

Longer Lists Yield Better Matchings

YURI FAENZA AND AAPALI VUORINEN, IEOR, COLUMBIA UNIVERSITY

Many centralized mechanisms for two-sided matching markets that enjoy strong theoretical properties assume that the planner solicits full information on the preferences of each participating agent. In particular, they expect that participants compile and communicate their complete preference lists over agents from the other side of the market. However, real-world markets are often very large and agents cannot always be expected to even produce a ranking of all options on the other side. It is therefore important to understand the impact of incomplete or truncated lists on the quality of the resultant matching.

In this paper, we focus on the Serial Dictatorship mechanism in a model where each agent of the proposing side (students) has a random preference list of length d , sampled independently and uniformly at random from n schools, each of which has one seat. Our main result shows that if the students primarily care about being matched to any school of their list (as opposed to ending up unmatched), then all students in position $i \leq n$ will prefer markets with longer lists, when n is large enough. Schools on the other hand will always prefer longer lists in our model. We moreover investigate the impact of d on the rank of the school that a student gets matched to.

Our main result suggests that markets that are well-approximated by our hypothesis and where the demand of schools does not exceed supply should be designed with preference lists as long as reasonable, since longer lists would favor all agents.

CONTENTS

Abstract	0
Contents	0
1 Introduction	1
1.1 Our Contributions	1
1.2 Organization of the Paper	2
1.3 Related Literature	3
1.4 Notation	4
2 Models and Results	4
2.1 A Discrete Random Market Model	4
2.2 A Continuous Market Model	6
2.3 From Continuous to Discrete	7
2.4 Implications and Extensions	10
2.5 Discussion of the Discrete Model	11
3 Proofs of Main Results	11
3.1 Connections Between the Discrete and Continuous Markets	11
3.2 Continuous Market	15
3.3 Discrete Market: Probability of Being Matched	16
3.4 Discrete Market: Impact on Rank	18
4 Numerical Experiments	18
5 Conclusions	19
References	20
A Background on Differential Equations and Stochastic Processes	21
B Extension to Schools Having Multiple Seats	24
C Missing Proofs	26
D Plots for Numerical Experiments	34

1 Introduction

One of the central objectives of theoretical research in matching markets is to design mechanisms and rigorously establish whether they, along with their outcomes, satisfy desirable properties. Depending on the specific market, the focus may be on concepts such as strategy-proofness, one- or two-sided optimality, or equilibria. These and other properties are often viewed as justifications for applying such mechanisms in practice, under the expectation that the valuable theoretical features will persist in the real world. However, in many applications, the assumptions that are necessary for such properties to hold may not be satisfied. For example, while in a static two-sided market one can compute a stable assignment efficiently, in real-world markets arising in public school allocation, a substantial number of agents often enter or leave the market after the first assignment has been decided (see for example [Ashlagi et al., 2023]). The goal of the central planner becomes therefore less well-defined and, depending on the mathematical formalization, may lead to computationally hard problems, see for example [Bampis et al., 2023, Faenza et al., 2024, 2025, Miyazaki and Okamoto, 2019].

In this paper, we focus on one of the main restrictions encountered in real-world markets: imposing a limit on the maximum length of preference lists of agents. Such limits are common, for instance in school markets. While Gale and Shapley’s Deferred Acceptance algorithm [Gale and Shapley, 1962] assumes that students are allowed to list all schools they deem acceptable in their preference list, education departments traditionally impose a strict maximum limit on the number of schools an applicant may list¹. This restriction is motivated by practical concerns such as the additional burden that acquiring more information poses to schools and students alike. However, the effects of restricting the length of preference lists are significant on the properties of the mechanism and of its outcome. For instance, while the original Deferred Acceptance mechanism is strategy-proof for the proposing side, its implementation with bounded-size lists leaves room for strategic behavior of students. This is not merely a theoretical concern, but has practical implications [Calsamiglia et al., 2010] which are well known to central planners. For instance, until the school year 2024/2025, the Department of Education of New York City limited² the number of schools in student preference lists to 12 and suggested that the applicants be strategic by reserving some of the slots for “safe schools” (schools where the student has high priority, and as a result, a high probability of being accepted) [DOE, 2023].

When deciding the length of preference lists in a matching market, it is therefore essential to carefully balance practical concerns with the potential efficiency losses that can result from limiting information exchange. A correct estimation of the efficiency loss induced by short(er) lists can therefore guide a market designer in striking the right trade-off.

1.1 Our Contributions

In this paper, we focus on the impact that the length of preference lists has on the quality of the output matching in the Serial Dictatorship mechanism for two-sided matching markets. In our theoretical model, we assume that the proposing side (students) have preference lists drawn uniformly at random and that the disposing side (schools) have only one seat (we later discuss how these assumptions can be relaxed in our computational experiments).

The concept of “quality of a matching” can be defined in multiple ways; we mostly focus on the probability that a student will be matched to any school of their preference list as opposed to remaining unmatched. This is justified by the fact that if a student cannot be matched to any

¹These include programs in Spain and Hungary [Calsamiglia et al., 2010], Australia [Artemov et al., 2017], as well as some cities in the United States.

²This rule was changed for the school year 2025/2026, and applicants can now list as many schools as they want.

school of their preference list, they completely lose control on the school they are matched to (in many markets, such students are assigned a remaining seat arbitrarily at the whim of the central planner). This seems by far the worse outcome for a student. We moreover study the probability that a student is matched to their top k choices, a popular measure of quality of a matching³.

1. Comparative analysis of balanced markets. In our model, the probability p_i that a student i is matched to any school is determined by two opposite drivers. On one hand, the longer the preference lists, the more schools will have been matched to students with priority higher than i , decreasing p_i . On the other hand, longer preference lists imply that a student has a larger list of acceptable schools, leading to an increase in p_i . Our main result is that under the Serial Dictatorship mechanism, when the number of schools n is large enough, the latter driver dominates for all students in positions $i \leq n$: that is, p_i increases with the length of lists. This is our Theorem 2.1.

In particular, if the demand of schools does not exceed supply, all students will prefer longer lists. It is not hard to see that in such a market, all schools are matched with higher probability when preference lists are longer (see Lemma 2.2). Therefore our results suggest that markets that are well-approximated by these hypotheses should be designed with preference lists as long as reasonable, since longer lists favor all agents.

2. Comparative analysis of general markets. Interestingly, it is not true in general that longer lists will increase the probability of being matched to some school for every student. Indeed, for n large enough, all students in position $i \geq \lceil 1.22n \rceil$ will be matched with a higher probability when the length of preference lists is 1 (students randomly choose one school to apply to), as opposed to when the length of preference lists is 2, see Section 2.3 and Figure 3. More generally, for all $d < \ell$ and n large enough, there exists some $i > n$ such that student i is matched with higher probability when preference lists are of length d , as opposed to when preference lists are of length ℓ , see Lemma 2.3.

3. Absolute bounds on the probability of being matched. While the previous results compare the probabilities of a given student getting matched to any school between markets with different list length, they do not give us any information on the absolute probability. We show in Theorem 2.4 that as the length of preference lists increases, the probability that the student in position n is matched (recall that n is the number of schools) quickly approaches $1/2$. This is therefore also a lower bound (resp., an upper bound) to the probability of a student $i \leq n$ (resp., $i \geq n$) of being matched. In Lemma 2.5, we further prove bounds on the probability that a student in position $i \leq n$ gets matched to one of their top- k choices as a function of the length of the market.

4. Numerical results. In Section 4, we numerically study two extensions of the model. We first focus on the case when the preference lists are not sampled uniformly at random, but rather, schools are sampled i.i.d. from one of 5 distributions that place different weights on different schools. With the exception of one such distribution (a “degenerate” one), numerical experiments confirm our main result: that every student in position $i \leq n$ will be matched with higher probability when lists are longer. We then focus on the case when schools have $q > 1$ available seats each, and confirm via simulations, that in this case too, students in balanced markets always prefer longer lists.

1.2 Organization of the Paper

We conclude this introductory section with further pointers to the literature. Section 2 is devoted to introducing the main models and ideas, and formally stating many results (including those discussed above), without proofs. In particular, in Section 2.1, we introduce the main (discrete)

³For instance, the National Resident Matching Program, while using the Deferred Acceptance algorithm, explicitly reports the number of applicants matched to their first choice [NRMP, 2023], while the Boston Mechanism directly aims at maximizing the number of students matched to their first choice.

model that we investigate in this paper. To analyze its properties for n large enough, it will be useful to consider a continuous model that can be interpreted as a limit of the discrete model. This continuous model is introduced in Section 2.2. In Section 2.3 we state the connections between the two models, as well as some relevant properties of the continuous model. We moreover discuss some implications of our results for Random Serial Dictatorship and extend our models to the case of schools having multiple seats in Section 2.4. Last, we discuss the relevance of our model and our hypothesis in Section 2.5.

Proofs of results stated in Section 2 appear in Section 3: we start with a limit theorem rigorously establishing the connection between the continuous and the discrete models (Section 3.1), followed by proofs of properties of the continuous market (Section 3.2). Results from Section 3.1 and Section 3.2 will allow to deduce properties of the discrete market (Section 3.3 and Section 3.4).

Numerical experiments testing the validity of our theoretical results when some assumptions are relaxed appear in Section 4. We conclude in Section 5.

1.3 Related Literature

Random models for matching markets have a long and rich history. The model that is closest to ours is the one by [Immorlica and Mahdian, 2015], where the authors consider random preference lists and assume that one of the two sides has preference lists of bounded length, providing an investigation of the model for large markets. In particular, [Immorlica and Mahdian, 2015] show that with high probability and market size large enough, the core (i.e., the set of stable matchings) has small size. [Kanoria et al., 2021] study a similar model, investigating the change in the quality of the matchings in the core as a function of the preference list length. Similar questions have also been answered in the case when preference lists are complete and markets are balanced [Pittel, 1989, 1992] or unbalanced [Ashlagi et al., 2017, Pittel, 2019].

The effect of short lists on the behavior of agents and on the quality of the output matching has been the subject of experimental studies [Artemov et al., 2017, Bó and Hakimov, 2020, Calsamiglia et al., 2010]. While specifics differ, they all share the common message that a loss of efficiency is experienced when agents are asked to report preference lists of bounded length, as opposed to preference lists of arbitrary length.

Many theoretical studies have been devoted to Serial Dictatorship and Random Serial Dictatorship, with goals different from ours – see, for instance, [Abdulkadiroğlu and Sönmez, 1998, 2003, Bade, 2020, Bampis et al., 2023, Bogomolnaia and Moulin, 2001] and the references within the survey article [Ehlers and Klaus, 2023].

To investigate our discrete model, we introduce a continuous model that can be thought of as a limit of the former. Continuous or semi-continuous models for two-sided markets have recently received quite some attention, since they often allow for tighter analysis, see, e.g., [Arnosti, 2022, 2023, Azevedo and Leshno, 2016]. In particular, the work in [Arnosti, 2023] is closely aligned in its methods with our work. They similarly introduce a description of the limit of a discrete market via a solution to an initial value problem; then utilize Markov chain methods via an analogue of our Theorem 2.6 to show convergence of certain quantities relating to that market. Their differential equation is very general and allows for arbitrary distributions and varying lengths of preferences across students. The price one pays for such generality is tractability, and they do not produce expressions ripe for analytical solutions. In contrast, our main technical contribution is to carefully construct an initial value problem for our specific case but that produces a simple interpretation connected to the preference of students and can be readily analyzed to yield insights for the original discrete model.

The outcome of the Serial Dictatorship mechanism in our model can also be understood as the application of a randomized version of the greedy algorithm on an online bipartite matching

problem where, as usual, nodes from one side of the graph are given, while the others arrive one at the time, have degree exactly d , and are matched to one of their currently unmatched neighbors uniformly at random (or discarded if no such neighbor exists). Starting from the seminal work by [Karp et al., 1990], many versions of online matching problems in bipartite graphs have been studied (see [Devanur and Mehta, 2023, Huang and Tröbst, 2023] for recent surveys). To the best of our knowledge, such models mostly focus on the competitive ratio of global objective functions (such as the number of matches or some profit or cost function associated to the matching), while our node-by-node analysis of the probability of being matched appears to be new.

1.4 Notation

We denote by $\mathbb{N} = \{1, 2, 3, \dots\}$ and, for $k \in \mathbb{N}$, $[k] = \{1, 2, \dots, k\}$. We use $n \in \mathbb{N}$ for the number of schools, $d \in \mathbb{N}$ and $\ell \in \mathbb{N}$ for lengths of preference lists, and $m \in \mathbb{N}$ as the number of students (in cases where we do not assume infinite students). Superscripts denote properties that are fixed for the market (except in the continuous realm where we place the d of $x_d(t)$ in the subscript for convenience), and subscripts are used for running indices. We use $i \in \mathbb{N}$ for students, $j \in [n]$ for schools. Random variables are denoted by uppercase letters.

2 Models and Results

2.1 A Discrete Random Market Model

Model description. Consider a random market consisting of $n \in \mathbb{N}$ schools and infinitely many students indexed by $i = 1, 2, 3, \dots$, where each school has exactly one seat. Each student has a strict *preference list*, consisting of $d \leq n$ schools, chosen uniformly at random from the set of n schools. We call d the *preference list length*. If a occurs before b in this preference list, the student strictly prefers being matched to a rather than b . If a school c does not appear in a student's preference list, the student finds this school unacceptable and prefers remaining unmatched to being matched to that school. A school that appears in a student's preference list is *acceptable* to the student.

Students are matched to schools via a *Serial Dictatorship* mechanism as follows. Students are listed in some deterministic order; and each student in their turn picks their most preferred school that has an available seat. If the student finds none of the remaining schools acceptable, then that student goes unmatched. For a student $i \in [m]$, we denote by $K_i^{n,d}$ the random variable denoting the position of the school they get matched to within their preference list, setting $K_i^{n,d} = \infty$ if the student is unmatched (so $K_i^{n,d} \in \{\infty, 1, \dots, d\}$). We define the random vector $\mathcal{K} = \{K_1^{n,d}, K_2^{n,d}, K_3^{n,d}, \dots\}$ as the realization of one run of the mechanism.

A student i wishes to minimize $K_i^{n,d}$. That is, they want to be ranked to the most preferred school in their list. On the other hand, not getting matched at all ($K_i^{n,d} = \infty$) is by far worse than getting matched to a school that the student finds acceptable but ranks lower. Define $M_i^{n,d} = \mathbb{1}_{\{K_i^{n,d} < \infty\}}$ to be the indicator random variable for the event that student i gets matched at all. The probability that they get matched to any school is then $\mathbb{P}(M_i^{n,d} = 1)$.

For a given market with a fixed number of schools, only d is chosen by a centralized planner. It is therefore natural to ask what impact varying d has on the distributional properties of \mathcal{K} , and on the distribution of individual students' ranks ($K_i^{n,d}$), as well as their probabilities of getting matched to any school, $\mathbb{P}(M_i^{n,d} = 1)$.

Main results. As our main result⁴, we show that for a number n of schools large enough, every student in position $i \leq n$ will be matched to a school with higher probability in markets with longer lists. Formally, we have the following result.

