in the product, and each term is at most $(1 - \frac{m-1}{mn-1})$.

Finally, we prove one minor technical lemma to argue that when $\ell - 1 \leq n/2$, $1 - (1 - \frac{m-1}{mn-1})^{\ell-1} \geq (1 - \ln(2)) \cdot (\ell - 1)/(2n)$.

Lemma 19. *Let* $\ell - 1 \leq n/2$ *and* $n \geq 2$*. Then* $1 - (1 - \frac{m-1}{mn-1})^{\ell-1} \geq (1 - \ln(2)) \cdot (\ell - 1)/(2n)$.

Proof. We start by searching for a constant c such that for all $x \in [0, 1/2]$, and all y such that $xy \in [0, 1/2]$, we have $1 - (1 - x)^y \geq cxy$. We will then use this constant to prove the lemma statement. Let's first fix x and c, and minimize $1 - (1 - x)^y - cxy$ over all $y \in [1, 1/(2x)]$ (this search is well-defined unless $x = 0$, in which case $1 - (1 - x)^y - cxy = 0$ for all y). The derivative with respect to y is $-\ln(1-x)(1-x)^y - cx$, and the second derivative is $\ln(1-x)^2(1-x)^y \geq 0$. So if the first derivative is positive at $y = 1$, it is positive on the entire interval $[1, 1/(2x)]$. So now we wish to see how small the first derivative can be, as a function of x, when $y = 1$.

This is again single-variate optimization: minimize $-\ln(1-x) \cdot (1-x) - cx$ on $[0, 1/2]$. The derivative is $1 + \ln(1 - x) - c$. Observe that if $c \leq 1 - \ln(2)$, then the derivative is ≥ 0 on the entire range $(0, 1/2)$. This means that for $c \leq 1 - \ln(2)$, the minimizer occurs at $x = 0$, which is $-\ln(1) \cdot 1 - c \cdot 0 = 0$. So at this point, we conclude that when $c \leq 1 - \ln(2)$, the first derivative with respect to y is non-negative at $y = 1$, and therefore positive on the entire interval $[1, 1/(2x))$. This means that the minimum occurs at $y = 1$ (and we will restrict ourselves from now on to $c \leq 1 - \ln(2)$).

Now that we know the minimum occurs at $y = 1$, our remaining minimization is trivial: minimize $1 - (1 - x) - cx = (1 - c)x$, which is achieved at $x = 0$, and indeed $c \cdot 0 \geq 0$. So we have proven that when $c \leq 1 - \ln(2)$, the minimum value for $1-(1-x)^y - cxy$ over all $x \in [0,1/2]$ and $y \in [1, 1/(2x))$ occurs at $x = 0$, $y = 1$, and the value is ≥ 0. This proves that $1-(1-x)^y \geq (1-\ln(2))xy$ whenever $x \in [0,1/2]$ and $xy \in [0,1/2]$.

Now we wish to apply the above fact when $x = \frac{m-1}{mn-1}$ and $y = \ell - 1 \leq n/2$. Indeed, observe first that when $n \geq 2$, we have that $x \leq 1/2$. Moreover, when $y \leq n/2$ we have $x \cdot y = \frac{nm-n}{2mn-2} \leq 1/2$. Applying the previous work gives $1 - (1-\frac{m-1}{mn-1})^{\ell-1} \geq (1-\ln(2)) \cdot (\ell-1) \cdot \frac{m-1}{mn-1}$. Note that for $m \geq 2$, $\frac{m-1}{mn-1} \geq 1/(2n)$. Therefore we can conclude the lemma statement. □

Now with Lemma 19, we can conclude that when $\ell - 1 \leq n/2$, we have:

$$\Pr[E_2] \in [(1-\ln(2))(\ell-1)/(2n), (\ell-1)/n].$$

Now, let's analyze the probability of $E_3 \wedge E_1$. Observe that first, $\Pr[E_1] = 1/m$, and E_3 and E_1 are independent. So let's look at $\Pr[E_3|E_1]$. Observe that $P_{\ell'}$ is *exactly* the probability that E_3 does *not* occur, conditioned on E_1, because the ℓ'^{th} highest value has $j(\ell') = j(\ell)$ and $\vec{v}_{i(\ell')} \notin R_{j(\ell')}$. If none of these events occur, then certainly event E_3 occurs. Therefore:

$$\Pr[E_3|E_1] \geq 1 - \sum_{\ell'<\ell} P_{\ell'}, \Pr[E_3 \wedge E_1] \geq (1 - \sum_{\ell'<\ell} P_{\ell'})/m.$$

And therefore we can conclude that:

$$P_\ell \geq \Pr[E_1 \wedge E_3] \cdot \Pr[E_2] \geq 1/m \cdot \left(1 - \sum_{\ell'<\ell} P_{\ell'}\right) \cdot (1-\ln(2))\,(\ell-1)/(2n). \quad \text{(A.1)}$$

Observe also that $P_\ell \leq \Pr[E_1 \wedge E_2] = \Pr[E_1] \cdot \Pr[E_2] \leq \frac{\ell-1}{mn}$. In particular, we can use this *upper bound* on $P_{\ell'}$, $\ell' < \ell$ to derive a cleaner *lower bound* on P_ℓ.

Indeed, for any $\ell < \sqrt{mn} \leq n/2$ (the last inequality is because we assume that $n \geq 4m$. This is the only place we use this assumption, but it is a key step), we have that:

$$\sum_{\ell' < \ell} P_{\ell'} \leq \sum_{\ell' < \sqrt{mn}} (\ell' - 1)/mn \leq \frac{\sqrt{mn}^2}{2mn} = 1/2.$$

By plugging the above in inequality A.1 we can conclude that for all $\ell < \sqrt{mn}$, we have:

$$P_\ell \geq \frac{(1 - \ln(2)) \cdot (\ell - 1)}{4nm}.$$

To recap: we have now shown that, independent of the particular values drawn, the ℓ^{th} highest of nm draws is equal to $\max_{i \in [n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \in R_j)\}$ with probability at least $P_\ell \geq \frac{(1-\ln(2))\cdot(\ell-1)}{4mn}$, for all $\ell < \sqrt{mn}$. So the only remaining step is to compute the expected value of the ℓ^{th} highest of nm draws from $\mathcal{ER}$.

Lemma 20. *Let $V_{x,y}$ denote the x^{th} highest of y i.i.d. draws from $\mathcal{ER}$. Then $\mathbb{E}[V_{x,y}] = y/(x-1)$.*[1]

Proof. One approach would be to explicitly write out the integral for the x^{th} highest of y draws. This is tedious. Instead, we will count the revenue of an auction in two different ways. Consider an $(x-1)$-unit auction with $x-1$ copies of the same item. There are y unit-demand bidders with values drawn i.i.d. from $\mathcal{ER}$. Then as all virtual values are non-negative and there is no ironing, the revenue-optimal auction is to set the x^{th} highest bid as the price and let the $(x-1)$ highest bidders get the item. The expected revenue of this auction is exactly $V_{x,y} \cdot (x-1)$.

---

[1] Observe that when $x = 1$, the expected value of the highest of ≥ 1 draws from $\mathcal{ER}$ is indeed $+\infty$.

On the other hand, we claim that the optimal revenue for this setting is exactly y. To see this, observe that clearly the optimal revenue is at most y, as even with y copies of the item, we could not get revenue more than 1 per bidder. On the other hand, consider the mechanism that posts a price p and lets any bidder willing to pay p get the item, for $p \to \infty$. Then the probability that at least one bidder chooses to pay is at least $y/p - \binom{y}{2}/p^2$. So the revenue is at least $y - \binom{y}{2}/p$, which approaches y as $p \to \infty$. So the optimal revenue for this setting is indeed y.

Therefore, we get that $y = V_{x,y} \cdot (x-1)$ and $V_{x,y} = \frac{y}{x-1}$. □

Now we can put everything together. We have argued that $\max_{i \in [n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}$ is equal to the ℓ^{th} highest of nm draws with probability $P_\ell \geq \frac{(1-\ln(2))(\ell-1)}{4mn}$ for all $\ell < \sqrt{mn}$, and also argued that the expected value of the ℓ^{th} highest of mn draws is $\frac{mn}{\ell-1}$. Therefore, for all j we get:

$$\begin{aligned}
\mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n}\left[\max_{i \in [n]}\{v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\}\right] &\geq \sum_{\ell < \sqrt{nm}} P_\ell \cdot \mathbb{E}[V_{\ell,nm}] \\
&\geq \sum_{\ell < \sqrt{nm}} \frac{(1-\ln(2))(\ell-1)}{4nm} \cdot \frac{nm}{\ell-1} \\
&\geq \sqrt{nm}/14.
\end{aligned}$$

Summing over all j yields the proposition statement. □

Now with Lemma 18 and Proposition 17, we can conclude the following:

$$\sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n}\left[\max_{i \in [n]}\left\{\varphi_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) + v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\right\}\right] \geq nm + m\sqrt{nm}/14.$$

This immediately implies the following corollary:

Corollary 10. *If one compares to the EFFTW benchmark (which upper bounds the expected revenue) instead of the expected revenue, then the competition complexity of* n *bidders with additive valuations over* $m \leq n/4$ *i.i.d., regular items is at least* $\sqrt{nm}/14$.

Proof. Simply recall that $\textsc{Rev}_{n+c}(\mathcal{ER}) = n + c$, and $\text{VCG}_{n+c}(\mathcal{ER}^m) = m \cdot \textsc{Rev}_{n+c}(\mathcal{ER}) = m(n+c)$. So in order to have $\text{VCG}_{n+c}(\mathcal{ER}^m) \geq nm + \sqrt{nm}/14$, we must have $c \geq \sqrt{nm}/14$. □

A.2 Competition complexity is not independent of n.

