From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2023

Essays on Patching Markets

Maxey, Tyler Tucker

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01w9505372q

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on Patching Markets

Tyler Maxey

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Economics
Adviser: Leeat Yariv

May 2023

© Copyright by Tyler Maxey, 2023.

All Rights Reserved

Abstract

A patching market is a market that involves either piracy or matching. In this dissertation, I study both types of patching markets. In Chapter 1, I assess losses from piracy by merging novel title-level data on pirated and legal viewing for 660 titles on two popular streaming platforms (Netflix and Disney+) from January 2021 to June 2022. I estimate that during this period, pirates lowered platforms' revenues by 0.3%, equal to about $70 million. In Chapter 2, I study a model of centralized school choice in which students engage in costly search over schools before submitting preference reports to a clearinghouse. I show that common preferences generate search externalities, and the intensity of which vary by search cost and search protocol. In Chapter 3, I study a model of dynamic matching with transfers where I provide conditions under which transfers increase or decrease efficiency relative to a setting without transfers. I also provide asymptotics on equilibrium welfare as waiting costs vanish.

Acknowledgements

I could not have completed this dissertation without the guidance and assistance of numerous people, most prominently my main advisor Leeat Yariv. I cannot imagine a better advisor. She met with me countless times, always had new insights, meticulously reviewed draft after draft, did a practice presentation with me before every workshop, and was extremely supportive and kind. I could write an entire chapter about how great an advisor she is, and without her, this dissertation would have been completed with probability zero. I thank my other committee members, Jakub Kastl and Alessandro Lizzeri, for their superb advising. Their comments and insights were always helpful, and they greatly strengthened my work and my capacity as a researcher. I also thank Sylvain Chassang, Faruk Gul, Adam Kapor, Pietro Ortoleva, and the workshop/seminar participants who always kept me on my toes and offered such constructive feedback.

I thank my two main undergraduate advisors who encouraged me to go to grad school in the first place. They are Rhonda Righter—who gave me my first research assistantship and advised me on my first paper and conference presentation—and Shachar Kariv—whose game theory class is what first got me interested in economics. Shortly thereafter, Shachar convinced me to do an economics PhD, and since then he has continued to be an invaluable mentor.

Thank you to Kris Ayers, Ellen Graff, Laura Hedden, Christina Lipsky, Laura Sciarrotta, and Stu Sternbach for keeping the department running so smoothly. Thank you to Guy Oliver and MUSO for so much assistance and generosity with the piracy data, which made my first chapter possible.

I have so many friends to thank, in particular, Eduard K. Böhm and Carolyn Tsao, without whom I could not have finished this PhD. They are incredibly supportive, thoughtful, generous, and they have wonderful senses of humor that made graduate school so much more enjoyable. I also thank Christine Cai, Sharada Dharmasankar, Jacob Dorn, Jonas Jin, Faizaan Kisat, Leon Musolff, Karl Schulze, and Jesse Silbert for our personal relationships, out professional relationships, but most of all, for listening to me talk endlessly about Berkeley. (Go bears!) I also thank my Atlanta and Berkeley friends (Alec, Andrew, Bailey, Ben, CJ, Freya, Josh, Nicole, and Zelia) for their long-lasting friendships and for repeatedly telling me that no matter what, they will never call me "doctor."

Thank you to the Lewis Center for the Arts for letting me take undergraduate writing classes throughout grad school. Thank you to the members of the Graduate Improv Group for the many laughs and for consistently attending practice once the club started providing pizza. Thank you to everyone in the skit last year who let me run around yelling and throwing food at you for entertainment. It was an experience I had dreamt of but never thought would come to fruition.

Thank you to my A95 officemates, who I always looked forward to seeing and have created such a fun, productive working environment. (Also thank you to whoever left the dart board in that office; we have used it so many times that it is now broken. Sorry.) Thank you to Tara Ronda and my co-instructors who made my time with the Prison Teaching Initiative so rewarding and streamlined.

I thank Capital Preferences for being so flexible with my graduate studies as I worked part-time. I am particularly grateful for Trent Jorgensen and Nick Smith, who are fantastic colleagues that have made the job so much easier and more fun.

I am thankful for all the amazing comedy television writers (especially Tina Fey) for writing such hilarious shows. Their work is why I wanted to do a project on streaming. I am also thankful for the late Stephen Sondheim, whose music and lyrics, during many hours of solitude, kept me in good company.

Finally, I thank my family. They put up with so much from me. Thank you to my parents, who have been extremely supportive of everything I have done. They have given me so much, and I would not be anywhere without them. And thank you to my brothers for our many calls that have made us closer despite being hundreds of miles apart.

To dedications.

Contents

Abstract . . . 3

Acknowledgements . . . 4

1 Estimating Losses from Online Piracy **1**

1.1 Introduction . . . 2

1.1.1 Related Literature . . . 4

1.2 Data and Context . . . 6

1.2.1 Data . . . 6

1.2.2 Context . . . 6

1.3 Model . . . 8

1.3.1 Preferences for Titles . . . 9

1.3.2 Preferences for Platforms . . . 10

1.4 Estimation . . . 11

1.4.1 Model Discussion and Identification . . . 11

1.4.2 Estimation Strategy . . . 15

1.5 Results . . . 16

1.6 Counterfactual: Changing Piracy Disutility . . . 19

1.7 Conclusion . . . 21

2 School Choice with Costly Information Acquisition **23**

2.1 Introduction . . . 24

2.2 Related Literature . . . 27

2.3 Model . . . 28

2.3.1 Utility Specifications . . . 29

2.3.2 Search Protocols . . . 29

2.3.3 Matching Procedure . . . 30

2.3.4 Equilibrium . . . 31
2.4 Idiosyncratic Preferences . . . 31
2.4.1 Simultaneous Search . . . 32
2.4.2 Sequential Search and Welfare Comparison . . . 35
2.5 Common Preferences . . . 37
2.5.1 Simultaneous Search and Sequential Search . . . 37
2.5.2 Welfare Comparison . . . 39
2.5.3 Arbitrary Number of Schools . . . 41
2.5.4 Example with Hybrid Preferences . . . 42
2.6 Coordinating Search Through Tie-Breaking . . . 43
2.6.1 Simultaneous Search . . . 44
2.6.2 Sequential Search . . . 46
2.7 Conclusion . . . 48
2.8 Appendix . . . 49
2.8.1 Proofs of Main Results . . . 49
2.8.2 Other Priorities . . . 59

3 Dynamic Matching with Transfers **61**
3.1 Model . . . 62
3.1.1 Setup . . . 62
3.1.2 Timing . . . 62
3.1.3 Ordering Rules . . . 63
3.1.4 Equilibrium . . . 63
3.1.5 Transfers . . . 63
3.2 Results . . . 64
3.2.1 Equilibrium Strategies . . . 64
3.2.2 FIFO . . . 64
3.2.3 Welfare . . . 66
3.3 Related Literature . . . 68
3.4 Conclusion . . . 68
3.5 Appendix . . . 69

Chapter 1

Estimating Losses from Online Piracy

The digital release of media on streaming platforms has made it relatively easy to upload and find online pirated version of platforms' content, resulting in potentially large revenue losses. To assess losses from piracy, I merge novel title-level data on pirated and legal viewing for 660 titles on two popular streaming platforms (Netflix and Disney+) from January 2021 to June 2022. I combine this with subscription data to estimate a structural model of demand for titles and platforms that accounts for households' heterogeneous disutilites of piracy. I estimate that roughly 11% of U.S. households are willing to pirate though there are substantial disutilities to doing so. During the period I study, pirates lowered platforms' revenues by 0.3%, equal to about $70 million. Through my counterfactual analysis, I find that making piracy 25% less appealing would have generated only $33 million in displaced revenues while making piracy 25% more appealing would have generated $146 million in displaced revenues. This suggests that slightly easing the policing of piracy could have substantial effects.

1.1 Introduction

Piracy—the act of illegally reproducing or disseminating copyrighted material—undermines copyright protections and reduces copyright holders' incentives to innovate. The digital release of movies and television programs over streaming platforms has made that content relatively easy to pirate. This is potentially quite harmful to platforms; the Department of Justice recently indicted members of an international piracy ring that "caused tens of millions of dollars in losses to film production studios" (Department of Justice, 2021). This is one particular case, and it prompts the question: How large are the industry's losses across all sources of pirated content? This paper aims to answer that question.

Assessing piracy's losses requires knowing consumers' preferences, which can be difficult to measure. Rather than use surveys, which can suffer from problems of respondent selection and incentive incompatibility, I take a revealed-preference approach and combine subscription data with novel title-level data on pirated and legal viewing for 660 titles on two leading streaming platforms (Netflix and Disney+) from January 2021 to June 2022. To my knowledge, I am the first to estimate a structural demand model for titles and platforms. I incorporate piracy in my model, which is important for two reasons. First, it allows me to estimate what fraction of households are willing to pirate, estimate pirates' willingness to pay for content, and run counterfactuals on how subscriptions would respond to changes in piracy disutility. Second, omitting consumers' option to pirate biases preference estimates because it ignores an avenue by which consumers value a platform's content but do not subscribe to it. In this case, estimating the model without piracy overestimates the coefficient on library platform quality by 15%.

I estimate that 11% of households are willing to pirate content but that piracy is associated with substantial costs to the pirates themselves. In utility terms, pirating a brand-new title is approximately equal to legally streaming a title that is 21 weeks old, holding all else equal. This is large, as households strongly prefer new content: 95% of titles are in the top 30 most-streamed titles for less than 21 weeks. Consequently, many households who are willing to pirate still subscribe to streaming platforms. During the period I study, pirates displaced 0.3% of platforms' revenues, equal to roughly $70 million. Through my counterfactual analysis, I find that there is an asymmetry in how displaced revenue changes with piracy disutility. A 25% increase in piracy disutility (that is, making piracy 25% less appealing) would have generated only $33 million in displaced revenues while a 25% decrease in piracy disutility would have generated $146 million in displaced revenues. Thus, although the estimated displaced revenues are relatively small, slightly easing piracy policing

could generate substantial losses.

I merge data on pirated and legal viewing from two main sources. The pirated-viewing data come from a piracy-tracking firm. These data contain weekly title-level observations on the number of times a title is streamed directly on a piracy website and torrented off a peer-to-peer network.[1] To get these data, the firm monitors traffic on a database of 750,000 piracy domains and tracks activity at the title level for those which are most popular, covering about 97% of all traffic on their database's domains. Their traffic data come from various sources including direct first-party measurement from sites, anonymous web traffic from a clickstream panel of 3 million, anonymized web traffic data from internet service providers, and public sources. The legal-viewing data come from Nielsen. These data contain the number of minutes U.S. households spend legally streaming the most popular 10 movies and 20 television shows each week. Being the most popular titles, this dataset captures the movies and television programs most important to households' subscription decisions. My analysis focuses on the two most popular streaming platforms in the United States: Netflix and Disney+. These platforms together have roughly 70 million U.S. subscribers and make up 93% of viewing in the Nielsen data.

At the heart of my model is heterogeneity in propensity to pirate, which I capture through differential piracy disutility. Pirating content online is free but can still come at a non-monetary costly to the consumer. For example, pirated titles can exhibit lower video and audio quality, require more search time, and expose the user to cybersecurity risks. The piracy disutility parameter in my model captures these costs. I model consumer heterogeneity with two latent household types: willing to pirate (with some finite piracy disutility) and unwilling to pirate (with infinite piracy disutility). To estimate households' preferences from the data, I must disentangle what fraction of viewings and subscriptions come from each household type. The fraction of households willing to pirate is identified by joint changes in piracy and subscriptions. For example, suppose a platform gains subscribers. If, following the gain, the platform's titles are pirated less often, those new subscribers were likely pirating that platform's titles prior to subscribing. If the platform's titles have no change in pirated viewing, the new subscribers likely were not pirating its titles. To estimate this model, I implement a nested fixed-point procedure where I numerically solve a system of equations that matches share- and model-predicted platform average utilities by latent household type.

This model and estimation strategy allow me to then run counterfactuals on how subscriptions would change with changes in piracy disutility. For example, the government could increase efforts

[1]The piracy ring mentioned above distributed movies and television shows across multiple channels including streaming websites and peer-to-peer networks.

to remove piracy sites, which would make piracy less accessible and lead pirates to substitute toward purchasing platform subscriptions or not watching their content. For this counterfactual, I restrict attention to local changes in piracy disutility and hold prices and other supply-side responses constant. Platform prices are quite sticky hence unlikely to change for small changes in piracy levels.[2] As stated earlier, I find that a small decrease in piracy disutility can generate substantial platform losses.

This paper is a first step in assessing the displacement of streaming subscriptions. While my analysis focuses on the losses of two streaming platforms, one may also be concerned about the losses to other streaming platforms and other online distribution methods of content such as online video rentals. Additionally, piracy is not the only method by which subscriptions may be displaced. Platforms have recently voiced concerns over password sharing, in which more users utilize an account than intended. My analysis does not capture these other avenues of displaced subscriptions; thus, my results serve as a lower bound.

1.1.1 Related Literature

This paper builds on the empirical piracy literature by analyzing a new setting with novel data and a novel approach. My biggest contribution is that I take a structural approach and merge data on legal streamings, pirated streamings, and pirated torrents for 660 of the most popular movies and television programs on two large streaming platforms. This approach allows me to estimate the fraction of households willing to pirate and their average disutility for pirating. I can also run counterfactuals on how pirates would respond to changes in piracy disutility. Regarding setting, the previous literature focuses mostly on music and box office sales, which involves a series of one-shot decisions, whereas I focus on subscriptions to streaming platforms, a relatively new distribution channel that involves making decisions over bundles of goods, distinguishing it from previous work.

This literature's earliest work relies on data from surveys that ask respondents how often they pirate content and how they trade off legally purchasing and pirating content. Peitz and Waelbroeck (2004); Rob and Waldfogel (2006); Zentner (2006); and Leung (2015) use survey data to estimate how much piracy displaces music sales; Rob and Waldfogel (2007) and Bai and Waldfogel (2012) do so for film.

Later work estimates industry losses with reduced-form techniques by exploiting periods before and after piracy's introduction. Danaher and Waldfogel (2012) use variation in domestic and international release dates of Hollywood movies to assess whether longer lags, which allow for more

[2]For larger changes, one would want to account for supply-side responses.

international pre-release piracy, reduce box office sales abroad. They analyze this relationship before and after the advent of a popular peer-to-peer file-sharing network and find that international box office sales are significantly lower than they would have been had there been no pre-release piracy. Waldfogel (2012) compares the quality of music before and after the introduction of the popular peer-to-peer file sharing application Napster. He finds that music quality did not decrease following Napster's introduction.

More recent reduced-form work is able to assess piracy's losses well after piracy's introduction by directly using data on piracy observations, with Oberholzer-Gee and Strumpf (2007) and Oberholzer-Gee and Strumpf (2016) as the first to do so. They analyze the extent that file sharing affected music sales in 2002, a period during which torrenting was relatively popular. They use data containing observations on a peer-to-peer network's torrent downloads and find that piracy had insignificant effects on music sales.

Similar to this paper, Nolan et al. (2022) look at piracy and streaming platforms but take a reduced-form approach and focus more on how piracy affects the incentives of multiple-service operators rather than how it affects streaming platforms. They study a 10,000-household panel and observe how much time households spend on various internet applications along with whether the household adopted a specific technology that makes it easier to search for pirated titles. My data is less granular but covers a larger sample, namely the entire United States. Additionally, I am able to observe piracy at the title level and via channels that are not confined to the specific pirating software they focus on.

To my knowledge, Liu (2018) is the only other paper studying piracy that takes a structural approach. He focuses on demand for movies and DVDs. Because of our different settings, our papers are distinct both in terms of the model and estimation procedure. He models heterogeneity in piracy disutility with a random coefficient whereas I model it as a latent type in which some households have infinite piracy disutility and others have a finite piracy disutility. Consequently, he uses estimation techniques similar to Berry et al. (1995) whereas I numerically solve for a nested fixed point specific to my model then run GMM to estimate remaining parameters.

1.2 Data and Context

1.2.1 Data

I merge data on pirated and legal viewing from two main sources. The pirated-viewing data come from MUSO.com, a piracy-tracking firm. These data contain weekly title-level observations on the number of times a title is streamed directly on a piracy website and torrented off a peer-to-peer network for a total of 87,000 titles from July 2019 to June 2022. To get these data, the firm monitors traffic on a database of 750,000 piracy domains and tracks activity at the title level on the most popular domains, covering about 97% of all traffic on their database's domains. Their traffic data come from various sources including direct first-party measurement from sites, anonymous web traffic from a clickstream panel of 3 million, anonymized web traffic data from internet service providers, and public sources.

The legal-viewing data come from Nielsen. These data contain the number of minutes U.S. households spend legally streaming the most popular 10 movies and 20 television shows each week. Being the most popular titles, this dataset includes the movies and television programs most important to households' subscription decisions. These data run from January 2021 to June 2022, totaling 78 weeks, and contains observations for approximately 260 movies and 230 television shows.

The supplementary data come from a variety of sources. The number of platform subscriptions come from streaming platforms' quarterly reports. Prices come from the platforms' sites. Title characteristics come from IMDb. For the population of potential consumers, I use the National Center for Education Statistics' most recent figure for the number of U.S. households with internet access (104 million).

1.2.2 Context

Table 1.1 shows the relative popularity of various platforms according to the Nielsen data. The first column shows the number of title-week observations in the top 30 by streaming platform. The second column shows fraction of household viewing time by platform. Because Netflix and Disney cover the overwhelming majority of viewing, my analysis focuses only on those two platforms. I exclude Amazon because its streaming service is bundled with other benefits (for example, Prime shipping), making it difficult to assess consumers' valuation for only the streaming component.

During the period I study, Netflix and Disney's libraries were disjoint. Table 1.2.2 highlights several key differences between the two platforms. Netflix has substantially more television programs

Table 1.1: Title-week pairs in Nielsen's top 30 from January 2021–June 2022.

Platform	Count	Frac. viewing time
Netflix	1,837	81.0%
Disney	348	12.0%
Amazon	107	4.3%
Hulu	64	1.9%
HBO Max	15	0.5%
Apple TV	9	0.3%

in the Nielsen data than Disney, both in terms of the number of titles and number of title-week observations.[3] Additionally, while Netflix has more movies than Disney in the Nielsen data, Disney's movies are on average much more popular than Netflix's movies. Last, platforms' prices differ: Netflix costs about twice that of Disney.

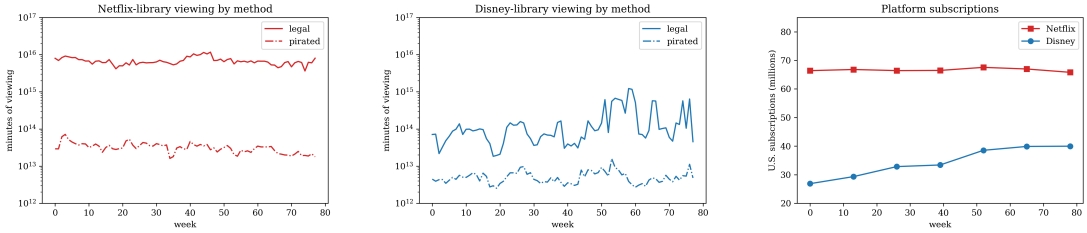
Table 1.2: Platform library characteristics from January 2021–June 2022.

		Netflix	Disney
TV	Unique titles	362	19
	Title-week pairs in Nielsen	2,853	101
	Avg. # wks in Nielsen per title	7.88	5.32
Movies	Unique titles	230	29
	Title-week pairs in Nielsen	481	263
	Avg. # wks in Nielsen per title	2.09	9.07
	Average monthly price	\$14.45	\$7.82

The first two panels of Figure 1.2.2 shows the total number of minutes watched each week, separate by platform library and method of viewing (legally streamed or pirated). In aggregate, households spend substantially more time legally streaming than pirating, so I draw the figure's y-axis with a log scale. For each platform library, legal viewing and pirated viewing follow similar patterns, just with different magnitudes. This suggests that legal streaming and pirated viewing are driven by similar sets of titles. The third panel of Figure 1.2.2 shows the number of subscriptions per quarter. Netflix subscriptions are relatively stable while Disney subscriptions increase by roughly 50% over the 18-month period.

[3]Movies have only one release, but television shows are released over time, often with multiple seasons. Because households may value old and new seasons of a television program differently, I log them as separate "titles." I define a *title* as either a movie, a single season of a television program, or a bundle of seasons of a television program. I group all season that premiered before January 2020 of a television program into one title. Each subsequent season of a television program is its own title.

Figure 1.1: Viewing and subscriptions for Netflix and Disney from January 2021–June 2022.



1.3 Model

To understand households' willingness to pay for content, I develop a model of consumer demand in the form of a discrete choice model for legally streaming titles, pirating titles, and subscribing to streaming platforms. Let j denote a title, which is either a movie or television show. I distinguish between two types of viewing: legal streaming (denoted by ℓ) and pirating (denoted by p). I assume that there are two platforms A and B.[4] In each period t, there exists a collection of titles that can be legally streamed on platform A, $\mathcal{L}^{\ell}_{At}$, a collection of titles that can be legally streamed on platform B, $\mathcal{L}^{\ell}_{Bt}$, and a collection of titles that can be pirated $\mathcal{L}^{p}_{t}$.[5]

I study two types of households: those willing to pirate and those unwilling to pirate. A willing-type household can both legally stream and pirate titles. As explained later in this section, a household's utility of watching title j can differ based on whether they legally stream or pirate it. An unwilling-type can only legally stream titles.

Time occurs in discrete periods $t = 1, \ldots, T$.[6] At the beginning of each period, households purchase platform subscriptions. Household i's choice in period t is denoted by $c_{it} \in \mathcal{K} = \{\varnothing, A, B, AB\}$, where $\varnothing$ denotes subscribing to neither, A denotes subscribing only to A, B denotes only subscribing to B, and AB denotes subscribing to both. Given subscriptions, households watch titles either via legally streaming, pirating (if they are a willing type), or both. Let $h_{i,t-1}$ denote the history of titles household i has watched by the end of period $t-1$. I assume that households do not rewatch titles. My data's panel runs for 18 months, a sufficiently short timespan such that it is unlikely that a large fraction of titles are rewatched. I relax the no-rewatch assumption for titles with more views than the number of the corresponding platform's subscribers.[7]

While pirating a title is free, it has many drawbacks—for example, poor video and audio quality

[4] While this analysis could be done for more than two platforms, for the purposes of the application, I focus on the two dominant platforms: Netflix and Disney+.

[5] For presentational simplicity, I present the model as if titles are available throughout the entire period. In reality, titles may not be available for the entire period, which I address in the estimation section.

[6] In the subscription data, a period is three months, which is what I use for estimation.

[7] See Section 1.4.

and an increased risk of getting a computer virus. Hence, I assume that households do not pirate titles available on their subscribed platforms. I denote a willing household i's composite library in period t by L_{it}. Specifically, $L_{it} = \left(L_{it}^{\ell} \cup L_{it}^{p}\right) \setminus h_{i,t-1}$, where $L_{it}^{\ell} = \cup_{k \in c_{it}} \mathcal{L}_{kt}^{\ell}$ is the set of titles available on their subscribed platforms and $L_{it}^{p} = \mathcal{L}_{t}^{p} \setminus L_{it}^{\ell}$ is the set of titles they can pirate in period t but are not available on their subscribed platforms. Thus, $L_{it}^{\ell} \cap L_{it}^{p} = \varnothing$. An unwilling-type household cannot pirate, so their composite library is simply $L_{it} = L_{it}^{\ell} \setminus h_{i,t-1}$.