THEOREM 2.1. *Let $d, \ell \in \mathbb{N}$ with $\ell \geq d$. For every n large enough and $i \leq n$, we have*

$$\mathbb{P}\left(M_i^{n,\ell} = 1\right) \geq \mathbb{P}\left(M_i^{n,d} = 1\right). \quad (1)$$

As we show in the following lemma, for each school, the probability of being matched increases with the length of preference lists. For a school $j \in [n]$, define $H_j^{n,d}(i)$ to be the indicator random variable for the event that school j gets matched to some student just before student i has had their turn. Then the following holds.

LEMMA 2.2. *Let $d, \ell, n \in \mathbb{N}$ with $d \leq \ell \leq n$. For every $j \in [n]$ and every $i = 1, 2, \dots$, we have:*

$$\mathbb{P}\left(H_j^{n,\ell}(i) = 1\right) \geq \mathbb{P}\left(H_j^{n,d}(i) = 1\right).$$

Theorem 2.1 and Lemma 2.2 imply that if the number of students does not exceed the number of schools, for n large enough all agents will be matched with higher probability in the market where preference lists are longer. Interestingly, Theorem 2.1 does not necessarily hold for i larger than n : as we discuss in Section 2.3, for any n large enough, students in position $i = \lceil 1.22n \rceil$ and beyond have a higher probability of being matched when preference lists are of length 1 as opposed to length 2. A similar phenomenon happens also for longer preference lists, as formalized by the next lemma.

LEMMA 2.3. *Let $d, \ell \in \mathbb{N}$ with $\ell \geq d$. For every n large enough, there exists $i > n$, such that*

$$\mathbb{P}\left(M_i^{n,\ell} = 1\right) \leq \mathbb{P}\left(M_i^{n,d} = 1\right).$$

While Theorem 2.1 allows us to compare probabilities of getting any school for students under two different list lengths, it does not give us any information on the absolute probability, $\mathbb{P}(M_i^{n,d} = 1)$, that a student i will get matched. This probability decreases with i and eventually converges to 0 as i becomes large enough (since all schools will have been taken), so a natural question is to ask how this probability behaves for smaller values of i . It turns out that as d increases, $\mathbb{P}(M_n^{n,d} = 1)$ quickly approaches $1/2$. This is significant for all students $i \leq n$ (since $\mathbb{P}(M_i^{n,d} = 1)$ decreases in i) and for all students when the number of students is comparable to the number of schools.

THEOREM 2.4. *Let $d \in \mathbb{N}$. For every n large enough, we have*

$$\frac{d}{2d+1} \leq \mathbb{P}\left(M_n^{n,d} = 1\right) \leq \frac{2d}{4d+1}. \quad (2)$$

In particular,

$$\lim_{d \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P}(M_n^{n,d} = 1) = \frac{1}{2}.$$

We additionally prove the following lemma that bounds the change in probability of a given student getting matched to one of their top- k schools for $k \leq d$.

LEMMA 2.5. *Let $d \in \mathbb{N}$. For every n large enough, $k \leq d$ and $i \leq n$, we have*

$$\mathbb{P}\left(K_i^{n,d} \leq k\right) - \mathbb{P}\left(K_i^{n,d+1} \leq k\right) \leq \left(\frac{d+2}{2d+3}\right)^{k/(d+1)} - \left(\frac{2d+1}{4d+1}\right)^{k/d}.$$

⁴All proofs of the results from this subsection appear in Section 3.3 and Section 3.4.

The dynamics of the discrete model. To analyze the discrete model, it is useful to study the number of students matched to any school just before it is student i 's turn, given by

$$T_i^{n,d} = \sum_{j=1}^{i-1} \mathbb{1}_{\{K_j^{n,d} < \infty\}}.$$

Since each school has exactly one seat, $T_i^{n,d}$ coincides with the number of schools matched to the students $\{1, \dots, i-1\}$. The number of schools that remain unmatched before student i 's turn is therefore $n - T_i^{n,d}$.

We remark on one property of the random market, called the *principle of deferred decisions*. This has proved useful in analyzing other markets, see, e.g., [Immorlica and Mahdian, 2015]. Observe that the preference list of student i does not play a role in the output until the moment the i -th round of Serial Dictatorship, when student i is selected to pick their favorite school. We therefore do not need to specify i 's preference list until this moment, allowing us to defer this decision until the list is needed. In particular, note that the distribution of $K_i^{n,d}$ depends only on $T_i^{n,d}$: we only need to know the number of schools that remain unmatched at the start of the turn to know the distribution of the position that the student gets matched to. This is further illustrated below.

Suppose it is the i -th student's turn, and the students prior to i have been matched to k schools, so $T_i^{n,d} = k$. It is then straightforward to compute the probability that a list of d schools chosen uniformly at random from the set of n schools will overlap with any of the unmatched $n - k$ schools. This probability is given by the following formula.

$$\mathbb{P}\left(M_i^{n,d} = 1 \mid T_i^{n,d} = k\right) = \begin{cases} 1 - \frac{\binom{k}{d}}{\binom{n}{d}}, & k \geq d, \\ 1, & \text{otherwise.} \end{cases} \quad (3)$$

It is immediately clear from this formula that the probability of getting matched depends only on n , d , and $T_i^{n,d}$. Furthermore, all else held constant, $\mathbb{P}\left(M_i^{n,d} = 1 \mid T_i^{n,d} = k\right)$ decreases as $T_i^{n,d}$ increases. The principle of deferred decisions allows us to connect our discrete model to a continuous one, discussed next.

2.2 A Continuous Market Model

We now introduce a continuous market model, which we will show is equivalent, to the limit (uniformly in probability) of the previously introduced discrete market model in the next section.

Description. Consider a (deterministic) continuous matching model consisting of a unit mass of schools and a continuum of students represented by the interval $[0, \infty)$ and indexed by $t \geq 0$. Let $d \geq 1$ be a parameter of the market, and define $x_d(t)$ to be the proportion of schools matched to students in $[0, t)$. The market is now defined by the Initial Value Problem

$$x_d(0) = 0, \quad x'_d(t) = 1 - x_d(t)^d. \quad (4)$$

The interpretation of the Initial Value Problem (4) is as follows: start with $x_d(0) = 0$ and traverse students in order from $t = 0$. At the time of student t , they get matched to schools at rate $1 - x_d(t)^d$, which is also the rate by which $x_d(t)$ increases at t . We note $x'_d(t)$ depends only on the proportion of schools matched until that point, and that $x_d(t)$ is smooth in its argument and takes values in $[0, 1]$.

Dynamics. The dynamics of the continuous model are described by the (conceptually) simple initial value problem in (4). Note that x_d is everywhere increasing, and $x_d(t) \rightarrow 1$ as $t \rightarrow \infty$, but $x'_d(t) \rightarrow 0$ as $n \rightarrow \infty$. Further, one immediately observes that for $t \approx 0$, $x'_d(t) \approx 1$: at first all

students are getting matched at the highest possible rate. The parameter d is a measure of intensity of matching: if the proportion of schools matched were fixed, smaller d would yield a slower rate of matching.

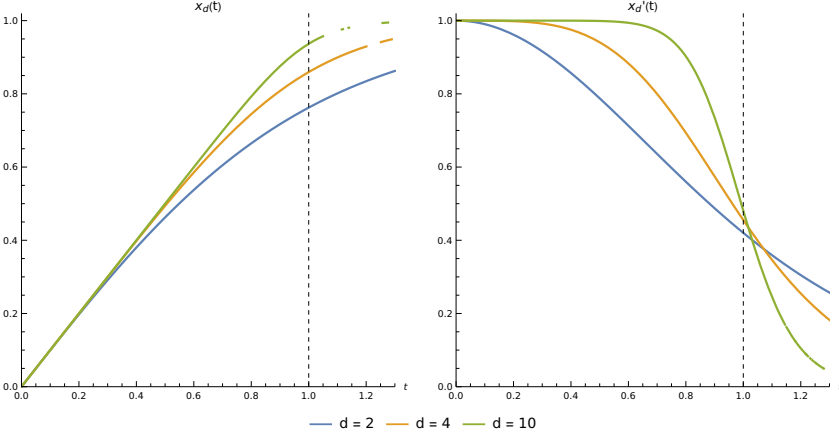


Fig. 1. x_d and x'_d for various values of d : note how the values for x'_d cross just after $t = 1$

Similarly to the discrete model (see the discussion in Section 1.1), we see that there are again two forces at play: for small d , the rate of getting matched is smaller, but simultaneously the proportion of schools taken up by earlier students is smaller, so later students may prefer small d as it gives them some schools to possibly get matched to. Indeed, for t exceeding 1, the probability of getting matched quickly vanishes as d increases. On the other hand, a large d increases the rate of getting matched, but also reduces the number of schools remaining by the time of student t 's turn.

We remark that although the description of the initial value problem might look deceptively straightforward, actually analyzing it proves to be quite challenging. The differential equation is highly non-linear, a special case of the *Chini type* differential equation studied at least since 1924 [Kamke, 1959], to which there is no known analytical solution for general d [Henk and Peyman, 2022]. To circumvent this, we must find implicit ways of proving properties about the solutions of the differential equation that do not require a full analytical solution.

2.3 From Continuous to Discrete

In this section we discuss the main theorem that connects the continuous and discrete markets, then state our main results within the continuous realm, and then describe how they carry over to the discrete case.

Connection to the discrete model. We now give an intuition on how the discrete market with fixed list length d approaches the continuous market when the number of schools $n \rightarrow \infty$. We make this limit rigorous via a functional law of large numbers in Theorem 2.6.

To see the connection, consider the discrete market described in Section 2.1 for a fixed list length d , and suppose we are at the turn of student $t = i/n$ for some fixed n . In Lemma 3.1, we show that

$$\mathbb{P}(M_i^{n,d} = 1) = 1 - (T_i^{n,d}/n)^d + O(d^2/n). \quad (5)$$

Here $T_i^{n,d}/n$ is the proportion of schools taken by students $\{1, \dots, i-1\}$ and so is an analogue of $x_d(t)$ (the proportion of schools taken by students $[0, t)$) in the continuous model. In the limit as

$n \rightarrow \infty$, the last term vanishes so when student $t = i/n$ has their turn, they get matched to a school with probability $x'_d(t) = 1 - x_d(t)^d$. Indeed, we show in Lemma 2.7 that $\mathbb{P}(M_i^{n,d} = 1) \rightarrow x'_d(i/n)$ in probability as $n \rightarrow \infty$. In the continuous model d is therefore the analogue of the list length (but is now allowed to be any real number in $[1, \infty)$).

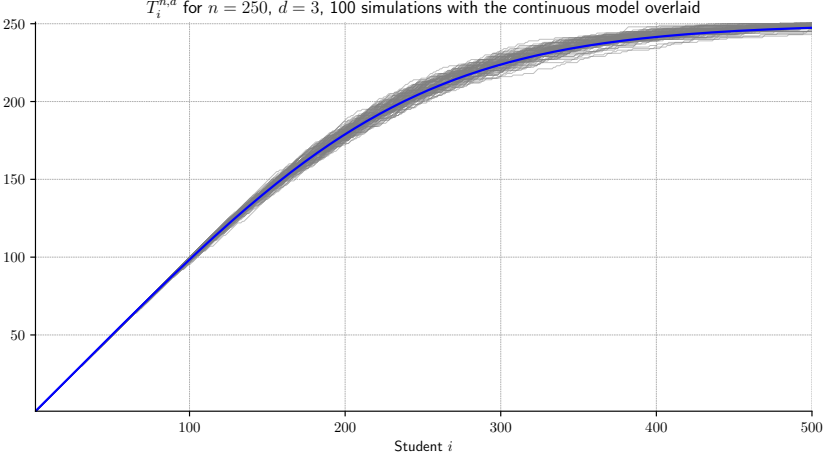


Fig. 2. 100 simulations of the discrete model with the continuous model overlaid.

The following theorem, proved in Section 3.1, describes rigorously the sense in which the continuous market is a limit of the discrete market.

THEOREM 2.6. Fix $d \in \mathbb{N}$ and define $p(x) = 1 - x^d$ for $x \in [0, 1]$. For $t \geq 0$, let $T_{\lfloor tn \rfloor}^{n,d}$ be the random variable denoting the number of schools matched by the first $\lfloor tn \rfloor$ students when there are n schools in the discrete random market with list length d . Then $n^{-1}T_{\lfloor tn \rfloor}^{n,d} \rightarrow x_d(t)$ uniformly in probability as $n \rightarrow \infty$, where $x_d(t)$ is the unique solution satisfying $x_d(0) = 0$ and $x'_d(t) = p(x_d(t))$ for $t \geq 0$. That is, for all $s \geq 0$ and $\varepsilon > 0$, as $n \rightarrow \infty$,

$$\mathbb{P} \left(\sup_{t \in [0, s]} \left| n^{-1}T_{\lfloor tn \rfloor}^{n,d} - x_d(t) \right| \geq \varepsilon \right) \rightarrow 0. \quad (6)$$

Moreover, as $n \rightarrow \infty$ for all $r \geq 1$, we have

$$\mathbb{E} \left(\left| n^{-1}T_{\lfloor tn \rfloor}^{n,d} - x_d(t) \right|^r \right) \rightarrow 0. \quad (7)$$

In particular, (6) implies that as $n \rightarrow \infty$, the number of schools taken prior to student $t = i/n$ converges in probability to the solution of the Initial Value Problem (4), i.e., the continuous market. Not only does the proportion of schools taken converge in probability for every point $t \geq 0$, but the maximum deviations of the random market around the continuous solution vanish in probability, and the mean converges to the continuous solution (by taking $r = 1$ in (7)). The theorem additionally posits that the initial value problem defining the continuous market, that is (4), has a unique solution satisfying the boundary data and that this solution extends for all time $t \geq 0$.

The next lemma rigorously connects the match rates of the discrete and continuous models in this limit. Recall that $M_i^{n,d}$ is the indicator random variable for whether student i gets matched to any school in the discrete random market with n schools and preference lists of length d .

LEMMA 2.7. For all $d \in \mathbb{N}$, $t \geq 0$, we have $\mathbb{P}\left(M_{[tn]}^{n,d} = 1\right) \rightarrow x'_d(t)$ as $n \rightarrow \infty$.

To understand whether $\mathbb{P}\left(M_i^{n,d} = 1\right)$ is larger in longer or shorter lists, we therefore need to understand whether $x'_d(t)$ is larger for larger or smaller d . That is, if $x'_{d+1}(t) \geq x'_d(t)$ for some $t \geq 0$, then student $t = i/n$ prefers lists of length $d + 1$ to those of length d for every large enough n . If we can show such a relationship holds for every $d \geq 1$ at t , then clearly student t always prefers longer lists to shorter ones. Our task of understanding match probability therefore becomes understanding $x'_d(t)$ in the continuous market.

