THE UNIVERSITY OF CHICAGO

STRATEGIC BEHAVIOR AND MECHANISM DESIGN WITH SEQUENTIAL INFORMATION REVELATION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
JOSHUA DAVID HIGBEE

CHICAGO, ILLINOIS
JUNE 2025

ABSTRACT

This dissertation consists of two essays on the role of incomplete information, and the timing with which it is revealed, in agents' strategic behavior and the mechanisms that are designed to account for this behavior. The first essay studies the question of optimal information provision by a monopolist platform when sellers learn about key features of the auction process as they gain experience. Online platforms often do not directly control users' pricing strategy, and instead offer analytics and other information to help steer user behavior. I study how information provision by an auction platform to sellers shapes platform fees and outcomes using data from eBay auctions of children's toys. I present evidence that new sellers face uncertainty about how to set optimal reserve prices: they set lower reserve prices and earn higher revenues as they gain more experience. I develop a model where new sellers learn to set reserve prices on an auction platform with selective participation, and I show that sellers choose reserve prices to both extract surplus from bidders and attract additional bidders to their auction. I provide conditions under which new sellers' beliefs about the effect of reserve prices on bidder arrival are semiparametrically identified. Estimates from the learning model indicate that new sellers underestimate the effect of high reserve prices on deterring bidder entry, which leads to higher reserve prices and more items listed than for fully-informed sellers. Counterfactual simulations show that platform information provision can help new sellers learn the true bidder arrival process, which increases bidder entry as well as seller and platform profits. However, when jointly choosing information provision and platform fees, it may be optimal to not fully reveal information to sellers.

The second essay studies the timing of contract design in large corporate transactions, where the suitability of various contract terms for a particular deal is not known until firms expend costly effort developing these terms. Significant corporate transactions (such as financing and acquisition agreements) are typically negotiated in stages, wherein core pricing terms are fixed early while most non-price provisions are relegated to subsequent

bargaining. This ordering stands in stark (and curious) contrast with canonical theories of contract design, which overwhelmingly counsel that non-price terms should be set first, saving price negotiations for last so as to fine tune the parties' net payoffs. This longstanding disjunction between theory and practice has become a celebrated puzzle for transactional design. We present an analytic framework that helps to reconcile the two, marrying a bargaining model and a search game over innovative contractual provisions. Our framework delivers a robust and tractable set of intuitions about when fixing price before other terms optimally incentivizes strategic search investments by the contracting parties. Our analysis is also amenable to making counterfactual comparisons of regimes where price is (and is not) set first, generating in the process several empirically testable implications.

TABLE OF CONTENTS

LIST OF FIGURES . . . vii

LIST OF TABLES . . . ix

ACKNOWLEDGMENTS . . . x

1 LEARNING AND INFORMATION DESIGN ON AN AUCTION PLATFORM . . . 1
- 1.1 Introduction . . . 1
  - 1.1.1 Related literature . . . 6
  - 1.1.2 Roadmap . . . 7
- 1.2 A simple model of platform fees and information provision . . . 8
- 1.3 Setting and auction data . . . 10
  - 1.3.1 Auction listing and bid data . . . 12
  - 1.3.2 Evidence for learning among new sellers . . . 13
  - 1.3.3 Alternative explanations for new seller behavior . . . 16
- 1.4 Auction platform model with seller learning . . . 18
  - 1.4.1 Bidder strategies . . . 20
  - 1.4.2 Seller strategies . . . 23
  - 1.4.3 Equilibrium definition . . . 29
- 1.5 Identification and estimation . . . 30
  - 1.5.1 Demand side estimation strategy . . . 31
  - 1.5.2 Supply side estimation strategy . . . 35
  - 1.5.3 Estimates from the platform model . . . 42
- 1.6 Counterfactual platform design with information provision . . . 46
  - 1.6.1 Revisiting the platform's problem of fees and information provision . . . 46
  - 1.6.2 Optimal fee structures without information provision . . . 48
  - 1.6.3 Optimal fees with information provision . . . 51
  - 1.6.4 Seller willingness to pay for information . . . 53
- 1.7 Conclusions . . . 55

2 FIX THE PRICE OR PRICE THE FIX? RESOLVING THE SEQUENCING PUZZLE IN CORPORATE CONTRACTING . . . 57
- 2.1 Introduction . . . 57
- 2.2 Contract production in the M&A market . . . 66
- 2.3 A model of two-stage contracting . . . 68
  - 2.3.1 Term innovation and alternate sequencing . . . 69
  - 2.3.2 Contract creation in the price-first game . . . 73
  - 2.3.3 Contract creation in the price-last game . . . 75
- 2.4 Characterizing the equilibrium contract . . . 77
  - 2.4.1 Simplifying assumptions and characterization of equilibrium search . . . 80
  - 2.4.2 Orthogonal (self-interested) search . . . 86
  - 2.4.3 Aligned (surplus-maximizing) search . . . 89

2.4.4 Partially contractible term search . . . 93
2.4.5 Equilibrium contract under alternative assumptions . . . 99
2.5 Implications . . . 104
2.6 Conclusion . . . 108

REFERENCES . . . 109

A APPENDIX FOR CHAPTER 1 . . . 117
A.1 Summary statistics and additional descriptive evidence . . . 117
A.2 Testing for common values . . . 123
A.3 Proofs . . . 124
A.4 Demand estimation details . . . 129
A.5 Orthogonalization of the likelihood function . . . 131
A.6 Likelihood derivation: demand side . . . 133
A.7 Likelihood approach: supply side belief estimation . . . 136

B APPENDIX FOR CHAPTER 2 . . . 141
B.1 Stochastic Nash bartering . . . 141
B.2 Proofs and derivations . . . 143
B.3 Additional figures . . . 148
B.4 Equilibrium contract with unrestricted term search and independent productivity shocks . . . 151

LIST OF FIGURES

1.1 Example of eBay search and item pages 11
1.2 Trends in variables of interest as new sellers gain experience 14
1.3 Coefficients from regressing reserve prices on lagged revenues, by binned seller experience 15
1.4 New seller trends by cohorts of the number of items listed (difference from experienced seller averages) 16
1.5 Optimal reserve price for varying reserve price coefficients $\delta_{0,2}$ 26
1.6 Intuition for identification of beliefs b_0 through variation in ψ and its derivatives 38
1.7 Estimated value distributions for auction participants 43
1.8 Estimated path of average new seller beliefs 45
1.9 Estimated average seller entry threshold conditional on baseline bidder arrival parameter $\delta_{j,0,1}$ 46
1.10 Percent changes in outcomes under alternative fee structures, $a = 0$ and $\omega = 0.5$ 49
1.11 Optimal fees and outcomes for varying ω with no information provision ($a = 0$) 50
1.12 Optimal fees with information provision 52
1.13 Estimated seller willingness to pay for information under different beliefs 54

2.1 Contract term components in firm payoff space 72
2.2 Firm payoffs in the bartering stage of the price-first game (perfect correlation in ϵ) 81
2.3 Relative search intensities for firms in price-first and price-last games 84
2.4 Comparative statics with respect to τ (orthogonal search) 88
2.5 Comparative statics with respect to γ_b (orthogonal search) 89
2.6 Comparative statics with respect to τ (aligned search) 91
2.7 Comparative statics with respect to γ_b (aligned search) 92
2.8 Comparative statics with respect to τ (endogenous angle search) 95
2.9 Comparative statics with respect to γ_a (endogenous angle search) 97
2.10 Comparative statics with respect to $\bar{a}$ (endogenous angle search) 99
2.11 Comparative statics with respect to τ (unrestricted term search) 103

A.1 Time trends of variables of interest with auction experience 118
A.2 Trends in the frequency of words in item descriptions 118
A.3 Trends in non-price variables among new sellers 120
A.4 Time trends of bids with auction experience 121
A.5 Distributions of reserve prices among new sellers in their first auction vs. experienced sellers 122
A.6 Coefficients for regressing lag auction sale outcomes on current effective reserve price, by experience bin 122
A.7 Estimated distributions of bidder valuations by the number of observed bidders per auction 124

B.1 Firm payoffs in the bartering stage of the price-first game (unrestricted model) . 142
B.2 Comparative statics with respect to τ (endogenous angle search) 149
B.3 Comparative statics with respect to γ_a (independent term shocks) 150

B.4 Firm payoffs in the bartering stage of the price-first game (unrestricted model) . 152
B.5 Term choice probabilities $\lambda_{j,G}$ 153

LIST OF TABLES

1.1 Estimated bidder arrival parameters and entry cost 42
1.2 Estimated new seller priors about the bidder arrival process 44

2.1 Timing of the two games 73

A.1 Summary statistics 117
A.2 Regression of log standardized bid values on number of bidders 123
A.3 Neural network architectures and performance 131
A.4 Simulations for maximum likelihood estimation (demand) 136

ACKNOWLEDGMENTS

I am grateful to Ali Hortaçsu, Dennis Carlton, Giovanni Compiani, and Günter Hitsch for their excellent comments and advising. I also thank Eric Richert, Brad Larsen, Avner Strulov-Shlain, Eric Budish, Chad Syverson, Devin Pope, Stephane Bonhomme, Kirill Ponomarev, Max Farrell, Cree Jones, Matthew Jennejohn, Eric Talley, Camilla Schneier, Arjun Gopinath, Nadia Lucas, Marco Loseto, Thomas Wiemann, and Samuel Higbee for their helpful comments and discussions. I am grateful to the participants of the IO Lunch Seminar, IO Reading Group, Implementation Matters, and Marketing Student Workshop for providing helpful feedback on my work. I am also grateful to my family, particularly Eliza, Christopher, Giselle, and my parents, for their support and encouragement.

CHAPTER 1

LEARNING AND INFORMATION DESIGN ON AN AUCTION PLATFORM

1.1 Introduction

Two-sided platforms help buyers and sellers transact by both providing a marketplace and accompanying infrastructure to successfully match users. These platforms earn profits by charging fees to buyers and sellers, and therefore seek to both attract users and ensure that they are generating revenue. However, these firms play a limited role in the decision problem of sellers that use their site: individual users choose which items to sell, as well as how to price and promote their listings within the platform's interface. While many platforms have a decentralized approach to on-site transactions, they often provide information to sellers to help them track and optimize their business (e.g., eBay Seller Hub, Amazon Seller Central, AirBnB Smart Pricing). This information—which can range from transaction data to fully automated tools—directly affects sellers' strategic decisions, since sellers may be uninformed or uncertain about how to maximize profits in a new setting.

Despite the prevalence of large platforms and the information services they provide, however, standard models of platform design ignore the use of information services by platforms. Both buyers and sellers are often assumed to have perfect information about how the platform functions, which they use to respond rationally to the fees chosen by the platform (Klein et al., 2005; Rochet & Tirole, 2003). However, recent empirical work shows potential gains to information in platform settings when it is used for more targeted advertising (Mela

I thank Ali Hortaçsu, Dennis Carlton, Giovanni Compiani, Günter Hitsch, Eric Richert, Brad Larsen, Avner Strulov-Shlain, Eric Budish, Chad Syverson, Devin Pope, Stephane Bonhomme, Kirill Ponomarev, Max Farrell, Camilla Schneier, Arjun Gopinath, Nadia Lucas, and Samuel Higbee, and seminar participants at the University of Chicago and BYU for their helpful comments. All errors are my own.

et al., 2023), personalized pricing (Wu et al., 2023), or suggested pricing (Fong et al., 2023). Ensuring that sellers have better information affects their behavior and can increase the total transacted revenue on the platform, allowing the platform to earn more commissions. Thus, a profit-maximizing platform must consider how to jointly choose the information and fees for its users.

In this paper, I study how information provision to new sellers by an auction platform determines the optimal fee structure and user welfare. I extend a two-sided auction platform model from Marra (2019) to a setting where sellers learn from past transactions whether to list items for sale and what reserve prices to set. Since new sellers update their beliefs with every item they list for auction, their beliefs—and consequently, their entry and reserve price decisions—vary with the information they observe. I apply this model to a large dataset of eBay auctions and find that new sellers are initially overoptimistic about bidder entry into their auctions, which leads to higher reserve prices and more entry. When the platform updates new sellers' beliefs toward the truth, however, it can increase bidder entry as well as platform and seller profits.

The question of optimal information provision is central to the platform's strategic problem. Within the auction platform, sellers face a problem similar to a standard monopoly pricing problem (Bulow & Klemperer, 1996). In this context, the demand faced by each individual seller is determined by both how many bidders enter each auction and how much those who enter value the item being sold; sellers must know these features to maximize profits. If sellers have incomplete information about their demand curve, the platform may wish to correct sellers' beliefs and benefit from their more informed choices. However, the platform may benefit from information asymmetries among its users if these increase either transaction volume or users' willingness to pay higher fees.[1]. Thus, sellers' beliefs about

1. That asymmetries in platform design may be optimal is not a new concept: both theoretical (Klein et al., 2005; Rochet & Tirole, 2003) and empirical (Gomes, 2014; Marra, 2019) work shows that different fees for two sides of a market may act as a form of cross-subsidization to induce optimal entry.

the arrival process—and the platform's role in influencing these beliefs—shape outcomes for sellers, bidders, and the platform itself.

This analysis is motivated by new evidence that new sellers learn to act optimally in both their choice of reserve prices and decision to list items for auction, suggesting that they may be influenced by new information about the auction process. I use rich item-level data from one million eBay auctions for children's toys to show that new sellers set lower reserve prices and earn higher revenues as they gain more experience. These new sellers' reserve prices are also more strongly correlated with past revenue than more experienced sellers' are, which is consistent with stronger early responses to new information. Reduced-form evidence alone, however, cannot isolate how information affects sellers' beliefs and behavior since sellers' private values also determine their decisions. Nor can this evidence reveal how much information is optimal for the *platform* to give to sellers or how information affects optimal fees.

To understand how new information shapes seller beliefs, and in turn determines sellers' reserve price and participation decisions within the platform, I develop a model of selective entry on an auction platform with seller learning. In equilibrium, both bidders and sellers participate in auctions if they believe it is profitable to do so, since each faces a fixed cost of either bidding on or listing an item. Since high reserve prices reduce the amount of expected surplus from an auction, I show that bidders are less likely to enter auctions with high reserve prices.[2] Sellers also decide to list an item for auction if their expected surplus from that auction is positive. However, new sellers do not fully understand the bidder entry process, and instead learn about bidder entry—and its effect on their expected profit, if they list an item for auction—through repeated transactions. In contrast, I show that a fully-informed seller chooses a reserve price to both attract and extract surplus from bidders, a generalization of the Myerson (1981) reserve price rule.

2. This effect is distinct from the mechanical effect of a high reserve price in an online auction, which excludes potential bidders from submitting lower bids even if they enter the auction.

I then estimate the model to quantify how bidders and sellers—particularly new, uninformed sellers—act on the platform and how receiving new information shapes sellers' beliefs. I use rich text descriptions for each item to flexibly model heterogeneity in item values with a neural network, and derive a likelihood that corrects for selection in observed bids on eBay. To avoid potential bias from this method due to the high dimensionality of the text data, I derive an orthogonal score from which I obtain consistent estimates of the parameters of interest (Chernozhukov et al., 2022; Farrell et al., 2020; Ichimura & Newey, 2022). I then show that seller beliefs about bidders' entry process are semiparametrically identified from the distribution of reserve prices, using results from the literature on identification of random coefficients models (Fox et al., 2012), and estimate the path of new sellers' beliefs as they transact on the platform. Since sellers' beliefs are very high dimensional due to the lack of a conjugate prior structure, I make this problem computationally tractable by approximating Bayesian learning via sequential update steps that I implement with neural networks.

I find that new sellers underestimate the effect of reserve prices on bidder entry, leading them to initially set too-high reserve prices and enter more than they should. Higher reserve prices have a strong negative effect on the expected number of bidders because bidders face a time cost of entering auctions. New sellers initially underestimate this effect, which causes them to be overoptimistic about the number of expected bidders. This means that new sellers are willing to set higher reserve prices, expecting to earn a higher revenue, which in turn causes more entry from overoptimistic sellers. However, new sellers' priors are relatively uninformative, so they rely heavily on new information to inform their beliefs. This allows many sellers to quickly learn to set prices that are broadly consistent with accurate beliefs about the bidder arrival process.

These results have important strategic implications for the platform and its decision of how much information to provide. The platform can choose to provide new sellers with data from past auctions, from which new sellers can learn about the true bidder entry process.

At the same time, the platform chooses what fees it will charge to both new and experienced sellers, who are already familiar with the bidder arrival process and set prices optimally. My estimates quantify how much and in what direction new sellers will adjust their beliefs about the bidder entry process in response to new information. When provided with a random sample of other auctions on the platform, the average new seller is able to update their beliefs toward the true parameters without incurring the time and monetary cost of starting to use the platform. This allows the platform to change its fee structure to increase platform profits from the baseline, though it may be optimal for the platform to further increase profits by exploiting new sellers' information gap.

Empirically estimating new sellers' beliefs about the bidder entry process is critical, as a platform's optimal information structure may be ambiguous without empirical evidence. First, while platforms may have increasing economies of scale in analyzing data, it is still costly to provide users with additional information. This may be exacerbated by users' low willingness to pay for additional information: sellers who do not anticipate changing their beliefs will see little value in purchasing additional data. Thus, despite the cost, platforms may prefer to offer such information for little to no fee, if they offer it at all. Second, it is possible for the platform to help one side of the market without harming the other: correcting sellers' beliefs yields improved profits for sellers while inducing more bidder entry. The effects of this may be magnified by the two-sidedness of online platforms: increasing the surplus of one side of the market may lead to more entry on both sides. Finally, since the users' decision problem may change with new information, the optimal fee structure on a two-sided platform (as studied under perfect information in e.g. Rochet and Tirole 2003 and Klein et al. 2005) may also change.

1.1.1 Related literature

This paper contributes to the large and growing literature on agent learning. Much of this literature focuses on learning-by-doing and the extent to which more experienced agents are able to leverage information to improve outcomes (Haggag et al., 2017; Simonsohn, 2010; Strulov-Shlain, 2021; Tadelis et al., 2023). In particular, Huang et al. (2020) shows how firms in a new market respond to market signals and set prices accordingly; I document similar behavior among new sellers on an auction platform. Cho and Rust (2010) also documents that car rental firms behave suboptimally in choosing rental prices, and as in my setting, lowering prices (at least for some cars) can increase overall profits. Other work addresses the theory of optimal behavior under uncertainty, especially in games with updating (Doraszelski et al., 2018; Hitsch, 2006; Keller & Rady, 1999; Rothschild, 1974). I contribute to the literature on estimating agent beliefs and learning process, as in Erdem and Keane (1996) and Kim (2020); I also provide semiparametric identification results for beliefs under learning as in Lu (2019) and Wang and Yang (2024).

I also address the field of auction design and optimal pricing in auctions. Auctions are a particularly well-suited setting in which to study information design in platform markets due to the rich literature on optimal bidder and seller behavior, including empirical tools to test and quantify theoretical results. In particular, I adapt the endogenous auction platform model of Marra (2019), which relates to other models of auctions with endogenous entry such as Levin and Smith (1996), to a setting with seller uncertainty. My model also develops an insight from Engelbrecht-Wiggans (1987) (that optimal reserve prices should account for their effect on bidder arrival) into a new reserve price condition that nests that of Myerson (1981); this reserve price also shows the tradeoff between the benefit of attracting another bidder (as in Bulow and Klemperer 1996) and expected surplus extraction when arrival is both endogenous and stochastic. Existing literature further shows that auction participants may act suboptimally, whether in laboratory settings or as large, sophisticated firms (Davis et al., 2011; Ostrovsky & Schwarz, 2016); my application examines small firms in a real marketplace. I further contribute the literature on empirical estimation of auctions by incorporating unstructured text data, similar to Bajari et al. (2023), Compiani et al. (2023), and Netzer et al. (2012).

This paper also relates to the growing literature on two-sided markets. Existing work has studied the question of cross-subsidization in optimal platform design both theoretically and empirically (Gomes, 2014; Jullien et al., 2021; Klein et al., 2005; Marra, 2019; Rochet & Tirole, 2003). Other work also examines the role of information provision and learning in online platforms (Foroughifar, 2023; Mela et al., 2023; Wu et al., 2023). In particular, Fong et al. (2023) finds that new sellers on a used goods platform set higher prices than experienced sellers, and evaluates in an experimental setting the extent to which providing a suggested price to sellers shifts their pricing strategies. I consider the fee and information provision problems jointly to understand the interaction between them, and in particular allow sellers to incorporate their private valuations in their pricing decisions. More broadly, studies of search and recommendation systems also examine how user behavior is influenced by non-price mechanisms, though these generally focus on buyers (Bronnenberg et al., 2016; Compiani et al., 2022; Hodgson & Lewis, 2023; Xu et al., 2022).

1.1.2 Roadmap

The rest of this paper proceeds as follows. Section 1.2 presents the problem of optimal platform design. Section 1.3 describes the eBay data used in this analysis and presents descriptive evidence that new sellers learn to set reserve prices. Section 2.3, combines a model of two-sided selective entry on an auction platform with a model of seller learning. Section 1.5 presents the identification and estimation strategies and estimates the structural model. Section 1.6 revisits the platform's optimal information and fee structure in light of the model estimates, and section 1.7 concludes.

