# Essays in Industrial Organization

by

Samuel Isaac Grondahl

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020



Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Department of Economics
August 15, 2020

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Nikhil Agarwal
Associate Professor of Economics
Thesis Supervisor

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Glenn Ellison
Gregory K. Palm Professor of Economics
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Amy M. Finkelstein
John & Jennie S. MacDonald Professor of Economics
Chairman, Department Committee on Graduate Theses

# Essays in Industrial Organization

by

Samuel Isaac Grondahl

Submitted to the Department of Economics
on August 15, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

## Abstract

This thesis is a collection of three chapters that investigate burgeoning empirical issues in industrial organization.

In the first chapter, I study platform fee policy with a specific focus on two-sided online marketplaces. The main contributions of the paper are threefold. First, I study a setting with coordinated price experimentation along the three different fee dimensions that are common to such marketplaces. Second, I describe the empirical impact of incomplete fee salience on equilibrium outcomes. Finally, I quantify the network externalities that must be present in order for observed fees to constitute an equilibrium. In the paper, I begin by developing a tractable model of the platform's problem that generates testable predictions and yields equilibrium conditions in terms of estimable quantities. Then, using estimates from experimental data obtained from a large online marketplace, I quantify the salience and network effects. To conclude, I consider the counterfactual level and composition of equilibrium platform fees under when these effects are muted or absent.

In the second chapter, using data from the same source as in chapter one, I study small sellers competing on the supply side of online marketplaces. As these platforms grow and markets become increasingly disintermediated, an important concern is whether small sellers, who may have limited experience or attention, can individually compete effectively with larger, often professional sellers operating on the same marketplaces. To answer this question, I develop and estimate a structural model that incorporates essential features of the empirical setting, including large and rapidly changing choice sets and buyer heterogeneity. Using the estimated model, I compute optimal pricing policies under various informational and computational restrictions. I find that small sellers adhering to a simple strategy can obtain nearly optimal expected revenue and that this strategy's information requirements are easily satisfied in the online setting. Additionally, I present suggestive evidence that sellers learn to approximate such a strategy through repeated market interactions.

In the third and final chapter, I investigate the industrial impacts of firm control rights, which confer discretion over firm policy and are usually shared between debt and equity holders. Control rights operate along a continuum and are difficult to measure. As a proxy, I consider the discontinuous shift in control from equity holders to creditors due to loan covenant violations, a common form of technical default. This paper contributes to the growing covenants literature in two ways. First, I consider the impact of and response to covenant violations at the industry level, inclusive of firms never in technical default. Second, I empirically document the effects of violations on contemporary product markets. I find that control rights transfers to creditors make firms tough in product markets, consistent with the predictions of a stylized model, and that markups decline at the industry level, though the declines are sharpest for firms directly affected.

Thesis Supervisor: Nikhil Agarwal
Title: Associate Professor of Economics

Thesis Supervisor: Glenn Ellison
Title: Gregory K. Palm Professor of Economics

## Acknowledgments

This thesis is the culmination of a long journey made possible only by the unwavering support of my advisors, family, and friends. I am blessed to count among them far too many to thank individually. Still, I would be remiss if I failed to recognize those who have had an especially profound impact on this journey.

My advisors, Nikhil Agarwal and Glenn Ellison, patiently taught me to appreciate and develop the skill for conducting and expositing original research. It's harder than it seems. Nikhil, Glenn, and Heidi Williams led me through the fog of fledgling research ideas, which can seem inescapable at times, and for this I am deeply grateful. Susan Athey, Sigridur Benediktsdottir, Ralph Bradburd, and Robert Shiller directed me forward at earlier points in my academic career. Without their encouragement I would not have inked the first word of this manuscript.

My classmates and friends helped me to maintain perspective in excitement and anxiety alike. Time spent together was a welcome respite from the rigors of the doctoral program. I will always cherish the memories made with Angelica Cristello, Or Elezra, Daniel Elkind, Elad Hirsh, Jon Pounds, Alexander Ramey, Cyprien Sarteau, Daniel Waldinger, Benjamin Winter, and Ariel Zucker during these years.

Most importantly, my family and loved ones gave me the energy and courage to succeed, just as they have so many times before. My two best friends, Shannon Conneely and our dog Zeke, inspired me to put forth my best effort every day, no matter the day or how successful the effort. For his part, Zeke was a most loyal and indefatigable companion, joining me on long walks as I collected my thoughts and at my desk as I put them to paper. Words of hope urging me to continue making progress, however incremental, regularly arrived in long form text messages that were always signed “Love, Dad.” My mom has forever been my biggest supporter and the strongest believer in my abilities and intentions. She worked without hesitation or complaint to encourage all of my endeavors, despite the outcomes, and to ensure that I had every opportunity for success. Finally, the memory of my grandfather, Melvin Plavin, continues to be the foundational spirit from which I draw guidance — academically, ethically, financially. I wish only that I could have shared this moment with him.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Fee Salience, Network Externalities, and Platform Policy

The growing landscape of online marketplaces has left few industries untouched. These marketplaces are likely to generate substantial incremental surplus by efficiently matching buyers and sellers who otherwise might not meet. These marketplaces generate revenue by charging fees on transactions that take place between end users on the platform. In this paper we provide an tractable analytical framework to study the forces that determine the overall level and composition of these fees. We then take the framework to the data, using rich experimental evidence from a large online platform. Finally, we perform a decomposition to quantify the relative impacts of platform externalities and incomplete fee salience on optimal fees. We determine that relative to counterfactual fees absent either effect, platform externalities reduce overall fees by 28–36% and incomplete salience raises overall fees by 11–25%. The latter effect is concentrated on the "buyer" side of the marketplace and consequently has a significant impact on the composition of fees: relative to a full salience baseline, we find that incomplete buy fee salience increases equilibrium buy fees by over 70% but *reduces* equilibrium sell fees by nearly 35%.

## 1.1 Introduction

Online marketplaces continue to transform the modern commercial landscape and fill a vital role in matching buyers and sellers (Einav et al. (2016)). These platforms provide alternative markets for nearly every conceivable category of goods, from niche, expensive sneakers (StockX) to smartphones (Swappa) to non-specific exchanges (eBay, Facebook Marketplace). As online marketplaces have proliferated and grown in reach in the quarter century since Craig Newmark first circulated his distribution list, so too has their sophistication. To remain viable as businesses, these platforms typically charge fees to both sides of the market, the overall level and composition of which vary widely both between platforms and over time.

The goal of this paper is to fill in a gap between theory and practice and provide insight about how platforms set fees and the factors that influence them. The general conclusion is that network effects and incomplete fee salience can have staggeringly large impacts on optimal fees, helping to explain both the high overall fee levels of many online platforms as well as the frequency and magnitude of fee updates which are familiar to users of these platforms. To develop intuition we describe a broad number of contemporary examples before focusing on one in particular, from which we obtain rich experimental data. To make sense of observed fees, we develop a model and solve for equilibrium fee levels as a function of economic primitives. To our knowledge ours is the first paper to directly estimate the quantities necessary to empirically test platform optimality conditions.

Our empirical analysis focuses on a large online marketplace for live event tickets (the "Platform") which believe to be broadly representative. Sellers may freely list individual tickets for sale on the Platform or one of its many competitors and in turn buyers browse among the available tickets and decided whether to transact. There are three fee levers across which the Platform may optimize: sell fees, which are deducted from payouts to sellers upon sale; buy fees, which are presented to buyers only once they have placed items in their carts and are prepared to checkout; and display fees, which are fees nominally charged to buyers that are fully incorporated into the prices that buyers see during browsing (before checkout). All three fees are charged as a percentage of the gross transaction price, which is the product of the per-ticket list price chosen by the seller and the quantity of tickets purchased. In principle one or several of these fees may be negative; in practice the platform often charges a "negative" display fee, marking the display price down below the seller's chosen list price (but still higher than the seller's *net* payout after adjusting for sell fees).

The Platform regularly performs fee experimentation in an effort to continually refine its fee structure. We obtain data covering three separate such experiments, one targeting each of the three fee levers, from which we estimate the economic primitives that characterize the Platform's optimal fee structure. Through a data use agreement, the Platform furnished us with rich transaction and browsing data very nearly describing the universe of interactions that both buyers and sellers have with the platform and, mediated by the platform, with one another. The data that we obtain are sampled from late 2018 and are restricted to users who were selected into the experiments; these data have been anonymized and scrubbed of any identifying information.

The remainder of the paper is structured as follows. We begin in Section 1.2 by surveying modern online marketplaces and establishing a set of stylized facts to be explained then continue in Section 1.3 to describe our empirical setting in closer detail. Classically, the economic incidence of taxes or fees within a competitive market is independent of the statutory incidence. Yet fee neutrality is rejected both by evidence from a survey of platform fees as well as experimental evidence that we shall present in this work. Two principal forces violate the neutrality result: (i) network effects, whereby equilibrium participation on each side influences demand on the other side(s), and (ii) limited salience, whereby agents may incompletely optimize with respect to certain fees. Depending upon certain economic primitives, these factors may amplify or mitigate one another, and a central goal of this paper is to compare the relative magnitudes of each.

In Section 1.4 we study a profit maximizing platform facing exogenous residual demand which incorporates the two highlighted departures from the classical model. Taking inspiration from the labor matching literature, we generalize the framework to allow for flexible matching between buyers and sellers. The model borrows from and is most similar in spirit to the general formulation of a monopoly platform introduced by Rochet and Tirole (2006). Using this model, we derive several key results that we use in later counterfactual analysis. The majority of the results are written in terms of "sufficient statistics" (Chetty (2009)) to the extent possible and are geared toward quantities that we can estimate directly. We derive equilibrium conditions completely characterizing the platforms level and mix of fees. From these, we derive comparative statics describing the general impact of changes to network effects or fee salience. We find that stronger network effects unequivocally reduce equilibrium (platform optimal) fees to *both* sides of the market while reducing fee salience to buyers (sellers) increases equilibrium fees charged to buyers (sellers) but *reduces* equilibrium fees charged to sellers (buyers). Finally, we specify preferences for representative agents on both sides of the market which in combination with the remainder of the model results establish closed form expressions for implied network externalities and salience effects.

In Section 1.5 we proceed to describe each of three separate experiments that allow us to cleanly observe demand response to variation along each fee margin. We estimate price elasticities (semielasticities) of transaction volume (in dollars) to be 2.41 (1.96), 2.67 (2.45), and 1.27 (1.46) when price changes are induced by display price, buy fees, and sell fees, respectively, with all three evaluated at their prevailing levels. Typical fees charged by this Platform are around 24% to buyers and 8% to sellers, therefore *fee* elasticities are substantially lower than price elasticities at approximately 0.479 and 0.180, respectively, for buy and sell fees. The implied platform externality elasticities — the change in membership on side A in response to an exogenous increase in side B — were 0.833 for buyers and 0.362 for sellers.

In Section 1.6 we use these estimates to quantify the contribution of each factor — platform externalities and incomplete salience — to the final platform fee policy. If initially the Platform were constrained to include all fees in the display price and there were no network externalities between the two sides, optimal fees would be (27.2%, 17.7%) to buyers and sellers, respectively, for total fees equal of 44.9%. Adding platform externalities reduces optimal fees to (18.4%,10.7%), an overall wedge of 29.1%. Allowing the firm to make use of both of its buy fee levers in order to reduce salience increases overall fees modestly to 32.4% (24.1%, 8.3%). These effects amplify one another: simultaneously eliminating network externalities and retaining incomplete salience raises optimal fees to 50.6%, higher than the sum of the two percentage changes independently. Finally, the salience impact is sensitive to what baseline price we consider consumers to be responding. In aggregate, platform externalities have a net effect of lowering overall fees by 28–36% while incomplete salience raises overall fees by 11–25%. Section 1.7 concludes.

Our work builds on several related literatures. It adds to a growing focus on platform dynamics and behavior. A number of theoretical analyses consider variants of a general platform structure, many of which are tailored to match certain settings of interest or describe certain stylized facts (Rochet and Tirole (2003, 2006); Armstrong (2006); Weyl (2010); White and Weyl (2016); Wang and Wright (2017); Peitz et al. (2017)). The empirical strand of this literature most similar to ours uses panels of end user interactions with competing platforms to evaluate the determinants of platform membership, policy, and competition (Jin and Rysman (2015); Cantillon and Yin (2011)). Similarly, Blake et al. (2018) study the reduced form impacts on transaction volume and composition (across good quality) of a platform significantly changing its fee structure. Also related is a vein focusing on markets for newspapers and magazines, though institutional differences relative to online markets limit direct comparisons with our setting (Argentesi and Filistrucchi (2007); Kaiser and Wright (2006)).

Additionally, our work on platform externalities builds on the network effects literature pioneered by Katz and Shapiro (1985, 1986) and Farrell and Saloner (1985, 1986). More recent work attempts to estimate network externalities across a variety of industries (Rysman (2004); Clements and Ohashi (2005); Ohashi (2003); Corts and Lederman (2009); Lee (2013)). Reshef (2019) uses discontinuous variation in platform membership on one side of a platform to directly identify the magnitude of network effects. We similarly contribute estimates of network externalities, identified in our case by experimental price variation.

Finally, we incorporate the widely studied impacts of incomplete salience into a platform context. Chetty et al. (2009) studies differences in demand response to taxes included in display prices relative to those charged at the register. Related work studies similar effects in other settings (Finkelstein (2009); Cabral and Hoxby (2012); Rivers and Schaufele (2015)). Ellison and Ellison (2009) contributes empirical evidence from online retailers of computer memory who obfuscate pricing to complicate consumer search and reduce effective own price elasticities. While we do not explicitly consider search, the mechanism in our setting is largely similar — the less salient fee lever is the one that is only presented after the consumer has already extensively browsed and found the tickets that she most prefers. Finkelstein (2009) considers the impact on highway tolls of reduced salience due to electronic tolling and finds that equilibrium tolls are 20-40% higher due the salience effect, consistent with our empirical findings.

## 1.2 Empirical Facts and Motivation

Platforms share the general property that surplus is derived from interactions between two distinct groups of agents. Early and well-studied examples of platforms include computer hardware architectures, operating systems, newspapers, and payment networks. More recently, advances in information technology and shifting consumer preferences have powered the ascent of large online marketplaces, platforms directly matching buyers and sellers. A summary of such platforms with their corresponding fee policies follows in Table 1.1.

In contrast with many legacy two-sided platforms, contemporary online marketplaces often charge non-trivial two part tariffs on a transaction basis but seldom charge membership fees. We make an additional distinction between online marketplaces broadly and "sharing economy" marketplaces specifically. The latter are those

Table 1.1: Fee policies (as of March 2020) of selected online marketplaces. "All-in" indicates whether the price that a prospective buyer sees while browsing matches the price seen at final checkout, excluding taxes and optional shipping charges, if applicable.

| Platform | Category | Seller Fee | Buyer Fee | All-in |
|---|---|---|---|---|
| Amazon | Retail | \$0.99 + 8-15% | None | Yes |
| eBay | Retail | \$0.35 + 10-12% | None | Yes |
| Walmart/Jet | Retail | \$0.00 + 8-15% | None | Yes |
| AbeBooks | Used Books | \$0.00 + 8% | None | Yes |
| Etsy | Crafts | \$0.20 + 5% | None | Yes |
| Swappa | Cell Phones | None | \$0.00 + 4-5% | Yes |
| Airbnb | Vacation Rentals | \$0.00 + 3% | \$0.00 + 14.2% | No |
| SeatGeek | Live Events | \$0.00 + 10-15% | \$2.50 + 20-35% | No |
| upWork | Freelance Work | \$0.00 + 5-20% | \$0.00 + 3% | No |
| TaskRabbit[1] | Freelance Work | \$0.00 + 15% | \$0.00 + 15% | No |

for which only a single unit of each platform good is available.[2] The following general observations apply broadly to online marketplaces:

(i) Fees are charged in proportion to transaction value (as in an ad valorem tax)

(ii) Membership fees are rare; when present they typically purchase additional service offerings (data feeds, accounting utilities, etc.)

(iii) When they exist, fees ostensibly charged to the buyer are typically obfuscated (presented only during the checkout process and not on "search" pages)

(iv) "Sharing economy" marketplaces often charge nonzero, differential fees to each side of the market

Based upon this evidence we make three conclusions which we investigate empirically based upon data from one such sharing marketplace. First, platforms work to "shroud" high fees and reduce fee salience, supported by the empirical fact that many

[2]As an example, Airbnb is a sharing economy marketplace because the platform good is an apartment/house-night and once transacted (booked) it cannot be booked again, but AbeBooks is not an sharing economy marketplace because when one copy of *Swann's Way* sells it does not *necessarily* preclude another identical copy from being sold after, even if in practice it is in principle possible for the book to temporarily sell out. Marketplaces for artistic goods are a middle ground. Etsy is not a sharing marketplace because most goods can be reproduced but an exchange for original oil paintings or sculptures like artprice.com would be because each work is globally unique and has no close substitute.

platforms charging fees to buyers choose to only present such fees when a transaction is initiated and not when prospective buyers are browsing platform inventory. Incomplete salience has strong implications for both the level and composition of fees. Second, network externalities, whereby additional membership on one side increases expected utility on the other side, have a similarly large impact on both platform fee policy as well as industry structure.[3] Finally, provided that platforms are optimizing, significant heterogeneity in fee structures, is indicative of significant heterogeneity both across platforms and within platform across sides. These three conclusions motivate the central focus of empirical work in Sections 1.5 and 1.6.

## 1.3 Setting and Data

We obtain data from an online marketplace for live event tickets (the "Platform") that has conducted large-scale experiments on each of its fee levers, described in more detail in Section 1.5. The data are anonymized and their contents are governed by a data use agreement. The Platform functions as a marketplace, matching individual buyers and sellers who trade in platform goods, which we refer to variously as "tickets" or "seats". Within the industry of live event ticket marketplaces, the Platform maintains a large market share but faces many close competitors. Though no precise estimates exist, most evidence suggests that the five largest marketplaces maintain an aggregate market share of over 90%.[4]

The basic unit of market organization within the Platform is the event, e.g. a

[3]Airbnb faces competition from VRBO, HomeAway, FlipKey, Wimdu, VacayHero, HomeToGo, HouseTrip, VayStays, TurnKey, ThirdHome, onefinestay, and 9flats, to name just the most popular, and has far from the cheapest fee structure, yet maintains a dominant market share, likely due in part exactly to this market share — it can offset its potentially higher fees by offering users more surplus due to more property availability and a larger pool of prospective buyers.

[4]Many exchanges are either private or part of larger companies that do not separately report financials of their individual business divisions, so precise reporting is not available. Approximate market share data are available for the several largest platforms, however. See, for instance, `https://www.ticketnews.com/2019/08/stubhub-marketplace-for-sale/` and `https://www.ftc.gov/system/files/documents/public_comments/2018/12/06720-163071.pdf`. The five largest secondary marketplaces are, in no particular order, Ticketmaster (also by far the largest primary ticket issuer), StubHub, SeatGeek, VividSeats, and GameTime. At the time of writing, all but Ticketmaster are privately held.

concert or a sports game. Within an event, with few exceptions, tickets are differentiated vertically with relatively significant dispersion. It is not uncommon to observe transactions to a single event range in price per seat by a factor of 50 or even 100: for the same event that $20 tickets transact we routinely observe $1000+ tickets transact as well; the former could be standing room only seats to Dodger Stadium while the latter might be seats in the San Miguel Club at the same stadium. We consider this vertical differentiation dimension extensively in a companion paper.

The platform charges three fees: sell fees, which are deducted from payouts to sellers upon sale; buy fees, which are presented to buyers only once they have placed items in their carts and are prepared to checkout; and display fees, which are fees nominally charged to buyers that are fully incorporated into the prices that buyers see during browsing (before checkout). All are charged as a percentage of total transaction volume. The platform also charges a small fixed fee that nominally covers logistics costs. This fee is very infrequently updated — we do not observe any changes since 2018 – and in conversations with decision makers it seems to be considered effectively immutable for various institutional reasons.

Given this fee structure, we interpret the $p$ in our model (Section 1.4) to be a percentage of transaction volume and accordingly we interpret the relation $T(\cdot)$ to map to units of dollar volume. Our question in this paper is not what non-linear or otherwise complex fee schedule would be globally optimal. Indeed, many variations would likely be an improvement over the status quo, such as multipart tariffs that differentially target per-ticket price and total transaction quantity. Rather, we assume that institutional factors constrain the structure of fees to the status quo (ad valorem based on total transaction volume) and consider the optimal fee structure subject to the constraints. Our focal estimates therefore quantify transaction volume responses to platform fees. However, for comparison we include corresponding estimates for both transaction counts and quantities.

### 1.3.1 Sellers

Sellers on the platform hold seats, most commonly obtained for personal consumption (e.g. as season ticket holders for a sports team), that they cannot or do not want to use. Speculators and larger brokers transact on the Platform as well. Users of all types may freely sell on any of the competing platforms. The selling experience and platform policies are broadly consistent across the entire industry.

An agent deciding to sell on any platform must first create an account and supply a valid payment instrument. This instrument is charged in the case that there is an issue with a sold ticket such as fraud or an issue with multiple sales (which would result in one of the buyers being refused entry). This latter feature presents a significant cost to multi-homing beyond account setup and verification. Even with a robust technological solution, which exists but is unavailable to most sellers, "double sales" are common due to a combination of technical factors. Sellers can rectify a double sale but only if they have close substitutes to offer the buyer, who may still reject them. The penalty on all platforms in such a case is the total transaction cost, which exceeds the seller's payout in all cases, plus a punitive sum. Consistent with these institutional details, qualitative evidence suggests that a vast majority of sellers single home at least at the listing level.

A potential seller must follow a sequential "sell flow" in order to create a listing. The sell flow comprises the following interactions/pages, in order: (1) event page, (2a) login page (if not already logged in), (2b) basic ticket details page (quantity, seat location), (3) pricing page, and (4) complete create listing page. It is only on stage (3) that a user observes her sell fee percentage. If she then successfully creates a listing, the sell fee presented in stage (3) is immutable in percentage terms, though it will change in absolute terms if she later updates the price of her listing.

A prospective seller on stage (3) is presented with the screen pictured in Figure 1-1. The sell fee percentage is nowhere directly displayed. Rather, the user selects a *per-ticket* list price and is presented with a message indicating her *total* net payout of all tickets in the listing sell. Hovering over a question mark indicator indicates the fee breakdown per ticket in dollar terms.

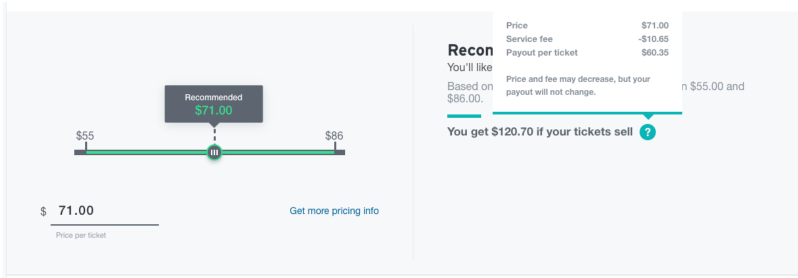



Figure 1-1: Screen presented to seller at the final stage before the listing is posted live. In order to see sell fees, the user must hover over the question mark, otherwise only total payout net of fees is displayed, though this too permits the seller to infer her fees.

After a listing is created it becomes immediately live and viewable to buyers, who may then execute transactions. Sellers are able to update listing prices and attributes or to delete listings after they are created. However, they have no control over whether or not a transaction occurs while the listing is active. At no time before or after a transaction do buyers and sellers learn each other's identities.

### 1.3.2 Buyers

Unlike sellers, buyers may freely use the Platform without registering for an account and must supply payment information only when completing a transaction. As with sellers, though, buyers must progress through a sequential "buy flow" in order to execute a transaction. In order, the stages are: (1) event page, (2) ticket view, (3) build order, (4) checkout initialize, (5) checkout final review, and (6) checkout complete.

The event page (1) allows users to browse a list of multiple listings at once, to sort and filter the listings in a variety of ways, and to see where each is spatially within the venue. Clicking on one listing in the list renders the ticket view (2), which includes additional information about the ticket as well as estimated fees. The fees presented are baseline fees and are not adjusted for experimentation, which is described in greater detail in Section 1.5.3. For instance, if an experiment were running that randomized buy fees between 23% and 27%, stage (2) would likely present estimated fees of 25%. Phase (3) technically allows users to add multiple listings to the same order, though in practice this is an experimental feature that was seldom available to users and was paired with phase (4), which begins the checkout process. In phase (4) the user must either login or supply guest credentials, including an email address, phone number, and valid payment instrument. Figure 1-2 provides a screenshot of phase (5), which is the first phase in which users presented with their true fees and total checkout price. As with sellers, fees are presented only in dollar terms and not in percentage terms. Clicking “buy now” immediately executes a transaction and proceeds to phase (6) with order confirmation information, though the transaction may later be rolled back, most commonly due to an automated fraud check.

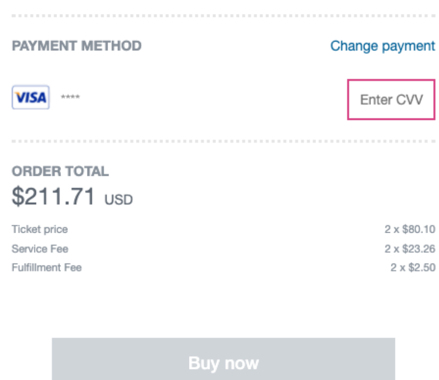



Figure 1-2: Screen presented to buyer at the final stage before the transaction is executed, her payment method is charged, and the tickets are transferred. In order to reach this screen the buyer must either be logged in or checkout as a guest but supply a valid email address and payment method. All previous screens display only estimated fees; the experimental assignment is not evaluated and final fees are not presented until she reaches this final stage.

### 1.3.3 Data

The data capture all essential user activities on the Platform, including listing behavior by sellers, browsing behavior by buyers, and all transactions between the two groups. We sample data from 176 unique days in the second half of 2018 covering a random sample of events. The latter sample is constructed to ensure that there is suitable experimental variation while still protecting the Platform's confidential information. We impose additional sample restrictions as necessary when evaluating each experiment. These are described in more detail in Section 1.5, where we also provide relevant summary statistics. The data do not include personally identifiable information, geolocation, telemetry, or any other user-level fields. However, the data do include unique identifiers that allow us to track user accounts and browsing sessions within our sample.

**Transaction Data** Each time a sale occurs on the Platform a record is written into the transaction table. The record includes the core information about the transaction — buyer and seller user identifiers, listing identifier, a unique transaction identifier, and a unix timestamp. Through appropriate joins, we also obtain the following additional fields: (i) quantity, (ii) list price, (iii) display price, (iv) sell fee charged, (v) buy fee charged, (vi) discounts applied, (vii) event identifier, and (viii) cancellation indicator and reason, if applicable. Transactions are most commonly canceled by automated fraud checks but may also be canceled if the seller is unable to deliver the tickets quickly enough. For simplicity and consistency, we eliminate all transactions that were canceled for any reason. Our data were pulled from raw sources well after the final sample event concluded, so no future cancellations are possible.

**Browsing Data** For each day in our sample period, we select all buyer browsing records for sampled events. A record is created whenever the user sends an event-specific request, for instance when scrolling through multiple listings on an event overview page, when viewing details about one specific listing, or when attempting to purchase a listing. A record observation in this dataset is uniquely identified by the 4-tuple of "session" identifier, event identifier, listing identifier, and unix timestamp. The session token (identifier) is a random string generated on each device the first time a user interacts with the Platform. Native apps (e.g. iOS or Android) generate a unique identifier when the app is installed. The token is independent of the user identifier, which we observe only if a user is logged in at the time of making a request.

As a data preprocessing step we back apply the user identifier to all requests within a session if a user is not initially logged in but later logs in.

Each record includes the following fields: (i) "session" identifier, (ii) "user" identifier (if the user is logged in at the time), (iii) event identifier, (iv) listing identifier, (v) interaction type (e.g. event, listing details, checkout), (vi) listing list price (per seat), (vii) listing quantity browsed, (viii) buy fee, (ix) and display price. If a user is browsing the event page, she typically will see multiple listings at once in a scrollable list. One record will be saved for each of these listings, but all records will indicate an event page view with the same unix time. An additional table gives a change history of all experimentation running on the Platform. Computing a hash of the session and event identifiers and jointing this latter table generates two additional fields which we append to each record: (x) display price experiment arm and (xi) buy fee experiment arm.

A user must log in in order to complete a transaction, so for all transactions we have in the browsing data a record matching that transaction on the 2-tuple of user identifier and listing identifier, both of which are globally unique across events.

**Listing Data** Using the same browsing data table as above, we observe for each event-day whether a given user began an attempt to list a set of tickets, whether or not the listing was ultimately created. A second table, separate from the experimentation table described previously, directly indicates whether each user (by unique identifier) is subject to experiment and if so, the corresponding treatment arm. Finally, a third table records all created listings. Merging these data gives, for each user-event-day, (i) an indicator for whether the user attempted to list a ticket, (ii) the user's sell fee treatment arm (if subject to experimentation), (iii) an indicator for whether the user created a listing, and (iv) unique listing identifiers (if any). Joining the transaction data in turn indicates whether or not each listing transacted at least once and the total transaction volume derived from each. Listings commonly transact incompletely in the sense that a listing with four available seats may only sell two. The remaining seats by default remain available under the same listing identifier but may or may not transact in the future.

## 1.4 Model

Section 1.2 highlights that platforms optimize their fees both in overall magnitude as well as *compositionally*. In this section, we specify a tractable model capturing the essential dimensions of the platform's problem that will prove useful in interpreting our empirical results. We approach the model exposition and statement of results with two goals in mind. First, we derive results in terms of "sufficient statistics" in the spirit of Chetty (2009) and limit our reliance on parametric assumptions where possible. Second, we turn a particular eye toward the class of "marketplace" platforms whereby agents must first choose whether and which platform to choose, then conditional on platform membership may choose whether or not to transact.

We model the platform as selecting its mix of usage fees to maximize profit. The platform is assumed to have some market power and to be a monopolist on its own residual demand, so our equilibrium concept will hold residual demand fixed. In the benchmark classical model studied in Appendix A.2.1, fees are "neutral", meaning the economic incidence of platform fees is independent of the statutory incidence. This result is inconsistent with the reality of modern online platforms. In order to better capture relevant dynamics, we amend the model in two important ways:

(i) Each side of the market influences the other beyond the effect through price such that, holding price constant, varying the mass and composition of agents on side $A$ influences preferences on side $B$.

(ii) Prices are incompletely salient, meaning certain fees have a different impact on demand or supply than equivalent changes in price but enter the platform's profit expression equally.

The remaining subsections outline the model setup and derive empirically relevant results in the following order. §1.4.1 establishes primitives and notation. §1.4.2

outlines assumptions on these primitives and introduces parametric forms where necessary. §1.4.3 defines and characterizes equilibrium. The focal results of the model — platform and incomplete salience effects — are presented in §1.4.4 and §1.4.5. We introduce preferences for representative agents in §1.4.6, which we use for later counterfactual analysis. We conclude in §1.4.7 with a discussion of our model in the context of the literature.

*Note.* We omit proofs of minor results from the main body of the text and instead include them in Appendix A.1.

### 1.4.1 Primitives

We study a two-sided profit maximizing platform, with buyers $B$ on the one side and sellers $S$ on the other. Sellers and buyers arrive stochastically and observe the platform's fee vector as well as current platform membership then decide whether or not to join. Let platform membership by side $i \in \{B, S\}$ be denoted $N^i$. In the case of a physical exchange — a flea market, for instance — the quantities $N$ indicate the masses of buyers and sellers physically present. In general, not every market participant will complete a transaction, but the total number of realized transactions is likely to be increasing in both sides' platform memberships. The same property extends to online marketplaces, though the interpretation of platform membership is more nuanced. For sellers, joining the platform means that they have registered for an account and listed items for sale. In the case of buyers, platform membership means surveying competing platforms and selecting the one yielding the highest expected utility. This activity is not costless and platforms take pains to generate lock-in. Once an agent chooses a platform, she is not required to complete a transaction.

Each consumer on each side has preferences that depend in general on platform fees charged to her side of the market $p^i$ as well as membership on the other side of the market $N^{-i}$. Platform fees are charged on the basis of ex-fee transaction volume, i.e. on the basis of the list price per item (excluding any fees) times the quantity, and are only charged when a transaction occurs. Upon a transaction, buyers pay a sum equal to the total list price plus buy fees while sellers receive a sum equal to the total list price less sell fees. We assume that fees $p^{-i}$ have no impact on side $i$ membership $N^i$ except through $N^{-i}$. Summing over all potential consumers, we therefore write aggregate membership functions $N^i(p^i, N^{-i})$. This function describes, for instance, the mass of buyers that will join the platform after observing fees charged to them $p^B$ and the mass of sellers $N^S$.[5]

Given aggregate memberships $(N^B, N^S)$, let total transaction volume be $T(N^B, N^S)$, which should be everywhere weakly increasing in both arguments and strictly increasing in at least one.[6] Concretely we interpret $T(\cdot, \cdot)$ to represent aggregate transaction volume in dollars. Fees $p^B, p^S$ represent percentage fees and $c$ represents variable per-transaction percentage costs.[7] Additionally, let $n^i(p^B, p^S)$ be the total side-$i$ membership $N^i$ subject to fees $(p^B, p^S)$. This quantity is the implicit solution to $n^i = N^i(p^i, N^{-i}(p^{-i}, n^i))$. Lemma 1.4.1 gives conditions under which this solution exists. Under the preceding definitions, platform profits are, as a function of $(p^B, p^S)$

$$\pi(p^B, p^S) = (p^B + p^S - c) \cdot T(n^B(p^B, p^S), n^S(p^B, p^S)) \tag{1.4.1}$$

In (1.4.1) we have implicitly assumed that platform fees influence total transaction volume only through platform memberships, i.e. that conditional on platform memberships, expected volume is (conditionally) independent of prices. We adopted this specification primarily for tractability and to maintain consistency with the broader platforms literature.

It shall prove convenient to define several additional objects. Let $\eta$ denote elastici-

[5]Returning to the physical market example, the decision of a potential buyer considering whether to travel to and browse a flea market is influenced by the announced fees charged to her (which should reduce her propensity to attend) as well as the number of sellers participating in the market (which should increase her propensity).

[6] The transaction function $T(N^B, N^S)$ says that, for instance, if 1000 sellers and 4000 buyers participate in a flea market on a given day, then in expectation a total transaction volume of $T(4000, 1000) = \$70,000$ will be realized.

[7]In our setting, variable costs principally account for (i) payment processing fees, (ii) costs associated with fraudulent or invalid tickets, (iii) payments to third parties for ancillary services or integrations, and (iv) additional customer support contacts. The first three costs, on a per-transaction basis, are directly proportional to total transaction value while the fourth is qualitatively increasing in total purchase price.

ties and $\sigma$ denote semielasticities signed in such a way that, subject to the assumptions that follow, all are weakly positive. This simplifies future analysis.

$$\sigma_p^i \equiv -\frac{\partial \log N^i}{\partial p^i} \tag{1.4.2}$$

$$\eta_p^i \equiv -\frac{\partial \log N^i}{\partial \log p^i} \tag{1.4.3}$$

$$\eta_N^i \equiv \frac{\partial \log N^i}{\partial \log N^{-i}} \tag{1.4.4}$$

The first two quantities are standard own price responses. The third is a participation elasticity and captures the effect of additional side $i$ participation on aggregate side $-i$ participation. As such, we commonly refer to these latter elasticities as "platform" or "network effect" elasticities.

*Remark.* The distinction between $n^i$ and $N^i$ is an important one, as is the distinction between their derivatives. $N^i$ is the primitive while $n^i$ is a derived correspondence that incorporates all feedback mechanisms. Consequently, in general $\frac{\partial N^i}{\partial p^i} \neq \frac{\partial n^i}{\partial p^i}$, as written in (1.4.8), and the quantities (1.4.2) and (1.4.3) represent partial own price effects *excluding network feedback*. These would be identified by local, random price variation but would not be identified by a permanent shift in overall platform fees, as these would induce feedback through platform effects.

### 1.4.2 Assumptions and Functional Forms

In this subsection we address assumptions under which an equilibrium exists and is unique, i.e. under which (1.4.1) has a unique optimum and each $n^i$ is well defined. We also supply a functional form for $T(\cdot,\cdot)$, the overall transaction volume. This structure is sufficient to characterize an equilibrium and to generate the results that follow.

If network effects are too large, platform size is highly unstable and may grow without bound. The feedback mechanism operates through platform membership: if $N^i$ increases, e.g. due to $p^i \downarrow$, then $N^{-i}$ increases in response, in turn inducing a further increase in $N^i$ *ad infinitum*. The following assumption ensures that not only is this feedback loop bounded but the functions $n^i(p^B, p^S)$ are well defined.

**Assumption 1.4.1.** *The demand responses $N^i$ everywhere satisfy $\frac{\partial N^i}{\partial N^{-i}} \in [0, \beta)$ for some $\beta \in (0, 1)$.*

In addition to ensuring that the $n^i$ are well defined, they must be well behaved as inputs into the platform's profit function in order to ensure that the platform's problem admits an interior maximum. The following assumption, in combination with the others, provides a sufficient condition.

**Assumption 1.4.2.** *The equilibrium demand relations $n^i : \mathbb{R}^2 \to \mathbb{R}_+$ are log-concave. That is, for any price vectors $\vec{p_1}, \vec{p_2}$ where $\vec{p_k} \equiv (p^B, p^S)'$,*

$$\log n^i(\theta \vec{p_1} + (1-\theta)\vec{p_2}) \geq \theta \log n^i(\vec{p_1}) + (1-\theta) \log n^i(\vec{p_2}) \quad \forall \vec{p_1}, \vec{p_2} \in \mathbb{R}^2, \forall \theta \in (0, 1)$$

*Remark.* Assumption 1.4.2 excludes constant elasticity demand specifications that are not log-concave though they are quasiconcave. The assumption guarantees that own price semielasticities are growing in absolute value with price. Simple of suitable specifications include, for instance, $n^B(p^B, p^S) = A - B \cdot (p^B)^k$ for $k \geq 1$, including linear demand. More generally, all suitable $n^i$ will be quasiconcave but may not necessarily be concave. Numerical simulation validates the results that follow and confirms that the assumption is tight: when log-concavity is violated, state vectors that meet the equilibrium conditions of Proposition 1.4.1 are not equilibria when log-concavity is violated, e.g. under constant elasticity specifications.

**Assumption 1.4.3.** *(i) Own price semielasticities are everywhere nonnegative*

$$\sigma_p^i \geq 0 \quad \forall p, N$$

*(ii) Network effect elasticities are constant and nonnegative*

$$\eta_N^i = \bar{\eta}_N^i \geq 0 \ \forall p, N$$

*Remark.* The first part of Assumption 1.4.3 requires that increasing the price $p^i$ charged to side $i$ weakly reduce that side's aggregate demand, beginning from any level of $p^i$, including $p^i < 0$, which describes net subsidies to side $i$. The second part of the assumption requires that a 1% increase in sellers generates a $k$% increase in buyers at all prices and participation levels. In the limiting case of no network externalities where $\eta^i_N = 0$ this is trivially satisfied. *Ex ante* it is difficult to trace out the elasticity surface if it is nonconstant. This does not imply that the membership elasticity transaction volume is constant, as this is mediated through $T(N^B, N^S)$.

Finally, we introduce the function form for total transaction value that we will use henceforth. The Cobb-Douglas specification that we introduce is commonly used in the labor matching literature and performs well empirically; for a survey refer to Petrongolo and Pissarides (2001). We consider alternative specifications in several of our later results and find that the results are qualitatively similar.

**Assumption 1.4.4.** *Total transaction volume is Cobb-Douglas in platform membership:*

$$T(N^B, N^S) = X \cdot (N^B)^{\alpha_B} (N^S)^{\alpha_S} \tag{1.4.5}$$

*for* $\alpha_B, \alpha_S \geq 0$ *and* $X > 0$.

This specification nests several interesting cases. When $\alpha_B = \alpha_S = 1$ the probability of an agent on one side of the market finding a match is proportional to the participation rate on the other side of the market. In this case the $N^i$ are interpretable as population fractions where $X$ is the total *potential* volume of transactions. In this case, complete participation on both sides of the market realizes the full potential transaction volume $X$ while if exactly one half of each side's potential members join the platform then total transaction volume is $\frac{X}{4}$. This is the same parameterization employed by Caillaud and Jullien (2003) and studied further by Rochet and Tirole (2006), though in both of those cases all agents on each side of the market interact with all agents on the other side of the market with constant probability $X$ such that each agent typically interacts multiple times. This latter interpretation is relevant for platforms such as credit card networks but must be amended in our empirical setting where unit demand prevails.

Another interesting case is captured by $\alpha_B > \alpha_S = 0$. This would correspond, for instance, to a stage game in which sellers must join the platform first based upon expectations of $N^A$ and subsequently buyers, observing $N^B$, decide whether or not to transact, given some outside option. In this case the number of sellers is assumed to be sufficient such that the number of realized transactions is independent of the number of sellers.

An alternative specification might consider $T$ to be a generalized mean:

$$T(N^A, N^B; \sigma) = \left[(N^A)^\sigma + (N^B)^\sigma\right]^{\frac{1}{\sigma}}$$

As $\sigma \downarrow$, this mean increases weight on the side with the lower participation level. This specification includes the special case in which transaction volume that is Leontief in participation levels (with $\sigma \to -\infty$). In practice the mass of available units for sale (sellers) typically outnumbers the mass of potential demand (buyers) by a significant margin — almost every event ends with significant quantities of unsold inventory still listed — so the Leontief specification would empirically coincide with the Cobb-Douglas case of $\alpha_B = 1,\ \alpha_S = 0$ considered above. Up to a scaling factor, the generalized mean with $\sigma = 0$ (the geometric mean) coincides with the pure multiplicative Cobb-Douglas case $\alpha_B = \alpha_S = 1$ also considered above. We further discuss this alternative parameterization when we derive equilibrium conditions in the following subsection.

### 1.4.3 Equilibrium

An equilibrium in the current setting is characterized by fee and participation levels such that no side deviates: no additional buyers (sellers) want to join or leave the platform after observing fees and the equilibrium seller (buyer) membership and the platform cannot increase profit by further adjusting its fees given realized platform membership. Our equilibrium concept holds fixed residual demand primitives, i.e. the functions $N^i(\cdot,\cdot)$ are independent of platform policy. This contrasts with some threads of the literature that study *general equilibria*, which would require incorporating effects of the platform's fees on other platforms' fees, which in turn influence users' outside options and impact the functions $N^i(\cdot,\cdot)$, which are primitives in our setup.

**Definition 1.4.1** (Partial Equilibrium)**.** A **partial equilibrium** of the platform model is a vector of platform fees and memberships $(p^{B\star}, p^{S\star}, N^{B\star}, N^{S\star})$ such that, subject to platform demand relationships $N^i(p^i, N^{-i}),\ i \in \{B, S\}$,

(i) Membership levels satisfy $N^{i\star} = N^i(p^{i\star}, N^{-i\star}),\ i \in \{B, S\}$

(ii) Platform fees satisfy $(p^{B\star}, p^{S\star}) \in \arg\max_{p^B, p^S} \pi(p^B, p^S)$

For fixed platform demand functions, it is reasonable to expect that each platform fee vector has associated with it unique platform membership levels. This condition is equivalent to asserting that $n^i(p^B, p^S)$ are well defined functions and is a requirement for unique equilibria. Indeed, mild regularity conditions captured in our assumptions guarantee this to be the case.

**Lemma 1.4.1.** *Let* $N([\cdot]\,;\vec{p}) \equiv \begin{bmatrix} N^A(\cdot;p^A) \\ N^B(\cdot;p^B) \end{bmatrix}$ *represent the stacked demand relationships and let Assumption 1.4.1 hold. Then* $N(\cdot)$ *has a unique fixed point for each pair* $(p^A, p^B)$.

**Proposition 1.4.1** (Partial Equilibrium: Existence and Uniqueness)**.** *Suppose Assumptions 1.4.1 – 1.4.4 are satisfied. Then an equilibrium in the sense of Definition 1.4.1 exists and is unique. Moreover, it is completely characterized by the following equations*

$$\frac{p^{B\star} - (c - p^{S\star})}{p^{B\star}} = \frac{1 - \eta_N^B \eta_N^S}{\eta_p^B(\alpha_B + \alpha_S \eta_N^S)} \tag{1.4.6}$$

$$\frac{p^{S\star} - (c - p^{B\star})}{p^{S\star}} = \frac{1 - \eta_N^B \eta_N^S}{\eta_p^S(\alpha_S + \alpha_B \eta_N^B)} \tag{1.4.7}$$

$$N^{B\star} = n^B(p^{B\star}, p^{S\star})$$

$$N^{B\star} = n^S(p^{B\star}, p^{S\star})$$

*where all quantities are evaluated at the equilibrium vector* $(p^{B\star}, p^{S\star}, N^{B\star}, N^{S\star})$.