In this section, we show that while indeed the "true" competition complexity in the big-n case may be better than what is achievable by comparing to the EFFTW benchmark, it is not independent of n. Specifically, we will argue that for all n and $m = 2$ items, the optimal mechanism for $(\mathcal{ER}^m)^n$ achieves revenue at least $2n + \ln(n)/10$.

Proposition 18. $\textsc{Rev}_n(\mathcal{ER}^2) \geq 2n + \ln(n)/10.$

Proof. Consider the following mechanism:

- Bidders select whether to be "high", "low", or "medium".
- First, all high bidders are processed in random order. When processed, a high bidder can get both remaining items for price p.
- Next, all medium bidders are processed in random order. When processed, a medium bidder will get a random remaining item for price q.
- Low bidders get no items.

We will think of $p \to \infty$, and take $q := \sqrt{n}$. First, it is easy to see that any bidder who values the grand bundle (of both items) at least as $2q$ will choose to be a high or medium bidder. Let's now count the expected number of such bidders.

Lemma 21.

$$\Pr_{(v_1,v_2)\leftarrow\mathcal{ER}^2}[v_1 + v_2 \geq 2q] = \frac{2}{2q-1} + \frac{(q-\frac{1}{2})\cdot\ln(2q-1) - q}{q^2(2q-1)}$$

Proof. There are two ways that that we can have $v_1 + v_2 \geq 2q$. First, maybe $v_1 \geq 2q - 1$. In this case, as $v_2 \geq 1$, surely $v_1 + v_2 \geq 2q$. Second, maybe $v_1 < 2q - 1$ and $v_2 \geq 2q - v_1$. So the probability of both cases together is: $\frac{1}{2q-1} + \int_1^{2q-1} \frac{1}{x^2} \cdot \frac{1}{2q-x} dx$. Also:

$$\begin{aligned}
&\int_1^{2q-1} \frac{1}{x^2} \cdot \frac{1}{2q-x} dx \\
=&\frac{1}{4} \cdot \left[\frac{x\ln(2q-x) + 2q - x\ln x}{q^2 x}\right]_1^{2q-1} \\
=&-\frac{1}{4} \cdot \left(\frac{2q-(2q-1)\cdot\ln(2q-1)}{q^2(2q-1)} - \frac{\ln(2q-1)+2q}{q^2}\right) \\
=&\frac{1}{2} \cdot \frac{-q + q(2q-1) + (2q-1)\cdot\ln(2q-1)}{q^2(2q-1)} \\
=&\frac{q^2 - q + (q-\frac{1}{2})\ln(2q-1)}{q^2(2q-1)}.
\end{aligned}$$

Adding back in the $\frac{1}{2q-1}$, we get the lemma statement. ☐

We will want to take q small enough so that in expectation $n^{\Omega(1)}$ bidders wish to be medium/high. This will let us claim that he expected number of medium/high bidders concentrates around its expectation. But we also want q to be big enough so that we can get some revenue from these cases.

Lemma 22. *Consider any set of $n-1$ bidders, and let $100 \leq q \leq \sqrt{n}$ and $p \to \infty$. Then with probability at least $1 - e^{-n\ln^2(q)/(128q^2)} = 1 - o(1/n)$, none of these bidders are high, and at least $\frac{n}{q} + \frac{n\ln(q)}{8q^2}$ are medium.*

Proof. The number of bidders with $v_1 + v_2$ exceeding $2q$ is a sum of independent $\{0, 1\}$ random variables whose expected value exceeds $\frac{n}{q} + \frac{n\ln(q)}{4q^2}$. (We have assumed $q \geq 100$ in order to simplify the second term).

We therefore wish to understand the probability that the actual number of medium/high bidders is at least its expectation minus $\frac{n\ln(q)}{8q^2}$. This is a simple application of the Chernoff bound with $\mu \geq \frac{n}{q}$ and $\delta \leq \frac{\ln(q)}{8q}$. So the probability of this deviation is at most $e^{-n\ln^2(q)/(128q^2)}$. When $q < \sqrt{n}$, this probability is $o(1/n)$.

Now we just need to recall that this is the probability that at least $\frac{n}{q} + \frac{n\ln(q)}{8q^2}$ bidders are high or medium. Observe further that the probability that any are high bidders is at most $4n/p$ (some bidder must value some item at at least $p/2$, which occurs with probability at most $4n/p$ by union bound), which approaches 0 as $p \to \infty$. $\square$

Now, let's see what value a bidder would need to have in order to prefer to be high instead of medium.

Lemma 23. *Assuming that all other bidders are medium or low, and there are* $k - 1$ *medium bidders, a bidder with* $v_1 + v_2 \geq \frac{pk-2q}{k-1}$ *chooses to be a high bidder instead of medium.*

Proof. If the bidder chooses to be medium, then they will get each item with probability $1/k$, and pay $2q/k$ in expectation. If they choose to be high, they will get each item with probability 1, and pay p. So in order to prefer this, they would need to have $(v_1 + v_2) \cdot (1 - 1/k) \geq p - 2q/k$. Rearranging yields the lemma. $\square$

Now we analyze the revenue from two possible cases. Below for notational simplicity, define $k := \frac{n}{q} + \frac{n\ln(q)}{8q^2}$.

- **Case One: Some bidder has $v_1 + v_2 \geq pk/(k-1)$.** In this case, with probability at least $1 - o(1/n)$, there are no high bidders and at least k medium bidders among the other $n-1$. Conditioned on this, the bidder will choose to be high, and we get revenue p. The probability that this case occurs is at least $2n(1-1/k)/p - \binom{2n}{2}(1-1/k)^2/p^2$ (as it occurs whenever anyone values some item above p). Conditioned on this, with further probability $1-o(1/n)$, this buyer chooses to be high and we get revenue p. So the total revenue from these cases is at least: $(1-o(1/n)) \cdot (p \cdot (2n(1-1/k)/p - \binom{2n}{2}(1-1/k)^2/p^2)) = 2(1-1/k)n - o(1)$ as $p \to \infty$.
- **Case Two: no bidder has $v_1 + v_2 \geq pk/(k-1)$.** In this case, we know that with probability $1-o(1/n)$, there are at least two (in fact, many more) bidders who are medium. So we get revenue $2q$ in these cases. Observe that as $p \to \infty$, we are in this case with probability approaching 1, so the total revenue from these cases is $2q \cdot (1-o(1/n)) = 2q - o(1)$ as $p \to \infty$.

So our revenue is $2n(1-1/k)+2q-o(1)$. Recall that $k := \frac{n}{q} + \frac{n\ln(q)}{8q^2} = \frac{8nq+n\ln(q)}{8q^2}$. Therefore our revenue is:

$$\begin{aligned}&2n - \frac{16nq^2}{8nq+n\ln(q)} + \frac{16nq^2+2nq\ln(q)}{8nq+n\ln(q)}\\ =&2n + \frac{2q\ln(q)}{8q+\ln(q)} \geq 2n + \ln(q)/5.\end{aligned}$$

Recall that we required $q \leq \sqrt{n}$ in order for our calculations to be valid, so by setting $q := \sqrt{n}$, our revenue is at least $2n + \ln(n)/10$.

$\square$

And now we can conclude that the competition complexity cannot be independent of n:

Corollary 11. *The competition complexity of n bidders with additive valuations over $m = 2$ i.i.d., regular items is at least* $\ln(n)/10$*. In particular, there exists no function $f(\cdot)$ only of m such that the competition complexity of n bidders with additive valuations over m i.i.d. items is $f(m)$.*

A.3 Omitted Proofs from Section 3.2

Proof of Fact 1. The proof follows quickly from Myerson's lemma, stated below:

Lemma 24 (Myerson's Lemma [49])**.** *Consider any Bayesian Incentive Compatible mechanism for a single item with payment rule $P(\cdot)$ and allocation rule $X(\cdot)$. That is, on bids $\vec{v}$, the mechanism charges bidder i $P_i(\vec{v})$ and awards bidder i the item with probability $X_i(\vec{v})$. Then for all i, the expected payment made by bidder i is equal to bidder i's expected virtual welfare. That is:*

$$\mathbb{E}_{\vec{v} \leftarrow D^n}\left[P_i(\vec{v})\right] = \mathbb{E}_{\vec{v} \leftarrow D^n}\left[X_i(\vec{v}) \cdot \varphi_D(v_i)\right],$$

$$\mathbb{E}_{\vec{v} \leftarrow D^n}\left[P_i(\vec{v})\right] \leq \mathbb{E}_{\vec{v} \leftarrow D^n}\left[X_i(\vec{v}) \cdot \overline{\varphi}_D(v_i)\right].$$

Now, consider the single-bidder auction that simply sets a price of v. Then it's expected revenue is simply $v \cdot \Pr_{w \leftarrow D}[w \geq v]$. The expected virtual welfare is $\mathbb{E}_{w \leftarrow D}[\varphi_D(w) \cdot \mathbb{I}(w \geq v)]$. Therefore, by Myerson's Lemma:

$$v = \frac{\mathbb{E}_{w \leftarrow D}\left[\varphi_D(w) \cdot \mathbb{I}(w \geq v)\right]}{\Pr_{w \leftarrow D}\left[w \geq v\right]} = \mathbb{E}_{w \leftarrow D_{\geq v}}\left[\varphi_D(w)\right].$$

The proof for ironed virtual values follows identically, after replacing the left-most equality with inequality. □