1.3.1 Preferences for Titles

Before getting to households' preferences for subscription bundles, I start with their preferences for titles. I follow from the canonical multinomial logit model as in McFadden et al. (1973). Household i's utility of *legally streaming* title j in period t is:

$$u_{ijt}^{\ell} = \underbrace{x_{jt}'\beta + \xi_{jt}^{\ell}}_{\delta_{jt}^{\ell}} + \varepsilon_{ijt}.$$

The x_{jt} term is a vector of title characteristics, which include genre, runtime, quality observed by the econometrician (given by IMDb score), number of weeks since the title's release, etc. The ξ_{jt}^{ℓ} term captures the objective quality for streaming title j that is observed by the consumer but not by the econometrician, and the ε_{ijt} term is an independently and identically distributed type I extreme value random variable. Like many logit models, $\delta_{jt}^{\ell} \equiv x_{jt}'\beta + \xi_{jt}^{\ell}$ captures the portion of utility common to all households, and ε_{ijt} captures households' idiosyncratic shocks. Note that a household's utility of streaming a title does not depend on the platform. As shown in the next subsection, I include subscription-bundle fixed effects at the platform-choice level to capture differences in platform characteristics.

A willing household i's utility of *pirating* title j in period t is:

$$u_{ijt}^{p} = \underbrace{x_{jt}'\beta + d + \xi_{jt}^{p}}_{\delta_{jt}^{p}} + \varepsilon_{ijt}.$$

The utility of pirating a title versus streaming a title is distinct in two ways. First, the term $d < 0$ captures willing households' average disutility of pirating. Second, title qualities observed by consumers but not by the econometrician may differ by viewing method. In addition to capturing preferences for the title's *content*, ξ_{jt}^{p} also captures a pirated title's idiosyncratic difference in piracy disutility. For example, pirated copies of title j may be of lower-than-average quality, offer limited

options to add or remove subtitles, etc. Last, it is without loss for ε_{ijt} to depend only on title, and not on viewing method, because a household does not simultaneously have a title's legal version and pirated version in their composite library.

I take a revealed preference approach and assume that, in a given period, households prefer watched titles to unwatched titles. The outside option $j = 0$ is to forgo watching titles, legally or pirated. A household watches titles in its composite library that are better than the outside option and does not watch those that are worse. I normalize $u_{i0t} = \varepsilon_{i0t}$.

This assumes that a household's decision to watch title j is independent of watching some other title j', ruling out time constraints. Each period is relatively long (three months), so time constraints and dependence across titles are not first-order concerns.

1.3.2 Preferences for Platforms

Households subscribe to platforms at the beginning of each period. I abstract away from quantity discounts and different tiers of services. I assume that each platform offers only one tier and one price for all consumers. On both platforms I study, all tiers gave users access to the same content, which is a critical dimension in which platforms differ.

Households choose the bundle that maximizes their utility. Household i's utility of subscribing to bundle $k \in \mathcal{K}$ at the beginning of period t is

$$v_{ikt} = \alpha_{ev} EV_{ikt} + \alpha_p p_{kt} + \alpha_k + \tilde{\xi}_{kt} + \tilde{\varepsilon}_{ikt},$$

where EV_{ikt} is household i's expected value of content under bundle k minus that of the outside option, p_{kt} is the total price of subscribing to the platforms in bundle k, α_k is a bundle-k fixed effect, $\tilde{\xi}_{kt}$ is the bundle's quality observed by consumers but not the econometrician, and $\tilde{\varepsilon}_{ikt}$ is an idiosyncratic shock independently and identically drawn from a type-I extreme value distribution. The outside option $k = \varnothing$ of subscribing to no platforms has a utility of $v_{i\varnothing t} = \tilde{\varepsilon}_{i\varnothing t}$.

A household's utility of content is the sum of watched titles' utilities times their respective runtimes. Specifically, let $L = L^\ell \cup L^p$ denote a hypothetical composite library of legal titles L^ℓ and a disjoint set of pirated titles L^p. Assume that willing household i has seen no titles in L. Then

$$V_{it}(L) = \sum_{j \in L^\ell} \left[\mathbf{1}(u^\ell_{ijt} > u_{i0t}) u^\ell_{ijt} y_j \right] + \sum_{j \in L^p} \left[\mathbf{1}(u^p_{ijt} > u_{i0t}) u^p_{ijt} y_j \right] + u_{i0t} * (\text{time spent on outside option}),$$

where y_j is title j's runtime in hours. An unwilling-type household's value is analogous but without

pirated titles. Taking expectations yields the library's expected value of content:

$$\mathbf{E}\left[V_{it}(L)\right] = \sum_{j \in L^{\ell}} \left[y_j \log(1 + e^{\delta^{\ell}_{jt}}) \right] + \sum_{j \in L^{p}} \left[y_j \log(1 + e^{\delta^{p}_{jt}}) \right] + H\gamma_{\text{Euler}}$$

where γ_{Euler} is Euler's constant and H is the number of hours in a period. This model assumes that households know the entire period's titles when purchasing bundles. Television follows a predictable annual schedule and movie trailers are released months in advanced of release, so this assumption is reasonable.

1.4 Estimation

1.4.1 Model Discussion and Identification

Using aggregate data on platform subscriptions and title viewing, I estimate parameters $\theta = (\alpha, \beta, \gamma)$, where α captures platform-bundle preferences, β captures title preferences, and γ is the fraction of willing-type households. I employ a nested fixed-point method to invert viewing and platform-bundle shares. For a given (α_{ev}, γ) pair, I solve a system of equations to deduce what fraction of subscriptions and legal streams come from each household type. I then invert those shares conditional on type to recover title and platform mean utilities, which I feed into a standard GMM framework to estimate α and β. I loop over (α_{ev}, γ) pairs until I find that which minimizes the GMM objective.

An important identifying assumption in the model is that households do not rewatch titles. Because my title-viewing share data are aggregated at the national level, the no-rewatch assumption allows me to pin down what fraction of households watch each title. Households make bundle-subscription choices with their viewing histories in mind. However, my data contain 660 titles across six periods, so the set of all possible viewing histories creates a high-dimensional discrete state space that is computationally infeasible to work with.

To make estimation computationally feasible, I employ a state-space simplification in which I estimate the model by grouping together households of the same type. I denote willing-type households by w and unwilling-type households by n. Within each type, I track what fraction of households watch each title in each period and assume that households make platform-bundle and viewing decisions based on the *probability* their type of household has already seen a title, rather than the binary outcome of having seen it or not. Suppose a type-h household, where $h \in \{w, n\}$, can watch title j in period t via method $m \in \{\ell, p\}$. Let $s^{m,h}_{j\tau}$ be the share of type-h households that watch title j in period τ via method m. The probability the household watches the title in period t

is

$$\underbrace{\left(1 - \sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,h}\right)}_{\substack{\text{Prob. type-}h \\ \text{has not yet seen } j}} \left(\frac{e^{\delta_{jt}^m}}{1 + e^{\delta_{jt}^m}}\right).$$

The first term is the probability the household has not yet seen the title. The second term is the probability that the title's utility exceeds that of the outside option. Consequently, type-h household's expected value of some composite library $L = L^\ell \cup L^p$ is

$$\begin{aligned}\mathbf{E}\left[V_t^h(L)\right] = &\sum_{j \in L^\ell} \left[\left(1 - \sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,h}\right) y_j \log(1 + e^{\delta_{jt}^\ell})\right] \\ &+ \mathbf{1}(h = w) \sum_{j \in L^p} \left[\left(1 - \sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,h}\right) y_j \log(1 + e^{\delta_{jt}^p})\right] + H\gamma_{\text{Euler}}.\end{aligned}$$

This expected-value formula mirrors that in the previous section but now accounts for the probability of having watched the title.

While the state-space simplification departs from reality, it maintains the important feature that as households watch more titles in a platform's library, they have a lower expected value of subscribing to that platform. In the extreme cases of all or no type-h households having seen a title, the simplified model's predictions coincide with those of a richer model that tracks households' viewing histories.

To deduce preferences from the data, one must know what fraction of legal streams and platform-bundle subscriptions come from each household type. However, the data do not contain household-type information. In what follows, I develop a nested fixed-point procedure that uses aggregate data to estimate (1) the fraction of willing-type households and (2) title and platform-bundle shares conditional on household type. After running this nested fixed-point procedure, estimating title and platform-bundle mean utilities is akin to that in standard logit models (McFadden et al. (1973)).

Let r_{kt}^h be the fraction of type-h households that subscribe to bundle k in period t and $r_{kt}^\bullet$ be the share of households that subscribe to bundle k in period t, unconditional on type.

The variables of interest for the nested fixed-point procedure are α_{ev}, γ, and probabilities $\pi = (\pi_{kt})_{k \in \mathcal{K}, t \in \{1,\ldots,T\}}$, where π_{kt} is the fraction of period-t, bundle-k subscribers who are willing-type households. These probabilities are not in the data but are important for estimation and the counterfactual exercises I perform later.

There is an outer loop and an inner loop. In the outer loop, I search over values of (α_{ev}, γ). In the

inner loop, I solve for the values of $\pi_t = (\pi_{\varnothing t}, \pi_{At}, \pi_{Bt}, \pi_{ABt})$ such that (1) each platform-bundle's mean utility predicted by share data matches that predicted by the model and (2) the implied fraction of willing households is γ. Formally, for each period t, I numerically solve the following nonlinear platform-bundle system of equations:

$$\begin{aligned}
&\underbrace{\log(r^n_{kt}(\pi_t)) - \log(r^n_{\varnothing t}(\pi_t))}_{\text{share-predicted } \tilde{\delta}^n_{kt} - \tilde{\delta}^n_{\varnothing t}} - \underbrace{\left[\log(r^w_{kt}(\pi_t)) - \log(r^w_{\varnothing t}(\pi_t))\right]}_{\text{share-predicted } \tilde{\delta}^w_{kt} - \tilde{\delta}^w_{\varnothing t}} \\
&\quad = \underbrace{\alpha_{ev}\left[EV^n_{kt}(\pi_t) - EV^n_{\varnothing t}(\pi_t)\right]}_{\text{model-predicted } \tilde{\delta}^n_{kt} - \tilde{\delta}^n_{\varnothing t}} - \underbrace{\alpha_{ev}\left[EV^w_{kt}(\pi_t) - EV^w_{\varnothing t}(\pi_t)\right]}_{\text{model-predicted } \tilde{\delta}^w_{kt} - \tilde{\delta}^w_{\varnothing t}} \quad \forall k \in \{A, B, AB\} \\
&\gamma = \pi_{\varnothing t} r_{\varnothing t} + \pi_{At} r_{At} + \pi_{Bt} r_{Bt} + \pi_{ABt} r_{ABt},
\end{aligned}$$

where $r^w_{kt}(\pi_t) = \frac{\pi_{kt} r^\bullet_{kt}}{\sum_{k'} \pi_{k't} r_{k't}}$ and $r^n_{kt}(\pi_t) = \frac{(1-\pi_{kt}) r^\bullet_{kt}}{\sum_{k'} (1-\pi_{k't}) r^\bullet_{k't}}$. For both household types, the mean utility of the outside option is normalized to $\delta^h_{\varnothing t} = 0$.

Once I have values of π_{kt} that solve the above system of equations, I can do standard logit share inversion to get platform-bundle mean utilities:

$$\tilde{\delta}^h_{kt} = \log(r^h_{kt}(\pi_t)) - \log(r^h_{\varnothing t}(\pi_t)) \quad \forall h \in \{w, n\}, k \in \{A, B, AB\}, t \in \{1, \ldots, T\}.$$

To infer mean utilities for titles, one must disentangle what fraction of legal streams come from each household type, as the share data includes information about the method of viewing but not household type. Given values of π_{kt}, one can solve a legal-streaming system of equations to deduce what fraction of title j legal streams come from willing- versus unwilling-type households. For presentational simplicity, I drop the π_t terms that determine the values of r^h_{kt}. The resulting system of equations is as follows:

$$\begin{aligned}
s^{\ell,\bullet}_{jt} &= \gamma s^{\ell,w}_{jt} + (1-\gamma) s^{\ell,n}_{jt} \\
\frac{s^{\ell,w}_{jt}}{\left(1 - \sum_{\tau=1}^{t-1} s^{\bullet,w}_{j\tau}\right)\left(\sum_{k: j \in \mathcal{L}^\ell_{kt}} r^w_{kt}\right)} &= \frac{s^{\ell,n}_{jt}}{\left(1 - \sum_{\tau=1}^{t-1} s^{\bullet,n}_{j\tau}\right)\left(\sum_{k: j \in \mathcal{L}^\ell_{kt}} r^n_{kt}\right)}.
\end{aligned}$$

The first equation follows from the law of total probability, conditioning on household type. The second equation follows from willing and unwilling types having the same mean utility of legally streaming title j in period t. Thus, the fraction of willing and unwilling households that legally stream title j is equal conditional on having access to it and not yet having seen it. Solving this system of equations yields the fraction of willing-type households that watched title j in period t

only in terms of the data, π, and γ:

$$s_{jt}^{\ell,w} = \frac{s_{jt}^{\ell,\bullet}}{\gamma + (1-\gamma)\frac{\left(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,n}\right)\left(\sum_{k:j\in\mathcal{L}_{kt}^{\ell}} r_{kt}^{n}\right)}{\left(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,w}\right)\left(\sum_{k:j\in\mathcal{L}_{kt}^{\ell}} r_{kt}^{w}\right)}}.$$

Calculating the share of willing households that pirated title j is considerably easier since only willing households can pirate. Specifically, $s_{jt}^{p,w} = \frac{s_{jt}^{p,\bullet}}{\gamma}$ and $s_{jt}^{p,n} = 0$.

With watch probabilities conditional on household type, one can deduce title mean utilities implied by shares:

$$\delta_{jt}^{\ell} = \log\left(\frac{s_{jt}^{\ell,w}}{(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,w})\sum_{k:j\in\mathcal{L}_{kt}^{\ell}} r_{kt}^{w}}\right) - \log\left(1 - \frac{s_{jt}^{\ell,w}}{(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,w})\sum_{k:j\in\mathcal{L}_{kt}^{\ell}} r_{kt}^{w}}\right)$$
$$\delta_{jt}^{p} = \log\left(\frac{s_{jt}^{p,w}}{(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,w})\sum_{k:j\in L_{kt}^{p}} r_{kt}^{w}}\right) - \log\left(1 - \frac{s_{jt}^{p,w}}{(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,w})\sum_{k:j\in\mathcal{L}_{kt}^{p}} r_{kt}^{w}}\right)$$

Given values for δ_{jt}^{ℓ} and δ_{jt}^{p}, one can calculate the expected value of each platform bundle $EV_{kt}^{h}(\pi)$ and confirm whether the platform-bundle system of equations is satisfied.[8]

The above assumes that titles are not rewatched; however, there are ten titles with more views than corresponding platform subscribers.[9] I call these titles *rewatchable* and allow households to rewatch these titles with the same utility for each viewing. I estimate these titles' mean utilities δ_{jt}^{m} *as if* no household has watched them in a previous period.[10] In the set of title characteristics, I include an indicator for whether a title is rewatchable to capture these titles' strong desirability.

The identification of parameters capturing title and platform preferences is relatively standard. For titles, there is substantial variation in genre, quality, and release times, identifying those parameters in β. For platforms, Disney has large increases in subscriptions and large increases in new titles, helping to identify households' sensitivity to library quality α_{ev}. The coefficient on price, however, is weakly identified given that each platform only has one price change during the study

[8] I write δ_{jt}^{ℓ} in terms of the fraction of willing-type households that watch the title. This system of equations is exactly identified, so the value of δ_{jt}^{ℓ} would be the same if written in terms of the fraction of unwilling-type households.

[9] Those titles are *The Book of Boba Fett*, *Cocomelon*, *Criminal Minds*, *Cruella*, *Encanto*, *Frozen II*, *Luca*, *Marmaduke*, *Raya and the Last Dragon*, and *Turning Red*.

[10] Mathematically, if title j is rewatchable, then

$$\delta_{jt}^{\ell} = \log\left(\frac{s_{jt}^{\ell,w}}{\sum_{k:j\in\mathcal{L}_{kt}^{\ell}} r_{kt}^{w}}\right) - \log\left(1 - \frac{s_{jt}^{\ell,w}}{\sum_{k:j\in\mathcal{L}_{kt}^{\ell}} r_{kt}^{w}}\right)$$
$$\delta_{jt}^{p} = \log\left(\frac{s_{jt}^{p,w}}{\sum_{k:j\in L_{kt}^{p}} r_{kt}^{w}}\right) - \log\left(1 - \frac{s_{jt}^{p,w}}{\sum_{k:j\in\mathcal{L}_{kt}^{p}} r_{kt}^{w}}\right).$$

I also remove the $(1-\sum_{\tau=1}^{t-1} s_{j\tau}^{\bullet,w})$ component from these titles' contributions to a platform's expected value.

period. Hence, I estimate the model both with and without price.

The identification of the fraction of willing-type households, γ, and piracy disutility, d, is more nuanced. These parameters are identified by joint changes in pirated viewing, legal viewing, and platform subscriptions. For example, suppose a platform gains subscribers. If, following the gain, the platform's titles are pirated less often, those new subscribers were likely pirating that platform's titles prior to subscribing. If the platform's titles have no change in pirated viewing, the new subscribers likely were not pirating its titles.

1.4.2 Estimation Strategy

Because my title-viewing data are at the weekly level, I allow title mean utilities to vary week to week. This gives more granular measurements of households' viewing preferences than if I aggregated viewing up to the quarterly level. A household's expected value of a library is the sum of her expected value of viewing over all weeks in the period. Let E_t be the set of weeks in period t. Then a type-h household's expected value of library L_{kt} is

$$\mathbf{E}\left[V_t^h(L_{kt})\right] = \sum_{e \in E_t} \sum_{j \in L_{kt}} \left(1 - \sum_{\epsilon < e} s_{j\epsilon}^h\right) h_j \log(1 + e^{\delta_{je}^m}) + H\gamma_{\text{Euler}},$$

where δ_{je}^m is the mean utility for watching title j in week e via method m.

I estimate the model using 2-step GMM. I assume that all title characteristics are exogenous; thus, there is no need for instruments. Platform utility has an endogenous object, namely price. However, there is very little price variation, as each platform only has one price change.[11] With only two price changes, I do not instrument. Because of the weak identification, I estimate the model under specifications with and without price. The moments I use for estimation are as follows

$$g(\theta) = \begin{pmatrix} \mathbf{E}\left[X_{je}^{m\prime} \xi_{je}^m(\theta)\right] \\ \mathbf{E}\left[\tilde{X}_{kt}^{h\prime} \tilde{\xi}_{kt}^h(\theta)\right] \end{pmatrix} = 0.$$

That is, title and platform unobserved qualities are orthogonal to title and platform characteristics, respectively.

Following standard 2-step GMM, I solve for the parameters that minimize the GMM objective $q_0(\theta) = \overline{g}(\theta)' W_0 \overline{g}(\theta)$, where $\overline{g}(\theta)$ are the sample analogues to the moments above. I let W_0 be the 2-by-2 identity matrix. After finding parameters that minimize the objective $\hat{\theta}_0$, I update the weight

[11] Disney's monthly subscription went from \$6.99 to \$7.99 in March 2021, and Netflix's went from \$13.99 to \$15.49 in January 2022.

matrix to be $W_1 = \left(\overline{g}(\hat{\theta}_0)'\overline{g}(\hat{\theta}_0)\right)^{-1}$ and re-estimate the model to get final estimates $\hat{\theta}$.

Note that if one omitted piracy in estimating consumer demand, estimates would be biased. This is because omitting piracy ignores an avenue by which consumers value a platform's content but do not subscribe to it. In the next section, I include demand estimates for specifications both with and without piracy.

1.5 Results

Table 1.3 shows the estimates for preferences over titles. There are four specifications. The first two specifications estimate the model without piracy, which uses only the legal viewing and subscription data and assumes that $\gamma = 0$. The second two specifications estimate the full model, which incorporates piracy. For both the restricted model without piracy and the full model with piracy, I estimate title preferences with and without quarterly fixed effects, which allow for the possibility that households had different propensities to watch over time. Including quarterly fixed had little effect on the results. The estimates with and without piracy have some differences (for example, on genre indicators) but generally are quite similar. All four specifications show that households have strong preferences for newer titles. As expected, a title's utility decreases as it ages. Additionally, households tend to prefer higher-quality titles, which is also expected.

The estimated piracy disutility is substantial. In utility terms, pirating a brand-new title is similar to legally streaming a title that is than 21 weeks old, holding all else equal. This is quite large given households' strong preferences for new content: 95% of titles are in the top 30 for less than 21 weeks, as shown by Figure 1.2. This suggests that even among willing-type households, piracy is relatively undesirable and not as strong a substitute for platform subscriptions as one might initially expect.

There are four platform-estimate specifications. Again, I look at two specifications for a restricted model without piracy and two specifications for the full model with piracy. For each of these models, I estimate platform preferences with and without price. I do so because each platform only has one price change, so the coefficient on price is weakly identified. As shown in Table 1.4, the estimate for the price coefficient is insignificant both with and without piracy, so I focus on the specification without price.

The coefficients on library expected value differ across models with and without piracy. This is because ignoring consumers' option to pirate biases preference estimates. Assuming that non-subscribers do not access a library's content overestimates consumers' sensitivity to platform quality.

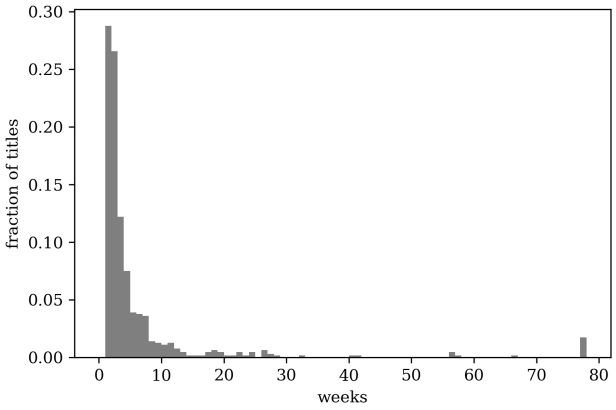
Table 1.3: Title preference estimates.