*Proof.* When the platform adjusts $p^i$ it has a direct effect on participation $N^i$ as well as an indirect effect through $N^{-i}$, which is influenced by the change in $N^i$, holding $p^{-i}$ fixed. We can compute the price derivatives by totally differentiating each side's participating relations:

$$\frac{\partial n^i}{\partial p^i} = \frac{dN^i}{dp^i} = \frac{\partial N^i}{\partial p^i} + \frac{\partial N^i}{\partial N^{-i}} \frac{\partial N^{-i}}{\partial N^i} \frac{dN^i}{dp^i}$$

therefore[8]

$$\frac{\partial n^i}{\partial p^i} = \frac{\frac{\partial N^i}{\partial p^i}}{1 - \frac{\partial N^i}{\partial N^{-i}} \frac{\partial N^{-i}}{\partial N^i}} \tag{1.4.8}$$

and similarly

$$\frac{\partial n^i}{\partial p^{-i}} = \frac{\frac{\partial N^{-i}}{\partial p^{-i}} \frac{\partial N^i}{\partial N^{-i}}}{1 - \frac{\partial N^i}{\partial N^{-i}} \frac{\partial N^{-i}}{\partial N^i}} \tag{1.4.9}$$

To evaluate the platform's side of the equilibrium, substitute (1.4.5) into the profit equation (1.4.1) to obtain the platform's problem

$$\max_{p^B, p^S} (p^B + p^S - c) \cdot X \cdot (N^B)^{\alpha_B} (N^S)^{\alpha_S}$$

Taking logs, the first order condition in $p^B$ yields

$$\begin{aligned}
-\frac{1}{p^B + p^S - c} &= \frac{\alpha_B}{n^B} \frac{\partial n^B}{\partial p^B} + \frac{\alpha_S}{n^S} \frac{\partial n^S}{\partial p^B} \\
-\frac{1 - \frac{\partial N^B}{\partial N^S} \frac{\partial N^S}{\partial N^B}}{p^B + p^S - c} &= \frac{\alpha_B}{N^B} \frac{\partial N^B}{\partial p^B} + \frac{\alpha_S}{N^S} \frac{\partial N^B}{\partial p^B} \frac{\partial N^S}{\partial N^B} \\
-\frac{1 - \eta_N^B \eta_N^S}{p^B + p^S - c} &= \frac{\alpha_B}{p^B} \eta_p^B + \frac{\alpha_S}{p^B} \eta_p^B \eta_N^S
\end{aligned}$$

where the second line follows from (1.4.8) and the final line applies definitions (1.4.3) and (1.4.4). Rewriting gives (1.4.6). (1.4.7) follows by symmetry. Assumption 1.4.2

[8] Assumption 1.4.1 guarantees that the derivative is finite, consistent with the result of Lemma 1.4.1. To see this another way, continue to expand the total derivative to obtain

$$\frac{\partial n^i}{\partial p^i} = \frac{\partial N^i}{\partial p^i} \sum_{k=0}^{\infty} \left( \frac{\partial N^i}{\partial N^{-i}} \frac{\partial N^{-i}}{\partial N^i} \right)^k$$

which is a convergent series under Assumption 1.4.1. This formulation highlights the feedback mechanism present in the platform setting, but is not necessary to establish the result.

guarantees that $\log \pi(p^B, p^S)$ admits a unique interior optimum and because log is a monotone transformation the maximizing arguments coincide. Uniqueness of $N^{i\star}$ is guaranteed by Proposition 1.4.1, given $p^\star$. This proposition also guarantees that $n^i(p^B, p^S)$ are well defined functions. □

**Corollary 1.4.1.1.** *Under the assumptions and conditions of Proposition 1.4.1, relative fees satisfy*

$$\frac{p^{B\star}}{p^{S\star}} = \frac{\eta_p^B}{\eta_p^S} \cdot \frac{\alpha_B + \alpha_S \eta_N^S}{\alpha_S + \alpha_B \eta_N^B} \tag{1.4.10}$$

*and the overall platform fee $p \equiv p^B + p^S$ is pinned down by*

$$\frac{p^* - c}{p^*} = \frac{1 - \eta_N^B \eta_N^S}{\eta_p^B(\alpha_B + \alpha_S \eta_N^S) + \eta_p^S(\alpha_S + \alpha_B \eta_N^B)} = \frac{1 - \eta_N^B \eta_N^S}{\alpha_B(\eta_p^B + \eta_N^B \eta_p^S) + \alpha_S(\eta_p^S + \eta_N^S \eta_p^B)} \tag{1.4.11}$$

*Remark.* For tractability we retain the Cobb-Douglas assumption throughout, including in the statement of Proposition 1.4.1. However, the generalized mean parameterization is informative as a point of comparison, particularly in considering the impact of how much influence the smaller side of the market has on total transaction volume. The optimal price charged to side $B$ in under the generalized mean setup satisfies

$$\frac{p^{B\star} - (c - p^{S\star})}{p^{B\star}} = \frac{1 - \eta_N^B \eta_N^S}{\eta_p^B} \frac{(N^B)^\sigma + (N^S)^\sigma}{(N^B)^\sigma + \eta_N^S (N^S)^\sigma}$$

When $N^B \ll N^S$ and $\sigma \ll 0$ the influence of the final term in the denominator is muted and the pricing expression approaches (1.4.6) with $\alpha_S = 0$. By contrast, when $N^B \gg N^S$, the price expression approaches

$$\frac{p^B - (c - p^S)}{p^B} = \frac{1 - \eta_N^B \eta_N^S}{\eta_p^B \eta_N^S}$$

which implies a lower markup when side $S$ is participation elastic but a higher markup when it is not, which is intuitive: in this example the platform cares principally about adding side $S$ agents and will move to extract nearly all revenue from side $B$. In limiting cases, the platform provides one side with subsidies to participate.

*Remark.* In the most general case, optimal pricing can be described by an expanded set of sufficient statistics, including transaction volume elasticities $\eta_{T,i} \equiv \frac{\partial \log T}{\partial \log N^i}$, $i \in \{B, S\}$. Platform optimality conditions are

$$\frac{p^B - (c - p^S)}{p^B} = \frac{1 - \eta_N^B \eta_N^S}{\eta_p^B (\eta_{T,B} + \eta_{T,S} \eta_N^S)}$$

and relative prices satisfy

$$\frac{p^B}{p^S} = \frac{\eta_p^B (\eta_{T,B} + \eta_{T,S} \eta_N^S)}{\eta_p^S (\eta_{T,B} \eta_N^B + \eta_{T,S})}$$

At the platform's optimum, prices are proportional to elasticities, subject to a network effects adjustment, as in (1.4.10). Taken directly, the indication might be that, counterintuitively, the more price elastic side of the market should optimally be charged a higher price. Indeed, parameterizing the ratio by $\eta_p^B$ and taking the partial derivative gives

$$\frac{\partial p^{B\star}(\eta_p^B; \cdot)}{\partial \eta_p^B} = \frac{1}{\sigma_p^S} \cdot \frac{\alpha_B + \alpha_S \eta_N^S}{\alpha_S + \alpha_B \eta_N^B}$$

where $\sigma_p^S \geq 0$, (incorrectly) indicating that, *ceteris paribus*, as own price elasticity on side $B$ increases (becomes more elastic), the optimal fee charged to side $B$ increases as well! This gives an indication as to why Assumption 1.4.2 is critical. In the constant elasticity case, e.g. in the classical model studied in Appendix A.2.1, the partial and total derivatives coincide and the above derivative at least in principle gives the correct comparative static. Yet constant elasticity demand violates the log-concavity assumption and is not admissible in our framework. The proper interpretation of the pricing rule is summarized in the following corollary.

**Corollary 1.4.1.2.** *For a platform with profit function (1.4.1) facing demand satisfying Assumptions 1.4.1 – 1.4.3, and with* $\eta_N \equiv 0$, *optimal fees are always higher to the side that is pointwise less price semielastic (equiv., elastic).*

Finally, we introduce a lemma that is essential in proofs of comparative statics that follow. It describes the local effect that a change in one fee has on the optimal level of the other fee. That is, beginning from an optimum we consider exogenously perturbing one of the fees slightly. To motivate the approach, suppose some exogenous parameter $\theta$ influences side $B$'s own price elasticity but influences no other quantities directly. Then starting from some initial optimum, a shift in $\theta$ requires that, under (1.4.6), $p^B$ be adjusted to compensate and regain the equilibrium conditions. However, adjusting $p^B$ violates (1.4.7), which is otherwise uninfluenced by the change in $\theta$. Bounding the degree to which "exogenous" variation in each price influences the another is essential in signing comparative statics.

**Lemma 1.4.2** (Seesaw Principle)**.** *Beginning from and maintaining the same assumptions as the equilibrium described by Proposition 1.4.1, consider exogenously varying fee $p^{-i}$. Let $p^{i\star}(p^{-i}) \equiv \arg\max_{p^i} \pi(p^i; p^{-i})$ be the platform's constrained (conditional) optimum. Then the optimal fees charged to side $i$ are decreasing in (exogenously imposed) fees charged to the other side $-i$. In particular,*

$$\frac{dp^{i\star}}{dp^{-i}} \in [-1, 0)$$

### 1.4.4 Platform Effects

A privately optimizing platform adjusts fees taking into consideration impacts via platform effects. Relative the a monopolist's problem absent platform effects, this feature tends to increase effective price elasticities, as is evident in (1.4.11), which should on balance reduce the optimal overall fee level. Less intuitive is how this overall decrease applies to each side of the market individually, especially in light of Lemma 1.4.2. If, for instance, buyers suddenly have much stronger preferences for variety, one possibility might be that in response the platform would lower fees to sellers in order to increase their participation levels, thereby increasing buyers' valuations and levels of platform membership, and in turn raise prices on buyers to capture some of this additional surplus. The following proposition establishes that this is not the case. On balance, the price effects will be shared by both sides of the market, though potentially unequally depending on model primitives. We confirm the result numerically.

**Proposition 1.4.2** (Platform Effect). *Beginning with an equilibrium as described by Proposition 1.4.1, consider perturbing a platform elasticity $\eta_N^i$ for some $i \in \{B, S\}$. Then equilibrium quantities respond as follows:*

*(i) Overall platform fees $p^* \equiv p^{B\star} + p^{S\star}$ are decreasing in the side $i$ network elasticity*

$$\frac{dp^*}{d\eta_N^i} \leq 0$$

*(ii) Optimal fees charged to each side of the market individually are both decreasing in the side $i$ network elasticity*

$$\frac{dp^{i\star}}{d\eta_N^i} \leq 0 \quad \text{and} \quad \frac{dp^{-i\star}}{d\eta_N^i} \leq 0$$

*Proof.* Without loss of generality focus on side $B$ and let

$$\hat{\sigma}_p^B(p^B; \eta_N^B, \eta_N^S) \equiv \sigma_p^B(p^B)\kappa(\eta_N^B, \eta_N^S)$$

such that

$$p^B + p^S - c = \frac{1}{\hat{\sigma}_p^B(p^B; \eta_N^B, \eta_N^S)}$$

where

$$\kappa(\eta_N^B, \eta_N^S) = \frac{\alpha_B + \alpha_S \eta_N^S}{1 - \eta_N^B \eta_N^S}$$

The multiplicative separability of $\sigma$ and $\kappa$ is guaranteed by Assumption 1.4.3. Totally differentiate

$$\frac{d\hat{\sigma}_p^B}{d\eta_N^B} = \kappa \frac{\partial \sigma_p^B}{\partial p^B} \frac{dp^B}{d\eta_N^B} + \sigma_p^B \frac{\partial \kappa}{\partial \eta_N^B}$$

where the partial and total derivatives of $\kappa$ coincide due to Assumption 1.4.3. Moreover, $\kappa$ is increasing in both of its arguments. Therefore

$$\frac{dp^B}{d\eta_N^B}\left[1 + \frac{dp^S}{dp^B}\right] + \frac{\partial p^S}{\partial \eta_N^B} = -\frac{1}{(\hat{\sigma}_p^B)^2}\left[\kappa \frac{\partial \sigma_p^B}{\partial p^B} \frac{dp^B}{d\eta_N^B} + \sigma_p^B \frac{\partial \kappa}{\partial \eta_N^B}\right] \tag{1.4.12}$$

We can compute the partial directly from (1.4.6):

$$
\begin{aligned}
\frac{\partial p^S}{\partial \eta_N^B} &= -\frac{\eta_N^S \sigma_p^S(\alpha_S + \alpha_B \eta_N^B) - \alpha_B(1-\eta_N^B\eta_N^S)\sigma_p^S}{[\sigma_p^S(\alpha_S + \alpha_B\eta_N^B)]^2} \\
&= -\eta_N^S \cdot \Lambda^S + \frac{a(1-\eta_N^B\eta_N^S)}{\alpha_S + \alpha_B\eta_N^B}
\end{aligned}
$$

while

$$
\frac{\sigma_p^B}{(\hat{\sigma}_p^B)^2}\frac{\partial \kappa}{\partial \eta_N^B} = \eta_N^S \cdot \Lambda^B
$$

where

$$
\begin{aligned}
\Lambda^B &= \frac{1}{\sigma_p^B(\alpha_B + \alpha_S\eta_N^S)} \\
\Lambda^S &= \frac{1}{\sigma_p^S(\alpha_S + \alpha_B\eta_N^B)}
\end{aligned}
$$

and $\Lambda^B = \Lambda^S$ by the initial optimality condition (1.4.10). Therefore

$$
-\frac{\partial p^S}{\partial \eta_N^B} - \frac{\sigma_p^B}{(\hat{\sigma}_p^B)^2}\frac{\partial \kappa}{\partial \eta_N^B} = -\frac{a(1-\eta_N^B\eta_N^S)}{\alpha_S + \alpha_B\eta_N^B} \leq 0
$$

By rearranging (1.4.12), noting from Lemma 1.4.2 that $1 + \frac{dp^S}{dp^B} > 0$, and substituting this result we can conclude that $\frac{dp^B}{d\eta_N^B} \leq 0$.

Proceed similarly for an impact on the other side of the market

$$
\frac{dp^B}{d\eta_N^S}\left[1 + \frac{dp^S}{dp^B}\right] + \frac{\partial p^S}{\partial \eta_N^S} = -\frac{1}{(\hat{\sigma}_p^B)^2}\left[\kappa\frac{\partial \sigma_p^B}{\partial p^B}\frac{dp^B}{d\eta_N^S} + \sigma_p^B\frac{\partial \kappa}{\partial \eta_N^S}\right] \tag{1.4.13}
$$

where all remarks still stand but we need to recompute

$$
\begin{aligned}
\frac{\partial p^S}{\partial \eta_N^S} &= -\eta_N^B \cdot \Lambda^S \\
\frac{\sigma_p^B}{(\hat{\sigma}_p^B)^2}\frac{\partial \kappa}{\partial \eta_N^S} &= \frac{b(1-\eta_N^B\eta_N^S)}{\sigma_p^B(\alpha_B + \alpha_S\eta_N^S)^2} + \eta_N^B \cdot \Lambda_B \\
\implies -\frac{\partial p^S}{\partial \eta_N^S} - \frac{\sigma_p^B}{(\hat{\sigma}_p^B)^2}\frac{\partial \kappa}{\partial \eta_N^S} &= -\frac{b(1-\eta_N^B\eta_N^S)}{\sigma_p^B(\alpha_B + \alpha_S\eta_N^S)^2} \leq 0
\end{aligned}
$$

As before, rearranging (1.4.13) and substituting this result gives $\frac{dp^B}{d\eta_N^S} \leq 0$. The final result follows immediately.

$\square$

**Proposition 1.4.3** (Implied Platform Externalities)**.** *Consider the equilibrium described by Proposition 1.4.1 and assume that the following quantities are known: (i) equilibrium platform fees $p^{B\star}, p^{S\star}$, (ii) platform marginal costs $c$, and (iii) price semielasticities of participation among sellers and buyers, $\sigma_p^B$ and $\sigma_p^S$. Then the implied platform elasticities $\eta_N^i,\ \ i \in \{B, S\}$ that rationalize the fee structure and satisfy the conditions for equilibrium are each independent of the own-side Cobb-Douglas parameters $\alpha_i$ but are a function of other side parameters $\alpha_{-i}$. In closed form, the implied elasticities are given by*

$$\eta_N^B = \frac{\sigma_p^B}{\sigma_p^S} - \alpha_S \sigma_p^B (p^* - c) \tag{1.4.14}$$

$$\eta_N^S = \frac{\sigma_p^S}{\sigma_p^B} - \alpha_B \sigma_p^S (p^* - c) \tag{1.4.15}$$

*where $p^* \equiv p^{B\star} + p^{S\star}$.*

*Proof.* For $\alpha_S, \alpha_B > 0$, (1.4.10) gives

$$\eta_N^B = \frac{\sigma_p^B}{\sigma_p^B} - \frac{\alpha_S}{\alpha_B} + \frac{\sigma_p^B}{\sigma_p^B} \frac{\alpha_S}{\alpha_B} \eta_N^S$$

Substituting into (1.4.11) we obtain

$$\frac{p - c}{p} \left( \eta_p^B - \frac{\sigma_p^B}{\sigma_p^B} \eta_p^S \right) (\alpha_B + \alpha_S \eta_N^S) = 1 - \left[ \frac{\sigma_p^B}{\sigma_p^B} - \frac{\alpha_S}{\alpha_B} \right] \eta_N^S - \frac{\sigma_p^B}{\sigma_p^B} \frac{\alpha_S}{\alpha_B} (\eta_N^S)^2$$

which has two real roots

$$\eta_N^S \in \left\{ \frac{\sigma_p^S}{\sigma_p^B} - \alpha_B L \sigma_p^S p \ , \quad - \frac{\alpha_B}{\alpha_S} \right\}$$

Only the first root can be positive and is as long as

$$\alpha_B \leq \left[\sigma_p^B \cdot (p - c)\right]^{-1}$$

Substituting this result into the first expression, or alternatively observing the symmetry in the model, generates the second participation elasticity.

To complete the proof, suppose without loss of generality that $\alpha_S = 0$ and $\alpha_B > 0$. From the first expression we obtain

$$\eta_N^B = \frac{\sigma_p^B}{\sigma_p^S}$$

so (1.4.11) is no longer quadratic but linear in $\eta_N^S$

$$\alpha_B L \sigma_p^B p = 1 - \frac{\sigma_p^B}{\sigma_p^S} \eta_N^S$$

yielding the same result. □

### 1.4.5 Incomplete Salience

Until this point we have assumed that the platform only has two fee levers, one applied to each side of the market. In practice, the platform has some flexibility in how it implements these fees. Consider a platform charging buyers a 10% fee in equilibrium and consider some item posted by a seller at a list price of \$100. One way that platform might implement this fee would be to incorporate the fee into the display price, so when buyers browse available items they see a price of \$110 and, conditional on checkout out, pay a total price of \$110 and receive the item. Another way the platform might extract the fee would be to leave the item displayed at \$100 but to charge a \$10 fee at checkout for a total checkout price of \$110. A third way that the platform might extract the same effective fee would be to display the item at \$95 but to charge a \$15 fee at checkout, yielding the same total price to the buyer and the same net fee to the platform.

If all buyers were fully rational and perfectly anticipated all fees, the breakdown would have no influence on equilibrium outcomes. In practice the demand responses to fees applied during browsing versus at checkout differ significantly in much the same spirit as Ellison and Ellison (2009) and Chetty et al. (2009). A platform that may freely choose between several fee levers will choose the combination that makes demand least elastic. Therefore, expanding the platform's policy space will in general *increase* overall platform fees.

To evaluate this impact in the context of the model, suppose that, *ceteris paribus*, the platform reduces salience of fees charged to side $B$ so $\eta_p^B \downarrow$ over the whole domain, for instance by allowing variation in display fee $p^{disp}$ where previously only checkout fees $p^{chk}$ were available or vice versa. The following proposition summarizes the effect of this change on all fee levels.

**Proposition 1.4.4** (Salience Effect). *Let aggregate platform demand on side $i \in \{B, S\}$ be parameterized by $\theta^i$ such that*

$$\frac{\partial \eta_p^i(p^B, p^S; \theta^i)}{\partial \theta^i} \geq 0 \quad \forall p^i, N^{-i}$$

*that is, increasing the index $\theta^i$ pointwise increases side $i$'s own price elasticity everywhere on its domain.*[9] *Beginning with an equilibrium as described by Proposition 1.4.1, consider perturbing $\theta$ (therefore perturbing own price elasticity on side $i$). Then equilibrium quantities respond as follows:*

*(i) The equilibrium fee charged to side $i$ is decreasing in $\theta^i$ while the fee charged to side $-i$ is increasing in $\theta^i$:*

$$\frac{dp^{i\star}}{d\theta^i} \leq 0 \quad \textit{and} \quad \frac{dp^{-i\star}}{d\theta^i} \geq 0$$

*(ii) The equilibrium overall platform fee $p^* \equiv p^{B\star} + p^{S\star}$ is declining in $\theta^i$*

$$\frac{dp^*}{d\theta^i} \leq 0$$

[9] The odd parameterization is required because constant price elasticities are prohibited in the model.

*Proof.* Begin as in Lemma 1.4.2, this time parameterized by $\theta$

$$p^i(\theta) + p^j(p^i(\theta)) - c = (\sigma^i_p(p^i(\theta);\theta))^{-1}\kappa$$

Totally differentiating gives

$$\frac{dp^i}{d\theta} + \frac{dp^j}{dp^i}\frac{dp^i}{d\theta} = -\frac{1}{(\sigma^i_p)^2}\left(\frac{\partial\sigma^i_p}{\partial p^i}\frac{dp^i}{d\theta} + \frac{\partial\sigma^i_p}{\partial\theta}\right)$$

or, rearranging,

$$\frac{dp^i}{d\theta}\left[\underbrace{1 + \frac{dp^j}{dp^i}}_{(1)} + \frac{1}{(\sigma^i_p)^2}\underbrace{\frac{\partial\sigma^i_p}{\partial p^i}}_{(2)}\right] = -\frac{1}{(\sigma^i_p)^2}\underbrace{\frac{\partial\sigma^i_p}{\partial\theta}}_{(3)}$$

Term (1) is bounded $\in [0,1)$ by Lemma 1.4.2, term (2) is positive due to Assumption 1.4.2, and term (3) is positive due to the parameterization over $\theta$. Therefore $dp^i/d\theta < 0$. Since $\theta^i$ only influences $p^j$ through $p^i$, applying Lemma 1.4.2 gives the second result. □

The proposition confirms the intuition that reducing fee salience to buyers results in higher fees charged to buyers and higher overall platform fees. However, fees to sellers are reduced. Intuitively, when buyers are made less elastic the platform can reduce fees charged to sellers, increasing their participation and increasing surplus on the buyer side, which the platform can then more easily extract.

Some care is required when evaluating whether the composition of fees charged to one side of the market is profit maximizing for the platform. Again considering the buyer side of the market and the two fee levers of display and checkout fees, $p^{B,disp}$ and $p^{B,chk}$, respectively, the additional equilibrium condition requires that semielasticities are equated when evaluated at equilibrium levels $\sigma^B_{p^{B,disp}} = \sigma^B_{p^{B,chk}}$. This is the proper test, as in general $p^{B,disp\star} \neq p^{B,chk\star}$, so the elasticities will not be equal. Conjecturing that the final checkout price is less salient than the display price, then if the platform is constrained to charge $p^{B,disp} \geq \underline{p}^{disp}$, optimality requires that $\sigma^B_{p^{B,disp}} \geq \sigma^B_{p^{B,chk}}$, which follows from combining the appropriate stationarity and dual feasibility conditions.

In our empirical setting, the platform self-imposes the constraint that $p^{B,disp} \geq -p^S$, that is, it cannot set fees such that the display an item is less than the seller's net payout.

### 1.4.6 Representative Agents

For the counterfactual analysis that follows we assume that both the buyer and seller sides of the market admit representative agents with utility representations that are linearly separable in wealth/expenditure $y$ and the platform good $x$, $U(x, y) = u(x) + v(y)$.

**Buyers** Assume that buyer utility is quasilinear in wealth (the numeraire) while $x$ units of the platform good generate utility $u(x)$ in units of the numeraire. For tractability, suppose that $u(\cdot)$ exhibits constant absolute risk aversion: $-u''(x)/u'(x) = \gamma_B$.[10] Then given platform prices $p$, utility is equivalently represented by

$$U_B(x; p) = M - e^{-\gamma_B x} - (1 + p^B)x$$

for some constant $M$. The price semielasticity is

$$\sigma_p^B(p) = \left((1 + p^B) \log \frac{\gamma_B}{1 + p^B}\right)^{-1} \tag{1.4.16}$$

which is positive and increasing for $\gamma_B/(1 + p^B) \in (1, e)$.

**Sellers** Let sellers have the following preferences over net payout $1 - p^S$ and the platform good $x$, where the platform good is the numeraire

$$U_S(x; p) = M - e^{-\gamma_S(1-p^S)x} - x$$

[10]If we instead let utility be quasilinear in the platform good and have constant absolute or relative risk aversion in residual wealth, semielasticities are generally declining in price, violating platform optimality conditions. Additionally, calibrating the wealth parameter is difficult and not innocuous. Maintaining monetary wealth/expenditure as the numeraire simplifies the analysis and exposition.

These sellers exhibit constant absolute risk aversion $\gamma_S$ in net payout. The corresponding price semielasticity is

$$\sigma_p^S(p) = \frac{1}{1-p^S}\left(\frac{1}{\log(\gamma_S(1-p^S))} - 1\right) \tag{1.4.17}$$

which is increasing in $p^S$.

### 1.4.7 Discussion

The model that we have developed in this section provides tractable equilibrium conditions as functions of "sufficient statistics" that may be directly estimated from the data. It also provides comparative statics relating to the two focal points of this work: incomplete fee salience and platform effects, and their influence on equilibrium fees. External to the core model, we develop preferences for representative agents that allow us to compute counterfactuals and quantify the magnitudes of the two highlighted effects.

The framework is broadly consistent with the related literature. It principally diverges in two ways. First, we are focused on partial equilibria and the platform's problem while much of the literature focuses on general equilibrium effects, competition between platforms, and total welfare. Second, we focus on a platform that matches buyers and sellers with unit demand and supply rather than one in which the two sides meet and transact repeatedly (e.g. ride sharing, payment platforms) or in which utility is primarily derived from platform membership (e.g. communication networks).

**Membership Fees** Rochet and Tirole (2006) contemplate this setting when they allow "payments between end users" and derive optimality conditions which amount to (i) choosing per-transaction costs in order to maximize expected surplus from end-user interactions and then (ii) choosing membership fees that optimally extract this (plus pure membership) surplus, balancing that the membership decision is fundamentally two-sided. Weyl (2010) relies on a similar approach in studying "insulating tariffs” in which, again, the two-sided dimension is restricted to membership fees.

However, the sharing marketplaces discussed in Section 1.2 and which are the focus of our empirical work are subject to institutional factors which often preclude them from charging membership fees. For instance, Airbnb has no idea how much a potential traveler would value use of its platform before that traveler joins, browses, and books trips. Even then, it would face significant difficulty in selecting the magnitude and frequency of membership fees, and such fees are not feasible. Subject to this constraint, rewriting the platform’s objective in the Rochet and Tirole (2006) formulation (merging eqs. (18) – (20)) gives

$$\pi = (a^S + a^B - c)Xn^S n^B$$

where $S, B$ index the two sides of the market, $a^i$ are pre-transaction charges, $X$ is the fraction of potential transactions that take place, and $n^i$ are platform memberships on each side. The latter two quantities are functions of $a^i$. When platforms lose this dimension of freedom (membership fees) in their policy space they can no longer separately optimize with respect to membership and (conditional) transaction margins. Therefore, even if transaction prices were *conditionally* neutral, they are typically non-neutral in the restricted setting, as we demonstrate. Yet beyond the institutional restrictions, incomplete fee salience generates non-neutrality. If platform fees are partially shrouded on one or both sides of the market, shifting the composition of fees without changing the level affects equilibrium outcomes — transfers, allocations, and platform profit. Schmalensee (2002) and Wright (2004) study interchange fees with a specific focus on institutional features, among them that fees are predominantly levied based upon usage. In Wright’s model, non-neutrality is driven by incomplete passthrough of the interchange fees by issuer and acquirer, which also generates many of the comparative statics in the paper, in much the same way that salience operates in the models that follow.

**Competitive Bottlenecks** Using the language of our model we can take a more general view of the "competitive bottlenecks" principle introduced by Armstrong (2006), who considers a pure membership model that despite not matching our empirical context is informative to consider. Demand relationships $N^i$ remain the same but platform profits are now

$$\pi(p^B, p^S) = n^B p^B + n^S p^S - c(n^B, n^S)$$

Subject to suitable regularity conditions, optimal fees charged to side $i$ satisfy an adjusted markup equation

$$\frac{p^i - \left(c_i - \overbrace{\frac{N^{-i}}{N^i}\eta_N^{-i}(p^{-i} - c_{-i})}^{(1)}\right)}{p^i} = \frac{1 - \overbrace{\eta_N^i \eta_N^{-i}}^{(2)}}{\eta_p^i}$$

where $c_i$ represents the partial derivative of the total cost function with respect to argument $i$. Expression (1) above adjusts the marginal cost by the incremental profit realized on side $-i$ due to platform effects, scaled appropriately by the "participation elasticity" $\eta_N^{-i}$ and the relative participation levels no each side of the market. The own price elasticity $\eta_p^i$ is inflated by the factor (2) to account for indirect price feedback through platform effects. In contrast with equilibrium under our model, in which prices were proportional to elasticities (i.e. semielasticities were equated), equilibrium in Armstrong's model is characterized by a condition closer to inverse elasticity pricing, where equilibrium fees are inversely proportional to own price elasticities. In spirit the result closely follows the intuition provided by Hermalin and Katz (2006): "the routing choice makes B's network membership decision more price sensitive *ceteris paribus* hence competition for B is fiercer, which leads B to be charged less in equilibrium."

In light of the above discussion, "competitive bottlenecks" are most accurately interpreted as factors which differentially affect price elasticities of participation on one side of the market. This contrasts with a theme in the literature that one side of the market is "ignored" because it multihomes. Reparameterizing the platform's problem as choosing participation rates rather than prices we can conditionally optimize $\pi(n^i; \bar{n}^{-i})$ which gives rise to Proposition 4 of Armstrong (2006). Yet this functions only as "coordinate descent" along one dimension and requires optimization along the second dimension as well (in this case $\bar{n}^{-i}$), hence it is inaccurate to conclude that either side of the market is "ignored". Nevertheless, the monopoly outcome will not in general coincide with the social optimum; for a thorough discussion of distortions, see Weyl (2010).

**Subsidies** Negative fees, or participation subsidies, occur commonly enough that they have received regulatory attention for at least the past several decades (see, e.g. Evans (2003), for a summary of the legal precedent). In the absence of network effects, negative fees (or fees below marginal cost) to either side of the market would correctly be interpreted as predatory. Yet in the platform context, negative fees to one side are natural, though overall fees will always exceed marginal cost. Rewriting (1.4.6) gives

$$p^\star - c = \frac{1}{\sigma_p^B} \cdot \frac{1 - \eta_N^B \eta_N^S}{\alpha_B + \alpha_S \eta_N^S}$$

where by assumption the numerator and denominator are both weakly positive, confirming that overall fees $p^*$ always exceed marginal cost. Yet taking (1.4.6) directly admits $p^{B\star} < 0$ in which case the own price elasticity $\eta_p^B < 0$ under the sign convention (1.4.3). The interpretation is curious — in the negative domain the interpretation of the elasticity is the percent change in demand due to a percent move in price away from zero (to more negative values) — but the equilibrium conditions apply over the full price domain $p^i \in \mathbb{R}$.

Consider a simple example with $\eta_N^A = \eta_N^B = 0$, $a = b = 1$, $c = 0$, and linear

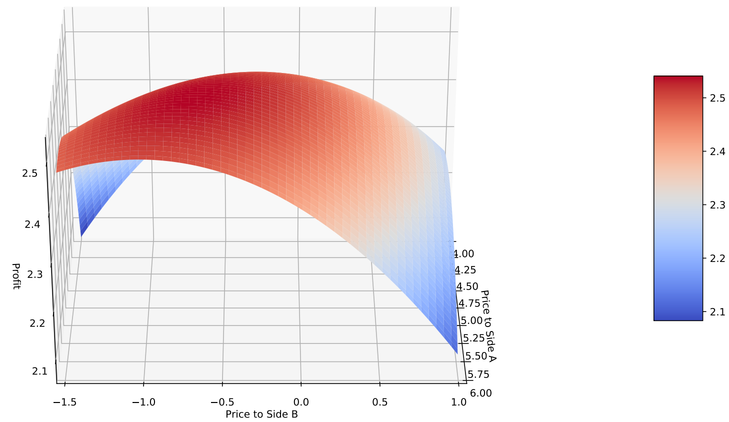



Figure 1-3: Simulated linear demand example: platform profit surface in the vicinity of the optimum.

demands

$$D^A(p^A) = 10 - p^A$$
$$D^A(p^A) = 10 - 2.5 \cdot p^B$$

Solving the system gives $p^{A\star} = \frac{16}{3}$ and $p^{B\star} = -\frac{2}{3}$. Therefore, at the optimum

$$\sigma_p^A = \frac{3}{14}; \quad \eta_p^A = \frac{16}{14}$$
$$\sigma_p^B = \frac{3}{14}; \quad \eta_p^B = -\frac{2}{14}$$

As expected, with zero marginal cost and not platform effects the sum of the price elasticities equals one at the optimum and the semielasticities are equated. We confirm that this is indeed a unique optimum by performing a broad grid search. Figure 1-3 shows the profit surface in the vicinity of the optimum.

## 1.5 Empirical Strategy and Experimental Results

The Platform has long had an interest in using empirical evidence to inform its fee policy. As a result, it frequently conducts large-scale pricing experiments along all fee dimensions. In this section we describe three independent experiments, on each per fee lever: seller fees, buyer checkout fees, and buyer display fees. We describe the operation of each of these fees in Section 1.3.

Experimentation with buyer fees (both display and checkout) is run on an ongoing basis using a purpose-built experimentation system. Seller fee experimentation is not incorporated into this system and is conducted less frequently. The largest and most recent sell fee experiment was run over approximately 100 days at the end of 2018. In an effort to capture experimental elasticities as they would have been measured at approximately the same point in time, we evaluate the buyer fee experimentation on samples covering roughly the same time period.

The section concludes with an aggregated readout and discussion of estimates from preferred specifications in §1.5.1. Each of the following three subsections describes, in order, seller, checkout, and display fee experiments and results. These may be skipped on a first reading. Within each subsection we first provide details of the experimental design, which is unique to each fee lever, as well as descriptive statistics of the relevant samples. We next describe the empirical specification, which is influenced by both the experimental design and institutional features, then provide estimates and implied price elasticities. Each subsection concludes by discussing robustness both with regard to the estimator as well as broader empirical challenges.

### 1.5.1 Summary of Results

In each of the three experiments we estimate regressions of the form

$$Y_{ijt} = f(\phi_{ijt}; X_{ijt}) + \epsilon_{ijt}$$

where $\phi_{ijt}$ is the fee charged to consumer $i$ on product $j$ at time $t$ and is the covariate of interest. Controls $X_{ijt}$ variously include event-day fixed effects or flexible time controls depending on the specification. The link function $f$ as well as the structure of the error term $\epsilon_{ijt}$ vary by regression. We discuss these variations in the following subsections. In general we prefer estimates from Poisson or logistic regression; we discuss the theoretical basis for this preference in Appendix A.3.

In all cases, outcomes of interest $Y_{ijt}$ are (i) total transaction volumes in dollars and (ii) total transaction counts. The former quantity (i) is the one that is directly interpretable within the context of the model but the latter quantity (ii) helps to understand whether the effect principally operates along the extensive margin (consumers stop buying tickets altogether due to higher fees) or whether intensive margin substitution plays a significant role (consumers partially substitute to fewer and/or less expensive tickets per transaction).

The regression point estimates give rise to multiple elasticities depending upon what we consider to be the baseline. Three will be informative to consider: price/fee semielasticity (the two are equivalent), fee elasticity, and overall price elasticity. These quantities are defined as follows (in sequence):

$$\begin{aligned}
\sigma_{\text{fee},y} &\equiv \frac{\partial \log y}{\partial \text{ (fee percentage)}} \\
\eta_{\text{fee},y} &\equiv -\frac{\partial \log y}{\partial \log(\text{fee percentage})} \\
\eta_{\text{price},y} &\equiv \left| \frac{\partial \log y}{\partial \log(\text{total amount paid in/out})} \right|
\end{aligned}$$

where $y$ is one of the two outcomes of interest (transaction volume or transaction count). The quantities $\sigma_{\text{fee},y}$ and $\eta_{\text{fee},y}$ match $\sigma_p^S$ and $\eta_p^S$, respectively, from the model in Section 1.4. For instance, $\eta_{\text{fee, trans.vol.}}$ is the percentage change in transaction volume due to a 1% change in fees, e.g. an increase from 20% to 20.2% ($1.01 \times 0.2$). The elasticity $\eta_{\text{price},y}$ by contrast represents the percent change in $y$ due to a 1% change in the total price — for a buyer facing a baseline 20% buy fee, this change would represent an increase to a 21.2% buy fee ($1.01 \times 1.2$). It does not apply to the model directly but is a useful point of comparison. The absolute value is notational convenience to ensure the signs match on both sides the market.

Table 1.2 summarizes the findings of this section. All elasticities are evaluated at mean empirical levels at the time of experimentation — approximately 10% for sell fees, 25% for buy fees, and a 6% display price markdown. That is, if a single ticket that some seller prices at $100 sells, the seller will receive $90, the buyer will observe a price of $94 while browsing but will pay a final checkout price of $119 (buy fees are calculated against *list* price, not display price).

The display price semielasticity is higher than the buy (checkout) fee semielasticity, indicating that the Platform must be hitting a constraint. Otherwise the Platform could lower the display price by percentage point and increase the buy fee by one percentage point which would increase volume without impacting total take rate per transaction. We confirmed with decision makers at the Platform that the display price policy was subject to this constraint. There was also a qualitative concern that consumers might learn to expect high checkout fees, i.e. long run checkout fee elasticities exceed those in the short run, but also that high state dependence might make it difficult for consumers to "unlearn" this believe if fees were to be reduced in the future. Both of these factors contribute to the shadow cost associated with the display price constraint.

We find in turn that the buy fee semielasticity exceeds the sell fee semielasticity, implying that buyers value additional sellers on the platform more than sellers value additional buyers. We discuss these platform effects in greater detail in Section 1.6.1.

Table 1.2: Summary of estimated elasticities from preferred specifications of three fee experiments. The final column ($\gamma$) gives implied risk aversion parameters defined in Section 1.4.6 and used in Section 1.6.

| | # Transactions | | | Transaction Volume ($) | | | |
|---|---|---|---|---|---|---|---|
| | Price | Fee | Semi- | Price | Fee | Semi- | $\gamma$ |
| Buy Fee | 2.092 | 0.390 | 1.648 | 2.665 | 0.479 | 1.959 | 1.872 |
| Sell Fee | 0.534 | 0.075 | 0.609 | 1.274 | 0.180 | 1.456 | 1.673 |
| Display Price | 2.059 | — | 2.133 | 2.412 | — | 2.450 | [1.504, 1.724] |

### 1.5.2 Sell Fees

**Experimental Design** In the fourth quarter of 2018, the Platform decided to run a large-scale experiment to directly estimate the impact of fees charged to sellers on market outcomes. A subset of approximately 200,000 users were randomly selected from among all users with at least one sale in the past year. Very large sellers — those with a specific designation based upon transaction volume — were excluded. These sellers typically had contracts negotiated on an ad hoc basis that could not be experimentally adjusted.

Of the 200,000 experimental users, equal proportions were randomly allocated to treatment and control arms. Randomization was stratified based upon transaction count and total transaction volume realized in the previous year. Experimental allocation was carried out prior to the preperiod and it was prespecified that users with one or more active listings at the beginning of the postperiod would be excluded. This restriction applied equally to both arms and was applied because the Platform specifies sell fees at the time of listing and as a matter of policy does not change them ex post. As a result of this restriction, a total of 124,153 total users were included in the experiment. Despite the restriction, the design achieved good balance even along dimensions not directly targeted as summarized in Table 1.3.

Table 1.3: Sell fee experiment: Preperiod summary statistics across treatment and control arms. Differences are Treatment minus Control; standard errors are computed from pooled variances.

| | | Treatment | Control | Difference | Std. Err. |
|---|---|---|---|---|---|
| Total | # Users | 62,010 | 62,143 | — | — |
| | Trans. Vol. | $10,204,182 | $10,517,096 | — | — |
| | Trans. | 44,457 | 43,954 | — | — |
| | Seats Sold | 90,559 | 89,288 | — | — |
| Per User | Trans. Vol. | $164.56 | $169.24 | −$4.68 | (7.941) |
| | Had Sale | 0.404 | 0.405 | −0.001 | (0.694) |
| | Seats Sold | 1.460 | 1.437 | 0.023 | (0.021) |
| | Seats per Trans. | 1.987 | 1.984 | 0.003 | (0.004) |
| | Sold Seat Price | $119.13 | $120.62 | −$1.50 | (0.862) |
| | # Transactions | 0.717 | 0.707 | 0.010 | (0.009) |

Both groups of users are observed for a total of 50 days before and after the experimental cutoff. Before the cutoff date, all users were presented with a 10% sell fee. After the cutoff date users in the treatment arm received a 15% sell fee while users in the control arm continued to receive a 10% sell fee. The experiment was applied on the basis of user ID. Because users selected into the experiment had existing accounts and users must be logged in in order to proceed through the sell flow, treatment was consistent regardless of which device a seller were using. The seller experience is described in more detail in Section 1.3.1; refer to Figure 1-1 for the screen presenting to each seller her sell fee. A seller who never reached the end of the sell flow had no way of observing her treatment. We directly observe whether each seller on each day entered the sell flow, in which case we infer that she had an intent to sell tickets. This institutional detail is an essential component of our empirical strategy that allows us to restrict focus to treatment effects instead of intent to treat effects.

The experiment was prespecified to run for equal length periods on either side of the treatment cutoff. Thereafter it was planned that most users, including both control users as well as users not selected for the experiment, were moved to different fee levels. Operational issues delayed this change by approximately two weeks, so we have an additional 14 days of unplanned experimental observation in the postperiod. We omit these days from much of our analysis in order to adhere closely to the prespecified test design, though we do run robustness checks with these data included.

**Data and Summary Statistics** Experimental data were collected over a 114 day period in late 2018. These include records of (i) all transactions occurring during the period, including quantities, prices, and all relevant fees, (ii) all listings created during the period, including initial list price and all list price updates and seller fees, and (iii) listing attempts at a page view or app interaction level, indicating whether a seller began the process of listing tickets for an event whether or not she ultimately created a listing.

Among the 124,153[11] total users in our experiment, 62,010 were in the treat-

[11] This total excludes users who had an active listing at the time of the experiment and whose exclusion was prespecified, as discussed above.

ment arm and 62,143 were control. We observe a total of \$39.5M in transaction volume, excluding fees, generating \$13.5M of total platform revenue across 160k total transactions and 331k seats, approximately balanced in the pre- and postperiods. Aggregating over all transactions, the mean seat sold for \$117 and the typical order had approximately 2.07 seats.

Aggregating over users, 54.7% had at least one transaction in either period. Conditional on having at least one transaction, the average user had 2.4 orders with 1.99 seats per order at an average price of \$124 per seat. This group of users had mean total transaction volume of \$584 and a median of \$250. Users with at least one preperiod transaction were 4.5% more likely to have a postperiod transaction than those without.

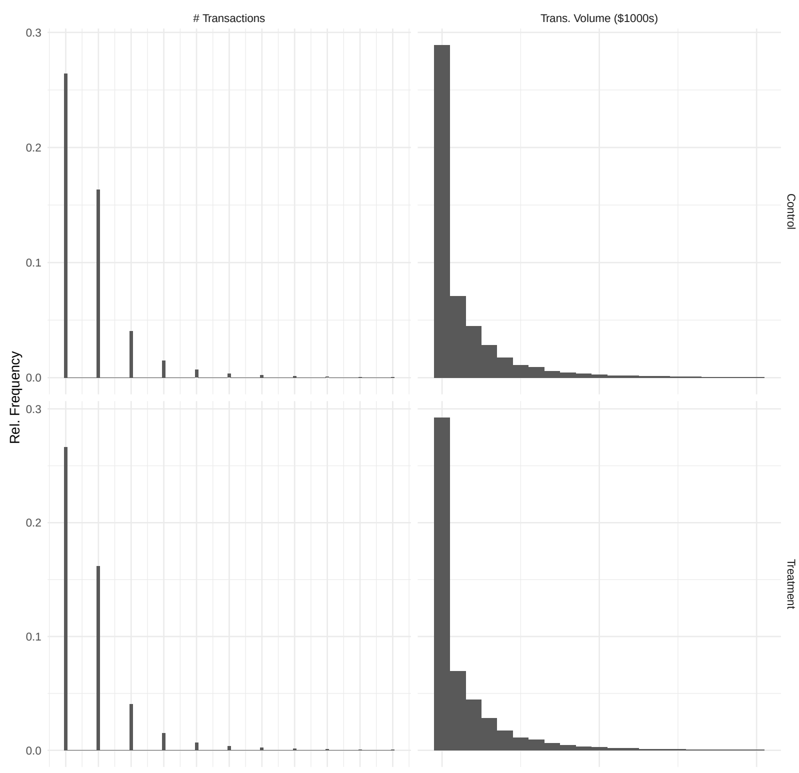



Figure 1-4: Sell fee experiment: Unconditional distributions of total transactions and total transaction volume in dollars, aggregated at the user level and unweighted. The modal user had zero transactions. Transaction volume bins are width \$100 and begin at zero.

Users were well balanced on observables across treatment arms during the preperiod, as summarized in Table 1.3. Only mean price of sold seats was meaningfully different (at the 10% level) across the groups relative to its standard error, though in absolute terms the difference was only approximately 1% of the mean value of either group. Applying the Bonferroni correction eliminates this significance and similarly a conservative Hotelling test fails to reject the null of preperiod balance.[12] Figure 1-4 shows the unconditional distributions of transaction counts and volumes in the preperiod.

**Regression Estimates and Elasticities** Figure 1-5 summarizes first differences between treatment and control groups in outcomes of interest, aggregated daily, and includes 14 additional days that we excluded from our analysis to maintain temporal balance, though estimates are robust to inclusion of these additional observations. We estimate regressions of the following form:

$$\log Y_{it} = \beta_0 \times \text{Treatment}_i + \beta_1 \times \text{Postperiod}_t + \beta_2 \times \text{Treatment}_i \times \text{Postperiod}_t + f(t) + \epsilon_{it}$$

where $f(\cdot)$ is a flexible time control and $\beta_2$ is the parameter of interest. We weight observations by the number of sellers entering the sell flow on each day, though results are robust to weighting by the number of sellers who reached by sell flow at least one on a prior date in each experimental period as well as by rerunning the regressions unweighted. For $f(\cdot)$ we consider linear time trends, second degree orthogonal polynomial time trends, natural cubic splines with four equally spaced interior knots, and day fixed effects.