A.4 Omitted Proofs from Section 2.4

A.4.1 Little n proofs

Proof of Observation 1. For all $\vec{v}$, if $\vec{v}_{(1)} \notin R_j$, then both random variables inside the expectation take value $v_{(1)j}$. If $\vec{v}_{(1)} \in R_j$, then the random variable for the left-hand expectation is at most $\max\{\overline{\varphi}_j(v_{(1)j}), v_{(2)j}\}$ (note that the $^+$ is no longer necessary as we're taking a maximum with $v_{(2)j} \geq 0$ anyway). The right-hand expectation is exactly this. Because the right-hand random variable is larger for all $\vec{v}$, the expectation is larger as well. □

Proof of Proposition 2. We again begin with a coupling. First, couple the draws $X_{i,n} := F_j(v_{ij})$ for all i. Next, couple the quantiles $Y_{\ell,m-1} = F_\ell(v_{(1)\ell})$ for $\ell < m, \ell \neq j$, and $Y_{j,m-1} = F_m(v_{(1)m})$ (if $j \neq m$, otherwise there is no $Y_{m,m-1}$ to define). Crucially, observe that this is a valid coupling, as values for items are drawn independently (in particular, conditioned on drawing v_{ij} for all j, the quantiles $F_\ell(v_{i\ell})$ are still i.i.d. and uniform from $[0,1]$). Observe now the following:

- Always, $v_{(\ell)j} = F_j^{-1}(X_{(\ell),n})$ for all ℓ.
- Whenever $\vec{v}_{(1)} \in R_j$, $X'_L(n,m) = X_{(1),n}$. Therefore:

$$\mathbb{E}_{\vec{v}\leftarrow D^n}\left[\max\{v_{(1)j}\cdot \mathbb{I}(\vec{v}_j \notin R_j), \overline{\varphi}_j(v_{(1)j}), v_{(2)j}\}\cdot \mathbb{I}(\vec{v}_{(1)} \in R_j)\right]$$

$$= \mathbb{E}\left[\max\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), F_j^{-1}(X_{(2),n})\}\cdot \mathbb{I}(X'_L(n,m) = X_{(1),n})\right].$$

- Conditioned on $\vec{v}_{(1)} \notin R_j$, $X'_L(n,m)$ is a uniformly random sample from $[X_{(1),n}, 1]$. This is because there is some strictly positive number of ℓ such that $Y_{\ell,m-1} > X_{(1),n}$. Conditioned on being $> X_{(1),n}$, each such value is drawn uniformly from $[X_{(1),n}, 1]$. And then $X'_L(n,m)$ picks one of them uniformly at random. Using Fact 1, we get:

$$\mathbb{E}_{\vec{v}\leftarrow D}\left[v_{(1)j} \cdot \mathbb{I}(\vec{v}_{(1)} \notin R_j)\right] \leq \mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))) \cdot \mathbb{I}(X'_L(n,m) \neq X_{(1),n})\right].$$

Summing the left-hand side of both equations is exactly

$$\mathbb{E}_{\vec{v}\leftarrow D^n}\left[\max\left\{v_{(1)j} \cdot \mathbb{I}(\vec{v}_{(1)} \notin R_j), \overline{\varphi}_j(v_{(1)j}), v_{(2)j}\right\}\right].$$

Summing the right-hand side is clearly upper bounded by

$$\mathbb{E}\left[\max\left\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), F_j^{-1}(X_{(2),n})\right\}\right].$$

□

For subsequent proofs, we will need one basic fact about maximums of random variables:

Fact 2. *For any three random variables X, Y, Y' such that for all x, $\mathbb{E}[Y'|Y = y, X = x] \geq y$, $\mathbb{E}[\max\{X, Y'\}] \geq \mathbb{E}[\max\{X, Y\}]$.*

Proof. In fact, we show that for all y, x, $\mathbb{E}[\max\{X, Y'\}|Y = y, X = x] \geq \max\{x, y\}$, which implies the desired statement. There are two cases to consider. First, perhaps $x \geq y$. In this case, we clearly have $\max\{x, Y'\} \geq x$ with probability 1. Therefore, $\mathbb{E}[\max\{X, Y'\}|Y = y, X = x] \geq x$ as well. Second, perhaps $y > x$. In this case, by hypothesis we have $\mathbb{E}[Y'|Y = y, X = x] \geq y$. So clearly $\mathbb{E}[\max\{X, Y'\}|Y = y, X = x] \geq y$. This covers both cases and proves the fact. □

Proof of Corollary 3. We first consider the three random variables $X = \overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), Y = F_j^{-1}(X_{(2),n}), Y' = \overline{\varphi}_j(F_j^{-1}(W_{2,n}))$. Then indeed, conditioned on $X = x$ and $Y = y$, $W_{2,n}$ is a uniformly random draw from

$[X_{(2),n}, 1]$. Fact 1 therefore concludes that $\mathbb{E}[Y'|X = x, Y = y] = y$, allowing an application of Fact 2 to conclude $\mathbb{E}\left[\max\left\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), F_j^{-1}(X_{(2),n})\right\}\right] \leq \mathbb{E}\left[\max\left\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), \overline{\varphi}_j(F_j^{-1}(W_{2,n}))\right\}\right]$. The corollary follows by observing that

$\max\left\{\overline{\varphi}_j(F_j^{-1}(X'_L(n,m))), \overline{\varphi}_j(F_j^{-1}(W_{2,n}))\right\} = \overline{\varphi}_j(\max\{X'_L(n,m), W_{2,n}\}) = \overline{\varphi}_j(X_L(n,m))$. ☐

Proof of Corollary 4. The proof is nearly-identical to that of Corollary 1 and omitted (the only difference is that we save a "+1" in the regular case because our current benchmark no longer has the "+" on the virtual values). ☐

A.4.2 Big n proofs

Proof of Proposition 3. Consider drawing quantiles $q_1, \ldots, q_n$, and also $\{q'_{i,k}\}_{i\in[\ell-1],k\in[m-1]}$. Couple draws from $\vec{v}$ from D^n and $\vec{w}$ from $D_j^{n+(\ell-1)(m-1)}$ as follows:

- Set $v_{ij} = w_{ij} = F_j^{-1}(q_i)$ for all i. Relabel the indices so that $v_{(1)j} \geq v_{(2)j} \geq \ldots \geq v_{(n)j}$, $w_{(1)j} \geq w_{(2)j} \geq \ldots \geq w_{(n)j}$, and $q_{(1)} \geq q_{(2)} \geq \ldots \geq q_{(n)}$.
- For $i \in [\ell-1]$, set $v_{(i)k} = F_k^{-1}(q'_{i,k})$ for all $k \neq j$. Set $v_{(i)m} = F_m^{-1}(q'_{i,j})$ (unless $j = m$, in which case $v_{(i)m}$ is already set, and there is no $q'_{i,m}$).
- For $i \in [\ell-1]$, set $w_{(i)k} = F_j^{-1}(q'_{i,k})$ for all $k \neq j$. Set $w_{(i)m} = F_j^{-1}(q'_{i,j})$ (unless $j = m$, in which case $w_{(i)m}$ is already set, and there is no $q'_{i,m}$). As all ws are drawn from D_j, interpret all $n + (\ell-1)(m-1)$ such draws as values of a single bidder for item j. Let $w_{(\ell)}$ denote the ℓ^{th} largest of these draws.
- Observe that $w_{(\ell)} \geq w_{(\ell)j} = v_{(\ell)j}$.

Now, we consider the random variables inside the left-hand and right-hand expectations. Let now i^* denote the minimum i such that $\vec{v}_{(i)} \notin R_j$. Observe that when $i^* < \ell$, this is also the minimum i such that there exists a $k \neq j$ with $F_j(w_{(i)j}) < F_j(w_{(i)k})$.[2] Therefore, $i^* \geq \ell$ only if no i, k exists for which $F_j(w_{(i)j}) < F_j(w_{(i)k})$. So first, consider the possibility that $i^* \geq \ell$. In this case, the contribution to the benchmark is upper bounded by $\max\{\overline{\varphi}_j(v_{(1)j}), v_{(\ell)j}\}$. But we have that $\overline{\varphi}_j(w_{(1)}) \geq \overline{\varphi}_j(v_{(1)j})$, and also $w_{(\ell)} \geq v_{(\ell)j}$. Therefore, we can conclude that:

$$\mathbb{E}_{\vec{v} \leftarrow D^n}\left[\left(\max_{i \in [n]}\left\{\overline{\varphi}_j(v_{ij})^+ \cdot \mathbb{I}(\vec{v}_i \in R_j) + v_{ij} \cdot \mathbb{I}(\vec{v}_i \notin R_j)\right\}\right) \cdot \mathbb{I}(i^* \geq \ell)\right]$$

$$\leq \mathbb{E}_{\vec{w} \leftarrow D_j^{n+(m-1)(\ell-1)}}\left[\left(\max\left\{\overline{\varphi}_j(w_{(1)}), w_{(\ell)}\right\}\right) \cdot \mathbb{I}(i^* \geq \ell)\right].$$

Now, consider the case that $i^* < \ell$. In this case, the contribution to the benchmark is exactly $\max\{\overline{\varphi}_j(v_{(1)j}), v_{(i^*)j}\}$. We now wish to argue that, conditioned on i^*, $v_{(i^*)j}$ and $v_{(1)j}$, the expected contribution to the right-hand side exceeds this. Indeed, observe that there exists at least one k for which $q'_{i^*,k} > q_{(i^*)}$. Conditioned on $q_{(i^*)}$ (and $v_{(1)j}$, which has no effect), $q'_{i^*,k}$ is simply a uniformly random draw from $[q_{(i^*)}, 1]$. That is, $w_{(i^*)k'}$, (where $k' = k$ if $k \neq j$ and $k' = m$ if $k = j$) is simply a draw from D_j, conditioned on exceeding $w_{(i^*)j}$. Therefore, $\mathbb{E}[\overline{\varphi}_j(w_{(i^*)k'}) | w_{(i^*)j} = x, w_{(i^*)k'} > w_{(i^*)j}] \geq x$ (by Fact 1). We can now apply Fact 2 to conclude that $\mathbb{E}[\max\{\overline{\varphi}_j(v_{(1)j}), v_{(i^*)j}\} \cdot \mathbb{I}(i^* < \ell)] \leq \mathbb{E}[\max\{\overline{\varphi}_j(v_{(1)j}), \max_k\{\overline{\varphi}_j(w_{(i^*)k})\}\} \cdot \mathbb{I}(i^* < \ell)]] \leq \mathbb{E}[\overline{\varphi}_j(w_{(1)}) \cdot \mathbb{I}(i^* < \ell)]]$. The last inequality follows simply because $w_{(1)} \geq v_{(1)j}$ and also $w_{(1)} \geq \max_k\{w_{(i^*)k}\}$.