1.2 A simple model of platform fees and information provision

To fix ideas, I present a simple model of an auction platform that can choose both fees and what information to provide to potential sellers. This stylized model highlights the important potential tradeoffs (or complementarities) between fees and information, which motivates a more detailed empirical study of these forces but cannot itself determine the platform's optimal decision. I assume sellers have potentially-biased beliefs about how to act optimally on the platform, and that receiving additional information makes sellers' beliefs and actions monotonically approach those that would hold under the full-information benchmark.

I consider three main decision variables that can be chosen by the platform: the amount of information a to provide to sellers, an insertion fee c^I for listing items, and a revenue fee c^R charged as a portion of successful transaction. Information provision $a \in \mathbb{R}_+$ is an index of how much the platform can shift sellers' beliefs, where $a = 0$ represents no information provision (i.e., sellers maintain their initial beliefs) and $a = \infty$ represents the limiting case of perfect information provision. For exposition, I also abstract from the two-sided nature of most platform fee structures and do not consider the incidence of either c^I or c^R on bidders or sellers.

The platform's profits depend on both whether sellers list items and the revenue they generate conditional on their participation. Both features are determined in equilibrium by bidder and seller behavior, so they are each functions of information a and fees c^I and c^R. I define the reduced-form objects $\mathcal{P}(a, c^I, c^R)$ as the probability that an item is listed and $\mathcal{R}(a, c^I, c^R)$ as the expected revenue from listed items. The platform's profit maximization problem is therefore

$$\max_{a,\, c^I,\, c^R} \quad \mathcal{P}(a, c^I, c^R) \cdot [c^I + c^R \cdot \mathcal{R}(a, c^I, c^R)]$$

For simplicity, I assume the platform faces zero marginal cost for facilitating auctions and for providing information, and that all other fixed costs are sunk.

Providing information affects seller behavior on both the extensive margin of participation and the intensive margin of expected revenue. The optimal platform design is characterized in part by the following first-order condition with respect to the amount of information provision:

$$0 = \underbrace{\mathcal{P}_a(a, c^I, c^R) \cdot [c^I + c^R \cdot \mathcal{R}(a, c^I, c^R)]}_{\text{Extensive margin}} + \underbrace{\mathcal{P}(a, c^I, c^R) \cdot c_S^R \cdot \mathcal{R}_a(a, c^I, c^R)}_{\text{Intensive margin}}$$

The signs of both $\mathcal{P}_a$ and $\mathcal{R}_a$ depend on sellers' initial beliefs about optimal participation and reserve prices, and have important implications for the platform's choice of how much information to provide to new sellers. I consider several cases informally here.

(i) *Overly pessimistic sellers.* Sellers may underestimate their expected profit per auction, listing fewer items and setting high reserve prices to avoid low sale prices. Here $\mathcal{P}_a > 0$ and $\mathcal{R}_a > 0$, and the platform optimally chooses to share information with sellers to increase the number and profitability of transactions.

(ii) *Overly optimistic sellers.* Sellers' optimism may cause them to list many items and set very low reserve prices that they do not expect to bind. Here $\mathcal{P}_a < 0$ and $\mathcal{R}_a < 0$, meaning that the platform may wish to take advantage of sellers' biased beliefs instead of providing them with helpful information.

(iii) *Uncertainty about the optimal reserve price.* Sellers may choose to set reserve prices to be either too high or too low. Here the signs of $\mathcal{P}_a$ and $\mathcal{R}_a$ may differ, so the platform must improve either the quantity of or revenue from transactions at the expense of the other.

While the details of the platform problem are important, these cases illustrate the potential factors that may be considered when determining optimal information provision.

Information provision also affects platform outcomes by determining how sellers respond to changes in fees. Consider as an example the case where $c^R = 0$, so the platform only earns revenue from charging insertion fees. Rearranging the platform's first-order condition with respect to insertion fees characterizes the optimal insertion fee as

$$c^I = \frac{-\mathcal{P}(a, c^I, 0)}{\mathcal{P}_{c^I}(a, c^I, 0)}$$

Though in general no closed-form solution exists for c^I, this equation provides some intuition for the platform's problem. Sellers who are likely to list items for auction—perhaps due to overoptimistic beliefs about their expected surplus—are willing to pay higher insertion fees since they believe they will recoup their loss. Alternatively, sellers who believe listing items is unprofitable must be enticed to participate with lower insertion fees. The effect of changing sellers' initial beliefs through information provision depends on both what these initial beliefs are and how they affect sellers' behavior.

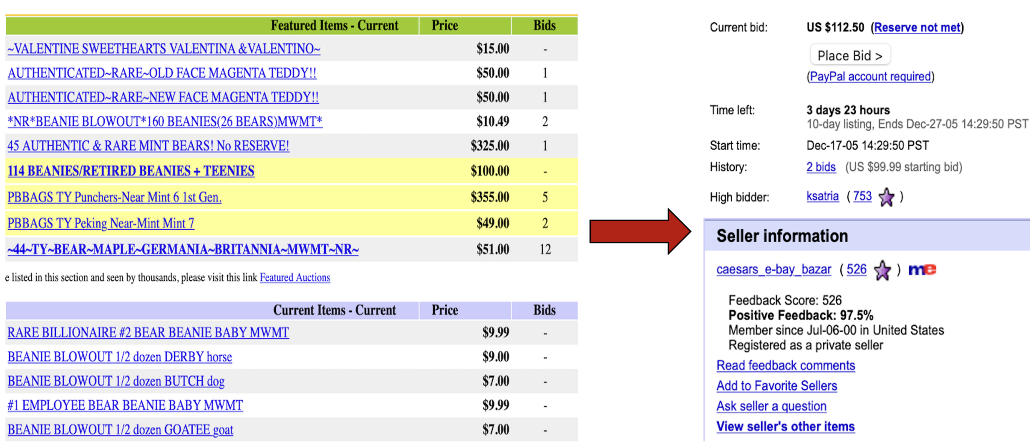
Ultimately, this simple model cannot solve the joint problem of determining information provision and platform fees. Both platform participation and expected revenue per auction are determined by bidders' and sellers' equilibrium behavior. Thus, taking this problem seriously requires a more careful look "under the hood" of the auction platform equilibrium. With this in mind, I turn to the setting of interest to highlight key features and trends in new seller behavior.

1.3 Setting and auction data

I use data from eBay, a well-known auction platform, to study how new sellers act and respond to new information. This data is a subset of the eBay auctions used in Resnick and Zeckhauser (2002), and spans from January to June 1999. During this time, eBay was relatively new (having been started in late 1995) and auctions were the only mechanism used to sell items, making this an ideal dataset in which to study how new sellers learn to operate in an unfamiliar setting. This data also preceeds the introduction of eBay's data analytics service "Seller Hub" in 2016.

To fix ideas, I present a simplified outline of the eBay auction process. Sellers choose to list an item for auction, and choose a starting minimum bid and (if desired) a secret reserve price along with an item description. Prospective bidders can find listed items on a search page, including the current minimum bid, and then choose to click into the item page. Bidders may then observe additional item details along with seller information and an indicator for whether the secret reserve price (if any exists) has been met. Bidders submit their bids to eBay, which proceeds as a second-price ascending auction where the current minimum bid is the second-highest of existing bids and the initial minimum bid. Examples of the search and item pages are presented in Figure 1.1.

Figure 1.1: Example of eBay search and item pages



Notes: This figure was retrieved from the Wayback machine, and has been edited to conserve space. The search and item pages are from 2001 and 2005, respectively. The sellers' ability to feature their item can be seen here - featured items are listed at the top of the page, while other items are highlighted and/or have bold titles. The seller's total feedback score and positive feedback rating are visible, as is the current bid, starting bid, and and indicator for the reserve not being met by the current bid.

1.3.1 Auction listing and bid data

Due to the prevalence of antique and custom items on eBay, I restrict attention to one of the more popular categories: a brand of stuffed animals called Beanie Babies (BBs). There are approximately 1 million BB auctions in the dataset, with about 2.7 million bids. All distinct varieties of BBs were produced by a single company (TY), which ensures some level of homogeneity among the listed items. Table A.1 presents various summary statistics for the analysis dataset.

Several features make this dataset attractive for empirical analysis. First, both the highest bids and secret reserve prices are recorded for each item in the dataset. Online auction datasets frequently impute secret reserve prices from observable indicators such as the "reserve not met" sign in Figure 1.1; since my focus is on sellers' choice of reserve price, it is helpful to obtain accurate measurements of this choice variable. Highest bids are similarly unobserved in many studies of online auctions due to the ascending minimum bid only depending on the second-highest bid. As will be shown later, the first-highest bid being observable is helpful for identifying and estimating seller beliefs. Other data such as the time of the auction, any promotional choices made by the seller, and seller-provided item descriptions are also included in the dataset.

Additionally, supplemental data on users' feedback reveals all positive, negative, and neutral ratings for accounts. Importantly, this feedback history dates back to the beginning of eBay in 1995; this allows me to construct the reputation variables that are visible to prospective bidders, specifically feedback scores (defined as the number of all feedback events) and rating (defined as the percent of feedback events which are positive). Unless otherwise noted, I use the inverse hyperbolic sine of feedback scores to address the significant skew in user feedback. While feedback scores are not a perfect measurement of the number of auctions in which sellers have participated, they are often used as a proxy for user experience on eBay (see e.g. Simonsohn 2010).

1.3.2 Evidence for learning among new sellers

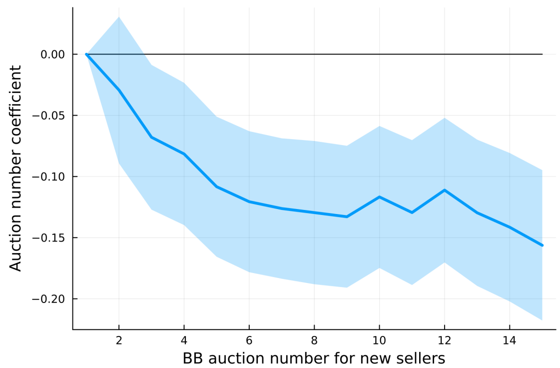
I now present evidence that new sellers are learning to run auctions as they gain more experience, and therefore can be influenced by observing information. I define new sellers as all accounts who have no recorded feedback before the start of the dataset and who list at least one item for sale. Similar to Kim (2020), I examine trends in new sellers' choices and outcomes as a reduced-form test for whether new sellers' behavior varies with experience. I also document how sellers' choice of reserve price becomes less strongly correlated with lagged revenues as they gain more experience, and that seller choices and outcomes are related to how many items they list.

I first show that as new sellers gain more experience, they change in both their choice of reserve prices and the outcomes they face. I plot time trends in seller revenue net of fees, the *effective reserve price* (defined as the maximum of the secret reserve price—if one exists—and the starting minimum bid), and the number of bidders per auction, while controlling for predicted item values, seller feedback, seller rating, and seller and month fixed effects. Throughout, I will homogenize variables such as reserve prices and revenues by dividing by the predicted item value. The process for estimating item values is detailed in section 1.5.1; additional plots in Appendix A.1 replicate these trends without depending on the predicted values.

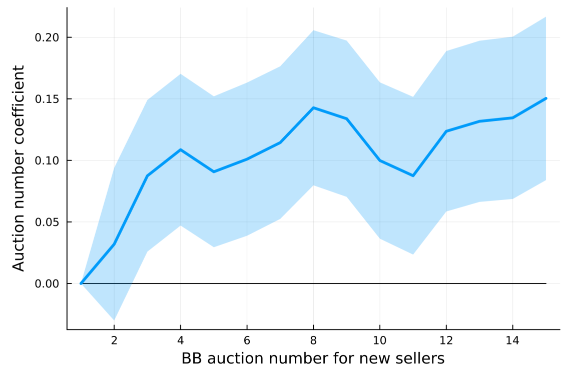
Figure 1.2 plots time trends when restricting the sample of new sellers to those with at least 15 auctions in the data.[3] These new sellers initially set higher prices, earn lower revenues, and attract fewer bidders than they do in later auctions. This pattern is consistent with seller learning when initial beliefs are biased toward setting higher reserve prices. Further, by examining only those who remain on the platform for at least 15 auctions and controlling for persistent seller heterogeneity, these trends do not simply reflect early exit by

3. I chose 15 auctions to avoid including sellers who may have few items in their possession and no interest in long-term trading, as well as to not have too short a panel for estimating seller fixed effects. The trends are similar for different windows.

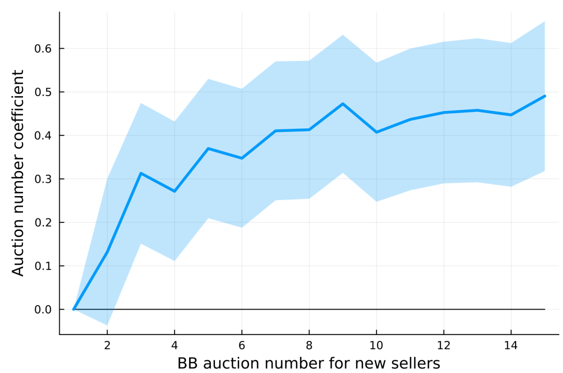
Figure 1.2: Trends in variables of interest as new sellers gain experience



(a) Effective reserve (homogenized)



(b) Net revenue (homogenized)



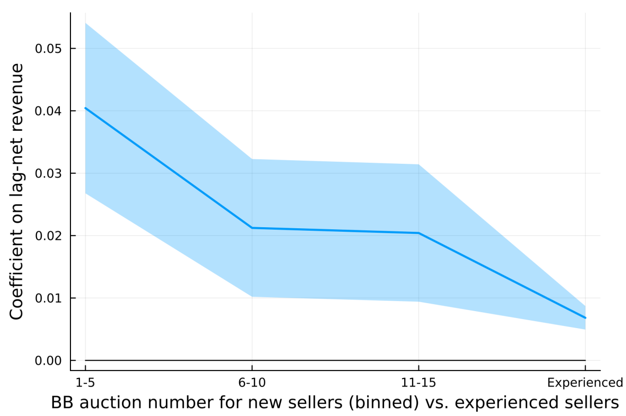
(c) Number of observed bidders

Notes: These regressions pool 1,639 new sellers' first 15 auctions with all auctions by 5,165 experienced sellers (defined as those with a feedback count of at least $\geq$47 at the start of the data, which is the 75th percentile of initial feedback count). The sample is limited to sellers with at least 15 auctions in the data; the results are similar when using different number of auctions. I plot 95% confidence intervals for the estimated coefficients.

high-value sellers.

Standard learning models also predict that the value of additional information decreases as sellers obtain more experience, and that seller beliefs converge toward toward the true parameter. To examine variability in seller choices over time, I regress current-auction reserve prices on lagged net revenue and lagged reserve prices (again conditioning on sellers with at least 15 auctions in the data). Figure 1.3 plots the coefficients of lagged net revenue from this regression, binned by the auction number among new sellers. New sellers' current prices are correlated with past revenue signals, and the magnitude of the lag coefficient is larger in early auctions. This is consistent with subsequent auctions containing relatively less information for more experienced sellers.

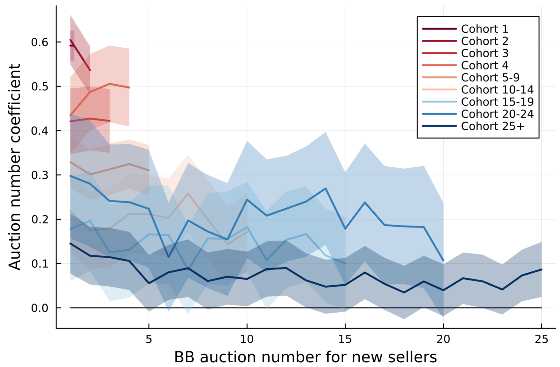
Figure 1.3: Coefficients from regressing reserve prices on lagged revenues, by binned seller experience



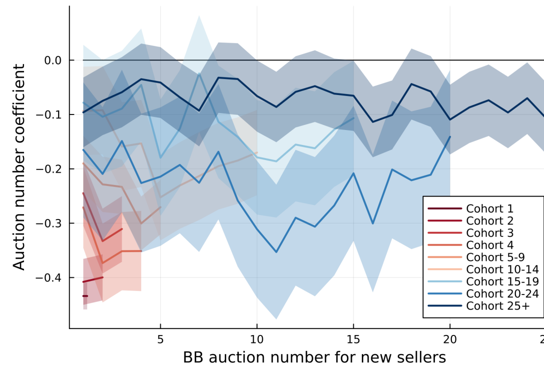
Notes: These are the coefficients when regressing current-period effective reserve price on lagged revenue, multiplied by indicator functions for new sellers being in the first 1-5, 6-10, and 11-15 auctions in the data (among new sellers with at least 15 auctions in the data and experienced sellers with >75th percentile of experience at the start of the data); these are compared with the coefficient of lagged revenue interacted with an indicator function for experienced sellers. The controls include month fixed effects, feedback percentage, predicted item value, and interactions of each experience indicator function with lagged effective reserve price. I plot 95% confidence intervals for the estimated coefficients.

Similar trends hold when examining new sellers more broadly, though the estimates are far noisier. To examine the effect of selection on new seller outcomes, I separate new sellers into cohorts based on the number of items they list. In Figure 1.4, I plot the average difference between new sellers' and experienced sellers' reserve prices and net revenue for each of new bidders' first k auctions. Consistent with selection, new sellers who list relatively few items are also those with significantly higher reserve prices and lower revenues.

Figure 1.4: New seller trends by cohorts of the number of items listed (difference from experienced seller averages)



(a) Effective reserve (homogenized)



(b) Net revenue (homogenized)

Notes: These figures represent a simple difference in means between new sellers in each cohort and experienced sellers (defined as sellers with >75th percentile of experience at the start of the data). This does not include any controls other than month fixed effects. I plot 95% confidence intervals for the estimated coefficients.

Additional results in Appendix A.1 highlight other patterns that are suggestive of sellers learning to set reserve prices. While sellers exhibit trends in some non-reserve price choice varibles (such as featuring items and timing of auctions), the magnitude of these trends are generally small, which motivates sellers' reserve pricing decision as the focus of my analysis. I also show that, while there are some trends in item descriptions, some of the most prominent focus on the item reserve price (in particular, noting the lack of a secret reserve price).

1.3.3 Alternative explanations for new seller behavior

I also consider mechanisms that might drive the trends shown above other than sellers responding to information about auctions. These other mechanisms include behaving strategically to sell higher-value items first, facing higher-value bidders for whom higher reserve price may be optimal, having systematically higher private values for all items when entering the platform, and having an initial endowment effect that decreases with selling experience. I discuss each in turn, considering either theoretical explanations or empirical tests for each possible mechanism.

First, sellers might strategically choose the order of their listings when starting their account, perhaps starting with items they value more highly conditional on an item's book value. However, this strategy runs counter to any dynamic considerations like those discussed in Foster et al. (2016). Upon entering the platform, sellers could instead list items they value less (and set a correspondingly lower reserve price) and yield higher sale probabilities. This would allow sellers to increase their reputation scores and potentially earn more in subsequent auctions. Thus, to the extent that items are listed in decreasing order of sellers' private value, this is unlikely to be done to optimize dynamic profits from listing auctions.

Sellers may face a different set of bidders in their initial auctions, for whom it may be optimal to set higher prices than for later auctions. It is difficult to directly analyze the distribution of bidder values without a model, since the increasing minimum bid and endogenous entry of bidders create a selection problem. To overcome this, I restrict attention to the first and second highest bids (where they exist) and control for the number of observed bids with fixed effects. Figure A.4(a) shows generally flat trends in the first and second highest bids in new sellers' auctions, with a slight but statistically insignificant increase for later auctions. This suggests that initially higher reserve prices are not driven by differentially higher bids.

Additionally, sellers may have systematically higher values for all items upon entering the platform. In this dataset, I can observe sellers whenever they bid for other items and test whether their bids change as they gain more experience on the platform. Figure A.4(b) shows the bids of new sellers for other listings, and illustrates that new bidders do not place higher (or lower) bids for other items upon entering the platform. This is consistent with the underlying distribution of seller valuations remaining constant throughout new sellers' first auctions on the platform.

Finally, new sellers may have an endowment effect that diminishes as they gain additional experience. This story is consistent with several experimental studies in which participating in more transactions decreases the endowment effect (List, 2003, 2004, 2011; Tong et al., 2016). In this observational data, I do not directly observe both measures of willingness to pay and willingness to accept for the same item; further, reserve prices may not reflect a true willingness to accept because sellers try to extract surplus from bidders. However, Figure A.5 shows the distributions of reserve prices for experienced and inexperienced sellers are quite similar for the approximately 50% of listings with lower standardized reserve prices. This is suggestive evidence against new sellers having a strictly higher willingness to accept for all their items. Figure A.6 also shows that new sellers' current reserve prices are more strongly (and negatively) correlated with failing to sell the previous item, while selling the previous item is uncorrelated with current reserve prices. Given evidence from Tong et al. (2016) that the endowment effect dissipates with successful transactions, this pattern does not seem to be driven by a diminishing endowment effect.