Regression estimates are summarized in Table 1.4. Our preferred specification is the direct diff-in-diff estimate, summarized in the "Permutation" row. This specification directly compares aggregate per-user behavior before and after the treatment is applied relative to the control group, pooling all days. Adherence was nearly perfect and yielded a 40.5 log point increase in fees, equivalent to a 49.9 percent increase

[12]The test is a multivariate extension of the $t-$test first described by Hotelling (1931). The "conservative" test applies the restriction that sample covariances are zero, which serves to bias the test statistic away from zero given that sample covariances must all be weakly positive.

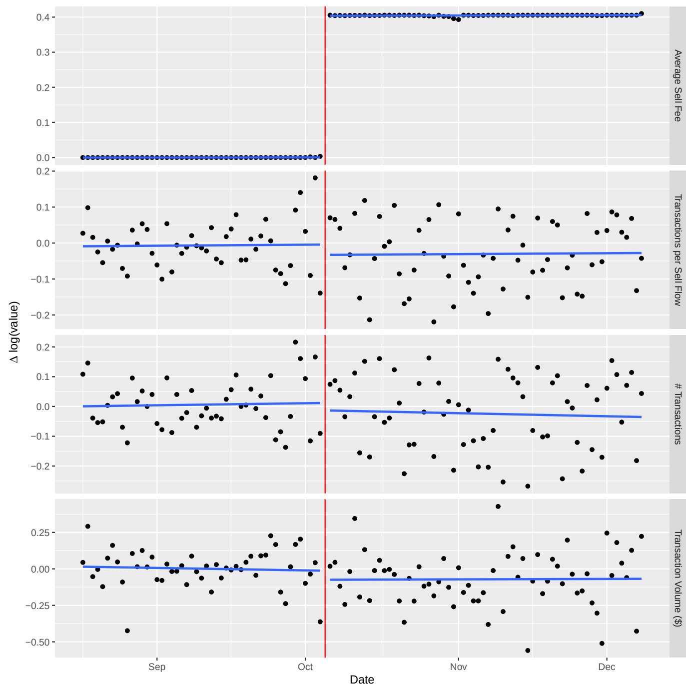



Figure 1-5: Sell fee experiment: Raw first differences (treatment minus control) in four outcomes (vertical panels) aggregated by UTC date. Vertical red line denotes the experimental cutoff. Postperiod includes 14 additional days relative to the preperiod.

which corresponds closely to the intended change from 10% to 15% fees. Column (3) transactions among users with an intent to list were reduced by around 2.6%. The treatment reduced total transaction volume by an even greater degree, approximately 5.7%, providing suggestive evidence that increased sell fees disproportionately impact higher dollar value listings — those with larger quantities and/or more expensive tickets — though we are underpowered in this experiment to perform that decomposition.

Point estimates from our preferred specifications give a price semielasticity of 0.609 for transactions and 1.456 for dollar volume, depending on specification. Similarly, the price elasticity of transactions of 0.534 and a price elasticity of dollar volume of 1.274. Finally, the fee elasticity of transactions is 0.075 and the fee elasticity of dollar volume is 0.180.

**Robustness** A number of concerns apply to our difference-in-differences design. While the clean experimental assignment and forced adherence guarantees uncon-

Table 1.4: Sell fee experiment: Regression estimates. Cells give difference-in-differences point estimates $\beta_2$ in log points (scaled up by 100 such that an estimate of 1.0 indicates 1 log point or approximately a 1% effect). Columns index outcomes; rows index time controls. Standard errors are given in parentheses below.

| | Mean Sell Fee (1) | Sell Flows (2) | Trans./Sell Flow (3) | Transactions (4) | Volume ($) (5) |
|---|---|---|---|---|---|
| Linear Time | $40.46^{***}$ | $-1.18$ | $-1.88$ | $-3.07$ | $-7.79$ |
| | (0.03) | (6.14) | (4.21) | (9.25) | (9.99) |
| Poly Time | $40.46^{***}$ | $-1.18$ | $-1.86$ | $-3.04$ | $-7.71$ |
| | (0.03) | (6.16) | (4.12) | (9.22) | (9.64) |
| Spline Time | $40.46^{***}$ | $-1.22$ | $-1.85$ | $-3.07$ | $-7.75$ |
| | (0.03) | (4.94) | (3.05) | (6.66) | (7.48) |
| Time FE | $40.46^{***}$ | $-0.68$ | $-2.37^{*}$ | $-3.05^{*}$ | $-7.28^{**}$ |
| | (0.03) | (1.04) | (1.55) | (2.04) | (3.14) |
| Permutation | $40.46^{***}$ | 0.309 | $-2.57^{*}$ | $-2.27$ | $-5.73^{**}$ |
| | (0.03) | (1.18) | (1.77) | (1.84) | (2.81) |

foundedness, we are concerned that serial correlation may significant bias our standard error computations as well as that the stable unit treatment value assumption may be violated.

To address the former concern we conduct a non-parameteric permutation test. On each iteration of the test we randomly assign each user to one treatment arm, maintaining the overall assignment proportions, then directly compute the second difference in each of our four outcomes of interest. We then compute sample variances directly from the empirical distribution of the estimates. Point estimates and computed standard errors from this procedure are presented in the final row of Table 1.4. We also plot the empirical distributions from the permutation test with 1000 iterations in Figure 1-6. Standard errors are similar to before and results are not materially different from the fixed effects specification.

The latter concern (SUTVA violation) is not one that we are able to address within this experiment. If a majority of all sellers were treated, plausibly some portion of the fee increase would be passed through and overall market prices would decline, partially counteracting some portion of our estimated effect. We have no *prima facie* expectation that the magnitude of this effect would be economically significant. After

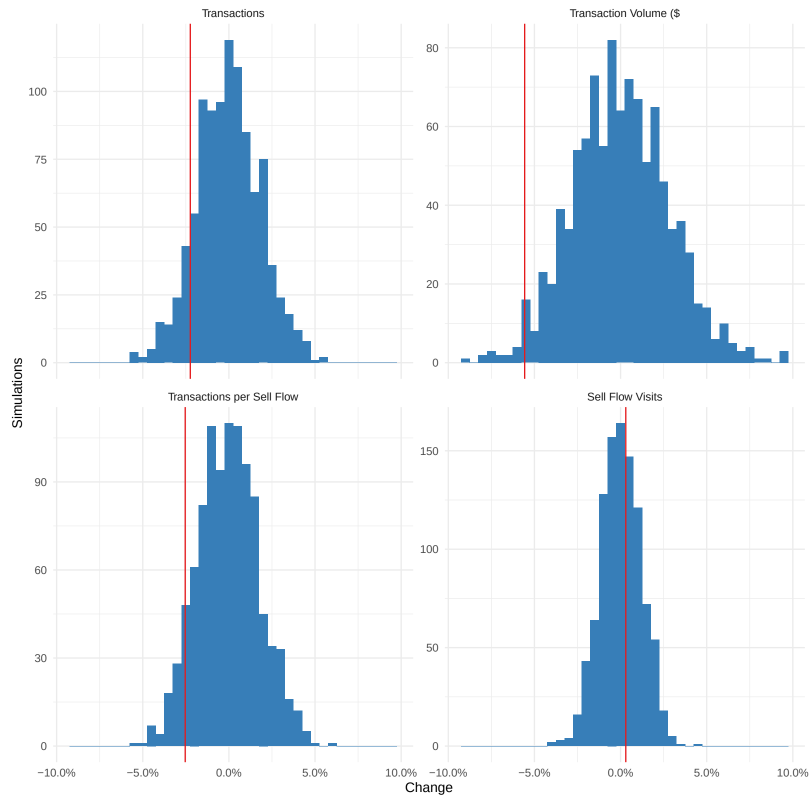



Figure 1-6: Sell fee experiment: Empirical distributions from permutation test with 1000 iterations. Horizontal axis gives estimates from each iteration; histogram bins have width 0.5%. Vertical axis gives count of samples falling within each bin. Vertical red lines indicate point estimate for true experimental allocations.

all, buyers may substitute to other platforms. For future work, we propose a future experiment with randomization at the event level, which would incorporate within-market general equilibrium effects and quantify the associated bias.

### 1.5.3 Buy Fee

**Experimental Design** The Platform operates a continually running buy fee experiment across a vast majority of its catalog. Each user is allocated a randomly generated session token during each new browsing session. This token is persistent within device but not within account and may be reset by clearing browser cookies or deleting and reinstalling a native app (iOS or Android). Unlike in the seller experiment, the same user receives a different session identifier on each device, even if she is logged in on both.

Buyer fees are set at the event level and are adjusted within business units based upon intuition about demand. A majority are set to default values, typically either 24 or 25%. Irrespective of the baseline fees per event, a fraction of users is served a random fee that is a constant percentage point level independent of the baseline fee. The experimental intensities is managed by a member of the business team on an ad hoc basis but generally followed a regular pattern. The modal event allocated 50.0% of users to the default fee and the remainder to experimental fees; the mean event allocated 55.4% and the median event allocated 53.5%. Among events running buy fee experiments, approximately 62.5% included two experimental variation while 37.5% included three experimental variations. Approximately 2.4% of events were excluded from experimentation.

Buyers are presented with event baseline fees, noted as "estimated fees" when they browse listings, though the default view presents prices exclusive of fees. After a user clicks through to the final checkout page, which requires them to either log in or provide a valid payment instrument (described in more detail in Section 1.3.2), her session token is sent along with the event token is sent to the experiment platform which computes a deterministic hash using the two strings. The binary representation generates a uniform random number on the unit interval with resolution $2^{-64}$ which is then compared against the intended experimental allocations for that event to determine experimental treatment. The information is passed back to the buyer's device and presented as in Figure 1-2.

**Data and Summary Statistics** We collect two separate samples at differing levels of aggregation. The first ("Sample A") is aggregated at the event-day level and includes 57,200 unique observations across 44 randomly selected days over 30 weeks between mid 2018 and early 2019 in order to approximately match the dates of the sell fee experiment. Each weekday appeared either 6 or 7 times in our sample, as did each month. The second sample ("Sample B") is aggregated at the user-event-day level and we obtain a sample of of 830,653 unique observations before sample restrictions. The sample covers eight UTC days over the same sample period as the sell fee experiment.

Both samples exclude special events, such as playoffs or exhibitions, and subsample at the event level. Due to concerns about the disclosure of confidential information, we have agreed not to disclose the sampling fraction or to report aggregates that might be indicative of overall business performance.

Both datasets were derived from browsing logs that record session tokens and other information required to compute buy fees as well as information about which tickets the user has viewed. These logs are are neither contingent on a transaction nor being logged in. We restrict attention to users who for each event-day arrived at the checkout page and would have observed their buy fee treatment. As discussed, baseline fees are variably adjusted in response to anticipated demand properties. We therefore exclude all users receiving these baseline fees and additionally exclude all events that for each event-day do not have at least two distinct experimental fee variations. That is, we intend to estimate exclusively using experimental variation in order to eliminate endogeneity concerns due to the variable baseline fee. After these restrictions were applied, the first sample retained 28,470 observations and the second 610,416.

Users with multiple transactions for the same event in the same day are treated as having made a single transaction but with all other quantities summed. That is, a user with two transactions of two seats and \$100 total price each is treated as having had one transaction for four seats and \$200. This convention only matters for the transaction regressions, where transacting is viewed as a binary decision rather than a count process. Indeed, over 91% of transacting users conducted only a single transaction for each event-day.

The mean experimental buy fee across the merged sample was 25.0% unconditionally and 24.4% conditional on having had at least one transaction. The mean order was for \$370 and 2.35 seats (median \$211 and 2, respectively) net of fees. Of users who reached the final checkout page at least one, the probability of at least one purchase was just under 20%. The median ticket viewed was listed at \$138 per seat exclusive of fees; the median quantity viewed was 2.

**Regression Estimates** Event level heterogeneity is significant, with total sales by event closely following a power law, as is day level heterogeneity within an event. Page visits, active listings, and conversion rate all increase significantly through the day of the event while list price tends to decline. To address this heterogeneity, we add event-day fixed effects in our preferred specifications. In OLS we achieve this by taking first differences

$$\Delta \log Y_{ijt} = \Delta \log(1 + \phi_{ijt})\beta + X_{ijt}\gamma + \epsilon_{ijt}$$

where $i$ indexes individuals, $j$ indexes events, and $t$ indexes time. Buy fees, denoted by $\phi$, and outcomes $Y$ are differenced from the daily mean

$$\Delta \log(1 + \phi_{ijt}) \equiv \log(1 + \phi_{ijt}) - \log(1 + \bar{\phi}_{jt}) \tag{1.5.1}$$

and similarly for $Y$. Therefore we interpret the fees as inducing a change in the *final price* rather than just the fee component. Covariates $X$ include, variously, the mean quantity of seats and per-seat price that a user is browsing in a given session, and are chosen to capture browsing intent in the spirit of Lewis et al. (2015). In addition to OLS, we estimate similar specifications by Poisson and logistic regression, which regress outcomes in levels, subject to the link function; in these cases we estimate regressions of the form

$$Y_{ijt} = \Delta \log(1 + \phi_{ijt})\beta + X_{ijt}\gamma + \alpha_{jt} + \epsilon_{ijt}$$

where $\alpha$ is an event-day fixed effect.

Where necessary we compute elasticities from the regression results. Both OLS run in logs as well as Poisson regressions give elasticities directly in their parameter estimates. For OLS this is easy to see directly. For Poisson with log link we have $\partial \log Y / \partial p = \hat{\beta}$ but our covariate is $\Delta \log(1 + \phi)$ hence the output is the elasticity. For the binary response model with logit link we have $\eta_p = \hat{\beta} \times (1 - P_0)$ where $P_0$ is the baseline choice probability. As a test, we ran both Poisson and logistic regression and computed the implied elasticities and standard errors from the differing point estimates and the two exactly coincided, consistent with the earlier findings.

Regression results from sample A are given in Table 1.5. OLS and Poisson regression results give different magnitude estimates for the elasticities – 95% confidence intervals around either point estimate fail to capture the other. We prefer the Poisson specification, which in column (1) would have been equivalent to running logistic regression observation by observation, though we pooled to the event-day-treatment level. All sample A regressions are weighted by the number of unique browsing session receiving a given experimental treatment for each event-day.

Table 1.6 contains results for regressions run on sample B. These regressions are unweighted as the unit of observation is user session-event-day. Alternative specifications and restrictions, for instance excluding users purchasing more than eight seats per event-day, minimally influenced the results. Expanding the sample to include users receiving baseline (dynamic) buy fees and similarly left the results essentially unchanged.

Table 1.5: Buy fee experiment: Regression results for Sample A. All regressions are weighted by number of unique browsing sessions. OLS is run in first differences; Poisson includes an event-day fixed effect.

| Model | Transaction | Qty. | Vol. | ATP | TPO |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| OLS | $-2.795^{***}$ | $-3.068^{***}$ | $-3.326^{***}$ | $-0.147^{*}$ | $-0.151^{*}$ |
| | (0.240) | (0.301) | (0.442) | (0.080) | (0.083) |
| Poisson | $-2.151^{***}$ | $-2.019^{***}$ | $-2.990^{***}$ | $-0.127$ | $-0.191$ |
| | (0.243) | (0.263) | (0.550) | (0.377) | (0.189) |
| Observation | 28,470 | 28,470 | 28,470 | 28,470 | 28,470 |
| *Note:* | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | | |

**Elasticities** Point estimates from the two samples matched closely despite being estimated on different samples and at different units of aggregation. The two samples selected a mutually exclusive set of days – no observation days included in sample A were included in sample B. Using inverse variance weighting to combine our esti-

Table 1.6: Buy fee experiment: Regression results for Sample B.

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Transaction | Quantity | Volume | ATP | TPO | Transaction | | ATP |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\Delta\log(1+\text{Buy Fee})$ | $-2.479^{***}$ | $-2.023^{***}$ | $-2.437^{***}$ | $-0.650^{**}$ | 0.020 | $-2.439^{***}$ | $-2.342^{***}$ | $-0.654^{**}$ |
| | (0.233) | (0.228) | (0.415) | (0.286) | (0.035) | (0.233) | (0.251) | (0.278) |
| Browse Quantity | | | | | $1.008^{***}$ | | | |
| | | | | | (0.0005) | | | |
| $\Delta\log(\text{Browse ATP})$ | | | | | | | $-0.127^{***}$ | |
| | | | | | | | (0.005) | |
| $\Delta\log(1+\text{Buy Fee}) \times \Delta\log(\text{Browse ATP})$ | | | | | | | $-0.033$ | |
| | | | | | | | (0.368) | |
| Implied Elasticity | $-2.056^{***}$ | $-2.023^{***}$ | $-2.437^{***}$ | $-0.650^{**}$ | 0.020 | $-2.056^{***}$ | $-1.985^{***}$ | $-0.654^{**}$ |
| | (0.160) | (0.228) | (0.415) | (0.286) | (0.035) | (0.166) | (0.180) | (0.278) |
| Observations | 610,416 | 610,416 | 610,416 | 104,106 | 104,106 | 610,416 | 610,416 | 104,106 |
| Browse Q FE | | | | | | × | × | × |
| Estimator | Logit | Poisson | Poisson | Poisson | OLS | Logit | Logit | Poisson |

*Note:* $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

mates our point estimate for the price elasticity of transactions is 2.09 while the price elasticity of transaction volume is higher at roughly 2.67.

These elasticities are implied price elasticities due to the fee change, not fee elasticities. The percentage change in fees is much larger than the percentage change in overall price due to the fee increase, since baseline fees are significantly lower than baseline all-in prices. We can compute the difference by noting that the mean buy fee in our sample is 24%, so for a 1% increase the ratio $\log(1.25/1.24)/\log(0.25/0.24) \approx 0.197$, implying fee elasticities for transactions and dollar volume of approximately 0.405 and 0.480, respectively. Alternatively, since there is some variability in overall fee levels by event, we can recompute our covariate as $\Delta \log(\phi)$ and reestimate, which gives buy fee elasticities directly and accounts for potential variability at the event-day level. Estimates using on sample B give elasticities of 0.390 and 0.479, respectively, for transactions and dollar volume, which closely coincide with the back of the envelope computation. Similarly reestimating using percentage point changes gives semielasticities of 1.648 and 1.959.

**Mechanism** To decompose the transaction volume effect we separately estimate effects on quantity of seats purchased (unconditionally), quantity of seats purchased conditional on transacting and average per seat price conditional on transacting. The price elasticity of quantity purchased (unconditionally) closely matches the price elasticity of transacting, confirmed by the tickets per order elasticity (column 5 in both tables). In the aggregated regression we cannot include controls for browsing behavior at the user session level; without these controls we obtain a small but negative point estimate. When run at the user level, though, this variation is entirely captured by the user's browsing behavior. By contrast, the average per-ticket price conditional on purchasing (column 4) is sharply lower at the user level when fees are higher. Taken in combination, we interpret these results to suggest that higher fees not only cause buyers to reduce overall buying intensity but additionally induce buyers to substitute to cheaper tickets and/or differentially affect higher dollar volume potential transactions.

Our results therefore capture both *intensive* and *extensive* margin responses to buy fee policy, similar to the findings in Blake et al. (2018); they term the two effects "quality upgrade" and "quantity", respectively. In contrast to their findings, we do not observe an impact of quality adjustments at the user level. That is, we find no evidence that a user browsing for a group of six tickets together will instead purchase five or seven tickets in response to higher or lower buy fees. However, we do find evidence that the composition of users responding to buy fee changes is nonrandom and disproportionately affects expensive transactions.

**Robustness** Just as in the sell fee analysis, we are concerned that serial correlation in our error terms may bias our standard error calculations. Therefore we conduct another permutation test using sample B with 1000 iterations. In each iteration we randomly counterfactually assign each user to a buy fee treatment arm and reestimate specification (1) — logistic regression of transaction probability — and record the point estimate. Of these, 0.8% returned more elastic estimates than our point estimate. The standard deviation of the simulated distribution is approximately 0.354, significantly larger than the computed standard error in the regression. We plot the implied null distribution from the regression along with the permutation distribution in Figure 1-7.

A remaining concern due to the experimental mechanism is that there may be some contamination if buyers substitute across devices. The same experimental extension suggested earlier — randomization at the event level — would suffice to quantify the impact of any potential contamination. Practically such an experiment is difficult to implement due to significant loss of power when we can no longer eliminate event level effects. Nevertheless, we have reason to believe that the magnitude of any resulting bias is minor given the growing prevalence of native app users, who are usually logged in and only use a single device (a smartphone). As a robustness check for this same type of contamination in the display price experiment that follows we restrict focus to these users and find that elasticity estimates are robust to the restriction.

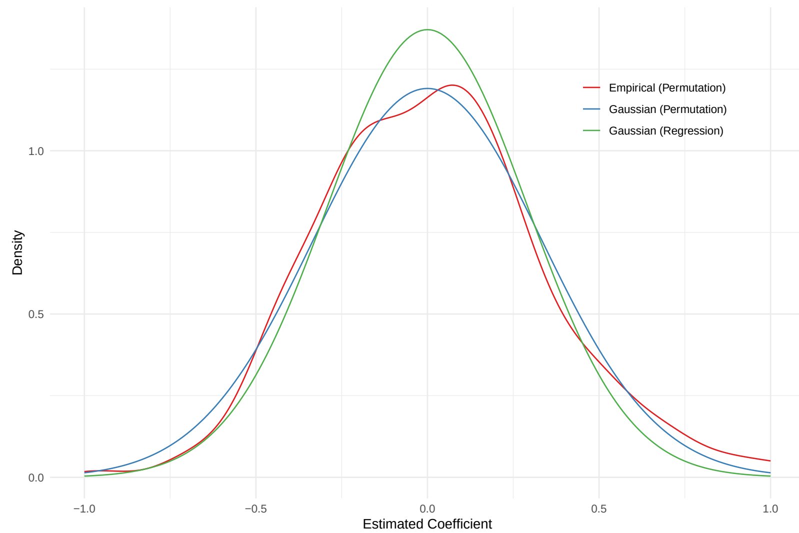



Figure 1-7: Buy fee experiment: Distribution of permutation test point estimates for specification (1) using sample B with 1000 iterations. True experimental point estimate not shown. Green distribution is implied null distribution based upon regression estimates. Blue is Gaussian fit with mean and standard deviation taken from the permutation test. Red is kernel smoothed empirical distribution from the permutation test.

### 1.5.4 Display Price

**Experimental Design** The display price experiment utilizes the same logic and testing platform as the buy fee experiment, relying on session tokens and event identifiers. When a user in a particular browsing session — often persistent within browsing device — arrives on an event page, her session token is hashed along with event information to yield a uniform random number which is then allocated a treatment arm. A smaller fraction of events are included in display price experimentation than in buy fee experimentation — approximately 20% of event-days in our sample was subject to display price experimentation versus nearly 80% subject to buy fee experimentation. Conditional on the event being subject to the experiment, modally 50% of users were allocated to the default treatment arm, which modally was set to a positive display price markdown, with the remainder assigned to a control arm receiving zero markdown. Other splits included 70%, 80%, or 90% of sessions assigned to the default arm, with allocations controlled by a member of the business team on an ad hoc basis. Display price is marked down by whole percentage point amounts, most commonly 5, 7, or 10%.

As a Platform tool, display price adjustments are specifically designed to exploit salience differences between list price and checkout price. Therefore the Platform typically "passes through" markdowns in the form of buy fees. However, the Platform operates under a self-imposed constraint that the total price charged to the buyer should be strictly increasing in the displayed price and therefore passes fees through incompletely. Modally, the Platform passes through 0.8 percentage points to buy fees, calculated relative to the baseline list price, for every 1 percentage point reduction in display price. Concretely, consider a $100 listing, and suppose control users are charged 25% buy fees, so see a list price of $100 and are charged $125 at checkout. Then a user in the display price markdown arm receiving a 5% markdown would see a list price of $95 and face a final checkout price of $124 ($95 + 29% × $100), resulting in a significantly higher effective buy fee relative to display price but a lower absolute all-in cost.

Display price markdown percentage and passthrough percentage are almost perfect collinear, precluding any direct estimation of their impacts. However, given our buy fee experiment, we can proxy for these effects and adjust our estimates accordingly. Let demand be given by $x(\text{display price}, \text{buy fee})$ where $\phi$ indicate fees and we use the convention that a display price markdown corresponds to $\phi_d < 0$. $\phi_b$ is the buy fee before any passthrough and $\alpha \geq 0$ is the passthrough fraction (empirically around 0.8). Then the observed effect is

$$
\begin{aligned}
&\log x((1+\phi_d)p, 1-\alpha\phi_d+\phi_b) \\
&-\log x(p, 1+\phi_b) \equiv \tilde{\eta}_d \log(1+\phi_d) = (\log x((1+\phi_d)p, 1-\alpha\phi_d+\phi_b) - \log x(p, 1-\alpha\phi_d+\phi_b)) \\
&\qquad\qquad\qquad\qquad\qquad\qquad - (\log x(p, 1+\phi_b) - \log x(p, 1-\alpha\phi_d+\phi_b)) \\
&\qquad \implies \tilde{\eta}_d \log(1+\phi_d) = \eta_d \log(1+\phi_d) - \eta_b \log\left(1 - \alpha\frac{\phi_d}{1+\phi_b}\right)
\end{aligned}
$$

where $\tilde{\eta}_d$ is the raw estimated display price elasticity before adjusting for the buy fee passthrough. In the above derivation we implicitly assume that all elasticities are locally constant. The true elasticity is therefore

$$\eta_d = \tilde{\eta}_d + \eta_b \frac{\log\left(1 - \alpha \frac{\phi_d}{1+\phi_b}\right)}{\log(1+\phi_d)} \tag{1.5.2}$$

where we apply the adjustment on the basis of empirically observed passthroughs $\alpha$.

**Data and Summary Statistics** We collect data covering a random sample of 36 browsing days, one day per week over 36 weeks between 2018 and 2019 covering the same period as the buy fee and sell fee experiments but with mutually exclusive observation dates. Each weekday is represented between 5 and 6 times in the sample. A total of 2,865 events are captured but after applying the restriction that each event-day had at least two experimental variations we are left with 635 unique events and 5,165 unique event-days. The mean event-day had $9,400 in transaction volume across 14.3 transactions.

The unit of observation is a browsing session-event-day, with an observation being recorded once a user views event page listings, irrespective of whether or not she clicks through the a checkout page. We group these sessions at the event-day-treatment level and for each record aggregate outcomes. A majority of event-days (70.2%) had two experimental variations while 26.2% had three and the remaining 3.6% had four. Of the included treatment-event-day observations, 53.0% were of control treatment arms while 11.1, 21.0%, and 14.9% were of 5, 7, and 10% display price markdown treatments. Weighted instead by the number of observed browsing sessions the respective shares were 36.0, 17.4, 23.8, and 22.7%, consistent with the observation that a number of experiments allocated well under 50% of participants to the control arm.

**Regression Estimates and Elasticities** We construct the display price treatment relative to the mean event-day treatment as in (1.5.1). We then estimate in the same way that we estimate the pooled (sample A) buy fee regressions — Poisson with log link, event-day fixed effects, and weights based upon the number of unique browsing sessions per event-day-treatment. Table 1.7 gives the result of the display price regressions. The raw elasticities due to variation in display price were 0.862 and 0.994 for number of transactions and dollar volume, respectively, yielding price elasticities of approximately 2.059 and 2.412. The corresponding semielasticities were 2.133 and 2.450. The effect on ticket price conditional on purchase is a noisy zero after adjusting for the buy fee passthrough effect.

Table 1.7: Display price experiment: Regression results estimated by Poisson regression with log link, including event-day fixed effects, weighted by number of unique browsing sessions. Implied elasticities were computed using (1.5.2).

| | *Dependent variable:* | | |
|---|---|---|---|
| | # Transactions | Volume ($) | Purchased Seat Price |
| | (1) | (2) | (3) |
| $\Delta$ log(Display Price) | $-0.862^{***}$ | $-0.994^{***}$ | $0.298^{*}$ |
| | (0.213) | (0.230) | (0.171) |
| Implied Elasticity | $-2.059^{***}$ | $-2.412^{***}$ | $-0.083$ |
| | (0.306) | (0.472) | (0.332) |
| Observations | 5,165 | 5,165 | 3,978 |
| *Note:* | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

**Robustness** The same source of contamination highlighted in the buy fee experiment applies to the display price experiment as well, likely to an even larger extent given that users observe their treatments before clicking through many pages. To evaluate the contamination we rely on a different data source that records browsing sessions from the "front end" — observing how users move through pages. The stages in the buy flow are summarized in Section 1.3.2. We observe stages 1–5 (initial browsing through final checkout page) but cannot observe stage 6 (checkout completed (post-transaction) page) due to data limitations. Nevertheless, session tokens and event information are recorded so we are able to determine the treatment applied to each session. As an additional precaution, we isolate our focus to iOS app users. These users are very likely to be logged in: 88% of iOS app events were conducted by a logged-in user versus 78% for the Android app and 36% for web. We additionally eliminate users unlikely to be browsing for tickets according to the following restrictions:

- Any user who viewed the “My Listings” page during the browsing session is excluded, as this is a likely seller simply monitoring the market(s) without an intention to buy.

- Any user who never viewed the event page is excluded (i.e. users who navigate directly to a single ticket without browsing).

- Any user who did not proceed through the checkout flow as expected is excluded. These may be brokers with automation or there may be an error in data collection.

We select a sample of all browsing sessions for all events of one professional sports team that occurred during a one-week period in late 2018. In total we observe just under 14,000 users of which approximately 4,800 were assigned to the control arm and 9,200 were treated with a 4% display price markdown but no passthrough.

Tables 1.8 and 1.9 describe flow through the buy funnel conditional on treatment; Figure 1-8 summarizes the corresponding relative conversion rates across treatment arms. Both indicate that users treated with the display price markdown were uniformly more likely to proceed to each stage in the buy flow relative to the control group, though as discussed our analysis excludes the very final checkout completion stage, which is the only stage that is revenue relevant for the platform. Therefore we present a regression summarizing final checkout behavior in Table 1.10. The point estimate implies a price elasticity of reaching the final checkout page (stage 5) of

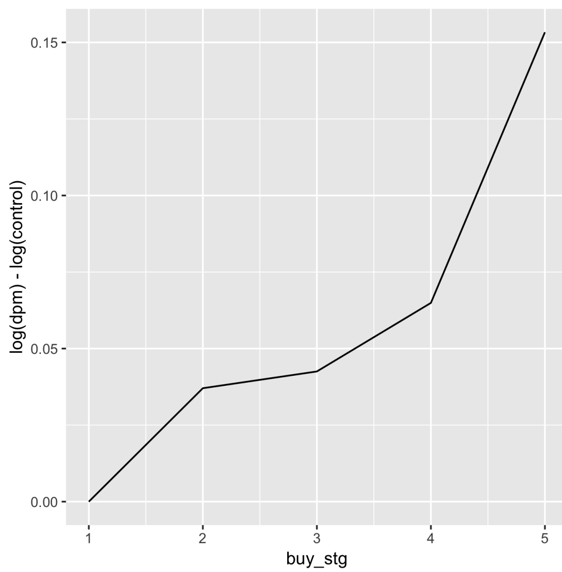



Figure 1-8: Relative conversion for DPM vs control group by buy funnel stage.

around 3.5, though the estimate is very noisy (95% CI $[0.39, 7.14]$), so the point estimate should not be taken too seriously. Critically, we are able to reject the null of no effect, which we would have observed given this restricted sample in the case that the DPM effect were driven mostly by users substituting across devices.

Table 1.8: Percentage of users at each stage in the buy flow, conditional on ever viewing the event page.

| Buy Stage | Treatment | Control |
|---|---|---|
| 1 | 100.0 | 100.0 |
| 2 | 41.6 | 40.1 |
| 3 | 39.8 | 38.2 |
| 4 | 16.9 | 15.9 |
| 5 | 7.38 | 6.33 |

Table 1.9: Percentage of users at each stage in the buy flow, conditional on reaching the previous stage.

| Buy Stage | Treatment | Control |
|---|---|---|
| 1 | 100.0 | 100.0 |
| 2 | 41.6 | 40.1 |
| 3 | 95.7 | 95.1 |
| 4 | 42.5 | 41.6 |
| 5 | 43.6 | 39.9 |

Table 1.10: Logistic regression of DPM on final checkout behavior. Includes controls for marketing channel, listing sort, device, and event. Excluded levels are direct (marketing), price (sort), computer (device), and an arbitrarily selected event.

| | | *Dependent variable:* |
|---|---|---|
| | | Checkout Review |
| Experiment: | Control Group | $-0.162^{**}$ |
| | | (0.074) |
| Marketing: | Organic Search | 0.017 |
| | | (0.082) |
| | Paid Search | $0.393^{***}$ |
| | | (0.076) |
| | Contextual | $1.071^{***}$ |
| | | (0.168) |
| | Email | $0.428^{***}$ |
| | | (0.164) |
| Sort: | Seat Quality | $1.346^{***}$ |
| | | (0.129) |
| | Best Value | $1.397^{***}$ |
| | | (0.103) |
| Device: | Smartphone | $-0.832^{***}$ |
| | | (0.077) |
| | Tablet | $-0.808^{***}$ |
| | | (0.218) |
| Constant | | $-2.599^{***}$ |
| | | (0.063) |
| Observations | 13,921 | |
| Log Likelihood | −3,315.624 | |
| Akaike Inf. Crit. | 6,651.247 | |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | |

## 1.6 Determinants of Platform Policy

In this section we consider the influence that incomplete salience and platform effects have on observed Platform fee levels. Operating under the assumption that our empirical context is in partial equilibrium, the equilibrium conditions of Section 1.4.3 and the calibrated representative agent preferences of Section 1.4.6 allow us to compute counterfactuals under perturbations to empirical platform effects and fee salience.

In the counterfactual analysis, we consider three sets of Cobb-Douglas parameters that govern total platform transaction volume as described in Eq. (1.4.5): $(\alpha_B, \alpha_S) \in \{(1,1), (.5,.5), (1,0)\}$. These cases are discussed at length in Section 1.4.2 and in the analysis that follows.

We calibrate the parameters governing the representative agent utilities by fitting the estimated semielasticities to (1.4.16) and (1.4.17). These calibrated parameters are supplied in the final column of Table 1.2. The implied $\gamma_B$ given these estimates 1.872. The corresponding estimate for $\gamma_D$ is bounded $\in [1.504, 1.724]$ depend on where we evaluate the semielasticity, as it is empirically nonconstant.[13] A similar exercise using the sell fee experiment gives an implied parameter value of $\gamma_S = 1.673$.

The final required inputs to the analysis that follows are empirically observed Platform fees and the Platform's variable cost structure. Aggregating over all users and all transactions during our sample period, including as well transactions with no experimentation, mean observed sell fees were 8.3% and mean buy fees were 24.1%. The discrepancy with the previously discussed typical levels of 10% and 25%, respective, are primarily explained by two factors: (i) a small number of large sellers have bilaterally negotiated sell fees and (ii) buyers occasionally use discount codes that reduce buy fees, though this is partially offset by buy fee passthrough. Internal accounting uses 5% of total all-in price (just over 6% of list price) as a default value

[13]It is not clear ex ante at which price level to evaluate the display price semielasticity in order to compare it with the buy fee elasticity. It directly affects prices between 90% and 100% of list price, suggesting that it should be evaluated at $1 + p^A \in [0.9, 1]$. On the other hand, buyers plausibly anticipate, albeit imprecisely, a significant buy fee at checkout, and are responding to expected all-in price, suggesting that we should evaluate at $1 + p^A \approx 1.25$. In practice we evaluate at $1 + p^A \in \{1, 1.25\}$ and consider the implications under both parameter values.

for variable transaction costs. The resulting Lerner index is between 0.809 and 0.821.

### 1.6.1 Impact of Platform Externalities

Recalling (1.4.11), the estimated short-run fee elasticities are far too low for the observed fee levels and overall markup to be profit maximizing for the Platform. Therefore platform effects are likely to play a significant role.[14] Proposition 1.4.3 gives closed form expressions for the implied platform elasticities, assuming again that the system is in partial equilibrium. We reproduce (1.4.14) and (1.4.15) below:

$$\eta_N^B = \frac{\sigma_p^B}{\sigma_p^S} - \alpha_S \sigma_p^B (p^* - c)$$
$$\eta_N^S = \frac{\sigma_p^S}{\sigma_p^B} - \alpha_B \sigma_p^S (p^* - c)$$

**Case:** $\alpha_B = \alpha_S = 1$ In the case when total transaction volume is some factor times the participation on each side, $\alpha_B = \alpha_S = 1$,[15] we obtain $\eta_N^B = 0.833$ and $\eta_N^S = 0.362$. Therefore increasing the number of sellers on the platform by 1% *ceteris paribus* must cause an increase in the number of buyers on the platform by 0.83% and similarly a 1% increase in the number of buyers must increase the number of sellers by 0.36% for observed fees to be optimal under this assumption.

Still assuming that $\alpha_B = \alpha_S = 1$, consider several shifts in platform elasticities. We may compute counterfactual equilibrium fees using (1.4.10), (1.4.11), (1.4.16), and (1.4.17), and the calibrated parameters $\gamma$ given at the beginning of the current section. In the extreme case of no platform effects ($\eta_N^S = \eta_N^B = 0$), equilibrium fees would be 34.9% and 15.7% to buyers and sellers, respectively, giving an overall wedge

[14]Several alternatives are possible. It is possible as well that long-run elasticities exceed short-run elasticities (Huang et al. (2018)). These would have to be approximately twice as large as short-run elasticities to rationalize the firm's fee policy, even under the most conservative Cobb-Douglas parameters in this respect ($\alpha_B = \alpha_S = 1$). As buyers can most easily substitute at the time of transaction, we find it likely that the relative long/short-run elasticities differ most significantly for sellers, the relative levels must be even starker. As such, we do not believe that long-run effects are sufficient to exclude significant platform effects. A second alternative explanation is that the transaction volume function $T$ exhibits significantly increasing returns to scale, but we view this to be even less likely.

[15]This is the case considered by Rochet and Tirole (2006) when payments between end users are permitted.

of 50.6%. Therefore the presence of platform effects nearly halves equilibrium sell fees while buy fees are reduced by over 30% (the total wedge is reduced by over one third).

As a less extreme example, hold fixed $\eta_N^S = 0.362$ but consider increasing $\eta_N^B$ modestly to 0.850. In this case optimal buy and sell fees are 24.1% and 8.1% — the buy fee decreases modestly but does not overcome rounding while the sell fee and overall platform fee both decrease, consistent with comparative static of Proposition 1.4.2. We obtain a qualitatively similar result when $\eta_N^B$ is held fixed while $\eta_N^S$ is modestly increased. This result provides context that the comparative static does not: at least in our empirical context, increasing the magnitude of the platform externality on buyers (sellers) disproportionately reduces sell (buy) fees. Though the result does not generalize, it may be more broadly applicable in similar contexts.

Two additional counterfactuals are informative to consider, both in which we maintain a constant product $\eta_N^B \times \eta_N^S$. This is the common ratio in the geometric series governing platform feedback, so holding this constant allows us to vary the relative platform elasticities without varying the overall feedback. When both elasticities are equal to the geometric mean of the baseline values, $\eta_N^B = \eta_N^S = 0.549$, optimal buy and sell fees are 18.4% and 12.4% and the overall fee level declines slightly. Still maintaining the product, when the buy fee elasticity increases significantly to $\eta_N^B = 1.000$ and $\eta_N^S = 0.302$ accordingly declines, optimal fees are 26.5% and 6.4% to each side and increase modestly overall. This qualitative evidence suggests that shifting platform elasticity toward the initially less (more) elastic side tends to reduce (increase) optimal overall platform fees.

**Case:** $\alpha_B = \alpha_S = 0.5$ In the case where total volume $T(N^B, N^S)$ exhibits constant returns to scale, implied platform elasticities are $\eta_N^B = 1.089$ and $\eta_N^S = 0.553$. Compared with the case when $\alpha_B = \alpha_S = 1$, elasticities are higher in absolute terms because additional users generate less additional transaction volume, hence higher implied elasticities are required rationalize the fee policy. However, relative elasticities are closer in the sense that the ratio $\eta_N^S/\eta_N^B$ moves toward 1. Intuitively, under the constant returns to scale specification, the platform must be acting to increase both sides of the market more equally as returns are decreasing to side $B$ growth holding side $S$ growth fixed.

**Case:** $\alpha_B 1,\ \alpha_S = 0$ In the limiting case, transaction volumes are simply proportional to the number of buyers, though this number is in general influenced by the number of sellers on the platform. This is equivalent to a transaction function $T$ that is globally Leontief but that typically has relative oversupply of sellers. In this case, $\eta_N^S = 0.362$ as before, consistent with Proposition 1.4.3, but $\eta_N^B = 1.345$ is sharply higher and buyers are elastic in seller participation. In fact, whenever $\alpha_S < 0.672$, equilibrium conditions are only satisfied if buyers are be elastic with respect to seller participation ($\eta_N^B > 1$). However, given the estimated semielasticities seller participation cannot be elastic in buyer participation for $\alpha_B \geq 0$. If both sides were elastic in other side participation then total participation would grow without bound as discussed in Proposition 1.4.1.

### 1.6.2 Impact of Salience

As discussed in Section 1.4, a platform with multiple fee levers affecting the same side of the market should marginally raise the fee with the lower price semielasticity and we would expect to observe equal semielasticities at prevailing fee levels. Our point estimates contradict this prediction, though we cannot reject that the semielasticities are equal at the 10% level. Additionally, as a matter of policy, the Platform self-imposes a constraint that display prices cannot be marked down more than 10% such that sellers never observe display prices less than their net payout.

Assume for the remainder of this analysis that $\alpha_S = \alpha_B = 1$, so the platform elasticities are as above: $\eta_N^B = 0.833$ and $\eta_N^S = 0.362$. From the equilibrium condition 1.4.10, *relative* fees satisfy $1.362 \cdot \sigma_p^B(p^B) = 1.833 \cdot \sigma_p^S(p^S)$. Substituting expressions for the calibrated representative agents (1.4.16) and (1.4.17), relative fees satisfy

$$1.362 \cdot \left((1+p^B)\log\frac{\gamma_B}{1+p^B}\right)^{-1} = 1.833 \cdot \frac{1}{1-p^S}\left(\frac{1}{\log(1.673 \cdot (1-p^S))} - 1\right)$$

To complete the system, we also need the overall markup equation

$$\frac{p-c}{p} = \frac{1 - 0.362 \times 0.833}{1.362 \cdot \frac{p^B}{1+p^B} \left(\log \frac{\gamma_B}{1+p^B}\right)^{-1} + 1.833 \cdot \frac{p^S}{1-p^S} \left[(\log(1.673 \cdot (1-p^S)))^{-1} - 1\right]}$$

The above two equations characterize $(p^{B\star}, p^{S\star})$ as a function of the buy fee parameter $\gamma_B$. We may evaluate the impact of incomplete salience by considering these equilibrium fee levels as we vary $\gamma_B$. Both equations are trivially satisfied under baseline values $(\gamma_B, p^B, p^S) = (1.872, .241, .083)$. If instead the platform were required to present fee-inclusive prices at all times, then using $\gamma_B = \gamma_D$ would violate both equations at prevailing fee levels.

First suppose that for institutional reasons the platform were constrained to leave sell fees unchanged ($p^S = 8.3\%$). Then under the conservative value $\gamma_D = 1.726$, the optimal buy fee would reduce by roughly three percentage points, to 21.0%. Evaluated at $\gamma_D = 1.504$, the optimal buy fee would be 15.9%.

If instead sell fees are permitted to fluctuate as well we need to solve both equations simultaneously. For $\gamma_D = 1.724$, optimal buy and sell fees are 18.4% and 10.7%; for $\gamma_D = 1.504$ these fees are 9.4% and 14.8%. Using these counterfactual fees as full salience baselines and for simplicity using the arithmetic mean to aggregate over the two values of $\gamma_D$, the results imply that incomplete salience increases equilibrium buy fees by 73.4% and reduces equilibrium sell fees by 35.1%. Both cases are consistent with Proposition 1.4.4, as the side of the market being made more salient is subjected to lower fees but the other side is charged higher fees. Overall fee burdens in the two cases are 29.1% and 24.2% compared with 32.4% at baseline, consistent with the second part of the proposition.

## 1.7 Conclusion

In this paper we developed a tractable model of two-sided platforms that lends itself well to the study of contemporary online platforms. From the model we then derived general optimality conditions for a profit maximizing platform. These are based upon a set of sufficient statistics — price and platform membership elasticities and semielasticities.

Using experimental data from one such online platform we directly estimate these price elasticities. Point estimates from our preferred specifications indicate that the price elasticities (semielasticities) of transaction volume (in dollars) are 2.41 (1.96), 2.67 (2.45), and 1.27 (1.46) for display price, buy fees, and sell fees, respectively, evaluated at their prevailing levels. Price elasticities of transaction counts lower, with elasticities of quantity transacted falling in between, indicating that fees disproportionately affect high dollar value platform goods (high price and/or high quantity). As additional evidence we examine the impact of fees on ticket price conditional on purchase and find buy fee elasticities between 0.2 and 0.6.

The overall level of observed fees was significantly too low to be optimal given the estimated price elasticities, provided that there were no network effects. The marginal benefit of lowering fees is greater in the presence of network effects relative to without as additional membership on side $A$ induced by lower fees increases membership on side $B$ due to the network effect, additionally increasing overall transaction volume. The implied platform elasticities — the percentage increase in side $A$ membership due to a 1% increase in side $B$ membership *ceteris paribus* — are 0.833 for buyers and 0.362 for sellers, i.e. if 1% more sellers exogenously joined the platform then 0.83% more buyers would join the platform. In the absence of platform effects, i.e. if platform elasticities were both zero, optimal fees to buyers and sellers would increase by 44.8% and 89.2%, respectively, an increase of 56.2% (18.2 percentage points) in overall fees.

Consistent with suggestive evidence from Section 1.2, fees presented at checkout are less salient than fees displayed during browsing, indicated by our estimated semielasticities of 1.96 and 2.45, respectively. Counterfactually, if the platform were forced to display all-in (fee inclusive) prices, optimal buy fees would be 23.7% lower and overall fees would be 10.2% lower (3.3 percentage points).