---

[2] If D_j has no point masses, we could instead just write $w_{(i)j} < w_{(i)k}$, but we write it like this to be careful in the case of point masses.

$$\mathbb{E}_{\vec{v}\leftarrow D^n}\left[\left(\max_{i\in[n]}\left\{\overline{\varphi}_j(v_{ij})^+\cdot\mathbb{I}(\vec{v}_i\in R_j)+v_{ij}\cdot\mathbb{I}(\vec{v}_i\notin R_j)\right\}\right)\cdot\mathbb{I}(i^*<\ell)\right]$$
$$\leq\mathbb{E}_{\vec{w}\leftarrow D_j^{n+(m-1)(\ell-1)}}\left[\overline{\varphi}_j(w_{(1)})\cdot\mathbb{I}(i^*<\ell)\right].$$

Summing up the two left-hand sides yields item j's contribution to the EFFTW benchmark. Summing up the two right-hand sides lower bounds the desired right-hand side.

□

Proof of Lemma 1. Again couple draws so that $w_i = F_j^{-1}(X_{i,n'})$. Then $\overline{\varphi}_j(X_{(1),n'}) = \overline{\varphi}_j(w_{(1)})$. Additionally, $\mathbb{E}[\overline{\varphi}_j(W_{\ell,n'})|X_{(1),n'} = x, X_{(\ell,n')} = y] \leq F_j^{-1}(y) = w_{(\ell)}$ by Fact 1. So the hypotheses of Fact 2 are satisfied, and we get that

$$\mathbb{E}_{\vec{w}\leftarrow D^{n'}}\left[\max\left\{\overline{\varphi}_j(w_{(1)}),w_{(\ell)}\right\}\right]\leq\mathbb{E}\left[\max\{\overline{\varphi}_j(F_j^{-1}(X_{(1),n'})),\overline{\varphi}_j(F_j^{-1}(W_{\ell,n'}))\}\right]$$
$$=\mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(\max\{X_{(1),n'},W_{\ell,n'}))\right]=\mathbb{E}\left[\overline{\varphi}_j(F_j^{-1}(X_B(n',\ell)))\right].$$

□

APPENDIX B

APPENDIX TO CHAPTER 4

B.1 Proof of Theorem 10

Proof of Theorem 10. We divide the proof of Theorem 10 into two lemmas as was explained in Section 4.3. (For ease of reading we repeat the statement of the lemmas here.)

Lemma 12 [Structure of equilibria] In a market with two tiers, a symmetric equilibrium always exists and is unique, and for each type of agents $i \in \{L, H\}$, the equilibrium strategy is either pure, or is a randomization between consecutive strategies $(k_i, K - k_i)$ and $(k_i + 1, K - k_i - 1)$ for some $k_i \in \{0, \ldots, K\}$.

Proof. We show that the structure is satisfied for the high doctors strategy. The proof for low doctors is essentially the same, but we must account for the probabilities $1 - \alpha_H$ and $1 - \alpha_L$, that a high and low hospital respectively are already taken by a high doctor (in that case, the low doctor gets automatically rejected).

Consider the market for high doctors and suppose that there is a symmetric equilibrium in which the probability that a doctor is accepted by a top hospital he applies to is p, and p' be the probability that he is accepted by a bottom hospital that he applies to. These probabilities are as defined in Section 5.2.

Now consider the best response of a single player. That player's goal is to choose $(y, K - y)$ to maximize his utility given what everyone else is doing. That is, he chooses (integer) $y \in [0, K]$ to maximize

$$v(1-(1-p)^y)+(1-p)^y(1-(1-p')^{K-y}) = v-(v-1)(1-p)^y-(1-p')^K\left(\frac{1-p}{1-p'}\right)^y.$$

Therefore, given the probabilities p, p' we define $f(y)$, the utility function of an agent when he lists y hospitals of the top tier as follows:

$$f(y) = v-(v-1)(1-p)^y-(1-p')^K\left(\frac{1-p}{1-p'}\right)^y \tag{B.1}$$

Note that, given p and p', the best response of a single agent must maximize $f(y)$. (By the large market assumption, p and p' remain unchanged regardless of the agent's action.) Note that the function f is concave when viewed as a function of a real variable y: It is of the form $c - a\alpha^y - b\beta^y$, where all of $a, b, c, \alpha, \beta > 0$, $\alpha < 1$, $\beta > 1$.[1] The second derivative of this function is

$$-a\ln^2(1/\alpha)\alpha^y - b\ln^2\beta\beta^y$$

which is negative for all y, so the function is strictly concave. Thus, the function either achieves its maximum over integers at some integer k, or it is equal at two adjacent integer values (and strictly smaller at all other integer values). $\square$

Lemma 13 [Existence of equilibria] A symmetric equilibrium always exists.

Proof. First, we show existence for the high ranked doctors. For a proposed solution $X := k+x$ ($k = \lfloor X \rfloor$, as usual), with x fraction of the top doctors listing $k+1$ high-type hospitals, and $K-k-1$ low-type hospitals, and the remaining $(1-x)$ fraction listing $(k, K-k)$ high-type/low-type hospitals respectively. The variable $X = k+x$ is the expected number of slots that a doctor uses to rank high-type hospitals. X ranges in $[0, K]$.

---

[1] Note that, at an equilibrium, we must have $p < p'$ as $v > 1$.

Let $p(X)$ (resp. $p'(X)$) be the probability that a single application of a doctor is accepted by an individual top (resp. low-type) hospital he interviews with, when the strategy is given by X. In Claim 1 (stated after this proof) we show that $p(X)$ is continuous and monotone decreasing, and $p'(X)$ is continuous and monotone increasing. Intuitively, the more slots used to rank high hospitals, the smaller the probability that a single application succeeds in getting accepted, and this relation is continuous. Clearly, $p'(K) = 1$ and $p(0) = 1$, as with no proposals sent there, the probability of a proposal being accepted is 1. If $vp(K) \geq 1$, then sending all proposals to the high-type hospitals is an equilibrium, as the added value of a proposal to the upper tier (conditioned on not yet being matched) is $vp(K) \geq 1$, more than the value of being accepted in a low hospital.

Now assume $vp(K) < 1 = p'(K)$. Using some ℓ slots to rank low-type hospitals, a doctor has a $1-(1-p')^{\ell}$ probability of getting accepted. If an equilibrium agents exists where agents apply to either k or $k+1$ high-type hospitals (by the above lemma, this are the only kind that can exists), then we must have that at the choice of using the $k+1$ slot for a high-type or a low-type hospital, the resident is neutral. Then, we need to have

$$pv + (1-p)(1-(1-p')^{K-k-1}) = (1-(1-p')^{K-k}).$$

Consider the function $g(X)$ defined as the difference of the two sides:

$$g(X) := pv + (1-p)(1-(1-p')^{K-k-1}) - (1-(1-p')^{K-k}), \tag{B.2}$$

with k being the integer part of X, and p, p' being $p(X), p'(X)$ respectively. When X is 0, then recall $p(0) = 1$, and hence this value is positive. When $X = K$ then this value is negative, as $p(K)v < 1 = p'(K)$. By continuity of p, p', the function g is continuous at fractional values of X but can be discontinuous when X is integer.

We first claim that, if $g(X) = 0$ for any X with integer part k, then this is an equilibrium. To see why note that the resident in this case is neutral between using k or $k + 1$ links to the upper part. Recall that the function $f(y)$ above expressing the value of using y links up is strictly concave, so if two neighboring integers have equal value these are the maximum.

Next, suppose the two sides are never equal. We claim that then a pure strategy Nash equilibrium must exists. Recall that $g(0) > 0$ and $g(K) < 0$. Note g is continuous from the right, but at integer points it is not continuous from the left, as at that point the integer part of X changes. If g is never zero (both parts of the above equation are never equal), there must be an integer $X = k$ such that $g(k) < 0$, but the lim-sup is positive when X approaches k from the bottom. We claim that, in this case, all doctors playing $(k, K - k)$ is an equilibrium. As $g(k) < 0$, sending $k + 1$ applications to high-type hospitals is worse than sending k such applications. Here we use that g is decreasing in X, which follows from monotonicity of p and p'. At the same time, if we consider a value $X < k$ but very close to k, p and p' are not affected (or barely), but k got replaced by $k - 1$, and thus the sign of g changed, so now we get that listing k high-type hospitals is better than listing $k - 1$. Again the function expressing the value of best response is concave, so if k is better than both $k - 1$ and $k + 1$, it must be the (unique) optimum.