The case of $d = 1$ and $d = 2$. For $d = 1$ and $d = 2$, we can in fact compute analytic solutions to the initial value problem (4) defining the continuous market. For the case of $d = 1$, we unsurprisingly get

$$x_1(t) = 1 - e^{-t}, \quad x'_1(t) = e^{-t},$$

and for $d = 2$ we get

$$x_2(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}}; \quad x'_2(t) = \frac{4e^{2t}}{(e^{2t} + 1)^2}.$$

The latter may also be written $x_2(t) = \tanh(t)$ and $x'_2(t) = \text{sech}(t)^2$ in terms of the hyperbolic functions. With these analytic expressions, one can solve to find $x'_2(t) \geq x'_1(t)$ if and only if $t \leq 1.219$. This means that for students before this cutoff, they prefer lists of length 2 to lists of length 1, and vice versa for larger t , students prefer the case of shorter lists: see Figure 3. We wish to highlight two key phenomena from this example: all students for $t \in [0, 1]$ have a higher probability of getting matched to any school with longer lists compared to shorter lists, and there is a cutoff (after $t = 1$) where this pattern reverses.

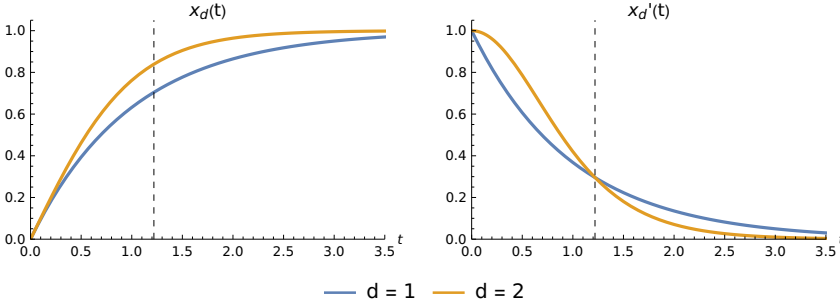


Fig. 3. x_d and x'_d for $d = 1$ and $d = 2$. Note that students before the highlighted cutoff at $t = 1.219$ have higher probability of being assigned to any school with longer lists, and vice versa for students after the cutoff.

The case of general d . In the case of general d , we cannot hope to solve the initial value problem analytically. Our main result in the continuous realm, proved in Section 3.2 is that all students for $t \leq 1$ always prefer longer lists.

THEOREM 2.8. For all $d, \ell \in [1, \infty)$ with $\ell \geq d$, and for all $t \in [0, 1]$, we have $x'_\ell(t) \geq x'_d(t)$.

2.4 Implications and Extensions

Random Serial Dictatorship. In the Random Serial Dictatorship mechanism, the order in which students propose is sampled from all possible permutations of students. When applied to our setting, it introduces an additional layer of randomness, on top of the one due to the preferences of agents.

Formally, consider a modification of our model, where there are n schools (each with one seat), $m \in \mathbb{N}$ students ordered randomly, and each of them has a strict preference list of d schools, chosen uniformly at random from the set of n schools. For $d, n \in \mathbb{N}$, we denote by $R^{n,d}$ the indicator random variable for the event that a student is matched to some school under the Serial Dictatorship algorithm, where students propose following the sampled order (note that this probability does not depend on the specific student, while it depends on m , but we are omitting this dependency for the sake of brevity). The following is an easy corollary of Lemma 2.2; its proof is given in Appendix C.

COROLLARY 2.9. *Let $d, \ell, n \in \mathbb{N}$ with $n \geq \ell \geq d$. We have*

$$\mathbb{P}(R^{n,\ell} = 1) \geq \mathbb{P}(R^{n,d} = 1).$$

Schools with multiple seats. In this section we discuss an extension of the discrete random market introduced in Section 2.1 to the case where schools each have $q \in \mathbb{N}$ seats. Like in the simpler market, we are interested in students $i = 1, 2, \dots$ as they each have their turn and are (possibly) matched to one of their remaining acceptable schools. Formally, denote by $M_i^{n,d,q}$ the indicator random variable of whether student i gets matched with n schools, d length lists, and q seats per school. Based on numerical experiments (some of which are reported in Section 5), we conjecture that the following extension of Theorem 2.1 holds.

CONJECTURE 2.10. *Let $d, \ell \in \mathbb{N}$ with $\ell \geq d$. For every n large enough and $i \leq nq$, we have*

$$\mathbb{P}(M_i^{n,\ell,q} = 1) \geq \mathbb{P}(M_i^{n,d,q} = 1).$$

We now sketch some ideas on how the techniques introduced in this paper could be used to attack Conjecture 2.10. More details are given in Appendix B.

We fix $q \in \mathbb{N}$ and do not include it in superscripts for brevity. Additionally, fix $n \in \mathbb{N}$ as the number of schools, and $d \geq 1$ as the list length. It is now not enough to keep track just of $T_i^{n,d}$, now defined to be the number of schools with all seats taken after students $\{1, 2, \dots, i-1\}$ have had their turn; instead, we must additionally keep track of the vector $S_i = (S_i^0, S_i^1, S_i^2, \dots, S_i^{q-1})$ counting the number of schools with $k = 0, 1, \dots, q-1$ seats taken at this point. (Note that $T_i^{n,d}$ is exactly the number of schools with q seats taken). Given this information, we can now see the market dynamics. When a student i has their turn, they pick d schools uniformly at random from the set of all schools. The probability that the student picks in this list any school that has any seats remaining (and thus the student gets matched) is still the same one given in (3), and is approximated by the same expression as Lemma 3.1. If the student picks a school with a remaining seat, it will have k seats taken with probability $S_i^k / \sum_{j=0}^{q-1} S_i^j$. The update rule for the counts is then as follows. If the school has, say k seats taken we then have $S_{i+1}^k = S_i^k - 1$, and if $k < q-1$ then $S_{i+1}^{k+1} = S_i^{k+1} + 1$ and if $k = q-1$ then $T_{i+1}^{n,d} = T_i^{n,d} + 1$.

In the case of $q = 1$ we observed interesting behavior near $i \approx n$. Then with $q > 1$ we would look for this at $i \approx nq$ due to the higher number of seats.

In Section 2.2 we introduced a continuous deterministic model with $q = 1$ and in Section 2.3 we showed how the discrete random model approaches the continuous model in probability as $n \rightarrow \infty$. Similarly, it is possible to construct a continuous model for arbitrary $q = 1, 2, 3, \dots$ and prove a similar limit result. We give an outline of this process in Appendix B and discuss strategies

for solving the resultant differential equations. The missing piece to prove Conjecture 2.10 is an analogous of Theorem 2.8. Again, we refer to Appendix B for details.

2.5 Discussion of the Discrete Model

Models where agents have preference lists that are sampled uniformly at random from the list of all permutations (possibly, as in our paper, of bounded size) form a very common theoretical assumption for the study of two-sided matching markets, see for example [Ashlagi et al., 2017, Immorlica and Mahdian, 2015, Pittel, 1989, 1992, 2019]. One of the reason for the popularity of these models is that they often lead to tractable expressions. Results can then be verified to hold numerically for models with correlated preferences, or this can motivate further analytical study of such more cumbersome markets. analytically. Our approach follows this general line of inquiry, with our main result (Theorem 2.1) proved for the case where lists are sampled uniformly at random, and verified numerically in Section 4 for preference lists sampled from more complex distributions.

The choice of Serial Dictatorship for the mechanism investigated in this paper is motivated by multiple facts. First, its relevance: it is broadly studied in theory (including its randomized version, see, e.g. [Bampis et al., 2023, Bogomolnaia and Moulin, 2001]) and applied in practice to many settings, most famously house allocation problems [Abdulkadiroğlu and Sönmez, 1999]. A Serial Dictatorship mechanism also naturally arises in the case of two-sided stable matching problems where schools have a shared (homogeneous) and strict preference list over the students which is, for instance, dictated by test scores, or by a lottery system where the lottery number determines priority [Arnosti, 2023]. In such cases, the Deferred Acceptance mechanism reduces to Serial Dictatorship, with students proceeding according to the shared order and picking their favorite remaining school. Both of these scenarios are common in school districts. For instance, in New York primary school matching the Department of Education has very limited information on students and so assigns priorities largely at random, and completely at random in cases such as city-wide gifted and talented schools (where the admissible students are restricted to those who are considered gifted or talented). On the other hand, the NYC specialized high schools are by law required to rank students uniquely by their score on the standardized Specialized High School Admissions Test (SHSAT) only [Faenza et al., 2023].

Another reason to study the Serial Dictatorship mechanism is that it endows students with an obvious order, allowing granular statements to be made about each student uniquely (such as in our case discussing outcomes based on order). In many other mechanisms where randomness comes into play, students become indistinguishable, and one can hope for at most aggregate level statements and results.

3 Proofs of Main Results

3.1 Connections Between the Discrete and Continuous Markets

In this section we prove the following theorem, which characterizes the behavior of the rate at which students get matched to schools as $n \rightarrow \infty$ via a solution to a differential equation.

THEOREM 2.6. *Fix $d \in \mathbb{N}$ and define $p(x) = 1 - x^d$ for $x \in [0, 1]$. For $t \geq 0$, let $T_{\lfloor tn \rfloor}^{n,d}$ be the random variable denoting the number of schools matched by the first $\lfloor tn \rfloor$ students when there are n schools in the discrete random market with list length d . Then $n^{-1}T_{\lfloor tn \rfloor}^{n,d} \rightarrow x_d(t)$ uniformly in probability as $n \rightarrow \infty$, where $x_d(t)$ is the unique solution satisfying $x_d(0) = 0$ and $x'_d(t) = p(x_d(t))$ for $t \geq 0$. That*

is, for all $s \geq 0$ and $\varepsilon > 0$, as $n \rightarrow \infty$,

$$\mathbb{P} \left(\sup_{t \in [0, s]} \left| n^{-1} T_{\lfloor tn \rfloor}^{n, d} - x_d(t) \right| \geq \varepsilon \right) \rightarrow 0. \quad (6)$$

Moreover, as $n \rightarrow \infty$ for all $r \geq 1$, we have

$$\mathbb{E} \left(\left| n^{-1} T_{\lfloor tn \rfloor}^{n, d} - x_d(t) \right|^r \right) \rightarrow 0. \quad (7)$$

We begin by computing an approximation to the probability of a student getting matched in the discrete case.

LEMMA 3.1 (PROBABILITY OF GETTING MATCHED). *In the discrete market with n schools, we have*

$$\left| \mathbb{P} \left(M_i^{n, d} = 1 \mid T_i^{n, d} = k \right) - \left(1 - (k/n)^d \right) \right| \leq d^2/n.$$

That is, when k schools have been matched prior to student i , then the probability that student i gets matched to any school is approximated by $1 - (k/n)^d$.

PROOF. Recall that when it is student i 's turn, they get matched to any school with probability given in (3), which for $k \geq d$ states $\mathbb{P} \left(M_i^{n, d} = 1 \mid T_i^{n, d} = k \right) = 1 - \binom{k}{d} / \binom{n}{d}$. This is the complement of the probability that a random list of d schools sampled from n without replacement will overlap entirely with the list of k taken schools. Consider now an alternative procedure where student i samples a list of d schools *with replacement*, where each school has a probability of k/n of being already taken by previous students. The probability of sampling a list with at least one school not taken by previous students conditioned on exactly k schools having been matched to students $1, \dots, i-1$ would be $1 - (k/n)^d$.

It remains to bound the difference between sampling the preference lists with replacement and sampling them without replacement. Observe that the outcomes differ exactly in the case that we choose at least one pair of schools that are the same while sampling with replacement. Otherwise the outcomes coincide. Formally, consider the probability space generated by the student sampling with replacement, assuming k schools are matched to students $1, \dots, i-1$. Let E be the event that student i gets matched to any school and let X be the event that the sample contains a repeated school. We have $\mathbb{P} \left(M_i^{n, d} = 1 \mid T_i^{n, d} = k \right) = \mathbb{P}(E \mid X^c)$, and $\mathbb{P}(E) = 1 - (k/n)^d$. Now with the law of total probability, write

$$\begin{aligned} |\mathbb{P}(E) - \mathbb{P}(E \mid X^c)| &= |\mathbb{P}(E \mid X) \mathbb{P}(X) + \mathbb{P}(E \mid X^c) (1 - \mathbb{P}(X)) - \mathbb{P}(E \mid X^c)| \\ &= \mathbb{P}(X) |\mathbb{P}(E \mid X) - \mathbb{P}(E \mid X^c)| \\ &\leq \mathbb{P}(X). \end{aligned}$$

Finally we bound $\mathbb{P}(X)$, the probability that a sample drawn with replacement contains duplicate items. Observe that for each of the $d(d-1)/2$ pairs of positions in the sample there is exactly a $1/n$ probability that that pair is the same. Using the union bound, this yields $\mathbb{P}(X) \leq \frac{d(d-1)}{2n} \leq d^2/n$. \square

Our aim is to show that the discrete market converges to the continuous one as n grows large. We assume here that reader is familiar with the theory of ODE and Markov processes. A gentle introduction to the relevant background can be found in Appendix A.

To connect the behavior of the discrete market with that of the continuous market, we apply the following theorem. See [Kroese et al., 2013, Theorem 17.3.1] for this version and [Ethier and Kurtz, 2009] for a proof.

THEOREM 3.2 (A FUNCTIONAL LAW OF LARGE NUMBERS). *Suppose $\{J_t^n\}_{n=1,2,\dots}$ is a family of continuous time Markov chains on finite state spaces where for $t \geq 0$, J_t^n takes values in $S^n \subseteq \mathbb{Z}$ and has transition rate matrix $Q^n = (q^n(i, j); i, j \in S^n)$. Suppose that there is a subset $\mathcal{D} \subseteq \mathbb{R}$ and a family $\{f^n\}_{n=1,2,\dots}$ of functions with $f^n : \mathcal{D} \times \mathbb{Z} \rightarrow \mathbb{R}$ which are bounded and continuous in the first argument such that, for each $i \in S^n$ and $k \in \mathbb{Z} \setminus \{0\}$ such that $i + k \in S^n$, we have*

$$q^n(i, i + k) = n f^n(i/n, k). \quad (8)$$

Define $g^n(x) = \sum_{k \in \mathbb{Z}} k f^n(x, k)$ for $x \in \mathcal{D}$, and suppose there exists a Lipschitz continuous function $g : \mathcal{D} \rightarrow \mathbb{R}$ such that the g^n converge uniformly to g on \mathcal{D} . Suppose $\lim_{n \rightarrow \infty} n^{-1} J_0^n = x_0$ for some x_0 . Then there exists a unique deterministic trajectory $x(t)$ satisfying $x(0) = x_0$ and $x'(t) = g(x(t))$, $x(t) \in \mathcal{D}$, $t \in [0, T]$, and $\{n^{-1} J_t^n\}_{t \geq 0}$ converges uniformly in probability on $[0, T]$ to $x(t)$.

We are now ready to prove Theorem 2.6.

PROOF OF THEOREM 2.6. As in the statement of the theorem, fix $d \geq 1$ and define $p(x) = 1 - x^d$ for $x \in [0, 1]$. For $t \geq 0$, define $X_t^n = T_{\lfloor nt \rfloor}^{n,d}$. Letting $\Delta = n^{-1}$ for convenience, note that $\{X_t^n\}_{t=0, \Delta, 2\Delta, \dots}$ is a discrete time Markov chain on the state space $S^n = \{0, 1, \dots, n\}$ with transition probabilities

$$\mathbb{P}(X_{t+\Delta}^n = j \mid X_t^n = i) = \begin{cases} p_i^n, & j = i + 1, \\ 1 - p_i^n, & j = i, \\ 0, & \text{otherwise.} \end{cases}$$

where $p_i^n = 1 - (i/n)^d \pm O(d^2/n)$ by Lemma 3.1.