To our knowledge this paper is the first to directly estimate platform relevant elasticities and directly test optimality conditions in the presence of incomplete fee salience. Future work estimating these quantities using data from other platforms would improve our understanding of these burgeoning industries. Estimates of residual demand elasticities are indicative of market power while estimates of platform elasticities are critical to understanding entry and how these industries organize more broadly.

Finally, policymakers require additional tools to effectively regulate these growing marketplaces. Many traditional analytical methods, such as tests for predatory pricing or anticompetitive tying, fail to directly apply to two-sided marketplaces. *Ex ante*, consolidation in such industries has ambiguous welfare impacts, which depend the relative magnitudes of platform effects, residual demand elasticities, and platform cost structure. Our analytical and empirical framework can be used to determine likely price impacts due to a change in residual demand elasticities and platform membership resulting from a shift in market structure, such as a merger. We hope that it may support or inspire future work in this space.

# Chapter 2

# Localized Markets and Optimal Seller Policy in Online Marketplaces

Online exchanges enable a long tail of small sellers to operate in markets directly, without engaging an intermediary. These sellers may lack the information, sophistication, or attention necessary to price effectively, especially when competing with much larger, established sellers. To quantify the impacts of these potential limitations, we develop a model of demand that incorporates rapidly changing choice sets and buyer heterogeneity at the choice level, features common in the online setting. We estimate this model using data from a large online marketplace for live event tickets and use the result to compute implied properties of seller behavior as well as optimal pricing policies for small sellers. The estimated model implies that markets are highly local — each buyer only seriously considers a small subset of goods — with the consequence that additional information about granular demand properties is of limited value from an expected revenue standpoint, though dynamic pricing remains highly valuable as supply conditions shift and especially if goods are perishable. We confirm that a simple strategy of pricing near the lower envelope of price among similar goods obtains a high fraction of unconstrained optimal expected revenue and demonstrate that this policy may be effectively learned by a seller observing a sequence of market realizations. We conclude with suggestive evidence that human sellers learn to approximate this policy as they repeatedly interact in the platform.

## 2.1 Introduction

Online platforms have the potential to democratize market access, opening the door for many small sellers to join the ranks of suppliers. The amount of surplus accruing to new sellers as a result of this disintermediation is a function of how meaningfully such sellers engage with the market. If small sellers significantly misprice their goods, e.g. due to limited information or lack of sophistication, then in principle disintermediation might even reduce economic surplus. As a concrete example, the extent to which a homeowner benefits from using Airbnb versus a local vacation rental firm to price and manage her second home on Cape Cod depends upon how well she can manage pricing herself (on Airbnb) relative to how well the professional can manage pricing, net of the latter's fees.

The central contention of this paper is that, due to high market localization, small sellers require only elementary information about highly similar goods in order to compete effectively. Markets are localized in the sense that most individual buyers have highly limited effective consideration sets. Given information about the local market — in our empirical setting along spatial and quality dimensions — we find that rudimentary pricing strategies that can be simply approximated obtain expected revenues near those obtainable by much more sophisticated sellers with richer market information.

We begin by developing a model and estimation strategy that incorporates several core features of online marketplaces, including frequently changing choice sets and rich buyer heterogeneity, but remains computationally tractable. We take this framework to data from a large online marketplace for live event tickets. This market has a number of notable features that both match well with many contemporary online platforms and also facilitate empirical work. We observe every transaction as well as the full market state at the time of purchase. Additionally, we have complete demand-relevant information: there is little to no unobserved product-level heterogeneity, as seats are physically fixed and vary continuously through venues. Finally, products exhibit pronounced vertical and horizontal differentiation.

A typical choice set in our setting is large. Even when restricted to a single event and a single quantity of available seats, the modal choice set in our data exceeds 600 unique items. Individual buyers are unlikely to consider or even see a vast majority of these. Therefore our demand model incorporates, at the buyer level, a posterior type distribution describing the buyer's relative preference for quality and price. Applying the posterior significantly reduces the size of the buyer's effective consideration set — imposing the population distribution increases the number of choices required to reach an aggregate choice probability of 50% by a factor of 1.5. The posterior consideration set tends to small and highly localized: our estimates indicate that most buyers only seriously consider a small set of competitively priced listings within a narrow vertical quality band.

We use the demand estimates to study the optimal pricing policy of isolated, small sellers in partial equilibrium, in the sense that their pricing policies have no impact on other sellers' policies. As a consequence of market localization, such sellers need only to compare their own goods with highly similar goods and price near the lower envelope set in price-quality space in order to effectively compete. This substantially reduces the informational requirements for a successful sales strategy relative to markets in which a detailed understanding of demand and market dynamics is necessary to realize near-optimal expected revenues. Goods are perishable in the live events setting that we study, so rudimentary dynamics do play a role but, to a close approximation, only insofar as sellers must understand that the continuation value of an unsold ticket declines as the event approaches and fewer selling periods remain.

In order to quantify these effects, we evaluate revenue optimal selling strategies under various information and policy restrictions. The overall findings are that, while additional flexibility and sophistication has a pronounced impact on pricing, the revenue impacts are substantially smaller. However, despite their simplicity, sellers must still learn effective strategies. As a benchmark by which to evaluate empirical seller behavior, we train a "learner" with the objective to maximize expected revenue using machine learning techniques to train a predictive model on historical data that would have been observable to small sellers. Relative to this, the median small seller in a typical period prices her tickets much higher than the revenue optimal level. While potential differences in residual valuations for unsold tickets and related endogeneity concerns preclude a causal interpretation, we find suggestive evidence of seller learning in that sellers move toward the revenue optimal price with each successive market interaction, which we define as an attempt to sell a ticket whether or not a transaction was realized.

The remainder of the paper proceeds as follows. We conclude the current section with a discussion of related literature. Section 2.2 provides institutional context and describes several aspects of our empirical setting. Section 2.3 develops the structural demand model and discusses details of the estimation procedure. In Section 2.4 we present the resulting estimates and discuss the implied properties of demand. We use the model and estimated structural parameters to compute optimal price paths and evaluate seller behavior in Section 2.5. Section 2.6 concludes.

Our work contributes to the literature on dynamic pricing policy. Sweeting (2012) uses both listing and transaction data to study the determinants of seller behavior. Using an instrumented regression to estimate price elasticities, he finds that changing opportunity costs drive most of the observed pricing patterns. Similar empirical papers use hotel data to evaluate dynamic policy Abrate et al. (2012); Cho et al. (2018). Theoretical approaches include earlier works such as Gallego and Van Ryzin (1994), and extensions by Bitran and Mondschein (1997); Bitran et al. (1998); McAfee and te Velde (2008) as well a more recent work such as Board and Skrzypacz (2016). Sweeting and Sweeney (2015) considers several sellers dynamically pricing against one another while McAfee and Te Velde (2006) simulate optimal price paths in the airline industry.

We also contribute evidence to the emerging literature on online marketplaces. A number of recent empirical papers have documented high own price elasticities among close substitutes in online settings (Ellison and Ellison (2009); Dinerstein et al. (2018)). Dinerstein et al. studies sellers on eBay and documents many of the same consumer behaviors that we observe in a different setting. Fradkin (2019) studies Airbnb listings and incorporates an adjustment to the error structure in the spirit of Ackerberg and Rysman (2005), mitigating the effect of an exploding choice set in the online marketplace. As consumers increasingly engage on both sides of the sharing economy, better understanding their decision and learning processes would enable platforms to better support their participation. Especially for highly differentiated goods, small sellers provide an essential source of supply and enabling them to most effectively participate in the marketplace is likely to have significant, positive impacts on welfare.

## 2.2 Setting, Data, and Motivating Facts

We obtain data from a large ticket marketplace (the “Platform”) covering one professional sports team (the “Team”) for the full 2017 and 2018 seasons. Under the terms of the data use and confidentiality agreement under which the data were procured, we cannot disclose the identity of the team or the marketplace. The Platform collects and maintains rich data describing not only transactions but pre-transaction seller and buyer behavior as well. These data include clickstream logs of browsing behavior as well as a log price, quantity, and ticket feature updates for every single listing. As a consequence we observe full choice sets at every moment in time, whether or not a transaction occurred.

**Descriptive Evidence** As is typical for many sports teams on the Platform, in the 2018 season a majority of the Team’s sales came from a minority of events: the top 20% of events generated over two-thirds of revenue (see Figure 2-1). Volume for any particular event was characterized by many small buyers each selecting a few seats, as described in Figure 2-2. The modal buyer selects a pair of seats. Nearly 90% of transactions were for quantities of four or fewer; transactions for groups beyond

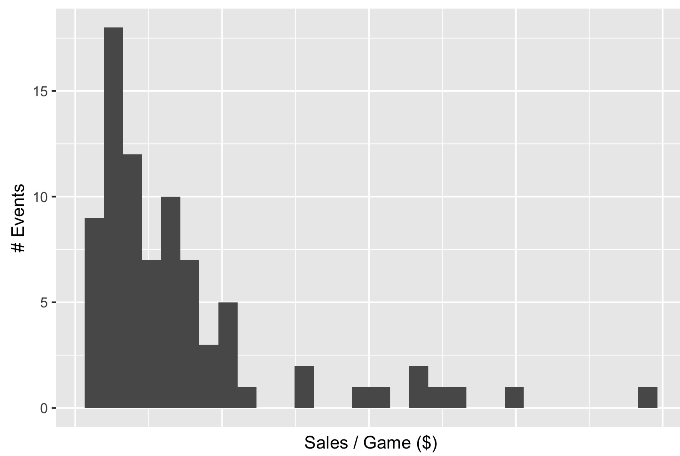



Figure 2-1: Sales per game, in dollars, for all games in 2018 season. While the horizontal axis is not labeled, it is scaled linearly, with the vertical axis intersection at $0.

four quickly and monotonically fall off in popularity. We do not observe significant evidence of shoppers searching and selecting across games. Modal buyers inspect only one event per browsing session, and very few investigate more than a few events, as in Figure 2-3.

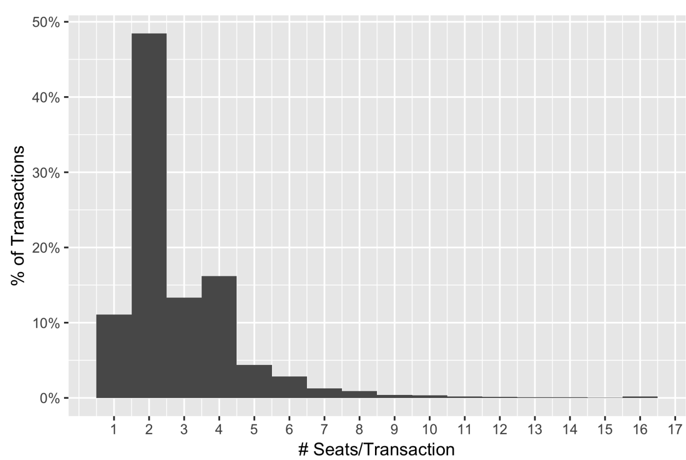



Figure 2-2: Share of transactions by number of seats per transaction. The histogram is winsorized at 16 seats.

We present a view of a typical market in Figure 2-4. In the figure we select a single event and plot every listing that was active at a given time slice as well as every transaction that occurred within an arbitrarily selected one-hour window. The plot is in quality-price space using vertical quality estimates developed in the following

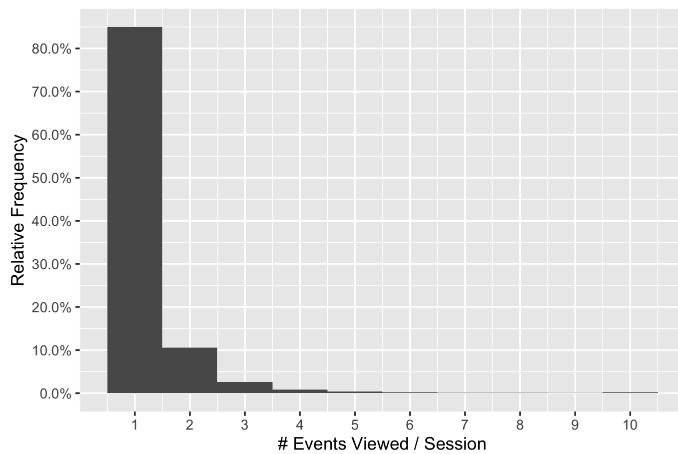



Figure 2-3: Relative frequency of browsing behavior. The horizontal axis gives the number of events viewed in a browsing session, winsorized at 10. Over 85% of browsing sessions included page views for only a single event; 99.5% of all sessions viewed six or fewer events. No sample restrictions were applied before computing relative frequencies, so the counts include, for instance, sellers with many listings checking on the status of multiple events as well as bots scraping market data.

sections. All transactions occur near the lower envelope set of listings, indicating that consumers are relatively good at search within an event. This is especially true at lower prices/seat qualities.

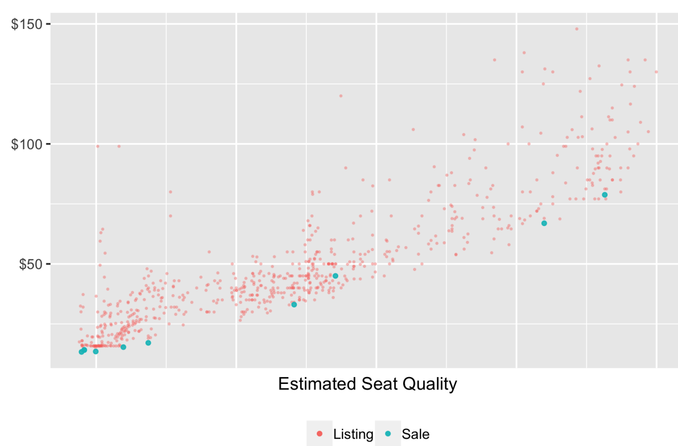



Figure 2-4: View of a typical market. Horizontal axis represents estimated seat quality, with higher quality seats to the right. Vertical axis is price in dollars. Orange dots represent listings for a single event that were active during a one-hour window on an arbitrarily selected day in 2018. Blue dots represent transactions that occurred in the same window.

The typical market for one of the Team's events exhibits a declining price path, i.e. the list price of a ticket at some given quality is expected to decline as the event approaches, at which time the unsold tickets expire worthless. As in Figure 2-4, most tickets that transact are near the "frontier" in quality-price space, therefore we build a index of event price using a quantile of quality adjusted price and plot it in Figure 2-5. The price decline is generated not only by price updates of existing listings but also by newly created listings.

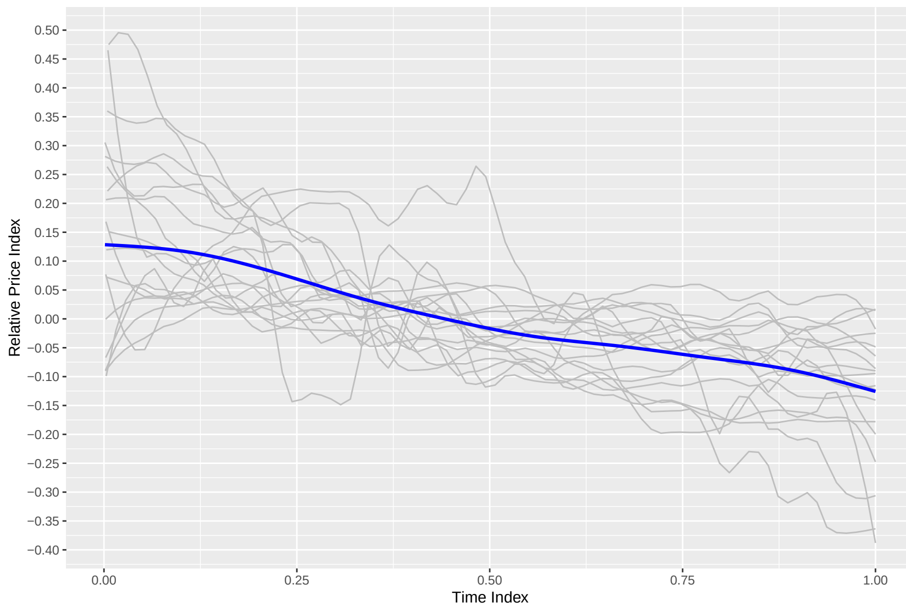



Figure 2-5: Price paths from a sample of events in 2018. Vertical axis gives thew $5^{th}$ percentile of quality-adjusted price in log points each period. Horizontal axis indexes time, from first transaction (0.0) to final transaction (1.0). Solid blue line gives a regression spline aggregated over all events; gray lines are LOESS fits by event.

**Data Description** The Platform provides transaction and listing data, including a detailed and timestamped record of every update to each (including, for instance, updates to listing price, delivery method, and quantity). In all of the analysis that follows, the basic unit of aggregation is the transaction, as the model is designed to capture intensive margin substitution, i.e. ticket choice conditional upon buying a ticket and in the case of sellers, sale probability conditional on some sale.

In each period, for each listing and each transaction, we observe anonymized seller and buyer identifiers as well as event, venue, section, row, seat(s), available quantities, and delivery method of the item. We merge on ancillary data based upon the venue-section-row tuple, which uniquely identifies a physical seat. Our setting is convenient in that there are no significant unobserved product attributes. For each venue-section-row, the available feature set comprises precise identifiers of (i) spatial row location within the venue, (ii) special (static) row features, and (iii) proxies for perceived location. The first group of features (i) includes angles and three distances to important venue features — line-of-sight distance, vertical, and horizontal displacement — which include, e.g., the front of the section, the closest point of the field, the closest point of the in-play area, the ordinal row rank, etc. The second set (ii) includes information about special zones which may, for instance, include free food, restrict alcohol, or be in the vicinity of kids' amenities. Group (iii) includes spatial indicators of similar properties as group (i) but based on the 2-dimensional projection presented to the buyer on the Platform's webpages.

At the time of each initiated transaction, we reconstruct the full choice set available to the buyer.[1] The choice set is restricted to listings with available purchase quantities matching the executed transaction. For instance, if a transaction were for two tickets then the choice set would include all listings from which a buyer could have chosen two tickets.[2] In the empirical work, we restrict attention to transactions for two, three, or four tickets. This subset captures approximately 80% of all transactions (see Figure 2-2) and practically ensures that choice sets are large. We estimate using events from the 2017 season while maintaining the 2018 data as a holdout sample for testing. In total across the 2017 season we capture 70,940 transactions for 184,575 total seats across 81 games, yielding \$6.83M in total sales, excluding fees. The mean ticket price was \$37.03 while the median was significantly lower, at \$25.25. The modal order was for two tickets, making up 61.4% of our sample, with orders for three and four tickets making up 17.0% and 21.6%, respectively. The mean choice set with available quantity two included 470 listings (median 453) and 30.6% of choice sets included

[1] This time may differ from the recorded transaction time due to automated fraud review or general network latency. In all cases we use the initiation time and not the completion time.

[2] Among these could be a listing with exactly two tickets as well as a listing with eight tickets that permits a buyer to buy two, leaving the listing with six tickets; in this case the two tickets that the buyer would receive is rule-based and not up to the discretion of the buyer.

more than 600 listings. Overall the mean choice set size was 399 (median 372).

## 2.3 Demand Model

In this section we describe a model of demand for the inside good. In the next section we estimate the model's parameters and consider its general empirical properties. We then use the empirical model to study optimal pricing policies in subsequent sections.

In an effort to match our empirical setting and to make estimation computationally tractable, the model that follows has three distinguishing features, each corresponding to different empirical facets. We briefly motivate and describe these below. Further exposition follows in Section 2.3.1, which outlines the core demand model, and Section 2.3.2, which outlines the estimation approach.

**Vertical Differentiation** Consistent with our empirical setting, seats within a venue are principally differentiated vertically. In order to generate a vertical quality index that will enter our demand model, we rely on the assumption that, at least locally, the difference in average transaction price between two seats should describe the dollar value quality difference after controlling for event and time factors. This index needs only to be approximately correct, as it enters the demand model through a flexible spline.

**Choice Sets** Live event ticket markets are well partitioned — a typical buyer has specific preferences for (i) a precise quantity of contiguous seats and (ii) a particular event (performer, venue, date). Even after this partitioning, choice sets are often large, typically ranging from several hundred to over one thousand distinct items available at any given time. Choice sets also change at high frequency due to purchases, price updates, and new listings. As a consequence we cannot rely on standard share based mixed logit estimators and must evaluate each choice set separately. However, we must also account for consumer heterogeneity at the choice level; to ignore it would impose overly restrictive substitution patterns. Incorporating such heterogeneity also substantially reduces the *effective* size of the choice set to any given consumer, which we discuss in more detail in Section 2.4.2.

**Consumer Types** While buyers approximately agree on the ordinal rankings of seats within a venue, the may disagree on their dollar values.[3] In the model, the distinction between a buyer selecting expensive seats behind home plate and one selecting cheap seats in the outfield is explained by different relative preferences for quality and expenditure. Users may also differ in the relative strengths of their idiosyncratic preferences for specific seats; the model incorporates this feature and allows it to covary with quality preference. At the choice level, consumer types are drawn from a posterior type distribution based upon the estimated population distribution and the buyer's observed choice.

### 2.3.1 Primitives

In general, a user $i$ with type $\theta_i$[4] has preferences represented by

$$u_{ijt} = \underbrace{v(x_j)}_{\text{quality}} - \underbrace{f(p_{jt};\theta_i)}_{\text{price}} + \underbrace{\tilde{\epsilon}_{ijt}}_{\text{idiosyncratic shock}}$$

where $j$ indexes products, $x_j$ are static product features, $v(\cdot)$ is a map from product features into quality index, $z_{jt}$ is a vector of time-varying market features, and $\tilde{\epsilon}$ is an idiosyncratic preference shock. In order to better match reality we relax homoscedasticity, so $\text{Var}(\tilde{\epsilon}_{ijt}) \not\perp \theta_i$, which we model in the following way

$$u_{ijt} = v(x_j) - \exp(\theta_i)p_{jt} + \exp\{\alpha_1(\theta_i - \bar{\theta})\}\epsilon_{ijt} \tag{2.3.1}$$

[3] The motivation for and contribution of each model element may best be described by example. Consider the market for tickets to a Yankees-Red Sox game at Fenway park in late August. The physical seats are invariant across games and are endowed with a consistent partial order. As a cheapskate, this paper's author might never purchase seats on the Green Monster or directly behind home plate, but nevertheless agrees that they are superior to the upper right field bleachers. Despite sharing the consensus quality ranking, the same cheapskate has a low absolute preference for quality and a high dispreference for spending money. He also has a low *relative* preference for quality and may frequently choose the cheapest available ticket when significantly higher quality tickets are available at only marginally higher prices.

[4] Throughout, a higher type $\theta$ corresponds to more price elastic agents.

where now $\epsilon \sim \text{Gumbel}(0, 1)$. The parameter $\bar{\theta} \equiv \mathbb{E}[\theta]$ is the mean consumer type and centers the elasticity distribution. The additional parameter $\alpha_1$ governs the relationship between consumer type and the variance of the idiosyncratic shock.

We model consumer types $\theta$ using a discrete distribution[5] with finite support $\theta_k \in \Theta$ and cardinality $\|\Theta\| = K$. The unconditional probability that any consumer is type $\theta_k$ is given by $s_k$. A practical difficulty is ensuring a property density: $\sum_k s_k = 1$. Therefore, we parameterize population shares $s_k$ as

$$s_k = \frac{\exp\{\phi_k\}}{\sum_m \exp\{\phi_m\}} \tag{2.3.2}$$

with the normalization $\phi_1 = 0$, which is non-restrictive since softmax is only defined up to an additive constant. This definition eliminates one parameter ($\phi_1$) and one constraint ($\sum_k s_k = 1$). Given $s_k$ and an observed choice $j(i)$ for consumer $i$, the posterior type distribution satisfies Bayes' rule

$$s_{ik} = \frac{s_k \Pr\{\text{choose } j \mid \theta_k\}}{\sum_m s_m \Pr\{\text{choose } j \mid \theta_m\}}$$

### 2.3.2 Estimation

#### Quality Index

We found it computationally prohibitive[6] to estimate the vertical quality map $v$ jointly with the demand model, as the feature space $X$ is high-dimensional and the class of functions $v$ over which we optimize during estimation must be rich.[7] As a solution, we develop a scalar "quality index" to initialize our estimation procedure. The details of this index construction are provided in Appendix B.1. To a first approximation, the quality index may be thought of as the output of a log price regression on rich

[5]We considered parametric families but concluded that the additional flexibility afforded by tracing out a nonparametric density outweighed the higher dimensionality of the parameter vector. Empirically, we have sufficient data to trace out the type density over broad support.

[6]For estimation we used a server with an Intel Xeon processor (48 logical cores, 4.4GHz operating frequency), 384GB (6×64) DDR4-2933 RAM, and an NVIDIA QUADRO RTX 8000 GPU (4608 CUDA cores, 576 tensor cores, 48GB dedicated ECC memory).

[7]For instance, due to its alignment, the first base side of Fenway Park is unbearably hot during afternoon games in the late summer, a preference consistently represented in the data.

seat features, flexibly estimated to incorporate non-linearities and interactions.

We expect this index to be ordinally but not cardinally correct (in the context of the demand model), so we apply a flexible spline as follows. Let the estimated quality indices be denoted $\hat{v}_r \in \mathbb{R}$ for each unique row in the stadium $r \in R$. Consider a natural cubic spline basis over the support of $\hat{V}$: let the function $\text{BS} : \mathbb{R} \to \mathbb{R}^m$ be a map from some estimated quality $\hat{v}$ into the spline basis and let $z_j$ be a representation of good $j$'s quality under the spline basis.[8] Then we approximate the cardinal quality for product $j$ in the demand model as

$$v(x_j) = z_j \equiv \text{BS}(\hat{v}_{r(j)})\beta$$

where $r(j)$ is a map from product $j$ to its unique (physical) row in the stadium.

### Demand Parameters

Analytically, it proved convenient to separately estimate the mean consumer type and the distribution of consumer type deviations. That is, we separately estimate the first and higher central moments of the type distribution, by adding the mean parameter $\gamma \equiv \bar{\theta}$ and the normalization

$$\mathbb{E}[\tilde{\theta}] = \sum_k s_k \theta_k = 0 \tag{2.3.3}$$

where $\tilde{\theta} \equiv \theta - \bar{\theta}$. This leaves the number of free parameters unchanged but has practical advantages. [9] Henceforth, for notational simplicity, we write $\theta$ in place of $\tilde{\theta}$.

After substituting the quality index spline and the type parameterization, utility

[8]For estimation we use 8 knots placed at equally spaced quantiles of $\tilde{V}$ and no intercept. The natural cubic spline is linear beyond its boundary knots hence the dimension of the basis is equal to the number of knots (in our case, 8).

[9]We experimented with relaxing this constraint, eliminating $\gamma$, and both simultaneously, and found that the true type distribution is more easily recovered with the constraint than without. Specifically, the empirical variation in Monte Carlo estimates was tighter and estimates converged more quickly.

is linear in parameters and our final estimating equation is

$$u_{ijt} = \exp\{-\alpha\theta_i\}\left(z_j\beta - \exp\{\theta_{ijt} + \gamma\}p_{jt} + \epsilon_{ijt}\right) \tag{2.3.4}$$

We complete the estimating specification with two normalizations. As a location normalization, we set the median within-venue quality $z_j$ to be zero, thus normalizing the willingness to pay for the average consumer. In turn, we normalize the idiosyncratic term $\epsilon$ to be standard Gumbel and the mean of $\theta$ to be zero, so the mean buyer obtaining this same free, median quality ticket draws utility from the standard Gumbel distribution.

We are left to estimate a total of 31 parameters: eight elements of $\beta$, two scalars $\alpha$ and $\gamma$, and 21 elements of $\phi$ corresponding to 21 types $\theta$. These types were selected to be centered at zero and to span 12 times the standard deviation of elasticities estimated by a simple representative agent model to ensure that we captured the entire relevant support.

Let $i$ index consumers, $j$ represent potential choices, and $l$ represent observed choices such that $l(i)$ is the observed choice of consumer $i$ facing choice set $J(i)$. The likelihood of an observation $i$ given utility parameters $\zeta \equiv (\beta, \alpha, \gamma)'$ and type distribution parameters $\phi$ is

$$L_i(\zeta, \phi) = \mathbb{P}(l(i)|\zeta) = \int_\Theta \mathbb{P}(l(i)|\theta, \zeta)\; dF(\theta|\phi)$$

For compactness we shall write $\mathbb{P}_{ij,\theta_m} \equiv \mathbb{P}_i(j|\theta_m, J(i), \zeta)$ to denote the conditional likelihood that consumer $i$ chooses alternative $j$ from $J$ given parameters $\zeta$ (suppressed in this notation) and type $\theta_i = \theta_m$. Subject to discrete $\theta$ and shares given by (2.3.2), the maximum likelihood estimator constrained by (2.3.3) solves

$$\begin{aligned} \max_{\zeta,\phi} \quad & \ell(\zeta, \phi) = \sum_i \log\left(\sum_m s_m \mathbb{P}_{il(i),\theta_m}\right) \qquad (2.3.5) \\ \text{s.t.} \quad & \mathbb{E}_{F(\phi)}[\theta] = \sum_m s_m \theta_m = 0 \end{aligned}$$

From (2.3.5) we form the following Lagrangian with two terms added to the log-likelihood

$$\mathcal{L}(\zeta, \phi, \lambda_1; \lambda_2) = \sum_i w_i \log \sum_m s_m \mathbb{P}_{il(i),\theta_m} \\ - \underbrace{\lambda_1 \sum_m s_m \theta_m}_{(i)} - \underbrace{\frac{\lambda_2}{2} \sum_{k=2}^{K-1} \left( \phi_k - \frac{1}{2} \left( \phi_{k-1} + \phi_{k+1} \right) \right)^2}_{(ii)} \tag{2.3.6}$$

where term $(i)$ is the Lagrangian penalty for our zero mean constraint on the population type distribution of $\theta$ and term $(ii)$ adds regularization that smoothes the type distribution to aid empirical estimation. We discuss the algorithmic procedure and develop the required analytical expressions to solve (2.3.6) in Appendix B.2.

## 2.4 Results

In the previous section we outlined and motivated a demand system and the required sub-components as well as estimation techniques. In the current section we first present the resulting estimates then discuss their implications within the empirical context. Our core findings are (i) that buyer heterogeneity is an essential model component for generating intuitively sensible choice probabilities and substitution patterns and (ii) that under the estimated model, implicit markets are highly local in quality space. In Section 2.5 we consider the implications of this latter finding on optimal seller behavior.

We estimate both model components using all listing, browsing, and transaction data from a large online live events marketplace covering a single professional sports team during the 2017 season. A superset of these data, covering both the estimation sample and an “out of sample” dataset used for later analysis, are further described in Section 2.2.

### 2.4.1 Estimates

We begin by estimating a quality index for each physical row $r \in R$ in the stadium with $\|R\| \approx 2000$. An adaptive grid search over hyperparameters yielded a final kernel parameter vector of $\vec{\vartheta} = (0.025, 0.003, 2.000, 0.200, 0.100)$, a tree depth restriction of three, a mild $L_2$ penalty, and a moderate pruning threshold. Figure 2-6 describes the estimated row quality distribution. The final out of sample root mean squared prediction error was approximately 15% when evaluated against observations in future seasons.[10] That is, given a model estimated on data from the 2017 season for the team of interest, then for the typical transaction in the 2018 season we are able to predict its price with 15% RMSE (roughly 12% mean absolute error) knowing only its section and row, subject to observing a sequence of previous transactions.

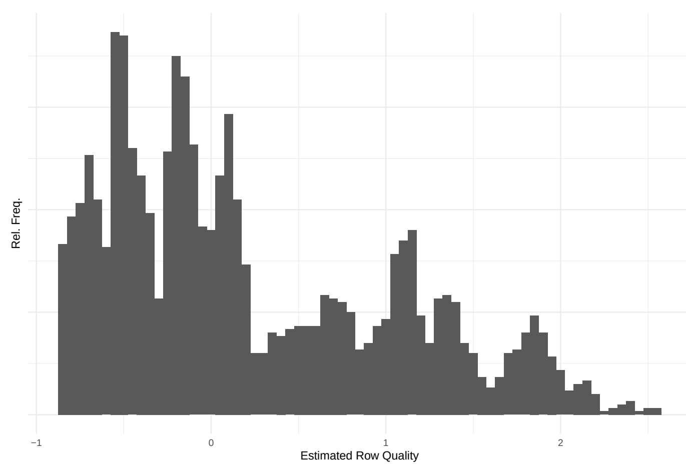



Figure 2-6: Estimated row quality distribution in hundredths of log point difference from median row by quality. Histogram plots relative frequency (fraction of rows within the stadium) by binwidths are five log points in quality. Overall, the stadium includes approximately 2500 unique rows.

Though these estimates must be rescaled to enter the demand model, the raw quantities are interpretable as approximate relative monetary values. Overall, 44.8% of all rows in the stadium are within 50 log points of the median row by estimated quality while 22.9% of rows are more than 100 log points better than the median. The worst row is approximately 85 log points worse than the median while the best

[10]The error metric is discussed in more detail at the end of Section B.1.

is approximately 255 points better. Note that many sports stadiums, including this one, expand approximately radially from the center of plate, and lower quality rows are typically further out. As a consequence, each of these rows tends to have more seats such that if we were to reweight by seat count rather than row count we would expect to have even greater density at lower quality rows.

We estimate the full demand model using this estimated quality index. In total the model includes 31 parameters, of which ten enter utility directly and the remaining 21 give the type distribution $\theta$. We present the utility parameters in Table 2.1 and the estimated type distribution in Figure 2-7. As a test of robustness we reestimate the same model with a finer grid of 63 discrete $\theta$ gridpoints as well as with a spline approximation to the density function. We find that utility parameters are robust to these various specifications: in the latter case the estimates for $(\gamma, \alpha)$ are $(-1.573, 0.228)$ which are nearly identical to our baseline estimates.

Table 2.1: Utility parameter estimates. Standard errors in parentheses are computed from the observed information matrix evaluated at the point estimates.

| Description | Parameter | Estimate |
|---|---|---|
| Mean Price Elasticity | $\gamma$ | $-1.573$ |
| | | (0.008) |
| Idiosyncratic Shock Scale | $\alpha$ | 0.234 |
| | | (0.014) |
| Quality Spline | $\beta_1$ | 7.019 |
| ↓ | | (0.187) |
| | $\beta_2$ | 9.026 |
| | | (0.216) |
| | $\beta_3$ | 10.733 |
| | | (0.218) |
| | $\beta_4$ | 12.930 |
| | | (0.222) |
| | $\beta_5$ | 14.868 |
| | | (0.217) |
| | $\beta_6$ | 18.272 |
| | | (0.147) |
| | $\beta_7$ | 26.201 |
| | | (0.303) |
| | $\beta_8$ | 22.498 |
| | | (0.174) |

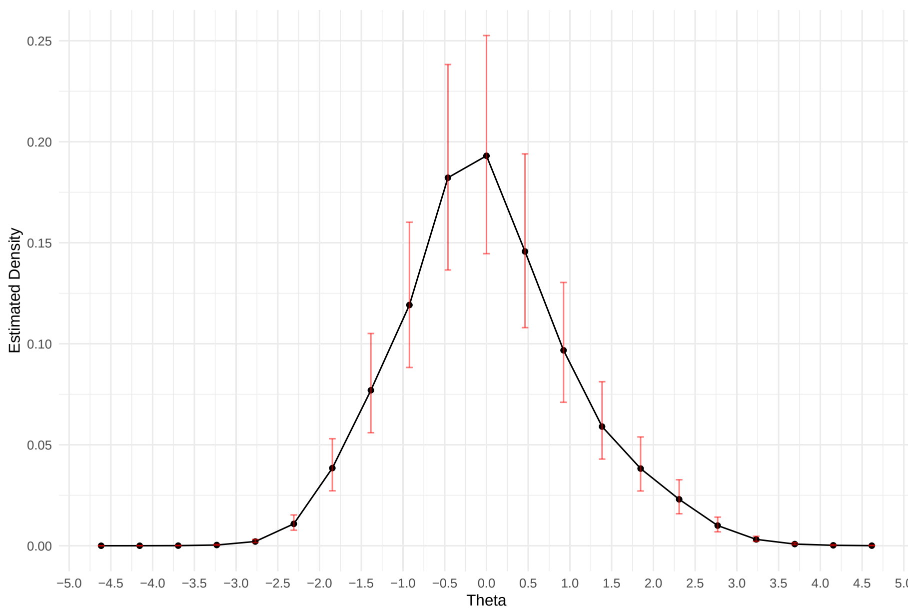



Figure 2-7: Estimated consumer type distribution over $\theta$ grid. Black points give point estimates. Red error bars give 95% confidence intervals based on estimated standard errors.

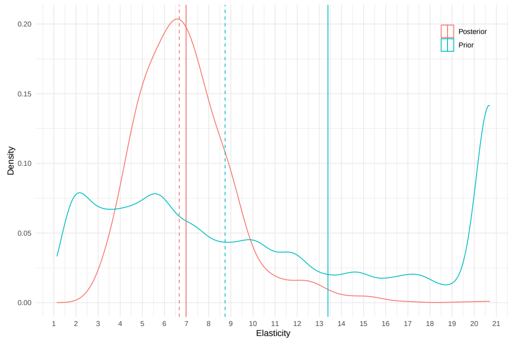
### 2.4.2 Implied Demand Properties

To evaluate price elasticities, we randomly sample 2000 transactions and choice sets and compute the implied own price elasticity of demand for the good that was ultimately chosen as well as for a randomly selected good from among the choice set.[11] Additionally, we evaluate each under both (i) the posterior type distribution of each consumer and (ii) the prior (population) type distribution. Mean own price elasticity, defined as $-\frac{\partial \log \mathbb{P}_{ijt}}{\partial \log p_{jt}}$, for the chosen good from this sample was 6.97 (median 6.67) under posterior share estimates and 13.39 (median 8.74) under prior shares. While these elasticities are qualitatively high, we believe that given the ease of search along the intensive margin in this setting, the estimates are quite plausible. Dinerstein et al. (2018) study a similar e-commerce setting and find own price elasticities in the range of 9.95 and 17.09.

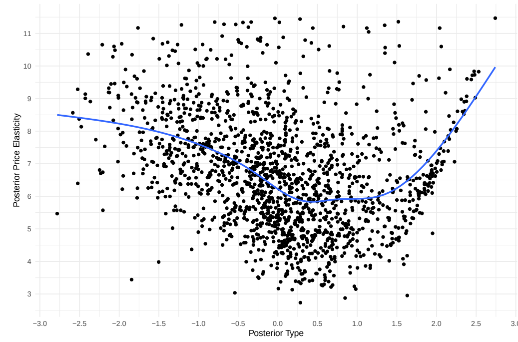
Panels (a) and (c) of Figure 2-8 describe the full distributions of these estimates among the selected sample. Under posterior shares, the price elasticity of the chosen listing has a distinct mode around 6.5 while the same distribution under prior shares is much more diffuse and has significant mass at elasticities above 20. By contrast, the distribution of price elasticities for a randomly selected listing from the choice set is much more similar (relative to that for the chosen listing) under the two type distributions; both have modes around 3 and both have substantial mass at even higher elasticities — above 30. This collection of facts highlights the function of our chosen parameterization: the posterior type distribution gives a clear definition for each consumer's *effective* consideration set, consistent with the empirical fact that a typical consumer only scrolls through a small subset of her available choices.

To quantify this effect, Figure 2-9 describes, for the sample described previously, the fraction of choices required to accumulate an aggregate choice probability above four thresholds. For instance, the top right panel of the figure gives (a distribution of) the percent of each consumer's choice set required to generate an aggregate 50% probability of purchase. Under the estimated posterior type distribution, the mean

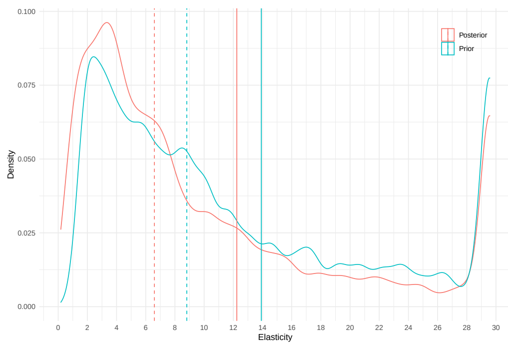
[11] The price elasticity parameter cannot be interpreted directly as it is dependent on typical qualities of available listings, overall choice set size, and posterior consumer type.



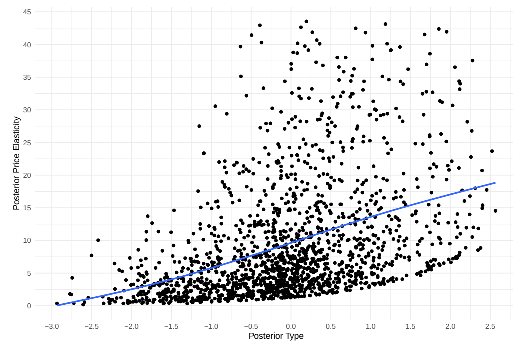

(a) Elasticity Distribution, Chosen Listing




(b) Type vs Elasticity, Chosen Listing




(c) Elasticity Distribution, Random Listing




(d) Type vs Elasticity, Random Listing

Figure 2-8: Implied price elasticities of demand for the chosen good from each choice set (panels (a) and (b)) and a random good from each choice set (panels (c) and (d)). Left panels give kernel smoothed densities over the implied elasticity distributions, winsorized at the $95^{th}$ percentile; means (solid vertical lines) and medians (dashed vertical lines) are computed prior to winsorization. Both left panels plot both the distribution is computed using the posterior type distribution for each consumer (red) and using the population prior (teal). Right panels plot individual elasticity estimates against posterior expected types for each consumer; solid blue line gives a regression spline.

consumer required 5.86% of her choice set to reach this aggregate choice probability while under the prior type distribution this quantity was was approximately 50% larger, at 8.73%. The relative shares of the choice set required to achieve a 20%, 90%, or 99% aggregate probability of being chosen follow a similar pattern.

Figures 2-10 and 2-11 give several concrete examples of effective consideration sets – in the case of the figures, these are sets of listings generating an aggregate choice probability of at least 50% – under prior and posterior consumer types. The figures highlight that the effective consideration sets not only contract significantly after the

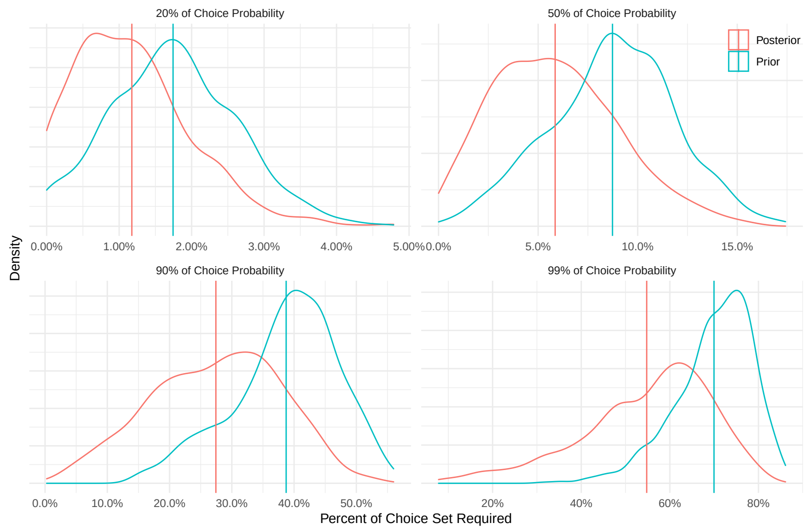



Figure 2-9: Kernel smoothed distributions of required shares of the overall choice set to generate aggregate choice probabilities given by each panel. In each, two distributions are drawn: that under the posterior type distribution (red) and the population prior (teal). Solid vertical lines give means of each distribution.

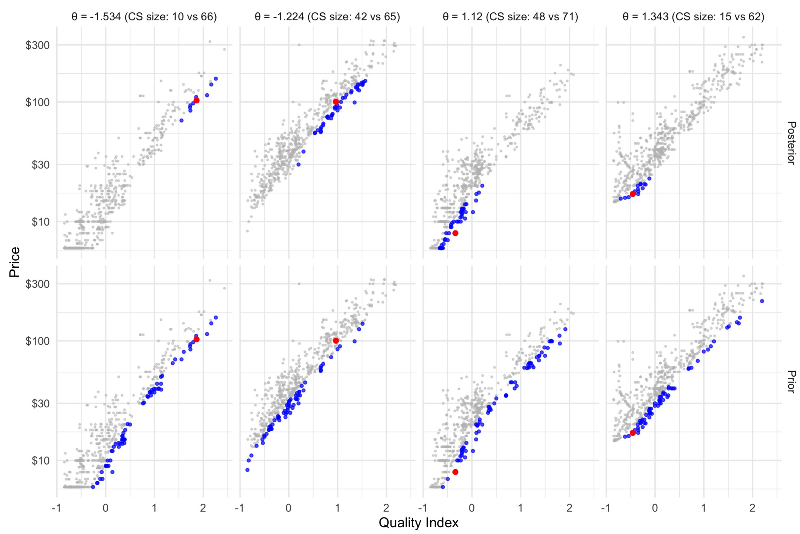



Figure 2-10: Quality space view of implied consideration sets for four sampled buyers (columns) and under estimated prior and posterior type distributions (rows). Blue points were in the effective consideration set – the set of listings cumulatively generating a 50% choice probability – while the red point was the observed choice; gray points are remaining available listings at the time of purchase. Horizontal axis gives quality index; vertical axis is log-scaled.

introduction of buyer heterogeneity at the transaction level but also localize both in quality space as well as spatially in the stadium. Figure 2-11 highlights the essential

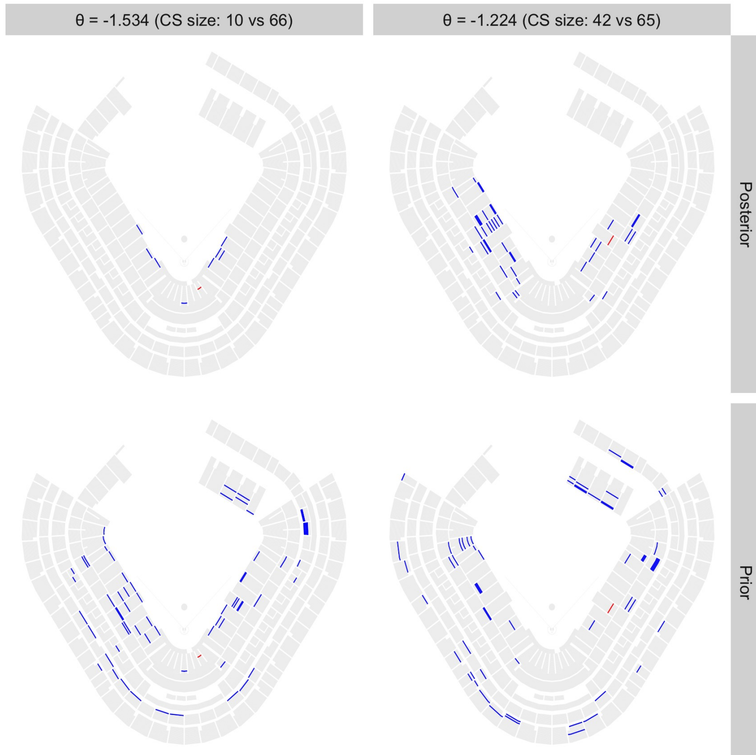



Figure 2-11: Spatial view of implied consideration sets for two sampled buyers (columns) under estimated prior and posterior type distributions (rows). Blue rows were in the effective consideration set – the set of listings cumulatively generating a 50% choice probability – while the red row was the observed choice.

role that buyer heterogeneity in the model plays in generating intuitive substitution patterns: the buyer who selected a seat in the first several rows of a premium section behind home plate (left column) is likely to consider inside substitution to other rows very near the field and home plate rather than seats in the right field bleachers (which would otherwise have been included in the restricted consideration set under the prior). Similar dynamics are evident in the right column but with a different seat locus. The examples in both figures are selected from a single game but describe consideration sets and choices for different buyers at different points in the event lifecycle. This is evident both from the overall change in listings and prices as well as the choice set under the prior distribution, which is identical in all panels and given the same choice set would generate the same effective consideration set. The coloring is consistent across both figures and the columns in Figure 2-11 correspond exactly to the first two columns of Figure 2-10.