The same argument with some slight modification holds for low tier doctors. Assume that the allocation of top tier doctors is fixed. For a proposed solution, consider same definition of X, k, x, p and p' for low tier doctors where X is the expected number of applications sent to top tier hospitals. We need to consider $(1 - \alpha_H)$ and $(1 - \alpha_L)$, the fraction of high and low hospitals respectively that are spoken for by high doctors. Therefore $p'(K)$ and $p(0)$, respectively change to α_L and α_H in this case. Here if $vp(K) \geq \beta$, then sending all proposals to the high-type hospitals is an equilibrium, as the added value of a proposal to the upper type (conditioned on not yet being matched) is $vp(K) \geq \alpha_L$, more than the maximum value possible on the lower type. And for the other cases $vp(K) < \alpha_L = p'(K)$ same argument holds for low doctors. □

Claim 1. *Let $X \in [0, K]$ be a the expected number of top applications in the strategy or high type doctors. Then:*

1. *$p(X, r)$ is strictly decreasing in X for all $r > 0$.*
2. *$p'(X, r)$ is strictly increasing in X.*

The properties are easy to see for a fixed (finite) n, additional applications, can only decrease the probability that one is accepted. Since this is true for all finite n, its also true in the limit in our large market model.

B.2 Proofs on Efficiency of Equilibrium

Proof of Proposition 7. As discussed in Section 4.4, under the assumption of $v > e/(e-1)$, no high doctor applies to a low tier hospital. Let x be the fraction of low tier doctors that apply to the top tier in equilibrium. As discussed previously, social welfare is $2vn(1-1/e) + n(v+1)(1-e^{-xr})/e + 2rn(1-e^{-(1-x)})$. There are three different cases for x: $x = 0$, $0 < x < 1$ and $x = 1$. If $x = 0$, the equilibrium welfare is same as SIMPLE. To show this for the case $0 < x < 1$, we are interested in finding a lower bound on the difference between the welfare at equilibrium and SIMPLE welfare. In equilibrium compared to SIMPLE there are less matches to low tier hospitals, and more matches to top tier hospitals. Less matches to low tier hospitals causes a decrease in welfare equal to

$$DEC = 2rn\left[1 - \frac{1}{e} - (1 - e^{-(1-x)})\right].$$

The matches from low tier doctors to high tier hospitals cause an increase of

$$INC = (v+1)n\frac{1}{e}(1 - e^{-xr}).$$

To bound decrease in welfare, we will lower-bound this increase by dropping the $+1$ from $(v+1)$. Now substituting v from the equilibrium condition above, we get that the increase is more than

$$n\frac{exr(1 - e^{-(1-x)})}{(1-x)(1-e^{-xr})}\frac{1}{e}(1 - e^{-xr}) = n\frac{xr(1-e^{-(1-x)})}{1-x}$$

The total difference $(DIFF)$ is greater than:

$$nr\left(\frac{x(1-e^{-(1-x)})}{1-x} - 2\left((1 - \frac{1}{e} - (1 - e^{-(1-x)})\right)\right) = nr\left(\frac{x(1-e^{-(1-x)})}{1-x} - 2\frac{e^x - 1}{e}\right)$$

We claim that this function is monotone decreasing in x, and hence its infimum occurs when x tends to 1. To see that the function is decreasing, consider the derivative with respect to x, which is:

$$nr\frac{e - e^x(x^2 - 3x + 3)}{e(1-x)^2}$$

and is indeed negative for all $0 \leq x < 1$. The value of the difference with fixed r and n is lowerbounded by the limit as x approaches 1, which is $rn(\frac{2-e}{e})$.

Recall that social welfare of SIMPLE is $2n(r+v)(1-\frac{1}{e})$. The ratio of equilibrium to SIMPLE is at least:

$$\frac{SW(equilibrium)}{SW(\textsc{simple})} \geq \frac{2n(r+v)(1-\frac{1}{e}) + rn\frac{2-e}{e}}{2n(r+v)(1-\frac{1}{e})}$$

$$= 1 + \frac{rn\frac{2-e}{e}}{2n(r+v)(1-\frac{1}{e})} \geq 1 + \frac{rn\frac{2-e}{e}}{2nr(1-\frac{1}{e})} = \frac{e}{2(e-1)} > 0.79$$

Finally, note that when $K = 1$, we have $x = 1$ only if $v(1 - e^{-r})/(er) \geq 1$. Then, if $x = 1$, we have $DEC = 2rn(1-1/e)$ and $INC = (v+1)n(1-e^{-r})/e \geq rn$, so here the overall difference is also lower-bounded by $rn(\frac{2-e}{e})$. This gives the same bound as the case of $0 < x < 1$.

$\square$

Proof of Theorem 11 We will show in Lemmas 27 and 28 that the efficiency of an equilibrium is at least a $\frac{1}{2(e-1)}$ of the SIMPLE benchmark. As shown previously SIMPLE generates a $\frac{e-1}{e}$ fraction of the OPTIMUM welfare. Therefore social welfare of equilibrium is at least $\frac{1}{2(e-1)}\frac{e-1}{e} = \frac{1}{2}$ of OPTIMUM.

Lemma 25. *Suppose $v \geq \frac{e}{e-1}$ and the high tier is allowed to submit a list of length $K \geq 1$. Then, the social welfare resulting from matches of high tier doctors with high tier hospitals is not less than their social welfare when $K = 1$.*

Proof. With $v \geq \frac{e}{e-1}$ and $K = 1$, everybody in high tier applies to high hospitals. Based on Proposition 8 the number of applications to top is increasing in K. Therefore there will be more matched doctors to high tier. This implies that social welfare from high tier doctors is more than in the $K = 1$ case. $\square$

Lemma 26. *Social welfare of equilibrium where high tier doctors have lists of size $K \geq 1$ and low tier doctors have lists size 1 is greater than or equal to $\frac{e}{2(e-1)}$ of the social welfare of SIMPLE.*

Proof. By Proposition 8, the number of matches of high tier to high tier (HH) is higher with $K \geq 1$. Also there might be matches from high tier of doctors to low tier of hospitals (HL). Suppose after high doctors' applications, f fraction of high tier hospitals and f' fractions of low tier hospitals are free.

Suppose that in equilibrium, x fraction of low tier doctors apply to top tier and $1-x$ fraction apply to low tier. At an equilibrium with $0 < x < 1$, the probability of acceptance on top times v must be equal to the probability of acceptance on low.

$$v\frac{f(1-e^{-xr})}{xr} = \frac{f'(1-e^{-(1-x)})}{1-x}$$

The DEC and INC terms (as defined in proof of 7) change to following quantities:

$$DEC = 2rn[1 - \frac{1}{e} - f' \times (1-e^{-(1-x)})]$$
$$INC = (v+1)nf(1-e^{-xr}) + rn(v+1)(1-f') + 2(1-f-\frac{e-1}{e})$$

With fixed f and f' we would like to find a lower-bound on $\frac{INC-DEC}{SW(\textsc{simple})}$. Let INC_{LH} be social welfare added by low to high matches. $INC - INC_{LH}$ is independent of x:

$$\begin{aligned} INC_{LH} &\geq vnf(1-e^{-xr}) \\ &= \frac{f'(1-e^{-(1-x)})xr}{f(1-e^{-xr})(1-x)}nf(1-e^{-xr}) \\ &= \frac{nxrf'(1-e^{-(1-x)})}{1-x} \end{aligned}$$

To find a lower-bound on $INC - DEC$, we first find a lower-bound on $INC_{LH} - DEC$.

$$INC_{LH} - DEC \geq nr\left(\frac{xf'(1-e^{-(1-x)})}{1-x} - 2(1-\frac{1}{e}-f'+f'e^{-(1-x)})\right)$$

By separating the part independent from x, we have:

$$= nr\left(\frac{xf'(1-e^{-(1-x)})}{1-x} - 2f'e^{-(1-x)}\right) - 2nr(1-\frac{1}{e}-f')$$

The derivative of this expression with respect to x is negative, therefore its infimum happens when x tends to 1.

$$INC_{LH} \geq -nrf' - 2nr(1 - \frac{1}{e} - f') = nr(f' + \frac{2}{e} - 2)$$

The total difference ($DIFF$) is greater than:

$$rn(f' + \frac{2}{e} - 2) + rn(v+1)(1-f) + 2nv(-f - \frac{e-1}{e})$$
$$\geq rn(f' + \frac{2}{e} - 2) + rn(1-f')$$

Therefore the ratio of welfare at equilibrium to welfare of SIMPLE is at least:

$$\begin{aligned}\frac{SW(equilibrium)}{SW(\text{SIMPLE})} &\geq 1 + \frac{rn(f' + \frac{2}{e} - 2) + rn(1-f')}{2n(r+v)(1-\frac{1}{e})} \\ &\geq 1 + \frac{rn\frac{2-e}{e}}{2nr(1-\frac{1}{e})} = \frac{e}{2(e-1)} > 0.79\end{aligned}$$

□

Lemma 27. *Suppose $v > \frac{e}{e-1}$ and doctors have a list of $K \geq 1$, and suppose in equilibrium, doctors in low tier send at least one application to low tier hospitals. The social welfare of equilibrium in this case is not less than* SIMPLE *benchmark.*

Proof. Since the behavior of high doctors is independent of low doctors because of their priority, as discussed in Lemma 25, their social welfare from top tier hospitals increases and is at least their social welfare in SIMPLE case.