	Without piracy		With piracy	
	(1)	(2)	(3)	(4)
constant	-4.79	-4.41	-4.69	-4.63
	(0.72)	(0.73)	(0.29)	(0.29)
tv indicator	-2.14	-2.14	-3.18	-3.18
	(0.06)	(0.06)	(0.03)	(0.03)
rewatchable	4.90	4.62	5.44	5.38
	(0.74)	(0.75)	(0.29)	(0.29)
animated	1.53	1.28	1.40	1.36
	(0.96)	(0.96)	(0.44)	(0.43)
documentary	2.62	2.34	-1.73	-1.74
	(1.26)	(1.26)	(0.47)	(0.47)
action	-0.04	-0.27	-0.41	-0.46
	(0.77)	(0.77)	(0.31)	(0.31)
romance	1.62	1.40	-0.04	-0.10
	(0.84)	(0.85)	(0.35)	(0.35)
drama	3.88	3.68	0.01	-0.03
	(0.83)	(0.83)	(0.30)	(0.30)
animated*IMDb	2.65	2.58	1.93	1.91
	(0.93)	(0.93)	(0.51)	(0.51)
documentary*IMDb	-0.37	-0.37	5.15	5.11
	(1.47)	(1.47)	(0.52)	(0.52)
action*IMDb	3.56	3.50	5.98	5.97
	(0.35)	(0.34)	(0.18)	(0.18)
romance*IMDb	0.87	0.80	4.79	4.79
	(0.64)	(0.64)	(0.30)	(0.30)
drama*IMDb	-1.63	-1.74	4.93	4.91
	(0.57)	(0.57)	(0.17)	(0.17)
comedy*IMDb	4.30	3.95	3.14	3.07
	(0.94)	(0.94)	(0.40)	(0.40)
1 week after release	-0.17	-0.18	-0.62	-0.61
	(0.13)	(0.13)	(0.10)	(0.10)
2 weeks after release	-0.82	-0.82	-1.11	-1.11
	(0.16)	(0.17)	(0.11)	(0.11)
3 weeks after release	-1.45	-1.44	-1.48	-1.48
	(0.22)	(0.22)	(0.11)	(0.11)
4 weeks after release	-1.88	-1.87	-1.79	-1.79
	(0.23)	(0.23)	(0.11)	(0.11)
5 weeks after release	-2.22	-2.21	-2.13	-2.13
	(0.23)	(0.23)	(0.11)	(0.11)
6 weeks after release	-2.30	-2.29	-2.26	-2.26
	(0.21)	(0.22)	(0.11)	(0.11)
7 weeks after release	-2.60	-2.59	-2.46	-2.46
	(0.22)	(0.22)	(0.11)	(0.11)
8 weeks after release	-2.71	-2.70	-2.61	-2.61
	(0.21)	(0.21)	(0.11)	(0.11)
9 weeks after release	-2.77	-2.76	-2.71	-2.71
	(0.22)	(0.23)	(0.11)	(0.11)
Quarter dummies	No	Yes	No	Yes
N	8,960	8,960	40,349	40,349
R-squared	0.371	0.373	0.544	0.544

	Without piracy		With piracy	
	(1)	(2)	(3)	(4)
10 weeks after release	-2.80	-2.78	-2.77	-2.77
	(0.23)	(0.23)	(0.12)	(0.12)
11 weeks after release	-3.00	-2.98	-2.97	-2.96
	(0.23)	(0.23)	(0.12)	(0.12)
12 weeks after release	-2.99	02.96	-3.00	-3.00
	(0.23)	(0.23)	(0.12)	(0.12)
13 weeks after release	-3.07	-3.04	-3.05	-3.05
	(0.23)	(0.23)	(0.12)	(0.12)
14 weeks after release	-3.19	-3.15	-3.13	-3.13
	(0.24)	(0.24)	(0.12)	(0.12)
15 weeks after release	-3.20	-3.16	-3.14	-3.14
	(0.24)	(0.24)	(0.12)	(0.12)
16 weeks after release	-3.11	-3.08	-3.21	-3.22
	(0.24)	(0.24)	(0.12)	(0.12)
17 weeks after release	-3.04	-3.01	-3.20	-3.20
	(0.23)	(0.23)	(0.12)	(0.12)
18 weeks after release	-3.24	-3.20	-3.34	-3.34
	(0.24)	(0.24)	(0.12)	(0.12)
19 weeks after release	-3.35	-3.32	-3.42	-3.42
	(0.23)	(0.23)	(0.12)	(0.12)
20 weeks after release	-3.58	-3.54	-3.48	-3.48
	(0.26)	(0.26)	(0.12)	(0.12)
21 weeks after release	-3.60	-3.56	-3.49	-3.49
	(0.28)	(0.26)	(0.13)	(0.13)
22 weeks after release	-3.65	-3.61	-3.56	-3.56
	(0.28)	(0.27)	(0.13)	(0.13)
23 weeks after release	-3.75	-3.72	-3.69	-3.69
	(0.27)	(0.27)	(0.13)	(0.13)
24 weeks after release	-3.83	-3.82	-3.71	-3.70
	(0.27)	(0.27)	(0.13)	(0.13)
25 weeks after release	-3.74	-3.69	-3.72	-3.72
	(0.26)	(0.26)	(0.13)	(0.13)
26 weeks after release	-3.83	-3.78	-3.74	-3.74
	(0.27)	(0.27)	(0.14)	(0.14)
27 weeks after release	-3.79	-3.74	-3.71	-3.70
	(0.29)	(0.29)	(0.14)	(0.14)
28 weeks after release	-3.74	-3.68	-3.74	-3.74
	(0.27)	(0.27)	(0.13)	(0.13)
29 weeks after release	-3.77	-3.72	-3.81	-3.80
	(0.28)	(0.28)	(0.13)	(0.13)
30 weeks after release	-3.75	-3.69	-3.82	-3.81
	(0.27)	(0.27)	(0.13)	(0.13)
31+ weeks after release	-4.18	-4.14	-4.56	-4.56
	(0.07)	(0.07)	(0.06)	(0.06)
d			-3.51	-3.52
			(0.03)	(0.03)
Quarter dummies	No	Yes	No	Yes

Notes: IMDb is the title's IMDb score normalized to a scale of 0 to 1, and d is an indicator for whether a title was pirated. Columns (1)-(2) estimate the model without piracy hence only use the legal viewing and subscription data. Columns (3)-(4) estimate the full model, which includes piracy. Title estimates are almost identical across platform specifications, so for brevity, all columns exclude price from platform preferences (which, as shown in Table 1.4, is insignificant). All standard errors are robust.

Figure 1.2: Number of weeks titles spend in the top 30.



As shown in Table 1.4, the model without piracy overestimates the coefficient on a platform expected value by 15% in the specification without price and by 6% in the specification with price. Turning to the model with piracy, estimates are as expected. The coefficient on the expected value of the library is significant and positive, indicating that households prefer better libraries. Additionally, all specifications have the same ordering of platform-bundle fixed effects.

Having a fixed effect for the "both" option captures the degree of complementarity between the two platforms. While it is initially surprising that the fixed effect for only Netflix is greater than that for both platforms, having to pay for both platforms can lower the fixed effect. Additionally, there could be some degree of substitutability between subscribing to both platforms. For example, a household may not have enough time to take full advantage of having access to both libraries.

One can directly use the estimates for π_{kt} calculated during estimation to disentangle platform revenues by household type. Willing-type households subscribe to fewer platforms than unwilling-type households. In total, I find that the average willing-type household provides 2.9% lower revenue to platforms than the average unwilling-type household. Moreover, there is an asymmetry across platforms in this revenue gap between household types. Willing-type households subscribed to Netflix at a rate 2.5% lower than unwilling-type households and 4.2% lower for Disney. This gap comes from the difference in library sizes. Disney's library is smaller. Thus, pirating Disney's library fewer instances of incurring the piracy disutility than pirating Netflix's library, making pirating Disney more favorable than pirating Netflix. In the next section, I explore how losses change with changes in piracy disutility.

Table 1.4: Platform preference estimates.

	Without piracy		With piracy	
	(1)	(2)	(3)	(4)
EV	0.0099	0.0083	0.0086	0.0078
	(0.0058)	(0.0067)	(0.0028)	(0.0035)
Netflix FE	-0.3010	-0.4627	-0.2478	-0.4172
	(0.2467)	(0.8346)	(0.1227)	(0.4430)
Disney FE	-1.9295	-2.0316	-1.9199	-2.0187
	(0.1020)	(0.5082)	(0.0500)	(0.2531)
Both FE	-0.6281	-0.8919	-0.5655	-0.8338
	(0.3426)	(1.3600)	(0.1597)	(0.6923)
price		0.0053		0.0046
		(0.0231)		(0.0117)
γ			0.1114	0.1114
N	18	18	36	36
R-squared	0.984	0.984	0.985	0.986

Notes: Columns (1)-(2) estimate the model without piracy hence only use the legal viewing and subscription data. Columns (3)-(4) estimate the full model that includes piracy. All standard errors are robust.

1.6 Counterfactual: Changing Piracy Disutility

I use the above estimates to calculate how subscriptions would change with piracy disutility. Specifically, I simulate willing households' viewing and subscription decisions for a range of counterfactual values for piracy disutility d. Running this counterfactual informs us of how households would react to changes in access to pirated titles. For example, if piracy were more heavily monitored and popular sites were shut down, pirated copies of titles would be more difficult to find, effectively increasing d.

I restrict my analysis to local changes in piracy disutility and assume that the fraction of willing-type households and platform prices are held constant. For large changes in piracy access, one could imagine that the set of pirates and platform prices would change. However, for small changes in d, holding these values constant is less egregious. Platform prices are quite sticky, so small changes in piracy access likely would not affect them.[12] Additionally, the fraction of willing-type households should not change with small changes in d.

Mathematically, I simulate willing households' decisions in which I add a constant Δ to δ^p_{je} for Δ on a grid from $\frac{-1}{4}\hat{d}$ to $\frac{1}{4}\hat{d}$. I leave the mean utilities of legally streaming titles, δ^ℓ_{je}, unchanged. Note that I can use titles' share-inverted mean utilities directly, which incorporates the utility

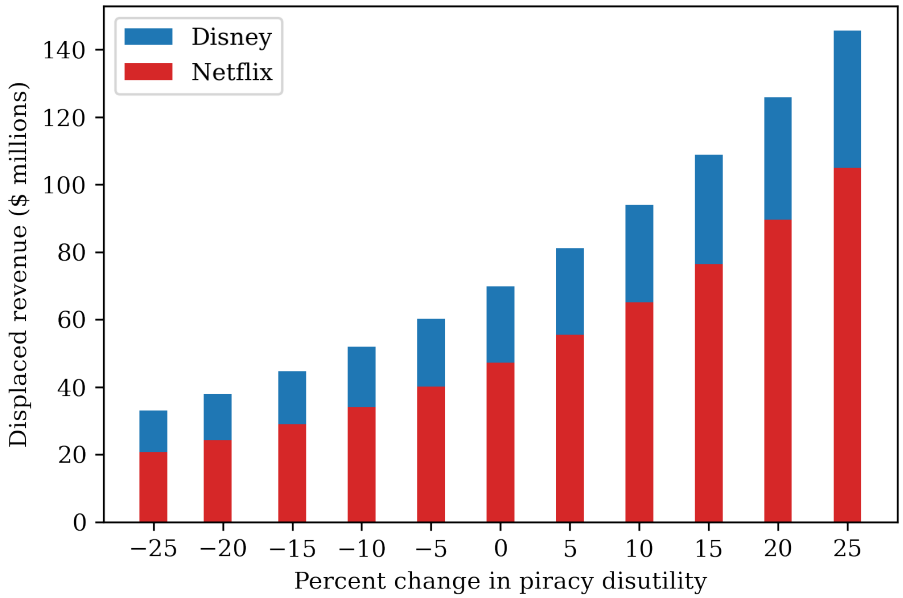
[12]Moreover, I do not have supply-side data that would allow me to estimate platforms' responses.

components observed to households but not the econometrician. For each counterfactual piracy disutility, I calculate the value of each library and use platform estimates $\hat{\alpha}$ and $\hat{\hat{\xi}}_{kt}$ to estimate what fraction of households would subscribe to each bundle. This requires numerically solving a fixed point of how many households subscribe to each bundle. A household's value of a library depends on the probability they have seen each title in the library each week, which depends on the number of like-type subscribers, which depends on the value of the library. The fixed-point procedure is similar to that for platform estimation, so I relegate details to the appendix.

Figure 1.6 illustrates results from the counterfactual. Each bar corresponds to the displaced revenue for a given counterfactual value of d. The bars are shaded by the each platform's share of the losses. The x-axis shows the counterfactual d as a percent difference from the estimated piracy disutility. A negative percentage means that piracy is made less appealing.

There is an asymmetry in how total displaced revenues change with piracy disutility. Making piracy 25% more appealing would increase total displaced revenue to $148 million during the period I study; making piracy 25% less appealing would decrease total displaced revenue to $33 million. Thus, although piracy seems to have relatively small effects, easing piracy policing could generate substantial losses.

Figure 1.3: Displaced revenue for counterfactual piracy disutilities.



As mentioned in Section 1.5, at the estimated piracy disutility, there is an asymmetry across platforms in the revenue gap between household types. This asymmetry persists with changes

in piracy disutility. Specifically, if piracy were made 25% less appealing, willing-type households would subscribe to Netflix and Disney 1.3% and 2.6% less often than unwilling-type households, respectively. If piracy were made 25% more appealing, willing-type households would subscribe to Netflix and Disney 5.5% and 7.6% less often than unwilling-type households, respectively. Again, this occurs because Disney has a smaller library; thus, pirating Disney's library incurs the piracy disutility less often than pirating Netflix's library.

1.7 Conclusion

To assess piracy's losses on the streaming industry, I merge novel title-level data on pirated viewing, legal viewing, and streaming subscriptions to estimate a structural demand model for titles and platforms. Importantly, my model accounts for household heterogeneity in propensity to pirate, and I estimate the fraction of households willing to pirate and their average piracy disutility. To disentangle what fraction of viewings and subscriptions come from households willing to pirate and households unwilling to pirate, I employ a nested fixed-point procedure that matches share- and model-predicted platform average utilities by latent household type.

I estimate that 11% of households are willing to pirate content but that piracy is associated with substantial costs to the pirates themselves. In utility terms, pirating a brand-new title is approximately equal to legally streaming a title that is 21 weeks old, holding all else equal. This is large, as households strongly prefer new content: 95% of titles are in the top 30 most-streamed titles for less than 21 weeks. Consequently, many households willing to pirate still subscribe to streaming platforms. During the period I study, pirates displaced 0.3% of platforms' revenues, equal to roughly $70 million. Through my counterfactual analysis, I find that a 25% increase in piracy disutility would have generated $33 million in displaced revenues while a 25% decrease in piracy disutility would have generated $146 million in displaced revenues. Thus, although the estimated displaced revenues are relatively small, slightly easing piracy policing could generate substantial losses.

This paper is a first step in assessing the displacement of streaming subscriptions. While my analysis focuses on the losses of two streaming platforms, one may also be concerned about the losses to other streaming platforms and other online distribution methods of content such as online video rentals. Additionally, piracy is not the only method by which subscriptions may be displaced. Platforms have recently voiced concerns over password sharing, in which more users utilize an account than intended. My analysis does not capture these other avenues of displaced subscriptions; thus,

my results serve as a lower bound.

Chapter 2

School Choice with Costly Information Acquisition

I study a model of centralized school choice in which students engage in costly search over schools before submitting preference reports to a clearinghouse. I consider three classes of preferences over schools—idiosyncratic, common, and hybrid—and characterize outcomes under two search protocols—simultaneous and sequential. With idiosyncratic preferences, there are no search externalities, and inefficiencies arise only because of uncoordinated search. Common preferences, however, generate search externalities: when high-priority students search, seats available to lower-priority students are adversely selected. Consequently, sequential search generates greater welfare than simultaneous search with idiosyncratic preferences but not necessarily with common. Additionally, with common preferences, welfare is nonmonotonic in search costs. I also show that the search protocol affects outcome inequality in important ways. For both protocols, I provide an instrument by which a designer can break students' indifferences in search strategies to coordinate search and increase welfare.

2.1 Introduction

Many large urban school districts—including New York, Los Angeles, Chicago, and Boston—use centralized clearinghouses to match students to high schools, and many foreign countries—for example, Brazil, China, Germany, and India—use them to match students to universities. There is an abundance of theoretical research on school assignment mechanisms. Most of the literature assumes students have complete information over their preferences.[1] However, numerous empirical studies suggest this may not be the case. Hoxby and Turner (2015) show that information interventions impact students' application and enrollment decisions. Moreover, acquiring information on schools can be costly, as shown by Hastings and Weinstein (2008). To address the tradeoffs of more information with the costs of search, I develop a theoretical model in which students endogenously acquire information about their preferences prior to participating in a centralized clearinghouse.

I study this model along two dimensions: preferences over schools and students' search protocol. For preferences, I consider idiosyncratic and common. Idiosyncratic preferences are relevant if students have heterogeneous preferences over programs offered, extracurriculars, etc. Common preferences capture environments in which students value the same characteristics, for example, academic quality. For each class of preferences, I consider simultaneous search, which applies to environments in which students commit to a tour of school visits, and sequential search, capturing settings in which students have many visit opportunities. I analyze the equilibrium effects preferences, search protocol, and search costs have on student welfare and inequality.

To match students to schools, I employ one of the most commonly used mechanisms: student-proposing deferred acceptance (henceforth DA).[2]

To see the forces in this model, consider a district with two schools where one school is strictly larger than the other. Both schools have the same priorities over a continuum of students, who know their precise rankings. Such priorities could be generated by performance on a standardized test or lottery, as done in many school-choice settings, and informing each student of her ranking.

First, suppose that students' preferences are idiosyncratic. Specifically, let students' utilities for schools be independently drawn from a standard uniform distribution.[3] Because preferences are independent, students do not impose search externalities on each other—each student's search problem is effectively an individual search problem. Because student-proposing DA is strategy-proof, students rank schools from highest utility realization to lowest, treating unvisited schools as their

[1] See next section for a review of related literature.
[2] See Gale and Shapley (1962) for details on the Deferred Acceptance algorithm.
[3] Here I consider a uniform distribution for simplicity. As shown in Section 4, the results generalize.

expected match utilities. As information costs increase, all students' expected values decrease.

Now consider common preferences. Each school has an objective quality drawn independently from a standard uniform distribution. Common values induce correlation across students' reports and alter the straightforward results above.

For simplicity, partition students into three categories: H, M, and L. H students have sufficiently high priorities that they are always matched to their reported first choice. M students have intermediate priorities and have guaranteed admission to the larger school if they list it as their first choice but not to the smaller school. L students have low priorities and do not have guaranteed admission to either school.[4]

Under common preferences, M and L students face a winner's curse—being admitted to a school they are not guaranteed admission to is bad news. This is because vacant seats are adversely selected. As H students collect more information, they are more likely to attend the better school, leaving seats open at the worse school.

This adverse selection affects search. At the outset, M and L students do not know which schools will have vacancies once H students have been matched. Consequently, in equilibrium, all M and L students refrain from search for *any* positive search cost.

There is a zero-sum nature to how match utilities are allocated under common preferences. In equilibrium, M students always report their guaranteed school as their first choice: it cannot be adversely selected, hence yields greater expected utility than the other school, which may be adversely selected. The remaining surplus is split between H and L students: gains in H students' match utilities equate to losses in L students' match utilities. A universal reduction in search costs allows H students to search more thoroughly and get better matches, hurting L students. Thus, decreasing search costs can lead to the undesirable effect of generating more unequal outcomes.

In practice, there are different search protocols by which students visit schools. Search is sequential if students have many opportunities to visit and no commitment. For example, high schools often offer shadow days over a range of dates with no restrictions. In other cases, search is simultaneous: students choose the schools they will visit at the outset. For instance, families may plan a tour of schools entailing nonrefundable travel and accommodations. Alternatively, students may partake in a bus tour of schools, hence will visit all schools scheduled on the tour regardless of their impressions of those they have already visited.

The search protocol impacts equilibrium outcomes. In individual search problems, sequential search is generally more efficient than simultaneous search. In a school-choice setting, however,

[4]As shown in Section 5, there are cutoffs in priorities that determine a student's type.

students are strategic, and their reports are siphoned into an algorithm where their matching depends on others' reports. Hence, it is not immediately clear which search protocol is more efficient. Under idiosyncratic preferences, there are no search externalities. Thus, as in individual choice, students are weakly better off under sequential search than under simultaneous search.

Common values, however, induce search externalities that yield more nuanced comparisons between search protocols. While M students have the same utility across both protocols since they always attend their guaranteed school without search, H students' search affects L students' outcomes. At high costs, sequential search is better than simultaneous search for H students but worse for L students: H students visit more schools under sequential search than under simultaneous search at high costs, so on average L students have access to worse seats under sequential search. At high costs, welfare is higher under *simultaneous* search than under sequential search, marking another distinction from the idiosyncratic case.

Despite match utilities being zero sum, at low costs, H and L students both have higher expected values under sequential search. While sequential search is *more efficient* than simultaneous search for H students, at low costs, it is also *less thorough*. H students sometimes attend the worse school, leaving open seats at the better school.

One can use the above analyses of these two polar environments to understand what would happen if students had preferences with both common and idiosyncratic components. While there are no search externalities under purely idiosyncratic preferences, the presence of an idiosyncratic component in this hybrid model *increases* search externalities spurred by the common component. The common component creates a winner's curse that discourages search for low-priority students, but potential gains in the idiosyncratic component incentivize more search than if preferences were purely common. Indeed, there are costs at which students do not search under common preferences but do search when an idiosyncratic component is added. An increase in high-priority students' search worsens the adverse selection faced by the lowest-priority student. Thus, despite idiosyncratic utilities inducing no search externalities on their own, the idiosyncratic component of a hybrid model exacerbates the search externalities brought on by the common component.

The results above are for symmetric equilibria in which ties are broken uniformly at random; however, symmetric tie-breaking can generate inefficiencies in this incomplete-information setting. The inefficiency from random tie-breaking stems from a coordination problem in which high-priority students unnecessarily crowd out seats that could be made available to lower-priority students.

Previous work has studied how to efficiently break ties in matching environments with complete information. A matching environment with information acquisition presents a new source of

indifference on the students' side: students may be ex-ante indifferent between visit sets. Under idiosyncratic preferences, schools are symmetric, so students are indifferent among *which* schools they visit conditional on the *number* of visits. For both search protocols, I provide a simple, noninvasive instrument by which a social planner breaks students' indifferences in search strategies to coordinate their actions and improve welfare. For a range of search costs, this instrument can generate substantial improvements.

My results have several implications for the design of school choice. Common preferences generate search externalities, which can make simultaneous search more efficient than sequential search, in equilibrium. Additionally, school districts can overcome coordination problems and increase welfare with a noninvasive instrument that breaks students' indifferences in search strategies.

2.2 Related Literature

Despite empirical evidence of students being partially informed of their preferences over schools—see Hastings and Weinstein (2008), Hoxby and Turner (2015), and Corcoran et al. (2018)—there is relatively little theoretical research on matching with incomplete information, and much of that literature is recent. Liu et al. (2014) and Liu (2020) formulate notions of stability under asymmetric information with transfers. Bikhchandani (2017), Chen and Hu (2018), and Fernandez et al. (2021) study environments with incomplete information and no transfers.

There has been some theoretical work on matching with information acquisition. Lee and Schwarz (2017) study costly interviews in one-to-one matching in which both sides are uninformed of their preferences at the outset and firms simultaneously search over candidates. Immorlica et al. (2020) find that when students sequentially inspect schools, information deadlocks preclude the existence of a mechanism that guarantees a regret-free stable outcome. Echenique et al. (2021) show that under interview capacity constraints, reported preferences should not be taken as an agent's true preferences because reports are distorted by the interview process.[5] Artemov (2021) studies a model in which schools' common values are known and one school has an idiosyncratic component that is costly to observe; he shows that under random serial dictatorship, students collect less information than is socially optimal. I study a school-choice setting with a centralized clearinghouse in which preferences range from idiosyncratic to common. I focus on how equilibrium behavior translates to different outcomes under multiple search protocols, paying particular attention to welfare, inequality,

[5] An assumption I do not impose but is common in the interview literature is that agents can match only with those they have interviewed. Imposing such an assumption would not substantively affect my results.

and comparative statics with respect to search costs.[6]

Much of this paper builds on Kloosterman and Troyan (2020), who show that in a setting with interdependent preferences, uninformed students face a winner's curse when they match with a school for which they have low priority. An important distinction between the models is that Kloosterman and Troyan assume that information is exogenous, whereas I study a setting of endogenous information acquisition. Additionally, their model gives student information a particular form (a student is either perfectly informed or completely uninformed), whereas I allow students to be partially informed. Thus, the degree of information, hence also the degree of the winner's curse, is determined through equilibrium. This enables me to gain insights on how common values impact students' search decisions, reporting decisions, and outcomes. Additionally, I show how the search protocol affects the severity of the winner's curse and how search costs impact which protocol yields greater welfare.