To highlight the impact of changing choice sets on demand properties, panels (b) and (d) of Figure 2-8 plot the mean posterior type of each consumer against implied elasticities, using the posterior type distribution, for the chosen (panel (b)) and one random (panel (d)) element from each choice set. In both cases there is significant dispersion even conditional on a fixed posterior type $\bar{\theta}_i$. Especially in the case of the chosen alternative, the implied elasticity is not mechanically related to $\bar{\theta}_i$. Additionally, the dispersion of implied elasticities is substantially higher among the randomly selected listings than among the realized choices (note that the two vertical scales differ). This elasticity is heavily influenced by both the chosen alternative as well as the available choice set. Low elasticities typically accompany choices of listings that are particularly good deals in relatively thin markets while high elasticities typically accompany listings in thick market segments (e.g. among bleacher tickets when all prices are clustered tightly around $10).

## 2.5 Optimal Seller Policy

In this section we use the demand model to consider optimal pricing policies for small sellers. The central finding is that, due to highly localized demand (discussed in the previous section), the optimal selling policy can be closely approximated by a much simpler policy, even under limited information. Nevertheless, dynamic pricing, even subject to the simplifying constraints, remains highly valuable. For an individual seller, this simple policy amounts to pricing against a small set of similar listings with the pricing "aggressiveness" adjusted for the amount of time remaining until the items expire (proxying for the continuation value).

Explicitly, our focus in on a seller who is small in the sense that her pricing policy has no impact on either (i) any other seller's pricing decisions or (ii) the aggregate choice probabilities (i.e. we exclude any effects on the extensive margin). Given the observed thickness and high transaction velocity of the typical market in our setting, we do not believe that these assumptions are restrictive for casual sellers, though they would necessary exclude brokers. In this way we diverge from the established literature on optimal selling policies for large sellers or retailers (e.g. Sweeting (2012); Board and Skrzypacz (2016)) our focus is of particular relevance for peer-to-peer marketplaces. We begin in Section 2.5.1 by studying the seller's dynamic problem and studying implied properties such as the evolution of continuation values. We consider in Section 2.5.2 the value of information of flexibility in selling policy for these types of sellers. Later, in Section 2.5.3, we build an "optimal learner" trained to approximate optimal behavior given information that human sellers are able to observe in practice. We benchmark empirical human performance against this standard and conclude with a discussion of observed learning among sellers who repeatedly interact with the platform.

### 2.5.1 Seller's Problem

In the most general form, a seller holding a single listing solves a dynamic program of the following form:

$$V_{it} = \max_{p_{it}} \mathbb{P}_{it}(p_{it})p_{it} + (1 - \mathbb{P}_{it})\beta\mathbb{E}_t[V_{it+1}] \tag{2.5.1}$$

where $\beta$ is a per-period discount factor, $\mathbb{P}(\cdot)$ represents the probability of selling the ticket in the current period, and $V_{iT+1} = \bar{V}_i$ is some seller-specific terminal value where $T$ is the final period of each market. By taking a first order condition of the Bellman equation, a necessary condition for optimality is

$$p^*_{it} = \sigma_{it}(p^*_{it})^{-1} \left[ 1 - \beta \frac{1 - \mathbb{P}^*_{it}}{\mathbb{P}^*_{it}} \left. \frac{\partial \mathbb{E}_t[V_{it+1}]}{\partial p_{it}} \right|_{p^*_{it}} \right] + \beta\mathbb{E}_t[V_{it+1}] \tag{2.5.2}$$

subject to the terminal value $\bar{V}_i$, where $\sigma(p) \equiv -\frac{\partial \log \mathbb{P}}{\partial p}$ is the own price semielasticity. In discrete time, as long as $\partial\mathbb{E}_t[V_{it+1}]/\partial p_{it} \geq 0$, substituting back into the Bellman equation and iterating expectations generates $\mathbb{E}_\tau[V_{it}] \geq \beta\mathbb{E}_\tau[V_{it+1}] \ \forall \tau \in \{0, \ldots, t\}$. In the limit with no discounting ($\beta = 1$) the implication is that continuation values must be declining (Sweeting, 2012, Proposition 1). In continuous time, $\dot{V}(x(t), t) \geq 0 \ \forall t$, independent of discounting.[12]

Given observed prices and listings in all periods, we can compute the terms that enter the optimality condition directly. For simplicity, we assume that $\beta = 1$ and that individual listing choice probabilities are sufficiently small in relative terms that $\frac{1-\mathbb{P}^*_{it}}{\mathbb{P}^*_{it}} \frac{\partial \mathbb{E}_t[V_{it+1}]}{\partial p_{it}}\Big|_{p^*_{it}} \approx 0$, and we restrict attention to sellers with listings that transacted at least once. The latter restriction eliminates those sellers with limited attention or sophistication, or with sufficiently high reservation prices that they are highly unlikely to match with a buyer. However, we do not impose any restrictions on terminal conditions, so sellers may have some residual value for unsold tickets. Letting $\eta \equiv -\frac{\partial \log \mathbb{P}}{\partial \log p}$ denote the own price elasticity, (2.5.2) becomes

$$\mathbb{E}_t[V_{it+1}] = p_{it} \cdot \underbrace{\frac{\eta(p_{it}) - 1}{\eta(p_{it})}}_{\text{Adj. Markup Term}} \tag{2.5.3}$$

If consumers are only solving the static problem and do not account for the dynamics, $\mathbb{E}_t[V_{it+1}]$ should be constant for all periods and price should respond only to static demand conditions summarized by $\eta$: when demand becomes more elastic, price should decline, and vice versa. We find empirically that own price elasticities, evaluated at observed prices, tend to *decline* in later periods. Therefore, implied continuation values are declining at a faster rate than prices, which are themselves declining significantly (see Figure 2-5). We describe the evolution of these three quantities in Figure 2-12, which indicates that due to the elasticity decline, prices would need to increase by approximately 4% over the event lifecycle if consumers were optimizing and the continuation value were nondecreasing. Given the observed price trajectories, though, the implied continuation values declined by approximately 30% over the event lifecycle, roughly consistent with the findings of Sweeting (2012).

The operating assumption both in the preceding analysis and in Sweeting (2012) is that consumers properly dynamically optimize. The implication is that, for instance,

[12] To see this, consider the stochastic Hamilton-Jacobi-Bellman equation where the state variable $x(t)$ indicates whether or not the seller still holds the ticket at time $t$, and consider the limiting case where $p(t) \to \infty$ $\forall t$ (which provides the bound but is unlikely to be optimal).

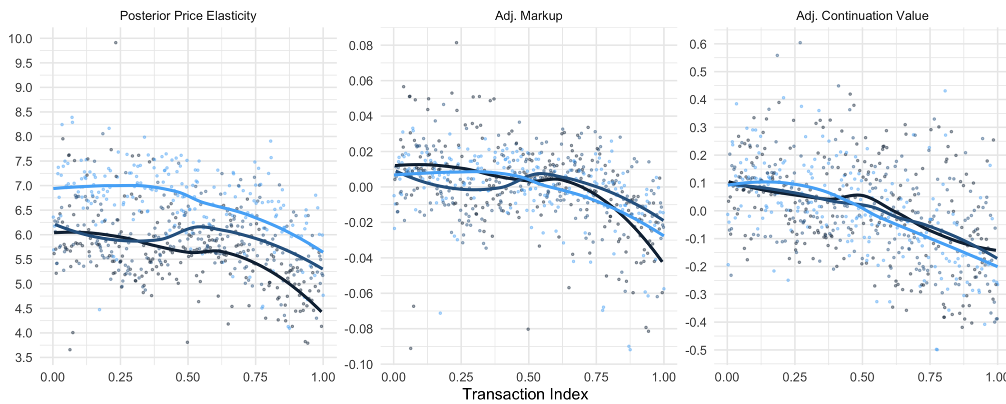



Figure 2-12: Evolution of price relevant quantities for a single market; horizontal axes index market periods $t$ from earliest to latest. The first panel describes the evolution in own price point elasticities at observed prices; the second describes the implied "adjusted markup" term from (2.5.3); the third gives the implied continuation value in each period. Dots represent individual observations, colored according to quality index (blue is highest quality; black lowest). Solid lines are LOESS smoothed analogs.

a consumer who listed a ticket at some arbitrary price for the event described by Figure 2-12 and never updated it must have had an opportunity cost of selling that exactly traced out the inverse of the middle panel (i.e. declined by approximately 4%). More likely, though, that seller simply forgot about or chose not to pay attention to her listing. The same optimization failure applies less pronounced examples as well.

As an alternative, we can instead solve backward for the seller's optimal policy in a given market using structural demand estimates to generate choice probabilities and subject to the assumption that the seller does not directly influence market aggregates, which allows us to use fixed empirical market configurations while perturbing the seller's policy. The seller is assumed to have all relevant information when optimizing, including a complete knowledge of buyer types in the current and all future periods as well as market configurations in the current and all future periods. We consider the impacts of relaxing these assumptions in the following subsections. Additionally, we assume that the seller is a revenue maximizer such that an unsold ticket has no residual value; without the assumption that sellers are properly optimizing, we cannot compute this quantity. Under this structure we can decompose the determinants of aggregate price decline. One factor is the purely dynamic one: the continuation value (opportunity cost) is declining over time as the good nears expiration. The other is a response to prevailing market conditions: sellers respond to observed price declines in similar listings and in aggregate market participants expect prices to decline, leading to fire sale dynamics in later periods and (if this effect dominates) likely admitting a multiplicity of equilibria. Many online marketplaces, the Platform included, offer seller tools for approximate "price matching", facilitating this latter behavior.

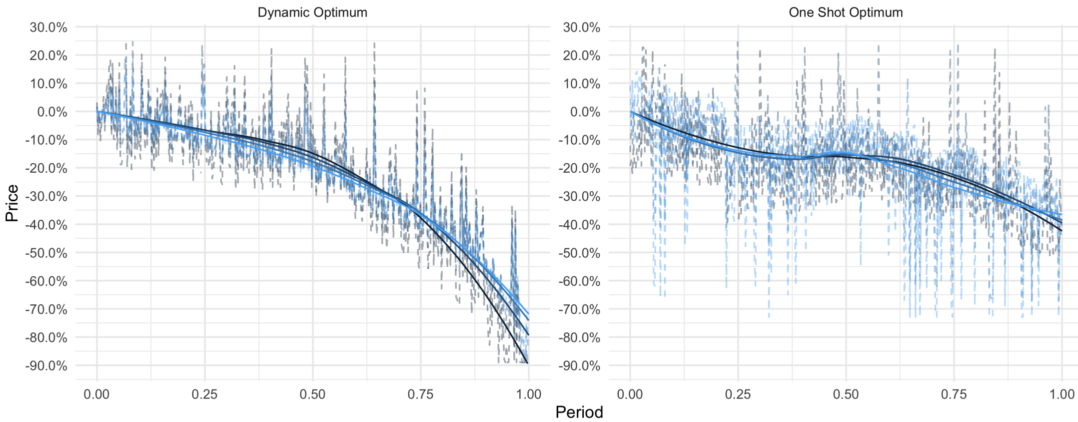
To perform the decomposition, we first compute the revenue optimal one shot (static) prices over the lifecycle of one sampled event, i.e. assuming that ever period were the final period. We then compute the dynamic optimum with the same terminal condition but solving backward so that continuation values in earlier periods are positive. The resulting optimal price paths as well as the own price elasticities evaluated along these paths are given in Figure 2-13. We find that pure price response does have a significant impact: ignoring dynamic concerns, the optimal price for the mean listing declines by approximately 40% between the first and the final periods.[13] Given that in the one shot case we are solving a static problem, the own price elasticity is tightly clustered around 1 in every period (bottom right panel).[14] However, the dynamics have a pronounced impact as well, and in combination the optimal price declines by over 80% over the event lifecycle. We recompute these price paths over a sample of eight events and find that the decomposition is relatively stable, with one shot prices declining by 52.7% of the total percentage decline including dynamic considerations.[15]

A final regularity among optimal price paths across all eight sampled events is that the period-to-period variability in optimal prices is much higher in the fully dynamic case. This is evident from the deviations between the dashed and solid lines

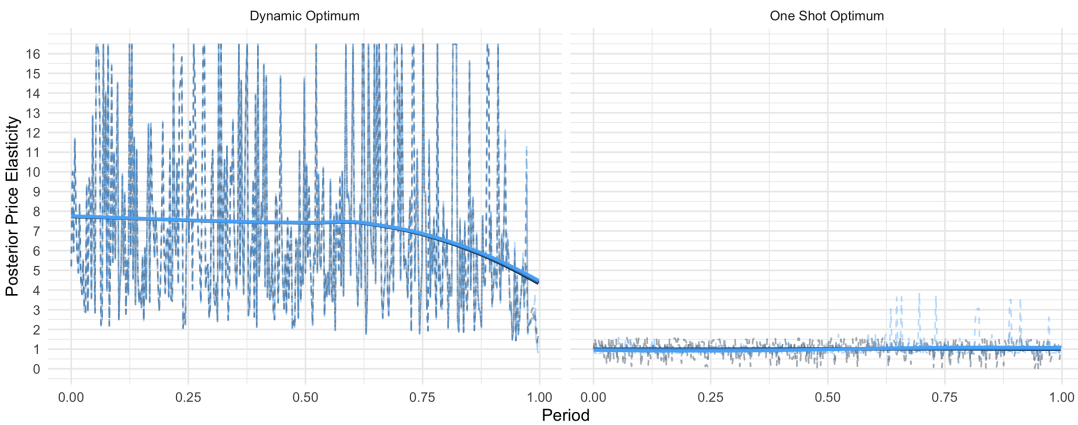
[13]This finding does not contradict the previous finding that own price elasticities decline in later periods. The former were point elasticities evaluated at observed prices; if sellers are not optimizing, these will differ from the own price elasticities evaluated at the optimal prices described by Figure 2-13.

[14]Slight variability is due to numerical imprecision when the choice probabilities are very low, which is usually caused when the "wrong" type of buyer shows up, e.g. when a very elastic buyer arrives but the seller holds a high quality ticket.

[15] Optimal prices in the two cases exactly coincide in the final period, therefore the absolute dollar price declines are significantly smaller in the one shot versus dynamic case, as in the former case the early period prices would be markedly lower.




(a) Price Paths




(b) Elasticity Paths

Figure 2-13: Optimal price paths for one event (top panels) and own price elasticities, evaluated at optimal prices (bottom panels). Horizontal axis indexes periods (early to late) and vertical axis gives relative prices (top panels) or own price elasticities (bottom panels). The left panels describe optimal relative price paths in the dynamic setting; the right panels describe the same price paths when every period is treated statically. Dashed lines describe quantities evaluated in individual periods; solid lines are the LOESS smoothed equivalents. Paths are drawn from four sampled row qualities, colored from highest (light blue) to lowest (black).

in the two panels of Figure 2-13. Closer examination reveals that in the dynamic case, optimal dynamic prices tend to be high when a consumer of the wrong "type" for the ticket arrives. That is, if a low elasticity buyer arrives and the seller is holding a low quality, cheap listing, she should increase the price: if the listing sells in the current period, the buyer is likely to have drawn a very high idiosyncratic shock for the seller's listing, and if it does not (which is highly likely), the seller reasonably expects to match with a more appropriate buyer type for her listing in some future period. By contrast, in the one shot optimum under the same information regime, the seller optimally reacts much less strongly to demand conditions. The result is analogous to the theoretical result that a seller with two listings of different qualities optimally prices them equally in a one period setting – the seller responds to buyer type to the extent that it influences her own price elasticity but has no additional strategic targeting consideration.[16]

### 2.5.2 Flexibility, Information, and Revenue

The preceding analysis established that a seller's revenue optimal price path is significantly declining as the event approaches, even ignoring dynamic considerations. However, that analysis assumed that sellers had complete information about all payoff relevant primitives in the current and all future periods. In this subsection we relax that assumption and consider the impacts on pricing and revenue. The central question is whether, on balance, a seller would do better to have full information about demand (i.e. every buyer's type and the full demand system) or to have full information and flexibility to price dynamically in response to other supply. We find that the latter dominates and that while the former has a significant influence on prices in any given period, it has a subdued effect on expected revenue. The finding is consistent with the view that markets are highly local: knowing the buyer's type may have a strong influence on the policy (pricing) response, but the buyer is only likely to consider a small set of relevant listings, so this price response has a minimal effect on realized outcomes.

In light of this question, we consider three regimes that differ along information and policy dimensions. Along the information dimension, sellers may either have perfect knowledge of every arriving buyer's type ("full" or "posterior" information) or may only know population shares ("limited" / "prior"), i.e. they may either perfectly

[16] To see this potentially counterintuitive result, consider two goods $i$ and $j$ with sale probabilities $q_i(p_i, p_j)$ and a seller with zero opportunity or other marginal cost. Then parameterize the problem as $\max_{p,\delta} p \cdot q_1 + (p + \delta) \cdot q_2$ then take FOCs, establishing that $\delta^* = 0$.

observe time varying demand or may only know the properties of mean demand. Along the policy dimension, sellers may either freely update every period ("flexible" or "dynamic" pricing) — as before, a period is associated with a single buyer arriving and competing *some* transaction — or may be constrained to chose a single price to maintain throughout the event lifecycle ("fixed" / "static"). In all cases we study a hypothetical seller who entered the market at or before the time that the first transaction were realized.

To clarify the following discussion, we define three specific information-policy regimes as well as

(FIDP) Full Information, Dynamic Price: Before each transaction takes place, each seller is notified of the posterior type distribution of the arriving buyer *as well as of all future buyers* and may update her price.

(LIDP) Limited Information, Dynamic Price: Sellers may update prices before every transaction but only know the population type distribution and have no information about specific buyers.

(FISP) Full Information, Static Price: Before any selling begins, the seller is told the posterior type distribution of each successive arriving buyer and must choose a static price for the whole event lifecycle.

Under each regime, we compute the optimal price path iterating backward from the final period. We continue to assume that the seller is sufficiently small that market configuration is independent of her pricing policy, hence we use empirically observed market configurations in every period to compute optima. This approach simplifies the computation because we can keep track of continuation values and evaluate the Bellman equation directly. By precomputing posterior type distributions and implied inclusive values we are able to vectorize the evaluation over a grid of relative prices and quality indices. As before, we select a random sample of eight events from the 2018 season on which to perform this computation and restrict attention to the market for pairs of tickets, as this is the single largest market in our data. Therefore our hypothetical seller has a pair of tickets to sell and enters the market before the first transaction is realized. Unless otherwise specified, we assume that the seller obtains no residual value after the final period for a ticket that goes unsold.

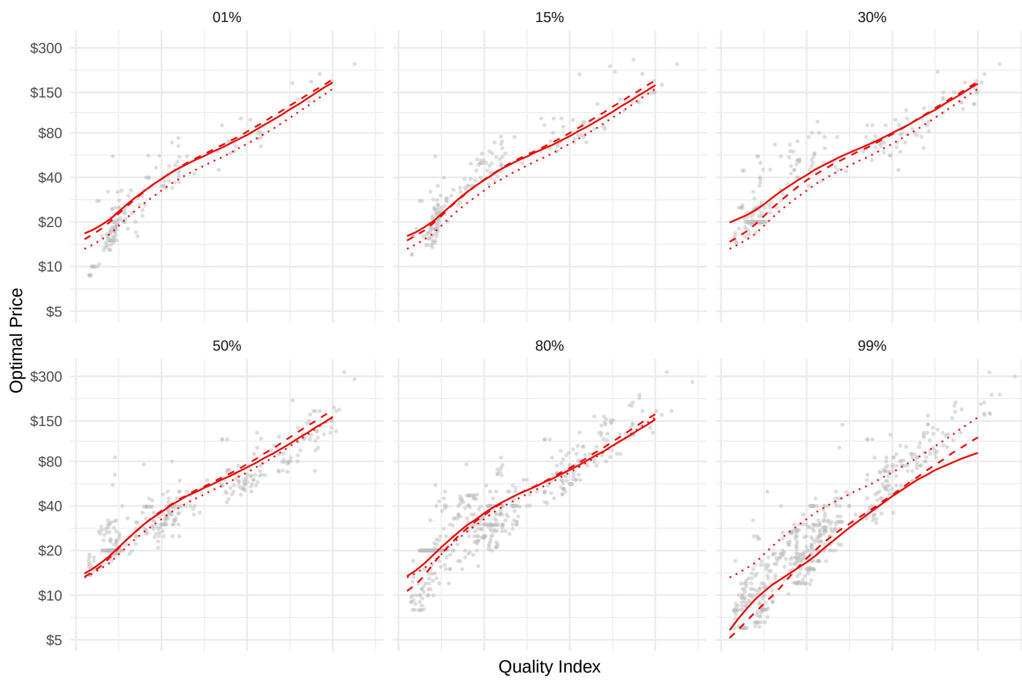



Figure 2-14: Temporal evolution of one market. Horizontal index quantifies the relative quality of each good; vertical index is price. Each panel is a different selling period. Panel titles give the approximate period as a percentage of the overall event life cycle (i.e. 50% is the middle period and 99% is nearly the final period). Gray dots are empirical listings in each period. Red lines are computed revenue optimal prices across the quality continuum in each period: (i) the solid line is the full information benchmark (FIDP), (ii) the dashed line is the limited information benchmark (LIDP), and (iii) the dotted line is the fixed price full information benchmark (FISP), when the seller must select a single price at the outset of the market.

Consistent with intuition, the optimal static price tends to be lower than the optimal dynamic price in early periods but in later periods dynamic prices tend to be significantly lower. Figure 2-14 shows the evolution of optimal prices for a single event. The panels show sequential periods during the event lifecycle and depict several general facts about the market. First, listing prices tend to decline toward the time of the event and more listings tend to appear as time progresses, net of sales. These factors tend to pressure optimal price down even absent any dynamic considerations by increasing own price elasticity. Second, optimal prices under the three regimes

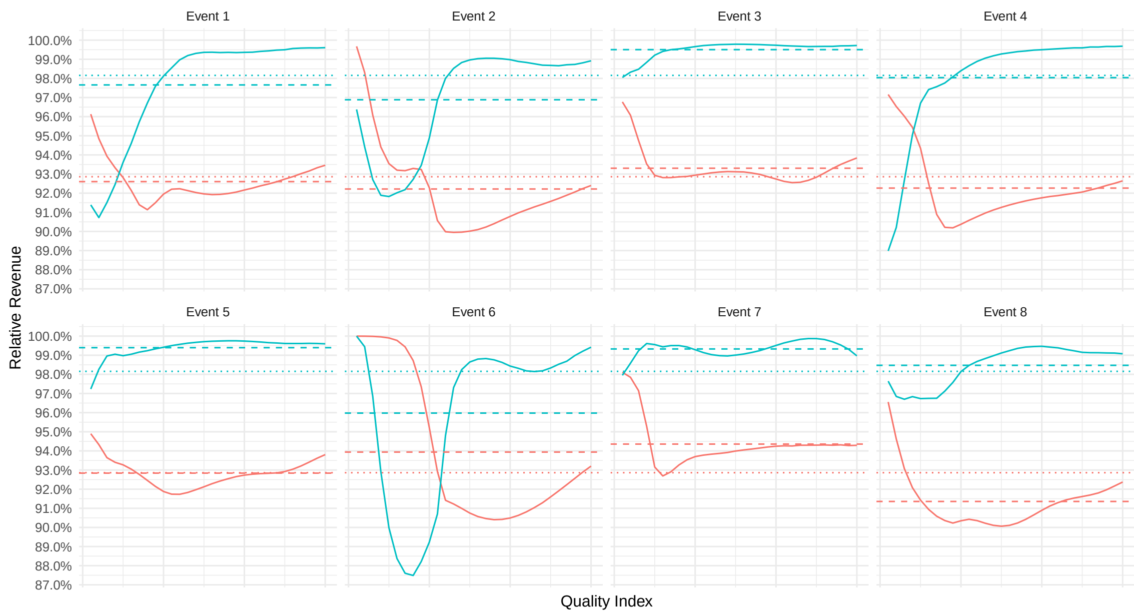



Figure 2-15: Maximized (optimal) revenue based upon two information and policy regimes for eight sampled events from 2018. Vertical axis gives the percent of expected revenue each policy generated relative to the full information, dynamic price optimum (FIDP); horizontal axis gives relative row quality. Teal lines describe limited information but dynamic price paths (LIDP). Red lines describe full information but static optima (FISP). Dashed lines are means within each event; dotted lines represent means aggregated across all events.

coincide closely during many periods and tend only to significantly diverge in late periods. Finally, both absolute price policy and relative policy (across the three regimes) exhibits significant heterogeneity across the quality spectrum.

For each of the eight events in our sample, and under each pricing regime, we compute the expected revenue for a seller holding a listing of a given quality before the first transaction takes place. Relative to the full information, dynamic pricing benchmark (FIDP), the mean seller dynamic pricing under limited information (LIDP) loses only 1.8% of expected revenue, while the mean seller restricted to price statically but with full information (FISP) loses 7.1% of expected revenue. There is substantial heterogeneity across ticket quality index and event, though, as depicted in Figure 2-15. Sellers of the lowest quality tickets benefited most from knowledge of posterior buyer types even when constrained to price statically, obtaining 97.4% of benchmark revenue versus 95.8% for sellers of the same tickets operating under

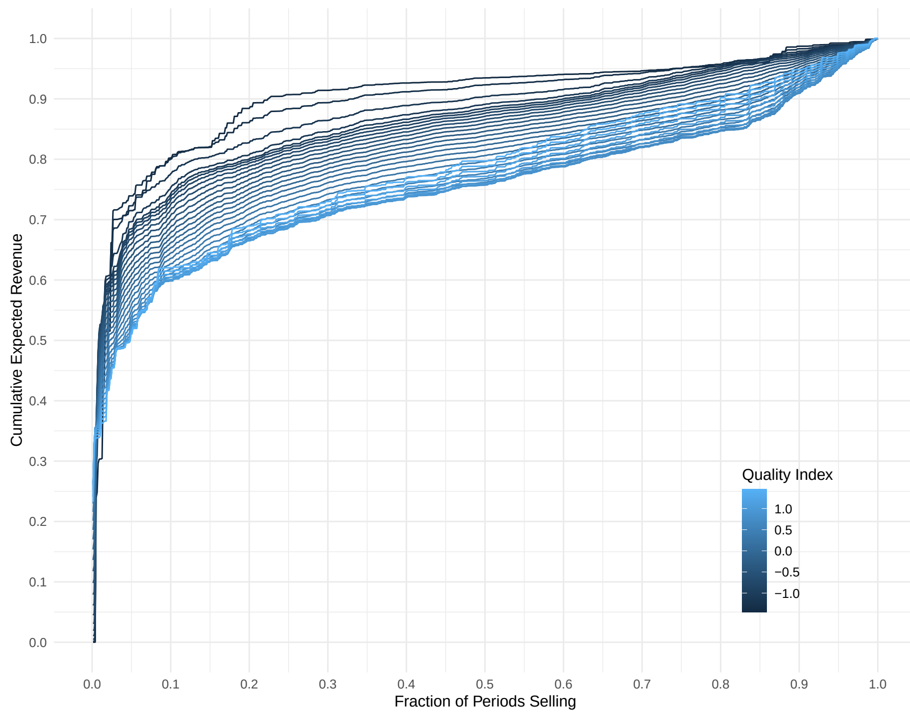



Figure 2-16: Fraction of expected revenue conditional on revenue optimal strategy, based upon number of periods selling for one representative market. Periods are counted from the final period, so a fraction of 0.1 corresponds to selling only in the final 10% of the event lifecycle. Curves are colored by quality index, from highest quality (blue) to lowest quality (black).

limited information but free to update price in every period.

Ex ante it is unclear whether a seller will optimally earn most expected revenue in early or late periods. On one hand, reducing price in early periods when other listing prices are typically high could be advantageous rather than remaining in the market (with higher probability) when prices decline. On the other, in these periods a seller derives significant option value from her ability cut price in later periods, increasing the value of keeping price high and remaining in the market with higher probability. Empirically, we find the latter effect to strongly dominate. Sellers of any quality listing optimally obtain over half of their revenue from the final 10% of periods in expectation. For high quality inventory this pattern was stable across games. Low quality inventory was subject to substantial variability, though, driven primarily by overall event quality as indicated total gross sales. For top events (high overall sales), low quality inventory generated disproportionate expected revenue in late periods —

in some events in our sample sellers of the lowest quality inventory obtain over 80% of expected revenue from the final 10% of periods and nearly 70% from the final 5% of periods while in others they obtain under 10% of expected revenue in the final 10% of periods (and 50% of revenue in just under 30% of periods).

### 2.5.3 Seller Learning

While they are useful to establish benchmarks of the various selling regimes, the assumptions maintained in the previous subsection are not plausible in practice. In this section we relax the assumptions and approximate how an optimizing seller might learn under additional information restrictions and use this to benchmark observed seller behavior. Additionally, we document that sellers improve their pricing policies with experience.

We develop a simple model for how a revenue maximizing seller might learn to price her tickets by observing past markets, subject to a severely limited information set. The resulting policy is qualitatively a simple one, and amounts essentially to pricing against similar listings with an adjustment for the time remaining before the event. Strikingly, the learned policy obtains expected revenue very near the benchmark established in the previous section, despite being far more rudimentary and subject to strong information constraints. Taken in the combination, and restricted to partial equilibrium, the evidence suggests that due to the implicit localization of markets, a pricing policy of approximately matching the most competitively priced similar listings nearly obtains the full theoretical optimum and simple tools may enable small consumer sellers to complete effectively even with highly sophisticated agents.

**Approximating a Learner** With an eye toward realism, we assume that sellers can learn optimal selling policies only from observing historical markets. For simplicity, we restrict focus to a seller with a pair of tickets slightly above median quality, and train our "learner" on a sample of twenty events from 2017. We maintain the same eight events from 2018 as a holdout sample against which to test the learned policy.

At any time period $t$, the seller observes the following market information: (i) the total number of available listings matching her quantity, (ii) cardinal indices of the two quantiles of currently available listings by utility to the mean buyer ($50^{th}$ and $90^{th}$), (iii) the prior two sets of features but restricted locally (to listings within 20% of relative value in quality space), and (iv) cardinal indices of the mean and $90^{th}$ percentile utility to the mean buyer among the past 20 transactions, (v) the time until the event starts (and selling ceases) so in total the seller maintains a nine element state vector.[17] The Platform makes both listing and recent transaction easily accessible to sellers, including detailed information about transaction dates and physical listing location, so all of this information is plausibly in the seller's information set. Given this state vector $\vec{s}$, the seller must learn a pricing policy $g : \vec{s} \to \mathbb{R}_+$.

We again employ the gradient boosted tree framework `xgboost` to train our optimal learner. The learner obtains a design matrix with rows describing the state vector $\vec{s}_t$ in each period $t$. The learning objective is total expected revenue given the policy $g(\cdot)$, which is to be learned. The primary difficulty in programming the learner is that updating the price in period $t$ affects the continuation value in all periods $t-k$ for $k \geq 1$, complicating gradient computation. Matters are simplified by the fact that the tree learner only requires the gradient and Hessian to be computed with respect to *each prediction* (output from the tree), in this cased a price $p_t^*$, holding all other predictions fixed. Consider again the seller's problem summarized by the Bellman equation (2.5.1), where $V_t$ represents the value function in period $t$ and $\mathbb{P}_t(p_t)$ represent the probability of sale in period $t$ conditional on empirical market configuration and price $p_t$. The derivative of the value function with respect to price in an arbitrary period is

$$\frac{\partial V_t}{\partial p_t} = \frac{\partial \mathbb{P}_t}{\partial p_t} \left[1 - \mathbb{E}_t[V_{t+1}]\right] + \mathbb{P}_t$$

[17] We experimented with expanding this state vector, e.g. adding additional quantiles or adjusting the locality restriction, and found minimal impacts on overall revenue – under 0.5% in expectation in most cases. Conditional on having the time to event variable and the local price and listing quantity variables, in sample performance was uniformly within 1% (in revenue terms) of all expanded specifications.

and the impact on the $k$-period-ago value $V_{t-k}$ due to a change in period $t$ is

$$\frac{\partial V_{t-k}}{\partial p_t} = \frac{\partial V_t}{\partial p_t} \prod_{l=1}^{k} (1 - \mathbb{P}_{t-l})$$

However, using these expressions directly adds significant weight to certain observations, particularly those toward the end (but not too close to the end) of the selling period. The contribution of each period to overall expected revenue is the expected revenue in the period times the probability of reaching the period, so the unadjusted (baseline) total weight on period $t$ is

$$\underline{w}_t = \sum_{k \leq t} \underbrace{\left( \prod_{l=1}^{k-1} (1 - \mathbb{P}_l) \right)}_{Pr(\text{reach } k)} \underbrace{\left( \prod_{l=k}^{t} (1 - \mathbb{P}_l) \right)}_{Pr(\text{reach } t \mid \text{reaching } k)} = t \cdot \prod_{k=1}^{t} (1 - \mathbb{P}_k)$$

which is greatest at mid-to-late periods when $t$ is large but the likelihood of reaching is still moderately high.

When trained with these implicit weights, the learner focused on only those periods with high weights and reported a constant optimal price for others. Other periods had too little influence on the overall expected revenue for the greedy algorithm to learn their relevant policies. As a result, the learning algorithm often failed to converge to global optima and occasionally did not converge at all and failed even to monotonically improve overall expected revenue. In order to mitigate these issues, we reweight observations as if the seller entered in each period anew, so our overall objective is the weighted mean over periods $t$ of expected revenue conditional on entering in period $t$, weighted by the probability of ever reaching period $t$. Mechanically, let $\tilde{\mathbb{P}}_t \equiv \prod_{k=1}^{t-1} (1 - \mathbb{P}_k)$ be the probability of reaching period $t$ without first selling. Then our adjusted weighting vector is

$$w = \left( M \, \vec{1} \right) \odot \tilde{\mathbb{P}} \qquad \text{with} \qquad M \equiv \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \frac{\tilde{\mathbb{P}}_2}{\tilde{\mathbb{P}}_1} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \frac{\tilde{\mathbb{P}}_T}{\tilde{\mathbb{P}}_1} & \frac{\tilde{\mathbb{P}}_T}{\tilde{\mathbb{P}}_2} & \cdots & 1 \end{bmatrix}$$

where $\vec{1}$ represents the column vector of ones and $\odot$ denotes elementwise multiplication (the Hadamard product). If the probability of sale in every period were $\alpha$, the relative weights would roughly satisfy $\frac{w_t}{\underline{w}_t} \approx \frac{1}{(1-\alpha)t}$, achieving our desired weighting. We confirm that this weighting recovers the desired optimal price path in simple Monte Carlo trials.

**Trained Learner Results** We train the boosted tree over 500 epochs, constraining each tree's depth to be at most three and adding mild $L_2$ regularization. Additionally, we restrict price updates to be at most 5% per epoch and initialize all prices such that the probability of sale in each period is bounded away from $\{0, 1\}$, which improves convergence properties. The boosted tree framework learns optimal policy exceptionally well despite being restricted to a relatively limited information set. Figure 2-17 depicts the learner's policy relative to the full information optimal policy for a selected event. For this particular event, the learner obtains 97.9% of the expected revenue obtained by the full information optimum. Aggregated over four events that were held out of the training set but that were selected from the same season of the same team, the mean learner obtains 97.3% of the optimal revenue. Qualitatively, the learner seems to be (i) positioning its listing near the lower envelope of listings in quality-price space (cf. Figure 2-4) and (ii) learning the approximately optimal degree by which to undercut as a function of time to event (as it does not know precisely how many transactions remain or when the final transaction will take place). The results suggest that, at least in contextual sample and for a seller with no residual value for her tickets, a tractable rule-based system can attain high performance thresholds. In light of the above observations, though, results decidedly do not generalize to large sellers whose pricing activities are likely to have an impact on the policies of other sellers as well as extensive margin substitution.

**Human Seller Learning** Given the ease with which a machine is able to learn optimal selling policy, a natural question is whether humans can replicate the success. Specifically, we would like to empirically test whether humans learn optimal policy

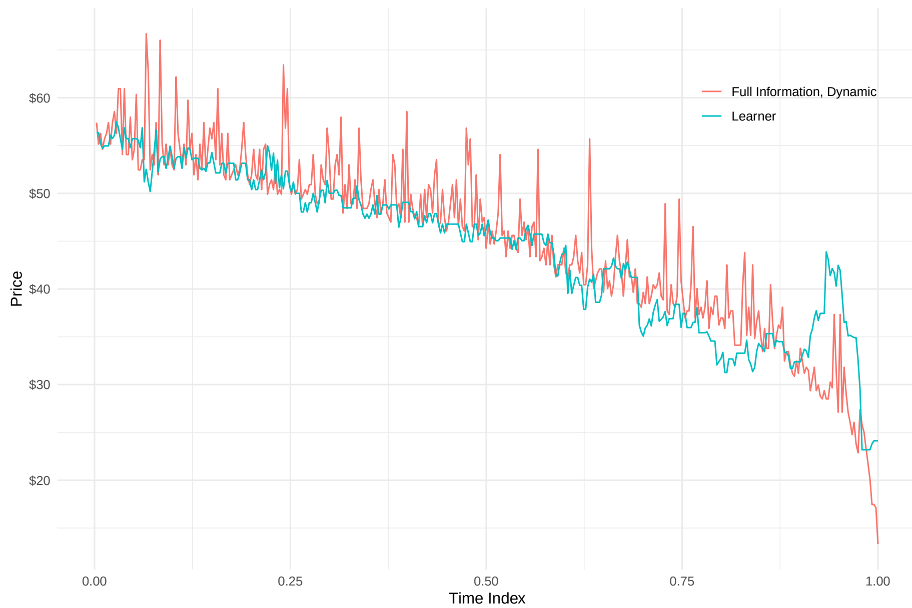



Figure 2-17: Optimal price paths for a pair of tickets in the $77^{th}$ percentile of the row quality index for a sampled event. The red line gives the optimal full information policy described in Section 2.5.2 while the teal path gives the path generated by the learner described in Section 2.5.3.

over time. The test is complicated by the fact that certain sellers price their tickets unreasonably high — so high that there is effectively no chance of the tickets ever selling. This is similar to the observation made by Dinerstein et al. (2018) and is a perplexing one.

For the sample of holdout events we predict the optimal policy as learned by the `xgboost` framework, then for every seller and every period in which the seller had a listing available for sale we compute the relative listing price as a percentage of the (constrained) optimal price. For each seller we additionally compute, at each time period, the number of previous listings that the seller has created (whether they sold or not) for the team in question. We ignore listings created for events of other teams or performers. Additionally, we exclude all sellers with over 200 previous listings in total or 81 per observed season selling (one per event, per season), as these are likely to be brokers. Based on this sample, which includes a total of 762 distinct sellers and total of 1.16M listing-periods, regressing the log price deviation on a count of previous seller listings (winsorized at 50 listings) gives a slope and intercept of $-0.0043$ (0.0015) and 0.4575 (0.0420), respectively (standard errors in parentheses). The estimates suggest that the typical seller prices her tickets approximately 45% above the revenue maximizing level (under the information and learning constraints maintained in this subsection), but that with each additional market interaction (in the form of a listing) she moves toward the revenue optimal pricing policy by approximate one percent.

This evidence should be taken to be suggestive and not indicative of the precise mechanism and magnitude of learning, as the benchmark is against a hypothetical seller with no residual value for her tickets, while in practice some substantial fraction of sellers are likely to use their tickets if unsold. Additionally, sellers who repeatedly list may endogenously favor selling, so the effect may partly capture differences in objectives rather than differences in execution. Taken in combination, however, the results suggest not only that sellers can learn to sell effectively but that they do.

## 2.6 Conclusion

In this paper we developed a model of demand for the inside good in a market with highly vertically differentiated goods. The model accommodated choice sets that changed with every buyer and allowed buyer type heterogeneity based upon realized behavior. We then estimated the model using data from a large online exchange for live event tickets.

The empirical results confirm that explicitly incorporating buyer type information significantly reduces each buyer's *effective* consideration set. In the context of large online marketplaces with a multitude of vertically differentiated goods, it is common for buyers only to consider a small subset. Including all available items in the choice set tends to misstate own price elasticities, as we confirm empirically. Our approach may equally be applied when the econometrician directly observes the consumer's choice set (e.g. with clickstream data indicating precisely the items viewed) as well as when this information is not known. Even when the data are available, e.g. for an online furniture retailer, the items presented on the screen may not all enter the buyer's choice set equally.

Based upon the structural estimates, we describe optimal selling behavior under various information and policy regimes and quantify the value of additional information relative to additional flexibility. In partial equilibrium, empirical results indicate that, in large part due to high localization of buyers' consideration sets, simple selling strategies suffice to generate expected revenue very nearly as high as that which could be obtained by more sophisticated strategies and with more information, though we do not investigate general equilibria (which may not be unique) in which all sellers adhere to these strategies. To demonstrate that such a policy is practicable, we train a prediction model to approximate an optimal selling policy under information sets empirically available to sellers. Though these information sets are limited, the approach nearly matches the full information benchmark in expected revenue. Using this learned policy as a benchmark for human sellers, we find that typical sellers price higher than the revenue optimal level but that each additional selling experience (attempting to sell one ticket, whether or not it transacted) is associated with a reduction in the degree of mispricing.

We believe that this area is rich for future work. Our approach to benchmarking learning is only cursory and rich data on seller behavior should be able to shed additional light not only on how sellers learn how who in particular learns and what might influence these sellers. Additionally, our estimated own price elasticities were quite high, consistent with the literature. If at least some fraction of small sellers have little to no residual valuation for a good that does not sell, it is perplexing that consumer sellers do not very aggressively undercut. Future work quantifying these elasticities and exploring institutional drivers across industries would help inform platform design and improve policy to support small sellers in the sharing economy.

# Chapter 3

# Capital Structure, Investment, and Competition: Evidence from Loan Covenant Violations

Creditors, facing an agency problem when lending to firms, often include financial covenants in debt contracts. Upon covenant violation, creditors assume some control rights and generally "tighten the reins." Resulting contractions in capital investment and R&D spending are well-documented. We investigate downstream effects on product-market competition and using a regression discontinuity design document that the transfer of control rights makes firms tough. The results shed light on the impact of capital structure and control rights on product markets.

## 3.1 Introduction

Especially in times of acute stress, a firm's capitalization structure can have profound impacts on it's short term operations. Loan covenants — contractual provisions designed to protect lenders when the firm's condition worsens — are disproportionately triggered when broader economic conditions deteriorate. Upon technical default, when covenant thresholds are crossed, creditors obtain control rights and exert influence over firm activities, most typically reducing short term expenditures (Chava and Roberts (2008); Roberts and Sufi (2009); Falato and Liang (2016)). During the pandemic, firms and their lenders temporarily waived covenant provisions in order to avoid these knock-on economic impacts associated with technical default (Bakewell and Lee (2020)).

In this work, we investigate the contemporaneous impact of covenant defaults on product market conduct. The related literature focuses on dynamic firm decisions that affect overall firm state, such as investment, employment, and R&D, whereas our principal focus is on pricing and static product market competition. To our knowledge this work is the first to investigate this dimension of technical default.

In practice, control rights operate along a continuum, with bargaining power shifting from borrower to lender as a firm's capacity to finance its operations and debt becomes more tenuous. These interactions take place in a repeated game, and for highly leveraged firms the creditors' threat of withholding future funding can influence firms' management decisions. This continuity, especially in the presence of correlated shocks at the industry level, complicates empirical work. Short-term debt needs often coincide with demand contractions or other contemporaneous effects, so firm performance and financial structure are typically endogenous. A good instrument would therefore be one that is discontinuous at some threshold. This solves the endogeneity problem by introducing a regression discontinuity estimator, since the structure of the error terms would have to be highly unusual with locally arbitrary mass points for correlated errors to remain an issue.