Consider now the low tier doctors matches in equilibrium. Let us call this matching M. Note that we are not considering high tier doctors in this matching. We claim that social welfare by low tier doctors is at least their social welfare if we do not consider their applications to top tier. Let the matching by low tier doctors in this case be M'. The low tier doctors matched in M' remain matched in M. Either they are matched to a high hospital, which contributes more to the social welfare, or if their applications on top all got rejected, they still have a spot in low tier hospital, because they had this spot with more competition when we ignored every low tier doctor matches to top. This satisfies the claim.

Now consider the high doctors matches in equilibrium and the matching M' for low doctors. We claim this has a higher social welfare than SIMPLE. Consider high-to-high (HH), high-to-low (HL) and low-to-low (LL) matches. where the first letter refers to doctor's tier and the second to hospitals' one. HH links in equilibrium make more social welfare than HH links in SIMPLE, due to Lemma 25. Since every low tier doctor sends at least one application to the low tier in M', the number of matches is at least $nr(1-1/e)(1-f)$ where f is the fraction of low tier hospitals occupied by high tier doctors. Therefore, the number of matches from $LL + HL$ is at least $nrf + nr(1-1/e)(1-f) \geq nr(1-1/e)$ and hence social welfare is more than SIMPLE. □

Lemma 28. *Suppose* $v \geq \frac{e}{e-1}$ *and doctors have a list of* $K \geq 1$*, and suppose in equilibrium, doctors in low tier send less than one application to low tier hospitals in expectation. The social welfare of equilibrium in this case is better than equilibrium when top tier has list of size* K *and low tier has list of size* 1.

Proof. Suppose that, for the low doctors, the expected number of applications they send to the lower tier is less than 1. (In other words $X^* > K-1$, where X^* is the expected number of application a low tier doctors sends to the high tier.) Based on equilibrium characterization, $pv \geq p'$, where p is the success of a link from lower tier on top and p' is the probability success of a link from low tier to top tier.

Consider the equilibrium when top hospitals have lists of size K and low hospitals have lists of size 1. Let q and q' be the success probability of an application from low tier doctors to high and low tier hospitals respectively. Call the equilibrium solution of this case E', and the equilibrium solution with K applications E.

In equilibrium E with low tier doctors having $K > 1$ applications, and sending in expectation less than 1 application to low tier hospitals, they are sending more application to high tier, so the high tier success probability p is less than q.

We claim that the number of LL applications in E' is less than in E. Suppose the opposite, this means for utility of an application to low we have $q' < p'$ due to higher chance of acceptance with less number of overall applications. Combining this with $p' \leq vp$, and $p \leq q$, we get $q' < p' \leq vp \leq vq$, which implies $vq > q'$. In this case sending applications down is not equilibrium in E', which contradicts with the assumption that LL applications in E is less than in E'. Therefore there must be more applications from low tier to low tier in equilibrium for K rather than 1. Which means social welfare is higher for $K > 1$. □

B.3 Appendix to Section 4.5

B.3.1 Proof of Proposition 8

Proof. We first prove that X_H^* increases in K. Suppose that $X^*(K)$ is not an integer. Let $g(X, K)$ be defined as

$$g(X, K) := pv + (1 - p)(1 - (1 - p')^{K-k-1}) - (1 - (1 - p')^{K-k}),$$

where k is the integer part of X, and p and p' is the acceptance probability of a single application with X of the applications going to the top part in expectation. This definition of g agrees with the one in Eq. (B.2), where we made the dependence on K explicit. Here

$$p'' = (1 - (1 - p')^{K-k-1})$$

is the probability that $K - k - 1$ applications sent to the lower tier results in success. Using this additional notation, we can write

$$g(X, K) := pv + (1 - p)p'' - (p'' + (1 - p'')p') = p(v - p'') - p'(1 - p'')$$

As $X^* = X^*(K)$ is an equilibrium, we have $g(X^*, K) = 0$. Recall that $g(X, K)$ is decreasing in X for every K. Hence, to show that $X^*(K + 1) > X^*(K)$ it suffices to show that $g(X^*, K + 1) \geq 0$. Let $\hat{p}'$ and $\hat{p}''$ denote the probabilities corresponding to p' and p'' in the last expression for $g(X^*, K + 1)$, and note that p is unchanged, as its only effected by X and not by K.

We claim that $\hat{p}' \leq p'$, as with more applications to the lower part, each has a smaller change of succeeding.

Second, we also claim that $\hat{p}'' \geq p''$. This is true, as with more applications to the bottom part, more hospitals in the bottom will secure an application, and hence overall more applicants will be accepted to some hospital.

This now implies that both $1 - \hat{p}'' \leq 1 - p''$ and $v - \hat{p}'' \leq v - p''$, but since $v > 1$ we also know that

$$\frac{v - \hat{p}''}{1 - \hat{p}''} \geq \frac{v - p''}{1 - p''}$$

Combining these gives us the bound as

$$\begin{aligned} g(X^*, K+1) &:= p(v-\hat{p}'') - \hat{p}'(1-\hat{p}'') \geq p(v-\hat{p}'') - p'(1-\hat{p}'') \\ &= (1-\hat{p}'')(\frac{v-\hat{p}''}{1-\hat{p}''}p - p') \geq (1-\hat{p}'')(\frac{v-p''}{1-p''}p - p') \\ &= \frac{1-\hat{p}''}{1-p''}((v-p'')p - (1-p'')p') = \frac{1-\hat{p}''}{1-p''}g(X^*, K) = 0 \end{aligned}$$

Now, to show that it is increasing in v, it suffices to note that $g(X)$ is strictly increasing in v. As g is decreasing in X, this implies that if X^* is an equilibrium when $v_H = v$ —thus $g(X^*) = 0$—, then for $v' > v$ we have $g(X^*) > 0$, which implies that the new equilibrium for v' must be greater than X^*.

In addition, it is decreasing in r. To see why, note that for a fixed strategy X, the probability p of matching to a high hospital (as defined in Section 5.2) remains the same. However, p' increases; although the proportion of high-doctors matched to high-hospitals is the same regardless of r, the number of low hospitals also increased when r increased. This causes p' to increase. Rewriting g as $p(v-1) - (1-p')^{K-k-1}(p'-p)$, it is easy to see that this is a decreasing function of p'. This implies that, everything else fixed, as p' increases, the X^* such a that $g(X^*) = 0$ will decrease and thus the equilibrium strategy decreases. □

B.3.2 Simulation Results

		$rH = 0.1,\ rL = 0.9$				$rH = 0.3,\ rL = 0.7$				$rH = 0.5,\ rL = 0.5$			
v_H	K	X_H^*	X_L^*	$w(S)$	$w(N)$	X_H^*	X_L^*	$w(S)$	$w(N)$	X_H^*	X_L^*	$w(S)$	$w(N)$
1.001	1	0.094	0.095	1.264	1.223	0.252	0.266	1.265	1.163	0.402	0.430	1.265	1.132
1.001	2	0.125	0.143	1.544	1.521	0.344	0.445	1.544	1.500	0.609	0.781	1.544	1.505
1.001	3	1.000	0.000	1.681	1.667	0.500	0.438	1.682	1.644	0.688	1.031	1.682	1.650
1.001	4	2.000	0.000	1.763	1.759	1.000	0.109	1.764	1.712	0.750	1.188	1.764	1.724
1.001	5	3.000	0.000	1.817	1.816	2.000	0.000	1.817	1.781	1.000	0.813	1.818	1.768
1.001	6	4.000	0.000	1.855	1.855	4.000	0.000	1.855	1.853	3.000	0.000	1.855	1.812
1.001	7	5.000	0.000	1.882	1.883	5.000	0.000	1.883	1.883	4.000	0.000	1.883	1.857
1.01	1	0.109	0.092	1.266	1.224	0.264	0.260	1.268	1.167	0.410	0.426	1.271	1.138
1.01	2	0.500	0.945	1.545	1.529	0.500	0.313	1.548	1.509	0.625	0.750	1.551	1.512
1.01	3	1.000	0.000	1.683	1.668	1.000	0.031	1.686	1.641	0.750	0.875	1.690	1.657
1.01	4	2.000	0.000	1.765	1.760	2.000	0.000	1.768	1.749	2.000	0.000	1.772	1.726
1.01	5	4.000	0.000	1.819	1.823	3.000	0.000	1.822	1.815	3.000	0.000	1.826	1.805
1.01	6	5.000	0.000	1.856	1.860	4.000	0.000	1.860	1.858	4.000	0.000	1.864	1.854
1.01	7	6.000	0.000	1.884	1.887	5.000	0.000	1.888	1.888	5.000	0.000	1.891	1.887
1.1	1	0.258	0.073	1.277	1.239	0.381	0.203	1.302	1.208	0.504	0.324	1.328	1.202
1.1	2	1.000	0.000	1.559	1.559	1.000	0.004	1.590	1.579	1.000	0.242	1.621	1.615
1.1	3	2.000	0.000	1.698	1.702	2.000	0.000	1.732	1.739	2.000	0.000	1.765	1.763
1.1	4	3.000	0.000	1.781	1.786	3.000	0.000	1.816	1.827	3.000	0.000	1.851	1.860
1.1	5	4.000	0.000	1.835	1.840	4.000	0.000	1.871	1.883	4.000	0.000	1.908	1.920
1.1	6	5.000	0.000	1.873	1.877	5.000	0.000	1.910	1.921	5.000	0.000	1.947	1.960
1.1	7	6.000	0.000	1.901	1.904	6.000	0.000	1.938	1.948	6.000	0.000	1.976	1.988
1.5	1	0.883	0.000	1.328	1.320	0.895	0.000	1.454	1.433	0.914	0.000	1.580	1.552
1.5	2	1.000	0.088	1.621	1.636	1.094	0.320	1.775	1.810	1.328	0.547	1.930	1.956
1.5	3	2.000	0.070	1.765	1.782	2.000	0.266	1.933	1.980	2.000	0.656	2.102	2.173
1.5	4	3.000	0.066	1.851	1.866	3.000	0.250	2.028	2.069	3.000	0.531	2.204	2.266
1.5	5	4.000	0.070	1.908	1.920	4.000	0.250	2.089	2.125	4.000	0.500	2.271	2.325
1.5	6	5.000	0.078	1.947	1.958	5.000	0.250	2.133	2.163	5.000	0.500	2.318	2.364
1.5	7	6.000	0.094	1.976	1.986	6.000	0.313	2.164	2.192	6.000	0.500	2.352	2.392
3	1	1.000	0.121	1.517	1.512	1.000	0.350	2.023	2.010	1.000	0.572	2.529	2.519
3	2	2.000	0.215	1.852	1.860	2.000	0.875	2.470	2.489	2.000	1.000	3.087	3.179
3	3	3.000	0.289	2.017	2.026	3.000	1.000	2.690	2.728	3.000	1.000	3.362	3.449
3	4	4.000	0.352	2.116	2.124	4.000	1.000	2.821	2.862	4.000	2.000	3.526	3.573
3	5	5.000	0.406	2.180	2.187	5.000	1.000	2.907	2.945	5.000	2.000	3.634	3.685
3	6	6.000	0.469	2.225	2.231	6.000	1.250	2.967	2.993	6.000	2.000	3.709	3.757
3	7	6.875	0.531	2.258	2.264	6.875	1.500	3.011	3.035	6.875	2.750	3.764	3.802
10	1	1.000	0.508	2.402	2.342	1.000	1.000	4.678	4.889	1.000	1.000	6.953	7.600
10	2	2.000	1.000	2.933	2.915	2.000	1.281	5.712	6.015	2.000	2.000	8.490	8.785
10	3	3.000	1.000	3.194	3.232	3.000	2.000	6.220	6.413	3.000	2.078	9.246	9.712
10	4	4.000	1.000	3.350	3.396	4.000	2.000	6.523	6.726	4.000	3.000	9.697	10.019
10	5	5.000	1.156	3.452	3.484	5.000	3.000	6.722	6.809	5.000	3.688	9.993	10.212
10	6	6.000	1.438	3.524	3.541	6.000	3.000	6.862	6.960	6.000	4.000	10.199	10.404
10	7	7.000	1.688	3.576	3.587	7.000	3.000	6.963	7.058	7.000	4.750	10.351	10.493