1.4 Auction platform model with seller learning

To directly study the extent to which new sellers are influenced by new information, I develop an auction platform model that allows for seller learning. The model adapts the two-sided endogenous entry model from Marra (2019) by allowing seller actions to vary with their individual beliefs about their profit function. I begin by introducing the following notation and assumptions.

A large number $\mathbf{N}_S$ of sellers and $\mathbf{N}_B$ of bidders can choose to participate on a monopoly auction platform. Each potential bidder i has valuation v_{ij} for each item j, where $v_{ij} \sim F_B$. Potential sellers s have outside option values v_{0sj} for each item j they possess, where $v_{0sj} \sim F_S$. Each item has auction-level observables X_j. I assume all prospective sellers and bidders know the valuation distributions F_B and F_S, which satisfy the following assumption.

Assumption 1. The value distributions F_B and F_S are absolutely continuous and have connected support. Bidder values v_{ij} are independent from $v_{i'j}$ for all $i \neq i' \in$

$\{1, ..., \mathbf{N}_B\}$, and seller values $v_{0\ell j}$ are independent from v_{ij} for all $i \in \{1, ..., \mathbf{N}_B\}$ and $\ell \in \{1, ..., \mathbf{N}_S\}$.[4] Further, the bidder value distribution F_B satisfies the strict monotone hazard rate property (i.e., $\frac{f_B(x)}{1-F_B(x)}$ is strictly increasing in x on the support of F_B). Finally, dependence on item j's characteristics takes the form $v \cdot \exp(\gamma(X_j))$, where X_j is exogenous to all values v, which are drawn from the players' respective distributions (F_B or F_S).

Throughout the discussion of the model, I focus on homogenized item values, which is equivalent to setting $\gamma(X_{jt}) = 0$ for all items. Since each item j is sold by a single seller s, I omit dependence of values and other terms on j and s where possible.

I assume all sellers know their values v_0 for the item they own. Sellers can choose to list the item for auction after incurring an item-specific entry cost $c_S^E \overset{\text{i.i.d.}}{\sim} F_{c_S^E}$, which is independent from sellers' private values v_0. After deciding to list the item, sellers also choose an effective reserve price r and minimum bid m. I assume m is exogenously drawn between 0 and r, and unlike Marra (2019), I assume that r is observable to all prospective bidders.[5] Potential bidders can see the item on a listing page and decide whether or not to enter the auction; if they do so, they incur entry cost c_B^E and only then learn their value v_i for the item and costlessly submit a bid. The bidder with the highest bid exceeding the reserve price wins the item, and the transaction price p is equal to the highest of the effective

4. Quint (2017) explores the relationship between reserve prices and outcomes in private and common values settings, and shows that in settings with a common value component the reserve price may be lower than under pure private values. I examine the assumption of independent private values in more detail in Appendix A.2. While this setting (as with many auction settings) likely has both a common value and private value component, I simplify the bidding model to allow more tractability in the higher-level problem of interest, i.e. the platform's decision of how much information to provide to shift seller behavior. In particular, this facilitates a more flexible specification of mean item utilities $\gamma(X_j)$ in estimation (see section 1.5 for more details).

5. In this dataset, only 23.9% of auctions have a secret reserve price, while 89% have an minimum bid higher than the eBay default of $1. Thus, this simplifying assumption is reasonable in the present setting. The "reserve not met" indicator disappears when at least one bid passes the reserve price, at which point the minimum bid has often increased from its starting value. Further, Katkar and Reiley (2007) documents that the "reserve not met" indicator still affects bidder entry even when the reserve price is still secret. Higher reserve prices are less likely to be met, which on averages makes this indicator last for a longer period. A broader literature studies the choice between public and secret reserve prices (Hasker & Sickles, 2010).

reserve price and the second highest bid.

Bidder and seller behavior is also affected by the cost structure of using the platform, denoted by a vector c of all associated costs and fees. In particular, entry costs c_B^E and c_S^E for both bidders and sellers are decomposed into time costs c_B^T and c_S^T and insertion fees c_B^I and c_S^I, respectively. The insertion fees are paid directly to the platform when a seller lists an item or when a bidder enters an auction, regardless of whether the item is sold. The platform can also impose bidder and seller fees c_B^R and c_S^R. If the item is sold, the highest bidder pays $(1 + c_B^R)p$ and the seller receives $(1 - c_S^R)p$, so the platform also receives revenue $(c_B^R + c_S^R)p$ from each successful sale.

The rest of the section is divided into three parts. I first describe bidders' optimal strategy, conditioning on seller behavior. I then examine the seller strategies and how they depend on sellers' beliefs about bidder behavior. Finally, I review the conditions on both bidder and seller behavior that must hold simultaneously in equilibrium.

1.4.1 Bidder strategies

I first focus on the the bidding strategy of actual bidders who enter the auction; these bidders learn their private valuation for each item and then submit bids. This implies an expected continuation value of entering in an auction, which determines bidders' decision of whether to enter an auction.

Bidding stage

After entering an auction and learning their private valuation for the listed item, each bidder's optimal strategy is to bid if they are able to do so. All $\tilde{N}$ bidders who have entered the auction face no cost to submitting their bid, but may be prevented from doing so by the minimum bid m. Following Marra (2019) and Vickrey (1961), all bidders with private value v_i will submit bids $\frac{v_i}{1+c_B^R}$ as long as their bid exceeds the current value of m. Any bidder with a sufficiently low private valuation ($v_i < (1 + c_B^R)m$) will not bid at all, since bidding a higher amount would yield negative expected utility.

I abstract from any potential learning and uncertainty by bidders about how to submit bids. This is because eBay provides an automatic bidding tool that increments the minimum bid up to the maximum value a bidder reveals they are willing to pay. Thus, eBay already implements the optimal bidding rule for all bidders via algorithm. This also allows for a more tractable model of bidder behavior, both for the researcher and for the sellers' mental model of bidder behavior.

Bidder entry stage

Potential bidders determine their expected surplus from entering any given auction, and they enter if there is positive expected surplus. In this setting, in contrast to Marra (2019), I assume the reserve price is public. Thus, bidders make their entry decision based in part on the fee-adjusted reserve price $r^B \equiv (1 + c_B^R)r^*$, which is the winning bidder's payment if the reserve price is binding.

For tractability, I assume that the number of bidders per auction is drawn from a Poisson distribution. This is formalized in the following assumption.

Assumption 2. Assume the expected number of bidders per auction is written as

$$\Lambda(r \mid \delta_0) = \exp(\delta_{0,1} + \delta_{0,2}\rho(r)) \tag{1.1}$$

for some vector δ_0. Further, the number of bidders per auction is Poisson with mean $\Lambda(r \mid \delta_0)$, yielding the following probability mass for n bidders entering an auction with reserve price r

$$p_n(r \mid \delta_0) = \frac{\Lambda(r \mid \delta_0)^n \exp(-\Lambda(r \mid \delta_0))}{n!} \tag{1.2}$$

The Poisson mean Λ, parameterized by the vector δ_0, characterizes bidder arrival for any auction as a function of its reserve price. Note that equation (1.1) is quite general: without any restrictions on ρ, this allows for any strictly positive Poisson mean, which fully characterizes the bidder arrival process. However, this parameterization will prove convenient in the following sections. While it is not conventional for the expected number of bidders in an auction to depend on the reserve price, I allow for this possibility and consider it formally below.

The equilibrium bidding strategy and the assumed bidder arrival process imply an expression for a potential bidder's *ex ante* expected surplus from entering an auction. This is given by

$$\begin{aligned}\pi_B(r \mid \delta_0, c) = \sum_{n=1}^{\mathbf{N}_B - 1} \underbrace{\frac{1}{n}}_{\text{(i)}} \cdot \underbrace{\mathbb{E}\Big[v_{n:n} - (1 + c_B^R) \max\{v_{(n-1):n}, r^*\} \Bigm| v_{n:n} \geq r^B\Big]}_{\text{(ii)}} \\ \cdot \underbrace{(1 - F_B(r^B)^n)}_{\text{(iii)}} \cdot \underbrace{p_n(r \mid \delta_0)}_{\text{(iv)}}\end{aligned} \tag{1.3}$$

where $v_{k:n}$ is the kth highest out of n realizations of v_i. The four components of π_B are (i) the probability that any given bidder has the highest value, (ii) the expected surplus when the highest bidder wins, (iii) the probability that the highest bid exceeds the fee-adjusted reserve price, and (iv) the probability that n bidders enter at the auction.

Under the previous assumptions, the bidder entry game has a unique equilibrium in which bidder entry depends directly on each item's reserve price. The existence and uniqueness of equilibrium follows from Marra (2019), but the presence of a public reserve price yields a new result: the expected number of bidders decreases in the reserve price.

Proposition 1. Let assumptions 1 and 2 hold. Then the bidder entry equilibrium exists and is unique. Further, $\frac{\partial \Lambda(r|\delta_0)}{\partial r} < 0$. (Proof in Appendix A.3)

The intuition for this result is as follows. Potential bidders have zero expected profit (net of their entry cost) from entering an auction, i.e.

$$0 = \pi_B(r \mid \delta_0, c) - c_B^E \tag{1.4}$$

since any positive profit will induce additional entry and negative profit will cause excess bidders to leave. The expected surplus in an any auction decreases in the number of competing bidders, since additional bidders both increase the expected price paid and lower the probability that a new bidder will win the item. Expected surplus is decreasing in the equilibrium number of expected bidders, Λ, and properties of the Poisson distribution imply that a unique value of Λ satisfies the expected zero profit condition for each reserve price r. Finally, expected bidder surplus declines with r, so fewer bidders enter when they expect more competition from the seller.

In light of Proposition 1, I revisit the functional form of the Poisson mean given in equation (1.1). Specifically, the function ρ is now determined by the zero profit condition (1.4) for each possible reserve price r. Without loss of generality, I assume that ρ is a strictly increasing function. Proposition 1 therefore implies that the bidder arrival parameter corresponding to the reserve price, $\delta_{0,2}$, is strictly negative. That is, increasing the reserve price will reduce the expected number of bidders that enter the auction.

1.4.2 Seller strategies

I now examine sellers' equilibrium behavior. I first show how the platform's pricing structure and endogenous bidder entry shape sellers' choice of optimal reserve price when sellers face no uncertainty. I then extend this result to allow for seller uncertainty and learning, and conclude with sellers' entry problem. Throughout, I treat each seller's item-specific entry cost c_S^E as fixed for a given auction.

Optimal reserve price on the platform under perfect information

I begin by writing the general form of the seller profit function to fix ideas. Sellers' expected profit, conditional on the bidder arrival parameter vector δ_0, is given by

$$\Pi(v_0, r \mid \delta_0, c) = (1 - c_S^P) \cdot \underbrace{R(r \mid \delta_0)}_{\mathbb{E}[\text{Revenue}|r,\delta_0]} + v_0 \cdot \underbrace{K(r \mid \delta_0)}_{\mathbb{P}[\text{Keep}|r,\delta_0]} - c_S^E$$

where the functional forms of the expected revenue R and and probability K that a seller keeps the item follow from the second-price auction literature and the Poisson arrival function.[6] The optimal interior reserve price r^* satisfies the first-order condition

$$\psi(r^* \mid \delta_0, c) \equiv \frac{-(1 - c_S^R) R_r(r^* \mid \delta_0)}{K_r(r^* \mid \delta_0)} = v_0$$

where ψ is the virtual type function mapping bids to the space of seller values. Applying the Poisson assumption and the relevant functional forms, and then rearranging the terms of R_r and K_r, yields the following characterization of the optimal reserve price:

Proposition 2. Assume bidder arrival is Poisson with mean $\exp(\delta_{0,1}+\delta_{0,2}\rho(r))$. Then, recalling that the fee-adjusted reserve price faced by bidders is $r^B \equiv (1 + c_B^R)r^*$, the optimal interior reserve price satisfies

$$\frac{v_0}{1 - c_S^R} = \left[r - \frac{1 - F_B(r^B)}{(1 + c_B^R) \cdot f_B(r^B)} - \frac{W_R(r \mid \delta_0)}{f_{\max}^B(r \mid \delta_0)} \right] \frac{f_{\max}^B(r \mid \delta_0)}{f_{\max}^B(r \mid \delta_0) + W_K(r \mid \delta_0)} \tag{1.5}$$

where $f_{\max}^B(r \mid \delta_0) \equiv (1 + c_B^R) \cdot \sum_{n=1}^{\bar{N}_B} p_n(r \mid \delta_0) F_B(r^B)^{n-1} f_B(r^B) n$ is the scaled

6. Formally, these functions are written $R(r \mid \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B} p_n(r \mid \delta_0) R_n(r)$ and $K(r \mid \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B} p_n(r \mid \delta_0) K_n(r)$, where $r^B = (1 + c_B^R) r$ and

$$R_n(r) \equiv nr(1 - F_B(r^B))F_B(r^B)^{n-1} + \frac{n(n-1)}{1 + c_B^R} \int_{r^B}^{\infty} z(1 - F_B(z))F_B(z)^{n-2} f_B(z) dz$$
$$K_n(r) \equiv F_B(r^B)^n$$

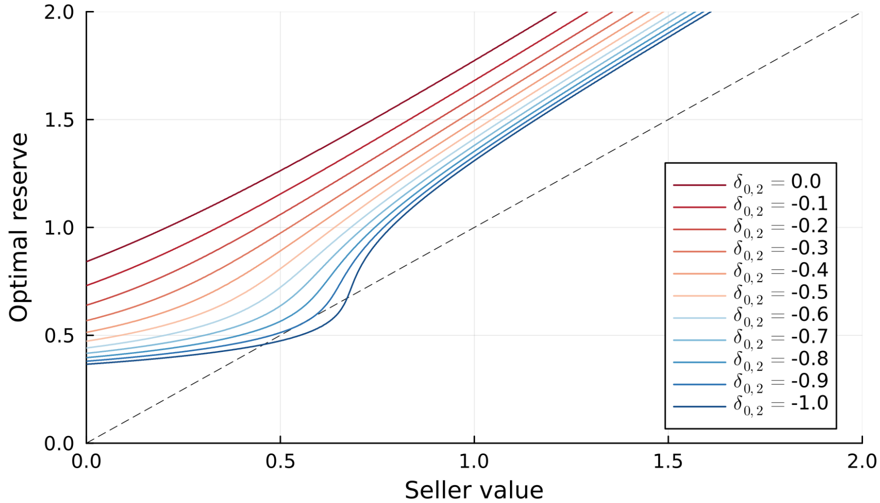
density of the highest bid at r given δ_0, and $W_R(r \mid \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B}(\frac{\partial p_n(r|\delta_0)}{\partial r})R_n(r)$ and $W_K(r \mid \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B}(\frac{\partial p_n(r|\delta_0)}{\partial r})K_n(r)$ are weighted averages of the expected revenue and keep probabilities for each n.

Note that when the bidder arrival process does not depend on the reserve price (that is, $\delta_{0,2} = 0$), it holds that $W_R = W_K = 0$. When the bidder and seller fees are also zero ($c_S^R = c_B^R = 0$), equation (1.5) reduces to the Myerson (1981) optimal reserve price formula.

To provide intuition for this pricing rule, Figure 1.5 plots the implied reserve price from Proposition 2 for different values of the bidder arrival coefficient $\delta_{0,2}$. The figure starts from the baseline case of $\delta_{0,2} = 0$ and shows increasingly negative values of $\delta_{0,2}$ in comparison. As $\delta_{0,2}$ decreases, sellers should optimally lower the markup in their reserve price. In this particular case, it may even be optimal to set a reserve price lower than the seller's own value if the bidder deterrence effect is sufficiently strong. The precise shape of the optimal reserve price function depends on the intercept $\delta_{0,1}$ of the log-mean of average bidders, as well as the shape of the bidder value distribution F_B and the function ρ that is determined by bidders' zero profit condition (1.4).

This equation generalizes several results from the related auction literature. As previously noted, it nests the Myerson (1981) reserve price formula in the case where $\delta_{0,2} = 0$ and carries with it the same intuition: sellers may set a higher reserve to extract additional surplus from bidders. Sellers' influence over the arrival process means they explicitly weigh the expected benefit of surplus extraction against the expected benefit of a potential additional bidder (though an extra bidder would be better in expectation, as in Bulow and Klemperer 1996, this arrival is not guaranteed). Thus, this equation echoes the argument in Engelbrecht-Wiggans (1987), in that sellers may wish to lower the reserve to attract more bidders. Since there is some probability that few bidders enter each auction, however, the seller may still "protect" some expected surplus by setting a non-zero reserve price.

Figure 1.5: Optimal reserve price for varying reserve price coefficients $\delta_{0,2}$



Notes: These figures show optimal reserve price functions for varying parameters of $\delta_{0,2}$, keeping $\delta_{0,1} = 1.0$ fixed, where log-bidder values are normally distributed with mean 0 and variance 0.4, and $\rho(r) = r \cdot F_B(r)$. The dashed 45-degree line represents all values where the seller value is equal to the optimal reserve price.

Optimal reserve price on the platform under seller uncertainty

I now examine sellers' strategies when they do not know the true vector δ_0 that parameterizes the bidder arrival process. This uncertainty arises because δ_0 is a part of bidders' equilibrium play in the "search" stage rather than "item" stage of the game. Any uncertainty about or inattention to the platform's search algorithm, bidders' search strategies, or bidders' search costs could therefore contribute to uncertainty about sellers' own effect on prospective bidders' search process. For example, Simonsohn (2010) documents that eBay sellers may not understand the impact of competition on their own profits and over-enter when market activity is high. This competition neglect is related to seller behavior in this setting, as sellers may set reserve prices too aggressively (in essence, competing) for their own items. These sellers may fail to realize the extent to which this behavior crowds out potential bidders, each of whom may choose another way to spend their time instead of entering an auction with an uncertain payoff.

The previous derivations can be extended straightforwardly in the case where sellers have some belief density b about the true value of δ_0. In an abuse of notation, I define subjective expected profit as $\Pi(v_0, r \mid b, c) \equiv \int \Pi(v_0, r \mid \delta, c) b(\delta) d\delta$. The respective profit function components $R(r \mid b)$ and $K(r \mid b)$ are defined similarly, implying the subjective virtual type function $\psi(\cdot \mid b, c)$ from rearranging the first-order condition of $\Pi(v_0, r \mid b, c)$ with respect to r. For tractability, I assume all sellers are myopic, so they only maximize current-period profits conditional on their beliefs b and do not actively experiment.[7]

As is standard in models of learning, sellers also have a model of the true data-generating process, from which they learn about the unknown parameter δ_0. I assume sellers update their beliefs after every auction they run. All sellers believe profit draws y are generated by the process

$$y = \Pi(v_0, r \mid \delta_0, c) + \epsilon \tag{1.6}$$

where ϵ is drawn i.i.d. from some distribution $F_{\epsilon|r,v_0}$ that is known to sellers. This randomness comes from the stochastic nature of bidders' arrival and their idiosyncratic valuations for each item. Each auction, sellers observe the associated data $\mathbf{D} = \{y, v_0, r\}$, and use the likelihood function implied by equation (1.6) and the noise distribution $F_{\epsilon|r,v_0}$. After each auction, sellers use a deterministic transition function $\mathcal{T}$ (such as Bayes' rule) to update their prior beliefs b to the posterior b':

$$b'(\delta) = \mathcal{T}(b(\delta), \mathbf{D}) \tag{1.7}$$

Thus, the evolution of sellers' reserve price strategies depends entirely on their beliefs b as updated after each auction, which in turn are driven by variation in data observed after each

7. This assumption is used by Huang et al. (2020) to simplify the analysis of firms learning from demand signals. This also relates to the "anticipated utility model" discussed in Cogley et al. (2007); in their setting, the model without active experimentation is a good approximation for the fully dynamic Bayesian model. While this assumption is not strictly necessary, it is useful to reduce the computational burden of estimating this model in the presence of other setting-specific computational challenges, and so I impose it from the outset in building the model; see section 1.5 for more details on the estimation procedure and its implementation.

auction. I note that while sellers may be Bayesian, this is not necessary for the conclusions of the model.

Seller entry stage

Prospective sellers will enter the platform as long as their net expected surplus (relative to never listing the item) is positive. This yields the following inequality for sellers with beliefs b:

$$\Pi(v_0, r^* \mid b, c) - v_0 \geq 0 \tag{1.8}$$

Sellers' expected surplus depends crucially on their choice of reserve price conditional on their beliefs and their private value for the item sellers may list. The following proposition provides a condition under which the optimal reserve price r^* is increasing in sellers' private value v_0 for any seller with beliefs b.

Proposition 3. Assume $K_r(\cdot \mid b) > 0$. Then r^* is increasing in v_0 and the virtual type function $\psi(\cdot \mid b, c)$ is increasing. (Proof in Appendix A.3)

This monotonicity is common and easily verified under the functional forms assumed in much of the related literature; I make this explicit because it depends on sellers' underlying beliefs b and it is important for subsequent propositions.

Proposition 3 implies a threshold rule for sellers, such that sellers with a private value above that threshold will not list an item for auction. Given the functional form of R and the monotonicity of r^* in v_0, sellers' gains from trade $\Pi(v_0, r^* \mid b, c) - v_0$ are strictly decreasing in v_0. This implies the following characterization of the seller entry problem.

Proposition 4. There exists a unique threshold $\bar{v}(b, c)$ for each belief density b such that sellers with beliefs b will list their item for auction if and only if $v_0 \leq \bar{v}(b, c)$.