Our empirical strategy exploits this discontinuity in the context of large commercial loans – debt issued directly to businesses by banks. These loans are typically *syndicated* between multiple banks or other financial entities/instruments (e.g. hedge funds, CLOs) with one or several *lead arrangers* (almost always banks) which are generally both underwriters and primary monitors. In the US, total outstanding commercial loan commitments stand at about \$2.5-3 trillion whereas total outstanding corporate bonds stands at about \$8.5 trillion.[1]

[1] For a current tally, refer to the Federal Financial Institutions Examination Council (FFIEC) Call Reports, available at `https://cdr.ffiec.gov/public/`.

At their core, these loans are simply contractual agreements often negotiated bilaterally, hence there is wide scope for varied and specific provisions. They are typically the most senior obligations of the firm and often include *covenants*, which are pairs of well defined events and consequences that are written into the loan contract. Two types of covenants prevail:

- *Nonfinancial (affirmative) covenants* require the borrower to, e.g., comply with laws and make certain disclosures; failure to adhere is breach of contract.
- *Financial (negative) covenants* define specific financial thresholds; if they are breached, the borrower is in *technical default* and the lender is allowed to call or accelerate the loan. Often lenders will instead renegotiate.

An example of a covenant follows, taken from Lifetime Brands' Oct. 31, 2006 SEC filing:

> **Section 7.13 Interest Coverage Ratio** The Borrower shall not permit the Interest Coverage Ratio as of the last of any fiscal quarter to be less than 4.00 to 1.00.
> **Section 8.02 Contract Remedies** [...] in the case of an Event of Default [...] the Loans, all accrued and unpaid interest thereon and all other amounts owing under the Loan Documents shall immediately become due and payable

We will exploit these financial covenants to instrument to control rights transfers. Chava and Roberts (2008) give a synopsis:

> Violations of financial covenants are often referred to as "technical defaults," which correspond to the violation of any covenant other than one requiring the payment of interest or principal. Upon breaching a covenant, control rights shift to the creditor, who can use the threat of accelerating the loan to choose her most preferred course of action or to extract concessions from the borrower to choose the borrower's most preferred course of action.

The overall narrative of this paper is that, upon covenant default, senior creditors maximize expected cash flows conditional on debt service default (as their payoff is constant if the firm is able to service the debt) while equity holders maximize expected discounted cash flows conditional on solvency (as they enjoy limited liability). We develop a simple, tractable model to formally describe these dynamics, and confirm that the general model predictions are uniformly consistent with the empirical evidence. Therefore, the impacts of adverse economic conditions in a given industry are likely to be amplified when a large share of constituent firms have binding loan covenants. While we do not specifically focus on leveraged buyouts (LBOs), our work may contribute to an understanding on their product market impacts, as it suggests that the high and restrictive debt loads that accompany many buyouts tend to reduce markups, contradicting the findings of Chevalier (1995).

Our work spans several literatures with the unifying focus on debt and its impact on industrial conduct. One of the largest and first to develop studies the role of debt in mediating agency problems and informational asymmetries between firm management and other stakeholders. Among the earlier work was Ross (1977), who studied the debt as a signaling mechanism for quality. Further work on informational asymmetries expanded the scope to study capital structures more broadly (Leland and Pyle (1977); Myers and Majluf (1984)). Aghion and Bolton (1992) investigate the implications for optimal financial contracting with attention toward control rights. The later literature analyzes explicitly the effect of debt on directing a firm's efforts and investment. Stulz (1990) identifies the fundamental tradeoff of a higher debt load as between a more disciplined approach to investment and the possibility therefore of missing out on good projects. Chakraborty and Ewens (2012) challenges this conclusion empirically, finding that debt-financed startups exhibit the opposite effect, and are much less likely to be successful than their equity-funded counterparts.

From a strategic standpoint, debt structure may influence management's investment decisions, and can add discipline where necessary (Jensen and Meckling (1976); Jensen (1986)). Brander and Lewis (1986) strategic impacts in product markets and find that debt may act as a commitment device to ensure that firms are tough in product markets. Chevalier (1995) provides an empirical test of these results, using a wave of LBOs in the supermarket industry to evaluate whether highly indebted firms become tougher or softer, as well as knock-on impacts at the local industry level. In contrast with typical theoretical predictions, she found that LBOs softened competitions and induced positive pressure on prices. Our results mirror, albeit imperfectly, and challenge this empirical finding.

A separate literature has emerged in recent years with an institutional focus on loan covenants. Across a wide cross section of firms, a transfer of control rights from debtor to creditor has persistent, negative effects on investment and similar activities for afflicted firms (Chava and Roberts (2008); Roberts and Sufi (2009); Falato and Liang (2016)). On the other hand, there is little to no effect on successful research and development for a similar cross section of firms despite reduced R&D spending, suggesting judiciousness on the part of the creditor (Chava et al. (2016)). Work emphasizing the impacts of control rights transfers on agency costs arrives at similar conclusions in support of the Jensen view (Bharath and Hertzel (2019); Nini et al. (2012, 2009)). Ersahin et al. (2020) use firm microdata to investigate potential mechanisms and find that firm activities in response to covenant violations are judicious and highly targeted toward non-core and low return business activities.

Our work merges and extends these literatures with a focus on product market conduct. In this way we fill a gap in the existing analysis by describing likely contemporaneous market impacts due to technical defaults in an industry. Additionally, we contribute focus on the impacts at the industry level, to firms not directly affected by the covenant default. From a regulatory standpoint it is important to understand the impacts of large and restrictive debt loads on end consumers at the industry level. We believe that this work provides a starting point for that analysis.

The remainder of the paper is structured as follows. Section 3.2 develops a simple model and generates predictions for both static and dynamic effects of control rights transfers. Section 3.3 gives an overview of the data and final dataset construction, as well as some summary statistics to motivate the empirical work. The empirical work is partitioned by dynamic and static effects and again by impacts directly on violating firms versus impacts on firms operating in industries with one or more violators. We focus first on dynamic effects, in Section 3.4, which provides estimates for the effect of a control rights transfer on the affected firm; Section 3.5 then performs the same exercise but for unafflicted firms in an industry with at least one violating firm. We study product market effects in Section 3.6, which requires us to estimate production functions. The section describes the estimation approach, production function estimates, and the impacts of covenant violations on markups. Section 3.7 summarizes and provides context for our results our results, then suggests extensions and discusses possible issues in our identification strategy. Section 3.8 concludes.

## 3.2 Theoretical Motivation

To motivate the empirical results, we develop a simple framework under which to consider the implications of credit shocks on product markets. We take motivation from a large literature that explores endogenous capital structure choices from a strategic perspective. However, our focus is instead on the likely impacts on both static and dynamic competition due to a shift in control rights, holding fixed some debt structure. We find that when creditors gain increased control rights they are likely to (i) become tougher and price more aggressively in the product market and (ii) reduce investment in future output.

### 3.2.1 General Framework

Bank loans and related credit facilities typically hold senior claims on a company's cash flows and assets. Their expected payoff is therefore maximized when the firm is highly likely to make its debt payments and minimally dependent on further profitability of the firm. Specifically, if a bank issues a debt instrument in the amount $D$ with per-period coupon $c$ and $\tau$ periods to maturity to a firm in state $s$, its value function is

$$V_b(s; D, \tau) = \max_i \int_{\pi(z;s)-i \geq c} [c + \beta V_b(s(z;i); D, \tau - 1)] \, dF(z) + \int_{\pi(z;s)-i<c} \pi(z;s) \, dF(z)$$

where $z$ is a stochastic component, $\pi$ is free cash flow, and $\beta$ is a per-period discount factor. We abuse notation and let $s(z;i)$ denote the next period state, though in reality $s$ is an argument to this function. The value function has terminal condition

$$V_b(s;D,0) = \int \min\{\pi(z;s) + \bar{D}(s(z)),\ D + c\}\ dF(z)$$

where $\bar{D}(s(z))$ is the maximum of the amount of liquidity a firm in state $s$ can obtain. If the firm can repay the debt obligations then the creditor receives $D + c$ terminally. If it cannot, the senior creditor has a claim on all cash flows in the final period plus whatever liquidity the firm can generate by any means. These expressions would grow more complex if the debt contracts lacked call protections, had a complex amortization schedule, or had accelerated repayment provisions. None of these changes have an effect on the qualitative results regarding the competing incentives of firms and their creditors. Firm managers have a fiduciary duty and often a direct financial incentive in maximizing firm value — the discounted present value of future dividend flows plus terminal liquidation value — and their liability is limited. Their value function is

$$V_f(s;D,\tau) = \max_i \int_{\pi(z;s)-i\geq c} [\pi(z;s) - c + \beta V_f(s(z;i);D,\tau-1)]\ dF(z)$$

where $i$ is investment and with terminal/rollover debt condition

$$V_f(s;D,0) = \max_{i,D',M'} \int \max\{0,\ \pi(z;s) - c - D + \beta V_f(s(z;i);D',M')\}\ dF(z)$$

where $M$ is the maturity at issue.

### 3.2.2 Static Incentives

The above framework is general, at least when restricting debt contracts to coupon/interest-only and limiting contractual provisions. It is also dynamic, which we will investigate next. However, even in a static setting the creditors have incentives that diverge from those of shareholders.

Following Brander and Lewis (1986), we consider single-period debt contracts

($M = 1$) with two competing firms. They study Cournot competition between the firms and find that, under intuitive regularity conditions in the payoff function, that in symmetric Nash equilibrium shareholder-optimal output is increasing in the debt load. To better match the empirical focus of this paper, we consider a slight adjustment to the Brander and Lewis framework. Instead of competing in quantities, consider two firms competing in price, and suppose product markets with enough differentiation that profit functions are twice continuously differentiable. Let $\pi^i(p_i, p_j; z_i)$ be free cash accruing to firm $i$ under prices $p_i$ and $p_j$ and subject to stochastic component $z_i$, and let subscripts on $\pi$ denote partial derivatives. Suppose, without loss of generality, that $z$ increases revenue *ceteris paribus*, i.e. $\pi^i_z > 0$. Define the threshold $\hat{z}(\cdot)$ implicitly by

$$\pi^i(p_i, p_j; \hat{z}_i(p_i; p_j, D)) = D \tag{3.2.1}$$

This is the threshold at which the firm is exactly solvent, i.e. when it is exactly able to pay off its one-period debt obligation $D$, conditional on prices. The creditor (loan or bondholder, $b$) and equity holder (firm management, $f$) objectives are, respectively

$$V_b(p_i, p_j) = \int_0^{\hat{z}_i(p_i;p_j,D)} \pi^i(p_i, p_j; z_i)\ dF(z_i) \quad + \quad D \cdot (1 - F(\hat{z}_i(p_i^*; p_j, D))) \tag{3.2.2}$$

$$V_f(p_i, p_j) = \int_{\hat{z}_i(p_i;p_j,D)}^{\infty} [\pi^i(p_i, p_j; z_i) - D]\ dF(z_i) \tag{3.2.3}$$

where the latter excludes part of the support due to limited liability. The optimal price responses satisfy the following first order conditions.

$$\frac{\partial V_b}{\partial p_i} = 0 = \int_0^{\hat{z}_i(p_i^*;p_j,D)} \pi^i_i(p_i^*, p_j; z_i)\ dF(z_i) \tag{3.2.4}$$

$$\frac{\partial V_f}{\partial p_i} = 0 = \int_{\hat{z}_i(p_i^*;p_j,D)}^{\infty} \pi^i_i(p_i^*, p_j; z_i)\ dF(z_i) \tag{3.2.5}$$

Both of the above expressions include terms additional terms involving effects through the threshold $\partial\hat{z}/\partial p_i$ but these are all identically zero due to the definition of $\hat{z}$, so we omit them.

**Proposition 3.2.1.** *Consider the static model developed above. If* $\pi_{ii} < 0$ *and* $\pi_{iz} > 0$

*then the creditor optimal price is less than the equity holder optimal price.*

**Corollary 3.2.1.1.** *Consider the static model developed above. If $\pi_{ii} < 0$, $\pi_{iz} > 0$, and $\pi_{ij} > 0$ then firm i's optimal price reduces when its competitor (firm j) shifts from equity holder control to creditor control.*

The hypotheses of Proposition 3.2.1 are intuitive and likely to hold empirically. The first, $\pi_{ii} < 0$, is a standard condition requiring that incremental price increases yield successively lower marginal revenues. This would be satisfied, for instance, if price elasticity were increasing in own price, which is the case in most demand systems of which we are aware. The second requires that, at each price vector, the marginal revenue due to a price increase is higher when demand is stronger. This condition would be satisfied if price elasticities were everywhere decreasing in $z$. The additional condition in Corollary 3.2.1.1 is a standard one, and guarantees that the marginal revenue due to an own price increase is increasing in competitor price, i.e. own price elasticity is a declining function of competitor price.

### 3.2.3 Dynamic Incentives

The static case fails to capture investment, which is inherently dynamic in nature. Following the above analysis, define a threshold $\hat{z}(\cdot)$ implicitly by

$$\pi(\hat{z}(s;C);s) = C \tag{3.2.6}$$

that is, the $z$ such that the firm in state $s$ makes profit $C$.

Assume for tractability also that debt can be infinitely rolled over or, equivalently, that debt contracts are issued in perpetuity. Therefore the creditor and shareholder value functions for fixed periodic payment $c$ are, respectively

$$V_b(s;c) = \max_i \int_{\hat{z}(s;c+i)}^{\infty} [c + \beta V_b(s(z;i);c)]\ dF(z) + \int_0^{\hat{z}(s;c+i)} \pi(z;s)\ dF(z) \tag{3.2.7}$$

$$V_f(s;c) = \max_i \int_{\hat{z}(s;c+i)}^{\infty} [\pi(z;s) - c - i + \beta V_f(s(z;i);c)]\ dF(z) \tag{3.2.8}$$

The corresponding first-order conditions are

$$\frac{\partial V_b}{\partial i} = 0 = \hat{z}_C(s; c+i^*) \cdot [V_b(s(\hat{z}(\cdot); i^*); c) - i^*] + \int_0^{\hat{z}(\cdot)} \frac{\partial V_b}{\partial s}\frac{\partial s}{\partial i}\, dF(z) \tag{3.2.9}$$

$$\frac{\partial V_f}{\partial i} = 0 = \hat{z}_C(s; c+i^*) \cdot V_f(s(\hat{z}(\cdot); i^*); c) + \int_{\hat{z}(\cdot)}^{\infty} \frac{\partial V_f}{\partial s}\frac{\partial s}{\partial i}\, dF(z) \tag{3.2.10}$$

where $\hat{z}_C$ denotes its first partial derivative with respect to the second argument and in the first FOC we have used that by definition, $\pi(\hat{z}(s; c+i); s) = c+i$

Debt financing may be used for various corporate purposes; let $k$ be a component of $s$ denoting the firm's total capital stock without consideration of its particular use and let $\partial V/\partial k$ denote the marginal value return on capital. Under similar conditions as above, therefore, the optimal investment level chosen by creditors is strictly lower than the optimal investment level chosen by shareholders, which we summarize in the following proposition.

**Proposition 3.2.2.** *Assume (i)* $V_f \geq V_b - i$ *and (ii)* $\partial V_f/\partial k \geq \partial V_b/\partial k > 0$ *when evaluated at the firm's optimal investment level* $i_f^*$*. Then the creditor optimal investment level is weakly lower than the shareholder optimal investment level; the comparison is strict when at least one of (i) or (ii) is strict.*

Again we believe that the hypotheses are likely to be satisfied in practice. The first says that evaluated at the firm's optimal investment choice, the shareholder value function after being rebated the cost of the investment weakly exceeds the creditor's value function. The second says that the marginal value accruing to shareholders due to an increase in its own capital stock exceeds that marginal value to creditors, an intuitive condition since firms are residual claimants to any excess profit while creditors benefit only from an longer expected lifespan of the firm (and of the perpetuity). Again this is evaluated at the shareholder's optimum.

By imposing additional structure on the problem we can derive tighter conditions under which the result holds, but such specificity is beyond the scope of this paper. To hint at what conditions on the primitives might resemble, consider the special case where a firm's own capital stock $k$ is a sufficient state variable and it depreciates at rate

$1-\delta$ each period. Envelope conditions for banks (creditors) and firms (shareholders), where superscripts indicate the agent and subscripts indicate partial derivatives, are:

$$V_k^b = \int_0^{\hat{z}} \pi_k + \beta\delta \int_{\hat{z}}^{\infty} V_k^b - \beta \hat{z}_k V^b - c\hat{z}_k$$
$$V_k^f = \int_{\hat{z}}^{\infty} \pi_k + \beta\delta \int_{\hat{z}}^{\infty} V_k^f - \beta \hat{z}_k V^f$$

So, for instance, if the first condition held and

$$Pr\{\pi > c+i\} \cdot \mathbb{E}[\pi_k|\pi > c+i] \geq Pr\{\pi < c+i\} \cdot \mathbb{E}[\pi_k|\pi < c+i]$$

were satisfied, where the expectation is taken over the relevant ranges of $z$, then the result would hold. In practice this condition could mean that the marginal value of capital is higher when the firm gets more favorable shocks, i.e. when the firm is relatively likely to remain solvent.

## 3.3 Data

To study the financial-real linkage, we shall require three datasets: one for each of the financial and real sides, plus a third linking them.

### 3.3.1 Data Sources

The loan data – our financial side – come from Thompson Reuters' DealScan, a proprietary dataset constructed with the intent to capture the universe of primary market loan arrangements. Loans are an important part of corporate financing, and account for over half of all corporate debt in the United States, disproportionately so for small firms with limited access to public debt markets.[2] The DealScan data cover a majority of the primary loan volume in the US, including the universe of loans

[2] According to the Loan Syndications and Trading Group (LSTA), an organization of lenders, there is around \$1.2 trillion outstanding (`http://www.lsta.org/news-and-resources/news/fact-sheet-leveraged-loans-in-the-us`), while the total outstanding corporate paper is slightly under \$1 trillion (`https://fred.stlouisfed.org/series/COMPOUT`).

to publicly traded companies (Carey and Hrycray (1999)). They are collected from public filings (SEC forms 10-K, 10-Q, and 8-K), directly from creditors (especially in the case of syndicated loans), and various other internal sources.

We are primarily interested in the signaling effects of a control rights transfer, just as in Chevalier (1995); if there were no public signal, the only response mechanism would be through price, which would be highly lagged and noisy. As in the Ellison and Ellison (2011) setup, competitor response is dependent upon such observation. We therefore restrict our attention to violations that are unambiguously observable. In particular, we focus on covenants that target one of three financial variables: (i) current ratio, (ii) net worth, or (iii) tangible net worth. Other commonly used covenants are complicated by accounting inconsistencies, which makes them more difficult to use and we therefore prefer the latter, which in any case have a total value of over $1 trillion (Dichev and Skinner (2002)).[3]

As an check that our violations dataset is accurate, we check it against a public dataset constructed by Roberts and Sufi (2009); Roberts maintains an updated version of the dataset on his faculty website.[4] It purports to include all covenant violations reported between 1996 and 2012 for publicly traded firms in the US, but we find that coverage is much sparser in later years. For this reason, and to more closely match Chava and Roberts, we truncate our sample at the end of 2004.

We compare our imputed violations in each quarter with violations reported in public SEC filings in Figure 3-2. From the figure it seems that our data qualitatively match those from the SEC filings. Our counts are strictly higher, though likely this is attributable to reporting frequency. Moreover, we do not observe contract renegotiation so it is possible that we overstate the frequency of violations. Fortunately our panel is relatively balanced and the raw violation counts closely match the relative proportion of firms observed in each quarter that are in violation, as given in Figure 3-1.

[3]This figure includes revolving credit facilities which typically are not fully drawn, hence the inconsistency with the total loan market value reported earlier.

[4]Available at `http://finance.wharton.upenn.edu/~mrrobert/styled-9/styled-11/index.html`

Our firm-level data is drawn from Compustat, which provides rich fundamental data for both public and private companies, though as one might expect the firm mix is skewed toward relatively large firms. Therefore, despite our covenant data being drawn only from public firms, product market effects drawn on private firms as well. Though we do not use it, we pull linked CRSP data as well, which could be valuable in extending the results, for instance to link with the patent dataset compiled by Kogan et al. (2016).

To link the two datasets, we rely upon a joint project between Wharton Research Data Services (WRDS) and Michael Roberts, undertaken as part of Chava and Roberts (2008), and maintained on both his faculty website and on WRDS. Chodorow-Reich (2014) performs a similar exercise to link the DealScan data with the Longitudinal Business Database (LBD).

### 3.3.2 Sample Construction

We take the full set of firm-quarter observations from the Compustat data between 1996Q1 and 2004Q4 and onto each merge an indicator for whether a covenant violation was reported for the firm in the specified quarter. Unlike Chava and Roberts, we do not restrict our sample only to firms with a covenant violation. We construct variables from the fundamentals that capture most directly the financial status of each firm and that are normalized such that their range is independent of firm size or type. A full description of these composite variables is provided in Appendix C.1. For purposes of completeness, we take only firms with eight or more quarters of data. This dataset is the basis for the results of Section 3.4.

In order to study market-level effects, we group firms into industries by their Standard Industrial Classification (SIC) codes. The code is four digits: the division (11), major group (83), industry group (416), and industry (1005), respectively (parentheses indicate the number of unique codes at each level). As an example, 2013 is "Sausages and Other Prepared Meat Products" while 2015 is "Poultry Slaughtering and Processing"; the hierarchy is "Manufacturing" → "Food and Kindred Products" → "Meat Products."

For each SIC and in each quarter, we compute the ratio total revenue of current firms in covenant violation to the total revenue of the industry. We construct the same variable but for violators in the previous quarter, as well as for violators in any of the previous four quarters. In order to eliminate markets in which one firm is highly dominant, we consider only industries for which no firm captures over 80% of the market by revenue in any given quarter, though our results are generally robust to this choice. This dataset is used to generate the results in Section 3.5.

### 3.3.3 Summary Statistics

Our final sample includes 20,312 unique firms; by contrast, there are approximately 4,000 publicly traded companies in the United States (Henderson (2019)). Table 3.1 gives summary statistics for variables of central importance in the paper. In total we have around 520,000 firm-quarter observations, though some fields are missing for some quarters, as indicated in the table.

The modal firm is never in violation. In total we record 11,896 firm-quarters of active covenant violations across 1,769 unique firms. Of those, 1,510 firms were in violation for more than one quarter and 947 were in active violation during more than four of the 36 quarters in our sample period.

The median firm has a 2.5% return on assets and invests roughly 5% of the outstanding value of its property, plant, and equipment on capital. The median (likely research intensive) firms that reports R&D expenditures spends 10% of its revenue on research and development, though the latter term likely overstates the overall median because it excludes observations with empty/missing R&D values.

Table 3.2 gives the same summary statistics for the subset of firms that were in covenant violation in at least one quarter over our sample. As the moments of the net worth and revenue distributions show, these firms tend to be smaller than those of the full sample. Market to book and Macro Q – both important factors in investment decisions – are relatively stable across samples.

Table 3.1: Full sample summary statistics.

| Statistic | N | Mean | St. Dev. | Median |
|---|---|---|---|---|
| $\frac{\text{CapEx}}{\text{PP\&E}}$ | 323,497 | 0.083 | 0.101 | 0.051 |
| Capital Asset Ratio | 413,081 | 0.275 | 0.263 | 0.181 |
| Current Ratio | 361,297 | 3.152 | 5.757 | 1.739 |
| Leverage | 271,187 | 0.348 | 0.452 | 0.273 |
| Net Worth (mil.) | 440,822 | 710.5 | 3689.4 | 49.8 |
| Revenue (mil.) | 520,425 | 322.4 | 1959.4 | 26.9 |
| Market To Book | 268,733 | 0.670 | 0.212 | 0.723 |
| Macro Q | 262,854 | 1.832 | 1.102 | 1.457 |
| ROA | 385,756 | 0.024 | 0.066 | 0.022 |
| $\frac{\text{R\&D}}{\text{Revenue}}$ | 126,076 | 0.167 | 0.184 | 0.106 |
| NumViol | 20,312 | 0.588 | 2.534 | 0 |

Table 3.2: Violators sample summary statistics.

| Statistic | N | Mean | St. Dev. | Median |
|---|---|---|---|---|
| $\frac{\text{CapEx}}{\text{PP\&E}}$ | 49,423 | 0.076 | 0.084 | 0.051 |
| Capital Asset Ratio | 53,220 | 0.293 | 0.243 | 0.215 |
| Current Ratio | 51,257 | 2.243 | 2.642 | 1.690 |
| Leverage | 40,747 | 0.361 | 0.326 | 0.289 |
| Net Worth (mil.) | 56,010 | 382.0 | 2338.1 | 66.9 |
| Revenue (mil.) | 59,296 | 212.0 | 778.8 | 36.526 |
| Market To Book | 40,582 | 0.704 | 0.166 | 0.737 |
| Macro Q | 39,909 | 1.871 | 1.114 | 1.495 |
| ROA | 48,779 | 0.027 | 0.048 | 0.027 |
| $\frac{\text{R\&D}}{\text{Revenue}}$ | 15,566 | 0.131 | 0.151 | 0.082 |
| NumViol | 1,769 | 6.756 | 5.664 | 5.0 |

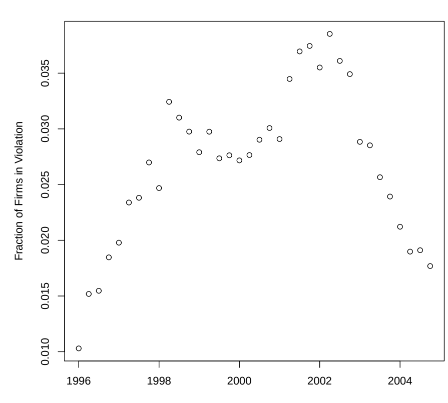
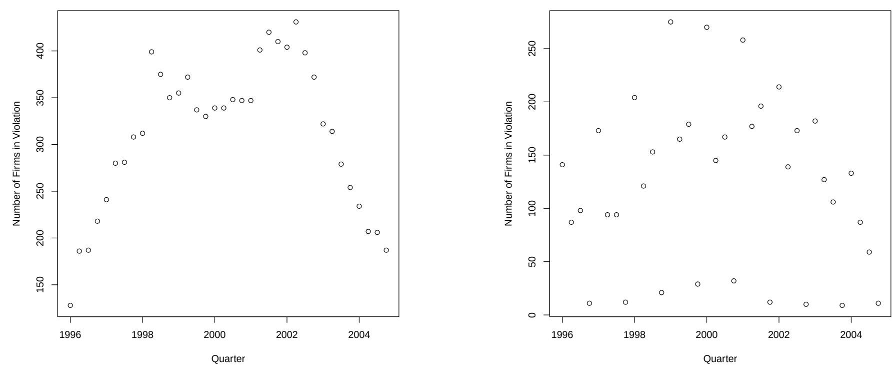



Figure 3-1: Fraction of firms observed in violation of one or more covenants in each quarter of our sample, 1995Q1–2004Q1.




(a) Dealscan data (b) SEC filings

Figure 3-2: Count of firms in violation of one or more covenants in each quarter, 1996Q1–2004Q1.

## 3.4 Firm Response to Covenant Violations

Debt covenants provide a convenient institutional context by which to study the impact of control rights on firm behavior. Covenants in general and the subset on which we focus in particular lend themselves to a regression discontinuity design, as the covenant triggers at a fixed threshold (typically a round number) on the basis of one or more continuous financial variables.

We begin by examining the raw data for insight and to motivate our empirical strategy. Figure 3-3 plots the mean of selected financial variables of interest against time in quarters, where $t = 0$ corresponds to the first covenant violation for each firm.

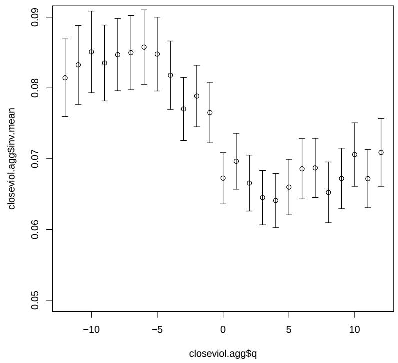
The first panel shows investment. There appears to be a pre-trend beginning several quarters prior to the violation in which investment is drawn down. The pairwise difference between any of the two periods is only on the cusp of significance at the 5% level but the joint test of whether the mean investment in the three quarters preceding a violation and the three quarters preceding those is equivalent is rejected at the 1% level. Nonetheless, even after controlling for the pre-trend, there is a discontinuous drop at $t = 0$ that we will exploit in our later analysis. We further analyze the discontinuity in Figure 3-4, which plots investment against the normalized distance to the violation threshold. For firms bound by multiple covenants we take this to be the most tightly binding of the set; normalization is necessary since thresholds and units of the various covenants differ.

The second panel of Figure 3-3 maintains the same time indexing but plots the current ratio. Notably, the current ratio is one of the financial variables that are contracted on in financial covenants, so the pre-trend in this case is in some sense by construction — without any pre-trend it would be impossible for firms not to be in violation initially and later find themselves in violation. However, the sharpest decline in the current ratio comes one to two quarters *after* the violation, after which it quickly recovers then slowly moves back toward the pre-violation level. Only 3,126 of our 13,908 observed violations were due to current ratio covenants, so a majority of firms contributing data to this figure were violators of an unrelated covenant (note that 15% of violation-quarters had more than one type of covenant in violation, but a vast majority of these were two net worth covenants). Combined with the first panel, the suggestive evidence is that firms reduce investment in order to ballast their short term financial condition and avoid a covenant violation if possible.

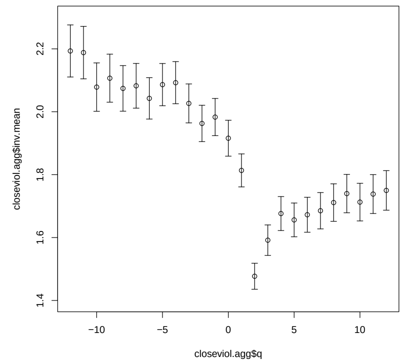
The third and fourth panels describe the behavior of revenue around a covenant violation. As a percentage of assets, revenue is extremely stable across the covenant violation. However, assets may also be changing around the threshold — that the current ratio moves down sharply suggests that assets are likely to decline. Therefore in the fourth panel we plot the change in log revenue — roughly the percentage change in revenue from quarter to quarter — and observe what appears to be a sharp reduction in revenue growth rate in the several periods following the covenant violation. In subsequent periods, though, the revenue trajectory approximately regains its pre-violation rate of growth.

### 3.4.1 Primary Empirical Specification

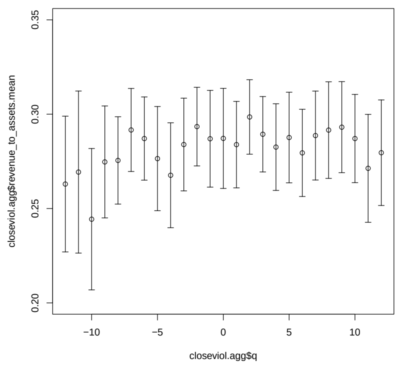
Inspired by the above evidence, we construct the variable $QSinceViol$, the number of fiscal quarters since the covenant violation was reported, which we motivate by noting that empirically the strongest investment responses are several quarters lagged after a violation. Identification of this variable comes only from firms with one or more



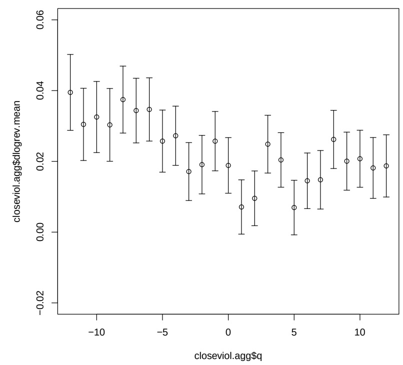

(a) Investment




(b) Current Ratio




(c) Revenue/Assets




(d) $\Delta log$(Revenue)

Figure 3-3: Conditional means of selected financial variables for the subset of firms with at least one covenant violation over our sample, 1996Q1 – 2004Q4. The horizontal axis corresponds to quarters before/after each firm's *first* violation; $t = 0$ is the first period in which the firm was observed to be in violation of one or more covenants. Error bars correspond to a 95% confidence interval on the point estimates of the conditional means.

violation. We also include the contemporaneous indicator $InViol_t$ which indicates that the firm is current in violation of one or more covenants as well as an indicator for whether the firm was not in violation in the previous quarter but has since violated:

$$NewViol_t = \mathbb{1}\{QSinceViol_{t+1} - QSinceViol_t = 1\}$$

where $\mathbb{1}$ is the indicator function. Anticipating that responses to future violations after the first might differ from the first due to a number of factors, we also include as a regressor $NumViol$, the number of violations beyond the first, which is identified

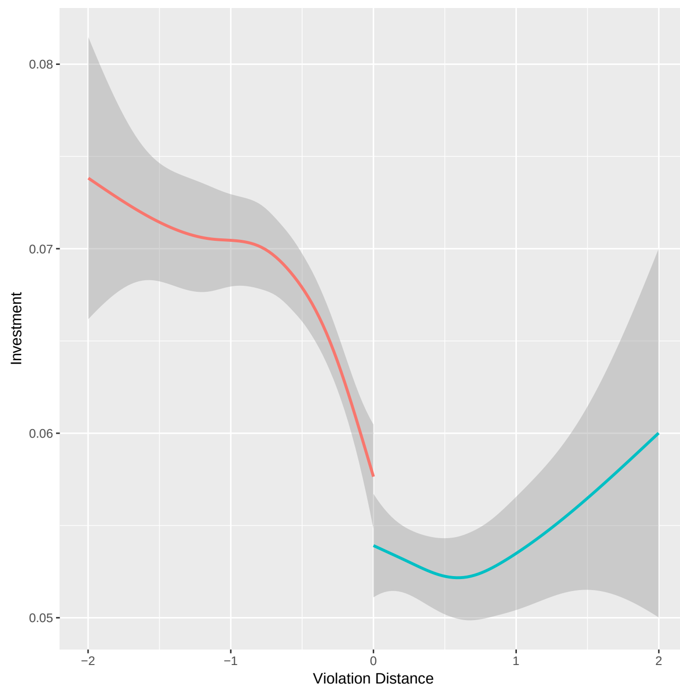



Figure 3-4: Investment as a function of normalized distance to the covenant threshold for the subset of firms with at least one covenant violation over the sample period. Smoothing is performed by local regression on a rectangular window (uniform kernel) of width 0.1.

off of firms with multiple violations. Controls $X$ are cash flow, capital asset ratio, current ratio, leverage, net worth, ROA, and Tobin's $q$, chosen to mirror the literature and also to cover the fundamentals upon which covenants are based. All are lagged with the exception of cash flow, since investments are financed contemporaneously with cash inflows. Finally, we include the normalized distance to the tightest binding covenant. For each of the three types of covenants and each period in which a firm has one or more of the three, we can compute the ratio of the distance between the underlying financial variable and the covenant threshold, and the threshold itself. Therefore a negative ratio indicates violation and a ratio of precisely zero indicates that the firm is right on the threshold. We then divide each of these three ratios by the standard deviation of each across all firms such that the distribution of all three metrics are qualitatively very similar. Finally, for each firm-quarter we record the tightest (most negative) of the three as that firm's distance to technical default.

Our main estimating equation for this section is

$$Investment_{it} = \alpha_i + \lambda_t + Viol'_{it}\beta + +X'_{it}\gamma + \varepsilon_{it} \tag{3.4.1}$$

where $\alpha_i$ and $\lambda_t$ are firm and quarter-year fixed effects, respectively. An immediate concern is whether a fixed effects estimator is appropriate. We are additionally concerned that there may be persistence in investment decisions if, for instance, capital expenditure is required to maintain equipment and adjustment costs are high. We are therefore inclined to include lagged outcome variables. However, for serially correlated $\varepsilon_{it}$ the naïve approach of including a lagged outcome within a fixed effects model is inconsistent if the number of within observations (time periods) is fixed, even if the number of firms grows very large (Nickell (1981)).

In light of these concerns, we consider two approaches. First, we can run two separate regressions: one with fixed effects and one with the lagged dependent variable, but neither having both. These models will generally bracket the true effect (Angrist and Pischke (2008)). Second, we can use a dynamic panel model to consistently estimate the true effect in the presence of both autocorrelated outcomes and errors as well as firm-level fixed effects (Arellano and Bond (1991)). In practice, we take the former approach, which yielded stable point estimates for parameters of interest across varied specifications. The Arellano-Bond estimator was quite unstable, which is likely due in part to the weak instrument problem highlighted by Blundell and Bond (1998).[5]

The results are given in Table 3.3. The point estimate on the violation indicator, constructed as above, is stable across specifications at around $0.8 - 1\%$. The outcome is quarterly investment divided by book value of property, plant, and equipment. Relative to the baseline of around 8% quarterly investment, this represents a reduction

[5] The standard `R` implementation in `plm` library failed to converge after running for over 12 hours on a single core; an ad hoc GMM estimator based upon the Arellano-Bond moment condition did converge but was highly sensitive to initial conditions and seemed to have a pretty flat objective. Computational issues aside, consistency in the dynamic panel setting is highly contingent upon properly specifying the functional form. This is prohibitive in our setting: if the proper specification is AR($n$) with $n$ large, we simply lack sufficient data for estimation (we have about ten years of quarterly data available).

of $10 - 15\%$. The magnitude of the effect is very close to that estimated by Chava and Roberts (2008), so we are confident in our sample construction and preliminary analysis despite not closely following their methodology.

Our preferred specification is (8), which includes time and firm fixed effects, plus the full set of controls. The estimated coefficients on the controls are all sensible, with the exception of that on Tobin's Q, which we should expect to be positive. The latter is a gross measurement (by necessity, due to the available accounting data), when in practice investment decisions should be made on the basis of the *marginal* product of capital, for which *marginal* Q is a proxy. In specifications (4) and (7), which have the full set of fixed effects but lack to full set of controls, macro Q has the expected sign. That it becomes negative in the final specification could point to collinearity with contemporaneous and lagged financial variables. We find it reassuring that lagged return on assets remains a very strong predictor of investment in both specifications in which it is included, consistent with intuition about investment decisions.

All specifications with the exception of (4) include flexible controls for the normalized default distance. In our preferred specification (8) investment is increasing in default distance, or, equivalently, it is decreasing as the firm gets closer to default. The third degree polynomial permits two stationary points plus a point of inflection, so it could in principle fit a highly nonlinear response as long as there is no discontinuity at the point of violation. To quantify the point estimate is about 1% on the linear term, consider a firm with a current ratio covenant fixed at 1.0. Then if that firm were investing 8% when its current ratio were at 2.0 it would have cut its investment to 7% when it got very near to the threshold. Our estimates then indicate that the firm would discontinuously cut another 1% of investment if its current ratio continued to deteriorate past the covenant threshold.

### 3.4.2 Additional Evidence

Recent work has suggested an important link between covenant violations and R&D spending, but has failed to find downstream losses in innovation by various measures (Chava et al. (2016); Bellucci et al. (2014); Mann (2016)). The indication is that

Table 3.3: The effect of a covenant violation on the violator's capital expenditure. Estimation is on the sample of firms that were in technical default for at least one quarter in our data.

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\frac{\text{Investment}}{\text{PP\&E}}$ | | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Violation | −0.006** | −0.008** | −0.011*** | −0.009** | −0.004 | −0.004 | −0.008** | −0.009** |
| | (0.003) | (0.003) | (0.004) | (0.004) | (0.003) | (0.003) | (0.004) | (0.004) |
| Lag(Investment) | | | 0.028*** | | | | | |
| | | | (0.008) | | | | | |
| Default Dist. | 0.011*** | 0.006* | 0.006 | | 0.006* | 0.002 | −0.002 | 0.011** |
| | (0.003) | (0.003) | (0.004) | | (0.004) | (0.004) | (0.004) | (0.005) |
| (Default Dist.)$^2$ | −0.001 | 0.007*** | 0.002 | | −0.001 | 0.007** | 0.006 | −0.001 |
| | (0.001) | (0.003) | (0.003) | | (0.001) | (0.003) | (0.004) | (0.004) |
| (Default Dist.)$^3$ | | −0.002*** | −0.0005 | | | −0.001** | −0.001 | −0.00002 |
| | | (0.001) | (0.001) | | | (0.001) | (0.001) | (0.001) |
| Cash Flow | | | −0.075*** | | | | | −0.064*** |
| | | | (0.002) | | | | | (0.002) |
| Lag(Curr. Ratio) | | | 0.003** | | | | | 0.007*** |
| | | | (0.001) | | | | | (0.002) |
| Lag(Cap. Asset Ratio) | | | −0.007 | | | | | −0.184*** |
| | | | (0.008) | | | | | (0.031) |
| Lag(Leverage) | | | −0.043*** | | | | | −0.011 |
| | | | (0.008) | | | | | (0.016) |
| Lag(ROA) | | | 0.401*** | | | | | 0.230*** |
| | | | (0.036) | | | | | (0.048) |
| Lag(Macro Q) | | | 0.0002*** | 0.003*** | | | 0.002*** | −0.005*** |
| | | | (0.0003) | (0.001) | | | (0.001) | (0.001) |
| Lag(Macro Q)$^2$ | | | | −0.00000 | | | −0.00000 | 0.0001*** |
| | | | | (0.00000) | | | (0.00000) | (0.00001) |
| Time FE | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Firm FE | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Observations | 23,168 | 23,168 | 15,806 | 18,025 | 23,168 | 23,168 | 18,025 | 16,082 |
| $R^2$ | 0.002 | 0.003 | 0.097 | 0.003 | 0.0004 | 0.001 | 0.003 | 0.080 |
| F Statistic | 17.405*** | 15.449*** | 153.844*** | 17.143*** | 3.081** | 3.636*** | 9.443*** | 116.087*** |
| *Note:* | | | | | | | | *p<0.1; **p<0.05; ***p<0.01 |

creditors are less likely to impose strict rationing on highly innovative firms, and moreover that creditors tend to make such choices judiciously. We therefore estimate three first-difference specifications of the form

$$\Delta \ln(Y_{it}) = \beta \; InViol_{it} + X_{it}'\gamma + \varepsilon_{it} \tag{3.4.2}$$

where controls $X$ are as above, plus quarterly fixed effects. The outcomes of interest $Y$ are R&D spending, total value of common equity, and return on assets (ROA). We omit firm effects because they are differenced out.

The results are presented in Table 3.4. Our broad conclusion is that shareholders lose some value in response to a covenant violation and subsequent transfer of control rights, but that the loss is more modest than the magnitude of adjustment on the investment side. We also find that financial performance is impaired and R&D declines in response to the covenant violation, but both of these point estimates are too noisy to be distinguishable from zero after flexibly controlling for distance to the covenant threshold. These results are consistent with the qualitative conclusions of Chava et al., mirroring the above results; the same discussion applies here.

As expected, R&D spending and common equity value are both increasing in the distance to the covenant threshold. This distance represents both likelihood of technical default *and* overall financial condition, so cannot be interpreted as the endogenous response in advance of a violation. On the other hand, return on assets is declining in distance to the threshold. Intuitively, return on assets is strictly declining in the *level* of assets for fixed profit; default distance is strictly increasing in a firm's asset level, all else equal, which could explain the effect.

One might expect that equity holders of more highly levered firms suffer proportionally less severe losses in response to a violation because ex ante creditor protections are stronger, which would be consistent with, for instance, LBO theory (Jensen (1986); Chevalier (1995)). We test this empirically in Table 3.5 and estimate it on the full sample. Concordantly, we omit flexible controls for default distance because a majority of firms are never in or near default. However, we find insufficient evidence to make such a conclusion. On the other hand, firms with stronger ROA suffer much more severely: our point estimate indicates every 3% higher quarterly lagged ROA is associated with a 1% greater loss in equity value in response to a violation. The estimate on the net worth interaction is also negative, indicating again that shareholders in financially stronger companies lose more ground upon a control rights transfer. Taken in whole, the set of evidence is consistent with findings on post-violation debt issuance (Roberts and Sufi (2009)), indicating that our results are likely not driven by cyclical deleveraging (Flannery and Rangan (2006); Kayhan and Titman (2007)).

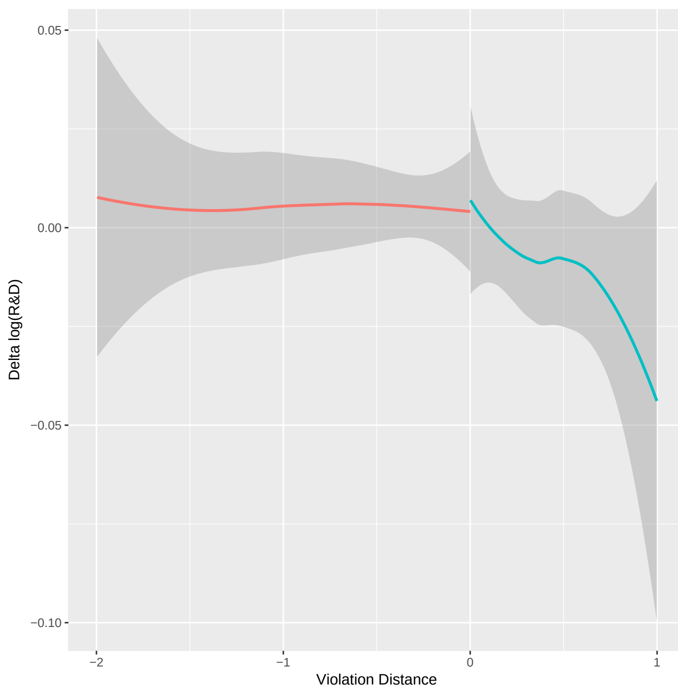



Figure 3-5: Change in log R&D spending as a function of normalized distance to the covenant threshold for the subset of firms with at least one covenant violation over the sample period. Smoothing is performed by local regression on a rectangular window (uniform kernel) of width 0.1.

## 3.5 Competitor Response to Covenant Violations

We now turn our attention to the effects of a covenant violation on other firms in the same industry as the violator, grouping industries at the level of the SIC code. As is standard with industry-level analysis, we must pay careful attention to correlated

Table 3.4: The effect of a covenant violation on the violator's R&D expenditure and financial status. Estimation is on the sample of firms with at least one covenant violation in our sample period.