Table B.1: Equilibria and welfare outcomes for different combinations of values v_H and length of list K. We assume that the number of high hospitals and doctors agrees, as well as the number low doctors and low hospitals. However, we assume three different market compositions: the proportion of high doctors (r_H) is 0.1, 0.3, and 0.5; the proportion of low doctors (r_L) is always $1 - r_H$).

X_H^*	X_L^*	r	v	w(S)/w(N)
0.094	0.095	1	1.001	1.117491166
1	0.3	10	6.171953564	1.003290867
1	0.4	10	8.328930557	1.01088794
1	0.5	10	10.76816087	1.018342694
1	0.6	10	13.47584871	1.025452035
1	0.7	10	16.45401694	1.032211033
1	0.8	10	19.71625006	1.038664643
1	0.9	10	23.28395803	1.04486198
1	0.1	1000	179.2345456	1.007484894
1	0.2	1000	374.2197676	1.014518933
1	0.3	1000	586.4670089	1.02102172
1	0.4	1000	817.6380872	1.027133691
1	0.5	1000	1069.560558	1.032951674
1	0.6	1000	1344.244542	1.038544177
1	0.7	1000	1643.901282	1.043960656
1	0.8	1000	1970.9636	1.049237358
1	0.9	1000	2328.108456	1.054401137

Table B.2: Ratio between social welfare of Simple and Equilibrium with different parameters for v and r and $K = 1$, showing that social welfare of Simple are in cases higher than equilibrium. The first row shows this phenomenon with v close to 1, while other instances have rather high v.

APPENDIX C

APPENDIX TO CHAPTER 5

C.1 No Limit for Granting Interviews

In this section, we recall the analysis for the baseline case, studied in [8] with n doctors and n hospitals in the market. In this case there is no limit on the number of interviews by hospitals and doctors are allowed to list only up to k hospitals, for $k << n$.

Example 4. *We show when* $k = 1$*, the social welfare of the matching is* $(1 - 1/e) \approx 0.63$ *in equilibrium. In a matching between hospitals and doctors the size of matching equals the number of matched hospitals, therefore the social welfare the ratio of the number of matched hospitals to the size of maximum matching which equals the fraction of matched hospitals. As Lemma 15 implies, each doctor applies to his/her favorite hospital. The probability of a hospital receiving at least one application is* $1-(1-1/n)^n$ *which tends to* $1-1/e$ *with* n *approaching* ∞*. Since in the deferred acceptance algorithm each doctor proposes to exactly one hospital, each hospital who receive an application in the first stage receives a proposal and become matched.*

When doctors apply to more than one hospital ($k > 1$), Arnosti [1] proposes the following formula for the probability p of a single proposal resulting to a permanent match as n grows to ∞. The main idea is that in the limit as n goes to infinity, the probability of a proposal resulting in a match is independent of the previous proposals of the applicant being rejected.

Proposition 19. *[1] Using the above large market approximation with each doctor* *applying to k hospitals, the probability p of a single proposal resulting in a permanent match in the deferred acceptance procedure, satisfies the following equation.*

$$(1-p)^k = e^{-(1-(1-p)^k)/p} \tag{C.1}$$

proof idea. Since the outcome of the deferred acceptance algorithm does not depend on the order in which doctors propose, we may hold out a single doctor d and run the deferred acceptance algorithm on the remainder of the market. Now consider d proposing to her favorite position. Her first few proposals may get rejected. Once a hospital accepts her proposal, it may reject a different doctor, who may propose for her next position, etc. We call the resulting sequence of rejections a *rejection chain.* The probability that a proposal of doctor d causes a rejection chain that gets doctor d rejected from the hospital that first accepted her vanishes as the market grows, therefore we may assume that d's first tentatively accepted proposal will lead to a permanent match. Also, in a large market, learning that d's first m proposals got rejected does not affect the probability of acceptance of other proposals. Thus, from d's perspective, each hospital that she applied to in the first stage, should be available to her with some probability p, and their availability should be independent. With this argument, the probability that d matches is $1-(1-p)^k$, and the expected number of hospitals d proposes to is

$$1 + (1-p) + (1-p)^2 + ... + (1-p)^{k-1} = \frac{1}{p}(1-(1-p)^k)$$

From the point of view of each hospital, each of these proposals is sent to them roughly with probability $1/n$; thus the probability that a hospital receives at least one proposal is

$$1-(1-1/n)^{\frac{n}{p}(1-(1-p)^k)} \approx 1 - e^{-(1-(1-p)^k)/p}.$$

Since doctors match with probability $1-(1-p)^k$, and the number of doctors and hospitals that match must be equal, we have that

$$(1-p)^k = e^{-(1-(1-p)^k)/p}.$$

□

Figure C.1 shows the expected size of matching as a function of the expected number of applications by doctors. We extended the figure for a fractional number of applications. When the expected number of applications is x, such that $k < x < k+1$, doctors have either k or $k+1$ applications. As seen in the figure, the matching size is increasing in the number of applications.

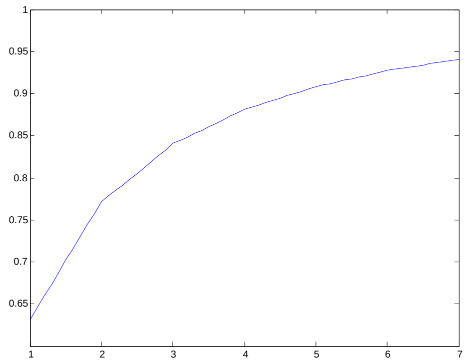

Figure C.1: Size of matching, with no limit on interviews. The horizontal axis is the expected number applications and the vertical axis is the size of matching. The size of the matching is increasing in the expected number of applications.