That is, sellers will only select into the platform if their values are sufficiently low given their beliefs, in which case it appears profitable to list items for auction. This holds regardless of heterogeneity across seller beliefs, since each auction's reserve price is public knowledge and bidder entry is determined by their zero profit condition. Further, heterogeneity in beliefs does not directly impact the bidder arrival problem: bidders are only impacted by the reserve price r^* instead of the underlying values of v_0 and b.

This threshold condition is similar to other results in the literature, and corresponds to Marra (2019) when beliefs b are common across all sellers and a point mass on the true parameters. I weaken the assumption that this threshold is objectively correct under the true item values: it needs only be optimal according to each seller's private beliefs about the true arrival process. This ensures that a unique equilibrium exists for the seller entry game even under heterogeneous and potentially biased beliefs among sellers.

1.4.3 Equilibrium definition

Before proceeding, I review the conditions for both bidders and sellers that must hold in equilibrium. Given a distribution F_B of bidder values, a distribution F_S of seller outside option values, a fixed bidder entry cost c_B^E, a distribution $F_{c_S^E}$ of seller entry costs, bidder and seller transaction fees c_B^R and c_S^R, initial ($t = 0$) prior beliefs b_0 about the parameters of bidders' entry process, and an updating rule $\mathcal{T}$ by which sellers update their beliefs as in (1.7), equilibrium consists of

(i) the threshold bidding rule as defined in 1.4.1

(ii) the bidder arrival parameter δ_0 that determines mean bidder arrival $\Lambda(r \mid \delta_0)$ as in (1.1)

(iii) the seller reserve price rule defined by $\psi(r^* \mid b, c) = v_0$ for any seller beliefs b

(iv) the seller entry threshold $\bar{v}(b, c)$ implied by (1.8)

such that actual bidders maximize expected profit from participating in an auction, potential bidders earn zero expected profit from entering any auction, sellers maximize expected profits given their beliefs about the bidder entry process, and the marginal seller earns zero expected profit given these beliefs.

The equilibrium conditions highlight the importance of seller beliefs and information on the auction platform. Sellers' beliefs determine both their entry decision and their choice of reserve price conditional on entry. Further, the degree of dispersion in seller beliefs affects how they update their beliefs when receiving new information. In turn, bidders' entry decisions and whether they submit bid depend on sellers' choice of reserve price. Thus, all actions on the platform are shaped by sellers' beliefs and the information they receive about bidder behavior.

1.5 Identification and estimation

The model from the previous section highlights two important ways in which sellers' beliefs and information affect platform outcomes. First, sellers' beliefs determine both their decision of whether to list an item. Second, sellers' beliefs determine how they choose a reserve price to attract bidders and extract bidder surplus. I now estimate the model to evaluate the role these features play in the empirical patterns, and how sellers' information affects their actions and platform outcomes.

Estimation proceeds in two parts. First, I estimate the demand side of the model, consisting of bidder arrival parameters, the distribution of bidder valuations F_B, and heterogeneity in mean item valuations. To do this, I derive a likelihood function for the demand side that corrects for selection in which bids are observed due to the increasing minimum bid rule. I also use results from the literature on debiased machine learning to flexibly estimate observable item heterogeneity. Using bidders' zero profit condition and the estimated parameters, I then construct the expected surplus per auction and recover bidders' time cost.

Second, I estimate the supply side of the model using data from both experienced and inexperienced sellers. Assuming experienced sellers have more accurate beliefs about bidder arrival, I use their first-order condition to recover seller values for all listed items. I then combine these estimated seller values with the seller entry condition to obtain the distributions of seller values and entry costs. Using these estimated distributions and the distribution of new seller reserve prices, I identify and estimate new sellers' prior beliefs about the bidder arrival process.

1.5.1 Demand side estimation strategy

I construct a likelihood to estimate the parameters of the demand model, which include both a high-dimensional component and a low-dimensional component. The high-dimensional parameter γ accounts for observable heterogeneity in item values and is a flexible function of the item text descriptions. The low-dimensional parameter vector ϑ_d consists of the bidder arrival parameters, as well as the distributions F_B of bidder values distribution and the distribution F_r of reserve prices. I first describe key components of the likelihood, including how to correct for bias due to the presence of high-dimensional parameters, and then present the likelihood itself.

Heterogeneity in item values

The high-dimensional parameter γ accounts for heterogeneity in item values as a function of each item's text description. Specifically, the mean of bidders' log-valuations for item j is written as $\gamma(X_j)$ for item-specific data X_j. The data X_j is a large collection of indicator variables for each month in which an item may be listed and for whether the text description includes one of the 5,368 words that appear in at least 10 item descriptions (more details on the data construction and estimation process may be found in Appendix A.4). This data, though unstructured, provides an extremely detailed view of item characteristics that are relevant to potential bidders.

I model γ as a neural network, which can flexibly account for item heterogeneity in the demand model but also introduce bias. This bias appears because the function of interest may be very complex relative to the amount of data, and it is possible to overfit the neural network. While model selection and regularization can be used to avoid overfitting γ, this also adds bias in the estimated item heterogeneity. This bias distorts the demand-side likelihood—shifting where in the parameter space the likelihood is maximized—which in turn biases the estimates of low-dimensional demand parameters ϑ_d. In order to consistently estimate ϑ_d, it is necessary to correct the estimation problem to make it insensitive to errors in the high-dimensional parameter γ.

I derive an orthogonal score to both flexibly estimate item heterogeneity and avoid the resulting bias in the structural parameters of interest. Since it is known how the high-dimensional parameter γ enters the structural model, its effect on the likelihood can be explicitly solved for. As in Chernozhukov et al. (2022), Farrell et al. (2020), and Ichimura and Newey (2022), I derive an influence function that measures the effect of varying γ on the first-order condition that determines ϑ_d. This influence function is then used as a correction term to make the estimation process insensitive to errors in the estimated high-dimensional parameter.

The orthogonal score contains two parts. The first part is the original score, or the gradient of the likelihood with respect to ϑ_d. In standard settings, without complications from a high-dimensional term, this first-order condition for the likelihood is used to pin down the parameters of interest. The second part of the orthogonal score is an adjustment term, or the influence function. This term involves a projection of how the likelihood varies with γ onto the space of the original score; this "partials out" the influence of γ when estimating ϑ_d in an analogue to the Frisch-Waugh-Lovell theorem. I use this orthogonal score to estimate ϑ_d via generalized method of moments. Further details on the debiasing procedure are shown in Appendix A.5.

Bidder arrival process

I assume bidders' entry equilibrium varies with observable seller and item characteristics. Specifically, I parameterize the expected number of bidders to enter an auction as

$$\Lambda_j \equiv \Lambda(r_j, Z_j \mid \lambda, \delta_0) = \exp(\delta_{0,1} + Z_j'\lambda + \delta_{0,2}\rho(r_j)) \tag{1.9}$$

where Z_j contains the seller feedback score, seller ratings, and the average log item value $\gamma(X_j)$. For tractability, I assume the additional arrival parameter λ is known to all sellers, regardless of their level of experience. To streamline notation, I combine the known arrival shifter $Z_j'\lambda$ with the intercept to write the item-specific arrival parameter $\delta_{j,0,1} \equiv \delta_{0,1} + Z_j'\lambda$, with $\delta_{j,0} \equiv \{\delta_{j,0,1},\ \delta_{0,2}\}$. I also use similar notation for beliefs, where b_j represents sellers' item-specific beliefs about bidder arrival with $\mathbb{E}_{b_j}[\delta_1] = \mathbb{E}_b[\delta_1] + Z_j'\lambda$.

The function ρ is determined by the bidder zero-profit condition at each point in the support of r. Since evaluating the high-dimensional index γ is itself computationally challenging, I opt for a reduced-form representation of the entry process instead of solving the fixed-point problem for ρ while estimating γ. I set ρ to be the product of the identity function and the CDF of the bidder value distribution evaluated at r, i.e. $\rho(r) = rF_B(r)$. The intuition for this choice comes from the zero profit condition for bidder entry and Proposition 1. The expected bidder surplus from entering the auction depends on the probability that their bid will be below the reserve price, $F_B(r)$; it is also strictly decreasing in r, which is the price paid when the reserve is binding. I also estimate the model with alternative specifications for ρ and find similar results.

Likelihood function for demand side

eBay's use of an increasing minimum bid creates a selection problem in which not all bids are observed for any given item. Due to randomness in the order of bidder arrival, some bidders will enter the auction after others have placed bids, learn their valuation, and not submit a bid because the minimum bid is already higher than their value for that item. In this and other online auction settings, the minimum bid will increase to the second highest of the existing bids and "lock out" subsequent arrivals (Freyberger & Larsen, 2022; Platt, 2017). This implies that only the two highest bids in an auction are known to correspond to the two highest-value bidders.

To counteract this problem, I derive a likelihood that explicitly models the selection process (see Appendix A.6 for a detailed derivation). I first denote $v_j^{(k)}$ as the kth highest homogenized log-bid. Using the Poisson assumption on bidder arrival, I model the distributions of the reserve price r_j and the highest two bids $v_j^{(k)}$ when these bids are observed.[8] This likelihood also conditions on the starting minimum bid m_j (which is binding when there are fewer than two bids), the number N_j of bids observed, arrival shifters Z_j, and components of observable heterogeneity X_j. Together, the demand likelihood contribution for a single auction is

$$\begin{aligned}
\ell_j^{\text{demand}}(\vartheta_d, \gamma \mid N_j, \{v_j^{(k)}\}, X_j, Z_j, m_j, r_j) = \\
& f_r(r_j \mid \vartheta_d, \gamma) \cdot \left[e^{-\Lambda[1-F_B(m_j|\vartheta_d,\gamma)]}\right]^{\mathbb{1}[N_j=0]} \\
& \cdot \left[f_B(v_j^{(1)} \mid \vartheta_d, \gamma)\Lambda e^{-\Lambda_j[1-F_B(m_j|\vartheta_d,\gamma)]}\right]^{\mathbb{1}[N_j=1]} \\
& \cdot \left[f_B(v_j^{(1)} \mid \vartheta_d, \gamma) f_B(v_j^{(2)} \mid \vartheta_d, \gamma)\Lambda_j^2 e^{-\Lambda_j[1-F_B(v_j^{(2)}|\vartheta_d,\gamma)]}\right]^{\mathbb{1}[N_j\geq 2]}
\end{aligned} \tag{1.10}$$

Appendix A.6 shows how this likelihood approach, together with the orthogonalization step, performs with simulated data.

After estimating the bidder value distribution F_B and the other bidder arrival parameters, I estimate the bidder entry cost c_B^E. For each auction, I compute the ex ante expected surplus $\pi_B(r \mid \Lambda, c)$ and use the bidder zero-profit condition in (1.4) to estimate bidders'

8. Note that r_j and $v_j^{(k)}$ are not observed directly; rather, each is a residual representing its counterpart in the data after being homogenized using the item value index $\gamma(X_j)$.

entry cost as the mean of the expected bidder surplus across all auctions. Since bidder insertion fees c_B^I are zero in the dataset, this implies that the full bidder entry cost is in fact the time cost c_B^T.

1.5.2 Supply side estimation strategy

I now turn to the problem of identifying and estimating the supply side of the model. The remaining parameters of interest are the distribution F_S of sellers' outside option values, sellers' entry cost distribution $F_{c_S^E}$, and new sellers' beliefs about the bidder arrival parameter δ_0. While seller values and costs have been estimated in many similar settings, estimating seller beliefs is key to understanding the role of information and learning on the auction platform.

Seller value and entry cost distributions

I first estimate all non-belief seller parameters (denoted ϑ_s) using data from experienced sellers. I make the simplifying assumption that the most experienced sellers have perfect information about the arrival process, so plugging the experienced sellers' reserve prices into the virtual type function $\psi(\cdot \mid \delta_0, c)$ yields the imputed seller values $\hat{v}_{0j}$ for all listed items.

In order to estimate the seller parameters ϑ_s, I make several functional form assumptions. I assume sellers' entry costs are i.i.d. Exponential, their values v_0 are from a two-component i.i.d. Gaussian mixture, and the nuisance distribution of entry parameters $\delta_{j,0,1}$ are i.i.d. Gaussian.[9] I also introduce some additional notation. First, denote $\tilde{c}(z)$ as the vector of seller costs where c_S^E is replaced with z and $\Pi^*(v_0 \mid \delta_0, c)$ as the maximized profit given seller

9. Since neither sellers' entry costs nor the full population of potentially-listed items are observed, I rely on stronger parametric assumptions to pin down the right tail of seller values. I do not assume sellers and bidders have the same value distribution, since the bidder value distribution consists of all bidders with a weakly positive valuation for the items, and the sellers are a selected sample of individuals who have by some means already acquired items.

value v_0 and costs c. Further, denote by $\bar{v}(\delta_{j,0}, c)$ the experienced seller entry threshold with known bidder entry parameter $\delta_{j,0}$ and costs c.

The likelihood contribution of a single auction run by an experienced seller is the density of the observed data, multiplied by a selection correction term to account for sellers' entry problem. The observed data includes both the implied seller values $\hat{v}_{0j}$ and the baseline arrival parameter $\delta_{j,0,1}$. There is random truncation in which data is observed due variation in entry costs c_S^E: each item is only listed if the entry cost is below the maximum expected seller surplus (that is, with zero entry cost). The probability that this occurs for a specific imputed item value $\hat{v}_{0j}$ is divided by the probability that any seller lists their item, for all possible entry costs and arrival parameters. Taken together, this yields the likelihood contribution

$$\ell_j^{\text{supply-e}}(\vartheta_s \mid \hat{v}_{0j}, \delta_{j,0,1}) = \frac{f_s(\hat{v}_{0j} \mid \vartheta_s) \cdot f_{\delta_{j,0,1}}(\delta_{j,0,1} \mid \vartheta_s) \cdot F_{c_S^E}(\Pi^*(\hat{v}_{0j} \mid \delta_{j,0}, \tilde{c}(0)) - \hat{v}_{0j} \mid \vartheta_s)}{\int_0^\infty \left[\int_{-\infty}^\infty F_S(\bar{v}(l, \tilde{c}(z)) \mid \vartheta_s) f_{\delta_{0,1}}(l \mid \vartheta_s) dl\right] f_{c_S^E}(z \mid \vartheta_s) dz}$$

Since seller entry costs are unobserved, I further assume that the mean of the seller entry cost distribution is equal to $c_S^I + c_B^T$, where the mean platform entry fees are observed in the data and bidders' time cost is identified from demand-side data. Variation in $\delta_{j,0,1}$ shifts the resulting distribution of sellers' participation thresholds and traces out the distribution of seller values.

Identifying new sellers' prior beliefs

Having obtained the sellers' value distribution, I use variation in new sellers' reserve prices to identify their underlying beliefs about the bidder arrival process. At a high level, a seller's choice of reserve price is determined by both their beliefs and their private underlying value for the item. New information, including from previous auctions, affects current reserve prices only through beliefs, so variation in sellers' private values for sellers who receive the same information yields useful variation in reserve prices. The distribution of reserve prices can then be inverted to recover the underlying distribution of seller beliefs.

The following proposition offers formal conditions under which seller beliefs b_{δ_2} about the effect of reserve prices on arrival can be semiparametrically identified from reserve price data.

Proposition 5. Denote history $\mathcal{H}$ as a collection of data $\mathbf{D}$ from auctions. Let Assumption 1 hold, assume F_S is known, and further assume

(i) The prior b is common to all sellers with a shared history $\mathcal{H}$.

(ii) b is composed of independent marginal densities b_{δ_1} and b_{δ_2}, where both the marginal b_{δ_1} corresponding to the intercept and the seller entry threshold $\bar{v}(b, c)$ are known.

(iii) b_{δ_2} satisfies the Carleman condition, i.e. the absolute moments of b_{δ_2} (written as $\mu_j = \int |\delta|^j b_{\delta_2}(\delta_2) d\delta_2$) are finite for all $j \geq 1$ and satisfy $\sum_{j=1}^{\infty} \mu_j^{-1/j} = \infty$.

(iv) The reserve price r has full support on a positive-measure interval including some x^* for which $\rho(x^*) = 0$ and $\psi(x^* \mid b, c) < \bar{v}(b, c)$.

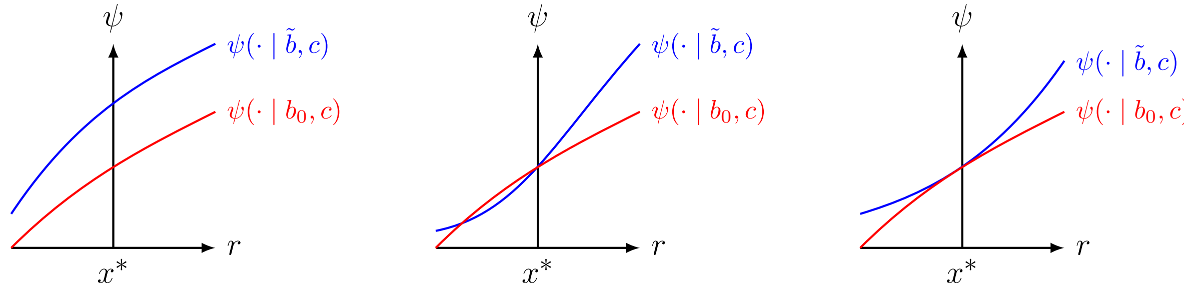
(v) For each function $\xi_k(\beta, x^*, c)$ defined as $\frac{\partial^{k-1}}{\partial x^{k-1}} \psi(x^* \mid b, c)$ with all terms of the form $\sum_{n=0}^{\infty} G_n(x^*) \frac{\partial^k}{\partial x^k} p_n(x^* \mid b)$ replaced with $\beta \sum_{n=0}^{\infty} G_n(x^*) \frac{\partial^k}{\partial x^k} p_n(x^* \mid b)$ for $G \in \{R, K\}$, ξ_k is invertible in β.

Then the marginal density b_{δ_2} of a seller with history $\mathcal{H}$ is identified up to its first $\bar{k}$ moments. (Proof in Appendix A.3)

The proof proceeds in two parts, for which I give a heuristic explanation here. The first step is to recover the virtual type function ψ shared by all sellers with the same information and therefore same beliefs. All reserve prices satisfy the first-order condition $v_0 = \psi(r \mid b, c)$, where the distribution F_S of v_0 is known. Given sellers' participation threshold $\bar{v}(b, c)$, the distribution of observed reserve prices can be written as a known, invertible function of $\psi(\cdot \mid b, c)$.

The next step of the proof is to invert the virtual type function to recover the marginal

Figure 1.6: Intuition for identification of beliefs b_0 through variation in ψ and its derivatives



(a) Identifying $\mathbb{E}_{b_0}[\delta_2]$ by evaluating $\psi(x^*|b, c)$

(b) Identifying $\mathbb{E}_{b_0}[\delta_2^2]$ by evaluating $\psi'(x^*|b, c)$

(c) Identifying $\mathbb{E}_{b_0}[\delta_2^3]$ by evaluating $\psi''(x^*|b, c)$

Notes: Each panel shows a thought exercise of comparing the virtual type function under the true belief density b_0 with the virtual type function under various candidate belief densities $\tilde{b}$ that share the same lower-order moments (i.e., in panel (b) the two densities have the same first moment, and in panel (c) the two densities have the same first two moments). Examining the virtual type function at x^* discards all candidates $\tilde{b}$ such that $\psi(x^* \mid \tilde{b}, c) \neq \psi(x^* \mid b_0, c)$, while examining the first derivative of the virtual type function of x^* discards all candidates $\tilde{b}$ such that $\psi'(x^* \mid \tilde{b}, c) \neq \psi'(x^* \mid b_0, c)$, and so on.

density b_{δ_2} of seller beliefs about the effect of the reserve price on bidder arrival. Figure 1.6 provides some intuition for this identification problem: the behavior of the virtual type function around a particular reference point $r = x^*$ reveals the moments of the belief density of interest. A key insight from Fox et al. (2012) is that evaluating a function and its derivatives at some carefully chosen point x^* (that is, where $\rho(x^*) = 0$) can greatly simplify these functions. This yields known functions of *only* model primitives that do not depend on b_0 (such as $F_B(x^*)$) and various raw moments of the true density b_0. These functions can then be inverted to recover the moments of the underlying belief density. Higher-order derivatives of ψ with respect to the reserve price can be inverted to obtain higher-order moments of b_{δ_2} and reject other densities that do not yield the same virtual type function.

The identification result is general since an arbitrary number of moments of the marginal belief distribution are identified. However, it is important to note that these beliefs are only identified in the context of the larger parametric model. It is not necessary to assume that sellers are Bayesian: identification of b_{δ_2} is achieved with data from a single period for all sellers that observe the same data history $\mathcal{H}$. In principle this allows researchers to test whether sellers follow different candidate learning rules, though this is outside the scope of this paper.