Starting from $\{X_t^n\}_{t=0, \Delta, 2\Delta, \dots}$, we will now construct a sequence of continuous time Markov chains that satisfy the conditions of Theorem 3.2, then bring back the result to the discrete time chain.

Associate to each discrete time chain $\{X_t^n\}_{t=0, \Delta, 2\Delta, \dots}$, a coupled continuous time Markov chain $\{J_t^n\}_{t \geq 0}$ as follows: let H_t^n be a homogeneous Poisson process with rate $\lambda = n$, and let $\{J_t^n\}_{t \geq 0}$ be defined by $J_t^n = X_{\Delta H_t^n}^n$. This is a well-known embedding technique (see Theorem A.13 in Appendix A for a statement and discussion) and yields the transition rate matrix

$$q^n(i, j) = \begin{cases} n p_i^n, & j = i + 1, \\ -n p_i^n, & j = i, \\ 0, & \text{otherwise.} \end{cases}$$

Set $f^n(x, 1) = p_{xn}^n$ and $f^n(x, \cdot) = 0$ otherwise. We now verify that we can apply Theorem 3.2. First, let $S^n = \{0, 1, 2, \dots, n\}$ and $\mathcal{D} = [0, 1]$. Note q^n, f^n satisfy (8): we only need to check $k = 1$ in (8), which we verify as

$$\begin{aligned} q^n(i, i + 1) &= n p_i^n \\ &= n f^n(i/n, 1). \end{aligned}$$

For each n , f^n is clearly bounded and continuous in the first argument. We moreover have for all $x \in [0, 1]$

$$\begin{aligned} g^n(x) &= \sum_{k \in \mathbb{Z}} k f^n(x, k) \\ &= p_{xn}^n \\ &= p(x) \pm O(d^2/n), \end{aligned} \quad (\text{by Lemma 3.1})$$

hence $g^n \rightarrow p$ uniformly, since $|g^n(x) - p(x)| = O(d^2/n)$ is independent of x . Observe that p is Lipschitz-continuous with Lipschitz-constant d . Since $J_0^n = 0$, we have $n^{-1} J_0^n = 0$.

We can therefore apply Theorem 3.2 and deduce that $\{n^{-1} J_t^n\}_{t \geq 0}$ converges uniformly in probability to the unique solution $x_d(t)$ of the initial value problem (4).

To carry over convergence uniformly in probability to the discrete Markov chain $\{X_t^n\}_{t=0,\Delta,2\Delta,\dots}$, fix some $s \geq 0$. Recall that, for $\tau \geq 0$, H_t^n is the count of events in the underlying Poisson process up to time τ . Since every event of H_t^n corresponds to either a unit jump in $\{X_t^n\}_{t=0,\Delta,2\Delta,\dots}$ (that is, $X_{t+\Delta}^n = X_t^n + 1$) or no transition (that is, $X_{t+\Delta}^n = X_t^n$), we have for each $t \geq 0$.

$$\begin{aligned} |X_t^n - J_t^n| &= |X_t^n - X_{\Delta H_t^n}^n| \\ &\leq |\lfloor nt \rfloor - H_t^n| \\ &\leq 1 + |nt - H_t^n|. \end{aligned}$$

This yields

$$\begin{aligned} \sup_{t \in [0,s]} |n^{-1}X_t^n - x_d(t)| &\leq \sup_{t \in [0,s]} |n^{-1}X_t^n - n^{-1}J_t^n| + \sup_{t \in [0,s]} |n^{-1}J_t^n - x_d(t)| \\ &\leq \frac{1}{n} + \sup_{t \in [0,s]} |t - n^{-1}H_t^n| + \sup_{t \in [0,s]} |n^{-1}J_t^n - x_d(t)|. \end{aligned} \quad (9)$$

Note that $n^{-1}H_t^n - t$ is a martingale⁵, and applying Doob's martingale inequality (see Theorem A.15 of Appendix A) we have

$$\begin{aligned} \mathbb{P} \left(\sup_{t \in [0,s]} |n^{-1}H_t^n - t| \geq \varepsilon \right) &\leq \varepsilon^{-2} \mathbb{E}((n^{-1}H_s^n - s)^2) \\ &= \frac{s}{\varepsilon^2 n}. \end{aligned} \quad (10)$$

To see that the latter inequality holds, observe $\mathbb{E}(n^{-1}H_s^n) = s$, so $\mathbb{E}((n^{-1}H_s^n - s)^2)$ equals the variance of $n^{-1}H_s^n$. Since H_s^n has a Poisson distribution with rate ns , (10) follows.

Note that the second (by (10)) and third (by Theorem 3.2) terms of the right-hand side of (9) vanish in probability as $n \rightarrow \infty$. So we must have that $n^{-1}T_{\lfloor tn \rfloor}^{n,d} \rightarrow x_d(t)$ uniformly in probability as $n \rightarrow \infty$. The convergence in r -mean follows since $n^{-1}T_{\lfloor tn \rfloor}^{n,d}$ is bounded, see Lemma A.7 in Appendix A. \square

We connect the solution to preferences of students via the following lemma.

LEMMA 2.7. *For all $d \in \mathbb{N}$, $t \geq 0$, we have $\mathbb{P}(M_{\lfloor tn \rfloor}^{n,d} = 1) \rightarrow x'_d(t)$ as $n \rightarrow \infty$.*

PROOF. Fix $d \in \mathbb{N}$, $t, \varepsilon > 0$. We will show that there exists $N \in \mathbb{N}$ such that for $n > N$,

$$\left| \mathbb{P}(M_{\lfloor tn \rfloor}^{n,d} = 1) - x'_d(t) \right| < \varepsilon.$$

To proceed, observe that Lemma 3.1 implies that there exists N_1 be such that for $n > N_1$ and all $x \in [0, 1]$,

$$\left| \mathbb{P}(M_{\lfloor tn \rfloor}^{n,d} = 1 \mid n^{-1}T_{\lfloor tn \rfloor}^{n,d} = x) - (1 - x^d) \right| \leq \frac{\varepsilon}{4}.$$

⁵For $s \geq t$, using the independence of increments property and the fact that $\mathbb{E}(H_{s-t}^n) = n(s-t)$, we have

$$\mathbb{E}(n^{-1}H_s^n - s \mid H_t^n) - (n^{-1}H_t^n - t) = \mathbb{E}(n^{-1}(H_s^n - H_t^n) \mid H_t^n) - (s - t) = \mathbb{E}(n^{-1}(H_{s-t}^n)) - (s - t) = 0.$$

Next, note that $1 - x^d$ is continuous in x , so there exists some $\delta > 0$ such that for all y with $|y - x| \leq \delta$, $|(1 - x^d) - (1 - y^d)| < \varepsilon/4$. Combining these two facts, we have that

$$\left| \mathbb{P} \left(M_{[tn]}^{n,d} = 1 \mid n^{-1}T_{[tn]}^{n,d} = y, |y - x| \leq \delta \right) - (1 - x^d) \right| \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}.$$

By Theorem 2.6, there exists now $N_2 \in \mathbb{N}$ such that for $n > N_2$,

$$\mathbb{P} \left(\left| n^{-1}T_{[tn]}^{n,d} - x_d(t) \right| > \delta \right) < \frac{\varepsilon}{2}.$$

Let $N = \max \{N_1, N_2\}$, then putting all these together with the law of total probability, we have for $n > N$

$$\begin{aligned} \left| \mathbb{P} \left(M_{[tn]}^{n,d} = 1 \right) - x'_d(t) \right| &= \left| \mathbb{P} \left(M_{[tn]}^{n,d} = 1 \mid \left| n^{-1}T_{[tn]}^{n,d} - x_d(t) \right| \leq \delta \right) \mathbb{P} \left(\left| n^{-1}T_{[tn]}^{n,d} - x_d(t) \right| \leq \delta \right) \right. \\ &\quad \left. + \mathbb{P} \left(M_{[tn]}^{n,d} = 1 \mid \left| n^{-1}T_{[tn]}^{n,d} - x_d(t) \right| > \delta \right) \mathbb{P} \left(\left| n^{-1}T_{[tn]}^{n,d} - x_d(t) \right| > \delta \right) - x'_d(t) \right| \\ &\leq \left| \mathbb{P} \left(M_{[tn]}^{n,d} = 1 \mid n^{-1}T_{[tn]}^{n,d} = y, |y - x_d(t)| \leq \delta \right) + \mathbb{P} \left(\left| n^{-1}T_{[tn]}^{n,d} - x_d(t) \right| > \delta \right) - x'_d(t) \right| \\ &< \left| 1 - x_d(t)^d + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} - x'_d(t) \right| \\ &= \varepsilon. \end{aligned}$$

□

3.2 Continuous Market

In this section we prove that the match rate in the continuous market increases with d for all students $t \in [0, 1]$.

THEOREM 2.8. *For all $d, \ell \in [1, \infty)$ with $\ell \geq d$, and for all $t \in [0, 1]$, we have $x'_\ell(t) \geq x'_d(t)$.*

PROOF. We use the notation $x(d, t)$ and simplify to x when it is clear from the context. We wish to show that for all $1 \leq d < \ell$

$$\frac{\partial}{\partial t} x(d, t) \leq \frac{\partial}{\partial t} x(\ell, t),$$

which follows if for all $d \in [1, \infty)$ and $t \in [0, 1]$ we have

$$\frac{\partial^2}{\partial d \partial t} x(d, t) \geq 0.$$

Recall that the ODE is defined by $x(d, 0) = 0$, $\frac{\partial}{\partial t} x(d, t) = 1 - x^d$. Note that x is twice continuously differentiable on its domain so its partial derivatives commute. We claim that $x(d, t)$ is a solution to this differential equation if and only if it solves the integral equation

$$t = \int_0^{x(d,t)} \frac{1}{1 - u^d} du. \quad (11)$$

To observe this, note that $x(d, 0) = 0$, and by implicitly differentiating (11) with respect to t , we recover the required condition on the derivative of x with respect to t :

$$\frac{\partial x}{\partial t} = 1 - x(d, t)^d.$$

Now we can use the Leibniz integral rule to implicitly differentiate (11) with respect to d and obtain

$$\begin{aligned} 0 &= \frac{\partial}{\partial d} \left(\int_0^{x(d,t)} \frac{1}{1-u^d} du \right) \\ &= \int_0^{x(d,t)} \frac{\partial}{\partial d} \left(\frac{1}{1-u^d} \right) du + \frac{1}{1-x(d,t)^d} \frac{\partial x}{\partial d}. \end{aligned}$$

Rearranging for $\frac{\partial x}{\partial d}$ and computing the derivative in the integrand yields

$$\begin{aligned} \frac{\partial x}{\partial d} &= -(1-x(d,t)^d) \int_0^{x(d,t)} \frac{u^d \log u}{(1-u^d)^2} du. \\ &= -\frac{\partial x}{\partial t} \int_0^{x(d,t)} \frac{u^d \log u}{(1-u^d)^2} du. \end{aligned} \tag{12}$$

Note that $\frac{\partial x}{\partial t}$ is positive, and the integrand is negative (since $0 \leq u \leq x \leq 1$, so $\log u \leq 0$). This means $\frac{\partial x}{\partial d} \geq 0$ (intuitively, fixing t and increasing d will increase the number of schools taken). Next compute

$$\begin{aligned} \frac{\partial^2 x}{\partial d \partial t} &= \frac{\partial}{\partial d} (1-x(d,t)^d) \\ &= -x^d \log x - dx^{d-1} \frac{\partial x}{\partial d} \\ &= -x^{d-1} \left(x \log x + d \frac{\partial x}{\partial d} \right). \end{aligned}$$

We want $\frac{\partial^2 x}{\partial d \partial t} \geq 0$, which means we need $x \log x + d \frac{\partial x}{\partial d} \leq 0$. That is,

$$\begin{aligned} x \log x - d(1-x^d) \int_0^x \frac{u^d \log u}{(1-u^d)^2} du &\leq 0 \\ \iff \frac{x \log x}{1-x^d} - \int_0^x \frac{du^d \log u}{(1-u^d)^2} du &\leq 0. \end{aligned}$$

Expressing both terms as integrals, we get

$$\begin{aligned} \frac{x \log x}{1-x^d} - \int_0^x \frac{du^d \log u}{(1-u^d)^2} du &= \int_0^x \left(\frac{1}{1-u^d} + \frac{du^d \log u}{(1-u^d)^2} + \frac{\log u}{1-u^d} \right) du - \int_0^x \frac{du^d \log u}{(1-u^d)^2} du \\ &= \int_0^x \frac{1 + \log u}{1-u^d} du. \end{aligned}$$

The rest of the proof is completed in Theorem C.1 of Appendix C where we show that this integral is non-positive for $x(d, t)$ with $t \leq 1$. \square

3.3 Discrete Market: Probability of Being Matched

In this section we prove the main results in the discrete market, often bringing back results from the continuous market via an application of Lemma 2.7.

The following theorem succinctly states that all students in a balanced market (where $i \leq n$) prefer longer lists for n large.

THEOREM 2.1. *Let $d, \ell \in \mathbb{N}$ with $\ell \geq d$. For every n large enough and $i \leq n$, we have*

$$\mathbb{P}\left(M_i^{n,\ell} = 1\right) \geq \mathbb{P}\left(M_i^{n,d} = 1\right). \quad (1)$$

PROOF. This follows directly from Theorem 2.8 and Lemma 2.7. \square

LEMMA 2.2. *Let $d, \ell, n \in \mathbb{N}$ with $d \leq \ell \leq n$. For every $j \in [n]$ and every $i = 1, 2, \dots$, we have:*

$$\mathbb{P}\left(H_j^{n,\ell}(i) = 1\right) \geq \mathbb{P}\left(H_j^{n,d}(i) = 1\right).$$

PROOF. We have

$$\begin{aligned} \mathbb{P}\left(H_j^{n,d}(i) = 1\right) &= \frac{1}{n} \sum_{r=1}^n \mathbb{P}\left(H_r^{n,d}(i) = 1\right) && \text{(by symmetry)} \\ &= \frac{1}{n} \sum_{r=1}^n \mathbb{E}(H_r^{n,d}(i)) \\ &= \frac{1}{n} \mathbb{E}\left(\sum_{r=1}^n H_r^{n,d}(i)\right) \\ &= \frac{1}{n} \mathbb{E}\left(T_i^{n,d}\right). \end{aligned}$$

Indeed, the random variable $\sum_{r=1}^n H_r^{n,d}(i)$ counts the number of schools matched to a student in positions $\{1, 2, \dots, i-1\}$ in the discrete market with n schools, where every student has a preference list of length d . This equals the number of students from $\{1, 2, \dots, i-1\}$ that are matched in that market, which is precisely $T_i^{n,d}$.

Recall from (3) that

$$\mathbb{P}\left(M_i^{n,d} = 1 \mid T_i^{n,d} = k\right) = \begin{cases} 1 - \frac{\binom{k}{d}}{\binom{n}{d}}, & k \geq d, \\ 1, & \text{otherwise.} \end{cases}$$

It is clear from this expression that the left hand side monotonically increases in d : for all $i = 1, 2, 3, \dots$, a longer list length increases the probability that student i is matched to any school. Since $T_i^{n,d} = \sum_{j=1}^{i-1} M_j^{n,d}$ we deduce that $T_i^{n,\ell} \geq T_i^{n,d}$ in the sense of stochastic dominance, and therefore in expectation. \square

We next state a lemma that we prove in Appendix C.