Proposition 20. *Suppose that either all the doctors apply to the same number of hospitals or a fraction of them apply to k while others apply to $k+1$ hospitals. In a setting where there is no limit on the number of interviews by hospitals and the preferences are uniformly random, the social welfare of the matching is increasing in the expected number of applications.*

Proof. We show that if one of the doctors who previously had k applications, now has $k+1$ applications and the number of applications of other doctors remain the same, the size of the matching can only increase. By lemma 15, both doctors and hospitals are truthful, therefore we are comparing the size of matching as the result of deferred acceptance algorithm when doctor d has an extra application and all other doctors have the same number of applications. Since the deferred acceptance algorithm is oblivious to the order in which doctors proposes, we may hold out the last application of doctor d and find the outcome when the doctor d does not have this application in his/her list. The result of the deferred acceptance algorithm, without this application is the same as the case where doctor d had k applications. We show that this last application can only increase the number of matching. If doctor d is matched with one of his/her first k proposals, the last application does not change. If doctor d proposes to $k + 1_{st}$ hospital, it can only increase the number of hospitals who have received any proposal. □

BIBLIOGRAPHY

[1] Nick Arnosti. Centralized clearinghouse design: A quantity-quality tradeoff. *Working Paper*, 2016. 69, 71, 76, 94, 95, 99, 110, 159

[2] Arash Asadpour and Hamid Nazerzadeh. Maximizing stochastic monotone submodular functions. *Management Science*, 62(8):2374–2391, 2015. 50, 54, 60, 61

[3] Chen Attias, Robert Krauthgamer, Retsef Levi, and Yaron Shaposhnic. Stochastic selection problems with testing. 2017. 52

[4] Christopher Avery and Jonathan Levin. Early admissions at selective colleges. *The American Economic Review*, 100(5):2125–2156, 2010. 72, 94

[5] Moshe Babaioff, Yannai A. Gonczarowski, and Noam Nisan. The menu-size complexity of revenue approximation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 869–877, 2017. 17

[6] Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S. Matthew Weinberg. A simple and approximately optimal mechanism for an additive buyer. In *the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2014. 12, 16, 17

[7] Hedyeh Beyhaghi and Robert Kleinberg. Pandora's problem with nonobligatory inspection. *Proceedings of the twentieth ACM conference on Economics and Computation, EC '19*, 2019. 11

[8] Hedyeh Beyhaghi, Daniela Saban, and Éva Tardos. Effect of selfish choices in deferred acceptance with short lists. *Proceedings of the 4th International Workshop on Matching Under Preferences, MATCH-UP '17*, 2017. 11, 94, 119, 120, 159

[9] Hedyeh Beyhaghi and Éva Tardos. Two-sided matching with limited interviews. *Manuscript*, 2019. 11

[10] Hedyeh Beyhaghi and S. Matthew Weinberg. Optimal (and benchmark-optimal) competition complexity for additive buyers over independent items. *Proceedings of the 51st ACM Conference on Theory of Computation, STOC '19*, 2019. 11

[11] Patrick Briest, Shuchi Chawla, Robert Kleinberg, and S. Matthew Weinberg. Pricing lotteries. *J. Economic Theory*, 156:144–174, 2015. 12

[12] Jeremy Bulow and Paul Klemperer. Auctions versus negotiations. *The American Economic Review*, pages 180–194, 1996. 5, 13, 16, 19

[13] Yang Cai, Nikhil Devanur, and S. Matthew Weinberg. A duality based unified approach to bayesian mechanism design. In *Proceedings of the 48th ACM Conference on Theory of Computation(STOC)*, 2016. `https://arxiv.org/abs/1812.01577`. 14, 17, 19, 20, 21

[14] Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 170–183, 2017. 12, 16, 17

[15] Hector Chade, Gregory Lewis, and Lones Smith. Student portfolios and the college admissions problem. *The Review of Economic Studies*, 81(3):971–1002, 2014. 72

[16] Shuchi Chawla, Jason D. Hartline, and Robert D. Kleinberg. Algorithmic Pricing via Virtual Valuations. In *the 8th ACM Conference on Electronic Commerce (EC)*, 2007. 12, 16, 17

[17] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-Parameter Mechanism Design and Sequential Posted Pricing. In *the 42nd ACM Symposium on Theory of Computing (STOC)*, 2010. 12, 16, 17

[18] Shuchi Chawla, David L. Malec, and Balasubramanian Sivan. The power of randomness in bayesian optimal mechanism design. *Games and Economic Behavior*, 91:297–317, 2015. 12, 14, 16, 17, 20

[19] Shuchi Chawla and J. Benjamin Miller. Mechanism design for subadditive agents via an ex ante relaxation. In *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016*, pages 579–596, 2016. 12, 16, 17

[20] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767, 2017. 12

[21] Brian C. Dean, Michel X. Goemans, and Jan Vondrák. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Mathematics of Operations Research*, 33(4):945–964, 2008. 50, 54

[22] Laura Doval. Whether or not to open Pandora's box. *Journal of Economic Theory*, 175:127 – 158, 2018. 7, 48, 52

[23] Joanna Drummond, Allan Borodin, and Kate Larson. Natural interviewing equilibria for stable matching. 2016. 72, 94

[24] Alon Eden, Michal Feldman, Ophir Friedler, Inbal Talgam-Cohen, and S. Matthew Weinberg. The competition complexity of auctions: A bulow-klemperer result for multi-dimensional bidders. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, page 343, 2017. vii, 6, 13, 14, 15, 16, 17, 18, 20, 21, 27, 29

[25] Alon Eden, Michal Feldman, Ophir Friedler, Inbal Talgam-Cohen, and S. Matthew Weinberg. A simple and approximately optimal mechanism for a buyer with complements: Abstract. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, page 323, 2017. 12, 16, 17

[26] Michal Feldman, Ophir Friedler, and Aviad Rubinstein. 99% revenue via enhanced competition. In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, pages 443–460, 2018. viii, 6, 15, 16, 17, 28, 122, 123

[27] Hu Fu, Christopher Liaw, Pinyan Lu, and Zhihao Gavin Tang. The value of information concealment. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 2533–2544, 2018. 17

[28] D. Gale and L. S. Shapley. College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69(1):9–15, 1962. 67, 99

[29] David Gale and Marilda Sotomayor. Ms. machiavelli and the stable matching problem. *The American Mathematical Monthly*, 92(4):261–268, 1985. 68

[30] J. C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society*, 41(2):148–177, 1979. 51

[31] J. C. Gittins and D. M. Jones. A dynamic allocation index for the sequential design of experiments. pages 241–266, 1974. 51

[32] Eric Grimson, Eva Tardos, Debra Richardson, and Jeff Vitter. Cra guidelines for enhancing faculty recruitment. 2010. 90

[33] Sudipto Guha, Kamesh Munagala, and Saswati Sarkar. Information acquisition and exploitation in multichannel wireless networks. 05 2008. 7, 8, 50, 52

[34] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Algorithms and adaptivity gaps for stochastic probing. In *Proc. 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1731–1747. SIAM, 2016. 50, 54

[35] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Adaptivity gaps for stochastic probing: Submodular and xos functions. In *Proc. 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1688–1702. SIAM, 2017. 50, 54

[36] Isa E Hafalir, Rustamdjan Hakimov, Dorothea Kübler, and Morimitsu Kurino. College admissions with entrance exams: Centralized versus decentralized. *Working Paper*, 2014. 72

[37] Sergiu Hart and Noam Nisan. The menu-size complexity of auctions. In *the 14th ACM Conference on Electronic Commerce (EC)*, 2013. 12

[38] Sergiu Hart and Noam Nisan. Approximate revenue maximization with multiple items. *J. Economic Theory*, 172:313–347, 2017. 12, 16, 17

[39] Sergiu Hart and Philip J. Reny. Maximizing Revenue with Multiple Goods: Nonmonotonicity and Other Observations. *Theoretical Economics*, 10(3):893–922, 2015. 12

[40] Jason D. Hartline. *Approximation and Mechanism Design.* 2011. 19

[41] Nicole Immorlica and Mohammad Mahdian. Marriage, honesty, and stability. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 53–62 (electronic). ACM, New York, 2005. 71, 75, 94, 99, 117

[42] Sangram Vilasrao Kadam. Interviewing in matching markets. *Working Paper*, 2015. 72, 94, 95

[43] Robert Kleinberg, Bo Waggoner, and E. Glen Weyl. Descending price optimally coordinates search. In *Proc. 17th ACM Conference on Economics and Computation (EC)*, pages 23–24, 2016. arXiv:1603.07682 [cs.GT]. 51, 53, 58, 59

[44] Robert Kleinberg, Bo Waggoner, and E. Glen Weyl. Descending price optimally coordinates search. In *Proceedings of the ACM Symposium of Ecomomic and Computation*. ACM, New York, 2016. 72

[45] Fuhito Kojima and Parag A. Pathak. Incentives and stability in large two-sided matching markets. *American Economic Review*, 99(3):608–27, 2009. 94

[46] Robin S Lee and Michael Schwarz. Interviewing in two-sided matching markets. Working Paper 14922, National Bureau of Economic Research, April 2009. 94

[47] Xinye Li and Andrew Chi-Chih Yao. On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237, 2013. 12, 16, 17

[48] Siqi Liu and Christos-Alexandros Psomas. On the competition complexity of dynamic mechanism design. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 2008–2025, 2018. 16, 17

[49] Roger B. Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73, 1981. 13, 19, 138

[50] Gregory Pavlov. Optimal mechanism for selling two goods. *The B.E. Journal of Theoretical Economics*, 11(3), 2011. 12

[51] Peter Postl. Delegated search: Procedure matters, 2004. Working paper. 52

[52] Alvin E Roth. The economics of matching: Stability and incentives. *Mathematics of Operations Research*, 7(4):617–628, 1982. 68

[53] Alvin E Roth and Marilda Sotomayor. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541, 1992. 67

[54] Tim Roughgarden, Inbal Talgam-Cohen, and Qiqi Yan. Supply-limiting mechanisms. In *13th ACM Conference on Electronic Commerce (EC)*, 2012. 16, 17

[55] Aviad Rubinstein and S. Matthew Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. In *Proceedings of the 16th ACM Conference on Electronic Commerce*, 2015. 12, 16, 17

[56] Sahil Singla. The price of information in combinatorial optimization. In *Proc. 29th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2523–2532, 2018. 51, 53

[57] George J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):213–225, 1961. 46

[58] John Thanassoulis. Haggling over substitutes. *Journal of Economic Theory*, 117:217–245, 2004. 12

[59] Martin L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–654, 1979. 6, 46, 51, 57

[60] Andrew Chi-Chih Yao. An n-to-1 bidder reduction for multi-item auctions and its applications. In *the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2015. 12, 16, 17