This paper also touches on the tie-breaking literature in matching. Existing work has focused on complete-information environments in which agents exhibit indifferences in their preferences or in which schools have coarse priorities over students—see Erdil and Ergin (2008) and Abdulkadiroğlu et al. (2009).[7] I contribute to this literature by studying how to efficiently break students' ex-ante ties in search strategies, prior to students participating in a centralized clearinghouse. Breaking ties at the outset can have important welfare effects.

2.3 Model

School-choice markets tend to be large. For example, the New York City Department of Education annually assigns tens of thousands of eighth graders to high schools; centralized college admission in China involves about 10 million applicants a year.[8] To approximate these large markets, I study a continuum of students with unit mass.

The school district has J schools. School or college j, denoted k_j, has capacity $q_j > 0$. The school district has sufficient capacity: $\sum_{j=1}^{J} q_j = 1$. I assume that all students are acceptable to all schools and that the outside option of not attending a school in the set $\{k_1, \ldots, k_J\}$ has a utility of 0, which implies that all schools are acceptable to all students. Without loss of generality, assume that

[6] For lab experiments on school choice with costly information acquisition, see Chen and He (2021b) and Bochet et al. (2021)—who study settings in which students can acquire costly signals of their preferences—and Hakimov et al. (2021)—who study a setting in which students pay a cost to observe the ordinal ranking of searched schools. The companion paper Chen and He (2021a), Bochet et al. (2021), and Hakimov et al. (2021) all derive theoretical results for their particular settings.

[7] Erdil and Ergin (2017) extend their earlier work to allow for indifferences in preferences on both sides of the market.

[8] See Corcoran et al. (2018) and Chen and Kesten (2017).

$q_1 \geq q_2 \geq \cdots \geq q_J$. Schools have the same priority rankings over students. This assumption applies to settings with a standardized test or some other objective measure governing schools' preferences. Alternatively, such a priority ranking can be the consequence of a global lottery over students.

Students are indexed by their priorities. Each student $i \in [0,1]$ knows her priority. A higher number means higher priority. This applies to settings in which students know their test score/lottery number.[9]

In many school-choice settings, schools have limited discretion over their admission policies, so I do not consider strategic behavior on their part. I focus only on students, as is common in the literature.[10]

2.3.1 Utility Specifications

Student i receives a utility of U_{ij} for attending school k_j. I start with two extreme cases of utility: pure idiosyncratic utilities and pure common values. For idiosyncratic preferences, $U_{ij} = S_{ij}$. Each S_{ij} is drawn independently according to a cumulative distribution function $F(\cdot)$, which has full support on $[0,b]$ for finite $b > 0$. I assume that F is continuously differentiable on $(0,b)$ and has probability density function $f(\cdot)$, which is continuous.[11] Idiosyncratic preferences capture heterogeneity in students' preferences for schools and serve as a useful benchmark that is easy to generalize.

For pure common values, $U_{ij} = \delta_j$. That is, each student receives the same utility δ_j for attending school k_j. We can interpret δ_j as an objective measure of school quality. Each δ_j is drawn independently from a uniform distribution on $[0,1]$.

I study these two antipodal models because they are instructive of the forces they bring.[12] After analyzing these two settings and the equilibrium patterns they generate, I study a setting that combines them and show how their forces interact.

2.3.2 Search Protocols

Unlike many matching models that assume perfect information over preferences, this model assumes that students do not know their preferences at the outset. Rather, each student i must pay a cost

[9] The qualitative results for idiosyncratic preferences generalize to a setting in which priorities are independent and identically distributed across students and schools, as shown in Section 8.2. For expositional simplicity, I focus on schools having common priorities over students.

[10] See Pathak and Sönmez (2008) and Abdulkadiroğlu et al. (2015).

[11] These conditions are sufficient for a well-behaved distribution of the match utility corresponding to the student's favorite visited school.

[12] Because I study a continuum of students, there is no aggregate uncertainty in preferences. In Section 5.3, I present a 3-student example that exhibits aggregate uncertainty and maintains the forces highlighted in the continuum model.

$c > 0$ to visit school k_j and observe U_{ij}.[13] Students cannot observe other students' search or utility realizations.

I study both simultaneous search and sequential search, as both search protocols are plausible in a school-choice or centralized-college-admission setting. If parents and students schedule a tour of schools with commitment, for example, by booking a trip with nonrefundable travel and accommodations, they are effectively searching simultaneously. In contrast, if families visit schools over time and have no commitment, they are essentially searching sequentially.

Under simultaneous search, each student chooses which set of schools to visit. This is students' only opportunity to search. Under sequential search, a student either visits a school in $\{k_1, \ldots, k_J\}$ they have not yet visited or stops search altogether. If she visits a school, she immediately observes her match utility for that school. She continues until she either stops search or has visited every school in the district. Under both protocols, students may choose to not visit any schools. Let L_i denote the set of all schools that student i visited once search has concluded.

2.3.3 Matching Procedure

After search, each student i submits a complete, strictly-ordered preference list $R_i \equiv R_i(\{U_{ij}\}_{j \in L_i})$ to the school district.[14] I place no additional restrictions on submitted reports. For example, a student can report an unvisited as her first choice. This could realistically occur if a student visits few schools and does not like any of them.

After students submit reports, the school district runs student-proposing DA and obtains matching ν. Because schools have the same preferences over students, student-proposing DA is effectively serial dictatorship. The algorithm starts at the top of the list of students ($i = 1$) then works its way down, assigning each student to her first choice that is still available. After the algorithm is finished running, each student i attends the school to which she is matched and receives utility

$$V_i(L_i, R) = U_{i\nu(i)} - c|L_i|,$$

where R is the collection of all students' reports.

[13] While I assume that c is constant across all students, the results under idiosyncratic preferences are maintained when students' costs are independently and identically distributed.

[14] It is not unreasonable to rank all schools if there are relatively few or if schools are in the same tier. See Chen and Kesten (2017) and Kesten and Kurino (2019).

2.3.4 Equilibrium

I focus on Nash equilibria. Under simultaneous search, the standard best-response conditions for a Nash equilibrium in this setting translate to:

(1) Students maximize expected utility when choosing visit strategies:

$$\mathbf{E}\left[V_i(L_i, (R_i(\{U_{ij}\}_{k_j \in L_i}), R_{-i}))\right] \geq \mathbf{E}\left[V_i(L', (R_i(\{U_{ij}\}_{k_j \in L'}), R_{-i}))\right] \ \forall L';$$

(2) Students maximize expected utility when submitting reports:

$$\mathbf{E}\left[V_i(L_i, (R_i, R_{-i}))|\{U_{ij}\}_{k_j \in L_i}\right] \geq \mathbf{E}\left[V_i(L_i, (R', R_{-i}))|\{U_{ij}\}_{k_j \in L_i}\right] \ \forall R',$$

where R_{-i} is the collection of other students' reports. The expectations are taken with respect to match-utility realizations and other students' reports. Under sequential search, the best-response conditions are analogous but with students maximizing over visit strategies (an ordering rule and a stopping rule) rather than a fixed L_i. Because the district uses DA and all students are acceptable to all schools, the market clears. All seats are filled, and there are no stray students left unmatched.

Most of the analysis is on symmetric Nash equilibria, in which students employ a uniform-random tie-breaking rule, which I refer to as an *uniform tie-breaking equilibrium* (UTB-equilibrium).[15] Specifically, if multiple search strategies are utility maximizing, the student chooses uniformly at random among those that involve the most visits. Similarly, if multiple reports to the clearinghouse are utility maximizing, the student chooses from them uniformly at random. This refinement simplifies my discussion but is not necessary for the qualitative results below.

2.4 Idiosyncratic Preferences

Idiosyncratic preferences apply to settings in which students have varying preferences over schools' characteristics: location, classes offered, and extracurriculars, for example. I show that under idiosyncratic preferences, students have straightforward reporting strategies and no search externalities. As a result, each student's expected utility decreases in costs and is weakly higher under sequential search than under simultaneous search. Idiosyncratic preferences serve as a useful benchmark and have results that contrast nicely with results under common preferences.

[15] There are many Nash equilibria in this game, but I focus on symmetric equilibria because they are realistic in school choice where it is unlikely that a sizable fraction of students coordinate strategies. I return to a discussion of asymmetric Nash equilibria in Section 6.

2.4.1 Simultaneous Search

I solve for the equilibrium via backwards induction. First, I solve for the reports students submit to the clearinghouse conditional on visits. Then, I solve for which schools students visit. Lemma 1 provides the equilibrium reporting strategy.

Lemma 1. *Under idiosyncratic preferences, in equilibrium, student i's report ranks visited schools from the highest realized S_{ij} to the lowest. She ranks unvisited schools as if they have realizations equal to $\mathbf{E}[S_{ij}]$. In a UTB-equilibrium, she also orders schools that are tied uniformly at random among themselves.*

Lemma 1 establishes that, under idiosyncratic preferences, students maximize utility by employing the natural reporting strategy of ranking schools from the highest realization to the lowest. This result follows from strategy-proofness of proposers in DA. Under idiosyncratic preferences, conditioning on others' actions does not affect a student's expected match utility of each school. Hence, she reports truthfully. The tie-breaking rule implies she orders unvisited schools uniformly at random.

The refinement that students randomize among utility-maximizing reports uniformly at random is fairly innocuous.[16] Because schools are symmetric, there is no reason students should systematically break ties to favor one school over another. Even so, the qualitative results below would still hold if students broke ties asymmetrically across schools in some independent and identically distributed fashion.

Now I solve for the set of schools each student visits. Importantly, this is not a standard search problem. When choosing visit sets, students account for the fact that their matching is determined by DA, which is affected by others' reports. As a note, it is sufficient to consider only positive-probability events, so for the remainder of the paper I ignore knife-edge cases (for example, two schools having the same realization).

Before getting to the equilibrium forces, consider a problem in which student i is given some exogenous feasible set of schools D_i and is matched to her highest-ranked school in D_i. Lemma 2 establishes that the student determines the number of schools she visits using a simple threshold strategy.

Lemma 2. *Consider an individual search problem with simultaneous search in which student i is matched to her highest-ranked school among schools in the feasible set D_i. Define*

$$\ell^*_{D_i}(c) = \max\left\{\underset{\ell\in\{0,\dots,|D_i|\}}{\operatorname{argmax}} \mathbf{E}\left[\max\{\mathbf{E}\left[S^{(\ell)}\right], \mathbf{1}(\ell < |D_i|)\mathbf{E}[S_{ij}]\}\right] - c\ell\right\},$$

[16] In Section 6, I analyze Nash equilibria that break this refinement.

where $S^{(\ell)}$ is the ℓ^{th}-highest order statistic of ℓ-many S_{ij} variables. Then $\ell^*_{D_i}(c)$ is the number of school visits that maximizes student i's expected utility in the individual search problem. It is unique and weakly decreasing in c.

This lemma is relatively straightforward but is convenient in characterizing the solution to the full game. The right-hand side is student i's expected match utility minus search costs. If the student visits all schools in her feasible set, she attends the best one she visits, yielding a match utility of $S^{(|D_i|)}$. If the student visits $\ell < |D_i|$ schools, her match utility is either the best school she visited or the expected value of an unvisited schools, whichever is greater. Thus, $\ell^*_{D_i}(c)$ maximizes the student's expected utility in the individual search problem. Uniqueness is given by the refinement, in which the student chooses the highest ℓ that maximizes her expected utility.

As costs increase, the student visits fewer schools. As shown in the appendix, school visits have diminishing returns in match utility. Thus, students increase the number of schools they visit until match-quality returns fall below cost c, implying that $\ell^*_{D_i}(c)$ is a threshold rule and is weakly decreasing in c.

Definition 1. *Consider simultaneous search. Let student i have feasible set D_i. Define the* Sim-Max *search strategy as follows:*

- *If $\ell^*_{D_i}(c) < |D_i|$, student i visits $\ell^*_{D_i}(c)$ schools uniformly at random among schools in D_i.*
- *If $\ell^*_{D_i}(c) = |D_i|$, student i visits all schools in D_i.*

The Sim-Max search strategy maximizes the expected utility of the individual search problem under simultaneous search. Proposition 1a shows the straightforward result that when preferences are idiosyncratic and students' matches are determined by the clearinghouse, students' search problems are effectively individual search problems. Because there is a continuum of students, there is no aggregate uncertainty. If all students employ the above strategy, they can deduce their feasible sets endogenous to the equilibrium.

Proposition 1a. *Under idiosyncratic preferences and simultaneous search, there is a unique UTB-equilibrium. In equilibrium, student i employs the Sim-Max strategy using the feasible set $D_i = \{k_j : i \geq r_j\}$, where r is the vector of admission cutoffs. Admission cutoffs are unique and independent of the cost c.*

For intuition, it is helpful to use an admission cutoff characterization of DA. Let $r \in [0,1]^J$ be a vector of admission cutoffs such that if student i lists k_j as her first choice, she is matched to it

$\iff i \geq r_j$. Thus, given admission cutoffs r, she can restrict her attention to the schools in her feasible set $D_i \equiv \{k_j : i \geq r_j\}$, simplifying her decision problem.

For illustration, let $q_1 > q_2 > \cdots > q_J$. By symmetry, each student has an equal ex-ante probability of attending each school in her feasible set. Students with the same feasible set matriculate into those schools at identical rates. This is true regardless of how many schools each student visits since they choose their visit sets uniformly at random within D_i. Hence, matriculation into schools is independent of the c. School k_J reaches capacity after Jq_J mass of highest-priority students have been matched. Therefore, $r_J = 1 - Jq_J$. The next group of students matriculate evenly across the remaining $J-1$ schools. Hence, $r_{J-1} = 1 - r_J - (J-1)(q_{J-1} - q_J)$. One can continue defining these cutoffs recursively until reaching $r_1 = 0$, implying that all students have school k_1 in their feasible set.

Corollary 1. *Under simultaneous search, all students' expected values are decreasing in cost c in the unique UTB-equilibrium. Hence, expected equilibrium welfare*

$$\mathbf{E}\left[W^{sim}(c)\right] = \int_0^1 \left(\mathbf{E}\left[V_i^{sim}(L_i^*, R^*)\right] - c\ell_{D_i}^*(c)\right) di$$

is also decreasing in c.

Corollary 1 establishes that, in equilibrium, higher costs make students worse off. As costs increase, students either (1) visit the same number of schools but pay more in information costs, or (2) visit fewer schools, reducing their match utility to the point that they are worse off despite the cost savings of fewer visits. Complete-information welfare corresponds to $\mathbf{E}\left[W^{\text{sim}}(0)\right]$ and is strictly greater than $\mathbf{E}\left[W^{\text{sim}}(c)\right]$ for any $c > 0$.

While decreasing c helps all students, it helps them differentially. Students with larger feasible sets visit weakly more schools. Let $\bar{\ell}(c)$ be the largest feasible set size such that $\ell_{D_i}^*(c) = |D_i|$. To highlight the distributive effects of decreasing search costs, I decompose welfare by sorting students into two groups.

I label students in the first group *admission-constrained students.* Such students have $D_i \in \mathcal{D}_{ac} \equiv \{D : |D| \leq \bar{\ell}(c)\} \cup \{k_1\}$ and do not get a higher expected match utility from a decrease in c. They either already visit every school to which they can gain admission or have only one school in their feasible set (hence do not visit it at any cost). While a decrease in c can increase admission-constrained students' expected values by decreasing their information costs, their match qualities cannot increase because they already attend the best school available to them.

In contrast, *cost-constrained students* have $D_i \in \mathcal{D}_{cc} \equiv \{D : |D| > \bar{\ell}(c)\}\backslash\{k_1\}$ and benefit two-fold from a decrease in c: they pay less per school visit and may be incentivized to visit more schools.

Decreasing c has important distributive effects. Decreasing c improves the match quality only of cost-constrained students. Moreover, for any fixed c, cost-constrained students are already better off than admission-constrained students. Hence, decreasing c makes those students better off by an even greater margin.

We can use results from the constant-cost model to draw similar conclusions when students' costs vary. If each student's cost is drawn independently from some arbitrary distribution $G(\cdot)$, a student's decision problem depends only on her draw of c_i and is otherwise unaffected by the cost distribution. Welfare would then correspond to a weighted expectation of the welfare terms in Corollary 1. We can still decompose students into admission constrained and cost constrained but with the nuance that students with the same feasible set may be sorted into different groups for sufficiently disparate realizations of c_i.

2.4.2 Sequential Search and Welfare Comparison

In this subsection, I study the model under sequential search. For presentation simplicity, I assume that each S_{ij} is drawn from a standard uniform distribution, which is assumed throughout the analysis. The following techniques could be applied to a large class of distributions. Uniformity allows for closed-form solutions amenable to clear comparisons across search protocols.

In standard individual search problems, the conventional wisdom is that sequential search is more efficient than simultaneous search (Wald, 1945). However, because students search strategically and submit reports to a clearinghouse, it is not immediately obvious which search method provides higher welfare. I show that because there are no search externalities under idiosyncratic preferences, sequential search is more efficient for all students. If a student knows her feasible set, her search problem is akin to the Pandora's problem introduced by Weitzman (1979) with the additional option to "take a box unopened," i.e., attend a school unvisited. This type of search problem has been studied by Doval (2018).

Definition 2. *Consider sequential search. Let student i have feasible set D_i. Define the* Seq-Max *search strategy as follows. Student i does not search if* $[|D_i| = 1$ *and* $c > 0]$ *or* $[|D_i| \geq 2$ *and* $c > \frac{1}{8}]$. *Otherwise, she employs the following search policies:*

Order *If visiting a school, visit uniformly at random among unvisited schools in D_i.*

Stopping

1. *If there is more than one unvisited school in D_i, stop only if the best visited school has match utility greater than $1-\sqrt{2c}$.*
2. *If there is only one unvisited school in D_i, stop if the best visited school has a match utility greater than $1-\sqrt{2c}$ or less than $\sqrt{2c}$.*

If a student visits a school that is really good (i.e., $S_{ij} > 1-\sqrt{2c}$), she stops search, as subsequent schools are expected to be worse. Per the utility-maximizing reporting strategy, she ranks that school ahead of all other schools in D_i and ultimately attends it. If, on the other hand, she visits all but one school and dislikes them (i.e., always has $S_{ij} < \sqrt{2c}$), she attends the last school without visiting it, as it is expected to be better than the schools she has already visited. If the utility of her best-visited school is intermediate (in $[\sqrt{2c}, 1-\sqrt{2c}]$), she visits that last school in her feasible set and attends her most preferred.

The above thresholds come from indifference conditions in the student's search. Threshold $\sqrt{2c}$ is the value of the outside option that makes a student indifferent between visiting a school and attending it unvisited (this value is sometimes referred to as the *backup value*). Threshold $1-\sqrt{2c}$ is the is the value of the outside option that leaves the student indifferent between visiting a school and stopping search to match with that school (this value is sometimes referred to as the *reservation value*).[17] As c increases, students visit fewer schools in expectation. When costs are prohibitively high, a student does not visit any schools. In contrast, if $c = 0$, she visits all schools, and we are in the realm of complete information.[18] As costs increase, a student's expected value decreases.

Proposition 1b shows that under sequential search, students again are essentially solving individual search problems.

Proposition 1b. *Under idiosyncratic preferences and sequential search, there is a unique UTB-equilibrium. In equilibrium, student i employs the Seq-Max strategy using the feasible set $D_i = \{k_j : i \geq r_j\}$, where r is the vector of admission cutoffs. Admission cutoffs are unique and independent of the cost c.*

Using the Seq-Max search strategy, equilibrium admission cutoffs from the simultaneous-search equilibrium apply to the sequential-search equilibrium, hence again are independent of c.[19]

[17] Mathematically, the backup value is the number x_B such that $c = \int_0^{x_B}(x_B - x)dF(x)$. The reservation value is the number x_R such that $c = \int_{x_R}^{\infty}(x - x_R)dF(x)$.

[18] This strategy can be adapted to distributions other than standard uniform as long as they satisfy conditions laid out in Theorem 1 of Doval (2018).

[19] As is the case for simultaneous search, the equilibrium strategies above extend to settings in which costs are distributed independently. Student i's decision problem depends only on her realized costs. If costs are distributed independently according to $G(\cdot)$, welfare is given by $\mathbf{E}[\mathcal{W}^{\text{seq}}(G)] = \int \mathbf{E}[W^{\text{seq}}(c)]\ dG$.

Remark 1. *Students have the same feasible sets in the unique simultaneous-search UTB-equilibrium as in the unique sequential-search UTB-equilibrium. Thus, all students do weakly better under the unique sequential search UTB-equilibrium than under the unique simultaneous search UTB-equilibrium in expectation. Hence,* $\mathbf{E}\left[W^{seq}(c)\right] \geq \mathbf{E}\left[W^{sim}(c)\right]$.

This result stems from the symmetry across schools and students having equal ex-ante probability of attending each school in their feasible sets. Sequential search maintains the set of schools to which students can gain admission but allows students to search more efficiently than simultaneous search. The latter part of Remark 1 jibes with the intuition that sequential search is more efficient than simultaneous. There are no externalities. Changing the protocol does not change the students' feasible sets.

2.5 Common Preferences

In this section, I analyze strategies and outcomes when students share preferences over schools. For example, students may all value academic quality and school amenities. Under such preferences, there is an externality of search: as high-priority students search, lower-priority students are left with seats that are adversely selected, making those students worse off. This externality of search affects students' visits. Moreover, there is a range of costs for which simultaneous search provides higher welfare than sequential search, underscoring the equilibrium effects of the search protocol and utility specification. As I show at the end of this section, the search externalities generated in a common-preferences model are also exhibited in a hybrid model, making this analysis useful in understanding more complicated settings in which preferences are comprised of both common and idiosyncratic components.

Let utilities for schools be common across students. Specifically, assume that $U_{ij} = \delta_j \ \forall i$. Let each δ_j be drawn independently from a standard uniform distribution.[20] For presentational simplicity, I consider only two schools and assume that $q_1 > q_2 > 0$. At the end of this section, I address what would happen with more than two schools.

2.5.1 Simultaneous Search and Sequential Search

It is convenient to partition students into one of three types. An *H student* has $i \geq 1 - q_2$. Such a student has high enough priority to gain admission to her first choice regardless of what others report.

[20] My analysis applies to other distributions, but focusing on Uniform[0,1] makes for clear comparisons to the idiosyncratic case.

An *M student* has $i \in [1-q_1, 1-q_2)$. Such a student has guaranteed admission to school 1 if she lists it as her most-preferred school but does not have guaranteed admission to school 2. For example, if all H students report $k_2 \succ k_1$, then school 2 fills to capacity before DA reaches M students. An *L student* has $i < 1 - q_1$. Her priority is so low that she does not have guaranteed admission to either school. These types are exogenously determined; they are not equilibrium objects. I define them based on the worst-case scenario of a student's feasible set, as students' reports may be correlated. With more than two schools, there would be more types, but the main forces discussed below would be the same.

When utilities are common, a partially informed student does *not* rank unvisited schools by their unconditional means, marking a clear distinction from the idiosyncratic case. A student is matched to a school for which she has low priority only if higher-priority students pass it over. Because all students have the same preferences, an informed high-priority student passing over a school means a lower-priority student should do so as well. Thus, a vacancy indicates that a school is not desirable—there is a winner's curse.