LEMMA 3.3. *For $d \geq 1$, we have $\left(\frac{2d+1}{4d+1}\right)^{1/d} \leq x(d, 1) \leq \left(\frac{d+1}{2d+1}\right)^{1/d}$.*

LEMMA 2.3. *Let $d, \ell \in \mathbb{N}$ with $\ell \geq d$. For every n large enough, there exists $i > n$, such that*

$$\mathbb{P}\left(M_i^{n,\ell} = 1\right) \leq \mathbb{P}\left(M_i^{n,d} = 1\right).$$

PROOF. Consider $y(t) = x_\ell(t) - x_d(t)$, it is clear that $\lim_{t \rightarrow \infty} y(t) = 0$. By Lemma 3.3, it is also clear that $y(1) > 0$. It therefore follows that there exists some $t \geq 1$ such that $y'(t) < 0$. The thesis then follows from Lemma 2.7. \square

We finally prove a corollary of Lemma 3.3 that is an intriguing and surprising result.

THEOREM 2.4. *Let $d \in \mathbb{N}$. For every n large enough, we have*

$$\frac{d}{2d+1} \leq \mathbb{P}(M_n^{n,d} = 1) \leq \frac{2d}{4d+1}. \quad (2)$$

In particular,

$$\lim_{d \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P}(M_n^{n,d} = 1) = \frac{1}{2}.$$

PROOF. From Lemma 2.7 and (4), we have $\mathbb{P}(M_n^{n,d} = 1) \rightarrow x'_d(1) = 1 - x_d(1)^d$. The claim then follows from the bounds in Lemma 3.3. \square

3.4 Discrete Market: Impact on Rank

In this section, we prove the following lemma.

LEMMA 2.5. *Let $d \in \mathbb{N}$. For every n large enough, $k \leq d$ and $i \leq n$, we have*

$$\mathbb{P}(K_i^{n,d} \leq k) - \mathbb{P}(K_i^{n,d+1} \leq k) \leq \left(\frac{d+2}{2d+3}\right)^{k/(d+1)} - \left(\frac{2d+1}{4d+1}\right)^{k/d}.$$

PROOF. Fix $d \in \mathbb{N}$ and $k \leq d$. Similar to Lemma 3.1, it is easy to show that probability of getting matched to the first k schools is given by

$$\mathbb{P}(K_i^{n,d} \leq k \mid T_i^{n,d} = k) = 1 - (k/n)^k \pm O(d^2/n).$$

By Theorem 2.6, we have that $n^{-1}T_i^{n,d} \xrightarrow{\mathcal{P}} x_d(i/n)$, so

$$\mathbb{P}(K_i^{n,d} \leq k) - \mathbb{P}(K_i^{n,d+1} \leq k) \xrightarrow{\mathcal{P}} x_{d+1}(i/n)^k - x_d(i/n)^k.$$

Consider the function $x_{d+1}(t)^k - x_d(t)^k$ for $t \in [0, 1]$, and observe that

$$\begin{aligned} \frac{\partial}{\partial t} (x_{d+1}(t)^k - x_d(t)^k) &= k(x_{d+1}(t)^{k-1}x'_{d+1}(t) - x_d(t)^{k-1}x'_d(t)) \\ &\geq k(x_d(t)^{k-1}x'_d(t) - x_d(t)^{k-1}x'_d(t)) \\ &= 0. \end{aligned}$$

The second last line follows from Theorem 2.8 since $x'_{d+1}(t) \geq x'_d(t)$ and $x_{d+1}(t) \geq x_d(t)$ (from the former since $x_{d+1}(0) - x_d(0) = 0$).

We therefore have that $x_{d+1}(i/n)^k - x_d(i/n)^k$ is increasing in i , so applying Lemma 3.3, we have

$$\begin{aligned} \max_{i \leq n} (x_{d+1}(i/n)^k - x_d(i/n)^k) &= x_{d+1}(1)^k - x_d(1)^k \\ &\leq \left(\frac{d+2}{2d+3}\right)^{k/(d+1)} - \left(\frac{2d+1}{4d+1}\right)^{k/d}, \end{aligned}$$

which completes the proof. \square

4 Numerical Experiments

We perform numerical experiments to evaluate how well our results generalize beyond the one-to-one Serial Dictatorship mechanism with uniformly random preferences.

Schools not being sampled uniformly at random. In our first set of numerical experiments we investigate the impact of students not picking their preference list uniformly at random from all schools. In particular, we now assume there is some probability distribution $p : \{1, \dots, n\} \rightarrow [0, 1]$ with $\sum_{j=1}^n p(j) = 1$, with $p(j)$ dictating the probability that students sample school j at their turn. The case of uniform sampling occurs when $p \equiv 1/n$.

We experiment with five different distributions for p : a uniform distribution (p_1), two types of Pareto-like distributions where students' preference of schools is concentrated at certain schools (p_2 and p_3 with low, and high concentration respectively), a setting of two classes (high demand and low demand) of schools (p_4), and finally a “degenerate” distribution where students sample from the first half of schools almost exclusively (p_5). Note that we would expect the last distribution to behave as if we had half as many schools, so that only students with $i \leq n/2$ would be guaranteed to prefer longer lists. The distributions are as follows:

$$p_1(j) = \frac{1}{n}, \quad p_2(j) \propto \frac{1}{2 + j/n}, \quad p_3(j) \propto \frac{1}{(1 + j/n)^{10}},$$

$$p_4(j) \propto 1 + 3\mathbb{1}_{\{j \leq 200\}}, \quad p_5(j) = \begin{cases} 2/n - 1/100, & j \leq 500, \\ 1/100, & j > 500. \end{cases}$$

We perform all numerics with $n = 1000$ schools, and compare the cases of d being 1, 2, 4, 10, and 20. All experiments are done with 100 000 repetitions.

Figure 5 in Appendix D shows the outcome of these experiments. We graph first the distribution of $p(\cdot)$ for all schools on the left, then the proportion of schools taken by students up to student i in the middle, and the probability that a given student i is matched to a school on the right.

Our main thesis – that all students in balanced markets prefer long lists – holds under all distributions other than the degenerate one (which is almost equivalent to having half as many schools). In particular, $\mathbb{P}(M_i^{n,d} = 1)$ increases in d under all distributions other than the degenerate. We conclude that our main result, Theorem 2.1, seems to be robust to i.i.d. sampling via reasonable non-uniform distributions.

Schools with multiple seats. In Section 2.4 and Appendix B we discuss the extension of the discrete and continuous models to the case when schools have $q \in \mathbb{N}$ seats. We now perform some numerical experiments to verify Conjecture 2.10. To do so, we compute numerical solutions to the differential equations describing the continuous model using the Euler method with step size 5×10^{-5} , and compare the probability of getting matched for successive values of $d \in \mathbb{N}$. We verified that the conjecture holds for all pairs of $q = 1, 2, \dots, 20$ and $d = 1, 2, \dots, 15$. This leads us to believe that the conjecture holds for all $q \in \mathbb{N}$ and $d \geq 1$. In Appendix D we show some plots of the resultant output.

5 Conclusions

In this paper, we investigated the impact of truncating preference lists in a two-sided matching market where students choose schools following a Serial Dictatorship order. Our main result is that if the market is balanced, all agents increase their probability of being matched when lists are longer. This result is proved for preference lists that are sampled uniformly at random and school quotas equal to 1 and then shown to hold numerically for more complex preferences and larger quotas. We believe these valuable insights can be used to support the expansion of the length of preference lists, which are often set to small values. Investigating similar questions for other mechanisms is a relevant open question.

Acknowledgments. The authors would like to acknowledge the support of NSF Grant 2046146 and the valuable feedback provided by three anonymous EC'25 reviewers.

References

- Atila Abdulkadiroğlu and Tayfun Sönmez. 1998. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica* 66, 3 (1998), 689–701.
- Atila Abdulkadiroğlu and Tayfun Sönmez. 1999. House allocation with existing tenants. *Journal of Economic Theory* 88, 2 (1999), 233–260.
- Atila Abdulkadiroğlu and Tayfun Sönmez. 2003. School choice: A mechanism design approach. *American economic review* 93, 3 (2003), 729–747.
- Nick Arnosti. 2022. A continuum model of stable matching with finite capacities. *arXiv preprint arXiv:2205.12881* (2022).
- Nick Arnosti. 2023. Lottery design for school choice. *Management Science* 69, 1 (2023), 244–259.
- Georgy Artemov, Yeon-Koo Che, and Yinghua He. 2017. Strategic ‘mistakes’: Implications for market design research. *NBER working paper* (2017).
- Itai Ashlagi, Andrei Graur, Irene Lo, and Kaleigh Mentzer. 2023. Overbooking with Priority-Respecting Reassignment. *Presentation at the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO'23)* (2023).
- Itai Ashlagi, Yash Kanoria, and Jacob D Leshno. 2017. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy* 125, 1 (2017), 69–98.
- Eduardo M Azevedo and Jacob D Leshno. 2016. A supply and demand framework for two-sided matching markets. *Journal of Political Economy* 124, 5 (2016), 1235–1268.
- Sophie Bade. 2020. Random serial dictatorship: the one and only. *Mathematics of Operations Research* 45, 1 (2020), 353–368.
- Evripidis Bampis, Bruno Escoffier, and Paul Youssef. 2023. Online 2-stage stable matching. *Discrete Applied Mathematics* 341 (2023), 394–405.
- Inácio Bó and Rustamdjan Hakimov. 2020. Iterative versus standard deferred acceptance: Experimental evidence. *The Economic Journal* 130, 626 (2020), 356–392.
- Anna Bogomolnaia and Hervé Moulin. 2001. A new solution to the random assignment problem. *Journal of Economic theory* 100, 2 (2001), 295–328.
- Caterina Calsamiglia, Guillaume Haeringer, and Flip Klijn. 2010. Constrained school choice: An experimental study. *American Economic Review* 100, 4 (2010), 1860–1874.
- Nikhil Devanur and Aranyak Mehta. 2023. Online matching in advertisement auctions. In *Online and Matching-based Market Design*, Nicole Immorlica, Federico Echenique, and Vijay V Vazirani (Eds.). Cambridge University Press.
- NYC DOE. 2023. *The Pre-K application is now open!* https://groups.io/g/NYCEducationnews/topic/fw_the_pre_k_application_is/96547480
- Lars H Ehlers and Bettina Klaus. 2023. Normative properties for object allocation problems: Characterizations and trade-offs. In *Online and Matching-based Market Design*, Nicole Immorlica, Federico Echenique, and Vijay V Vazirani (Eds.). Cambridge University Press.
- Stewart N Ethier and Thomas G Kurtz. 2009. *Markov processes: characterization and convergence*. John Wiley & Sons.
- Yuri Faenza, Ayoub Foussoul, and Chengyue He. 2024. Two-Stage Stochastic Stable Matching. In *International Conference on Integer Programming and Combinatorial Optimization*. Springer, 154–167.
- Yuri Faenza, Ayoub Foussoul, and Chengyue He. 2025. Minimum Cut Representability of Stable Matching Problems. *arXiv preprint arXiv:2504.04577* (2025).
- Yuri Faenza, Swati Gupta, and Xuan Zhang. 2023. Discovering Opportunities in New York City’s Discovery Program: Disadvantaged Students in Highly Competitive Markets. In *EC'23: Proceedings of the 24th ACM Conference on Economics and Computation*. ACM.
- David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15.
- Geoffrey Grimmett and David Stirzaker. 2020. *Probability and random processes*. Oxford university press.
- Koppelaar Henk and Nasehpour Peyman. 2022. Series Solution of High Order Abel, Bernoulli, Chini and Riccati Equations. *Kyungpook Mathematical Journal* 62, 4 (2022), 729–736.
- Zhiyi Huang and Thorben Tröbst. 2023. Applications of Online Matching. (2023).
- Nicole Immorlica and Mohammad Mahdian. 2015. Incentives in large random two-sided markets. *ACM Transactions on Economics and Computation (TEAC)* 3, 3 (2015), 1–25.
- Erich Kamke. 1959. *Differentialgleichungen lösungsmethoden und lösungen*. Chelsea Publishing Company.
- Yash Kanoria, Seungki Min, and Pengyu Qian. 2021. In which matching markets does the short side enjoy an advantage?. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 1374–1386.

- Richard M Karp, Umesh V Vazirani, and Vijay V Vazirani. 1990. An optimal algorithm for on-line bipartite matching. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*. 352–358.
- Dirk P Kroese, Thomas Taimre, and Zdravko I Botev. 2013. *Handbook of Monte Carlo methods*. John Wiley & Sons.
- Shuichi Miyazaki and Kazuya Okamoto. 2019. Jointly stable matchings. *Journal of Combinatorial Optimization* 38, 2 (2019), 646–665.
- National Resident Matching Program NRMP. 2023. National Resident Matching Program, Results and Data: 2023 Main Residency Match. (2023).
- Boris Pittel. 1989. The average number of stable matchings. *SIAM Journal on Discrete Mathematics* 2, 4 (1989), 530–549.
- Boris Pittel. 1992. On likely solutions of a stable marriage problem. *The Annals of Applied Probability* 2, 2 (1992), 358–401.
- Boris Pittel. 2019. On likely solutions of the stable matching problem with unequal numbers of men and women. *Mathematics of Operations Research* 44, 1 (2019), 122–146.
- Gerald Teschl. 2024. *Ordinary differential equations and dynamical systems*. Vol. 140. American Mathematical Society.

A Background on Differential Equations and Stochastic Processes

This appendix contains some preliminaries that are required to follow the proofs in the main text, in particular on the theory of ordinary differential equations (ODEs) and initial value problems (IVPs), as well as probability and Markov theory.

Ordinary Differential Equations

An ordinary differential equation is a description of the local dynamics of a system in terms of its current state and derivatives; in particular, a first order ordinary differential equation describes the rate of change of a quantity x as a function of the current time t and the current state $x(t)$. That is, $x'(t) = f(t, x(t))$ for some f . A problem with such a description along with boundary data (e.g., $x(t_0) = x_0$) is called an *initial value problem*. A main theme in differential equations is to understand when such local descriptions give rise to solutions, either in some small neighborhood of a point, or globally. Lipschitz continuity is an important property in differential equations, as it is a sufficient condition for the existence of such solutions as we will argue next.

Definition A.1 (Lipschitz continuity). A function $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is *Lipschitz continuous* under norm $\|\cdot\|$ on $U \subseteq \mathbb{R}^m$ if there exists some $L \geq 0$ such that

$$\|f(x) - f(y)\| \leq L \|x - y\|,$$

for all $x, y \in U$. We call L the Lipschitz constant of f on U .

The standard result for the existence and uniqueness of solutions to initial value problems of ordinary differential equations comes from variations of the Picard-Lindelöf Theorem, which roughly state that if f is Lipschitz at a point, a solution exists and is unique in some neighborhood of that point. The following is a version that extends this fact to the positive reals for the one dimensional case, adapted from [Teschl, 2024, Theorem 2.2 and Corollary 2.6].