Though Proposition 5 offers semiparametric identification of b_{δ_2}, the necessary assumptions are somewhat restrictive. First, Assumption 1 imposes that all sellers' outside option values are drawn i.i.d. from the same distribution, which rules out time-invariant heterogeneity in sellers' value distributions. Also, assumption (i) of the proposition rules out unobserved determinants of beliefs to allow comparisons between sellers that receive identical information. Assumption (ii) is also restrictive: other components of the sellers' decision problem must be known and separable in some sense (here, independence of marginal beliefs) to isolate the effect of beliefs about any one parameter. Assumptions (iii) and (iv) are similar to assumptions in the random coefficients literature, principally Fox et al., 2012; these use variation in a linear index to recover population densities, while I study an individual's belief density. Assumption (v) is also technical, and requires the derivatives of the virtual type function to be invertible in weighted sums of the derivatives of bidder arrival probabilities. This is a joint restriction on the seller beliefs about $\delta_{0,1}$ and the bidder value distribution F_B at the point x^* from assumption (iv).

This result is related to others in the literature on identifying individual beliefs in structural models. Lu (2019) shows that state-dependent beliefs can be identified in a setting with finite support and Bayesian updating; in contrast, I do not require Bayesian updating and allow for absolutely continuous density functions. Wang et al. (2024) likewise adopts a finite-support approach with Bayesian updating, which is used to identify beliefs about time-varying macroeconomic trends; Wang and Yang (2024) offers more general results in finite-support settings for both myopic and forward-looking agents. Aguirregabiria and Magesan (2020) semiparametrically identifies firm beliefs within a game, and similarly relies on a finite support. While they do not require Bayesian updating, they require beliefs to align with the truth in some cases.

Estimating new sellers' prior beliefs

Though beliefs are semiparametrically identified for each history $\mathcal{H}$ of auction data, in practice I make several additional assumptions on the prior and updating process for computational tractability. The model as written has no conjugate prior structure, so the posterior depends on an increasingly large state space of all possible profit signals, reserve prices, private seller values, and covariates for all sellers' previous auctions. Further, solving for each seller's path of posterior distributions is computationally costly, especially when searching for the parameters of the true prior distribution.

I simplify the state space by using a parametric approximation to Bayesian updating, and I solve the computational bottleneck by approximating this updating process with neural networks. I assume sellers' initial beliefs about δ_0 are bivariate normal, with parameters jointly denoted as ϑ_b. I also assume beliefs are updated according to a modified Laplace approximation to Bayes' rule: each period, sellers' beliefs are a bivariate normal with mean equal to the maximum a posteriori estimate of the true Bayesian posterior and covariance matrix given by the curvature of the true Bayesian posterior at the maximum a posteriori estimate. This assumption helps solve a major computational challenge, since the prior is not conjugate with the posterior due to the nonlinear dependence of expected profit on the bidder arrival parameters.[10] Rather than solving each updating step for all individuals and auctions in the data (a prohibitively slow process), I simulate possible beliefs, covariates, and signals and fit a neural network to approximate the updating process for new sellers' beliefs. Appendix A.7 discusses additional details of the estimation procedure, particularly the steps taken to approximate the updating rule and other associated functions.

I also exploit the structural model in estimation to determine other variables of interest.

10. In addition to this nonlinearity, each sellers' path of beliefs about the arrival parameter can evolve differently according to their signals. The Laplace approximation is one of several posterior approximations used in Bayesian statistics, and imposing that beliefs are updated in this manner ensures that learning follows a computationally tractable Markov process with a relatively low-dimensional state.

I use $\psi(r^* \mid b, c)$ as a control function for seller values v_0 since they are an unobservable but critical component of sellers' updating and decision processes. Assuming that all beliefs are bivariate normal helps pin down b_{δ_1}, which allows me to relax the independence assumption in Proposition 5. Together, parametric assumptions on the entry cost distribution and sellers' prior yield the distribution of sellers' entry threshold $\bar{v}(b, c)$.

Having established the identification of new sellers' prior parameters, I now explain the likelihood approach used for estimation. I use a change of variables to obtain the density of the new sellers' reserve prices from their underlying value distribution. The density of sellers' reserve prices, conditioning on beliefs b_t and costs $\tilde{c}(0)$,[11] is obtained from the distribution of seller values

$$\frac{\partial}{\partial x}\mathbb{P}[r \leq x] = \frac{\partial}{\partial x}\mathbb{P}[v_0 \leq \psi(x \mid b_t, \tilde{c}(0))] = f_s(\psi(x \mid b_t, \tilde{c}(0)) \mid \vartheta_s) \cdot \psi'(x \mid b_t, \tilde{c}(0))$$

where $\psi'(x \mid b_t, \tilde{c}(0))$ is the first derivative of the virtual type function with respect to x. As with the experienced-seller likelihood, there is a selection term to account for the probability of each item being listed by any given seller: the numerator is the probability that an item is listed conditional on $\hat{v}_{0j}$ (i.e., the seller has a favorable-enough entry cost to list the item) and the denominator is the probability that any seller with beliefs b_t lists an item. The likelihood contribution of a single auction of item j run by an inexperienced seller in their tth auction is therefore

$$\ell_{jt}^{\text{supply-i}}(\vartheta_b \mid \mathbf{D}_{jt}, \vartheta_s) = f_s(\hat{v}_{0j} \mid \vartheta_s) \cdot \psi'(r_j \mid b_t, \tilde{c}(0)) \cdot \frac{F_{c_S^E}(\Pi^*(\hat{v}_{0j} \mid b_t, \tilde{c}(0)) - \hat{v}_{0j} \mid \vartheta_s)}{\int_0^\infty F_S(\bar{v}(b_t, \tilde{c}(z)) \mid \vartheta_s) \cdot f_{c_S^E}(z \mid \vartheta_s) dz}$$

$$\begin{aligned} s.t. \quad & b_{t+1} = \mathcal{T}(b_t, \mathbf{D}_{jt} \mid \vartheta_b) \; \forall t \\ & \hat{v}_{0j} = \psi(r_j \mid b_t, \tilde{c}(0)) \end{aligned}$$

11. The entry cost is irrelevant for the virtual type function and can be treated as zero; see 1.4.2 for reference.

Though beliefs in this model can be identified from data within a single period, additional variation across and within different sellers is pooled via the assumed learning rule to aid in estimation.

1.5.3 Estimates from the platform model

Table 1.1 presents the estimated bidder arrival parameters. As predicted by the model, the reserve price coefficient $\delta_{0,2}$ is negative, indicating that higher reserve prices deter potential bidders from entering auctions. The relationship between bidder entry and the covariates is intuitive as well: more bidders enter when sellers have better ratings and more feedback, as well as when the estimated mean item value is higher.

Table 1.1: Estimated bidder arrival parameters and entry cost

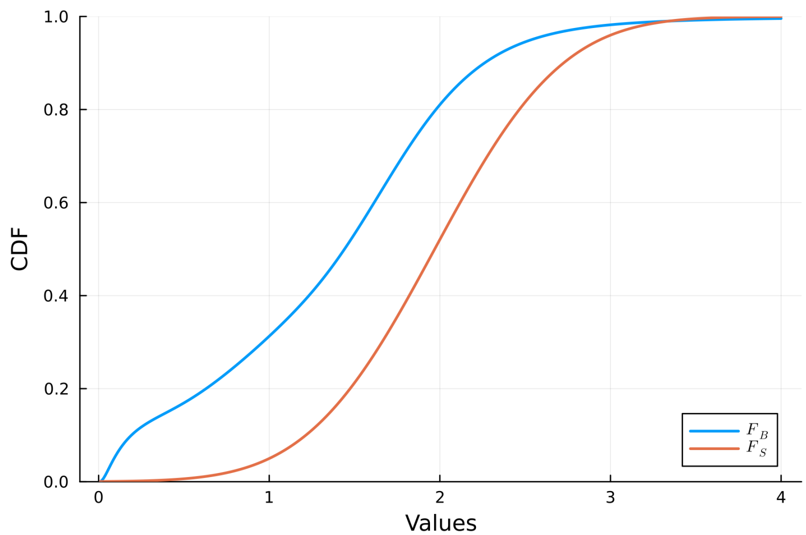
			λ_0				
	$\delta_{0,1}$	$\delta_{0,2}$	Rating	IHS(Feedback)	ln(Pred. Item Value)	c_B^T	$\mathbb{E}[c_S^E]$
Estimate	0.871	-0.245	0.055	0.146	0.409	0.054	0.131
Std. Err.	(0.059)	(0.007)	(0.026)	(0.010)	(0.022)	(0.004)	(0.004)

Notes: The estimated coefficients are obtained via debiased GMM; further details are presented in Appendix A.5. Estimates and standard errors for all but the entry costs are computed using the optimal weighting matrix. The entry cost parameters are computed using mean insertion fees for sellers and the average expected bidder surplus per auction; standard errors for both are adjusted for first-step estimation of the other demand side parameters.

The estimated bidder entry cost comes from the equilibrium bidder entry condition and the other estimated bidder arrival parameters. Evaluating the expected zero profit condition in equation (1.4) yields an estimated homogenized bidder time cost c_B^T of 0.056, since there are no bidder insertion fees. For the average (median) item in the dataset, this is approximately \$0.48 (\$0.33). This represents a moderate but not prohibitive time cost to entering each auction and inspecting the listing. The average homogenized seller insertion fee is approximately $c_S^I = 0.075$, though sellers' overall entry costs are assumed to be heterogeneous for different items.

As described in the previous section and as is common in the auction literature, the estimated demand parameters yield the optimal reserve pricing rule for sellers with perfect information. I use the sample of reserve prices chosen by experienced sellers (defined as those in the top 25% of sellers by experience at the start of the data) to impute the sellers' outside option for each item, under the assumption that experienced sellers have perfect information about the bidder arrival process. Figure 1.7 plots the estimated value distribution F_S along with the bidder value distribution F_B obtained in the demand-side estimation.

Figure 1.7: Estimated value distributions for auction participants



Notes: The bidder value distribution F_B is estimated via the maximum likelihood approach in 1.5.1 using a 5-component Gaussian mixture model for log values. The seller outside option distribution F_S is fit to imputed seller values among experienced sellers, as in 1.5.2, and uses a 2-component Gaussian mixture model for seller values.

The estimated seller value distribution in Figure 1.7 largely first-order stochastically dominates the estimated bidder value distribution. Since the population of sellers is the group of users who have previously acquired Beanie Babies, it is reasonable for them to have a higher value distribution for these items than any random bidder. However, this difference in value distributions is not unreasonably large: the seller value distribution largely falls between distributions of the maximum value of two bidders and that of three bidders (which are not plotted here), so it may be profitable in expectation for a seller with a large $\hat{v}_{0j}$ to list an item for sale.

I now estimate new sellers' beliefs about the bidder arrival process. I estimate the model on a sample of all sellers with at least 5 auctions in the data, to restrict attention to "serious" sellers. I also limit the sample to the first 5 auctions of all such sellers to focus on the early beliefs of new sellers. The prior mean for the reserve price coefficient $\delta_{0,2}$ is higher than the estimated parameter -0.245, implying that new sellers' beliefs about bidder arrival are upwardly biased, and particularly so for auctions with high reserve prices. The variance of beliefs about $\delta_{0,2}$ is moderately large, allowing sellers to update their beliefs about the effect of reserve prices on bidder entry.

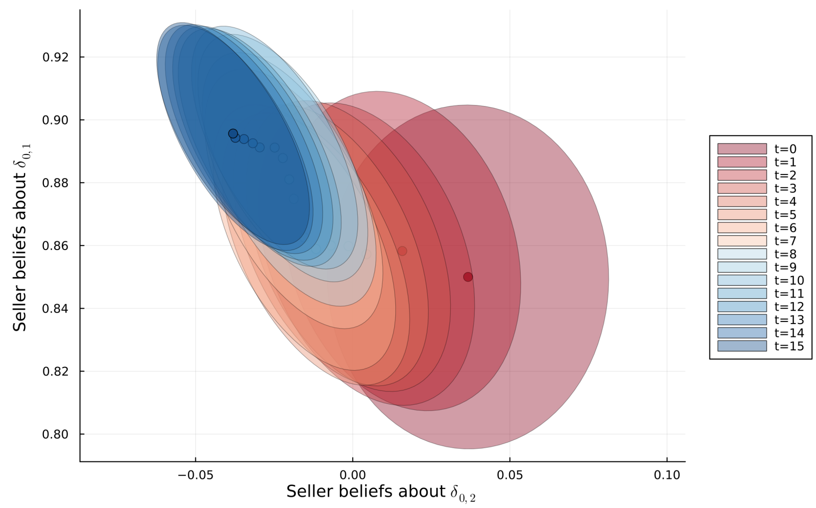
Table 1.2: Estimated new seller priors about the bidder arrival process

	$\mathbb{E}[\delta_{0,1} \mid b_0]$	$\mathbb{E}[\delta_{0,2} \mid b_0]$	StdDev($\delta_{0,1} \mid b_0$)	StdDev($\delta_{0,2} \mid b_0$)	Cor($\delta_{0,1}$, $\delta_{0,2} \mid b_0$)
Estimate	0.85	0.037	0.547	0.448	-0.006
Std. Error	(3e-5)	(2e-5)	(3e-5)	(3e-5)	(4e-5)

Notes: The model is estimated on the first 5 auctions (where applicable) of all 3,975 new sellers that list at least 5 auctions for sale. Standard errors are naive standard errors, treating seller value distribution parameters as known (*standard error correction for two-step estimation in progress*).

To help interpret the estimated prior parameters in Table 1.2, I evaluate the implied path of new sellers' average beliefs about the unknown parameter δ_0. Figure 1.8 plots ellipses corresponding to the estimated beliefs of these new sellers for up to their first 15 auctions; the contours represent 0.1-standard deviations around the mean. The prior mean shifts with successive auctions, moving toward a lower arrival coefficient $\delta_{0,2}$.

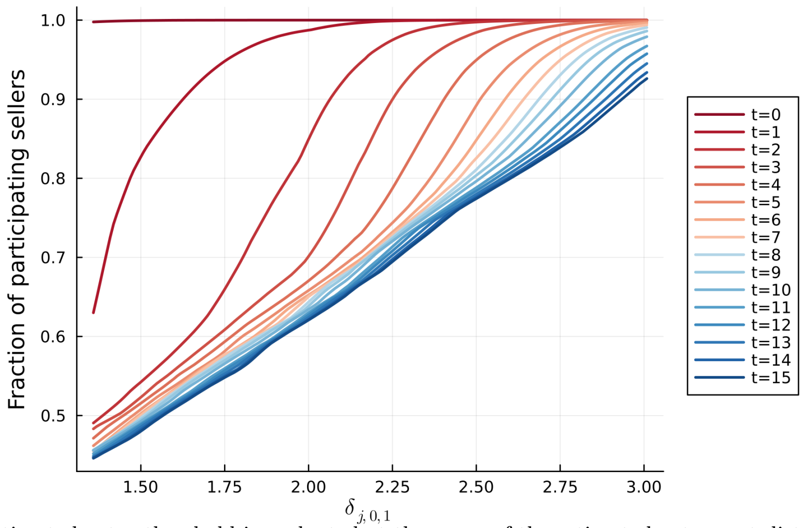
Figure 1.8: Estimated path of average new seller beliefs



Notes: The ellipses represent the average beliefs of new sellers in their first 15 auctions, conditioning on all 3,975 new sellers who list at least 5 auctions. The contours represent a 0.1-standard deviation around the mean of the belief distribution.

Finally, I directly show the effect of seller beliefs on selective entry by plotting the average seller entry threshold for new sellers relative to the experienced seller entry threshold. Figure 1.9 shows how the experienced sellers' entry threshold is increasing in the expected arrival rate, where the entry threshold is evaluated at the mean of the estimated entry cost distribution $F_{c_S^E}$. The entry threshold is significantly higher among new sellers in their first auctions. This is consistent with Table 1.2: in spite of the lower prior mean for $\delta_{0,1}$ (relative to the true parameter), its high variance combined with the higher prior mean for $\delta_{0,2}$ makes entry attractive to new sellers. This relative ordering of entry thresholds (most visible for items with a high baseline entry parameter $\delta_{j,0,1}$) is consistent with the selection pattern in Figure 1.4, where new sellers that exit early also set higher reserve prices in their first auctions.

Figure 1.9: Estimated average seller entry threshold conditional on baseline bidder arrival parameter $\delta_{j,0,1}$



Notes: The estimated entry threshold is evaluated at the mean of the estimated entry cost distribution, conditional on the log of the expected number of bidders for $r = 0$. I plot these entry thresholds between the 2.5% and 97.5% quantiles of the observed baseline entry parameters $\delta_{j,0,1}$.

1.6 Counterfactual platform design with information provision

I now study how information provision by the platform affects the optimal fee structure, platform profits, and both bidder and seller welfare. I simulate platform outcomes under alternative information structures and show that the platform can improve its own profits, as well as bidder entry and seller surplus, from providing information. I also estimate the difference between new sellers' expected return to information and the true return to information, which helps explain the prevalence of free information services provided by platforms.

1.6.1 Revisiting the platform's problem of fees and information provision

The estimated structural model characterizes the full platform problem, taking into account how bidders and sellers act in equilibrium. As in the expositional model in section 1.2, sellers choose whether to participate on the platform and what reserve prices to set as a function of their beliefs and the fees they face. I assume the platform faces both inexperienced sellers (with mass ω) that update their beliefs when they receive additional information, and experienced sellers (with mass $1-\omega$) that are not affected by information provision. The choice of fees affects sellers of both types, and when the platform cannot provide more information to inexperienced sellers, the optimal fee structure depends on the fraction of sellers who are new.

The auction platform chooses information provision and fees jointly to maximize profits. The platform sets seller-facing fees[12] c_S^I and c_S^R, where the former shifts the average entry cost to potential sellers.[13] The platform can provide information in the form of a sample auctions for each seller, comprising a dataset $\mathbf{D}_a$ that is drawn i.i.d. from the distribution $F_{\mathbf{D}}$ of auctions run by experienced sellers under the true data-generating process. All new sellers use this dataset to update their priors from b_0 to $\mathcal{T}(b_0, \mathbf{D}_a)$ before their first auction, and all sellers have the option to list T items.[14] Formally, the platform's profit maximization problem for each potentially listed item is

$$\begin{aligned}
\max_{a,\, c_S^I,\, c_S^R} \quad \int\int \Bigg(& \omega \cdot \sum_{t=1}^{T} \frac{1}{T}\, \mathbb{E}_{b_{t-1}} \Bigg[\int \underbrace{\mathbb{1}[v_0 \leq \bar{v}(\mathcal{T}(b_{t-1}, \mathbf{D}_a), \tilde{c}(z))]}_{\mathbb{P}[\text{Entry} \mid \text{inexperienced}]} \\
& \cdot \underbrace{\left[c_S^I + c_S^R \cdot \int R(r^*(v_0 \mid \delta, c_S^R) \mid \delta_0) \mathcal{T}(b_{t-1}(\delta), \mathbf{D}_a) d\delta \right]}_{\text{Platform revenue} \mid \text{entry, inexperienced}} \,\Bigg|\, b_0,\ \mathbf{D}_a \Bigg] dF_{\mathbf{D}}(\mathbf{D}_a) \\
& + (1-\omega) \cdot \underbrace{\mathbb{1}[v_0 \leq \bar{v}(\delta_0, \tilde{c}(z))]}_{\mathbb{P}[\text{Entry} \mid \text{experienced}]} \cdot \underbrace{\left[c_S^I + c_S^R \cdot R(r^*(v_0 \mid \delta_0, c_S^R) \mid \delta_0) \right]}_{\text{Platform revenue} \mid \text{entry, experienced}} \Bigg) dF_S(v_0) dF_{c_S^E}(z \mid c_S^I)
\end{aligned} \tag{1.11}$$

where b_t evolves from b_0 according to the updating rule $\mathcal{T}$ and the path of observed profit

12. I treat bidder-facing fees c_B^I and c_B^R as being fixed at zero. This is motivated both by the fee being equal to zero for bidders in the data, as well as analysis by Marra (2019) on a wine auction platform indicating that increasing revenue while maintaining transaction volume requires setting $c_B^R < 0$.

13. Recall that the distribution of entry costs is i.i.d. Exponential, with mean $\hat{c}_S^I + c_B^T$ where $\hat{c}_S^I$ is the average insertion fee in the dataset. I assume that choosing $c_S^I \neq \hat{c}_S^I$ affects only the mean of the shifted Exponential distribution $F_{c_S^E}(\cdot \mid c_S^I)$; all other moments are determined by the original fee structure.

14. This assumption simplifies the seller arrival process over time. In practice, sellers may differ in how many items they wish to sell, and new waves of sellers may arrive at different times. This particular setup highlights the tradeoffs between focusing on new sellers that learn over time and experienced sellers that are unaffected by information provision.

signals. As in the simple model, I assume there is zero marginal cost to providing data and hosting auctions on the platform. Though it is likely that there is some fixed cost in providing sellers with data (moving from $a = 0$ to $a > 0$), I assume this cost is sunk and not relevant for the platform's problem. I continue to use standardized average item values, abstracting away from heterogeneity in the dollar value of each potentially-listed item.