The following example illustrates the effect the winner's curse has on search and lower-priority students' reports. H students are not affected by other students' search. If costs are low, in equilibrium, all H students visit both schools and are matched to the better school.

Next, consider an M student with priority $i = 1 - q_2 - \varepsilon$ for a small $\varepsilon > 0$, placing her at the cusp of being an H student. Suppose she visited school 1, thus knows δ_1, but did not visit school 2. What should she report to the clearinghouse?

If she reports $k_1 \succ k_2$, she is matched school 1, which yields a match utility of δ_1.

If she reports $k_2 \succ k_1$ and is matched to school 1, it must be that school 2 is fully occupied, which occurs only if $\delta_2 > \delta_1$. The student receives utility δ_1, which is the same utility as if she had reported $k_1 \succ k_2$.

If she reports $k_2 \succ k_1$ and is matched to school 2, then school 2 must have vacant seats, meaning H students attended school 1. In that case, $\delta_1 > \delta_2$, so the student is strictly worse off reporting $k_2 \succ k_1$ than reporting $k_1 \succ k_2$. That is, reporting $k_2 \succ k_1$ and being matched to her first choice is bad news. Thus, the student should report $k_1 \succ k_2$, even for arbitrarily low δ_1.

Reporting $k_2 \succ k_1$ is never a best response since she is matched to school 2 if and only if it is worse than school 1. Taking this a step further, because the student's report to the clearinghouse does not change based on the realization of δ_1, she has no incentive to visit school 1 in the first place.

Importantly, the winner's curse emerges only when students visit a strict subset of schools and

must infer the value of unvisited schools. By strategy-proofness of DA, any student with complete information over her preferences reports truthfully. Because of search externalities, a partially-informed student ranks an unvisited school according to its expected value *conditional on being matched to it*, not the school's unconditional expected value.

The severity of the winner's curse depends on how much information higher-priority students gather. As H students acquire more information, they are more likely to attend the better school, leaving M and L students with more adversely selected seats. The above example assumes H students visit both schools, but the intuition for the winner's curse is similar when H students visit only one—seats are still adversely selected at the school for which M students do not have guaranteed admission. This adverse selection discourages M students from reporting $k_2 \succ k_1$ and renders them unwilling to visit only school 1.

Proposition 2 illustrates that the externality of search is so prominent that in the unique equilibrium, M and L students do not search for *any* positive information cost, marking another distinction from the idiosyncratic case (where a positive fraction of M students are willing to search).

Proposition 2. *If $c \in \left(0, \frac{1}{8}\right]$, there is a unique UTB-equilibrium under both search protocols:*

- *H students employ the respective Sim-Max or Seq-Max search strategy.*
- *M students visit neither school and report $k_1 \succ k_2$.*
- *L students visit neither school and uniformly randomize between reporting $k_1 \succ k_2$ and reporting $k_2 \succ k_1$.*

If $c > \frac{1}{8}$, no student searches in equilibrium.

To avoid occupying a seat that is adversely selected, M students report $k_1 \succ k_2$ since they are guaranteed admission to school 1. L students lose out and occupy vacancies left by H and M students, yielding an expected utility lower than $\mathbf{E}[\delta_j]$.[21,22]

2.5.2 Welfare Comparison

While the two search protocols have similar equilibria, they yield very different welfare implications. There is a range of costs such that simultaneous search yields higher welfare than sequential search,

[21] If a district has many schools that can be divided into fully separated tiers of two schools each, students can implement an equilibrium analogous to that above under appropriate utilities and search costs. Students calculate the highest tier to which they can gain admission, which schools in that tier to which they have *guaranteed* admission, and employ the respective search and reporting strategies.

[22] If information costs were heterogeneous across students, the strategy in Proposition 2 would not necessarily be an equilibrium. Because of search externalities, a student's decision problem depends on her realized costs as well as others'.

running contrary to the case of idiosyncratic preferences in which sequential search is weakly better for all students no matter the cost. This also contrasts with insights in individual decision making.

The district has exactly enough seats for students. Utilities are pure common value. Therefore, total welfare from match utility is constant across all allocations. The only way to change welfare is by changing the number of schools searched. Proposition 2 states that only H students search. Thus, welfare comparative statics are driven by H students' equilibrium search.

The search protocol has clear distributive effects. Figure 2.1 plots the expected values for each student type under both search protocols.

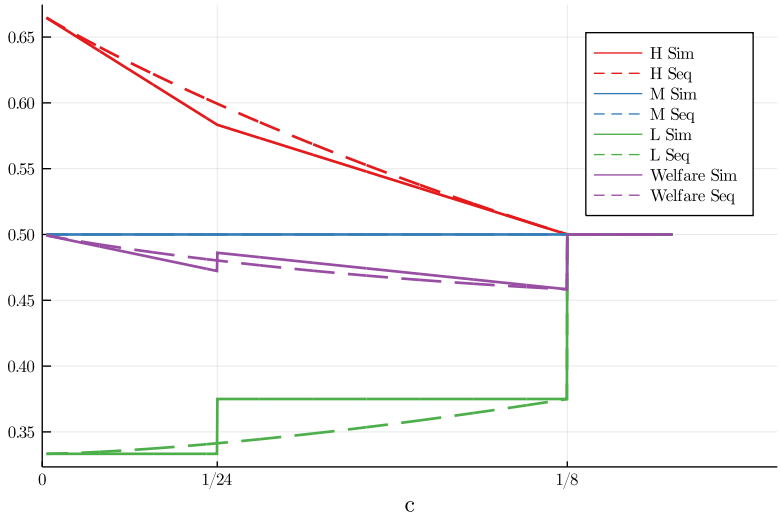
Under simultaneous search, H students visit two schools if $c \leq \frac{1}{24}$, one school if $c \in \left(\frac{1}{24}, \frac{1}{8}\right]$, and zero schools if $c > \frac{1}{8}$. In contrast, under sequential search, each H student visits one or two schools for $c \in \left(0, \frac{1}{8}\right)$. Because H students face no externalities, they do better under sequential search than under simultaneous search. M students get $\frac{1}{2}$ in expectation under both search protocols since they always attend school 1 without search.

How L students fare is more subtle. An increase in match utility of the H students comes at the direct expense of the L students. However, at low costs, H students *and* L students both have better outcomes under sequential search than under simultaneous search, as shown in Figure 2.1. While H students search *more efficiently* under sequential search at low costs, they also search *less thoroughly*. Sometimes they stop search after visiting one school despite a positive probability that the other school is better. There is a chance an H student attends the worse school, allowing an L student to attend the better school. Thus, at low costs, all students are weakly better off under sequential search.

At intermediate costs $\left(c \in \left(\frac{1}{24}, \frac{1}{8}\right]\right)$, L students do better under simultaneous search than under sequential search because H students visit fewer schools under simultaneous search. Under simultaneous search, H students go from visiting two schools to only one at $c = \frac{1}{24}$. Since H students visit one fewer school, L students gain access to seats that are less adversely selected, hence the positive jump in L students' expected value. Likewise, H students suddenly spend less in information costs at $c = \frac{1}{24}$, explaining the positive jump in welfare.

This observation highlights an important implication that an increase in information costs discourages H students from searching, which benefits L students. A school district may be interested in both the welfare and the distribution of utilities across students. For example, it may dislike unequal student outcomes, especially if priorities are determined by pure lotteries. We see that if search is simultaneous and costs are just a hair above $\frac{1}{24}$, a small decrease in costs marginally benefits H students but greatly harms L students. Hence, a school district may refrain from subsidizing

Figure 2.1: Expected values for H students, M students, L students, and welfare under both search protocols. "Sim" refers to simultaneous search. "Seq" refers to sequential search. In equilibrium, all students of the same type have the same ex-ante expected value.



visit costs to spare L students.

I focus on a two-school environment for simplicity. The main forces and search externalities are maintained in settings with more than two schools, as I show in the next subsection.

2.5.3 Arbitrary Number of Schools

I now show that the main forces in the two-school model extend to settings with an arbitrary number of schools $J \geq 2$. Define H^* *students* as those with priority $i \geq 1 - q_J$ and L^* *students* as those with priority $i < 1 - q_J$. Then H^* students are matched to their highest-reported school regardless of other students' reports, so they employ the Sim-Max and Seq-Max strategies in equilibrium. For reasons analogous to the two-school case, seats made available to L^* students are adversely selected. The degree of adverse selection depends on the strategies of higher-priority students. Proposition 3 shows that the main comparison between simultaneous and sequential search holds in environments with J schools.

Proposition 3. *Let $J \geq 2$ and $c \in \left(0, \frac{1}{8}\right)$. For sufficiently small c, total expected L^* students' match utility is greater under any equilibrium with sequential search than for any equilibrium with simultaneous search. For sufficiently large c, total expected L^* students' match utility is greater for any equilibrium with simultaneous search than for any equilibrium with sequential search.*

Proposition 3 follows from the search externalities that H^* students impose on lower-priority

students. An increase in H^* students' match utility translates to a decrease in L^* students' match utility. For sufficiently low costs and simultaneous search, H^* students visit every school and match with the best one. Under sequential search, however, there is a positive probability that they match with a school that is not the best: for example, because they visited a sufficiently good school early on and concluded search. For sufficiently high costs, this ordering is reversed: H^* students visit only one school under simultaneous search and visit at least one school under sequential search, hence have better matches under sequential search.[23]

While common preferences are a polar environment, this case of the model underscores how lower-priority students are affected by the actions of higher-priority students. Next, I show that adverse selection on the common component can be exacerbated by the presence of an idiosyncratic component.

2.5.4 Example with Hybrid Preferences

In this subsection, I examine equilibrium behavior for an example in which students' preferences have both idiosyncratic and common components. This captures preferences in which students have homogeneous preferences over some characteristics (for example, school quality) but heterogeneous preferences over others. Having both components exacerbates the externalities present the pure common-value case.

Let there be three students (h, m, and ℓ) and two schools (with sizes $q_1 = 2$ and $q_2 = 1$). Both schools prefer student h to student m to student ℓ. This setup is analogous to the two-school pure common-value case: student h has guaranteed admission to both schools, student m has guaranteed admission to school 1 but not school 2, and student ℓ does not have guaranteed admission to either. Students break indifferences uniformly at random, but that is not important for my results. The search protocol is simultaneous search. Students' preferences are given by $U_{ij} = S_{ij} + \delta_j$, where the idiosyncratic component S_{ij} is independently drawn uniformly from $\{0, 1, 2\}$ and the common component δ_j is independently drawn uniformly from $\{0, 1\}$. These discrete distributions are chosen for tractability and simplicity.

Similar to the pure common-value case, in equilibrium, student ℓ never searches, and student h visits two schools if costs are sufficiently low (namely, if $c \leq \frac{4}{36}$). Contrary to the pure common-value case, there are positive costs for which student m searches. Specifically, student m visits two schools if $c \leq \frac{2}{36}$ and visits one school if $c \in \left(\frac{2}{36}, \frac{4}{36}\right]$. Despite student h visiting both schools and student m

[23] Solving for equilibrium search and reporting strategies becomes challenging with an arbitrary number of schools because the space of relevant types grows rapidly.

consequently facing adverse selection on the common component, potential gains in the idiosyncratic component incentivize search.

Student m's search negatively impacts student ℓ's expected value. When student m searches, the expected value of δ of her match increases. The zero-sum nature of the common component causes student m's expected gains in δ to be exactly offset by student ℓ's expected losses in δ. If there were no idiosyncratic component, student m would not search. Hence, adding an idiosyncratic component can incentivize search that exacerbates the adverse selection of student ℓ's *common component*. Welfare implications are stated in Corollary 2.

Corollary 2. *Equilibrium expected welfare is nonmonotonic and discontinuous in search costs under hybrid preferences. Discontinuities in expected welfare coincide with costs at which students change the number of schools they visit.*

While there are costs such that student m's search can reduce student ℓ's expected utility and overall welfare, there are also costs for which student m's search increases welfare. Since the total match utility from the δ_j's is constant, a social planner maximizing welfare matches students to schools based on their draws of S_{ij}. However, student m internalizes gains in δ_j and S_{ij}, but the planner internalizes only student m's gains in S_{ij} since the sum of expected common components is constant across matchings. Thus, it is ambiguous whether search increases or decreases welfare. As in the common-value case, discontinuities in welfare can occur only at breakpoints in search. Table 2.4 in the appendix decomposes each student's expected match utility into their idiosyncratic and common components, highlighting the misaligned incentives of student m and the social planner.

2.6 Coordinating Search Through Tie-Breaking

While the symmetric equilibrium is compelling under idiosyncratic preferences, it can generate inefficiencies. These inefficiencies stem from a coordination problem in which high-priority students crowd out seats that could be made available to low-priority students. High-priority students matriculate into schools at equal rates. Hence, schools fill to capacity from smallest to largest, and low-priority students are shut out from small schools. However, if high-priority students matriculate into larger schools at greater rates than to smaller schools, small schools do not fill up as quickly, granting some lower-priority students access to them.

For a range of costs under both search protocols, a planner can increase welfare by breaking indifferences in students' search strategies that accounts for students' endogenous search. The literature

has shown the importance of how clearinghouses break ties in reported preferences in complete-information settings.[24] In my setting, there are no indifferences in students' reports.[25] Rather, indifferences lie in which schools students visits. A planner can use a *noninvasive* design instrument in which they prescribe students with visit recommendations to select a particular (asymmetric) Nash equilibrium. I focus on a noninvasive instrument as a practical matter—it may be much easier for a school district to prescribe visit recommendations than to change how it operates the clearinghouse.[26] Even with simple recommendations, the school district can substantially decrease inequality and generate welfare improvements.[27]

Formally, consider idiosyncratic preferences and a school district with J schools, where $q_1 > q_2 > \cdots > q_J$.[28] The planner's recommendations are used to break students' ties in search. The planner maximizes expected welfare such that, in equilibrium, it is incentive compatible for each student to follow the planner's recommendation. The planner knows visit cost c. Given the planner's recommendations, students choose how many schools they visit and what preferences they submit to the clearinghouse.

2.6.1 Simultaneous Search

Under simultaneous search, students are indifferent to many feasible sets, and the planner prescribes each student with a particular one. The planner's prescriptions are constrained to constitute "true" feasible sets. That is, in equilibrium, if school k_j is in student i's feasible set, then student i is matched to school k_j if she reports it as her first choice. A student treats schools outside her feasible set as unattainable, hence does not visit them and ranks them at the bottom of her report to the clearinghouse. I retain the assumption that if a student is assigned feasible set D and if $\ell^*_D(c) < |D|$, the student visits $\ell^*_D(c)$ schools in D uniformly at random. This allows the planner to know how many students matriculate into each school given the feasible sets she assigns. Without knowing what fraction of students ultimately attend each school, the planner would not know the equilibrium set of schools to which each student can gain admission.

If costs are sufficiently low, the planner cannot overcome the coordination problem (without binding instructions). Student i is indifferent to all feasible sets of sizes $\ell^*_{D_i}(c) + 1$ or greater. If

[24] See Erdil and Ergin (2008) and Erdil and Ergin (2017).

[25] The coordination problem would exist even if students could submit reports with indifferences.

[26] Abdulkadiroğlu et al. (2006) describe a lengthy process to get the Boston School Committee to approve and implement changes to how it ran its clearinghouse.

[27] Because the instrument relies on indifferences, it is imperative that school utilities be drawn independently from the same distribution. This is plausible if schools are tiered and we restrict attention to schools within a tier that are ex-ante symmetric.

[28] There is little room for coordination in the common-value setting.

$\ell^*_{\{k_1,\ldots,k_J\}}(c) = J$ or $J-1$, the planner cannot change high-priority students' feasible sets without making them worse off. However, if costs are not too low, namely if $\ell^*_{\{k_1,\ldots,k_J\}}(c) \leq J-2$, the planner can shrink high-priority students' admissions sets to have size $\ell^* + 1$ without making them worse off. This allows the planner to control matriculation across schools and open up seats previously unavailable to low-priority students.

We can run a linear program (as is done for sequential search in the next subsection) to solve for the most efficient asymmetric Nash equilibrium subject to the planner's constraints. It turns out that if search costs are not too low and the largest school is not too large, the planner can induce a Pareto-improving Nash equilibrium in which *all* students have an feasible set containing $J-1$ schools. Moreover, the Coordinated Feasible Set Algorithm (CFS Algorithm) below returns the optimal feasible sets, as stated in Proposition 4a.[29] Figure 2.2 illustrates the algorithm's output for an example with four schools. The grid's x-axis denotes students' priorities. Student i's feasible set is the set of schools that lie above her priority.

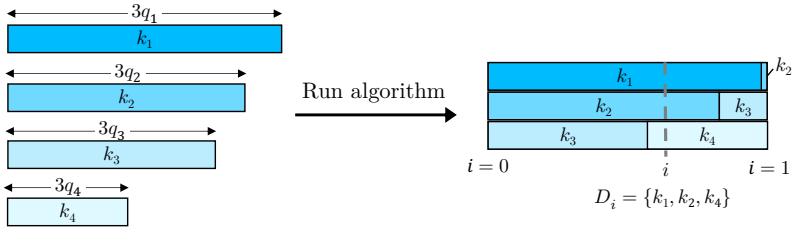
Coordinated Feasible Set Algorithm

Step 0. Consider $m \in \{2, 3, \ldots, J-1\}$ such that $\ell^*_{\{k_1,\ldots,k_J\}}(c) = m-1$. Assume $q_1 \leq \frac{1}{m}$. Assign all students empty feasible sets. Set a counter t equal to 0.

Step 1. Append school 1 to the feasible sets of students with $i \in [0, mq_1]$. Update $t \leftarrow t + mq_1$.

Step $n \geq 2$. If $t + mq_n \leq 1$, append school k_n to the feasible sets of students with $i \in (t, t+mq_n]$ then update $t \leftarrow t + mq_n$. If $t + mq_n > 1$, append school k_n to the feasible sets of students with $i \in (t, 1] \cup [0, t + mq_n - 1]$ then update $t \leftarrow t + mq_n - 1$.

Figure 2.2: Example feasible sets after running the CFS Algorithm for $J = 4$, $\ell^*_{\{k_1,\ldots,k_4\}}(c) = 2$, and $q_1 \leq \frac{1}{3}$.



Proposition 4a. *Consider costs such that $\ell^*_{\{k_1,\ldots,k_J\}}(c) = m-1$ for some $m \in \{2, 3, \ldots, J-1\}$ under idiosyncratic preferences and simultaneous search. If $q_1 \leq \frac{1}{m}$, the Coordinated Feasible Set*

[29] A district that generates students' priorities using a random lottery may find this result particularly appealing since it implies that a student with an low draw does no worse than one with a high draw.

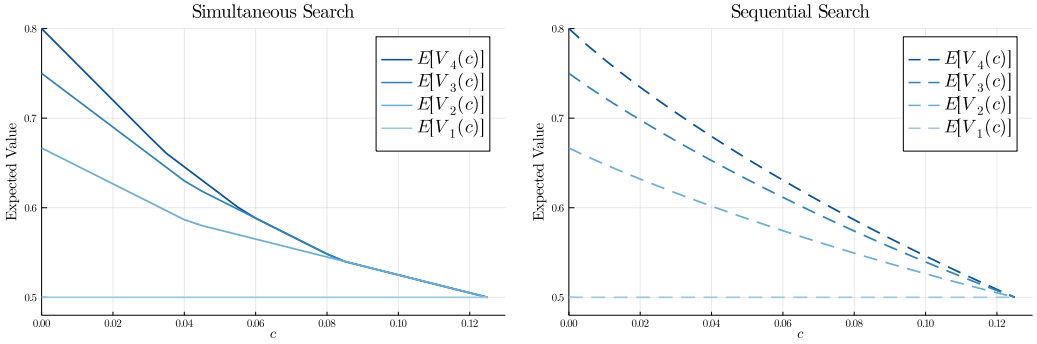
Algorithm outputs expected welfare-maximizing feasible set recommendations.

To see why the CFS Algorithm maximizes welfare, consider the example in Figure 2.2. No school is appended to a student's feasible set multiple times since $(J-1)q_j \leq 1 \ \forall j$. Also, all $J-1$ rows of the grid fill up since $\sum_{j=1}^{J}(J-1)q_j = J-1$. Hence, each student has $J-1=3$ different schools in her feasible set. Since $\ell^*_{\{k_1,\ldots,k_4\}}(c) = 2$, no student is strictly better off with an feasible set of more than three schools.

For costs and school capacities such that we can employ the CFS Algorithm, the improvement in welfare (over the symmetric equilibrium) is weakly decreasing in c. Students visit weakly more schools in the induced asymmetric Nash equilibrium than in the symmetric equilibrium. Hence, increasing c decreases the welfare difference.

The left panel of Figure 2.3 plots a student's expected value conditional on feasible set size, illustrating the benefits of giving students larger feasible sets. The figure also highlights the reduction in inequality in student outcomes. At costs greater than $\frac{11}{192} \approx 0.0573$, all students' expected values are equal to $\mathbf{E}[V_3(c)]$ instead of some being $\mathbf{E}[V_3(c)]$, some $\mathbf{E}[V_2(c)]$, and others $\mathbf{E}[V_1(c)]$.[30]

Figure 2.3: Students' expected values by feasible set size for $J=4$ and $S_{ij} \sim^{iid} \text{Uniform}[0,1]$.



2.6.2 Sequential Search

Under sequential search, students are indifferent to the order by which they visit schools, hence the planner prescribes them with visit orders. The planner cannot reduce students' feasible set sizes, as under simultaneous search, because that would not constitute a Nash equilibrium—ex ante, students strictly benefit from additional schools in their feasible sets.

The planner's objective and constraints are linear; hence, the planner can find recommendations

[30] Because the asymmetric Nash equilibrium's welfare improvement depends on the symmetric equilibrium's admission cutoffs, which depend on the *difference* in school sizes, comparative statics with respect to changes in q are ambiguous.

by solving the following linear program, which I refer to as the Visit Order Linear Program. Let $p_c^{|D|}(t)$ be the probability a student attends the t^{th} school on her visit order list when she has an feasible set of size $|D|$ and faces information cost c. We can calculate these probabilities using the equilibrium visit strategy in Proposition 1b. Let $\Sigma^j = \{\sigma_1^j, \sigma_2^j, \ldots, \sigma_{j!}^j\}$ be the set of permutations of $1, \ldots, j$. Let $\sigma_s^j(k)$ denote the position k takes in permutation σ_s^j. For example, if $\sigma_s^4 = (2, 3, 1, 4)$, then $\sigma_s^4(3) = 2$ since 3 occupies the second position. Let x_s^j denote the fraction of students with an feasible set of size j and execute a visit strategy according to the order given by $\sigma_s^j \in \Sigma^j$. The linear program is:

$$\begin{aligned}
\text{maximize } & \sum_{j=1}^{J} \left(\sum_{\sigma_s^j \in \Sigma^j} x_s^j \right) \mathbf{E}\left[V_j(c)\right] \\
\text{subject to } & \sum_{k=j}^{J} \left[\sum_{\sigma_s^j \in \Sigma^k} x_s^k p^k(\sigma_s^j(k)) \right] = q_j \quad \forall j = 1, \ldots, J \\
& \sum_{j=1}^{J} \sum_{\sigma_s^j \in \Sigma^j} x_s^j = 1 \\
& x_s^j \geq 0 \quad \forall s = 1, \ldots, j! \quad \forall j = 1, \ldots, J.
\end{aligned}$$

From a practical perspective, numerically solving this linear program is simple. The above program assumes schools fill to capacity from smallest to largest, providing a lower bound on welfare improvement.[31] To find the maximum welfare in general, one would need to check all possible orderings schools fill to capacity. As shown in Proposition 4b, the above linear program yields positive welfare gains for parameters such that the coordination problem is nontrivial.