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | Δlog(R&D) | | Δlog(Equity) | | Δlog(ROA) | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Violation | −0.012 | −0.011 | −0.020*** | −0.019*** | −0.009 | −0.012 |
| | (0.008) | (0.009) | (0.002) | (0.003) | (0.007) | (0.008) |
| Macro Q | −0.0001 | 0.008*** | 0.001*** | −0.004*** | −0.0001 | −0.004 |
| | (0.001) | (0.003) | (0.0002) | (0.001) | (0.001) | (0.002) |
| (Macro Q)$^2$ | | −0.0002*** | | 0.00005*** | | 0.00003 |
| | | (0.0001) | | (0.00001) | | (0.00003) |
| Default Dist. | 0.014* | 0.009 | 0.030*** | 0.054*** | −0.019*** | −0.028*** |
| | (0.008) | (0.011) | (0.003) | (0.004) | (0.007) | (0.010) |
| (Default Dist.)$^2$ | −0.003 | −0.007 | −0.011*** | −0.020*** | −0.001 | −0.003 |
| | (0.007) | (0.009) | (0.002) | (0.003) | (0.006) | (0.008) |
| (Default Dist.)$^3$ | 0.001 | 0.001 | 0.001*** | 0.002*** | 0.001 | 0.002 |
| | (0.001) | (0.002) | (0.0004) | (0.001) | (0.001) | (0.001) |
| Cash Flow | −0.001 | −0.007*** | 0.062*** | 0.069*** | 0.035*** | 0.058*** |
| | (0.001) | (0.003) | (0.002) | (0.002) | (0.009) | (0.012) |
| Lag(Curr. Rat.) | 0.003* | 0.006 | −0.001* | −0.006*** | 0.003 | 0.002 |
| | (0.002) | (0.004) | (0.001) | (0.001) | (0.002) | (0.005) |
| Lag(C-A Ratio) | −0.023 | 0.099 | 0.003 | −0.059*** | −0.003 | −0.104* |
| | (0.023) | (0.088) | (0.005) | (0.020) | (0.012) | (0.059) |
| Lag(Leverage) | 0.003 | −0.070* | −0.005 | 0.075*** | 0.021* | 0.120*** |
| | (0.019) | (0.040) | (0.006) | (0.012) | (0.013) | (0.030) |
| Lag(ROA) | 0.305*** | 0.385*** | 0.505*** | 0.388*** | | |
| | (0.051) | (0.089) | (0.027) | (0.035) | | |
| Constant | −0.003 | | −0.008** | | 0.004 | |
| | (0.010) | | (0.003) | | (0.009) | |
| Time FE | No | Yes | No | Yes | No | Yes |
| Firm FE | No | Yes | No | Yes | No | Yes |
| Observations | 2,612 | 2,612 | 12,949 | 12,949 | 8,076 | 8,076 |
| $R^2$ | 0.028 | 0.026 | 0.174 | 0.128 | 0.004 | 0.009 |
| *Note:* | | | | | | *p<0.1; **p<0.05; ***p<0.01 |

Table 3.5: Investigating the channels by which covenant violations operate and are exacerbated. Effects are for the violating firm and are estimated on the full sample.

| | *Dependent variable:* | |
|---|---|---|
| | Δlog(Equity) | |
| | (1) | (2) |
| Violation | −0.038*** | −0.014 |
| | (0.007) | (0.034) |
| Lag(Capital Asset Ratio) | −0.009* | −0.009* |
| | (0.005) | (0.005) |
| Lag(Current Ratio) | 0.001** | 0.001** |
| | (0.001) | (0.001) |
| Lag(Leverage) | −0.024*** | −0.022** |
| | (0.009) | (0.009) |
| Lag(Net Worth) | −0.050*** | −0.046*** |
| | (0.008) | (0.008) |
| Lag(ROA) | 0.479*** | 0.494*** |
| | (0.020) | (0.021) |
| Lag(Macro Q) | 0.00002 | 0.00002 |
| | (0.00004) | (0.00004) |
| Violation×Lag(Leverage) | | 0.010 |
| | | (0.053) |
| Violation×Lag(Net Worth) | | −0.083* |
| | | (0.043) |
| Violation×Lag(ROA) | | −0.296*** |
| | | (0.100) |
| Constant | 0.035*** | 0.032*** |
| | (0.006) | (0.006) |
| Firm Effects | No | No |
| Quarter Effects | Yes | Yes |
| Observations | 82,102 | 82,102 |
| $R^2$ | 0.009 | 0.009 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

shocks, both on the demand side and on the upstream supply side. We therefore restrict our attention to the sample of firms that have never had a covenant violation. Another concern is that industry concentration moves endogenously with covenant violations, reflecting default or other restructuring. To address this in part, we take only the sample of industries for which no firm in any year captured over 80% of the industry-level revenues, though the results are generally robust to other, similar cuts.

### 3.5.1 Empirical Strategy

For firm $i$ in industry $I$ (a set of firms in that industry), we construct the following measure of industry-level covenant violations for each period $t$:

$$SICViolRev_{it} = \frac{\sum_{j \in I \cap V_{t-1}} Rev_{jt}}{\sum_{k \in I \setminus i} Rev_{kt}}$$

where $V_t$ is the subset of all firms that were in violation of one or more of their covenants in period $t$. The measure captures the proportion by revenue of the remainder of firm $i$'s industry (holding $i$ out) that is in technical default. We take this to be a good proxy for how strongly the control rights transfer should affect the industry: a violation by a firm that accounts for a third of the industry by revenue should have a greater impact on competition than if the violation were by a firm that controls only a few percent of the market. We will also define the dummy

$$SICViol_{it} = \mathbb{1}\{SICViolRev_{it} > 0\}$$

which captures any firm-quarter-year that knows one (or more) of its competitors is in technical default, independent of that competitor's relative size, the extent of the violation, or the number of such competitors.

We estimate specifications of the form

$$Y_{it} = \alpha_i + \lambda_t + \beta \; SICViolRev_{it} \; + \; X_{it}'\gamma + \varepsilon_{it} \tag{3.5.1}$$

In $X$ are included all of the firm-level fundamentals as before, and we again include firm and time fixed effects when appropriate. Therefore we should expect that the partial effect of one or more violations being activated at the margin is driven by the consequence of the violation. The violation variable is constructed as before, that is, it turns on only in the first period following a violation. If a correlated shock drove both the outcome and the covenant violation, it should have struck the industry in the period that the covenant was violated, that is, in the period before our violation indicator activates (since we take reported violations, which are reported at the ends of fiscal quarters).

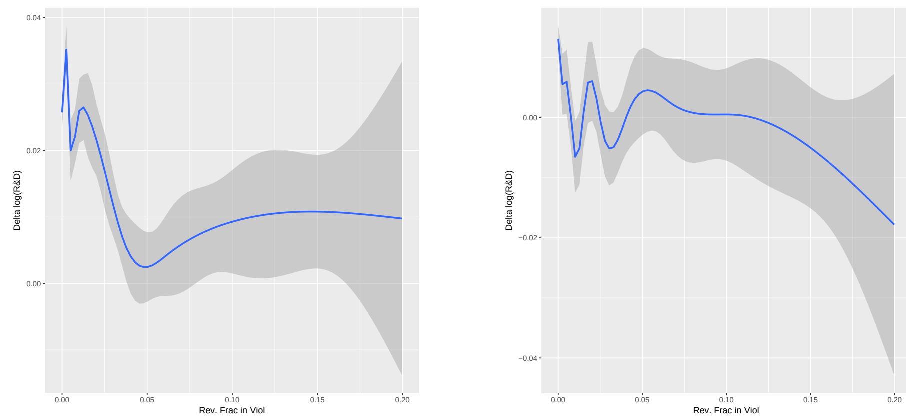
### 3.5.2 Results

Our results are given in Table 3.6. The outcomes $Y$ of interest are log changes in investment, R&D, common equity value, and ROA. The first differences regressions have quarterly effects in place of the firm-year-quarter fixed effects. To translate the point estimates into absolute levels, note that the mode and median of SICViolRev are 0%; the mean is 1.8% while the third quartile is 1.3%.

We find that investment reduces by about 1.5% per 10% of revenue in violation; R&D reduces by nearly 2% in response to the same level of violation, equity by just over 0.5%. Return on assets reduces by roughly the same amount as common equity at the point estimate, but is too noisy to distinguish from zero. We take these results as preliminary evidence of intensifying industry-level competition as a consequence of a covenant violation, an idea to which we will return briefly.

Among covariates, the best predictor across all specifications was lagged return on assets. Unsurprisingly, its point estimate in specification (4) is extremely negative: an inordinately high lagged return on assets predicts strong mean reversion. On the other hand, high lagged ROA is associated with increases in investment, R&D spending, and common equity value.

Across all specifications, the coefficients on the default distance covariates are very nearly zero. It is for this reason that we omitted these variables from the final specification in the previous section. Intuitively covenant violations that are very far from binding are unlikely to affect firm decision making, so including large values of default distances would likely generate only spurious point estimates. We also faced the issue of how to handle firms lacking the types of covenants that we study. To address both of these issues, we top coded the default distance to 5.0 which is roughly twice the value at the third quartile of firms with covenants. Therefore there is a large point mass of firms with the same default distance and it is unsurprising that the point estimates would be very near to zero.




(a) R&D (b) Investment

Figure 3-6: Nonparametric evidence on the effect of the relative proportion of competitors in technical default on the growth rate of R&D and investment spending by firms that were never in technical default.

## 3.6 Production Function Estimation

### 3.6.1 Estimation Procedure

We assume that the true production function in levels is given by

$$Y_{it} = F(X_{it}^1, \ldots, X_{it}^V, K_{it}; \beta_{j(i)}) \exp\{\omega_{it}\}$$

where $X$ are variable inputs, $K$ is capital, $\omega$ is a firm-level productivity shock, and production parameters $\beta$ are fixed at the industry level $j$. In logs this gives

$$y_{it} = f(x_{it}, k_{it}; \beta_{j(i)}) + \omega_{it} + \epsilon_{it}$$

Table 3.6: Effect of a covenant violation on other firms in the industry. Data are estimated on the sample of firms that were never in technical default during our sample, subject to the constraint that no firm controlled over 80% of the market by revenue. Industries are defined at the SIC level.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Δlog(Investment) | Δlog(R&D) | Δlog(Equity) | Δlog(ROA) |
| | (1) | (2) | (3) | (4) |
| SICViolRev | −0.155*** | −0.199*** | −0.055*** | −0.041 |
| | (0.033) | (0.034) | (0.011) | (0.027) |
| Default Dist. | −0.000 | −0.00000 | 0.000 | 0.000 |
| | (0.00000) | (0.00000) | (0.00000) | (0.00000) |
| $(\text{Default Dist.})^2$ | 0.000 | 0.000 | −0.000 | −0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $(\text{Default Dist.})^3$ | −0.000 | −0.000 | 0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Cash Flow | −0.0001 | −0.0002*** | 0.001*** | 0.003*** |
| | (0.0005) | (0.0001) | (0.0001) | (0.0005) |
| Lag(Curr. Ratio) | −0.001** | 0.003*** | −0.001*** | 0.0001 |
| | (0.0003) | (0.0002) | (0.0001) | (0.001) |
| Lag(C-A Ratio) | 0.002 | −0.016** | −0.002 | 0.001 |
| | (0.004) | (0.007) | (0.001) | (0.003) |
| Lag(Leverage) | 0.007** | 0.0003 | −0.004** | 0.009*** |
| | (0.003) | (0.0003) | (0.002) | (0.003) |
| Lag(ROA) | 0.087*** | 0.032*** | 0.326*** | −1.827*** |
| | (0.012) | (0.004) | (0.004) | (0.034) |
| Lag(Macro Q) | 0.00003 | 0.00002 | −0.00001*** | −0.00003 |
| | (0.0001) | (0.00004) | (0.00000) | (0.00003) |
| Constant | 0.007*** | 0.015*** | 0.007*** | 0.082*** |
| | (0.002) | (0.002) | (0.001) | (0.003) |
| Observations | 66,927 | 33,848 | 118,295 | 73,773 |
| $R^2$ | 0.001 | 0.009 | 0.056 | 0.038 |
| *Note:* | | | | *p<0.1; **p<0.05; ***p<0.01 |

where lower case variables are taken in logs. Due to data limitations, we have $y_{it}$ as the log of deflated sales rather than a quantity. Our estimates may therefore be biased in levels but we have strong reason to believe that this bias, if any, acts only to weaken our results. Moreover, the control function approach that we highlight next ought to absorb endogeneity in price that is correlated with variation in unobserved productivity. We address these issues in the robustness section that follows; for further discussion, see §VI.A of De Loecker and Warzynski (2012).

We follow the estimation procedure described by Ackerberg et al. (2015). An immediate downside relative to Olley and Pakes (1996) or Levinsohn and Petrin (2003) is that we cannot directly recover production function parameters in the first stage. In return, however, we are bound by much less restrictive functional forms. In particular, we will prefer a Cobb-Douglas production function of the form

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \epsilon_{it}$$

One advantage of this functional form is that the output elasticity with respect to labor is no longer constant across time. The estimation procedure then proceeds as follows. First we must assume that capital accumulation is a deterministic function:

$$k_{it} = \kappa(k_{it-1}, i_{it-1})$$

where investment $i$ is chosen in the previous period. Next we must assume that there is some *unobserved* intermediate input $m$ that does not enter the production function directly and is chosen according to

$$m_{it} = h_t(k_{it}, l_{it}, \omega_{it})$$

where $h_t$ is strictly increasing in $\omega_{it}$. Ackerberg et al. suggest that the "material input" $m$ ought to be Leontief inputs to production but do not give much further interpretation. We think of these inputs as administrative positions, e.g. managers and accountants. Clearly these positions do not scale linearly in the size of the business, but it is plausible that they do grow monotonically, at least within industry. This view is productive also because it is testable: in financial data, costs of variable inputs to production are recorded under the line item "Cost of Goods Sold" while these material inputs under consideration are separately recorded under the line item "Selling, General, and Administrative". This accounting item includes a number of expenses that would fall outside of our definition of material inputs but nonetheless provides a rough empirical test of the identifying assumption in our setup.

Subject to those two assumptions we can define

$$\Phi_t(k_{it}, l_{it}, m_{it}) \equiv \beta_0 + \beta_k k_{it} + \beta_l l_{it} + h_t(k_{it}, l_{it}, m_{it}) + \epsilon_{it}$$

such that

$$y_{it} = \Phi_t(k_{it}, l_{it}, m_{it}) + \epsilon_{it}$$

We estimate $\Phi$ using the first stage moment condition

$$\mathbb{E}[\epsilon_{it}|I_{it}] = 0$$

Clearly we are unable to identify any structural parameters $\beta$ in the first stage. To obtain identification in the second stage, we must additionally assume that productivity shocks are first-order Markov, implying that they follow a process

$$\omega_{it} = g(\omega_{it-1}) + \xi_{it}$$

where $\mathbb{E}[\xi_{it}|I_{it-1}] = 0$. Conditional on a first stage (possibly non-parametric) estimate for $\hat{\Phi}(\cdot)$, the conditional moment restriction in our second stage is

$$\mathbb{E}[\xi_{it} + \epsilon_{it}|I_{it-1}] = 0$$

$$\mathbb{E}\Big[y_{it} - f(l_{it}, k_{it}; \beta) - g\Big(\hat{\Phi}(k_{it-1}, l_{it-1}, m_{it-1}) - f(l_{it-1}, k_{it-1}; \beta)\Big) \mid I_{it-1}\Big] = 0$$

### 3.6.2 Computing Markups

As discussed, we use the SG&A accounting line item as our "material input" though this choice is not entirely innocuous. Qualitatively, SG&A should captures many administrative expenditures which ostensibly are necessary for the firm to function but contribute nothing *directly* to the to output if an excess are employed. With a fixed number of managers, one could easily imagine that the marginal product of each additional unit of labor or capital would remain strictly positive but might be diminishing. However, with a fixed number of productive units of labor and capital, an additional manager likely has near zero marginal product (subject to the Leontief threshold being reached).

Identification requires monotonicity in this quantity, an assumption that we cannot test directly, for $\omega$ is unobserved. However, we assess empirically whether SG&A has the claimed property. If it were not Leontief – for instance, if most of the spending were on advertisements and revenue were therefore likely to be increasing in SG&A spend – then we would naturally expect some dispersion: some firms may have a high ad spend but produce low quality products (spend relatively little on ads) while others might rely on word of mouth and produce higher quality products. In practice, though, SG&A, COGS, and revenue are very highly rank correlated.

To check, for each (SIC) industry-quarter we compute the within-industry percentile of SG&A and percentile of revenue. We plot this relationship in 3-7. The regression of the percentiles against one another gives the relationship

$$Percentile(Revenue) \ = \ 1.287 \ \times \ Percentile(SG\&A)$$

with an $R^2$ of 0.837.

Next, we estimate production functions separately for each SIC industry using the moment condition above. For tractability we assume Cobb-Douglas production technology. In practice we found that an adaptive grid search was computationally more stable than gradient-based methods. As a check, we estimated the Cobb-Douglas parameters separately by quarter and jointly across all quarters (but still unique to the SIC) and found that the resulting estimates differed only minimally. We prefer the former specification for the added flexibility: if an industry underwent significant changes over time the pooled estimation would mask them.

Given the production function estimates, we compute markups based upon the optimality condition of the firm's cost minimization problem (De Loecker and Warzynski (2012); De Loecker et al. (2020)). Without reproducing the derivation, we use the following expression to compute the markup:

$$\mu_{it} = \beta_{vs(i)} \frac{R_{it}}{C_{it}}$$

where $i$ indexes firms, $t$ indexes time, and $s(i)$ indicates firm $i$'s industry. Here $\beta_{vts(i)}$ is the output elasticity with respect to the variable input for industry $s(i)$ in period $t$; since we assumed Cobb-Douglas production technology this is just the parameter on the variable input. The quantities $R_{it}$ and $C_{it}$ are revenues and costs, respectively. Following De Loecker et al. (2020) we plot quarterly markups in Figure 3-8. Qualitatively these results match De Loecker et al., including the sharp increase in markups beginning in approximately the first quarter of 1999.

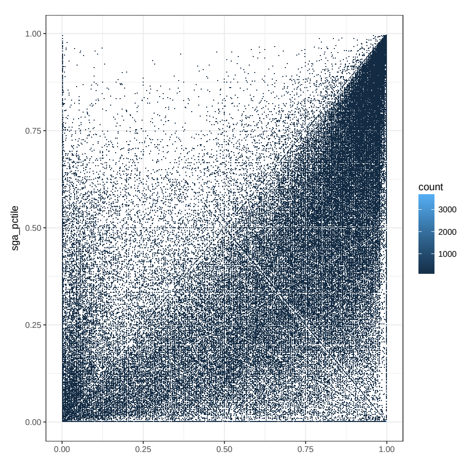



Figure 3-7: Revenue vs. SG&A percentiles, computed within SIC-quarter.

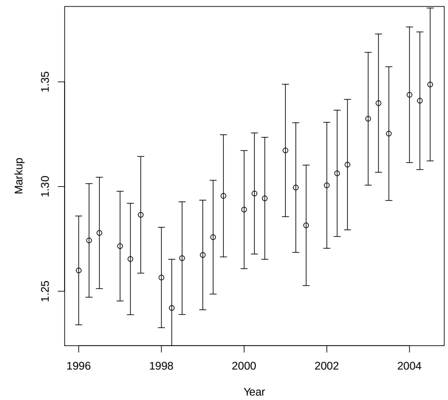



Figure 3-8: Mean quarterly markups aggregated across all industries. Error bars represent a 95% confidence interval around the mean.

### 3.6.3 Effect of Covenant Violations on Own Markups

With markups computed for every firm-quarter we can evaluate the effect of a covenant violation on a firm's own markups. We present the results in Tables 3.7 and 3.8. The former performs estimation on the full sample while the latter uses only the subset of firms with at least one covenant violation during our sample period. In both cases the sample is restricted to those firms for which we perform production function estimation and compute markups; this sample selection procedure is described in the previous section.

Our preferred specification in both cases is the within estimator using both firm and time fixed effects, specification (4). We control both for broad industry effects using the $\texttt{SICViolRev}_{it}$ variable, which is computed holding out firm $i$, as well as the firm's own distance to the default threshold. Adding fixed effects tends to reduce the point estimate, but even in our preferred specification we obtain a point estimate of approximately 2%. This reduction, relative to a baseline markup of $20-30\%$ for the typical firm, corresponds to a nearly 10% reduction in markup in relative terms.

The sign on the default distance is as expected after including firm fixed effects. Additionally, the percentage of (revenue-weighted) competitors in violation is consistent with the analysis that we perform in the following section, even when only estimating on the subset of violating firms and including the own violation indicator.

In total, we interpret this as strong evidence that firms in technical default become more competitive in the sense that they sharply reduce their markups in response to the default. This is consistent with the prediction of Proposition 3.2.1, the story being that firms in such a position get tough in the short term in an effort to ensure that they meet their debt obligations. In addition, the partial effect of competitors being in technical default is to reduce the firm's own markup, suggesting a feedback mechanism by which deteriorating industry state may be amplified depending upon capital structures on the constituent firms. The consistent narrative is that competition intensifies in the product market in response to a covenant violation by one or several firms in an industry.

Table 3.7: Effect of covenant violation on own markups. Estimation is done using the full sample, including firms that were never in technical default.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Markup | | | |
| | (1) | (2) | (3) | (4) |
| $\mathbb{1}\{\text{Own Violation}\}$ | $-0.057^{***}$ | $-0.056^{***}$ | $-0.021^{***}$ | $-0.020^{***}$ |
| | (0.006) | (0.006) | (0.003) | (0.003) |
| % SIC Rev. in Viol. | $-0.056^{***}$ | $-0.056^{***}$ | $-0.035^{***}$ | $-0.029^{***}$ |
| | (0.014) | (0.014) | (0.008) | (0.009) |
| Default Dist. | $-0.006^{***}$ | $-0.006^{***}$ | $0.001^{**}$ | $0.001^{**}$ |
| | (0.001) | (0.001) | (0.0004) | (0.0004) |
| Constant | $1.226^{***}$ | | | |
| | (0.003) | | | |
| Firm FE | No | No | Yes | Yes |
| Time FE | No | Yes | No | Yes |
| Observations | 109,988 | 109,988 | 109,988 | 109,988 |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | | |

Table 3.8: Effect of covenant violation on own markups. Estimation is done using the only the sample of firms that violated a covenant at least once during our sample period.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Markup | | | |
| | (1) | (2) | (3) | (4) |
| $\mathbb{1}\{\text{Own Violation}\}$ | $-0.029^{***}$ | $-0.029^{***}$ | $-0.018^{***}$ | $-0.015^{***}$ |
| | (0.006) | (0.006) | (0.003) | (0.003) |
| % SIC Rev. in Viol. | $-0.087^{***}$ | $-0.086^{***}$ | $-0.022$ | $-0.009$ |
| | (0.029) | (0.030) | (0.017) | (0.017) |
| Default Dist. | $-0.002^{*}$ | $-0.002^{*}$ | $0.003^{***}$ | $0.002^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Constant | $1.199^{***}$ | | | |
| | (0.004) | | | |
| Firm FE | No | No | Yes | Yes |
| Time FE | No | Yes | No | Yes |
| Observations | 19,996 | 19,996 | 19,996 | 19,996 |
| *Note:* | | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

### 3.6.4 Effect of Covenant Violations on Competitors' Markups

As a final exercise, we can examine companies that were never in violation of any covenants during our sample period but that nonetheless were bound by covenants. We can therefore flexibly control for distance to the technical default threshold but isolate the effect to those firms that never breached contractual agreements with their creditors. The full results are given in Table 3.9. The row of interest is the first — percent of revenue in violation — which is the percent of revenue within each industry-quarter generated by firms in technical default, relative to the total revenue for that industry-quarter. For this sample the quantity did not have to be computed with any firms held out, since the sample lacks violating firms. Otherwise computation is just as above.

We find that having firms in one's own industry that are in technical default tends to toughen competition: firms tend to reduce markups when product market competitors are in technical default. The result is consistent with Corollary 3.2.1.1 and the intuition follows that of the previous subsection. When a firm is in technical default and its creditors gain increased control rights, its incentives shift and the firm becomes more competitive (reduces price) in the product market in order to realize higher revenues in worse states of the world, which coincide with higher likelihoods of debt service default. In turn, since prices are typically strategic complements, the best response for competitors is to match the price decline. It is likely that competitors not in technical default and whose equity holders retain stronger control rights would respond to a lesser degree, which is what we find empirically, though the comparative static in general depends on structural primitives that are beyond the scope of our theoretical treatment.

To quantify the reduced form effect, note that the mean revenue in violation was 2.902%, aggregated across all industries and quarters. The median was 0.290% while the first and third quartiles were 0.000% and 2.809%, respectively. A typical markup over this period was around 1.3. Therefore, relative to the baseline, these effects are quantitatively small. However, markups grew by approximately 0.1 units over our period; relative to this baseline, and according to our preferred specification (4), a small firm in an industry with one third of all firms (weighted by revenue) in technical default faces a markup decline of about 10% relative to its decade-long secular increase.

Table 3.9: Effect of covenant violations on competitors in the same industry that were never in technical default.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Markup | | | |
| | (1) | (2) | (3) | (4) |
| % SIC Rev. in Viol. | $-0.059^{***}$ | $-0.036^{***}$ | $-0.057^{***}$ | $-0.029^{***}$ |
| | (0.014) | (0.008) | (0.014) | (0.009) |
| Def. Dist | $-0.003^{***}$ | $0.002^{***}$ | $-0.003^{***}$ | $0.002^{***}$ |
| | (0.001) | (0.0004) | (0.001) | (0.0004) |
| Constant | $1.212^{***}$ | $1.201^{***}$ | $1.206^{***}$ | |
| | (0.003) | (0.031) | (0.006) | |
| Firm FE | No | Yes | No | Yes |
| Time FE | No | No | Yes | Yes |
| Observations | 109,988 | 109,988 | 109,988 | 109,988 |
| F Statistic | $23.349^{***}$ | $44.755^{***}$ | $2.236^{***}$ | $16.654^{***}$ |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | | |

## 3.7 Discussion

The theoretical and empirical results present a consistent narrative describing the effect of firm control rights on both dynamic decisions (investment, R&D spending, &c.) as well as static product market competition. The simple but general model generated, under reasonable assumptions, unambiguous predictions for both policy dimensions that matched the evidence aggregated over a broad group of industries. In this section we succinctly summarize the results and discuss their implications as well as potential shortcomings.

### 3.7.1 Product Market Effects

Senior loans are typically among the least expensive financing options available to a firm, primarily because such creditors are contractually guaranteed to be paid back first and have a high likelihood of recouping payment even in adverse conditions. When conditions deteriorate, though, the agency problem often becomes more acute. If the firm is approaching failure, management, put in place by equity holders and often with aligned interest, may be inclined to make a last ditch effort to save the firm. Senior creditors, on the other hand, have less interest in the long term viability of the firm and more interest in generating short term cash flow to service the firm's debt. To mitigate this agency problem, many loan contracts include "covenants" linked to financial variables such that if certain accounting metrics deteriorate beyond a certain point then creditors may exercise partial control over the firm.[6]

As studied in Section 3.2.2, as the senior claimant, the creditor maximizes expected cash flows conditional on debt service default whereas the equity holder maximizes expected cash flows conditional on solvency (due to limited liability, equity holders owe and receive nothing in default). Therefore, creditors prefer more aggressive pricing in product markets than do equity holders, and a transfer of control rights to creditors, for instance in the case of technical default, are expected to reduce optimal markups. In response, if prices are strategic complements, product market competitors are likely to toughen competition as well.

Our results confirm these predictions and were remarkably consistent across a number of specifications estimated on different samples, consistently indicating that competition in an industry in which a large fraction of firms are in technical default is intense. Markups, returns on assets, and common equity values all decline in response to these covenant violations. We use three specific types of covenants in our analysis, though there is reason to believe that the absolute magnitudes of these effects would be amplified if we included a broader set of covenant types.

[6] Often the penalty is that the loan becomes callable. The firm, not wanting or able to service or refinance the loan, negotiates with creditors and informally transfers control rights.

### 3.7.2 Dynamic Effects

Overall, our estimates indicate that in response to a control rights transfer to creditors, firms reduce their investment. This result is consistent with the predictions of our basic model: investment decisions are inherently dynamic and influence both the mean and variance of future outcomes. As residual claimants on profit, shareholders derive value from both aspects, while creditors only care about the probability that they will be repaid. When control rights transfer, therefore, the level of investment moves from the shareholders' optimum to the creditor's optimum and empirically tends to decline.

Moreover, independent of the direction of the change, we should unambiguously expect the value of common equity to discontinuously decline when control rights are transferred to creditors, since the level of investment (and other choice variables) moves away from the shareholders' optima, which maximized their own value function. Empirically we found that the declines were largest for the *ex ante* most successful firms (in terms of net worth and return on assets), which is again consistent with intuition. For high growth and/or margin firms, the upside of deepening investment is higher so there is likely to be a wider gap between the shareholder and creditor optima.

### 3.7.3 Correlated Shocks

The primary concern we have concerning the validity of our estimates is that correlated demand shocks may be prevalent and significant. One potential mechanism by which these could confound our results is as follows. A negative shock to demand impairs the financial position of the firm by reducing current period cash flows while also reducing the net present value (NPV) of forward-looking projects. These effects yield disproportionately high levels of covenant violations (pegged to financial variables) and also weigh on the equity value and reduce the marginal value of investment. The demand shocks are correlated at the industry level, so the effect will hit competitors as well.

We believe that our empirical strategy sufficiently mitigates these concerns. Suppose indeed that common shocks were present, and that they do behave as described. Then the outcome (equity value, capital investment, R&D expenditure) should be correlated with continuous markers of financial health and market status. Moreover, the effect should be smooth and monotone: barring contractual stipulations, there is no reason to expect that any particular margin induces a discontinuous effect on the outcome (more strongly: on all outcomes, at the same threshold). Our empirical specifications test the categorical effect at such cutoffs and include continuous controls.

Further, managers have little incentive to cross the covenant threshold. As we have found, equity loses significant value upon a covenant violation. The incentive, then, is to take action to avoid a covenant violation, which might include such actions as cutting investment to improve the short term financial status of the firm. As a consequence, endogenous managerial effects should if anything work in the opposite direction of our effects (investment being cut locally on the *non-violating* side). We contend, therefore, that these effects are unlikely to bias our estimates in the direction of the effect.

### 3.7.4 Extensions

While this work firmly establishes a reasonably robust set of reduced form effects, additional work could help to empirically confirm the narrative and to identify potentially heterogeneous effects based on industry and firm characteristics.

One reasonable extension would be to identify the subset of industries that primarily sell durable goods and re-run the same set of specifications on this sample of firms. We should expect that the competitive effects would be even more pronounced if indeed creditors optimize for short term capacity to repay. For durables producers, reducing price in the current period will increase quantity but a large component of the increase will be due to early replacement, so later quantities will decline. A firm in such an industry facing an upward-sloping marginal cost schedule (e.g. from running machines without enough downtime) can increase its short-term performance by cutting price at the cost of long-term performance. Yet this may very well be optimal for creditors holding short maturity debt. For select, well-studied markets we may also be able to evaluate price effects due to these technical defaults. This analysis would require a precise structural model of the demand side, however, so might only be possible in select industries.

A number of additional extensions are immediate. First, if we restrict attention to long-lived firms for which there is reliable data, we may test the persistence of the effects. We could also, for publicly traded names, test whether equity values appropriately respond to the covenant announcement. Additionally, to take our hypothesis seriously, we could instrument in some other way for debt load (rather than directly for control rights) and test whether it has similar bearing on the rest of the industry. Finally, we could instrument for firm-level variables with our covenant indicator as a first stage, and in the second state test effects on other firms in the industry (appropriately lagged). This exercise would permit us to study the mechanism by which the violating firm becomes tougher in the product market.

## 3.8 Conclusion

In this paper we study the product market impacts of shifts in firm control rights. These rights govern which agents have the power to make decisions affecting the firm and are derived from a combination of formal and informal mechanisms. The shifts we study are from equity holders to creditors, who have different payoff structures and correspondingly different preferences, and in particular we study shifts triggered by loan covenant violations (technical defaults).

We join an emerging literature on loan covenants and their impacts on firm decision making as well an older vein of the industrial organization and corporate finance literatures studying strategic capital structure choice and its impacts on product markets. Our work merges the two to evaluate the impact of covenant violations on product markets. Given that covenant violations are more likely to be triggered when broader economic conditions are deteriorating, this work suggests an additional mechanism in the real-financial linkage.[7]

Work on covenant-related linkages has proliferated in recent years and we anticipate that it will continue to be a fruitful area of inquiry. In the midst of the COVID-19 pandemic, firms often turned to corporate loans, many of which included restrictive financial covenants, as their first line of financing to bridge the immediate liquidity crisis (Ng (2020)). As the recent literature has established, the contractual structure of these agreements has significant implications for the real economy, particularly in a protracted downturn. As the relative size of corporate loan financing continues to grow, including a significant risky leveraged segment, additional work to further understand the impacts and risks will only prove more essential (Faria-e Castro and Bharadwaj (2019); Wirz and Timiraos (2020)).

[7] There are multiple mechanisms by which financial shocks impact the covenants, violations, and creditor responses which are beyond the scope of the paper. As one example, consider the lender health channel studied by Chodorow-Reich (2014); Chodorow-Reich and Falato (2020, forthcoming). Given the financial shocks and the covenant impacts, our main focus is on the resulting product market impacts.

# Appendix A

# Appendix for Chapter 1

## A.1 Omitted Proofs

*Proof of Lemma 1.4.1.* Let $\preceq$ represent the product order over $\mathbb{R}^2$, $|\cdot|$ represent the coordinatewise absolute value, and let $x, y \in [0,1]^2$ each represent price vectors for notational convenience, hence both the domain and codomain are compact. By construction $x \preceq y + |x-y|$ hence $N(x) \preceq N(y+|x-y|) \preceq N(y) + \beta|x-y|$ for some $\beta \in (0,1)$ where the first relation follows from monotonicity and the second from the bound on the derivative applied coordinatewise. By the same argument, $N(y) \preceq N(x) + \beta|y-x|$, thus squeezing the coordinates such that $|N(x) - N(y)| \preceq \beta|x-y|$. Let $\|\cdot\|$ be any norm. By subadditivity and absolute homogeneity of norms, the elementwise comparison thus gives $\|N(x)-N(y)\| \leq \beta\|x-y\|$ thus the map $N$ is a contraction under metric space induced by the norm. By the Banach Fixed Point Theorem, $N$ admits a unique fixed point. □

*Proof of Corollary 1.4.1.2.* Proceed by contradiction and suppose $\sigma_p^A(p) \geq \sigma_p^B(p)\ \forall p$ but $p^{A\star} > p^{B\star}$. Then decrementing $p^A$ and incrementing $p^B$ both by $\epsilon > 0$ gives incremental log profit

$$\epsilon \cdot \left[\sigma_p^A(p^A) - \sigma_p^B(p^B)\right] \geq \epsilon \cdot \left[\sigma_p^A(p^B) - \sigma_p^B(p^B)\right] \geq 0$$

where the first inequality follows from Assumption 1.4.2, a contradiction, since $(p^{A\star}, p^{B\star})$

must not have been optimal. □

*Proof of Lemma 1.4.2.* Beginning with expression (1.4.6), multiply through by $p^i$ to get a price semielasticity on the right hand side:

$$p^i + p^j - c = (\sigma_p^i)^{-1}\kappa$$

where $\kappa = \frac{1-\eta_N^i \eta_N^j}{(\alpha_B + \alpha_S \eta_N^j)}$. Totally differentiating then gives

$$\frac{dp^i}{dp^j} = -\left(1 + \frac{\kappa}{(\sigma_p^i)^2} \cdot \frac{\partial \sigma_p^i}{\partial p^i}\right)^{-1}$$

By Assumption 1.4.2, $\frac{\partial \sigma_p^i}{\partial p^i} \geq 0$, establishing the result. □

## A.2 Model Variants

### A.2.1 Full Rationality, Constant Elasticity

In this section, we consider the classical model in which all agents are fully rational hence it doesn't matter on which side the tax is levied. Consistent with the main model and our empirical setting, we assume that the platform, facing supply and demand relationships, chooses ad valorem fees, which we assume are applied to the demand side of the market. This assumption is without loss of generality because fees in this classical setting, as we will show. Additionally, for tractability, we assume that price elasticities of supply and demand are both constant. Therefore, given platform fees $\tau$ and price elasticities of demand and supply $\epsilon_D, \epsilon_S$, our demand and supply relations satisfy

$$
\begin{aligned}
D(p) &= A_D[p(1+\tau)]^{-\epsilon_D} \\
S(p) &= A_S p^{\epsilon_S}
\end{aligned}
$$

where $A_D, A_S$ are scale constants. In equilibrium the market is balanced, giving a unique equilibrium price as a function of demand primitives and the tax:

$$
\begin{aligned}
D(p) &= S(p) \\
\implies p^* &= \left(\frac{A_S}{A_D}\right)^{\frac{-1}{\epsilon_D+\epsilon_S}} (1+\tau)^{\frac{-\epsilon_D}{\epsilon_D+\epsilon_S}}
\end{aligned}
$$

As a function of the platform fee $\tau$, the platform's log profits (evaluated at equilibrium quantities) are

$$
\begin{aligned}
\ln R(\tau) &= \ln(S^* p^* \tau) \\
&= C - \frac{\epsilon_D(1+\epsilon_S)}{\epsilon_D+\epsilon_S}\ln(1+\tau) + \ln\tau \qquad C := \ln A_S - \frac{1+\epsilon_S}{\epsilon_D+\epsilon_S}\ln A_S/A_D
\end{aligned}
$$

hence the revenue maximizing tax $\tau^*$ satisfies

$$\tau^* = \frac{\epsilon_D + \epsilon_S}{\epsilon_S(\epsilon_D - 1)}$$

When the platform has a marginal cost proportional to monetary volume $\omega$, e.g. in the case of interchange fees, the optimal fees net of costs satisfy

$$\tau^* - \omega = (1+\omega)\frac{\epsilon_D + \epsilon_S}{\epsilon_S(\epsilon_D - 1)}$$

In the limiting case where supply is perfectly competitive and has constant marginal cost $c$ we have $\epsilon_S \to \infty$ in which case

$$\tau^* = \lim_{\epsilon_S \to \infty} \frac{\epsilon_D + \epsilon_S}{\epsilon_S(\epsilon_D - 1)} = (\epsilon_D - 1)^{-1}$$

by L'Hôpital. Since $\tau$ is a an ad valorem tax, let $\tilde{p} \equiv (1+\tau)p$ be the all-in price, in which case we obtain the standard Monopolist markup

$$\begin{aligned} \tau = \frac{\tilde{p} - c}{c} &= (\epsilon_D - 1)^{-1} && p = c \\ \implies \frac{\tilde{p} - c}{\tilde{p}} &= \epsilon_D^{-1} \end{aligned}$$

*Remark.* In this full rationality model, changing the statutory incidence of the tax has no impact on equilibrium quantities. Writing

$$D(p) = A_D \hat{p}^{-\epsilon_D}$$
$$S(p) = A_S[\hat{p}(1-\tau)]^{\epsilon_S}$$

and noting that $p = (1+\tau)\hat{p}$ yields an identical expression for the equilibrium price, as expected. In the full rationality model, therefore, there is no scope for the platform to optimize its *composition* of fees, only its levels.

## A.3 Elasticity Estimation

The core of this work relies on accurately estimating the impact of experimental variation in some fee level on an outcome of interest. In all cases our theoretical results suggest that we are interested in elasticities in semielasticities, therefore a natural approach would be to take logs of each side and run OLS. However, this procedure does not consistently recover the conditional expectation $\mathbb{E}[Y|X]$ except under a restrictive structure on the error term — approximately, the variance of the error must scale at exactly $\exp\{2X'\beta\}$, as shown by Silva and Tenreyro (2006). The intuitive result is an application of Jensen's inequality and in general we should expect that $\mathbb{E}[\exp\{\hat{y}\}] \leq y$ when the log-linearized equation is estimated by OLS. The nonlinear least squares estimator

$$\min_b \sum_i [y_i - \exp\{x_i b\}]^2$$

is consistent but may be inefficient since it adds weight $\exp\{x_i b\}$ to each observation, significantly increasing weight on large observations. Alternatively, Silva and Tenreyro (2006) suggest using an estimator based on the assumption that the variance is approximately proportional to the outcome, $V[y_i|x_i] \propto \mathbb{E}[y_i|x_i]$, giving

$$\sum_i [y_i - \hat{y}_i] x_i \ = \ 0$$

where $\hat{y}_i = \exp\{x_i\hat{\beta}\}$. This expression is also the estimator based upon the conditional moment $\mathbb{E}[\epsilon|X] = 0$ when the estimating equation is $Y = \exp\{X\beta\} + \epsilon$.

In practice we estimate both OLS and the nonlinear least squares estimator above and in our setting the two tend to return similar estimates. We rely on the fact that the above estimator coincides exactly with the (pseudo) Poisson maximum likelihood estimator with exponential link, which is intuitive given the Poisson variance property. This simplifies estimation since most statistical languages include rich library support for Poisson regression (in the R language we use the `glm` package).

*Remark.* The pseudo Poisson regression should give identical results as logistic regression (binomial with logit link) in the sense that the conditional expectations should match. The parameter estimates will differ, though, due to the differing link functions. We confirm this empirically and rely on convention when deciding between the two. A proof follows.

*Proof.* Starting with the PMF we can derive the pseudo ML estimator of the Poisson regression:

$$\begin{aligned}
f_{pois}(y_i; \lambda = e^{x_i\hat{\beta}}) &= \frac{\exp\{y_i x_i \hat{\beta} - e^{x_i\hat{\beta}}\}}{y_i!} \\
l(y|x, \hat{\beta}) &= \sum_i [y_i x_i \hat{\beta} - \exp\{x_i\hat{\beta}\} - \log(y_i!)] \\
\frac{\partial l(\cdot)}{\partial \hat{\beta}} = 0 &= \sum_i [y_i x_i - x_i \exp\{x_i\hat{\beta}\}] \\
&= \sum_i [y_i - \hat{y}_i] x_i
\end{aligned}$$

This coincides exactly with the maximum likelihood logit estimator (i.e. binomial with sigmoid link), modulo the link function, characterized by the following

$$\begin{aligned}
Pr(y_i = 1; x_i\beta) &= \hat{y}_i = \sigma(x_i\beta) = (1 + e^{-x_i\beta})^{-1} \\
l(y|x, \hat{\beta}) &= \sum_i \log[\hat{y}_i^{y_i}(1 - \hat{y}_i)^{1-y_i}] \\
&= \sum_i y_i[\log e^{x_i\hat{\beta}} - \log(1 + e^{x_i\hat{\beta}})] + (1 - y_i)[\log(1) - \log(1 + e^{x_i\hat{\beta}})] \\
&= \sum_i [y_i \log e^{x_i\hat{\beta}} - \log(1 + e^{x_i\hat{\beta}})] \\
\frac{\partial l(\cdot)}{\partial \hat{\beta}} = 0 &= \sum_i [y_i - \hat{y}_i] x_i
\end{aligned}$$

□

*Remark.* Event-level heterogeneity is significant in our setting and as such we would like to use a within estimator, estimating the marginal impact of fees on outcomes of interest net of mean event effect. However, we have a short panel as sale periods for individual events are limited in length and therefore adding fixed effects to nonlinear models generally introduces the incidental parameters problem. Fortunately, both the Poisson and logistic regressions that we intend to run have conditional likelihoods that are independent of the incidental parameters (Lancaster (2002)).

# Appendix B

# Appendix for Chapter 2

## B.1 Quality Index Construction

We develop a vertical quality index separately from the core demand model, motivated by two empirical truths: (i) the feature space describing goods in our setting is high-dimensional and (ii) the number of *potential* transactions (choices) is several orders of magnitude larger than the number of *realized* transactions on the platform. To address (i) we make use of machine learning techniques for automated feature selection, including interactions between features and flexible non-linearities. In response to (ii) we focus only on realized transactions when constructing the quality index. As with many online commerce settings, despite large choice sets, consumers may only seriously consider some much smaller subset of goods — those that are competitively priced or match their tastes most closely.[1] *Ex ante*, without the demand model, we have no way to determine which elements from the choice set are "relevant" and cannot sufficiently reduce choice sets to make the problem computationally tractable. However, transacted goods must have been in this set, and subject to an identifying assumption we can consistently recover vertical quality indices from the transaction data alone.

The "quality index" developed in this section bears close resemblance to the he-

[1] Some goods are easily eliminated from the choice set, e.g. those that are strictly dominated by another listing in the same row of the venue at a lower price. However, these *a priori* restrictions on choice sets fail to sufficiently address the dimensionality problem.

donic model of Rosen (1974) and further studied by Bajari and Benkard (2005). The kernel approach employed draws inspiration from the Nadaraya-Watson estimator (Nadaraya (1964); Watson (1964)) as well as the "kernel trick" and the more general "representer theorems" in the machine learning literature (Schölkopf et al. (2000); Rasmussen and Williams (2006)). Matzkin (2003) studies an application of kernel methods to hedonic models. We amend these approaches in two ways due to empirical concerns. First, we slightly relax the assumptions of the hedonic model and aggregate across markets.[2] Second, because the feature space is much too large for us to estimate the price surface by kernel methods alone, we nest a nonlinear regression inside a lower-dimensional kernel and minimize the empirical mean squared error directly.

**Setup**

We maintain the generic description for the utility of consumer $i$ choosing seat $j$ at time $t$ from the full demand model:

$$u_{ijt} = v_j - f(p_{jt}) + \epsilon_{ijt}$$

The identifying assumption is that markets are sufficiently thick that choice sets do not change "too quickly" and the inclusive value is locally (temporally) stable conditional on consumer type.[3] Concretely, in the language of the demand model, consider that two consumers $i$ and $i'$ of the same type $\theta$ arriving sequentially in some neighborhood of time $t$ and choosing products $j$ and $j'$, respectively. Although consumer $i'$ cannot purchase the exact good that $i$ has already purchased (seats are globally unique), $i'$ has sufficiently many similar seats from which to choose, including those that may have been newly added in the intervening time, that $i$ and $i'$ have

[2] Identification in the hedonic model is typically predicated upon a large number of products per market having strictly positive demand. Empirically, however, we have at most one product per market with positive demand. Adding time or a "market index" as a product feature could bridge the two approaches but requires a "smoothness" assumption about how utility varies with time, as we discuss.