THEOREM A.2. *Consider the initial value problem*

$$x' = f(t, x), \quad x(t_0) = x_0, \tag{13}$$

where $(t_0, x_0) \in U$ for some open set $U \subseteq \mathbb{R}^2$ and $f : U \rightarrow \mathbb{R}$ is a Lipschitz continuous function. Then if $[t_0, \infty) \times \mathbb{R} \in U$ there exists a unique solution $x(t)$ to (13) for all $t \geq t_0$.

Probability and Markov theory

In this section we briefly cover some fundamentals of general probability and Markov theory. The treatment is not fully formal as we omit certain regularity conditions and assumptions – however such conditions are immediately satisfied by the objects we study in this paper. We refer to [Grimmett and Stirzaker, 2020] for further details.

Recall that a random variable X on a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ is a measurable real-valued function from (Ω, \mathcal{F}) to $(\mathbb{R}, \mathcal{L})$.

Definition A.3 (Stochastic process). A stochastic process on a state space $\mathcal{S} \subseteq \mathbb{R}$ is a set of random variables $\{X_t\}_{t \in \mathcal{T}}$ on the same probability triple that are indexed by some \mathcal{T} (thought of as time) and each take values in \mathcal{S} (that is, $X_t \in \mathcal{S}$ for all $t \in \mathcal{T}$).

Common cases are $\mathcal{T} = \{0, 1, 2, 3, \dots\}$ or $\mathcal{T} = \{0, \Delta, 2\Delta, 3\Delta, \dots\}$ for some $\Delta > 0$ for *discrete time* processes and $\mathcal{T} = [0, \infty)$ for *continuous time* processes.

Convergence. Consider a sequence $\{X^n\}_{n=1,2,\dots}$ of random variables on the same probability triple $(\Omega, \mathcal{F}, \mathbb{P})$. We recall some ways in which X^n may converge to some other random variable X .

Definition A.4 (Convergence in Probability). The sequence of random variables $\{X^n\}_{n=1,2,\dots}$ converges to the random variable X in probability, denoted $X^n \xrightarrow{\mathcal{P}} X$ if for all $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X^n - X| \geq \varepsilon) \rightarrow 0,$$

as $n \rightarrow \infty$.

A stochastic process can converge pointwise (for all $t \in \mathcal{T}$) in probability, but clearly a stronger condition is that of uniform convergence, where this convergence is uniform through \mathcal{T} .

Definition A.5 (Uniform convergence in probability). The sequence of stochastic processes $\{X_t^n\}_{t \in \mathcal{T}, n=1,2,\dots}$ on \mathcal{T} converges uniformly in probability to $\{X_t\}_{t \in \mathcal{T}}$ if for all $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{t \in \mathcal{T}} |X_t^n - X_t| \geq \varepsilon\right) \rightarrow 0,$$

as $n \rightarrow \infty$.

That is, the maximum deviation over the index set between the X_t^n and X_t is itself bounded and vanishes in probability.

Definition A.6 (Convergence in r -mean). The sequence of random variables $\{X^n\}_{n=1,2,\dots}$ converges to a random variable X in r -mean, denoted $X^n \xrightarrow{L^r} X$ if for all $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{E}(|X^n - X|^r) \rightarrow 0$, as $n \rightarrow \infty$.

Note in particular that for $r = 1$, this is convergence in mean, that is, $\mathbb{E}(X^n) \rightarrow \mathbb{E}(X)$. Convergence in probability does not automatically imply convergence in r -mean, though the converse is true. One special case where the converse holds is when X^n and X are bounded.

LEMMA A.7. Suppose $X^n \xrightarrow{\mathcal{P}} X$ and $|X| \leq M$ and $|X^n| \leq M$ for some $M \geq 0$. Then $X^n \xrightarrow{L^r} X$ for all $r \geq 1$.

Markov processes. A Markov process is a stochastic process which additionally satisfies a property of being *memoryless*, meaning that its future evolution is entirely dictated by its current state independent of the past. We restrict ourselves to time-homogeneous Markov chains with finite state spaces.

Definition A.8 (Discrete time Markov chain). A *discrete time Markov chain* is a stochastic process on a finite state space $\mathcal{S} \subseteq \mathbb{R}$ and index set $\mathcal{T} = \{0, 1, 2, 3, \dots\}$ that satisfies the property

$$\mathbb{P}(X_{t+1} = j \mid X_t = i_t, X_{t-1} = i_{t-1}, \dots, X_1 = i_1) = \mathbb{P}(X_{t+1} = j \mid X_t = i_t),$$

for all states $j, i \in \mathcal{S}$ and all $t \in \mathcal{T}$. The evolution of such a Markov chain is therefore determined by its (one step) *transition probabilities* $p(i, j)$ for $i, j \in \mathcal{S}$, arranged in a transition matrix P defined by

$$p(i, j) = \mathbb{P}(X_{t+1} = j \mid X_t = i).$$

Definition A.9 (Continuous time Markov chain). A continuous time Markov chain is a stochastic process on a finite state space $\mathcal{S} \subseteq \mathbb{R}$ and index set $\mathcal{T} = [0, \infty)$ that satisfies the property

$$\mathbb{P}(X_{t+s} = j \mid X_t = i, X_{t_l} = i_l, X_{t_{l-1}} = i_{l-1}, \dots, X_0 = i_0) = \mathbb{P}(X_s = j \mid X_0 = i), \quad (14)$$

for all $s \geq 0$, $t > t_l > t_{l-1} > \dots > t_1 > t_0 \geq 0$ and all $j, i, i_l \in \mathcal{S}$.

Note that (14) implies that the transition times must also be random and memoryless. It turns out that this property embodies continuous time Markov chains with a large amount of structure. In particular, the time between transitions (when the chain moves from one state to another) must be exponentially distributed (this is the only memoryless continuous distribution, satisfying $\mathbb{P}(X \leq x \mid X \geq y) = \mathbb{P}(X \leq x - y)$).

In the discrete case the evolution was defined by the one step transition probabilities P . In the continuous case, these are replaced by a Matrix-valued function of time $P(t)$. It turns out there is a compact way to represent the possible transitions via *transition rates* $q(i, j)$ that describe the instantaneous rate of moving from state i to state j as formalized by the following lemma.

THEOREM A.10 (TRANSITION RATES). For a continuous time Markov chain on a finite state space, let $p(i, j; t) = \mathbb{P}(X_t = j \mid X_0 = i)$ for $t \geq 0$. Denote by $P(t)$ the matrix with entries $p(i, j; t)$. Then there exists a transition rate matrix Q with entries $q(i, j)$ that is the unique solution to

$$Q = \lim_{t \rightarrow 0} \frac{P(t) - I}{t},$$

note that $q(i, i) = -\sum_{j \neq i} q(i, j)$.

We next introduce the concept of a Poisson point process in order to build continuous time Markov chains from discrete time ones. Intuitively, a homogeneous Poisson point process simply counts the number of events that have happened until time t , where the time between successive events is iid and exponentially distributed.

Definition A.11 (Homogeneous Poisson point process). Let $\lambda > 0$. Define the stochastic process $\{H_t\}_{t \geq 0}$ with state space $\{0, 1, 2, \dots\}$ as follows. For $i = 1, 2, \dots$ let $E_i \sim \text{Exp}(\lambda)$, and let $H_t = \sup_{i=0,1,2,\dots} \{\sum_{j=1}^i E_j \leq t\}$. We call $\{H_t\}_{t \geq 0}$ the homogeneous Poisson process with rate λ on $t \geq 0$. Note that $H_t \sim \text{Pois}(\lambda t)$, and in particular $\mathbb{E}(H_t) = \lambda t$ for all $t \geq 0$.

A Homogeneous Poisson point process is an example of a very simple Markov chain.

LEMMA A.12. A homogeneous Poisson point process with rate λ is a continuous time Markov chain on the countable state space $\mathcal{S} = \{0, 1, 2, \dots\}$ with transition rates

$$q(i, j) = \begin{cases} \lambda, & j = i + 1, \\ -\lambda, & j = i, \\ 0, & \text{otherwise.} \end{cases}$$

Given a continuous time Markov chain, one can decompose it into a discrete time Markov chain that describes the transition probabilities at jump times, and a per-state rate of an Exponential distribution that describes the duration that the chain stays in that state before jumping to the next.

The following theorem gives one variation of a theorem that allows easily constructing continuous time Markov chains from discrete time Markov chains.

THEOREM A.13 (CONSTANT-RATE MARKOV CHAIN EMBEDDING). *Let $\{X_n\}_{n=0,1,2,\dots}$ be a discrete time Markov chain on finite state space \mathcal{S} with $X_0 = s_0$ for some initial state $s_0 \in \mathcal{S}$. Let $\lambda > 0$ be some rate and let $\{H_t\}_{t \geq 0}$ be the homogeneous Poisson point process with rate λ on $[0, \infty)$.*

Define a stochastic process $\{Z_t\}_{t \geq 0}$ by $Z_t = X_{H_t}$. Then $\{Z_t\}_{t \geq 0}$ is a continuous time Markov chain whose state space is \mathcal{S} , initial state is s_0 , and that has rates $q(i, j) = \lambda p(i, j)$ for $i \neq j$ and $q(i, i) = -\lambda(1 - p(i, i))$. This continuous time Markov chain is called an embedded chain because its transition rates follow the transition probabilities of the discrete chain, that is, $\mathbb{P}(Z_s = j \mid Z_t = i) = p(i, j)$ if $H_t = H_s + 1$.

Martingales. A martingale is a stochastic process with the property that its expected value in the future is the current value.

Definition A.14 (Martingale). A stochastic process $\{X_t\}_{t \geq 0}$ is a *martingale* if $\mathbb{E}(X_s \mid X_t) = X_t$ for all $s \geq t$.

The homogeneous Poisson process corrected for its mean is a martingale. That is, $H_t - \lambda t$ is a martingale. The following is a standard result in martingale theory.

THEOREM A.15 (DOOB'S MARTINGALE INEQUALITY). *If $\{X_t\}_{t \geq 0}$ is a martingale, then for all $T \geq 0$ and $C > 0$,*

$$\mathbb{P}\left(\sup_{t \in [0, T]} X_t \geq C\right) \leq \frac{\mathbb{E}(\max(X_T, 0))}{C}.$$

Further, for $r \geq 1$, we have

$$\mathbb{P}\left(\sup_{t \in [0, T]} |X_t| \geq C\right) \leq \frac{\mathbb{E}(|X_T|^r)}{C^r}.$$

B Extension to Schools Having Multiple Seats

In Section 2.4, we set up an analogue of the discrete model of Section 2.1 in the case that each school has $q = 1, 2, 3, \dots$ seats. In this appendix, we roughly trace the steps of Section 2 for the case of arbitrary q . We first describe the continuous model for the multiple-seat case, then discuss how Theorem 2.6 may be extended to this case. In particular, we develop the initial value problem equivalents of (4) for arbitrary $q = 1, 2, 3, \dots$ and $d \geq 1$. We discuss the solutions to the initial value problem for $d = 1$, then discuss how we may use this solution for $d > 1$.

The continuous model with multiple seats. Recall that $T_i^{n,d}$ is the number of schools with q seats taken after students $\{1, 2, 3, \dots, i-1\}$ have had their turn, and S_i^k is the number of schools with k seats taken at this point for $k = 1, 2, \dots, q-1$. We again keep track of the continuous analogues, for $i = \lfloor tn \rfloor$, letting $x_d(t)$ be the proportion of schools with all seats taken, and $y_d^k(t)$ be the proportion of schools with k seats taken for $k = 0, \dots, q-1$. Initially all schools have no seats taken, so

$$x_d(0) = 0, \quad y_d^0(0) = 1, \quad y_d^k(0) = 0, \quad k = 1, 2, \dots, q-1. \quad (15)$$

As we discuss in the main body, the probability that student t is matched to any student goes to $1 - x_d(t)^d$ in probability, and the proportion of time they get matched to a school with k seats taken is proportional to S_i^k , we therefore get that the rate at which schools with k seats taken become schools with $k+1$ seats taken is

$$(1 - x_d(t)^d) \frac{y_d^k(t)}{\sum_{m=0}^{q-1} y_d^m(t)} = \frac{1 - x_d(t)^d}{1 - x_d(t)} y_d^k(t),$$

since $x_d(t) + \sum_{k=0}^{q-1} y_d^k(t) = 1$. Note further that this is a positive flow into y_d^{k+1} and negative for y_d^k . We have each of these flows from y_d^k to y_d^{k+1} for $k = 0, \dots, q-1$ and one from y_d^{q-1} to x_d . For convenience, define

$$\gamma_d(t) = \frac{1 - x_d(t)^d}{1 - x_d(t)}.$$

Denoting derivative with respect to time with a dot for clarity to avoid multiple superscripts, this gives us the following differential equation

$$\begin{aligned} \dot{y}_d^0(t) &= -\gamma_d(t)y_d^0(t), \\ \dot{y}_d^k(t) &= \gamma_d(t)(y_d^{k-1}(t) - y_d^k(t)), \quad k = 1, \dots, q-1, \\ \dot{x}_d(t) &= \gamma_d(t)y_d^q(t). \end{aligned} \tag{16}$$

Connection with the discrete model. Following the steps in Section 2.3, we can apply a slightly more general equivalent of Theorem 2.6 for this multi-dimensional case, again from [Kroese et al., 2013, Theorem 17.3.1] to prove the following lemma.

LEMMA B.1. Fix $d \in \mathbb{N}$, $q = 1, 2, \dots$ and define the initial value problem for $\{x_d, y_d^k\}$ for $k = 0, 1, \dots, q-1$ via (15) and (16). For $t \geq 0$, let $T_i^{n,d}$ and S_i^k for $k = 0, \dots, q-1$ be as in Section 2.4. Then $n^{-1}T_{\lfloor tn \rfloor}^{n,d} \rightarrow x_d(t)$ and $n^{-1}S_{\lfloor tn \rfloor}^k \rightarrow y_d^k(t)$ for $k = 0, 1, \dots, q-1$ uniformly in probability as $n \rightarrow \infty$, where $\{x_d(t), y_d^k(t)\}$ for $k = 0, 1, \dots, q-1$ is the unique solution satisfying the initial value problem of (15) and (16) for $t \geq 0$.

PROOF. The proof follows identically to the proof of Theorem 2.6 in Section 3, adapted to the case of multiple equations. \square

Dynamics of the multiple seats model for $d = 1$. Observe that for $d = 1$, we have $\gamma_d \equiv 1$. As in the earlier continuous model, we can again solve this explicitly. It's easy to see we get $y_1^0(t) = \exp(-t)$, and for $k = 1$, we now have $\dot{y}_1^1 = \exp(-t) - y_1^1$, which yields $y_1^1(t) = t \exp(-t)$. Following this pattern, we have

$$\begin{aligned} y_1^k(t) &= \frac{t^k}{k!} e^{-t}, \quad k = 0, \dots, q-1, \\ x_1(t) &= 1 - e^{-t} \sum_{k=0}^{q-1} \frac{t^k}{k!}. \end{aligned} \tag{17}$$

Readers familiar with probability or phase-type distributions will note that this is the cumulative distribution function of an Erlang distribution with parameters $k = q$ and $\lambda = 1$. Figure 4 shows the solution for $d = 1$, $q = 4$.