Proposition 4b. *Consider $J \geq 3$ and $c \in \left(0, \frac{1}{8}\right]$ under idiosyncratic preferences and sequential search. The solution of the Visit Order Linear Program induces a Nash equilibrium with strictly greater expected welfare than that of the symmetric equilibrium.*

When $J = 2$, students attend each school with $\frac{1}{2}$ probability regardless of their order, so there is no room for coordination. At extreme costs, the social planner has no impact on students. If $c = 0$, students visit every school. They matriculate evenly across schools in their feasible sets regardless of their visit orders. If $c > \frac{1}{8}$, students do not visit any schools. When $J = 3$ and $c \in \left(0, \frac{1}{8}\right]$, however, students are more likely to attend schools earlier in their visit order. This allows the planner to assign visit orders such that students visit larger schools at higher rates than smaller schools. The right panel of Figure 2.3 shows students' expected values by feasible set size and search cost.

[31] In numerous simulations, this turns out to be the optimal ordering.

Consider an example in which $q = (\frac{4}{10}, \frac{3}{10}, \frac{2}{10}, \frac{1}{10})$, $c = 0.06$, and $S_{ij} \sim^{iid}$Uniform[0,1]. In the symmetric equilibrium, 40% of students have four schools in their feasible set, 30% have three schools, 20% have two schools, and 10% have one.[32] Using optimal recommendations increases expected utility for more than half of students (55.3%), including all students with only one or two schools in their feasible set in the symmetric equilibrium. Optimal recommendations provide 65.3% of students with four schools in their feasible sets, 28.3% with three, 6.4% with two, and no student with only one. In total, welfare increases by 3.5%.

2.7 Conclusion

I present a model of centralized school choice in which students pay information costs to learn their utilities for attending schools. When preferences are idiosyncratic, there is no externality of search. All students have higher expected utility under sequential search than simultaneous search. Under both protocols, students' expected utilities decrease in costs.

When preferences over schools are common, there are search externalities, which influence equilibrium behavior. As high-priority students search, remaining seats are adversely selected, making lower-priority students worse off. This adverse selection disincentivizes lower-priority students from conducting any search in equilibrium.

The search protocol also has important equilibrium effects. Under common preferences, there are costs at which sequential search yields *lower* welfare than simultaneous search. At low costs, low-priority students have higher expected utility under sequential search, but at intermediate costs, they have higher expected utility under simultaneous search. Under both search protocols, welfare is nonmonotonic in costs.

Hybrid preferences generate an interplay between idiosyncratic and common components. Adding an idiosyncratic component can incentivize search that is not incentivized under pure common preferences. Search has ambiguous welfare effects. More search could increase welfare by inducing a matching that better aligns with idiosyncratic preferences but could decrease welfare when a student's gains in search come from the common component. In which case, search can increase student inequality by assigning a lower-priority student a seat that is more adversely selected.

Symmetric strategies present a coordination problem in which high-priority students crowd out seats, and I provide a simple design instrument a planner can use to induce an asymmetric Nash equilibrium with greater welfare than the symmetric equilibrium. The planner coordinates students'

[32] It is a coincidence that q and the fractions of students with each feasible set size are so similar.

strategies by providing them with visit recommendations. Under simultaneous search, the planner can run a simple algorithm to solve for feasible sets that give each student an expected utility equal to that of the highest-priority student. Under sequential search, the planner prescribes visit orders and can run a linear program to solve for prescriptions that maximize student welfare.

My analysis suggests the importance of considering students' incomplete information when studying centralized school choice. Under endogenous information acquisition, students' preferences, the search protocol, and information costs all affect equilibrium outcomes. There is a lot of interesting work to be done on these and other facets of school choice with incomplete information, which I leave for future research.

2.8 Appendix

2.8.1 Proofs of Main Results

Proof of Lemma 1

Consider student i with visit set L_i. Others students' utilities are independent of student i's utilities. Hence, $\mathbf{E}\left[S_{ij}|\{s_{im}\}_{k_m \in L_i}\right]$ is student i's expected value of school k_j conditional on being matched to it, regardless of other students' search and reporting strategies. If student i visited school k_j, then $\mathbf{E}\left[S_{ij}|\{s_{im}\}_{k_m \in L_i}\right] = s_{ij}$. If she did not, then $\mathbf{E}\left[S_{ij}|\{s_{im}\}_{k_m \in L_i}\right] = \mathbf{E}\left[S_{ij}\right]$.

By strategy-proofness of DA for proposers, it is a dominant strategy for student i to report schools truthfully, from highest $\mathbf{E}\left[S_{ij}|\{s_{im}\}_{k_m \in L_i}\right]$ to lowest. The student is indifferent between schools with the same $\mathbf{E}\left[S_{ij}|\{s_{im}\}_{k_m \in L_i}\right]$. Hence, tied schools can be ordered in any way among themselves. Per the equilibrium refinement, she orders ties uniformly at random.

Proof of Lemma 2

Consider the individual search problem in which student i is matched to her highest-ranked school in some feasible set D_i.

To reduce notational clutter, suppress the i subscripts on S_{ij}. Recall that S_j has cumulative density function $F(\cdot)$ on the support $[0, b]$, for finite b, and probability density function $f(\cdot)$. Both are well behaved. Let $\tilde{S}^{(\ell)}$ denote the highest order statistic of ℓ-many S_j variables. The probability density function of $\tilde{S}^{(\ell)}$ is given by $\tilde{f}_\ell(x) = \ell f(x) F(x)^{\ell-1}$.

Define

$$\ell^*_{D_i}(c) = \max\left\{ \underset{\ell \in \{0,\dots,|D_i|\}}{\operatorname{argmax}} \mathbf{E}\left[\max\{\mathbf{E}\left[S^{(\ell)}\right], \mathbf{1}(\ell < |D_i|)\mathbf{E}\left[S_j\right]\}\right] - c\ell \right\},$$

where $S^{(\ell)}$ is the ℓ^{th}-highest order statistic of ℓ-many S_{ij} variables. The maximization problem is well-defined and over a finite support, hence has a solution, given by $\ell^*_{D_i}(c)$. Because $\ell^*_{D_i}(c)$ is the largest among the solutions, it is unique.

To prove that $\ell^*_{D_i}(c)$ is decreasing, it is sufficient to show that visiting more schools has diminishing returns in match utility. I shorthand $\mathbf{E}\left[S_j\right]$ as $\mathbf{E}\left[S\right]$. Slightly abusing notation, let $\mathbf{E}\left[U(\ell)\right]$ be the student's expected match utility after visiting ℓ schools in D_i and reporting optimally. The student's expected utility for the strategy of visiting ℓ schools is $\mathbf{E}\left[\max\{\mathbf{E}\left[S^{(\ell)}\right], \mathbf{1}(\ell < |D_i|)\mathbf{E}\left[S_j\right]\}\right] - c\ell$. This can be broken into two cases. If $\ell < |D_i|$, the expected match utility is

$$\mathbf{E}\left[U(\ell)\right] = \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell)}\}\right] = b - \int_{\mathbf{E}[S]}^{b} F(x)^{\ell} dx.$$

If $\ell = |D_i|$, the expected match utility reduces to

$$\mathbf{E}\left[U(|D_i|)\right] = \mathbf{E}\left[\tilde{S}^{(|D_i|)}\right] = b - \int_{0}^{b} F(x)^{|D_i|}\, dx.$$

Let $\Delta_{\ell+1} = \mathbf{E}\left[U(\ell+1)\right] - \mathbf{E}\left[U(\ell)\right]$ be the marginal returns (of match utility) of visiting the $(\ell+1)^{th}$ school. For $\ell \in \{1, \dots, |D_i| - 1\}$, $\Delta_{\ell+1} \leq \Delta_\ell$ if and only if

$$\begin{aligned}
&\mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell+1)}\}\right] - \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell)}\}\right] \leq \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell)}\}\right] - \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell-1)}\}\right] \\
&\iff \int_{\mathbf{E}[S]}^{b} F(x)^{\ell}(1 - F(x))dx \leq \int_{\mathbf{E}[S]}^{b} F(x)^{\ell-1}(1 - F(x))dx,
\end{aligned}$$

which is true since $F(x) \in [0,1]\ \forall x \in [0,b]$.

For $\ell = |D_i| - 1$, we have that $\Delta_{\ell+1} \leq \Delta_\ell$ if and only if

$$\begin{aligned}
&\mathbf{E}\left[\tilde{S}^{(\ell+1)}\right] - \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell)}\}\right] \leq \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell)}\}\right] - \mathbf{E}\left[\max\{\mathbf{E}\left[S\right], \tilde{S}^{(\ell-1)}\}\right] \\
&\iff 2\int_{\mathbf{E}[S]}^{b} F(x)^{\ell} dx - \int_{\mathbf{E}[S]}^{b} F(x)^{\ell-1} dx - \int_{\mathbf{E}[S]}^{b} F(x)^{\ell+1} dx - \int_{0}^{\mathbf{E}[S]} F(x)^{\ell+1} dx \leq 0 \\
&\iff \int_{\mathbf{E}[S]}^{b} F(x)^{\ell-1} \underbrace{\left(2F(x) - 1 - F(x)^2\right)}_{=-(1-F(x))^2 \leq 0} dx \underbrace{- \int_{0}^{\mathbf{E}[S]} F(x)^{\ell+1} dx}_{\leq 0} \leq 0,
\end{aligned}$$

which is true. Hence, for all ℓ, returns in match utility diminish in the number of schools visited. Therefore, $\ell^*_{D_i}(c)$ is weakly decreasing in c.

Proof of Proposition 1a

I show existence of an equilibrium by construction. Since there is a continuum of students, there is no aggregate uncertainty. Since $\sum_{j=1}^{J} q_j = 1$ and all students and schools are acceptable, a mass of q_j students ultimately attend school k_j. By the symmetry refinements, students matriculate to schools evenly across all schools in their feasible sets. Then admission cutoffs are uniquely pinned down:

$$\begin{aligned}
r_J &= 1 - Jq_J \\
r_{J-1} &= 1 - r_J - (J-1)(q_{J-1} - q_J) \\
&\vdots \\
r_j &= 1 - r_{j+1} - j(q_j - q_{j+1}) \\
&\vdots \\
r_2 &= 1 - r_3 - 2(q_2 - q_3) \\
r_1 &= 0.
\end{aligned}$$

Changing any of these cutoffs would lead to at least one school being either over- or under-subscribed, hence not constitute an equilibrium.[33] Note that admission cutoffs depend only on q, hence are independent of c.

Given admission cutoffs r, Lemma 2 implies that student i maximizes utility by visiting $\ell^*_{D_i}(c)$ schools in their feasible set. By the refinement, she visits schools in her feasible set uniformly at random. Hence, her search strategy is uniquely identified to be the Sim-Max strategy. This is true for all students i, implying that the UTB-equilibrium is unique.

Proof of Corollary 1

Suppose student i has feasible set D_i and that $|D_i| \geq 2$. By the proof of Proposition 1a, student i's feasible set is constant across all costs. For notational convenience, I drop D_i from $\ell^*_{D_i}(c)$. Let $\mathbf{E}[V_c(\ell)]$ denote student i's expected value when she faces visit cost c and visits $\ell \leq |D_i|$ schools in

[33] One can check this recursively by looking at the mass of students attending each school. Start with school k_J, then school k_{J-1}, and so on. Any change in r must make at least one school over- or under-subscribed.

D_i, reporting optimally.

Consider two costs $c < c'$. If $\ell^*(c) = \ell^*(c')$, the student has the same expected match utility across both costs but pays more in visit costs under c'. Hence, $\mathbf{E}[V_{c'}(\ell)] \leq \mathbf{E}[V_c(\ell)]$, strict if $\ell^*(c) > 0$.

If $\ell^*(c) \neq \ell^*(c')$, then $\ell^*(c) > \ell^*(c')$ since $\ell^*(\cdot)$ is decreasing. Then

$$\begin{aligned}\mathbf{E}[V_{c'}(\ell^*(c'))] - \mathbf{E}[V_c(\ell^*(c))] &\leq \mathbf{E}[V_{c'}(\ell^*(c'))] - \mathbf{E}[V_c(\ell^*(c'))] \\ &= (c - c')\ell^*(c') \leq 0.\end{aligned}$$

Hence, $\mathbf{E}[V_c(\ell^*(c))]$ is decreasing in c.

Proof of Proposition 1b

The distribution of $S_{ij} - \frac{1}{2}$ is symmetric about 0, and all schools have the same visit cost. Using the terminology of Doval (2018), if $c \leq \frac{1}{8}$, the "backup value" of a school is $\sqrt{2c}$, which is weakly less than a school's expected utility. Hence, Proposition 5 of Doval (2018) implies a student uses the strategy in Theorem 1 of Doval (2018). Schools are symmetric, so visits are ordered uniformly at random, which implies that students matriculate evenly across schools in their feasible sets. Therefore, the equilibrium admission cutoffs are the same as in Proposition 1a, hence independent of c. Again, the refinement uniquely pins down students' search strategies, so this is a unique UTB-equilibrium.

Proof of Remark 1

Per the proofs of Propositions 1a and 1b, the admission cutoffs are the same across both protocols. Hence, students' feasible sets are the same under both search protocols. All strategies under simultaneous search are feasible under sequential search, hence all students do at least as well under sequential search in expectation. Therefore, $\mathbf{E}[W^{\text{seq}}(c)] \geq \mathbf{E}\left[W^{\text{sim}}(c)\right]$.

Proof of Proposition 2

I show the proof for simultaneous search and then for sequential search. Throughout, I consider $c > 0$. (If $c = 0$, we are in the complete-information case.) In all that follows, I ignore knife-edge events that occur with zero probability (for example, a realization of δ_j being equal to a particular number).

Simultaneous Search

H students have the same decision problem as in the idiosyncratic-utility case. They visit 2 schools if $c \in \left(0, \frac{1}{24}\right]$, 1 school if $c \in \left(\frac{1}{24}, \frac{1}{8}\right]$, and 0 schools if $c > \frac{1}{8}$.

To solve for M and L students' equilibrium strategies, I go through the three mutually exclusive, collectively exhaustive cases based on the number of schools H students visit. When H students visit zero schools or two schools, the strategies are straightforward, so I start with those.

If H students visit 0 schools, they uniformly randomize between reporting $k_1 \succ k_2$ and $k_2 \succ k_1$. Half attend k_1, and the other half attend k_2. Then M and L students with $i \geq 1 - 2q_2$ have the same decision problem as the H students, hence employ the same strategy of no search and reporting randomly. Then k_2 fills to capacity at priority $1 - 2q_2$. Thus, all students with $i < 1 - 2q_2$ attend k_1 regardless of their reports, hence do not search in equilibrium.

Suppose H students visit both schools. Consider an M student with $i = 1 - q_2 - \varepsilon$ for small $\varepsilon > 0$. By reasoning in Section 5, the student is best off visiting no schools and simply reporting $k_1 \succ k_2$. The same is true for all lower-priority students M students. Now consider an L student with $i = 1 - q_1 - \varepsilon$ for small $\varepsilon > 0$. Once the algorithm reaches this student, only one school has vacancies. Hence, she does not search in equilibrium. The same is true for all lower-priority L students.

The only case that remains is H students visiting exactly one school, which occurs when $c \in \left(\frac{1}{24}, \frac{1}{8}\right]$. By the refinement, half of H students visit k_1 and the other half k_2. If a student visits k_j, she reports $k_j \succ k_{-j} \iff \delta_j > \frac{1}{2}$.

There are 4 subcases to consider. Table 2.1 shows, for each subcase, the mass of vacant seats remaining at each school after H students have searched and been matched. I start by focusing on a high-priority M student with $i = 1 - q_2 - \varepsilon$ for small $\varepsilon > 0$. I can make ε small enough such that all nonzero entries in Table 2.1 remain nonzero regardless of the actions of the ε-mass of higher-priority M students.

Table 2.1: Mass of vacant seats after H students and M students have searched and been matched, in equilibrium. In each cell, the top entry is the mass of seats available at school 1 and the bottom entry is the mass of seats available at school 2.

	After H students are matched		After M students are also matched	
	$\delta_1 < \frac{1}{2}$	$\delta_1 > \frac{1}{2}$	$\delta_1 < \frac{1}{2}$	$\delta_1 > \frac{1}{2}$
$\delta_2 < \frac{1}{2}$	$q_1 - \frac{q_2}{2}$ $\frac{q_2}{2}$	$q_1 - q_2$ q_2	$\frac{q_2}{2}$ $\frac{q_2}{2}$	0 q_2
$\delta_2 > \frac{1}{2}$	q_1 0	$q_1 - \frac{q_2}{2}$ $\frac{q_2}{2}$	q_2 0	$\frac{q_2}{2}$ $\frac{q_2}{2}$

Some algebra shows that this student employs the following strategy conditional on H students visiting

one school: visit 2 schools if $c \in \left(0, \frac{1}{48}\right]$ and 0 schools if $c > \frac{1}{48}$. Since $c \in \left(\frac{1}{24}, \frac{1}{8}\right]$, this student does not search and simply reports $k_1 \succ k_2$, giving an expected utility of $\frac{1}{2}$. (Reporting $k_2 \succ k_1$ yields and expected utility of $\frac{3}{8}$.)

Now consider a student i who is an M student with priority lower than $1 - q_2 - \varepsilon$. If all M students with priority higher than i employ the above strategy, student i faces the same decision problem described above, hence in equilibrium employs the same strategy.

Turn to an L student with $i = 1 - q_1 - \varepsilon$ for small $\varepsilon > 0$. The right two columns of Table 2.1 show the mass of remaining seats at each after M students employ the above strategy. Some algebra shows that visiting zero schools yields student i an expected value of $\frac{3}{8}$, visiting one school yields $\frac{13}{32} - c$, and visiting both schools yields $\frac{5}{12} - 2c$.[34]

For $c \in \left(\frac{1}{24}, \frac{1}{8}\right]$, this students is best off visiting no schools. By symmetry, reporting $k_1 \succ k_2$ and reporting $k_2 \succ k_1$ yield the same expected utility. Hence, a student with $i = 1 - q_1 - \varepsilon$ mixes between the two reports. Such a student has an equal probability of attending either school. As a result, all subsequent L students face the same decision problem, hence employ the same strategy.

Sequential Search

The strategy of the H students is given by the idiosyncratic case. For M and L students, I keep track of various cases using Table 2.2, which shows the mass of seats remaining at each school conditional on the realizations of δ_1 and δ_2.

Table 2.2: Mass of vacant seats after H students and M students have searched and been matched, in equilibrium. In each cell, the top entry is the mass of seats available at school 1 and the bottom entry is the mass of seats available at school 2.

	$\delta_1 < \sqrt{2c}$	$\delta_1 \in \left(\sqrt{2c}, 1 - \sqrt{2c}\right)$	$\delta_1 > 1 - \sqrt{2c}$
$\delta_2 < \sqrt{2c}$	$q_1 - \frac{q_2}{2}$ $\frac{q_2}{2}$	$q_1 - q_2$ q_2	$q_1 - q_2$ q_2
$\delta_2 \in \left(\sqrt{2c}, 1 - \sqrt{2c}\right)$	q_1 0	H students attend the better school	$q_1 - q_2$ q_2
$\delta_2 > 1\sqrt{2c}$	q_1 0	q_1 0	$q_1 - \frac{q_2}{2}$ $\frac{q_2}{2}$

Consider an M student with $i = 1 - q_2 - \varepsilon$ for small $\varepsilon > 0$. Suppose she has visited school 1 and is considering whether to visit school 2. There are three cases.

Case 1: $\delta_1 < \sqrt{2c}$

Let $W(\cdot)$ denote the expected payoff for employing the specified strategy ignoring the sunk cost of already

[34] By symmetry, visiting only k_1 yields the same expected value as visiting only k_2.

visiting school 1.

$$\begin{aligned}
\mathbf{E}\left[W(\text{report } k_1 \succ k_2)|\delta_1\right] &= \delta_1 \\
\mathbf{E}\left[W(\text{report } k_2 \succ k_1)|\delta_1\right] &= P(\delta_2 < \sqrt{2c})\mathbf{E}\left[\delta_2|\delta_2 < \sqrt{2c}\right] + P(\delta_2 > \sqrt{2c})\delta_1 \\
&= c + (1-\sqrt{2c})\delta_1 \\
\mathbf{E}\left[W(\text{visit } k_2, \text{ report truth})|\delta_1\right] &= \sqrt{2c}\mathbf{E}\left[\max\{\delta_2, \delta_2\}|\delta_1; \delta_2 < \sqrt{2c}\right] + (1-\sqrt{2c})\delta_1 - c \\
&= \frac{1}{2}\delta_1^2 + (1-\sqrt{2c})\delta_1
\end{aligned}$$

If $\delta_1 < \frac{1}{2}\sqrt{2c}$, it is a best response to report $k_2 \succ k_1$. If $\frac{1}{2}\sqrt{2c} < \delta_1 < \sqrt{2c}$, the best response is reporting $k_1 \succ k_2$. In either case, she is not incentivized to visit k_2.

Case 2: $\sqrt{2c} < \delta_1 < 1 - \sqrt{2c}$

$$\begin{aligned}
\mathbf{E}\left[W(\text{report } k_1 \succ k_2)|\delta_1\right] &= \delta_1 \\
\mathbf{E}\left[W(\text{report } k_2 \succ k_1)|\delta_1\right] &= P(\delta_2 < \sqrt{2c})\mathbf{E}\left[\delta_2|\delta_2 < \sqrt{2c}\right] \\
&\quad + P(\delta_2 \in (\sqrt{2c}, 1-\sqrt{2c})) \\
&\quad\quad * \mathbf{E}\left[\min\{\delta_1, \delta_2\}|\delta_1; \delta_2 \in (\sqrt{2c}, 1-\sqrt{2c})\right] \\
&\quad + P(\delta_2 > 1 - \sqrt{2c})\delta_1 \\
\mathbf{E}\left[W(\text{visit } k_2, \text{ report truth})|\delta_1\right] &= \delta_1 - c
\end{aligned}$$

$\mathbf{E}\left[W(\text{report } k_1 \succ k_2)|\delta_1\right] \geq \mathbf{E}\left[W(\text{report } k_2 \succ k_1)|\delta_1\right]$ since the former has match utilities at least as large under every realization of δ_2. Thus, the student does not continue search; she simply reports $k_1 \succ k_2$.

Case 3: $\delta_1 > 1 - \sqrt{2c}$

$$\begin{aligned}
\mathbf{E}\left[W(\text{report } k_1 \succ k_2)|\delta_1\right] &= \delta_1 \\
\mathbf{E}\left[W(\text{report } k_2 \succ k_1)|\delta_1\right] &= \mathbf{E}\left[\delta_2\right] = \frac{1}{2} \\
\mathbf{E}\left[W(\text{visit } k_2, \text{ report truthfully})|\delta_1\right] &= \mathbf{E}\left[\max\{\delta_1, \delta_2\}|\delta_1\right] - c \\
&= \frac{1}{2}(1+\delta_1^2) - c
\end{aligned}$$

Since $\delta_1 > 1 - \sqrt{2c}$, the student is best off ending search and simply reporting $k_1 \succ k_2$.