[3] This assumption is analogous to Lipschitz continuity of $u$ in $t$ though that level of formalism is not a focal point of this section.

approximately equal inclusive values. The inclusive value is the expected utility conditional on purchase (the $n^{th}$ order statistic), so the following holds approximately

$$\begin{aligned}
\mathbb{E}[u_{ijt}|i \text{ purchase } j] &\approx \mathbb{E}[u_{i'j't+\Delta t}|i' \text{ purchase } j'] \\
\mathbb{E}[v_j|\text{purchase}] - f(p_j) + \mathbb{E}[\epsilon_{ijt}|\text{purchase}] &\approx \mathbb{E}[v_{j'}|\text{purchase}] - f(p_{j'}) + \mathbb{E}[\epsilon_{i'j't}|\text{purchase}] \\
\implies \mathbb{E}[v_j - v_{j'}|\text{purchase}] &\approx f(p_j) - f(p_{j'}) \qquad \text{(B.1.1)}
\end{aligned}$$

Quantifying relative qualities therefore reduces to choosing a link function $f$ and comparing relative transaction prices for two seats that sold around the same time for a given event; this is the standard kernel approach.[4]

The estimation target is a function $v : X \to \mathbb{R}$ mapping seat features into a scalar vertical quality index. Under the root mean squared error (RMSE) objective, the functional to be optimized is

$$L(v;p) = \frac{1}{2} \sum_{g \in G} \sum_{k \in g} w_k \left( u_k - \sum_{j \in g \setminus \{k\}} \omega_{kj} u_j \right)^2$$

where $k, j$ index transactions, $w$ are observation weights,[5] $u_k \equiv v(x_k) - f(p_k)$, and $\omega$ is the (scalar) kernel weight. The outer sum is over events/games, indexed by $g \in G$, as we consider each game to be a separate market (cf. Figure 2-3). For transactions within a game, which index the second sum, the summand penalizes utility deviations from the weighted mean utility of similar transactions.[6] We assume that qualities are identical across a physical section-row within a stadium, so for instance seats 13 and

[4] A complication in the analysis is that inclusive values for a given choice set differ by consumer type and the assumption above that the two buyers were of the same type is likely to be frequently violated in practice. However, given any choice set, any two buyers who choose seats of similar quality (especially if the seats are spatially close as well) are likely to be of similar types in expectation. The quality model therefore relies on kernel methods to quantify "closeness" between any two transactions, both on the basis of (i) choice sets being similar and (ii) consumer types being similar, describing the degree to which (B.1.1) holds in expectation.

[5] Weights are most reduced for thin markets with relatively small choice sets, in which case the inclusive value assumption is most likely to be violated. Typically this corresponds to transactions for odd quantities, such as a school group shopping for thirty tickets together, or transactions for niche items, such as single tickets in premium sections.

[6] The indexing does not restrict the comparison only to past transactions: the model is trained after an event is complete, so sequential transaction 200 is compared with both transactions 199 and 201 (weighted by their similarities).

14 in section 201 and row B of Fenway Park are assumed to have identical qualities. Since the loss function is linearly separable in games, we can write the loss for game $g$ in matrix form as

$$L_g(v; p) = \frac{1}{2} \left\| W_g^{1/2} (I - \Omega_g)(P_g - M_g V) \right\|_2^2 \tag{B.1.2}$$

where $W_g$ is a diagonal matrix of observation weights, $\Omega_g$ is a matrix of cross-observation weights, $P_g$ is a vector of observed log prices, $V \equiv V(X)$ are (to be estimated) quality indices for each section-row in the venue, and $M_g = [\delta_{r',r(i)}]_{i,r'}$ is a matrix of indicators for the row in the venue to which each observation corresponds; $\delta_{\cdot,\cdot}$ is the Kronecker delta. We define $\Omega_g$ to be the product over five exponential kernels, in (i) time, (ii) row quality (estimated), (iii) transaction quantity (number of seats), (iv) zone within venue, and (v) section within zone . Entry $(i, j)$ in $\Omega_g$ has value

$$\Omega_{g,(i,j)} = \kappa(x_i, x_j) \equiv \prod_d \exp\{-\vartheta_d \rho_d(x_i, x_j)\} \tag{B.1.3}$$

where $\kappa(\cdot, \cdot)$ is the kernel function and is positive definite, $x_i$ is a feature vector describing the section-row purchased in transaction $i$, $\rho_d$ is a distance metric along dimension $d$, and $\vartheta_d$ is a kernel weight.

**Estimation**

Given the objective (B.1.2) and definitions for all objects aside from $R \equiv v(X)$, $L_g : v \to \mathbb{R}_+$ is a well defined functional that we can optimize directly. Intuitively, we expect that a combination of nonlinearities and interactions in $v$ capture important aspects of quality, but *ex ante* we do not know exactly which to include. However, as we search over a richer class of functions the search becomes more difficult. As a solution, we make use of a "greedy search" approach using gradient boosted regression trees (GBTs) and modest regularization. The "greedy search" builds layers and trees sequentially maximize the incremental loss reduction due to each operation (a "split"). Though the greedy search cannot in general guarantee global optimality, in practice it often performs well.[7] In our case, with GBTs, the required inputs are gradients and Hessians computed with respect to *predictions* and conditional on some baseline predictions generated by the preceding trees. To solve for these quantities, first rewrite the objective to obtain

$$L(v;p,g) = \frac{1}{2}R'\left(\sum_g M_g'\Sigma_g M_g\right)R \;-\; \left(\sum_g Y_g'\Sigma_g M_g\right)R \;+\; \frac{1}{2}\sum_g Y_g'\Sigma_g Y_g \qquad \text{(B.1.4)}$$

where

$$\Sigma_g = (I - \Omega_g)'W_g(I - \Omega_g)$$

hence the gradient and Hessian are

$$\text{Grad} = R'\left(\sum_g M_g'\Sigma_g M_g\right) \;-\; \left(\sum_g Y_g'\Sigma_g M_g\right) \qquad \text{(B.1.5)}$$

$$\text{Hess} = \sum_g M_g'\Sigma_g M_g \qquad \text{(B.1.6)}$$

Given these quantities, optimal sequential tree splits take a convenient closed form (Chen and Guestrin (2016)) and training proceeds as follows. First, define feature primitives (i.e. raw features excluding interactions and nonlinearities) and build a design matrix $X$ and response vector $Y$. At each iteration $i$, accumulate the current prediction $\hat{R}$ by summing leaf weights for each observation. Compute the gradient and Hessian using $\hat{R}$ and evaluate a candidate split over a grid along each feature dimension, selecting the split that reduces the loss function by the greatest amount (incorporating regularization penalties). Update $\hat{R}$ and proceed to iteration $i+1$.[8]

For row quality estimation we use the `xgboost` library with the custom gradient (B.1.5) and Hessian (B.1.6) programmed in `C++` and dynamically linked. As both require several auxiliary matrices for computation we adapted the default bindings into the `R` language to accommodate and efficiently marshal the additional data.

[7] A similar problem applies to artificial neural networks with respect both to network structure construction and to training deep networks, as in Bengio et al. (2007)

[8] We omit discussion of the nuances of training, including how observations are aggregated as each tree grows depthwise, various approximations to speed split evaluation, and certain regularization approaches. For a more detailed description of training, refer to Chen and Guestrin (2016).

The above describes one iteration of the "inner loop" of our optimization procedure and gives estimates conditional on some kernel $\Omega$. Given that the kernel is a function of the row quality estimates themselves, estimation iterates through an "outer loop" with each iteration constructing the kernel with the quality estimates generated by the previous iteration. We initialize qualities with a simple log price regression which gives baseline row quality estimates that we use to construct our initial kernel matrix $\Omega_g^{(1)}$ for each event.[9] We then run one complete "outer" training iteration, yielding updated estimates for row qualities and giving rise to an updated $\Omega_g^{(2)}$. We complete this process through five outer iterations, at which point the estimates converge to an acceptable tolerance, which we set to a one log point root mean square deviation in estimated quality aggregated over all rows.

Hyperparameters include five kernel parameters $\vartheta_k$ plus seven regularization parameters. We perform an adaptive grid search across all twelve dimensions to calibrate these parameters. In each case we use five-fold cross validation to generate selection metrics. The evaluation metric under model $v(\cdot)$ is the one transaction ahead deviation between actual and predicted transaction price conditional on observed seat features. Quantitatively, this metric reduces to (B.1.4) with two adjustments: (i) we exclude the third term, which is independent of $R$ and (ii) we restrict $\Omega_g$ to be lower triangular, where rows and columns are ordered by temporal transaction index. We select the hyperparameter grid point with the best evaluation metric as well as points giving the best single-coordinate deviation along each dimension, yielding a cube, and repeat the process with a finer grid constrained to that cube.

[9] The regression is of the form

$$\ln p_{jt} = \beta_{section} + \beta_{row} \times \text{row\_rank}_j + \beta_{row2} \times \text{row\_rank}_j^2 + \beta_{game} \times f(t) + \epsilon_{jt}$$

where section and game are indicators, $f(\cdot)$ is a natural cubic spline with five knots at equally spaced empirical time quantiles, and row_rank is the row number within the section (1 indexed), for which we use an orthogonal polynomial basis.

## B.2 Maximum Likelihood Estimator Details

**Algorithmic Approach** In general, by the Karush-Kuhn-Tucker theorem, maximizing $f$ subject to equality constraints $g$ satisfies the stationarity condition at the optimum $(\zeta^*, \phi^*)$

$$\nabla\mathcal{L}(\zeta^*, \phi^*, \lambda_1; \lambda_2) = \nabla f(\zeta^*, \phi^*) - \lambda_1 \nabla g(\zeta^*, \phi^*) = 0$$

If we let $\lambda_1(\zeta^*, \phi^*)$ solve the stationary condition and define

$$F(\zeta, \phi) \equiv \nabla f(\zeta, \phi) - \lambda_1(\zeta, \phi)\nabla g(\zeta, \phi)$$

then the problem amounts to finding a zero of $F$. Direct gradient descent of $F$ over $(\zeta, \phi, \lambda_1)$ tends not to converge to the appropriate saddle point solving (2.3.6) (Platt and Barr (1988)), so we take the typical approach of using Newton's method, but with three adjustments that yielded practical benefits. First, we substitute and eliminate $\lambda_1(\zeta, \phi)$ based upon a necessary optimality condition. Second, we enable backtracking on each Newton-Raphson (NR) step to mitigate known numerical difficulties when solving the constrained problem, such as the Maratos effect — if the objective degrades or improves by less than a fixed tolerance, we recursively halve the step along all dimensions until an improvement is obtained.[10] Finally, we ensure that the NR steps are well behaved away from the optimum by constraining the Hessian to be negative definite. We achieve this by computing the eigendecomposition at each step and adjusting eigenvalues as necessary.

**Analytic Expressions** Each NR step requires us to compute the Jacobian of $F$, giving the update at step $t$

$$\Delta_t(\zeta, \phi)' = -D_F(\zeta_t, \phi_t)F(\zeta_t, \phi_t)$$

[10] During estimation, we seem numerically to obtain quadratic convergence in a neighborhood of the optimum; backtracking and Hessian adjustments were only necessary in early stages.

We now describe the terms of the vector valued $F$, which require some care to derive, but from which it is then immediate to compute the Jacobian. It shall prove analytically convenient to use the following definition and subsequent expressions. Let $s^*_{im}$ denote the posterior probability that consumer $i$ is of type $\theta_i = \theta_m$, $\mathbb{P}(\theta_m | l(i); J(i), \zeta, \phi)$:

$$s^*_{im} = \frac{s_m \mathbb{P}_{il(i),\theta_m}}{\sum_n s_n \mathbb{P}_{il(i),\theta_n}} = \frac{\exp\left(\phi_m + \ln \mathbb{P}_{il(i),\theta_m}\right)}{\sum_n \exp\left(\phi_n + \ln \mathbb{P}_{il(i),\theta_n}\right)} \tag{B.2.1}$$

where the second equality comes from the definition (2.3.2). Differentiating on element of (2.3.5) in the parameter vector $\zeta$ and using (B.2.1) gives the derivative

$$\frac{\partial \ell_i}{\partial \zeta} = \sum_m s^*_{im} \frac{\partial \log \mathbb{P}_{il(i),\theta_m}}{\partial \zeta} \tag{B.2.2}$$

where

$$\begin{aligned} \log \mathbb{P}_{ij,\theta_m} \quad &= \quad \exp(-\alpha\theta_m)\left(x_j\beta - \exp(\theta_m + \gamma)p_j\right) \\ &\quad - \log \sum_{k \in J(i)} \exp\Big\{ \exp\left(-\alpha\theta_m\right)\left(x_k\beta - \exp(\theta_m + \gamma)p_k\right) \Big\} \end{aligned}$$

Similarly, the elementwise derivatives in the type parameters $\phi$ are

$$\frac{\partial \ell_i}{\partial \phi_k} = \sum_m s^*_{im} \frac{\partial \log s_m}{\partial \phi_k} \quad = \quad s^*_{ik} - s_k \tag{B.2.3}$$

Summing (B.2.2) and (B.2.3) over observations $i$ gives direct maximum likelihood first-order conditions. Applying these results we obtain the following elements of $F$

$$\begin{aligned} \left.\frac{\partial \mathcal{L}}{\partial \zeta}\right|_{\lambda_1} &= \sum_i w_i \sum_m s^*_{im} \frac{\partial \log \mathbb{P}_{il(i),\theta_m}}{\partial \zeta} \\ \left.\frac{\partial \mathcal{L}}{\partial \phi_k}\right|_{\lambda_1} &= \sum_i w_i (s^*_{ik} - s_k) \\ &\qquad - \lambda_1 s_k \left(\theta_k - \sum_m s_m \theta_m\right) \\ &\qquad - \lambda_2 \left[\left(\phi_k - \frac{\phi_{k-1} + \phi_{k+1}}{2}\right) - \frac{1}{2}\left(\phi_{k-1} - \frac{\phi_{k-2} + \phi_k}{2} + \phi_{k+1} - \frac{\phi_k + \phi_{k+2}}{2}\right)\right] \end{aligned}$$

There is no entry corresponding to $\lambda_2$ because this is an *ad hoc* regularization parameter. At the optimum $F = 0$ so we can solve for $\lambda_1$ by multiplying the latter expression by $\theta_k$ and summing to obtain

$$\begin{aligned}
0 &= \sum_k \theta_k \sum_i w_i (s_{ik}^* - s_k) - \lambda_1 \sum_k \theta_k s_k (\theta_k - \bar{\theta}) & \bar{\theta} \equiv \sum_k s_k \theta_k \\
&= \sum_k \theta_k \sum_k w_i s_{ik}^* - \bar{\theta} \sum_i w_i - \lambda_1 \left( \sum_k \theta_k^2 s_k - \bar{\theta}^2 \right) & \\
\lambda_1 &= \frac{\sum_k \theta_k \sum_k w_i s_{ik}^*}{\sum_k \theta_k^2 s_k} & \bar{\theta} = 0 \quad (\text{FOC}_{\lambda_1}) \\
&\approx \frac{\sum_k s_k \theta_k}{\sum_k s_k \theta_k^2} \sum_i w_i = \lambda_1^* & \sum_i w_i s_{ik}^* \to s_k
\end{aligned}$$

The final approximation follows from the convergence condition that the posterior and prior shares match. At the optimum the $\theta$ distribution should be exactly centered at zero, so $\lambda_1^* = 0$ and the constraint is non-binding. We separately estimate the mean elasticity with the parameter $\gamma$. Substituting this expression into $F$ yields the following adjusted entry where we explicitly account for the effect of $(\zeta, \phi)$ on the constraint

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \zeta} &= \left. \frac{\partial \mathcal{L}}{\partial \zeta} \right|_{\lambda_1} - \frac{\partial \lambda_1^*}{\partial \zeta} \sum_m s_m \theta_m \\
\frac{\partial \mathcal{L}}{\partial \phi_k} &= \left. \frac{\partial \mathcal{L}}{\partial \phi_k} \right|_{\lambda_1} - \frac{\partial \lambda_1^*}{\partial \phi_k} \sum_m s_m \theta_m
\end{aligned}$$

We present closed-form expressions for all required elements of gradient and Hessian below.

The required components of the gradient $(F)$ are

$$
\begin{aligned}
d_{ij,\theta_m} &= \exp\left(-\alpha\theta_m\right)\left(x_j\beta - \exp\left(\theta_m+\gamma\right)p_j\right) \\
\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \beta} &= \exp\left(-\alpha\theta_m\right)\left(x_{l(i)} - \sum_j \mathbb{P}_{ij,\theta_m} x_j\right) \\
\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \gamma} &= -\exp\left(\left(1-\alpha\right)\theta_m+\gamma\right)\left(p_{l(i)} - \sum_j \mathbb{P}_{ij,\theta_m} p_j\right) \\
\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \alpha} &= -\theta_m\left(d_{il(i),\theta_m} - \sum_j \mathbb{P}_{ij,\theta_m} d_{ij,\theta_m}\right) \\
\frac{\partial \lambda_1^*}{\partial\left(z_k\phi\right)} &= \frac{s_k\left(\theta_k-\sum_m s_m\theta_m\right)}{\sum_m s_m\theta_m^2}\sum_i w_i - \lambda_1^*\frac{s_k\left(\theta_k^2-\sum_m s_m\theta_m^2\right)}{\sum_m s_m\theta_m^2} \\
&= s_k\theta_k\frac{\sum_i w_i-\theta_k\lambda_1^*}{\sum_m s_m\theta_m^2} \\
\frac{\partial \lambda_1}{\partial\left(z_k\phi\right)} &= \frac{\sum_i w_i s_{ik}^*\left(\theta_k-\sum_m s_{im}^*\theta_m\right)-\lambda_1 s_k\left(\theta_k^2-\sum_m s_m\theta_m^2\right)}{\sum_m s_m\theta_m^2} \\
\frac{\partial \lambda_1}{\partial \zeta} &= \frac{\sum_i w_i\sum_m \theta_m s_{im}^*\left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta}-\sum_n s_{in}^*\frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial\zeta}\right)}{\sum_m s_m\theta_m^2} \\
\frac{\partial s_{im}^*}{\partial\zeta} &= s_{im}^*\left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta}-\sum_n s_{in}^*\frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial\zeta}\right)
\end{aligned}
$$

Stacking the two sets of first order conditions, we get the following analytic formulae

$$
\begin{aligned}
\frac{\partial}{\partial\zeta'}\left[\sum_m s_{im}^*\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta}\right] &= \sum_m s_{im}^*\frac{\partial^2 \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta\partial\zeta'}+\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta}\frac{\partial s_{im}^*}{\partial\beta'} \\
&= \sum_m s_{im}^*\left(\frac{\partial^2 \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta\partial\zeta'}+\left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta}\right)^2\right) \\
&\quad -\left(\sum_m s_{im}^*\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\zeta}\right)^2
\end{aligned}
$$

where for convenience we use $(\cdot)^2$ to denote an outer product when applied to a vector, i.e. $\vec{x}^2 = xx'$. Second derivatives with respect to $\zeta$ are:

$$
\begin{aligned}
\frac{\partial^2 \ln \mathbb{P}_{il(i)}}{\partial\beta\partial\beta'} &= -\exp\left(-2\alpha\theta_m\right)\sum_j \mathbb{P}_{ij}x_{bj}\left(x_{bj} - \sum_k \mathbb{P}_{ij}x_{bk}\right)' \\
&= -\exp\left(-2\alpha\theta_m\right)\left(\sum_j \mathbb{P}_{ij}x_{bj}x'_{bj} - \sum_j \mathbb{P}_{ij}x_{bj}\left(\sum_j \mathbb{P}_{ij}x_{bj}\right)'\right) \\
\frac{\partial^2 \ln \mathbb{P}_{il(i)}}{\partial\beta\partial\gamma} &= \exp\left(\left(1-2\alpha\right)\theta_m + \gamma\right)\left(\sum_j \mathbb{P}_{ij,\theta_m}x_{bj}p_{lj} - \sum_j \mathbb{P}_{ij,\theta_m}x_{bj}\sum_j \mathbb{P}_{ij,\theta_m}p_j\right) \\
\frac{\partial^2 \ln \mathbb{P}_{il(i),\theta_m}}{\partial\gamma^2} &= \frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\gamma} + \exp\left(\left(1-\alpha\right)\theta_m + \gamma\right)\left(\sum_j \mathbb{P}_{ij,\theta_m}p_j\frac{\partial \ln \mathbb{P}_{ij,\theta_m}}{\partial\gamma}\right) \\
&= \frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\gamma} \\
&\quad -\exp\left(2\left(\left(1-\alpha\right)\theta_m + \gamma\right)\right)\left(\sum_j \mathbb{P}_{ij,\theta_m}p_j^2 - \left(\sum_j \mathbb{P}_{ij,\theta_m}p_j\right)^2\right) \\
\frac{\partial^2 \ln \mathbb{P}_{il(i)}}{\partial\beta\partial\alpha} &= -\theta_m\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\beta} \\
&\quad +\theta_m\exp\left(-\alpha\theta_m\right)\left(\sum_j \mathbb{P}_{ij,\theta_m}d_{ij,\theta_m}\left(x_j - \sum_k \mathbb{P}_{ik,\theta_m}x_k\right)\right) \\
\frac{\partial^2 \ln \mathbb{P}_{il(i)}}{\partial\gamma\partial\alpha} &= -\theta_m\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\gamma} \\
&\quad -\theta_m\exp\left(\left(1-\alpha\right)\theta_m + \gamma\right)\left(\sum_j \mathbb{P}_{ij,\theta_m}d_{ij,\theta_m}\left(p_j - \sum_k \mathbb{P}_{ij,\theta_m}p_k\right)\right) \\
\frac{\partial^2 \ln \mathbb{P}_{il(i),\theta_m}}{\partial\alpha^2} &= -\theta_m\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\alpha} + \theta_m\left(\sum_j \mathbb{P}_{ij,\theta_m}d_{ij,\theta_m}\frac{\partial \ln \mathbb{P}_{ij,\theta_m}}{\partial\alpha}\right) \\
&= -\theta_m\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial\alpha} - \theta_m^2\left(\sum_j \mathbb{P}_{ij,\theta_m}d_{ij,\theta_m}^2 - \left(\sum_j \mathbb{P}_{ij,\theta_m}d_{ij,\theta_m}\right)^2\right)
\end{aligned}
$$

Finally, we need to compute the Hessian with respect to the share parameters $\phi$ by

$$\frac{\partial}{\partial(z_l\phi)}\left[s_{ik}^* - s_k\right] = s_{ik}^*\left(\delta_{k,l} - s_{il}^*\right) - s_k\left(\delta_{k,l} - s_l\right)$$

$$\begin{aligned}
\lambda^* &= \frac{\sum_m s_m\theta_m}{\sum_m s_m\theta_m^2}\sum_i w_i \\
\frac{\partial\lambda_1^*}{\partial(z_k\phi)} &= s_k\theta_k\frac{\sum_i w_i - \theta_k\lambda_1^*}{\sum_m s_m\theta_m^2} \\
\frac{\partial^2\lambda_1^*}{\partial(z_k\phi)\,\partial(z_l\phi)} &= \frac{\partial\lambda_1^*}{\partial(z_k\phi)}\left(\delta_{k,l} - \frac{s_l\theta_l^2}{\sum_m s_m\theta_m^2}\right) - \frac{\partial\lambda_1^*}{\partial(z_l\phi)}\left(\frac{s_k\theta_k^2}{\sum_m s_m\theta_m^2}\right)
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial}{\partial(z_l\phi)}\left[-\lambda_1^* s_k\left(\theta_k - \sum_m s_m\theta_m\right) - \frac{\partial\lambda_1^*}{\partial(z_k\phi)}\sum_m s_m\theta_m\right] \\
&\quad= -\frac{\partial\lambda_1^*}{\partial(z_l\phi)}s_k\left(\theta_k - \sum_m s_m\theta_m\right) - \frac{\partial\lambda_1^*}{\partial(z_k\phi)}s_l\left(\theta_l - \sum_m s_m\theta_m\right) \\
&\qquad - \lambda_1\left(s_k\left(\theta_k - \sum_m s_m\theta_m\right)(\delta_{k,l} - s_l) - s_k s_l\left(\theta_l - \sum_m s_m\theta_m\right)\right) \\
&\qquad - \frac{\partial^2\lambda_1^*}{\partial(z_k\phi)\,\partial(z_l\phi)}\sum_m s_m\theta_m
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\lambda_1}{\partial(z_k\phi)(z_l\phi)} &= \frac{\partial\lambda_1}{\partial(z_k\phi)}\left(\delta_{k,l} - \frac{s_l\left(\theta_l^2 - \sum_m s_m\theta_m^2\right)}{\sum_m s_m\theta_m^2}\right) - \frac{\partial\lambda_1}{\partial(z_l\phi)}\frac{s_k\left(\theta_k^2 - \sum_m s_m\theta_m^2\right)}{\sum_m s_m\theta_m^2} \\
&\quad - \sum_i w_i s_{il}^* s_{ik}^*\frac{\theta_k + \theta_l - 2\sum_m s_{im}^*\theta_m}{\sum_m s_m\theta_m^2} \\
&\quad + \lambda_1\frac{s_l s_k\left(\theta_k^2 - \sum_m s_m\theta_m^2\right) + s_k s_l\left(\theta_l^2 - \sum_m s_m\theta_m^2\right)}{\sum_m s_m\theta_m^2}
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial}{\partial\zeta'}\left[-\lambda_1 s_k\left(\theta_k - \sum_m s_m\theta_m\right) - \frac{\partial\lambda_1}{\partial(x_k\phi)}\sum_m s_m\theta_m\right] \\
&\quad= -\frac{\partial\lambda_1}{\partial\zeta'}s_k\left(\theta_k - \sum_m s_m\theta_m\right) - \frac{\partial^2\lambda_1}{\partial(x_k\phi)\,\partial\zeta'}\sum_m s_m\theta_m
\end{aligned}$$

$$
\begin{aligned}
\left(\sum_m s_m \theta_m^2\right) \frac{\partial^2 \lambda_1}{\partial (x_k \phi) \partial \zeta'} = \sum_i w_i s_{ik}^* \Bigg[ \\
& \left(\theta_k - \sum_m \theta_m s_{im}^*\right) \left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_k}}{\partial \zeta} - \sum_n s_{in}^* \frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial \zeta}\right) \\
& - \sum_m \theta_m s_{im}^* \left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \zeta} - \sum_n s_{in}^* \frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial \zeta}\right) \\
& \Bigg] - \frac{\partial \lambda_1}{\partial \zeta} s_k \left(\theta_k^2 - \sum_m s_m \theta_m^2\right)
\end{aligned}
$$

$$
\left(\sum_m s_m \theta_m^2\right) \lambda_1 = \sum_m \theta_m \sum_i w_i s_{im}^*
$$

$$
\left(\sum_m s_m \theta_m^2\right) \frac{\partial \lambda_1}{\partial \zeta} = \sum_i w_i \sum_m \theta_m s_{im}^* \left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \zeta} - \sum_n s_{in}^* \frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial \zeta}\right)
$$

$$
\begin{aligned}
\left(\sum_m s_m \theta_m^2\right) \frac{\partial^2 \lambda_1}{\partial \zeta \partial \zeta'} = \sum_i w_i \sum_m \theta_m s_{im}^* \Bigg\{ \\
& \Bigg[ \left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \zeta} - \sum_n s_{in}^* \frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial \zeta}\right) \\
& \times \left(\frac{\partial \ln \mathbb{P}_{il(i),\theta_m}}{\partial \zeta'} - \sum_n s_{in}^* \frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial \zeta'}\right) \Bigg] \\
& + \frac{\partial^2 \ln \mathbb{P}_{il(i),\theta_m}}{\partial \zeta \partial \zeta'} \\
& \Bigg\} - \frac{\partial}{\partial \zeta'} \left[\sum_n s_{in}^* \frac{\partial \ln \mathbb{P}_{il(i),\theta_n}}{\partial \zeta}\right]
\end{aligned}
$$

# Appendix C

# Appendix for Chapter 3

## C.1 Variable Descriptions

Important variables are described in the body of the paper. Below we describe the construction of the set of controls that we commonly use in our specifications as functions of primitives.

**Capital Asset Ratio (C-A Ratio)** Ratio of net physical plant, property, and equipment to total assets.

**Cash Flow** Ratio of income before extraordinary items plus depreciation and amortization to net physical plant, property, and equipment.

**Current Ratio** Total assets divided by total liabilities.

**Investment** Capital expenditure divided by net physical plant, property, and equipment.

**Leverage** Total debt divided by total assets.

**Macro Q** Common equity plus total debt less inventories, all divided by net physical plant, property, and equipment.

**Net Worth** Total assets less total liabilities, divided by total assets.

**Return on Assets (ROA)** Operating income before depreciation extraordinary items divided by total assets.

## C.2 Omitted Proofs

*Proof of Proposition 3.2.1.* For notational simplicity let $p \equiv p_i$ and hold $p_j$ fixed. Then dividing the two first order conditions by $1 - F(\hat{z})$ and $F(\hat{z})$, respectively, reduces them to $\mathbb{E}[\pi_i(p; z)|z \leq \hat{z}]$ and $\mathbb{E}[\pi_i(p; z)|z \geq \hat{z}]$. Since $\pi_{iz} > 0$, by monotonicity $\mathbb{E}[\pi_i(p; z)|z \leq \hat{z}] < \mathbb{E}[\pi_i(p; z)|z \geq \hat{z}]$ $\forall p$. Evaluating the creditor's first-order condition at a solution to the equity holder's first-order condition, $p_f^*$, therefore gives $\mathbb{E}[\pi_i(p_f^*; z)|z \leq \hat{z}] < 0$. Therefore, because $\pi_{ii} < 0$, the optimal $p_b^*$ that satisfies $\mathbb{E}[\pi_i(p_b^*; z)|z \leq \hat{z} = 0$ must be less than $p_f^*$. □

*Proof of Corollary 3.2.1.1.* The condition that $\pi_{ij} > 0$ ensures that prices are strategic substitutes (response curves are upward sloping) since $\pi_{ii} < 0$. Applying Proposition 3.2.1 then establishes the result. □

*Proof of Proposition 3.2.2.* Use condition (i) to establish an inequality in the first terms of the respective FOCs then proceed exactly as in the proof of Proposition 3.2.1. □

# Bibliography


Abrate, Graziano, Giovanni Fraquelli, and Giampaolo Viglia (2012), "Dynamic pricing strategies: Evidence from european hotels." *International Journal of Hospitality Management*, 31, 160–168.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer (2015), "Identification properties of recent production function estimators." *Econometrica*, 83, 2411–2451.

Ackerberg, Daniel A and Marc Rysman (2005), "Unobserved product differentiation in discrete choice models: estimating price elasticities and welfare effects, forthcoming rand journal of economics."

Aghion, Philippe and Patrick Bolton (1992), "An incomplete contracts approach to financial contracting." *The Review of Economic Studies*, 59, 473–494.

Angrist, J.D. and J.S. Pischke (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arellano, Manuel and Stephen Bond (1991), "Some tests of specification for panel data: Monte carlo evidence and an application to employment equations." *Review of Economic Studies*, 58, 277–297.

Argentesi, Elena and Lapo Filistrucchi (2007), "Estimating market power in a two-sided market: The case of newspapers." *Journal of Applied Econometrics*, 22, 1247–1266.

Armstrong, Mark (2006), "Competition in two-sided markets." *The RAND Journal of Economics*, 37, 668–691.

Bajari, Patrick and C Lanier Benkard (2005), "Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach." *Journal of political economy*, 113, 1239–1276.

Bakewell, Sally and Lisa Lee (2020), "Wall street caves to companiesâĂŹ demands to stave off defaults." *Bloomberg*, URL `https://www.bloomberg.com/news/articles/2020-04-16/forget-2020-besieged-companies-seek-wall-street-loan-reprieve`.

Bellucci, Andrea, Ilario Favaretto, and Germana Giombini (2014), “Does innovation affect credit access? new empirical evidence from italian small business lending.” IAW Discussion Papers 104, Institut fÃijr Angewandte Wirtschaftsforschung (IAW).

Bengio, Yoshua, Pascal Lamblin, Dan Popovici, and Hugo Larochelle (2007), “Greedy layer-wise training of deep networks.” In *Advances in neural information processing systems*, 153–160.

Bharath, Sreedhar T and Michael Hertzel (2019), “External governance and debt structure.” *The Review of Financial Studies*, 32, 3335–3365.

Bitran, Gabriel, Rene Caldentey, and Susana Mondschein (1998), “Coordinating clearance markdown sales of seasonal products in retail chains.” *Operations research*, 46, 609–624.

Bitran, Gabriel R and Susana V Mondschein (1997), “Periodic pricing of seasonal products in retailing.” *Management science*, 43, 64–79.

Blake, Thomas, Sarah Moshary, Kane Sweeney, and Steven Tadelis (2018), “Price salience and product choice.” Technical report, National Bureau of Economic Research.

Blundell, Richard and Stephen Bond (1998), “Initial conditions and moment restrictions in dynamic panel data models.” *Journal of Econometrics*, 87, 115–143.

Board, Simon and Andrzej Skrzypacz (2016), “Revenue management with forward-looking buyers.” *Journal of Political Economy*, 124, 1046–1087.

Brander, James A. and Tracy R. Lewis (1986), “Oligopoly and financial structure: The limited liability effect.” *The American Economic Review*, 956–970.

Cabral, Marika and Caroline Hoxby (2012), “The hated property tax: salience, tax rates, and tax revolts.” Technical report, National Bureau of Economic Research.

Caillaud, Bernard and Bruno Jullien (2003), “Chicken & egg: Competition among intermediation service providers.” *RAND journal of Economics*, 309–328.

Cantillon, Estelle and Pai Ling Yin (2011), “Competition between exchanges: Lessons from the battle of the bund.”

Carey, M. and M. Hrycray (1999), “Credit flow, risk, and the role of private debt in capital structure.” Federal Reserve Board Mimeo.

Chakraborty, Indraneel and Michael Ewens (2012), “Does security choice matter in venture capital? the case of venture debt.” Mimeo.

Chava, Sudheer, Vikram K. Nanda, and Steven Chong Xiao (2016), “Lending to innovative firms.” Working Paper.

Chava, Sudheer and Michael R. Roberts (2008), "How does financing impact investment? the role of debt covenants." *The Journal of Finance*, 63, 2085–2121.

Chen, Tianqi and Carlos Guestrin (2016), "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Chetty, Raj (2009), "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods." *Annu. Rev. Econ.*, 1, 451–488.

Chetty, Raj, Adam Looney, and Kory Kroft (2009), "Salience and taxation: Theory and evidence." *American economic review*, 99, 1145–77.

Chevalier, Judith A. (1995), "Do lbo supermarkets charge more? an empirical analysis of the effects of lbos on supermarket pricing." *The Journal of Finance*, 50, 1095–1112.

Cho, Sungjin, Gong Lee, John Rust, and Mengkai Yu (2018), "Optimal dynamic hotel pricing." Technical report.

Chodorow-Reich, Gabriel (2014), "The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis." *The Quarterly Journal of Economics*, 129, 1–59.

Chodorow-Reich, Gabriel and Antonio Falato (2020, forthcoming), "The loan covenant channel: How bank health transmits to the real economy." *Journal of Finance*.

Clements, Matthew T and Hiroshi Ohashi (2005), "Indirect network effects and the product cycle: video games in the us, 1994–2002." *The Journal of Industrial Economics*, 53, 515–542.

Corts, Kenneth S and Mara Lederman (2009), "Software exclusivity and the scope of indirect network effects in the us home video game market." *international Journal of industrial Organization*, 27, 121–136.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020), "The rise of market power and the macroeconomic implications." *The Quarterly Journal of Economics*, 135, 561–644.

De Loecker, Jan and Frederic Warzynski (2012), "Markups and firm-level export status." *American Economic Review*, 102, 2437–71.

Dichev, Ilia D. and Douglas J. Skinner (2002), "LargeâĂŞsample evidence on the debt covenant hypothesis." *Journal of Accounting Research*, 40, 1091–1123.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan (2018), "Consumer price search and platform design in internet commerce." *American Economic Review*, 108, 1820–59.

Einav, Liran, Chiara Farronato, and Jonathan Levin (2016), "Peer-to-peer markets." *Annual Review of Economics*, 8, 615–635.

Ellison, Glenn and Sara Fisher Ellison (2009), "Search, obfuscation, and price elasticities on the internet." *Econometrica*, 77, 427–452.

Ellison, Glenn and Sara Fisher Ellison (2011), "Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration." *American Economic Journal: Microeconomics*, 3, 1–36.

Ersahin, Nuri, Rustom M. Irani, and Hanh Le (2020), "Creditor control rights and resource allocation within firms." *Journal of Financial Economics*.

Evans, David S (2003), "The antitrust economics of multi-sided platform markets." *Yale Journal on Regulation*, 20, 4.

Falato, Antonio and Nellie Liang (2016), "Do creditor rights increase employment risk? evidence from loan covenants." *The Journal of Finance*, 71, 2545–2590.

Faria-e Castro, Miguel and Asha Bharadwaj (2019), "Syndicated loans in the u.s." Technical report, Federal Reserve Bank of St. Louis, URL `https://www.stlouisfed.org/on-the-economy/2019/october/syndicated-loans-us`.

Farrell, Joseph and Garth Saloner (1985), "Standardization, compatibility, and innovation." *the RAND Journal of Economics*, 70–83.

Farrell, Joseph and Garth Saloner (1986), "Installed base and compatibility: Innovation, product preannouncements, and predation." *The American economic review*, 940–955.

Finkelstein, Amy (2009), "E-ztax: Tax salience and tax rates." *The Quarterly Journal of Economics*, 124, 969–1010.

Flannery, Mark and Kasturi P. Rangan (2006), "Partial adjustment toward target capital structures." *Journal of Financial Economics*, 79, 469–506.

Fradkin, Andrey (2019), "Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb."

Gallego, Guillermo and Garrett Van Ryzin (1994), "Optimal dynamic pricing of inventories with stochastic demand over finite horizons." *Management science*, 40, 999–1020.

Henderson, Richard (2019), "The incredible shrinking stock market." *Financial Times*, URL `https://www.ft.com/content/0c9c0b64-9760-11e9-9573-ee5cbb98ed36`.

Hermalin, Benjamin E and Michael L Katz (2006), "Your network or mine? the economics of routing rules." *The RAND Journal of Economics*, 37, 692–719.

Hotelling, Harold (1931), "The generalization of student's ratio." *Ann. Math. Statist.*, 2, 360–378, URL https://doi.org/10.1214/aoms/1177732979.

Huang, Jason, David H. Reiley, and Nickolai M. Riabov (2018), "Measuring consumer sensitivity to audio advertising: A field experiment on pandora internet radio.", URL http://davidreiley.com/papers/PandoraListenerDemandCurve.pdf.

Jensen, Michael C. (1986), "Agency costs of free cash flow, corporate finance, and takeovers." *The American Economic Review*, 76, 323–329.

Jensen, Michael C and William H Meckling (1976), "Theory of the firm: Managerial behavior, agency costs and ownership structure." *Journal of Financial Economics*, 3, 305–360.

Jin, Ginger Zhe and Marc Rysman (2015), "Platform pricing at sports card conventions." *The Journal of Industrial Economics*, 63, 704–735.

Kaiser, Ulrich and Julian Wright (2006), "Price structure in two-sided markets: Evidence from the magazine industry." *International Journal of Industrial Organization*, 24, 1–28.

Katz, Michael L and Carl Shapiro (1985), "Network externalities, competition, and compatibility." *The American economic review*, 75, 424–440.

Katz, Michael L and Carl Shapiro (1986), "Technology adoption in the presence of network externalities." *Journal of political economy*, 94, 822–841.

Kayhan, Ayla and Sheridan Titman (2007), "Firms' histories and their capital structures." *Journal of Financial Economics*, 83, 1–32.

Kogan, Leonid, Dimitris Papanikolaou, Amit Serua, and Noah Stoffman (2016), "Technological innovation, resource allocation and growth." Forthcoming, Quarterly Journal of Economics.

Lancaster, Tony (2002), "Orthogonal parameters and panel data." *The Review of Economic Studies*, 69, 647–666.

Lee, Robin S (2013), "Vertical integration and exclusivity in platform and two-sided markets." *American Economic Review*, 103, 2960–3000.

Leland, Hayne and David H Pyle (1977), "Informational asymmetries, financial structure, and financial intermediation." *Journal of Finance*, 32, 371–87.

Levinsohn, James and Amil Petrin (2003), "Estimating production functions using inputs to control for unobservables." *Review of Economic Studies*, 70, 317–341.

Lewis, Randall, Justin M Rao, and David H Reiley (2015), "Measuring the effects of advertising: The digital frontier." In *Economic Analysis of the Digital Economy*, 191–218, University of Chicago Press.

Mann, William (2016), "Creditor rights and innovation: Evidence from patent collateral." Working Paper.

Matzkin, Rosa L (2003), "Nonparametric estimation of nonadditive random functions." *Econometrica*, 71, 1339–1375.

McAfee, R. P. and V. Te Velde (2006), "Dynamic pricing in the airline industry." In *Handbook on Economics and Information Systems*, Elsevier.

McAfee, R Preston and Vera te Velde (2008), "Dynamic pricing with constant demand elasticity." *Production and operations Management*, 17, 432–438.

Myers, Stewart C. and Nicholas S. Majluf (1984), "Corporate financing and investment decisions when firms have information that investors do not have." *Journal of Financial Economics*, 13, 187–221.

Nadaraya, Elizbar A (1964), "On estimating regression." *Theory of Probability & Its Applications*, 9, 141–142.

Ng, Serena (2020), "Another problem for the fed: Banks pressured as clients scramble for cash." *The Wall Street Journal*, URL `https://www.wsj.com/articles/another-problem-for-the-fed-banks-pressured-as-clients-scramble-for-cash-11584356272`.

Nickell, Stephen (1981), "Biases in dynamic models with fixed effects." *Econometrica*, 49, 1417–1426.

Nini, Greg, David C Smith, and Amir Sufi (2009), "Creditor control rights and firm investment policy." *Journal of Financial Economics*, 92, 400–420.

Nini, Greg, David C Smith, and Amir Sufi (2012), "Creditor control rights, corporate governance, and firm value." *The Review of Financial Studies*, 25, 1713–1761.

Ohashi, Hiroshi (2003), "The role of network effects in the us vcr market, 1978–1986." *Journal of Economics & Management Strategy*, 12, 447–494.

Olley, G Steven and Ariel Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64, 1263–1297.

Peitz, Martin, Sven Rady, and Piers Trepper (2017), "Experimentation in two-sided markets." *Journal of the European Economic Association*, 15, 128–172.

Petrongolo, Barbara and Christopher A Pissarides (2001), "Looking into the black box: A survey of the matching function." *Journal of Economic literature*, 39, 390–431.

Platt, John C and Alan H Barr (1988), "Constrained differential optimization." In *Neural Information Processing Systems*, 612–621.

Rasmussen, Carl Edward and Christopher K. I. Williams (2006), *Gaussian Processes for Machine Learning*. MIT Press.

Reshef, Oren (2019), “Smaller slices of a growing pie: The effects of entry in platform markets.”

Rivers, Nicholas and Brandon Schaufele (2015), “Salience of carbon taxes in the gasoline market.” *Journal of Environmental Economics and Management*, 74, 23–36.

Roberts, Michael and Amir Sufi (2009), “Control rights and capital structure: An empirical investigation.” *Journal of Finance*, 64, 1657–1695.

Rochet, Jean-Charles and Jean Tirole (2003), “Platform competition in two-sided markets.” *Journal of the european economic association*, 1, 990–1029.

Rochet, Jean-Charles and Jean Tirole (2006), “Two-sided markets: a progress report.” *The RAND journal of economics*, 37, 645–667.

Rosen, Sherwin (1974), “Hedonic prices and implicit markets: Product differentiation in pure competition.” *Journal of Political Economy*, 82, 34–55.

Ross, Stephen A. (1977), “The determination of financial structure: The incentive-signalling approach.” *Bell Journal of Economics*, 8, 23–40.

Rysman, Marc (2004), “Competition between networks: A study of the market for yellow pages.” *The Review of Economic Studies*, 71, 483–512.

Schmalensee, Richard (2002), “Payment systems and interchange fees.” *The Journal of Industrial Economics*, 50, 103–122.

Schölkopf, Bernhard, Ralf Herbrich, Alex J. Smola, and Robert Williamson (2000), “A generalized representer theorem.”

Silva, JMC Santos and Silvana Tenreyro (2006), “The log of gravity.” *The Review of Economics and statistics*, 88, 641–658.

Stulz, Rene (1990), “Managerial discretion and optimal financing policies.” *Journal of Financial Economics*, 26, 3–27.

Sweeting, Andrew (2012), “Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for major league baseball tickets.” *Journal of Political Economy*, 120, 1133–1172.

Sweeting, Andrew and Kane Sweeney (2015), “Staggered vs. simultaneous price setting with an application to an online market.” In *Technical Report*, Working paper, University of Maryland, College Park, MD.

Wang, Zhu and Julian Wright (2017), “Ad valorem platform fees, indirect taxes, and efficient price discrimination.” *The RAND Journal of Economics*, 48, 467–484.

Watson, Geoffrey S (1964), “Smooth regression analysis.” *Sankhyā: The Indian Journal of Statistics, Series A*, 359–372.

Weyl, E Glen (2010), “A price theory of multi-sided platforms.” *American Economic Review*, 100, 1642–72.

White, Alexander and E Glen Weyl (2016), “Insulated platform competition.” Mimeo.

Wirz, Matt and Nick Timiraos (2020), “The next coronavirus financial crisis: Record piles of risky corporate debt.” *The Wall Street Journal*, URL `https://www.wsj.com/articles/the-next-coronavirus-financial-crisis-record-piles-of-risky-corporate-debt-11584627925`.

Wright, Julian (2004), “The determinants of optimal interchange fees in payment systems.” *The Journal of Industrial Economics*, 52, 1–26.