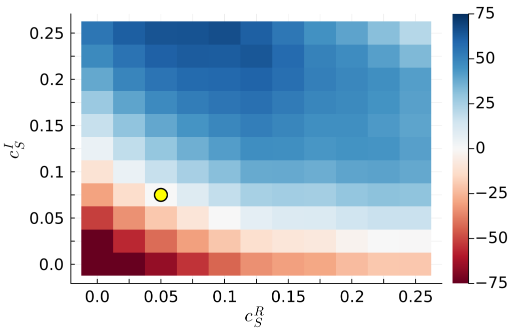
The maximization problem in (1.11) highlights the importance of sellers' initial beliefs in the platform's choice of fees and information provision. The first two lines correspond to the new sellers on the platform, each of whom observes a dataset $\mathbf{D}_a$ with a auctions before listing their first item. Each seller only chooses to list an item if their updated beliefs $\mathcal{T}(b_0, \mathbf{D}_a)$ about the bidder entry process imply it is optimal to do so; conditional on entry, they will also choose a reserve price according to their beliefs. Thus, knowledge of b_0 is critical for understanding how any additional information $\mathbf{D}_a$ will shift seller beliefs and therefore behavior. The third line corresponds to experienced sellers, whose selective entry and choice of reserve price are informed by their perfect information about the bidder entry process. While these sellers are not directly affected by the platform's information provision, they must pay fees c_S^I and c_S^R which apply to all sellers regardless of their beliefs and are chosen by the platform jointly with its information provision a.

1.6.2 Optimal fee structures without information provision

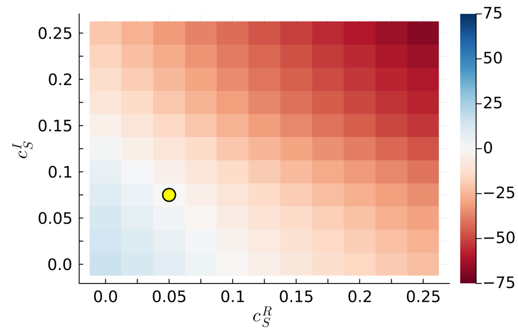
As a benchmark, I examine how platform outcomes change under alternative fee structures in the absence of information provision. The average seller fees on the platform when the data were collected are approximately $c_S^R = 0.05$ and $c_S^I = 0.075$. Since about half of all items in the data are listed by experienced sellers, I use $\omega = 0.5$ as the baseline fraction of new sellers, and I set $T = 15$ as the number of items that may be listed by each seller.

Changing insertion and revenue fees alone can significantly impact the platform's profits and bidder and seller outcomes. Figure 1.10 shows percent changes in platform revenue, the number of bidders, seller surplus, and the seller participation rate relative to the baseline fee structure. In general, increasing fees benefits the platform and reduces bidder and seller participation, as well as seller welfare. However, platform revenue in panel (a) is non-monotonic in the revenue fee: the reason for this can be seen in the severe entry effects for both bidders and sellers in panels (b) and (d). High fees reduce the probability that sellers list items on the platform and drive up the average reserve prices, which contributes to a lower number of bidders.

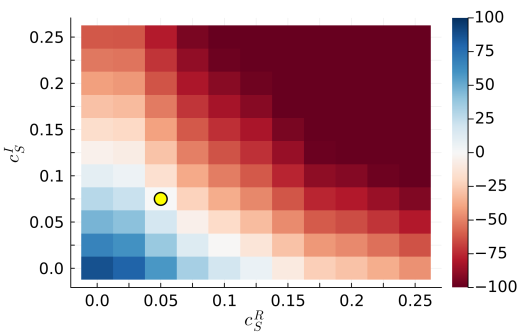
Figure 1.10: Percent changes in outcomes under alternative fee structures, $a = 0$ and $\omega = 0.5$



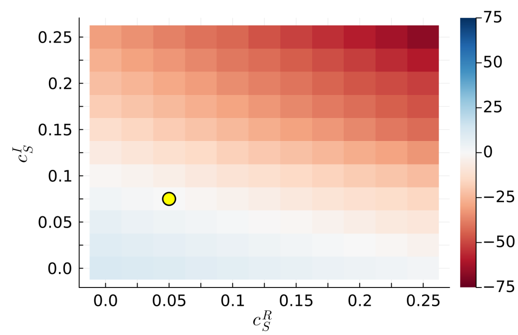
(a) Platform revenue



(b) Number of bidders



(c) Seller surplus

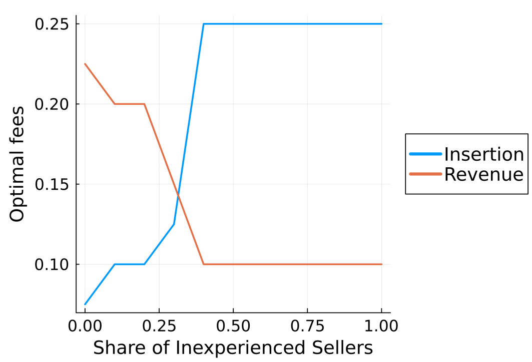


(d) Seller participation rate

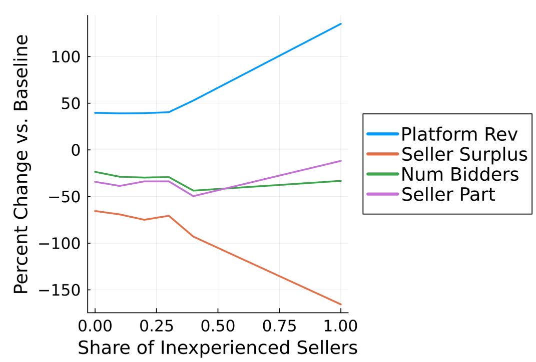
Notes: Each scenario was evaluated from a grid of possible fees (in increments of 0.025 from 0 to 0.25), simulating 250 sellers for each combination of parameter values. Blue (red) squares represent a percentage increase (decrease) in a given outcome relative to the outcome under the baseline values of c_S^R and c_S^I. The yellow dot indicates the baseline fee structure in the dataset.

The effect of fees on platform outcomes varies with the composition of sellers—whether inexperienced or experienced—on the platform. Figure 1.11 plots the optimal seller fees under varying fractions ω of new sellers on the platform, as well as changes in various platform outcomes (under the optimal fee structure) relative to the baseline fee structure at $\omega = 0.5$. I keep all other parameters the same as in Figure 1.10. The optimal fee structure changes as ω increases: when most sellers are experienced, it is optimal to set a low insertion fee and instead charge a higher revenue fee. However, when the platform faces less knowledgeable sellers it becomes optimal to set a higher insertion fee and a lower revenue fee. This is because new sellers are more optimistic and therefore willing to pay a higher insertion fee, though this leads to lower realized surplus. At the same time, lower revenue fees decrease the average reserve price and allow for more successful transactions. Note that as the fraction ω of new sellers increases, seller surplus goes down while the overall entry rate increases, since the greater portion of inexperienced sellers are behaving suboptimally.

Figure 1.11: Optimal fees and outcomes for varying ω with no information provision ($a = 0$)



(a) Optimal fees



(b) Platform outcomes

Notes: Each scenario was evaluated from a grid of possible fees (in increments of 0.025 from 0 to 0.25), simulating 250 sellers for each combination of parameter values. The platform outcomes of interest are platform revenue, seller surplus, number of bidders participating on the platform, and the seller participation rate. I simulate 250 sellers for each combination of parameter values.

Before proceeding, I briefly address possible reasons why the estimated optimal platform fees differs from the chosen platform fees in the data. First, many platforms pursue a growth-oriented strategy in their early years, during which they may forego short-term profits to develop a platform and gain additional users; for tractability I do not model these more complex dynamic incentives in addition to the detailed seller learning problem. I also treat the platform as a monopolist and ignore both existing competitors and the threat of competition (which limit the extent to which the platform may wish to increase fees) for the purposes of these counterfactuals. However, to address unmodeled constraints by the platform during this period, I will also consider the case where the platform can only implement Pareto-improving fees for any information structure.[15] Regardless of why eBay did not implement higher fees during the sample period, this policy has since changed: as of November 2024, eBay charges relatively low (or zero) insertion fees for most moderately priced items, with revenue fees of 13-15% across many item categories.[16]

1.6.3 Optimal fees with information provision

I now quantify the platform's joint problem of choosing fees and information provision. To do so, I simulate multiple auctions for inexperienced and experienced sellers under both alternative fee structures and alternative amounts a of information provided to new sellers. I assume this information (in the form of "sample auctions" provided from the platform to the sellers) is learned costlessly by all new sellers, and I maintain the share of new sellers at $\omega = 0.5$.

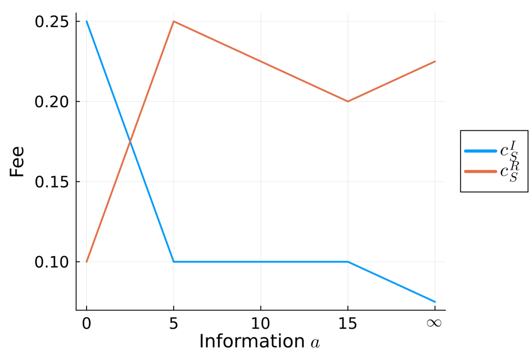
Figure 1.12 compares outcomes under different information regimes. I consider two cases: where the platform is choosing a fee and information structure that is optimal for its own profit maximization problem, and a second case where the platform constrains itself to search for only Pareto-improving choices. When the platform is unconstrained, it benefits most by *not* providing new sellers with any information and instead charging high insertion fees to naive sellers. Panel (a) of Figure 1.12 shows that the optimal insertion fee generally declines with the amount of information provided to new sellers, while the optimal revenue

15. Since bidders have a zero profit condition that determines entry, in expectation no bidder's welfare will be changed by these regimes. To compare bidder outcomes to those of sellers and the platform, I measure the average number of bidders that enter an auction on the platform. Any "Pareto" improvement in this exercise must maintain at least the same number of bidders as under the baseline fee structure—i.e., ensuring at least the same amount of possible bidder surplus is maintained—as well as weakly increasing both platform and seller surplus (with one strict improvement).

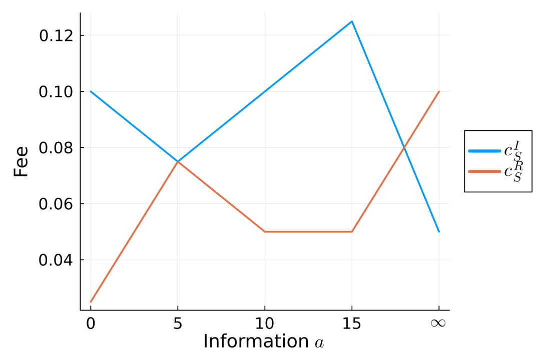
16. See eBay's website for details on its current fee structure.

fee increases relative to zero information provision. This leads to more platform revenue, though with lower seller and bidder entry onto the platform. In particular, when the platform is a monopolist that is not constrained in any way (e.g. by regulation or threat of entry), the platform may act in a way that significantly decreases sellers' expected surplus. The difference in optimal fees across information structures highlights how understanding sellers' beliefs is important for optimal platform design.

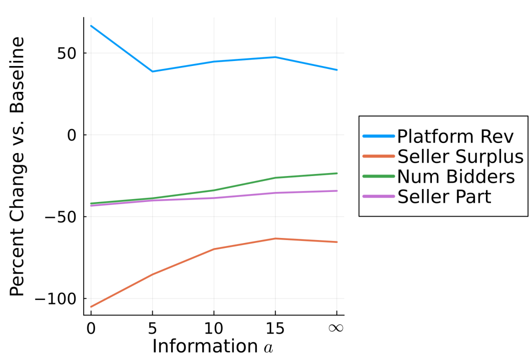
Figure 1.12: Optimal fees with information provision



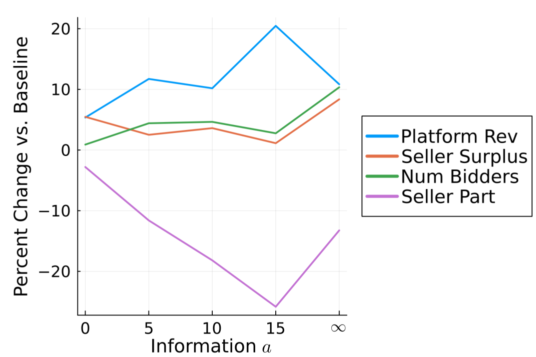
(a) Platform-optimal fees



(b) Pareto-optimal fees



(c) Outcomes under platform-optimal fees vs. baseline



(d) Outcomes under Pareto-optimal fees vs. baseline

Notes: Each scenario was evaluated from a grid of possible fees (in increments of 0.025 from 0 to 0.25) and information provision (for $a \in [0, 5, 10, 15, \infty]$, where ∞ represents perfect information provision i.e. the platform's maximum likelihood estimate being given to the sellers), simulating 250 sellers for each combination of parameter values. The platform outcomes of interest are platform revenue, seller surplus, number of bidders participating on the platform, and the seller participation rate. Panels (a) and (c) represent how optimal fees and platform outcomes change with the amount of information provision when the platform maximizes its own profits; panels (b) and (c) repeat this exercise with the constraint that the chosen fee structure must yield a Pareto improvement for the given level of information provision.

In contrast, when the platform is constrained to make only Pareto-improving changes, the platform wishes to provide some information to sellers to help guide their behavior. Panel (b) of Figure 1.12 indicates that platform profits are maximized by giving $a = 15$ observations to new sellers and optimizing fees accordingly. This leads to a positive change in bidder surplus that is competed away by an increase in the number of bidders. Seller participation is lower under the alternative information structure than in the baseline setting because new sellers with higher private values v_0 are more selective—and profitable—in their entry. The platform still withholds some information in this case, since doing so allows it to charge somewhat higher fees from naive sellers who over-participate relative to the full-information setting, while still benefitting from shifting new sellers' behavior.

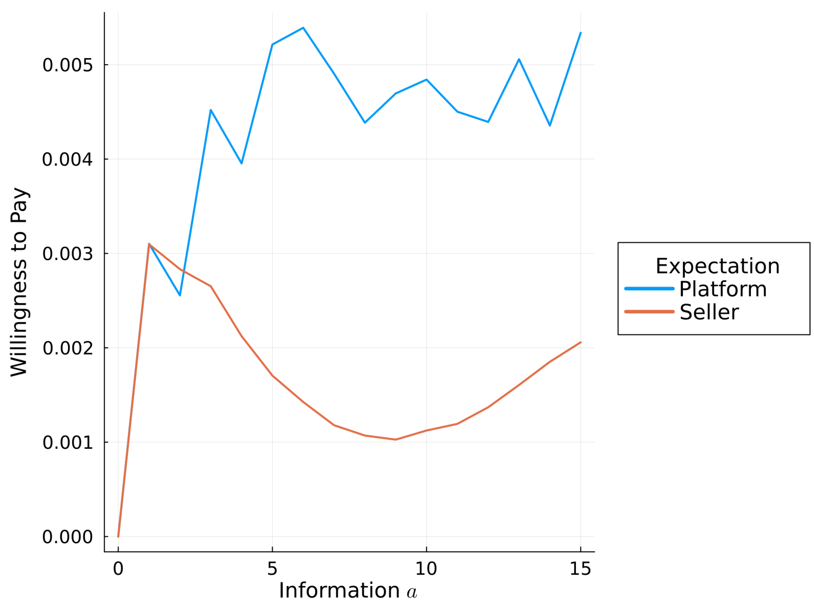
1.6.4 Seller willingness to pay for information

I now examine the difference between new sellers' *ex ante* expected value from receiving information from the platform and the true expected value from providing new sellers with information. As shown in the previous section, information provision can be used to improve platform profits by shaping new sellers' behavior. However, new sellers' biased beliefs mean they may not value their own learning as much as the platform. Throughout this exercise I assume seller fees are fixed at the baseline values of $c_S^R = 0.05$ and $c_S^I = 0.075$, and restrict attention only to new sellers ($\omega = 1$).

New sellers may not value information properly when their beliefs about the bidder arrival process are biased. A new seller's willingness to pay for information is the gap between expected surplus when receiving information from the platform and when learning unassisted. However, a new seller's expected evolution of their own beliefs depends on their current, biased beliefs about the true bidder arrival process. This means that they do not anticipate their beliefs or outcomes changing significantly from additional information. In contrast, the true expected returns from information are higher because new sellers' update their beliefs more significantly.

Figure 1.13 plots new sellers' willingness to pay for information under their initial, biased beliefs against the true expected gains from the provision of additional observations from the data-generating process. Note that the willingness to pay under sellers' expected path of future beliefs is lower than when conditioning on the platform's full information set.[17] Thus, new sellers' subjective expectation of the marginal value of information is lower than its true value. This difference helps explain why some platforms may offer seller tools for free rather than imposing additional fees. New sellers may undervalue these tools due to their biased beliefs, and the platform can benefit more from better-informed sellers' choices than from charging information fees to uninterested sellers, especially if there is some hassle cost for sellers in accessing and using this data.

Figure 1.13: Estimated seller willingness to pay for information under different beliefs



Notes: This figure plots the willingness to pay (as a fraction of mean item value) for various amounts of information a when sellers know they will list 15 items for auction. These are simulated differences between sellers' expected willingness to pay for different amounts of information under both new sellers' expectation of the value of information and the platform's expectation of the value of new information. The blue line is the expected gain from information to sellers when conditioning on the platform's knowledge of the data-generating process and new sellers' initial prior, while the orange line is sellers' expected gain from information when conditioning on new sellers' biased beliefs of how their own beliefs will evolve. I use 2,500 simulated sellers for each amount of information and for both types of expectations.

17. This holds with the exception of the first few units of information (observations) provided to sellers; the nonlinearity in expected surplus is due to taking expectations over profit which is itself nonlinear in the unknown parameters.

1.7 Conclusions

This paper studies the problem of information provision by an auction platform where sellers face uncertainty about the bidder arrival process. I first present evidence that new sellers learn to set optimal reserve prices as they gain more experience. I pair a model of seller learning with a model of two-sided endogenous entry onto an auction platform to investigate how new seller behavior is driven by both selection and learning. The model implies a new reserve price formula that is designed to both attract bidders to the auction and extract surplus from them; failure to account for the negative effect of high reserve prices on bidder entry leads new sellers to set higher-than-optimal reserve prices. I show that sellers' beliefs about the bidder arrival process can be semiparametrically identified from reserve price data under certain conditions, estimate new sellers' beliefs, and evaluate their implications for the platform design problem.

These results highlight platforms' ability to influence user behavior outside of its well-known ability to charge different fees to different sides of the market. Information provision improves the quality of the marketplace for new sellers, who are able to better optimize their entry decisions and pricing strategy. Sellers' improved profits increase the transaction volume on the platform, which increases platform profits while also inducing additional bidder entry through lower average prices on the platform. Thus, Pareto improvements can be achieved through more widespread access to important economic knowledge. However, this may not always align with platforms' incentives: withholding information may allow the platform to extract revenue from uninformed sellers (or possibly, in other settings, uninformed buyers). In particular, a large platform that faces little competition may not be incentivized to use additional information to improve seller or bidder outcomes.

More broadly, this paper speaks to the importance of information and its availability in the modern economy. Platforms are increasingly common in areas such as consumer products (eBay, Amazon), social media and advertising (Facebook/Meta, Twitter/X), financial services (NYSE, cryptocurrency exchanges), transportation (Uber, Lyft), and online betting (DraftKings, FanDuel). Users of each platform may lack crucial knowledge that affects their behavior on that platform, making them susceptible to the influence of information or misinformation in developing their beliefs. As the set of tools that can be used to shape information expands, helped in part by recent advances in generative AI, platforms face a more complex problem of how to shape their users' beliefs. Whether the platform benefits from doing so, and whether this comes at the expense of any of the platform's own users, is a central question in both the design and potential regulation of platform markets.

CHAPTER 2

FIX THE PRICE OR PRICE THE FIX? RESOLVING THE SEQUENCING PUZZLE IN CORPORATE CONTRACTING

In theory, there is no difference between theory and practice...

...while in practice, there is.

2.1 Introduction

The aphorism above has a notoriously contested provenance,[1] but its insights are hard to gainsay. One need not venture far into myriad academic fields of inquiry to find a landscape riddled with famous (if not spectacular) collisions between theory and practice. Examples abound in areas as diverse as education, labor, health, engineering, software, environmental policy, and medicine. In the light of these collisions, the above maxim represents a cautionary tale about the dangers of siloed thinking, admonishing theorists to remain mindful of institutional details while chiding practitioners to continually scrutinize ossified practices that might have outlived their original purposes.

A notably stark disjunction between theory and practice has long afflicted high-stakes corporate transactions, such as acquisition and financing agreements. These contracts are among the longest, most contemplated, and most heavily negotiated in modern markets, and their sophisticated designers pour considerable energy into calibrating their architecture. On the theory side, few subfields of economics are as well developed as optimal contract design.

This chapter is coauthored with Matthew Jennejohn, Cree Jones, and Eric Talley. We thank Jens Frankenreiter, Brennan Platt, attendees of the 2024 Winter Deals conference, and seminar participants at the University of Chicago for helpful comments and suggestions. All errors are ours.

1. This excerpt is widely (and erroneously) attributed by turns to Yogi Berra, Albert Einstein, Richard Feynman, Nassim Taleb, and several others. The earliest use of this phrase we can find appears substantially older. *See* Brewster (1862).