To recap, the strategy after visiting school 1 is to report $k_2 \succ k_1$ if $\delta_1 < \frac{1}{2}\sqrt{2c}$ and report $k_1 \succ k_2$ if

$\delta_1 > \frac{1}{2}\sqrt{2c}$. In either case, do not engage in more search. The expected value of this strategy is

$$\begin{aligned}\mathbf{E}\left[V(\text{visit } k_1, \text{ report optimally})\right] &= P\left(\delta_1 < \frac{1}{2}\sqrt{2c}\right)\mathbf{E}\left[c + (1-\sqrt{2c})\delta_1 | \delta_1 < \frac{1}{2}\sqrt{2c}\right] \\ &\quad + P\left(\delta_1 > \frac{1}{2}\sqrt{2c}\right)\mathbf{E}\left[\delta_1 | \delta_1 > \frac{1}{2}\sqrt{2c}\right] - c \\ &= \frac{1}{2} + \left(\frac{c}{2}\right)^{3/2} - c,\end{aligned}$$

which is less than $\mathbf{E}\left[V(\text{no visits, report } k_1 \succ k_2)\right] = \frac{1}{2}$ for $c \in \left(0, \frac{1}{8}\right]$. Hence, visiting k_1 at the outset is not an equilibrium strategy.

Repeating this exercise but for visiting k_2 first instead of k_1 yields an analogous best response: report $k_1 \succ k_2$ if $\delta_2 < 1 - \frac{1}{2}\sqrt{2c}$ and report $k_2 \succ k_1$ if $\delta_2 > 1 - \frac{1}{2}\sqrt{2c}$. In either case, do not engage in more search. By similar algebra, the student is not incentivized to visit school 2 in the first place. Rather, she is better off visiting no schools and blindly reporting $k_1 \succ k_2$.

The above is for an M student with $i = 1 - q_2 - \varepsilon$. In equilibrium, all lower-priority M students face the same decision problem, hence employ the same strategy.

Now consider an L student with $i = 1 - q_1 - \varepsilon$ for small $\varepsilon > 0$. She does not have guaranteed admission to either school. Table 2.3 shows the various cases of the mass of vacant seats at each school when the algorithm reaches this student. The student's match is already determined except when $[\delta_1, \delta_2 < \sqrt{2c}]$ or $[\delta_1, \delta_2 > 1 - \sqrt{2c}]$.

Table 2.3: Mass of vacant seats after H students and M students have searched and been matched, in equilibrium. In each cell, the top entry is the mass of seats available at school 1 and the bottom entry is the mass of seats available at school 2.

	$\delta_1 < \sqrt{2c}$	$\delta_1 \in \left(\sqrt{2c}, 1-\sqrt{2c}\right)$	$\delta_1 > 1 - \sqrt{2c}$
$\delta_2 < \sqrt{2c}$	$\frac{q_2}{2}$	0	0
	$\frac{q_2}{2}$	q_2	q_2
$\delta_2 \in \left(\sqrt{2c}, 1-\sqrt{2c}\right)$	q_2	Only lower school	0
	0	is vacant	q_2
$\delta_2 > 1\sqrt{2c}$	q_2	q_2	$\frac{q_2}{2}$
	0	0	$\frac{q_2}{2}$

Suppose this student has already visited k_1. There are three cases to consider.

Case 1: $\delta_1 < \sqrt{2c}$

This is the same as Case 1 for the M student with $i = 1 - q_2 - \varepsilon$.

Case 2: $\sqrt{2c} < \delta_1 < 1 - \sqrt{2c}$

Regardless of her report, her match utility is fixed at $\min\{\delta_1, \delta_2\}$, so there are no benefits of further search.

Case 3: $\delta_1 > 1 - \sqrt{2c}$

This is analogous to Case 1 for the M student with $i = 1 - q_2 - \varepsilon$ but for high values of δ_1 instead of

low. The student reports that she prefers k_1 if $\delta_1 > 1 - \frac{1}{2}\sqrt{2c}$; otherwise she reports that she prefers k_2. Again, it is not worth visiting k_2.

Thus, if a student has already visited k_1, she is not incentivized to visit k_2. Comparing $\mathbf{E}\left[V(\text{visit } k_1, \text{ report optimally})\right]$ to visiting no schools shows that

$$\begin{aligned}
&\mathbf{E}\left[V(\text{visit } k_1, \text{ report optimally})\right] - \mathbf{E}\left[V(\text{no visits})\right] \\
&\quad = P(\delta_2 < \sqrt{2c}) \int_0^{\frac{1}{2}\sqrt{2c}} \frac{1}{2}\sqrt{2c} - \delta_1 \; d\delta_1 \\
&\qquad + P(\delta_2 > 1 - \sqrt{2c}) \int_{1-\sqrt{2c}}^{1-\frac{1}{2}\sqrt{2c}} 1 - \frac{1}{2}\sqrt{2c} - \delta_1 \; d\delta_1 - c \\
&\quad = -c\left(1 - \sqrt{\frac{c}{2}}\right),
\end{aligned}$$

which is less than zero for $c \in \left(0, \frac{1}{8}\right]$, so the student is better off not visiting any schools rather than visiting k_1. By symmetry, $\mathbf{E}\left[V(\text{visit } k_1, \text{ report optimally})\right] = \mathbf{E}\left[V(\text{visit } k_2, \text{ report optimally})\right]$. Thus, the student does not search in equilibrium and randomizes between $k_1 \succ k_2$ and $k_2 \succ k_1$. In equilibrium, all lower-priority L students face the same decision problem, hence employ the same strategy.

Proof of Proposition 3

There are exactly enough seats for all students. Therefore, the total expected match utility across all students is constant at $\mathbf{E}\left[\delta_j\right] = \frac{1}{2}$. If H^* students have higher expected match utilities under simultaneous search than under sequential search, then L^* students have lower expected match utilities under simultaneous search than under sequential search and vice versa. H^* students are defined such that they match to their first choice regardless of other students' reports. Hence, their feasible sets are the set of all schools. H^* students employ the respective Sim-Max or Seq-Max strategy in equilibrium.

Consider $c > 0$ sufficiently low such that $\ell^*_{D_i}(c) = J$ for H^* students. In equilibrium under simultaneous search, H^* students visit every school and attend the best one. In equilibrium under sequential search, there is a positive probability an H^* student does not match with the best school. Hence, H^* students have higher expected match utility under simultaneous search than under sequential search, implying that L^* students have lower expected match utility under simultaneous search than under sequential search.

Now consider $c \in \left(\frac{1}{12}, \frac{1}{8}\right)$, which is sufficiently high for $\ell^*_{D_i}(c) = 1$ for H^* students. For search protocol $\alpha \in \{\text{sim}, \text{seq}\}$, denote H^* students' expected value by $\mathbf{E}\left[V^\alpha(c)\right]$, H^* students' expected match utility by $\mathbf{E}\left[U^\alpha(c)\right]$, and H^* students' expected number of schools visited $\mathbf{E}\left[L^\alpha(c)\right]$. One can show that $\mathbf{E}\left[L^{\text{seq}}(c)\right] > 1$

for $c \in \left(\frac{1}{12}, \frac{1}{8}\right)$. Moreover, $\mathbf{E}\left[V^{\text{seq}}(c)\right] \geq \mathbf{E}\left[V^{\text{sim}}(c)\right]$ for all c. Thus,

$$\begin{aligned}
\mathbf{E}\left[U^{\text{seq}}(c)\right] &= \mathbf{E}\left[V^{\text{seq}}(c)\right] + c\mathbf{E}\left[L^{\text{seq}}(c)\right] \\
&\geq \mathbf{E}\left[V^{\text{sim}}(c)\right] + c\mathbf{E}\left[L^{\text{seq}}(c)\right] \\
&> \mathbf{E}\left[V^{\text{sim}}(c)\right] + c \\
&= \mathbf{E}\left[U^{\text{sim}}(c)\right].
\end{aligned}$$

Because H^* students have higher expected match utilities under sequential search than under simultaneous search, L^* students have lower match utilities under sequential search than under simultaneous search.

Proof of Proposition 4a

Consider $m \in \{2, 3, \ldots, J-1\}$ such that $\ell^*_{\{k_1,\ldots,k_J\}}(c) = m - 1$. Then all students are indifferent to feasible sets of size m and greater. By reasoning stated in the text, if $q_1 \leq \frac{1}{m}$, the CFS Algorithm assigns each student an feasible set of size m. Hence, such feasible set recommendations maximize welfare.

Proof of Proposition 4b

Suppose $J = 3$. Let μ^n be the mass (or fraction) of students with an feasible set of size n in the symmetric equilibrium. Then $\mu^3 = 3q_3$ and $\mu^2 = 2(q_2 - q_3)$. It is sufficient to provide a feasible solution such that strictly more students have an feasible set of size 3 and the same fraction of students with an feasible set of size 2. Below I provide such a solution. Let $\bar{\mu}^n_j$ denote the mass of students with an feasible set of size n and ultimately attend school k_j for that proposed solution, and let $\bar{\mu}^n \equiv \sum_{j=1}^n \bar{\mu}^n_j$ be the total mass of students with an feasible set of size n.

Let $p_c^{|D|}(t)$ be the probability a student attends the t^{th} school on her visit order list when she has an feasible set of size $|D|$ and faces visit cost c. For $|D| \geq 3$:

$$p_c^{|D|}(t) = \begin{cases} (1-\sqrt{2c})^{t-1}(\sqrt{2c}) + \frac{(1-\sqrt{2c})^{|D|} - (\sqrt{2c})^{|D|}}{|D|} & t < |D| \\ \left[(1-\sqrt{2c})^{|D|-1} + (\sqrt{2c})^{|D|-1}\right](\sqrt{2c}) + \frac{(1-\sqrt{2c})^{|D|} - (\sqrt{2c})^{|D|}}{|D|} & t = |D| \end{cases}.$$

For $|D| \geq 3$ and $c \in \left(0, \frac{1}{8}\right]$, $p_c^{|D|}(1) > p_c^{|D|}(2) \geq p_c^{|D|}(3)$; students are more likely to ultimately attend schools early on their visit lists. Consider the following assignment for a mass of students with $|D| = 3$. Let $\frac{1}{6} + \varepsilon$, for small $\varepsilon > 0$, fraction each be assigned to visit orders (1,2,3) and (1,3,2). Let $\frac{1}{6} - \frac{\varepsilon}{2}$ fraction be assigned each remaining visit order: (2,1,3), (2,3,1), (3,1,2), and (3,2,1). By symmetry, $\bar{\mu}^3_2 = \bar{\mu}^3_3$. Because students are likely to attend the school they visit first, slightly more than $\frac{1}{3}$ of students with $|D| = 3$ attend school 1 and slightly fewer than $\frac{1}{3}$ attend school 3.

Sort students in this way until school 3 is filled to capacity. Then $\bar{\mu}^3_3 = q_3$. Since fewer than one-third

of $|D| = 3$ students attend school 3, $\bar{\mu}^3 > \mu^3$. The above fills schools from smallest to largest as long as ε is small enough such that $q_1 - \bar{\mu}_1^3 > q_2 - \bar{\mu}_2^3$, which is possible since $q_1 > q_2$.

Once school 3 is filled, students split evenly across school 1 and school 2 until school 2 fills to capacity since $p_c^2(1) = p_c^2(2) = \frac{1}{2}$. Then $\bar{\mu}^2 = 2(q_2 - \bar{\mu}_2^3) = 2(q_2 - \bar{\mu}_3^3) = 2(q_2 - q_3) = \mu^2$. That is, the same fraction of students have $|D| = 2$ in the proposed assignment as in the symmetric equilibrium.

Assign all remaining students attend to school 1. The proposed assignment is feasible and provides strictly greater welfare than the symmetric equilibrium. Hence, the linear program's solution must also provide strictly greater welfare. For $J > 3$, have students employ a symmetric strategy until school 4 is filled. Implement an assignment analogous to the above for remaining students.

Proof of Corollary 2 and Utility Decomposition of Hybrid Model

Table 2.4 decomposes each student's expected match utility into their idiosyncratic and common components for costs $c \in \left(0, \frac{4}{36}\right]$. In equilibrium at these costs, student h visits both schools, and student ℓ does not visit either. Table 2.4's rows correspond to the number of schools student m visits. Conditional on visiting one school, it does not matter which she visits—the table's entries remain the same. For $c \leq \frac{2}{36}$, student m is incentivized to visit both schools as opposed to just one.

Increased search at $c = \frac{2}{36}$ causes a discontinuous jump down in welfare (last column of the table) since the additional visit cost is larger than student m's gains in $\mathbf{E}\left[S_{2\nu(2)}\right]$. As under common preferences, welfare under hybrid preferences has discontinuities and is nonmonotonic in costs. Holding the number of visits constant, expected welfare is continuous in search costs. Hence, discontinuities may occur only at costs that align with breakpoints in search.

Table 2.4: Decomposition of expected utilities and expected welfare for $c \in \left(0, \frac{4}{36}\right]$.

Student m visits	$\mathbf{E}\left[\delta_{\mu(1)}\right]$	$\mathbf{E}\left[S_{1\mu(1)}\right]$	$\mathbf{E}\left[\delta_{\mu(2)}\right]$	$\mathbf{E}\left[S_{2\mu(2)}\right]$	$\mathbf{E}\left[\delta_{\mu(3)}\right]$	$\mathbf{E}\left[S_{3\mu(3)}\right]$	$\mathbf{E}\left[W\right]$
One school	$\frac{23}{36}$	$\frac{50}{36}$	$\frac{17}{36}$	$\frac{42}{36}$	$\frac{14}{36}$	$\frac{36}{36}$	$\frac{182}{36} - 3c$
Both schools	$\frac{23}{36}$	$\frac{50}{36}$	$\frac{18}{36}$	$\frac{43}{36}$	$\frac{13}{36}$	$\frac{36}{36}$	$\frac{183}{36} - 4c$

2.8.2 Other Priorities

Throughout the main text, I assume that schools have common priorities over students, but under idiosyncratic preferences, many of the results generalize to a model in which priorities are independent and identically distributed across students and schools. This captures settings in which schools differ in their priorities over students. For example, schools could prioritize students who live closer over those who live farther away. Formally, let x_{ij} denote student i's priority for school k_j. Let $x_{ij} \sim^{iid} \text{Uniform}[0, 1]$. That is, priorities are drawn identically and independently across students and schools. I define such priorities as *idiosyncratic priorities*.

Claim 1. *Consider idiosyncratic preferences and idiosyncratic priorities. Under each search protocol, there exists a unique UTB-equilibrium. In equilibrium, student i employs the respective Sim-Max or Seq-Max strategy using the feasible set $D_i = \{k_j : x_{ij} \geq r_j\}$, where r is the vector of admission cutoffs. Admission cutoffs are unique, independent of the cost c, and are the same across search protocols.*

Proof of Claim 1

The argument closely resembles that of Propositions 1a and 1b. Again, since there is a continuum of students, there is no aggregate uncertainty. Since $\sum_{j=1}^{J} q_j = 1$ and all students and schools are acceptable, a mass of q_j students ultimately attend school k_j. By the symmetry refinements of a UTB-equilibrium, students matriculate to schools evenly across all schools in their feasible sets. This is true for both search protocols. As a result, matriculation in this incomplete-information setting coincides with a complete-information setting in which students' preferences are drawn uniformly at random. Lemma 4 of Abdulkadiroğlu et al. (2015) establishes that such a setting admits a unique set of admission cutoffs. Therefore, in my setting, there exists a unique set of admission cutoffs, which are independent of c and the same across search protocols.

Chapter 3

Dynamic Matching with Transfers

I study a two-sided market of dynamic matching with sequential arrivals. Agents must trade off between waiting to match with a more desirable agent versus match immediately with a less desirable agent. I study a transfer environment and a transferless environment for two priority orders: first-in-first-out (FIFO) and last-in-first-out (LIFO). For both FIFO and LIFO, the equilibrium queue buffer is weakly shorter with transfers than without transfers. Transfers provide strictly shorter queue buffers only when high and low types are misaligned in willingness to wait for a congruent type absent transfers. For FIFO, welfare is higher with transfers than without while the reverse is true for LIFO. In the limit as waiting costs go to zero, the equilibrium for LIFO with transfers is asymptotically efficient while the equilibrium welfare for FIFO with transfers remains bounded away from the social optimum.

3.1 Model

3.1.1 Setup

I study an infinite-horizon dynamic matching market akin to that of Baccara et al. (2020). There are two kinds of agents: Squares and Rounds. A Square is a high type (H) with probability p and a low type (L) with probability $1-p$. Likewise, a Round is a high type (h) with probability p and a low type (ℓ) with probability $1-p$. Let $U_a(b)$ denote the surplus a type-a agent receives from matching with a type-b agent. I assume homophily, so agents receive a larger surplus from matching with a congruent type than from matching with an incongruent type:

$$U_H(h) > U_H(\ell) \qquad U_L(\ell) > U_L(h)$$
$$U_h(H) > U_h(L) \qquad U_\ell(L) > U_\ell(H).$$

Let $U_{Xy} \equiv U_X(y) + U_y(X)$ denote the joint surplus generated from an (X, y) match. Let

$$U \equiv U_{Hh} + U_{L\ell} - U_{H\ell} - U_{Lh}$$

capture the degree of preference super-modularity. That is, U captures the efficiency gains from matching congruent types versus matching incongruent types. I assume that each agent suffers a cost $c > 0$ for each period they spend in the market waiting to be matched and that agents can leave only after matching. For expositional simplicity, I assume symmetry within types: $U_H(h) - U_H(\ell) = U_h(H) - U_h(L)$ and $U_L(\ell) - U_L(h) = U_\ell(L) - U_\ell(H)$. I also assume some regularity conditions that guarantee uniqueness of the equilibrium. Specifically, I assume that $\frac{p(1-p)U}{2c}$, $\frac{p(U_H(h)-U_H(\ell))}{c}$, $\frac{(1-p)(U_L(\ell)-U_L(h))}{c}$, $\sqrt{\frac{p(1-p)U}{2c} + \frac{1}{4}} - \frac{1}{2}$, $\sqrt{\frac{p(1-p)(U_H(h)-U_H(\ell))}{2c} + \frac{1}{4}} - \frac{1}{2}$, and $\sqrt{\frac{p(1-p)(U_L(\ell)-U_L(h)))}{2c} + \frac{1}{4}} - \frac{1}{2}$ are not integers.

3.1.2 Timing

For each period t, there are three stages. First, a Square and a Round enter the market, with their types determined independently by the above probabilities. Second, agents are ordered according to a priority rule, which I describe in more detail below. Third, agents declare their demands, i.e., with which types of agents they are willing to match. For example, an L-Square could declare she is willing to match with an h-Round, an ℓ-Round, or both. A clearinghouse uses the priority rule and the agents' demands to clear the market. Matched agents exit the market. Unmatched agents remain in the market, incurring a waiting cost of c, and proceed to period $t+1$.

3.1.3 Ordering Rules

I study two priority rules that determine the order by which queued agents are allowed to match: first-in-first-out (FIFO) and last-in-first-out (LIFO). FIFO corresponds to a priority ordered by the time of arrival and LIFO is the reverse. For both protocols, there is a queue of Squares and a queue of Rounds. I focus on equilibria in which agent weakly prefer matching with their congruent type, absent waiting costs. Such equilibria exist under my set of assumptions. The clearinghouse first tries to make as many (H, h) and (L, ℓ) matches as possible. Then it resorts to making (H, ℓ) and (L, h) matches. Once the clearinghouse has made all possible matches, the game proceeds to the next period.

3.1.4 Equilibrium

I focus on stationary pure-strategy Nash equilibria. Let $\mathbf{s}^t = (s_H^t, s_L^t, s_h^t, s_\ell^t)$ be period t's state, where s_x^t is the number of queued x-type agents in period t. Let q_i^t be agent i's priority in period t. Let $\theta_i^t \equiv (\mathbf{s}^t, q_i)$ be the augmented state for player i. A strategy is a mapping from the augmented state θ_i^t to the set of types with which the agent is willing to match. I focus on stationary pure strategies, meaning that an agent's strategy does not depend on t and does not involve mixing.

I focus on equilibria characterized by queue buffers. Specifically, I focus on equilibria in which agents queue up to some threshold k to match with their congruent type and match immediately with any agent if their rank in queue is above k. Expected wait costs are increasing in the agent's rank in queue, so there is a threshold such that agents queue for congruent types up to some buffer k and are willing to match with any type otherwise. The market is balanced, so Squares and Rounds both use the same threshold k. If Squares and Rounds tried using unequal queue buffers, the "stubborn" side gets their way since a match can occur only if both sides agree to match. Hence, the market effectively operates with the larger of the two buffers.

3.1.5 Transfers

Throughout my analysis, I compare the results for settings with and without transfers. I focus on transfer strategies that are stationary pure strategies that are symmetric within type. That is, there is some transfer vector $\tau = (\tau_{Hh}, \tau_{H\ell}, \tau_{Lh}, \tau_{L\ell}) \in \mathbf{R}^4$ such that, if (X, y) match, then the y-type round gives a transfer of τ_{Xy} to the X-type square. For example, if (H, ℓ) match, then the H-square receives a utility of $V_H(\ell) = U_H(\ell) + \tau_{H\ell}$ and the ℓ-round receives a utility of $V_\ell(H) = U_\ell(H) - \tau_{H\ell}$. A *transfer equilibrium* is a stationary pure-strategy Nash equilibrium in which agents use some constant transfer vector τ. I also impose the additional refinement on the transfer equilibrium that both types wish to use the same queue buffer k absent what buffer the other side will use. This refinement is for mathematical simplicity and does not substantially affect results. Let a *transferless equilibrium* be a stationary pure-strategy Nash equilibrium that does not involve transfers.

3.2 Results

3.2.1 Equilibrium Strategies

3.2.2 FIFO

The following lemmas provide the equilibrium buffers for FIFO in the transferless and transfer environments.

Lemma 3a. *For FIFO without transfers, the unique equilibrium buffer is*

$$k^{transferless,FIFO} = \max\{\lfloor k^{H,FIFO} \rfloor, \lfloor k^{L,FIFO} \rfloor\},$$

where $k^{H,FIFO} = \frac{p(U_H(h)-U_H(\ell))}{c}$ *and* $k^{L,FIFO} = \frac{(1-p)(U_L(\ell)-U_L(h))}{c}$.

Lemma 3b. *For FIFO with transfers, the unique equilibrium buffer is*

$$k^{transfer,FIFO} = \lfloor \frac{p(1-p)U}{2c} \rfloor.$$

Proving these lemmas is relatively straightforward. In the transferless case, one simply has to find the (relaxed) buffers $k^{H,FIFO}$ and $k^{L,FIFO}$ such that the high types and low types are roughly indifferent between matching with an incongruent type immediately versus waiting for the next k congruent types to arrive. In the case with transfers, one can show that there exists a set of transfers such that agents are willing to wait in a queue with buffer k if and only if $U \in \left[\frac{2kc}{p(1-p)}, \frac{2(k+1)c}{p(1-p)}\right]$.

The conventional wisdom is that transfers increase efficiency in matching settings, as first shown by Shapley and Shubik (1971). However, there are relatively few results on how transfers affect efficiency in a dynamic setting. Proposition 6 illustrates that transfers can increase efficiency but not in all cases.

Proposition 5. *In equilibrium,*

- $k^{transfer,FIFO} \leq k^{transferless,FIFO}$.
- $k^{transfer,FIFO} < k^{transferless,FIFO} \implies \lfloor k^{H,FIFO} \rfloor \neq \lfloor k^{L,FIFO} \rfloor$.
- $\lfloor k^{H,FIFO} \rfloor = \lfloor k^{L,FIFO} \rfloor \implies k^{transfer,FIFO} = k^{transferless,FIFO}$.