Recall finally that a student t gets matched with probability that goes to $1 - x_d(t)^d$. Denote by $M_i^{n,d,q}$ be the indicator random variable of whether a student i gets matched with n schools, d length lists, and q seats per school. We therefore have

$$\mathbb{P}\left(M_i^{n,d,q} = 1\right) \xrightarrow{\mathcal{P}} 1 - x_d(t)^d,$$

as $n \rightarrow \infty$. This follows similarly to Lemma 2.7 via Lemma B.1. Note that the difference to Lemma 2.7 is simply that this time $x_d'(t)$ is not necessarily equal to $1 - x_d(t)^d$.

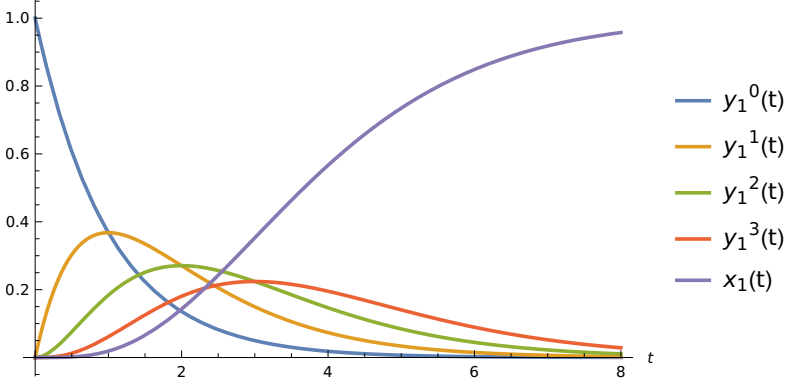


Fig. 4. Solution to the multiple-seat continuous market for $d = 1$, $q = 4$.

Discussion of dynamics for $q > 1$. The fact that γ_d multiplies every equation in (16) suggests the following idea. Define an initial value problem for τ_d as

$$\tau_d(0) = 0, \quad \dot{\tau}_d(t) = \frac{1 - x_1(\tau_d(t))^d}{1 - x_1(\tau_d(t))}. \quad (18)$$

It is not hard to show that the following now holds

$$\begin{aligned} y_d^k(t) &= y_1^k(\tau_d(t)), & k &= 0, 1, \dots, q-1 \\ x_d(t) &= x_1(\tau_d(t)). \end{aligned} \quad (19)$$

This now suggests a procedure for computing solutions to the initial value problem defined by (15) and (16): first compute the solution to (17), then solve (18) to compute $\tau_d(t)$ and plug it into (19).

We further remark that one can interpret $\tau_d(x)$ as describing the ratio of how much faster students are matched to schools when they get $d > 1$ tries compared to a single try. $1 - x_1^d$ is the probability of getting matched to any school given a list of length d and $1 - x_1$ is that probability with lists of length 1. The interpretation via (19) then tells us that the case for general d is exactly the same as for $d = 1$ except now the rate at which students are matched to schools is rescaled by a factor of τ_d .

C Missing Proofs

From Section 2.4.

COROLLARY 2.9. *Let $d, \ell, n \in \mathbb{N}$ with $n \geq \ell \geq d$. We have*

$$\mathbb{P}(R^{n,\ell} = 1) \geq \mathbb{P}(R^{n,d} = 1).$$

PROOF. By symmetry, we have

$$\mathbb{P}(R^{n,d} = 1) = \frac{n}{m} \mathbb{P}(H_1^{n,d}(m) = 1),$$

where m is the number of students in the market. From Lemma 2.2, $\mathbb{P}(H_1^{n,d}(m) = 1)$ increases with d , so $\mathbb{P}(R^{n,d} = 1)$ also increases with d . \square

THEOREM C.1. Suppose $x(d, t) \geq 0$ is defined implicitly via

$$\int_0^x \frac{1}{1-v^d} dv = t.$$

Then for $t \in [0, 1]$, $d \geq 1$, we have

$$\int_0^x \frac{1 + \log u}{1 - u^d} du \leq 0. \quad (20)$$

PROOF. Observe that $x(d, t)$ is monotonically increasing in t and that $x(d, t) \leq t$. The integrand in (20) is negative on $u \in [0, 1/e]$ and vanishes at $u = 1/e$ after which it increases monotonically. The results therefore follows immediately when $x(d, t) \in [0, 1/e]$, and since the integral increases for $x \in [1/e, 1]$, it suffices to show the result for $x(d, 1)$. In Lemma 3.3 we show that

$$x(d, 1) \leq \left(\frac{d+1}{2d+1} \right)^{1/d},$$

for $d \geq 1$ and in Lemma C.3 we show

$$\int_0^{\left(\frac{d+1}{2d+1}\right)^{1/d}} \frac{1 + \log(u)}{1 - u^d} du \leq 0,$$

for $d \geq 1$ which completes the proof. \square

LEMMA 3.3. For $d \geq 1$, we have $\left(\frac{2d+1}{4d+1}\right)^{1/d} \leq x(d, 1) \leq \left(\frac{d+1}{2d+1}\right)^{1/d}$.

PROOF. We first show the upper bound. Recall $x(d, 1)$ is defined via

$$1 = \int_0^{x(d,1)} \frac{du}{1 - u^d}.$$

Since $(1 - u^d)$ is strictly positive, it suffices to show

$$\int_0^{x(d,1)} \frac{du}{1 - u^d} \leq \int_0^{\left(\frac{d+1}{2d+1}\right)^{1/d}} \frac{du}{1 - u^d},$$

or equivalently

$$1 \leq \int_0^{\left(\frac{d+1}{2d+1}\right)^{1/d}} \frac{du}{1 - u^d}.$$

$(1 - u^d)^{-1}$ is the sum of a geometric series for $u^d \in (0, 1)$, so with $z \in (0, 1)$ write

$$\begin{aligned}
\int_0^z \frac{du}{1-u^d} &= \int_0^z \sum_{r=0}^{\infty} u^{rd} du \\
&= \sum_{r=0}^{\infty} \int_0^z u^{rd} du \\
&= \sum_{r=0}^{\infty} \frac{z^{rd+1}}{rd+1} \\
&= z \left(1 + \sum_{r=1}^{\infty} \frac{(z^d)^r}{rd+1} \right) \tag{21}
\end{aligned}$$

$$\begin{aligned}
&\geq z \left(1 + \sum_{r=1}^{\infty} \frac{(z^d)^r}{r(d+1)} \right) \\
&= z \left(1 - \frac{1}{d+1} \log(1-z^d) \right). \tag{22}
\end{aligned}$$

The validity of exchanging the order of summation and integration follows by the Tonelli theorem since the integrand is non-negative. Now substitute $z = \left(\frac{d+1}{2d+1} \right)^{1/d}$ and $q = 1/d$. We have $q \in (0, 1]$ and $z = \left(\frac{q+1}{q+2} \right)^q$, and we need to show that (22) is lower bounded by 1. After some rearranging, this reduces to showing

$$\log(2+q) \geq \left(\frac{q+1}{q} \right) \left(\left(1 + \frac{1}{q+1} \right)^q - 1 \right),$$

for $q \in (0, 1]$, which we show in Lemma C.7.

For the lower bound, we apply identical reasoning until (21), but must show that for $z = \left(\frac{2d+1}{4d+1} \right)^{1/d}$, we have

$$1 \geq z \left(1 + \sum_{r=1}^{\infty} \frac{(z^d)^r}{rd+1} \right).$$

We again substitute $q = 1/d$ and this time get $z = \left(\frac{q+2}{q+4} \right)^q$, and this yields

$$\begin{aligned}
z \left(1 + \sum_{r=1}^{\infty} \frac{(z^d)^r}{rd+1} \right) &= z \left(1 + \frac{z^d}{d+1} + \sum_{r=2}^{\infty} \frac{(z^d)^r}{rd+1} \right) \\
&\leq z \left(1 + \frac{z^d}{d+1} + \sum_{r=2}^{\infty} \frac{(z^d)^r}{rd} \right) \\
&= z \left(1 + \frac{z^d}{d+1} - \frac{z^d}{d} - \frac{1}{d} \log(1-z^d) \right) \\
&= z \left(1 - \frac{z^d}{d(d+1)} - \frac{1}{d} \log(1-z^d) \right) \\
&= \left(\frac{q+2}{q+4} \right)^q \left(1 - \frac{q^2(q+2)}{(q+1)(q+4)} - q \log \left(\frac{2}{q+4} \right) \right).
\end{aligned}$$

To complete the proof, we therefore need

$$\left(\frac{q+2}{q+4}\right)^q \left(1 - \frac{q^2(q+2)}{(q+1)(q+4)} - q \log\left(\frac{2}{q+4}\right)\right) \leq 1$$

for $q \in (0, 1]$, which we show in Lemma C.2. \square

LEMMA C.2. $1 - \frac{q^2(q+2)}{(q+1)(q+4)} - q \log\left(\frac{2}{q+4}\right) \leq \left(\frac{q+4}{q+2}\right)^q$ for $q \in [0, 1]$.

PROOF.

$$\begin{aligned} 1 - \frac{q^2(q+2)}{(q+1)(q+4)} - q \log\left(\frac{2}{q+4}\right) &\leq \left(\frac{q+4}{q+2}\right)^q \\ \iff \left(\frac{q+4}{q+2}\right)^q + \frac{q^2(q+2)}{(q+1)(q+4)} - q \log\left(2 + \frac{q}{2}\right) - 1 &\geq 0 \\ \iff \left(\frac{q+4}{q+2}\right) \left(\frac{q+4}{q+2}\right)^{q-1} + \frac{q^2(q+2)}{(q+1)(q+4)} - q \log\left(2 + \frac{q}{2}\right) - 1 &\geq 0. \end{aligned}$$

Now apply Claim 1 and Claim 2 to get

$$\begin{aligned} &\left(\frac{q+4}{q+2}\right) \left(\frac{q+4}{q+2}\right)^{q-1} + \frac{q^2(q+2)}{(q+1)(q+4)} - q \log\left(2 + \frac{q}{2}\right) - 1 \\ &\geq \left(\frac{q+4}{q+2}\right) \left(\frac{1}{2} + \frac{q}{8}(1 + 4 \log(2))\right) + \frac{q^2(q+2)}{(q+1)(q+4)} - q \left(\frac{1}{4} \left(1 - \frac{q}{10}\right) q + \log(2)\right) - 1 \\ &= \frac{q^2(q^4 - 3q^3 - 11q^2 - 20q^2 \log(2) + 53q - 100q \log(2) + 100 - 80 \log(2))}{40(q+1)(q+2)(q+4)} \end{aligned}$$

This is positive if $q^4 - 3q^3 - 11q^2 - 20q^2 \log(2) + 53q - 100q \log(2) + 100 - 80 \log(2) \geq 0$, but this is a polynomial with two real roots, both at $q > 1$, so since this equation is positive for $q = 0$, it must be non-negative for all $q \in [0, 1]$. \square

CLAIM 1. $\left(\frac{q+4}{q+2}\right)^{q-1} \geq \frac{1}{2} + \frac{q}{8}(1 + 4 \log(2))$ for $q \in [0, 1]$.

PROOF. Observe this holds at equality for $q = 0$ and apply Claim 2 to the the derivative of the difference to get:

$$\frac{\partial}{\partial q} \left[\left(\frac{q+4}{q+2}\right)^{q-1} - \frac{1}{2} - \frac{q}{8}(1 + 4 \log(2)) \right] \geq \frac{1}{16} \left((q+2)(q+4) \log\left(\frac{q+4}{q+2}\right) - 2(q+4 \log(2)) \right).$$

This is now true if and only if

$$\log\left(\frac{q+4}{q+2}\right) - \frac{2(q+4 \log(2))}{(q+2)(q+4)} \geq 0. \quad (23)$$

One easily checks

$$\frac{\partial}{\partial q} \left[\log\left(\frac{q+4}{q+2}\right) - \frac{2(q+4 \log(2))}{(q+2)(q+4)} \right] = \frac{4(-3q + 4q \log(2) - 8 + 12 \log(2))}{(q+2)^2(q+4)^2} \geq 0,$$

where the inequality holds since both the numerator and the denominator are nonnegative. Since one readily verifies that (23) is valid for $q = 0$, the claim follows. \square

CLAIM 2. $\left(\frac{q+4}{q+2}\right)^q \geq \frac{(q+4)^2}{16}$ and $\log\left(2 + \frac{q}{2}\right) \leq \log(2) + \frac{1}{4} \left(1 - \frac{q}{10}\right) q$ for $q \in [0, 1]$.

PROOF. For the first:

$$\begin{aligned} \left(\frac{q+4}{q+2}\right)^q &\geq \frac{(q+4)^2}{16} \\ \iff \left(\frac{q+4}{q+2}\right)^q &\geq \left(1 + \frac{q}{4}\right)^2 \\ \iff \log\left(\frac{q+4}{q+2}\right) &\geq \frac{2}{q} \log\left(1 + \frac{q}{4}\right). \end{aligned}$$

Note that the left hand side always exceeds $\log(5/3) > 1/2$, and the right hand side never exceeds $1/2$ because $\log(1 + q/4) \leq q/4$.

For the second, take the derivative of the difference of the left term and the right term:

$$\begin{aligned} \frac{\partial}{\partial q} \left(\log\left(2 + \frac{q}{2}\right) - \frac{1}{4} \left(1 - \frac{q}{10}\right) q - \log(2) \right) &= \frac{1}{q+4} + \frac{q-5}{20} \\ &= \frac{(q-1)q}{20(q+4)}, \end{aligned}$$

which is now clearly non-positive. \square

LEMMA C.3. For $d \geq 1$, we have

$$\int_0^{\left(\frac{d+1}{2d+1}\right)^{1/d}} \frac{1 + \log(u)}{1 - u^d} du \leq 0.$$

PROOF. Change variables with $q = 1/d$ (so $q \in [0, 1]$) and define $A = \frac{d+1}{2d+1} = \frac{q+1}{q+2}$, then write

$$\begin{aligned} \int_0^{A^q} \frac{1 + \log(u)}{1 - u^{1/q}} du &= \int_0^{A^q} \sum_{n=0}^{\infty} u^{n/q} (1 + \log(u)) du \\ &= \sum_{n=0}^{\infty} \int_0^{A^q} u^{n/q} (1 + \log(u)) du \\ &= \sum_{n=0}^{\infty} \left(\int_0^{A^q} u^{n/q} du + \int_0^{A^q} u^{n/q} \log(u) du \right) \\ &= \sum_{n=0}^{\infty} \left(\frac{qA^{n+q}}{n+q} + \frac{qA^{n+q} \log(A^q)}{n+q} - \frac{q^2 A^{n+q}}{(n+q)^2} \right) \\ &= A^q \left((1 + \log(A^q)) \sum_{n=0}^{\infty} \frac{qA^n}{n+q} - \sum_{n=0}^{\infty} \frac{q^2 A^n}{(n+q)^2} \right) \\ &= A^q \left((1 + \log(A^q)) \left(1 + \frac{qA}{q+1} + qA^2 \sum_{n=2}^{\infty} \frac{A^{n-2}}{n+q} \right) - 1 - \frac{q^2 A}{(q+1)^2} - \sum_{n=2}^{\infty} \frac{q^2 A^n}{(n+q)^2} \right) \\ &\leq A^q \left((1 + \log(A^q)) \left(1 + \frac{qA}{q+1} + 2qA^2(\log(4) - 1) \right) - 1 - \frac{q^2 A}{(q+1)^2} \right) \\ &\leq A^q \left((1 + q \log(A)) 2A(1 + qA(\log(4) - 1)) - 1 - \frac{q^2}{(q+2)^2} \right) \\ &\leq A^q \left(\left(1 - \frac{2q \log^2(2)}{q + \log(4)} \right) (1 + q \log(2)) - 1 - \frac{q^2}{(q+2)^2} \right). \end{aligned}$$

In the first inequality we removed non-positive terms and applied Lemma C.6. In the second we applied the facts that $1 + \frac{qA}{q+1} = 2A$ and $\frac{q^2}{(q+2)^2} \leq \frac{q^2 A}{(q+1)^2}$. The last inequality follows from Lemma C.4 and Lemma C.5.