Considerable ink has been spilled on the topic during the last eight decades, garnering multiple Nobel prizes[2] and suffusing pedagogy across economics departments, business schools and law schools alike. Given the appreciable economic stakes and the sophisticated players involved, large corporate transactions would seemingly be an ideal proving ground for contract design theory. Indeed, if there is *any* area where contract negotiators might behave like the rational actors inhabiting canonical economic models, this is surely it.

And yet, the norms and protocols of practitioners who structure large corporate transactions have long diverged materially from fundamental tenets of contract design theory: in short, both sides regard the other as operating "backwards" in assembling the price and non-price terms of a deal. According to conventional contract theory, price is a perfectly adjustable zero-sum mechanism—the consummate numeraire that fluidly transfers payoffs between parties in a welfare-neutral manner. Non-price terms, in contrast—such as covenants, conditions, warranties, and the like—are rarely zero-sum, and their contractual allocation has real welfare consequences. Accordingly, efficiency calculus counsels that non-price provisions should be structured first (prior to setting price), with a goal of maximizing expected joint surplus. Only after those terms are fixed should pricing enter *at the very end* of the process. Such a sequence makes eminent sense (at least in theory), since the transfer-payment aspect of price makes it an ideal tool for truing up any payoff imbalances left behind after aggregating the joint welfare maximizing non-price terms, "greasing the wheels" of a mutually beneficial optimal contract. This sequential prediction is so fundamental and well-supported, in fact, that it permeates virtually all of contract design theory (*See* Bolton and Dewatripont 2004).

Nevertheless, and in stark contrast with contract theory, transactional practice typically proceeds in the reverse direction, fixing core price terms at the onset, often via a succinct

2. In the last thirty-five years alone, Nobel laureates specializing in contract theory include Milgrom and Wilson (2020), Hart and Holmstrom (2016), Tirole (2014), Roth and Shapley (2012), Hurwicz, Maskin and Meyerson (2007), Mirrlees and Vickrey (1996), Harsanyi, Nash and Selten (1994), and Coase (1991).

term sheet produced by executives and insiders. Only after pricing is locked in does a coterie of outside lawyers and other transactional specialists sweep in to hammer out the non-price details. While these late-moving actors have significant negotiating latitude, one thing they are almost *never* permitted to do is to revisit pricing: Although price re-cuts sometimes happen, they are heavily discouraged by a variety of institutional factors and are therefore extremely rare.[3] Put simply, the practice of fixing price at the onset of bargaining means that transactional professionals are left with the task of assembling the remaining non-price components with nary a drop of the transactional grease that price adjustments can (theoretically) afford.[4]

The persistent deviation of contract theory (fix price last) from transactional practice (fix price first) has long puzzled scholars and practitioners, sparking debate and inquiry into why price—which is otherwise an ideal payoff re-leveling mechanism—remains an inflexible anchor on negotiations, especially in the realm of large corporate transactions where the monetary stakes are appreciable. Compounding this quandary further, perhaps, is a different norm in smaller transactions (such as used car sales or residential real estate) where—consistent with theory—pricing decisions typically remain more fluid throughout the negotiation process. Why would such contracts tend to conform to theoretical predictions while large corporate transactions (with billions of dollars at stake) diverge?

This paper seeks to reconcile theory and practice by amalgamating a contract negotiation framework with a search model. The key to our approach is to analogize contract design to a production process (S. J. Choi et al., 2021, 2022), whereby the choice set of non-price terms

3. Price renegotiation in the LVMH-Tiffany transaction is a highly-publicized exception to the rule, as described in Jennejohn et al. (2022). This broader phenomenon of reference dependence in shaping economic outcomes is discussed in O'Donoghue and Sprenger (2018); other work on the role of similar benchmarks or expectations in two-stage bargaining settings includes Basak and Khan (2024), Crawford (1982), and Muthoo (1992).

4. Note that certain less frequent M&A sales processes, namely multi-bidder auctions, are more amenable to non-price terms being set first. Full-blown auctions, however, are quite rare, and some auction processes even allow bidders some degrees of freedom to adjust nonprice terms. We discuss auction structures at greater length below.

available to the parties is endogenously revealed through the parties' efforts, expended at a private and non-contractible cost. Within our framework, fixing the deal price at the onset can emerge as an efficient design choice, both from an incentive compatibility perspective and for joint surplus maximization. Specifically, we show that fixing price first can better incentivize parties toward efficient search for and production of welfare-enhancing non-price terms, a sort of two-sided "hostage-taking" in the spirit of Williamson (1983). Moreover, to the extent that uncovering creative non-price provisions translates into greater value in high-stakes environments (Gabaix and Landier (2008)), our model predicts that "price-first" bargaining will tend to be more prevalent in large corporate transactions than in smaller-stakes deals.

The intuition behind our argument unfolds in four key steps. First, we posit (realistically) that contracting parties are heterogeneous, and consequently the optimal contract terms for a randomly chosen set of counterparties will differ from any other randomly chosen dyad. Second, we assume (again realistically) that bargaining power is an exogenous primitive, which itself cannot be bargained over. That is, if a negotiating party possesses the lion's share of the bargaining power, she cannot commit to *not* exploit that power at a later stage. Third, we argue that the universe of possible non-price terms (beyond standard-form "boilerplate" templates) is not obvious ex ante; rather, finding a bespoke non-price term that enhances payoffs requires discretionary and costly search process, undertaken by at least one (and possibly both) of the parties.

Fourth, and critically, we posit that the most skilled searcher for non-price terms need not also be the best negotiator. Thus, when search ability and bargaining power are not aligned, fixing price last (as conventional theory counsels) can disincentivize efficient search. The reason is simple: because the costs of searching for welfare-enhancing non-price terms will become sunk once such terms are unveiled, those efforts quickly become irrelevant in a last stage where the parties bargain over price; instead, the bargaining outcome from that point forward hinges centrally on each party's relative bargaining power. As such, a party contemplating searching for innovative non-price terms faces the prospect that even if she succeeds, her counterparty will marshal superior bargaining power to extract the newly-created value through pricing concessions, leaving the successful searcher with little more than nonpecuniary bragging rights (and a sunk cost). The searcher's anticipatory concern over expropriation becomes especially acute, moreover, as the non-searching party's relative bargaining power increases.

When, in contrast, price is fixed from the onset, the parties' search incentives change fundamentally, typically in the direction of *more* efficient search. When (for example) the most efficient searcher has little relative bargaining power, the rigidity of an immovable price metamorphoses from a bug into a feature,[5] incentivizing her to work harder to find a value-enhancing non-price term with less fear that her counterpart will later expropriate the added value by wheedling for price concessions (which are now no longer permitted). Moreover, even when the searching party also possesses significant relative bargaining power, her incentives remain roughly unchanged regardless of whether price is set first or last: in either case, she will be able to capture most of the value she brings to the table from a successful search. Aggregating across cases, our theoretical framework predicts that in a "large" set of parametric environments, fixing price *ab initio* can catalyze more efficient production of non-price terms overall, resulting in more advantageous expected outcomes for both parties.

More generally, our analysis illustrates a counter-intuitive possibility for creating value by transforming a seemingly frictionless bargaining problem with modest negotiation costs into a more rigid bartering problem, where the only available currency for negotiated exchange is by "horse trading" the newly identified non-price terms. While locking in price upfront no doubt introduces certain transactional frictions, it can simultaneously promote efficient

5. *Compare* Kafka (1915).

incentives, resulting actuarially in new, welfare-enhancing transactions and terms that would have been unlikely or impossible were pricing determied at the end.

In addition to developing a theoretical model capable of reconciling the longstanding disjunction between theory and practice, we also make three contributions to the contract theory literature. First, we develop a flexible contract design framework in two-dimensional payoff space—representing the buyer and seller—centered at their anticipated respective payoffs ex ante under the standard form contract. Working in polar coordinate space, our framework reduces the search strategy of each party to a decision over two variables: (1) the direction of search (measured by an angle in payoff space) and (2) the intensity of search (measured by the length of the ray associated with search in the chosen direction). By reducing the optimization problem of both buyer and seller to these two foundational dimensions, our framework delivers a tractable and powerful baseline that we believe can be deployed and extended in other contexts like trade negotiations or exclusive contracting in labor or real estate. These settings differ from smaller transactions in which the typical price-last formula is followed, highlighting the role of actively creating value-enhancing contract terms in these larger transactions.

Second, we embed a discrete "bartering" model (over non-price terms in stage 2) nested within a Nash bargaining model (over pricing, in stage 1). This unlikely marriage mimics the reality of real world, high-stakes transactions, but at the theoretical expense of introducing discontinuities that undermine the tractability of the model. Nevertheless, we are able to find closed-form characterizations of equilibrium behavior under certain simplifying assumptions and simulate numerical solutions across a full range of associated parameter values. We do so by using a choice probability approach taken from the discrete choice literature (see e.g. S. Berry et al. 1995; Eaton and Kortum 2002; McFadden 1972), under the assumption that deal-specific heterogeneity makes the proposed contract terms vary in their suitability across negotiating dyads. This framing has the added bonus of making the analysis of more complex contracts (with myriad specific terms which are themselves the byproducts of negotiations) empirically tractable using standard tools from industrial organization. While other studies of bargaining leverage detailed data on alternating offers (Backus et al., 2020; Dunn et al., 2024), our approach is particularly suited to settings with simultaneous and/or unknown procedures for contract development.

Third, we develop a standard sister Nash bargaining model with search over non-price terms in stage 1 and subsequent price setting in stage 2. This model formalizes the more conventional intuition for setting price last but highlights its dangers in the setting where efficient contract terms are not obvious *ex ante* and must be discovered. That is, the timing of this contracting game means that non-contractible efforts to develop contract terms before the price is fixed are sunk, leading to a two-sided analogue to the holdup problem in Klein et al. 1978. Moreover, the similarity between the two models enables us to compare them across different parametric settings. This comparison yields both empirical predictions and an explanation that reconciles the disjunction between practice and the currently prevailing theoretical models: setting the transaction price first–in many settings–incentivizes more efficient search for (and production of) contract terms.

The account we offer here does more than resolve a longstanding conundrum using an original and tractable model, however. It also sheds light on a variety of other norms that are commonly observed in large-stakes deal negotiations as well as important legal doctrines. For example, a direct implication of our framework is the possibility of equilibrium deal failure. In our model, bargaining parties may rationally sign up a preliminary deal featuring standard "boilerplate" terms that—at least when signed—makes them jointly worse off than the *status quo ante* without a transaction. Why would they do so? Because in equilibrium they expect that the ensuing search for non-price terms may yield new payoff-enhancing structures, thereby making the risk of subsequent deal failure worth the gamble.

In a similar vein, our approach reveals a plausible rationale behind enforcing even preliminary agreements that do not have all their key terms locked in.[6] This is an area where courts have grown increasingly willing to deploy enforcement tools (such as reliance or expectation damages) against a party who fails to deploy "good faith" efforts to finalize the terms of a preliminary agreement.[7] Fixing price ex ante in our framework is important precisely in situations where it is important to incentivize parties to expend good-faith efforts to find value enhancing terms. A party's failure and/or refusal to do so can be particularly harmful in our setting, since it can increase the odds of wasteful deal failure. Consequently, courts' enhanced willingness to enforce preliminary agreements with open terms can be interpreted as consistent with catalyzing efficient search incentives within our model.

Our account also yields predictions about the nature of the search for non-price terms. Significantly, our model gives the parties considerable discretion about what "direction" (in payoff space) to conduct their search. We show that in a large (and plausible) family of equilibria, the parties will tend to avoid engaging in purely "selfish" search—*i.e.,* developing bespoke terms that benefit themselves but impose costs on their counterparty. Selfish search is risky in our model, since any "bartering" of non-price terms (with no prospect of a price adjustment) must result in a (weak) Pareto improvement over the baseline boilerplate agreement to survive. A selfish search is destined to fail this test when analyzed in isolation; its prospective utility, then, is critically dependent on being combined with another discovered term so that their aggregation results in Pareto improvement over the boilerplate. While possible, such combinations are not reliably rendered in equilibria of our model, and in any event they require significant coordination to produce. In contrast, it tends to be more lucrative for each party to search for non-price terms in a (weakly) unselfish manner, so as to

6. In U.S. contract law, preliminary agreements with open terms are referred to as "Type II" agreements, differentiating them from "Type I" preliminary agreements, where all essential terms are settled and only certain formalities are lacking. *See* Teachers' Insurance v. Tribune Co., 670 F. Supp. 491 (SDNY 1987).

7. *Compare* Empro v. Ballco, 870 F.2d 423 (7th Cir. 1989) *with* SIGA v. PharmaThene, 67 A.3d 330 (Del. 2013), Pennzoil v. Texaco, 481 U.S. 1 (1987), Copeland v. Baskin Robbins, 96 Cal.App.4th 1251 (Cal. Ct. App. 2002).

ensure any term they uncover represents an acceptable improvement over the default.

Finally, our model helps reveal the critical importance that good lawyering can play in transaction design. Highly skilled lawyers in our model face lower search costs, plausibly reflecting a combination of greater creativity and more robust firm-level experience (their own and their partners). Consequently, good lawyers are also more skilled at smoking out value-enhancing non-price terms, which in turn expands the frontier of payoff possibilities that are available in equilibrium. In fact, when two high quality firms interact with one another, they may be in a better position to coordinate their searches, producing non-price terms that—while not individually Pareto improving—become strongly welfare enhancing when combined as part of a bartered *quid pro quo*.

Although we are not the first to observe the odd disjunction between the theoretical account of optimal contracting (where price is chosen last) and the practical reality (where price is fixed first), our framework is novel in several respects. For instance, our model provides microfoundations for observed phenomena, such as A. B. Badawi and de Fontenay (2019)'s of the "first mover" advantage in the design of non-price M&A terms. It also advances earlier efforts that explored the roles of bargaining power and asymmetric information on the design of non-price terms in M&A contracts (A. Choi & Triantis, 2012) by introducing the contractual innovation process—the non-trivial search for new terms in the context of a particular design problem (Jennejohn et al., 2022)—into its core model. Our model also uses existing tools from the literature on empirical choice models to present an empirically tractable model that can be used to analyze the value of chosen contract terms, even when there is no post-negotiation variation in price. This empirical tractability extends to other settings where firms or other agents may jointly choose among discrete options.

Our analysis unfolds as follows. In Part 2, we provide relevant context for the development of contracts in corporate transactions. In Part 3, we present the overview of our baseline model, with additional details deferred to Appendix 1. In Part 4, we describe the model's solution and present various comparative statics, as well as several extensions to and modifications of the baseline model. Part 5 discusses various implications of our analysis for both contract theorists and practitioners. Part 6 concludes.

2.2 Contract production in the M&A market

The modern M&A agreement is a complex piece of transactional technology, typically encompassing over 100 pages of obligations (Coates, 2016a; Hwang & Jennejohn, 2018; Jennejohn, 2018).[8] While many markets cope with similar levels of contractual complexity by standardizing terms across deals (Gulati & Scott, 2012a), M&A agreements are surprisingly resistant to rote use of boilerplate, and a significant amount of transaction-specific tailoring of terms often occurs in each negotiation (Coates, 2016a; Jennejohn, 2020; Talley, 2009). In short, there is space for creativity for the transaction designer, and, indeed, reputational benefits accrue to advisors who successfully innovate effective new terms.

The terms of these complex contracts can be sorted into several key categories. First, the operative terms of the agreement set forth the details of how the business combination will be accomplished, including the price for the acquisition and the nature of the consideration used (cash, the acquirer's stock, or a combination of the two). Second, the seller provides a series of representations and warranties relating to, most notably, the qualities of the target company, thereby addressing potential risks that are unobservable during the acquirer's due diligence process.[9] Third, a series of covenants, which apply to the behavior of either (or both) acquirer or seller between the time the contract is executed and the time the transaction closes,[10] address pre-closing risks, such as: interim operating covenants that

8. Often, ancillary agreements are also attached to the main agreement (Hwang, 2016). Our focus here is on the main M&A agreement that accomplishes the core transaction.

9. The buyer also typically provides a series of representations and warranties focused primarily upon its ability to execute the transaction, but these are usually less negotiated, especially in cash deals.

10. For most large transactions, a period of time between signing and closing is necessary in order to allow, for instance, for regulatory reviews or stockholder approvals.

require the seller to operate the target in the ordinary course of business, thereby precluding extraordinary decisions that would impair the value of the target company; regulatory provisions that address the possibility of, for instance, an antitrust or national security regulator attempting to prevent or force the restructuring of the transaction; and deal protection devices, like “no-shop" provisions, that constrain the seller's ability to pursue alternative bids. Fourth, and finally, conditions to closing and termination provisions connect breaches of the aforementioned terms to the parties' duty to close the transaction, thereby incentivizing performance.

To make that complexity manageable, the advisers to a transaction—the investment bankers and deal lawyers advising both buyer and seller—tend to bifurcate the negotiation process into two steps. First, the key operative—or “business"—terms, including the price and a smattering of important terms across the four categories above, are determined and reduced to a preliminary agreement, such as a term sheet or letter of intent. The principals of both buyer and seller are heavily involved at this stage since, as one hoary treatise in the field notes, “the usual topics of discussion at the outset are generally basic business areas, on which attorneys should defer to their clients" (Freund, 1975). After the core business terms are preliminarily agreed upon, the detailed “legal" terms of the agreement are then hammered out. Here, the division of labor shifts, with the deal lawyers taking the wheel.[11]

As a practical matter, the price and other key terms set in the first step of the negotiation process are typically quite sticky. Detailed accounts of such stickiness are not generally available in the public record, since secrecy in merger negotiations is jealously kept. Nevertheless, they arise from time to time, especially for imperiled deals that devolve into litigation. For

11. A notable exception to this ordering can be found in deals where the target company conducts an auction rather than negotiation with an exclusive bidder. Here, it is more conventional for the non-price terms to be fixed up-front, providing a “package” against which prospective bidders formulate their competing offers. Because auction deals mechanically give significant bargaining power to the seller, our model predicts that they will be most attractive when the seller faces a low cost in searching for non-price terms, *and* when such terms are valued relatively homogenously by prospective bidders. We return to this point in Section 5 below.

instance, the M&A deal at the heart of *Frontier Oil v. Holly* provides a glimpse of how resistant the initially-set price term can be to change.[12] That transaction involved, among other assets, the acquisition of an oil rig that had been sitting on the grounds of Beverley Hills High School for decades. When the negotiations were far advanced, news broke that environmental activist Erin Brockovich planned to bring a mass toxic tort suit against one of the target Frontier's company's subsidiaries, which operated the rig that allegedly harmed students. That potential liability risk had not been disclosed during due diligence, despite similar cases resulting in settlements worth hundreds of millions of dollars. Instead of adjusting the purchase price in light of the revelation though, the parties reworked a series of detailed terms in the agreement, resolutely remaining in the second—non-price—stage of the negotiating process rather than going back to square one and recutting the deal.

In the next section, we introduce a model that explains why this curious approach to contract design is pursued, shedding light on market practice and, in turn, informing the legal system's approach to enforcing contractual obligations as they emerge in this negotiating process.

2.3 A model of two-stage contracting

In the light of general industry practices that sequentially stage price and non-price negotiations, along with the importance of deal term innovation in complex agreements, in this section we develop a novel, tractable framework that incorporates both industry practice and term innovation. We begin in Subsection 2.3.1 by describing the transaction in the model (the sale of a "business asset" from a seller to a buyer[13]) that involves three steps, the order of which will change depending on the model: (1) the parties set the price, (2) the parties

12. See Frontier Oil Corp. v. Holly Corp., 2005 WL 1039027 (Del. Ch. April 29, 2005).

13. The precise legal mechanism for the transaction is not critical to our inquiry, and it thus could be a stock sale, an asset sale, a statutory merger, a negotiated tender offer, or any other bargained-for means for transferring ownership of the business asset.

search for new terms, and (3) the parties select the other terms of the contract based on the return of each party's search. In Subsection 2.3.2 we present the model in which price is set first, the search for other terms is second, and bartering over other terms occurs third. In Subsection 2.3.3 we present the model in which search for non-price terms occurs first, bartering over other terms occurs second, and price is set last. In Section 2.4 we compare closed form solutions across the two models, focusing first on three restricted cases to fix ideas, and then exploring numerical solutions to the general case.

2.3.1 Term innovation and alternate sequencing

Consider a potential transfer of a business asset from a representative seller s to a representative buyer b. Each respective party places a "baseline" valuation of π_i on the asset (where $i \in \{b, s\}$), and we assume these valuations to be common knowledge amongst the parties. The buyer's bargaining power is represented by an exogenous parameter $\tau \in (0, 1)$, and thus the seller enjoys complementary bargaining power $1 - \tau$.

To transfer the asset, the parties must enter a contract consisting of a price p paid from b to s, as well as a vector of non-price terms m. The non-price terms collectively give rise to an additional expected value $v_i(m)$ to each party i, independent of (and in addition to) the parties' baseline valuations π_i. Because our key results hinge on strategic dynamics within payoff space, our analysis need not characterize the full vector space of all possible non-price terms; we instead characterize any non-price term vector m by the expected payoffs it conveys to the parties, $v(m) \equiv (v_b(m), v_s(m)) \in \mathbb{R}^2$. With one exception, the non-price terms are assumed hidden from the parties, and discovering them requires costly search (described below). The sole exception is a "default" (or "standard form" or "boilerplate") set of non-price terms m_0, which are commonly known. We normalize the coordinates of m_0 in payoff space to be at the origin, so that the expected additional payoffs delivered by the default are normalized at $v_i(m_0) = 0$ for $i \in \{b, s\}$.