Proof. Using Lemmas 3a and 3b, one can express $k^{transfer,FIFO}$ as a weighted sum of $k^{H,FIFO}$ and $k^{L,FIFO}$:

$$k^{transfer,FIFO} = \lfloor (1-p)k^{H,FIFO} + pk^{L,FIFO} \rfloor,$$

which establishes all three results. □

The proof is in the appendix. Proposition 6 shows that the transfer equilibrium has weakly shorter queues than the transferless equilibrium. It also shows that for transfers to shorten queues, high types and low types must be misaligned in which queue buffers they would pick absent transfers. The converse does not

hold because it is possible for agents to be slightly misaligned but the equilibrium transfer buffer be equal to the transferless equilibrium buffer.[1] In circumstances where both types of agents are willing to wait with the same queue buffer without transfers, allowing transfers does not provide any efficiency gains. Moreover, even when transfers allow for shorter queues, the equilibrium buffer is still not at the social optimum, which is given by expressions in the next section.

I now turn to equilibrium strategies in the case of LIFO. Qualitatively, the results are similar to those in the FIFO case, but the precise expressions for the equilibrium queue buffers differ because agents face different expected wait times depends on the priority order.

Lemma 4a. *For LIFO without transfers, the unique equilibrium buffer is*

$$k^{transferless,LIFO} = \max\{\lfloor h^{H,LIFO} \rfloor, \lfloor h^{L,LIFO} \rfloor\},$$

where $k^{H,LIFO} = \sqrt{\frac{2p(1-p)(U_H(h)-U_H(\ell))}{c} + \frac{1}{4}} - \frac{1}{2}$ *and* $k^{L,LIFO} = \sqrt{\frac{2p(1-p)(U_L(\ell)-U_L(h))}{c} + \frac{1}{4}} - \frac{1}{2}$.

Lemma 4b. *For LIFO with transfers, the unique equilibrium buffer is*

$$k^{transfers,LIFO} = \lfloor \sqrt{\frac{p(1-p)U}{2c} + \frac{1}{4}} - \frac{1}{2} \rfloor.$$

Proposition 6. *In equilibrium,*

- $k^{transfer,LIFO} \leq k^{transferless,LIFO}$.
- $k^{transfer,LIFO} < k^{transferless,LIFO} \implies \lfloor k^{H,LIFO} \rfloor \neq \lfloor k^{L,LIFO} \rfloor$.
- $\lfloor k^{H,LIFO} \rfloor = \lfloor k^{L,LIFO} \rfloor \implies k^{transfer,LIFO} = k^{transferless,LIFO}$.

Proof. Using Lemmas 4a and 4b, one can rewrite $k^{transfers,LIFO}$ containing terms similar to those of $k^{H,LIFO}$ and $k^{L,LIFO}$:

$$k^{transfers,LIFO} = \lfloor \sqrt{\frac{p(1-p)\left[U_H(h) - U_H(\ell) + U_L(\ell) - U_L(h)\right]}{c} + \frac{1}{4}} - \frac{1}{2} \rfloor.$$

Then $\min\{\lfloor k^{H,LIFO} \rfloor, \lfloor k^{L,LIFO} \rfloor\} \leq k^{transfers,LIFO} \leq \max\{\lfloor k^{H,LIFO} \rfloor, \lfloor k^{L,LIFO} \rfloor\} = k^{transferless,LIFO}$. All three statements follow. □

Proposition 6 echoes results from Proposition 5 but for the LIFO priority order, suggesting that the role transfers play is somewhat robust to the priority order. Again, the transfer equilibrium exhibits weakly shorter queues than the transferless equilibrium. For cases in which transfers yield strictly shorter queues, high and low types must be misaligned in the transferless case. If high and low types are aligned in the transferless case, transfers do not affect the equilibrium queue buffer.

[1] For example, consider $p = \frac{1}{2}$. If $k^{H,FIFO} = 2.1, k^{L,FIFO} = 1.8$, then $k^{transfer,FIFO} = 1 < 2 = k^{transferless,FIFO}$. However, if $k^{H,FIFO} = 2.9, k^{L,FIFO} = 1.8$ then $k^{transfer,FIFO} = 2 = k^{transferless,FIFO}$.

Despite transfers having the same effect on the queue buffer for FIFO and LIFO, they have opposite welfare implications, as I show in the next section.

3.2.3 Welfare

In this section, I look at welfare for three cases: FIFO with transfers, LIFO with transfers, and the social optimum (OPT). One can calculate welfare given the equilibrium queue buffer. Transfers sum to zero, hence do not affect welfare except through the equilibrium queue buffer. As given by Baccara et al. (2020), the expected per-period welfare with a buffer of k is

$$W(k) = \underbrace{pU_{Hh} + (1-p)U_{L\ell}}_{\equiv W_{LHS}} \underbrace{-\frac{p(1-p)U}{2k+1} - \frac{2k(k+1)c}{2k+1}}_{\equiv W_{RHS}(k)},$$

where W_{RHS} captures losses from incongruent matching and wait costs. Note that W_{LHS} does not depend on k, so I focus my attention on W_{RHS}. The socially optimal queue buffer is $k^{OPT} = \lfloor \sqrt{\frac{p(1-p)U}{2c}} \rfloor$.

Lemma 5. *In equilibrium,* $k^{transfers,LIFO} \leq k^{OPT} \leq k^{transfers,FIFO}$.

Proof. Throughout the proof, let $x = \frac{p(1-p)U}{2c} > 0$. First, compare $k^{transfers,LIFO}$ and k^{OPT}.

$$k^{transfers,LIFO} \leq k^{OPT} \iff \lfloor \sqrt{x + \frac{1}{4}} - \frac{1}{2} \rfloor \leq \lfloor \sqrt{x} \rfloor,$$

which is true for all $x > 0$. Hence, $k^{transfers,LIFO} \leq k^{OPT}$.

Now compare k^{OPT} and $k^{transfers,FIFO}$.

$$k^{OPT} \leq k^{transfers,FIFO} \iff \lfloor \sqrt{x} \rfloor \leq \lfloor x \rfloor,$$

which is true for all $x > 0$. Hence, $k^{OPT} \leq k^{transfers,FIFO}$. □

Lemma 5 establishes that queues for LIFO are too short relative to the social optimum while those for FIFO are too long. This has important welfare implications. While Propositions 5 and 6 establish that transfers have the same affect on queue buffers across FIFO and LIFO, it turns out that transfers have opposite effects on welfare.

Proposition 7. *In equilibrium for FIFO, welfare is weakly greater with transfers than without. For LIFO, however, welfare is weakly greater without transfers than with. That is,*

$$W(k^{transfers,FIFO}) \geq W(k^{transferless,FIFO})$$
$$W(k^{transfers,LIFO}) \leq W(k^{transferless,LIFO}).$$

Proof. If $c \geq 2p(1-p)U$, equilibrium buffer for all four cases and the socially optimal buffer are all 0. If $c < 2p(1-p)U$, then $\frac{\partial^2 W_{RHS}(k)}{\partial k^2} = \frac{4(c-2p(1-p)U)}{(2k+1)^3} < 0$. Hence, $W_{RHS}(k)$ is concave in k. The result then follows from observing that k^{OPT} maximizes $W(k)$ and that $k^{transfers,LIFO} \leq k^{transferless,LIFO} \leq k^{OPT} \leq k^{transfers,FIFO} \leq k^{transferless,FIFO}$. □

It is difficult to rank FIFO and LIFO welfares in general. However, one can get a clear comparison on their welfares as waiting costs vanish.

Proposition 8.

$$\lim_{c\to 0} W_{RHS}(k^{OPT}(c)) = \lim_{c\to 0} W_{RHS}(k^{transfers,LIFO}(c)) = 0.$$

Meanwhile,

$$\lim_{c\to 0} W_{RHS}(k^{transfers,FIFO}(c)) < 0.$$

Proof. Note that $W_{RHS}(k) \leq 0 \ \forall k \geq 0$. Throughout, I consider only the transfers cases of FIFO and LIFO. Let $\tilde{k}^{\Gamma}(c)$ be the relaxed equilibrium queue buffer in which $\Gamma \in \{OPT, LIFO, FIFO\}$. Take the limit of $W_{RHS}(k)$ using relaxed buffers:

$$\begin{aligned}
&\lim_{c\to 0} W_{RHS}(\tilde{k}^{OPT}(c)) = \lim_{c\to 0} -\sqrt{2cp(1-p)U} = 0 \\
&\lim_{c\to 0} W_{RHS}(\tilde{k}^{LIFO}(c)) = \lim_{c\to 0} -\frac{2p(1-p)U}{\sqrt{\frac{2p(1-p)U}{c}}+1} = 0 \\
&\lim_{c\to 0} W_{RHS}(\tilde{k}^{FIFO}(c)) = \lim_{c\to 0} -\frac{p(1-p)U(4c+p(1-p)U)}{2(c+p(1-p)U)} = -\frac{p(1-p)U}{2} < 0.
\end{aligned}$$

Note that $k^{\Gamma}(c) \to \infty$ as $c \to 0$ for all three cases. Hence, $\lim_{c\to 0} -\frac{p(1-p)U)}{2k^{\Gamma}(c)+1} \to 0$ for all three cases. Therefore, $\lim_{c\to 0} W_{RHS}(k^{\Gamma}(c)) = \lim_{c\to 0} -\frac{2k^{\Gamma}(c)(k^{\Gamma}(c)+1)c}{2k^{\Gamma}(c)+1}$. Because $\tilde{k}^{\Gamma} \geq \lfloor \tilde{k}^{\Gamma} \rfloor = k^{\Gamma}$ and everything is continuous, $\lim W_{RHS}(\tilde{k}^{\Gamma}(c)) \leq \lim W_{RHS}(k^{\Gamma}(c))$, implying that $\lim W_{RHS}(k^{\Gamma}(c)) = 0$ for $\Gamma \in \{OPT, LIFO\}$.

For FIFO, because the limit under the relaxed buffer is strictly negative, the limit under the equilibrium buffer is also strictly negative. Suppose not. Then $\forall \varepsilon > 0 \ \exists \delta > 0$ such that $c < \delta \implies |W(k^{FIFO}(c))| < \varepsilon$. However, $\tilde{k}^{FIFO}(c) = k^{FIFO}(c)$ whenever $\tilde{k}^{FIFO}(c)$ is an integer. Moreover, $\tilde{k}^{FIFO}(c) \to \infty$. Thus, $\exists c < \delta$ such that $\tilde{k}^{FIFO}(c) = k^{FIFO}(c)$ and $W(k^{FIFO}(c)) = W(\tilde{k}^{FIFO}(c))$. But $W(\tilde{k}^{FIFO}(c)) \to -\frac{p(1-p)U}{2} < 0$, implying that we can bound $W(k^{FIFO}(c))$ away from 0. □

Proposition 8 shows that in the limit as $c \to 0$, the social optimum and LIFO protocols have arbitrarily small losses from incongruent matches and waiting. For both cases, the equilibrium queue buffer $k \propto \sqrt{\frac{1}{c}}$. Therefore, the waiting cost term $-\frac{2k^2c}{2k+1} - \frac{2k^2c}{2k+1}$ goes to 0. FIFO, however, maintains strictly negative wait costs in the limit. This is because $k \propto \frac{1}{c}$, so the buffer grows at a rate faster than the per-period waiting costs dissipate.

3.3 Related Literature

This paper lies at the intersection of two literatures: matching with transfers and dynamic matching. Most papers studying matching with transfers focus on static settings. Shapley and Shubik (1971) show that when agents have quasilinear utility and transfers are unrestricted, stables matchings are utilitarian efficient. Kelso Jr and Crawford (1982) show the existence of a stable match when agents have preferences satisfying gross substitutes. Hatfield and Milgrom (2005) generalize the Kelso-Crawford model to allow matches with contracts. They show that for environments of many-to-one matching, when workers are substitutes, profit-maximizing firms match with weakly more agents when more contracts are made available.

Pivoting to the dynamic matching literature, much of this paper adapts models and results from Baccara et al. (2020). They characterize the set of equilibrium strategies in a discretionary setting as well as the socially optimal mechanism. I have extended their setting to allow for transfers, and I impose slightly different assumptions on agents' preferences. In their paper, they assume that all agents prefer high types whereas my setting has all agents prefer congruent types. Because they assume that the clearinghouse matches as many congruent pairs as possible, our results on cases without transfers are similar. Akbarpour et al. (2020) explore how various matching algorithms perform in an environment of dynamic matching in a networked market. They provide conditions under which greedily matching agents is close to optimal. They also provide conditions under which local algorithms that do not exploit the global network structure are close to optimal. Leshno (2022) studies the welfare properties of various waiting list allocation mechanisms. He shows that limiting agents' information on their expected wait times can lead to shorter expected wait times and improve welfare.

The paper closest in topic to this one is Chen and Tian (2021). They study a setting of dynamic matching with transfers in a discrete-time game but where the pool of agents remains constant throughout the entire game. Agents interact with one-to-one matching and can gain new matches each period. Hence, the tradeoffs in their game are distinct from those in this paper, where agents trade off match quality and waiting costs.

3.4 Conclusion

I extend a model of two-sided dynamic matching to allow for transfers. In this model, transfers can reduce equilibrium queue buffers when high types and low types are misaligned in their willingness to wait for their preferred type. The welfare implications of lower queue buffers depend on the priority ordering: transfers can only increase welfare for FIFO and only decrease welfare for LIFO. If high and low types are aligned in their willingness to wait, transfers have no effect on welfare.

While transfers (weakly) increase welfare for FIFO, transfers do not get agents to the socially optimal queue buffer. Even as waiting costs go to zero, the limiting equilibrium welfare for FIFO with transfers is bounded strictly below that of the social optimum. Meanwhile, the limiting equilibrium welfare for LIFO is

equal to that of the social optimum, highlighting the efficiency differences between FIFO and LIFO.

3.5 Appendix

Proof of Lemma 3b

Proof. Given a buffer k and a vector of transfers τ, one can find the conditions such that all agents are willing to wait in a queue for their congruent type up to buffer k but not $k+1$. For FIFO, these conditions are

$$\begin{array}{ccccc}
U_H(h)+\tau_{Hh}-\frac{kc}{p} & \geq & U_H(\ell)+\tau_{H\ell} & \geq & U_H(h)+\tau_{Hh}-\frac{(k+1)c}{p} \\
U_h(H)-\tau_{Hh}-\frac{kc}{p} & \geq & U_h(L)-\tau_{Lh} & \geq & U_h(H)-\tau_{Hh}-\frac{(k+1)c}{p} \\
U_L(\ell)+\tau_{L\ell}-\frac{kc}{1-p} & \geq & U_L(h)+\tau_{Lh} & \geq & U_L(\ell)+\tau_{L\ell}-\frac{(k+1)c}{1-p} \\
U_\ell(L)-\tau_{L\ell}-\frac{kc}{1-p} & \geq & U_\ell(H)-\tau_{H\ell} & \geq & U_\ell(L)-\tau_{L\ell}-\frac{(k+1)c}{1-p}.
\end{array}$$

There exists such a $\tau \iff U \in \left[\frac{2kc}{p(1-p)}, \frac{2(k+1)c}{p(1-p)}\right]$. Hence, $k^{transfer,FIFO} = \lfloor \frac{p(1-p)U}{2c} \rfloor$. By the regularity conditions, $\frac{p(1-p)U}{2c}$ is not an integer, so the threshold is unique. □

Proof of Lemma 4b

Proof. Given a buffer k and a vector of transfers τ, one can find the conditions such that all agents are willing to wait in a queue for their congruent type up to buffer k but not $k+1$. For LIFO, these conditions are

$$\begin{array}{ccccc}
U_H(h)+\tau_{Hh}-\frac{k(k+1)c}{2p(1-p)} & \geq & U_H(\ell)+\tau_{H\ell} & \geq & U_H(h)+\tau_{Hh}-\frac{(k+1)(k+2)c}{2p(1-p)} \\
U_h(H)-\tau_{Hh}-\frac{k(k+1)c}{2p(1-p)} & \geq & U_h(L)-\tau_{Lh} & \geq & U_h(H)-\tau_{Hh}-\frac{(k+1)(k+2)c}{2p(1-p)} \\
U_L(\ell)+\tau_{L\ell}-\frac{k(k+1)c}{2p(1-p)} & \geq & U_L(h)+\tau_{Lh} & \geq & U_L(\ell)+\tau_{L\ell}-\frac{(k+1)(k+2)c}{2p(1-p)} \\
U_\ell(L)-\tau_{L\ell}-\frac{k(k+1)c}{2p(1-p)} & \geq & U_\ell(H)-\tau_{H\ell} & \geq & U_\ell(L)-\tau_{L\ell}-\frac{(k+1)(k+2)c}{2p(1-p)}.
\end{array}$$

There exists such a $\tau \iff U \in \left[\frac{2k(k+1)c}{p(1-p)}, \frac{2(k+1)(k+2)c}{p(1-p)}\right]$. Hence, we can solve for k and find that $k^{transfer,FIFO} = \lfloor \sqrt{\frac{p(1-p)U}{2c} + \frac{1}{4}} - \frac{1}{2} \rfloor$. By the regularity conditions, this term is not an integer, so the threshold is unique. □

Bibliography

Abdulkadiroğlu, A., Che, Y.-K., and Yasuda, Y. (2015). Expanding "choice" in school choice. *American Economic Journal: Microeconomics*, 7(1):1–42.

Abdulkadiroğlu, A., Pathak, P., Roth, A. E., and Sönmez, T. (2006). Changing the boston school choice mechanism. Technical report, National Bureau of Economic Research.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *American Economic Review*, 99(5):1954–78.

Akbarpour, M., Li, S., and Gharan, S. O. (2020). Thickness and information in dynamic matching markets. *Journal of Political Economy*, 128(3):783–815.

Artemov, G. (2021). Assignment mechanisms: common preferences and information acquisition. *Journal of Economic Theory*, page 105370.

Baccara, M., Lee, S., and Yariv, L. (2020). Optimal dynamic matching. *Theoretical Economics*, 15(3):1221–1278.

Bai, J. and Waldfogel, J. (2012). Movie piracy and sales displacement in two samples of chinese consumers. *Information Economics and Policy*, 24(3-4):187–196.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Bikhchandani, S. (2017). Stability with one-sided incomplete information. *Journal of Economic Theory*, 168:372–399.

Bochet, O., Magnani, J., et al. (2021). Limited strategic thinking and the cursed match. Technical report.

Chen, Y. and He, Y. (2021a). Information acquisition and provision in school choice: a theoretical investigation. *Economic Theory*, pages 1–35.

Chen, Y. and He, Y. (2021b). Information acquisition and provision in school choice: an experimental study. *Journal of Economic Theory*, 197:105345.

Chen, Y. and Kesten, O. (2017). Chinese college admissions and school choice reforms: A theoretical analysis. *Journal of Political Economy*, 125(1):99–139.

Chen, Y. and Tian, G. (2021). Stability and efficiency in dynamic matching with transfers.

Chen, Y.-C. and Hu, G. (2018). A theory of stability in matching with incomplete information. *Available at SSRN 3384439*.

Corcoran, S. P., Jennings, J. L., Cohodes, S. R., and Sattin-Bajaj, C. (2018). Leveling the playing field for high school choice: Results from a field experiment of informational interventions. Technical report, National Bureau of Economic Research.

Danaher, B. and Waldfogel, J. (2012). Reel piracy: The effect of online film piracy on international box office sales. *Available at SSRN 1986299*.

Doval, L. (2018). Whether or not to open pandora's box. *Journal of Economic Theory*, 175:127–158.

Echenique, F., Gonzalez, R., Wilson, A., and Yariv, L. (2021). Top of the batch: Interviews and the match. *American Economic Review: Insights*.

Erdil, A. and Ergin, H. (2008). What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review*, 98(3):669–89.

Erdil, A. and Ergin, H. (2017). Two-sided matching with indifferences. *Journal of Economic Theory*, 171:268–292.

Fernandez, M., Rudov, K., and Yariv, L. (2021). Centralized matching with incomplete information.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Hakimov, R., Kübler, D., and Pan, S. (2021). Costly information acquisition in centralized matching markets.

Hastings, J. S. and Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics*, 123(4):1373–1414.

Hatfield, J. W. and Milgrom, P. R. (2005). Matching with contracts. *American Economic Review*, 95(4):913–935.

Hoxby, C. M. and Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review*, 105(5):514–17.

Immorlica, N., Leshno, J., Lo, I., and Lucier, B. (2020). Information acquisition in matching markets: The role of price discovery.

Kelso Jr, A. S. and Crawford, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society*, pages 1483–1504.

Kesten, O. and Kurino, M. (2019). Strategy-proof improvements upon deferred acceptance: A maximal domain for possibility. *Games and Economic Behavior*, 117:120–143.

Kloosterman, A. and Troyan, P. (2020). School choice with asymmetric information: Priority design and the curse of acceptance. *Theoretical Economics*, 15(3):1095–1133.

Lee, R. S. and Schwarz, M. (2017). Interviewing in two-sided matching markets. *The RAND Journal of Economics*, 48(3):835–855.

Leshno, J. D. (2022). Dynamic matching in overloaded waiting lists. *American Economic Review*, 112(12):3876–3910.

Leung, T. C. (2015). Music piracy: Bad for record sales but good for the ipod? *Information Economics and Policy*, 31:1–12.

Liu, Q. (2020). Stability and bayesian consistency in two-sided markets. *American Economic Review*, 110(8):2625–66.

Liu, Q., Mailath, G. J., Postlewaite, A., and Samuelson, L. (2014). Stable matching with incomplete information. *Econometrica*, 82(2):541–587.

Liu, Z. (2018). Quantifying the heterogeneous effects of piracy on the demand for movies. Technical report, Working Paper, Available at publish. uwo. ca/ zliu328/jobmarketpaper. pdf.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

Nolan, Z., Williams, J. W., and Zhang, H. (2022). The impact of video piracy on content producers and distributors.

Oberholzer-Gee, F. and Strumpf, K. (2007). The effect of file sharing on record sales: An empirical analysis. *Journal of political economy*, 115(1):1–42.

Oberholzer-Gee, F. and Strumpf, K. (2016). The effect of file sharing on record sales, revisited. *Information Economics and Policy*, 37:61–66.

Pathak, P. A. and Sönmez, T. (2008). Leveling the playing field: Sincere and sophisticated players in the boston mechanism. *American Economic Review*, 98(4):1636–52.

Peitz, M. and Waelbroeck, P. (2004). The effect of internet piracy on cd sales: Cross-section evidence. *Available at SSRN 511763*.

Rob, R. and Waldfogel, J. (2006). Piracy on the high c's: Music downloading, sales displacement, and social welfare in a sample of college students. *The Journal of Law and Economics*, 49(1):29–62.

Rob, R. and Waldfogel, J. (2007). Piracy on the silver screen. *The Journal of Industrial Economics*, 55(3):379–395.

Shapley, L. S. and Shubik, M. (1971). The assignment game i: The core. *International Journal of game theory*, 1(1):111–130.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186.

Waldfogel, J. (2012). Copyright protection, technological change, and the quality of new products: Evidence from recorded music since napster. *The journal of law and economics*, 55(4):715–740.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654.

Zentner, A. (2006). Measuring the effect of file sharing on music purchases. *The Journal of Law and Economics*, 49(1):63–90.