This quantity is now non-positive if

$$\left(1 - \frac{2q \log^2(2)}{q + \log(4)}\right) (1 + q \log(2)) - 1 - \frac{q^2}{(q+2)^2} \leq 0.$$

This holds at equality for $q = 0$, so it suffices to show that the derivative is non-positive. Writing $x = \log(2)$, abbreviating by \dots the left-hand-side of the previous inequality, we have

$$\begin{aligned} \frac{\partial}{\partial q} (\dots) &= \frac{\partial}{\partial q} \left(\left(1 - \frac{2q \log^2(2)}{q + \log(4)}\right) (1 + q \log(2)) - 1 - \frac{q^2}{(q+2)^2} \right) \\ &= \frac{\partial}{\partial q} \left(q \log(2) - \frac{2q \log^2(2)}{q + \log(4)} (1 + q \log(2)) - \frac{q^2}{(q+2)^2} \right) \\ &= -\frac{\partial}{\partial q} \left(2x^2 \frac{q}{q+2x} + 2x^3 \frac{q^2}{q+2x} \right) + x - \frac{2q(q+2)^2 - 2q^2(q+2)}{(q+2)^4} \\ &= -2x^2 \left(\frac{q+2x-q}{(q+2x)^2} \right) - 2x^3 \left(\frac{2q(q+2x) - q^2}{(q+2x)^2} \right) + x - \frac{4q}{(q+2)^3} \\ &= \frac{(q+2x)^2 x - 4x^3 - 2x^3 q(q+4x)}{(q+2x)^2} - \frac{4q}{(q+2)^3} \\ &= \frac{xq^2 + 4qx^2 + 4x^3 - 4x^3 - 2x^3 q^2 - 8x^4 q}{(q+2x)^2} - \frac{4q}{(q+2)^3} \\ &= \frac{xq(q+4x - 2x^2(q+4x))}{(q+2x)^2} - \frac{4q}{(q+2)^3} \\ &= \frac{qx(1 - 2x^2)(q+4x)}{(q+2x)^2} - \frac{4q}{(q+2)^3} \\ &= q \left(\frac{\log(2) (1 - 2 \log^2(2)) (q + \log(16))}{(q + \log(4))^2} - \frac{4}{(q+2)^3} \right) \\ &\leq q \left(\frac{\log(2) (1 - 2 \log^2(2)) (1 + \log(16))}{(0 + \log(4))^2} - \frac{4}{(1+2)^3} \right) \\ &\leq -0.08q \\ &\leq 0, \end{aligned}$$

□

LEMMA C.4. For $q \in [0, 1]$ and $A = \frac{q+1}{q+2}$,

$$-\frac{\log(2)}{1 + q \log(2)} \leq \log(A) \leq -\frac{2 \log^2(2)}{q + 2 \log(2)}.$$

PROOF. For the first inequality from the thesis, we need

$$\begin{aligned} \log(A) + \frac{\log(2)}{1 + q \log(2)} &\geq 0 \\ \iff (1 + q \log(2)) \log\left(\frac{q+1}{q+2}\right) + \log(2) &\geq 0. \end{aligned}$$

One can readily show that the second derivative is negative, so the function is concave and exceeds a line between the endpoints, where the inequality also holds. This means the inequality holds for all $q \in [0, 1]$. More in detail, we have

$$\begin{aligned} \frac{\partial}{\partial q} \left[(1 + q \log(2)) \log\left(\frac{q+1}{q+2}\right) \right] &= \log(2) \log\left(\frac{q+1}{q+2}\right) + \frac{1 + q \log(2)}{(q+1)(q+2)}, \\ \frac{\partial^2}{\partial^2 q} \left[(1 + q \log(2)) \log\left(\frac{q+1}{q+2}\right) \right] &= \frac{\partial}{\partial q} \left(\log(2) \log\left(\frac{q+1}{q+2}\right) + \frac{1 + q \log(2)}{(q+1)(q+2)} \right) \\ &= \frac{\log 2}{(q+1)(q+2)} + \frac{\log(2)(q+1)(q+2) - (1 + q \log(2))(2q+3)}{(q+1)^2(q+2)^2} \\ &= \frac{2 \log(2)(q+1)(q+2) - (1 + q \log(2))(2q+3)}{(q+1)^2(q+2)^2} \\ &= \frac{\log(2)(2q^2 + 6q + 4) - (1 + q \log(2))(2q+3)}{(q+1)^2(q+2)^2} \\ &= \frac{\log(2)(3q+4) - (2q+3)}{(q+1)^2(q+2)^2} \\ &= \frac{q(3 \log(2) - 2) + 4 \log(2) - 3}{(q+1)^2(q+2)^2}. \end{aligned}$$

The numerator is linear in q and negative for $q = 0$ and $q = 1$, while the numerator is strictly positive for every q . So, the second derivative is negative for all q .

The second inequality from the thesis holds at equality for $q = 0$. Further, one can show that the derivative of the difference is negative for $q \in [0, 1]$:

$$\frac{\partial}{\partial q} \left(\log(A) + \frac{2 \log^2(2)}{q + \log(4)} \right) = \frac{1}{(q+1)(q+2)} - \frac{2 \log^2(2)}{(q + \log(4))^2}.$$

It is not hard to verify that for $q \in [0, 1]$

$$(q+1)(q+2) \geq \frac{(q + \log(4))^2}{2 \log^2(2)},$$

which completes the proof. □

LEMMA C.5. For $q \in [0, 1]$ and $A = \frac{q+1}{q+2}$, we have

$$2A(1 + qA(\log(4) - 1)) \leq 1 + q \log(2).$$

PROOF.

$$\begin{aligned}
& 2A(1 + qA(\log(4) - 1)) \leq 1 + q \log(2) \\
\iff & \frac{2(1 + q)(1 + qA(\log(4) - 1))}{2 + q} - 1 - q \log(2) \leq 0 \\
\iff & \frac{q^2(q(\log(8) - 2) - 3 + \log(16))}{(q + 2)^2} \leq 0 \\
\iff & q(\log(8) - 2) - 3 + \log(16) \leq 0.
\end{aligned}$$

This is increasing in q and one can verify this inequality holds for $q = 1$. \square

LEMMA C.6. For $A = \frac{q+1}{q+2}$, $q \in [0, 1]$, we have

$$\sum_{n=2}^{\infty} \frac{A^{n-2}}{n + q} \leq 2(\log(4) - 1).$$

PROOF. One can verify that the sum is decreasing in q , so the left hand side is maximized at $q = 0$. Then, using the fact that $\sum_{n=1}^{\infty} \frac{2^{-n}}{n} = \log(2)$ (this is the polylogarithm of order 1 evaluated at $1/2$), we have

$$\sum_{n=2}^{\infty} \frac{A^{n-2}}{n + q} \leq \sum_{n=2}^{\infty} \frac{2^{2-n}}{n} = -2 + 4 \sum_{n=1}^{\infty} \frac{2^{-n}}{n} = 2(\log(4) - 1).$$

\square

LEMMA C.7. For $q \in [0, 1]$,

$$1 + \left(\frac{q}{q+1} \right) \log(q+2) \geq \left(\frac{q+2}{q+1} \right)^q.$$

PROOF. We will show this inequality after taking a logarithm on both sides.

We begin with two bounds for the logarithm, for $x \in [0, 1]$, we have

$$\log(1+x) \geq \frac{2x}{2+x}, \tag{24}$$

$$\log(2+x) \geq \log(2) + \frac{x}{2} - \frac{x^2}{4(1+x)}. \tag{25}$$

Both are easy to verify as they hold for $x = 0$ and the difference of the left and the right side has positive derivative on $x \in [0, 1]$. In fact, the first holds for all $x \geq 0$ (note $2x/(2+x) = x - x^2/(2+x)$):

$$\begin{aligned}
\frac{\partial}{\partial x} \left(\log(1+x) - x + \frac{x^2}{2+x} \right) &= \frac{1}{1+x} - 1 + \frac{2x(2+x) - x^2}{(2+x)^2} \\
&= \frac{1}{1+x} - 1 + \frac{(4+x)x}{(2+x)^2} \\
&= \frac{-x(2+x)^2 + (4+x)x(1+x)}{(1+x)(2+x)^2} \\
&= \frac{x^2}{(1+x)(2+x)^2}.
\end{aligned}$$

The second for $x \in [0, \sqrt{2}]$:

$$\begin{aligned} \frac{\partial}{\partial x} \left(\log(2+x) - \log(2) - \frac{x}{2} + \frac{x^2}{4(1+x)} \right) &= \frac{1}{2+x} - \frac{1}{2} + \frac{8x(1+x) - 4x^2}{16(1+x)^2} \\ &= \frac{x(2-x^2)}{4(1+x)^2(2+x)}. \end{aligned}$$

Applying (24) and (25) to the logarithm of the left hand side, we have

$$\begin{aligned} \log \left(1 + \left(\frac{q}{q+1} \right) \log(q+2) \right) &\geq \log \left(1 + \left(\frac{q}{q+1} \right) \left(\log(2) + \frac{q}{2} - \frac{q^2}{4(q+1)} \right) \right) \\ &= \log \left(1 + \frac{4q(q+1)\log(2) + q^2(q+2)}{4(q+1)^2} \right) \\ &\geq 2 \times \frac{4q(q+1)\log(2) + q^2(q+2)}{4(q+1)^2} \times \frac{4(q+1)^2}{4q(q+1)\log(2) + q^2(q+2) + 8(q+1)^2} \\ &= \frac{8q(q+1)\log(2) + 2q^2(q+2)}{4q(q+1)\log(2) + q^2(q+2) + 8(q+1)^2}. \end{aligned}$$

Next apply Lemma C.4, to get

$$\frac{q \log(2)}{1 + q \log(2)} \geq q \log \left(\frac{q+2}{q+1} \right).$$

The proof now hinges on showing that

$$\frac{8q(q+1)\log(2) + 2q^2(q+2)}{4q(q+1)\log(2) + q^2(q+2) + 8(q+1)^2} \geq \frac{q \log(2)}{1 + q \log(2)}.$$

Via simple algebraic manipulation, one can show that this holds if and only if

$$q^2 \log(2) - 2(1 - \log(2))(2 \log(2) - 1)q + 4(\log(2) - 1)^2 \geq 0.$$

This is now a quadratic in q , and in fact it has negative discriminant, so it has no real roots. Further, it is positive at $q = 0$, so the quadratic is positive for all $q \in [0, 1]$. \square

D Plots for Numerical Experiments

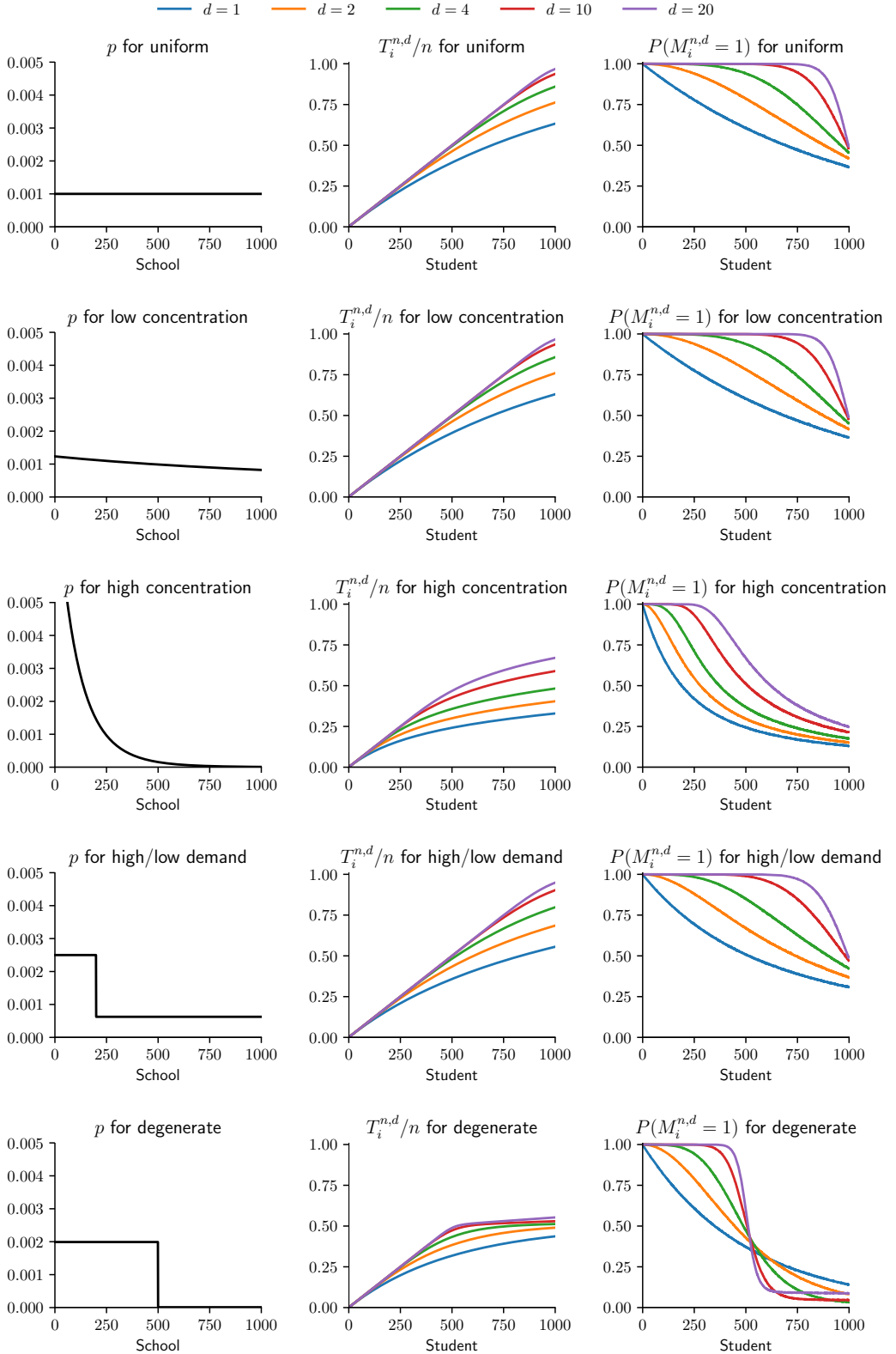


Fig. 5