For exposition purposes, it will frequently prove convenient to characterize non-price terms in payoff space using polar coordinates, with radius $r \in \mathbb{R}_+$ and angle $\theta \in [0, 2\pi]$. To further economize on notation, it will also be convenient to transform θ into $\theta(a) \equiv \pi(a + 0.25)$, where $a \in [-1, 1]$.[14] Thus, the contract terms create expected valuations of:

$$v_b(m) = r\cos(\theta(a))$$
$$v_s(m) = r\sin(\theta(a))$$

The final contract terms are chosen from the subset $\mathcal{M}^*$ of terms that are known to both firms at the time of bargaining, which include by default the standard-form terms m_0, any other non-price terms discovered by the parties, and the combination of such terms.

Prior to negotiation, each party i can search for one new term m_i. The parties' search decisions are made simultaneously, and search efforts are assumed (at least for now) to be non-contractible. When each player uncovers a new term m_i, that new price vector's coordinates in payoff space are added to the choice set of possible non-price terms. Since both parties search, the choice set minimally expands further to $\mathcal{M}^* = \{m_0, m_b, m_s\}$. We also assume that the new terms (if any) discovered by each party can be combined additively, so that the choice set expands further to $\mathcal{M}^* = \{m_0, m_b, m_s, (m_b + m_s)\}$; we write the latter term as m_{bs}.[15] Thus, each set of contract terms in $\mathcal{M}^*$ is indexed by the subscript j where

14. Under this normalization, a search along direction $a = 0$ corresponds to a thoroughly "selfless" search, where the parties benefit symmetrically from the discovered term. In contract, the cases of $a = 1$ and $a = -1$ correspond to "selfish" zero-sum search, where the non-price term enhances value for one side in the same amount as it reduces value for the other.

15. To fix ideas, we assume that the technology for combining terms is linear in the expected payoffs, i.e. $v_i(m_{bs}) = v_i(m_b) + v_i(m_s)$. Consequently, the new contract bs is characterized by

$$r_{bs} = \sqrt{r_b^2 + r_s^2 + 2r_b r_s \cos(\theta(a_s) - \theta(a_b))}$$
$$a_{bs} = \frac{1}{\pi}\left[\theta(a_b) + \tan^{-1}\left(\frac{r_s \sin(\theta(a_s) - \theta(a_b))}{r_b + r_s\cos(\theta(a_s) - \theta(a_b))}\right)\right] - 0.25$$

$j \in \{0, b, s, bs\}$.

Each player faces a cost $c_i(r_i, a_i)$ to search for non-price term innovations. The cost is assumed to be continuously differentiable, increasing and convex in search intensity r_i, but with $\frac{\partial c_i(0,a_i)}{\partial r_i} = 0$ so search is costless on the margin near the default contract. Costs are assumed to be weakly decreasing in $|a_i|$, and thus it is costlier to search in directions that are joint-surplus improving.

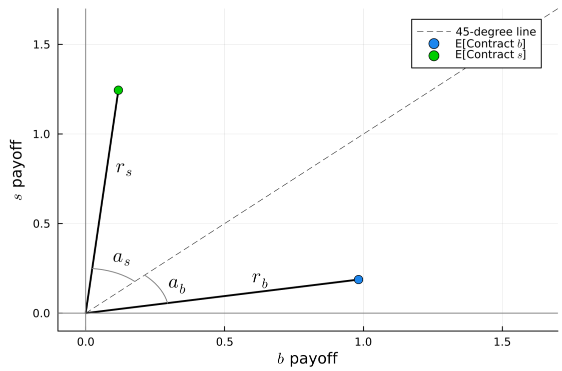
Because search does not always align with success, our framework also allows for *ex ante* uncertainty in firms' ultimate success in discovering a new term given their search intensity. This assumption mimics real-world variation in the challenge of finding new terms, since both search productivity and term values may vary across deals in difficult to observe ways. (This uncertainty also may apply to attempts to combine discovered terms, as the firms' success in combining potentially-conflicting terms may also differ across deals.) To capture this uncertainty, we denote the realized search radius for term m_j as $r_j \cdot \epsilon_j$ with $\mathbb{E}[r_j \cdot \epsilon_j | r_j] = r_j$. We call ϵ_j a term-specific productivity shock that is only observed after investment decisions $\{r_i^*, a_i^*\}$ are made. This implies that realized payoffs, denoted $v_i(m_j; \epsilon_j)$, are in expectation equal to the average payoffs $v_i(m_j)$ defined above.[16]

Before proceeding, we provide some graphical intuition for the set of possible contracts in the payoff space. In Figure 2.1, each component is added sequentially to build a representation of the possible choices of contract terms. Panel (a) first shows two hypothetical contracts in payoff space, using the respective radii and search angles (recentered around the 45-degree line) to characterize each. Panel (b) then includes the compound expected payoff of the combined term, and panel (c) further illustrates the uncertainty in term production by representing the payoff as a point along a ray, with the associated density of ϵ plotted (in

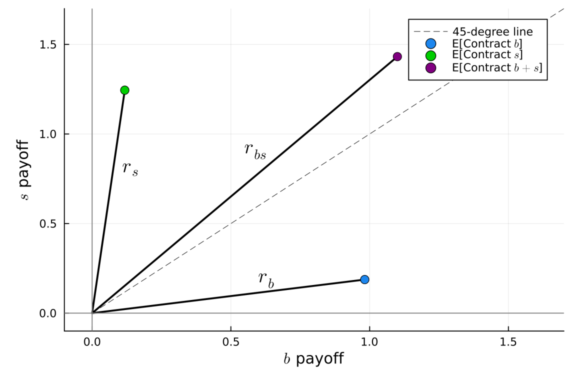
16. Note that while expected payoffs are linear in the component payoffs (i.e., $v_i(m_{bs}) = v_i(m_b) + v_i(m_s)$), the same does not hold for realized payoffs. We interpret ϵ as a shock to the firms' ability to implement the term in an actual contract, noting that frictions in the contract negotiation process may complicate the process of combining two distinct terms.

symmetric, "butterfly" fashion) along each ray. Though the expected contract term payoff is determined by the choice of r_i^* and a_i^*, the observed payoff will fall somewhere along the realization of its corresponding ray.

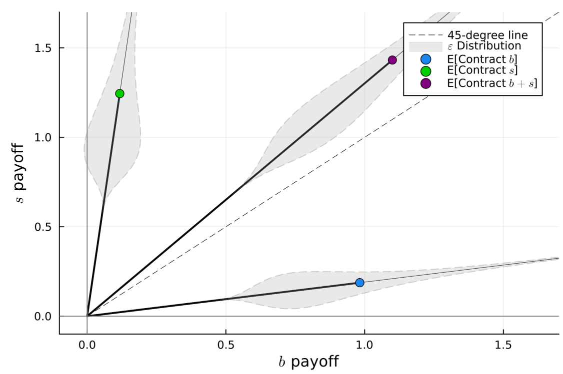
Figure 2.1: Contract term components in firm payoff space



(a) Representation of terms



(b) Creation of combined term



(c) Uncertainty in term production along angle of search

Notes: Each panel plots components of the contract term model to illustrate the choice set in payoff space. Successive panels add more features of the search game, but remove some notation to highlight the new features. Panel (a) begins with the payoffs of the proposed firm contract terms, and panel (b) adds the combined contract term from using both firms' terms. Panel (c) plots densities around the search radii corresponding to the densities of each term-specific shock ϵ, where the mean contract value is at the colored dots first plotted in the preceding panels.

With this framework in mind, we consider two possible games that are differentiated by when the two pieces of the contract are created. In the "price-first" game (PF), firms determine the price p_{PF}^* via Nash bargaining before searching for new terms $m_{i,PF}$, and then bargain over which terms m^*_{PF} to select. In the "price-last" game (PL), firms invest in and choose the contract terms m^*_{PL} first, and then the deal price p^*_{PL}. The productivity draws associated with the chosen contract in the two games are ϵ^*_{PF} and ϵ^*_{PL}.

Table 2.1: Timing of the two games

t	Price-first	Price-last
0	b and s decide to transact	b and s decide to transact
1	p^*_{PF} is chosen	$\{r^*_{i,PL}, a^*_{i,PL}\}$ are chosen
2	$\{r^*_{i,PF}, a^*_{i,PF}\}$ are chosen	m^*_{PL} is chosen
3	m^*_{PF} is chosen	p^*_{PL} is chosen

We now examine in detail how the prices and contract terms are determined in the two games. The timing of the two games is summarized in Table 2.1. Prices in both games are determined via Nash bargaining over expected equilibrium payoffs given available information when the price is chosen. We also assume the chosen contract maximizes the weighted Nash product of firms' continuation payoffs at the contract term stage. This is similar to the standard Nash bargaining framework (Nash, 1950) but with a discrete choice set $\mathcal{M}$ rather than a convex choice set; we call this Nash *bartering* to emphasize this distinction. While this modified framework does not have all the guarantees of standard Nash bargaining (in particular, the relationship of non-cooperative and cooperative bargaining), it provides a concise framing of the term bartering stage without taking a stand on the timing and rules of a sequential bargaining game.[17]

2.3.2 Contract creation in the price-first game

We now study the timing of the price-first contract game. We proceed by backward induction, first considering how contract terms are chosen, then firms' search choices for terms, and lastly the price bargaining game.

17. See Appendix B.1 for a more detailed discussion of the relationship of this model to standard Nash bargaining.

Bartering for terms. Since the contract price is fixed (under the default contract with standardized values $v_i(m_0) = 0$) before contract terms are chosen, the firms choose whichever term out of $\mathcal{M}^*$ yields the greatest Nash product:

$$\begin{aligned}
m^*_{PF} &= \underset{m_j \in \mathcal{M}^*}{\text{argmax}} \quad (v_b(m_j)\epsilon_j)^\tau \cdot (v_s(m_j)\epsilon_j)^{1-\tau} \\
&= \underset{m_j \in \mathcal{M}^*}{\text{argmax}} \quad \underbrace{\left[v_b(m_j)^\tau \cdot v_s(m_j)^{1-\tau}\right]}_{\delta_{j,PF}} \cdot \epsilon_j \\
&= \underset{m_j \in \mathcal{M}^*}{\text{argmax}} \quad NP_{j,PF} \\
s.t. \quad & v_i(m_j) \geq 0, \quad i \in \{b, s\}
\end{aligned}$$

The non-negativity constraint holds because neither firm will accept a contract term that reduces their individual surplus.[18] We also emphasize that the set of possible terms $\mathcal{M}^*$ that is considered under bargaining is itself an equilibrium object that was previously decided by firms' investment decisions. We consider this choice now.

Search for terms. Firms choose their search angle a_i and search radius r_i to maximize their expected net payoff from the term bartering stage. Each term-specific shock ϵ_j is crucial in determining which contract term is chosen, but these are not realized until after firms' decisions are made. Thus, the expected payoffs depend both on the Nash program in the term-bartering stage and the joint distribution of ϵ, which we leave unspecified for now.

$$\{r_i^*, a_i^*\} = \underset{r_i, a_i}{\text{argmax}} \quad \mathbb{E}[v_i(m^*; \epsilon^*) \mid m^* \in \mathcal{M}^* \text{ is chosen in game } PF] - c_i(r_i, a_i)$$

18. For the price-first game, this assumption functionally implies that neither party is made worse off than their status quo ante should a preliminary deal fail – so that neither party would become "damaged goods."

We write the equilibrium expected term-specific payoff (conditioning on the equilibrium firm choices r_i^* and a_i^*) as $U_{i,PF}^*$. These expected payoffs, and their associated costs, are considered by the forward-looking firms when deciding on the contract price.

Bargaining for prices. Firms condition on their expected equilibrium net payoffs $U_{i,PF}^* - c_i(r_{i,PF}^*, a_{i,PF}^*)$ in the continuation game as a reference point when bargaining. The equilibrium price is determined by

$$\begin{aligned} p_{PF}^* &= \underset{p\in\mathbb{R}_+}{\operatorname{argmax}} \quad (\pi_b - p + U_{b,PF}^* - c_b(r_{b,PF}^*, a_{b,PF}^*))^{\tau} \\ &\qquad\qquad\qquad\qquad \cdot (p - \pi_s + U_{s,PF}^* - c_s(r_{s,PF}^*, a_{s,PF}^*))^{1-\tau} \\ &= \tau(\pi_s - U_{s,PF}^* + c_s(r_{s,PF}^*, a_{s,PF}^*)) + (1-\tau)(\pi_b + U_{b,PF}^* - c_b(r_{b,PF}^*, a_{b,PF}^*)) \end{aligned}$$

In other words, the firms split both the expected surplus from the sale of the asset and the expected net surplus created by new contract terms according to their relative bargaining power.

2.3.3 Contract creation in the price-last game

We now study the timing of the price-last contract game. We proceed by backward induction, first considering how the price is set given contract terms, then how the contract terms are chosen, and lastly the firms' choice of term production.

Bargaining for prices. The contract price is chosen only after the contract terms are decided and the cost of finding these terms is sunk. Thus, the price for any chosen terms

m^*_{PL} (with the associated shock ϵ^*_{PL}) is:

$$\begin{aligned}
p^*_{PL} &= \underset{p\in\mathbb{R}_+}{\operatorname{argmax}} \quad (\pi_b - p + v_b(m^*_{PL};\epsilon^*_{PL}))^\tau \cdot (\pi_s + p - v_s(m^*_{PL};\epsilon^*_{PL}))^{1-\tau} \\
&= \tau(\pi_s - v_s(m^*_{PL};\epsilon^*_{PL})) + (1-\tau)(\pi_b + v_b(m^*_{PL};\epsilon^*_{PL}))
\end{aligned}$$

That is, setting the price after terms are decided means that firms negotiate the price to split the newly created value from the contract.

Bartering for terms. Both firms anticipate that the price negotiated in the last stage of the game splits the surplus from the contract according to each firm's relative bargaining power (disregarding sunk costs). Thus, the firms choose the terms that solve the Nash program:

$$\begin{aligned}
m^*_{PL} &= \underset{m_j\in\mathcal{M}}{\operatorname{argmax}} \quad [\tau(v_b(m_j)+v_s(m_j))\epsilon_j]^\tau \cdot [(1-\tau)(v_b(m_j)+v_s(m_j))\epsilon_j]^{(1-\tau)} \\
&= \underset{m_j\in\mathcal{M}}{\operatorname{argmax}} \quad \tau^\tau(1-\tau)^{(1-\tau)}[v_b(m_j)+v_s(m_j)]\cdot\epsilon_j \\
&= \underset{m_j\in\mathcal{M}}{\operatorname{argmax}} \quad \underbrace{[v_b(m_j)+v_s(m_j)]}_{\delta_{j,PL}}\cdot\epsilon_j \\
&= \underset{m_j\in\mathcal{M}}{\operatorname{argmax}} \quad NP_{j,PL} \\
s.t. \quad & v_b(m_j)+v_s(m_j) \geq 0
\end{aligned}$$

Note that this program is equivalent to maximizing the combined surplus generated by the new contract terms, regardless of who benefits from that term. This is an intuitive result: Because the last stage of the game will split the total surplus available according to exogenously given bargaining power, neither side benefits from selecting non-price terms that do not maximize the total expected "pie." In slight contrast to the price-first game, there is a non-negativity constraint implying that new terms will only be considered if they are a net *joint* improvement over the default contract m_0.

Search for terms. Each firm chooses search angle a_i and search radius r_i knowing how both the final contract terms and price will be chosen. In this case, the firms choose to maximize their share of the expected joint surplus minus the cost from finding these terms.

$$\begin{aligned}
\{r_b^*, a_b^*\} = \underset{r_b, a_b}{\operatorname{argmax}} \quad & \tau \cdot \mathbb{E}[v_b(m^*; \epsilon^*) + v_s(m^*; \epsilon^*) \mid m^* \in \mathcal{M}^* \text{ is chosen in game } PL] \\
& - c_b(r_b, a_b) \\
\{r_s^*, a_s^*\} = \underset{r_s, a_s}{\operatorname{argmax}} \quad & (1-\tau) \cdot \mathbb{E}[v_b(m^*; \epsilon^*) + v_s(m^*; \epsilon^*) \mid m^* \in \mathcal{M}^* \text{ is chosen in game } PL] \\
& - c_s(r_s, a_s)
\end{aligned}$$

This stage differs materially from the firm problem in the price-first stage: instead of receiving the full benefit of their search, firms only receive a share of the combined firms' surplus from the chosen term. Equivalently, both parties are aware that their search costs will become sunk (and thus disregarded) in subsequent stages. As in the price-first game, it is helpful to denote the equilibrium expected payoff to firm i from the chosen contract term as $U^*_{i,PL}$.

2.4 Characterizing the equilibrium contract

Having laid out the basic structure of our bargaining/bartering model in both the price-first and price-last structures, we now proceed to explore comparisons between the competing approaches. Our framing thus far has been deliberately general, which limits our ability to solve directly for firms' equilibrium choices, since that will tend to turn on specific functional forms related to productivity shocks (ϵ) and search costs (c_i). We now impose some additional restrictions in order to directly analyze the firms' decisions, developing core intuitions in the process.

We proceed by imposing a set of three simplifying assumptions that allow us to obtain closed-form solutions for equilibrium strategies, along with direct comparisons of the price-first and price-last structures. The first two assumptions limit the directionality of and correlation between search efforts, ensuring that the equilibrium contract always incorporates terms proposed by both firms. The third assumption imposes a general functional form for search costs. These assumptions will allow us to pin down each firm's search decisions and illustrate comparative statics with respect to bargaining power and search cost parameters.

Having imposed these assumptions, we then begin by exploring two special cases of the model that tightly constrain each firm's direction of search for new terms. In the first case, each firm is restricted to searching for terms that are value-enhancing for itself but value-neutral for the other firm (i.e., $a_b = -0.25$ and $a_s = 0.25$). We call this orthogonal, or self-interested, search. In the second case, we restrict each firm to search only for terms that are symmetrically value-enhancing for both itself and the other firm (i.e., $a_b = a_s = 0$). We call this case aligned, or surplus-maximizing, search. In both restricted settings, firms are allowed their discovered terms to create a composite/joint term m_{bs} that becomes part of the choice set along with the individually discovered terms m_b and m_s. These two special cases provide helpful intuition about equilibrium behavior when firms' directionality of search is exogenous; however, neither restriction is necessary to pin down equilibrium strategies.

In both cases, the price-first model (weakly) dominates the price-last model on efficiency grounds. In the "orthogonal search" case, the price first model is Pareto optimal relative to the price last model for both firms across the full range of values for the unrestricted parameters. Notably, for all interior values of the bargaining power, this is a strict improvement for both firms. In the "aligned search" case, the price-first model is Kaldor-Hicks optimal relative to the price last model across the full range of values for the unrestricted parameters. For all cases with symmetric search costs *except* the case of equal bargaining power, the price-first model is strictly dominant on efficiency grounds.

We then consider a third case in which we assume firms' search angles are endogenously determined but bounded. These bounds may arise due to technological constraints, professional norms, good-faith bargaining obligations, or other forces that prevent searches that are "too selfish" from occurring. This case provides additional intuition for firms' incentives when both their angle and direction of search are at least partially non-contractible: the price-first structure incentivizes more investment in novel contract terms than in the price-last game. When it is sufficiently costly to search for the most welfare-enhancing terms (i.e., along the 45-degree line, or $a_b = a_s = 0$), the price-first game creates more total surplus relative to the price-last game. More broadly, the two games yield different contracts in expectation as a result of the firms' differing incentives across the two games.

Finally, we discuss a generalization of this model that allows each firms' contract development efforts to affect the probability any individual terms are chosen, which in turn affects the expected payoff from the chosen contract. This generalization draws on standard tools in the discrete choice literature and allows for closed-form choice probabilities for each of the proposed contract terms. We allow for the firms' search for new contract terms to be unrestricted, though the firms' incentives to propose valuable contract terms lead them to behave similarly to the restricted setting.

This dominance of the price-first model in the first two cases and across a large range of values in the third case and generalized case is driven by properly incentivizing each firm's search for new terms. In the price-first model each firm captures more of their realized value of the discovered terms. In the price-last model, realized value of the discovered terms is redistributed based on the relative bargaining power of the firms, which leads firms to under-invest in the search for new terms. This difference is exacerbated when one firm is a more efficient searcher and can create more expected surplus than the other firm, regardless of the firms' relative bargaining power.

2.4.1 Simplifying assumptions and characterization of equilibrium search

In order to achieve tractability in our comparitive statics analysis, we begin by making the following simplifying assumptions regarding the term search stage:

A1 The direction of search is limited to the first quadrant so that all proposed contracts have weakly positive payoffs, i.e. $|a_i| \leq \bar{a} \leq 0.25$.

A2 The productivity shocks are perfectly correlated, i.e. $\epsilon_{bs} = \epsilon_b = \epsilon_s$.

Assumption **A1** introduces the possibility that the directionality of each firm's search (a_i) is partially contractible and can be constrained to (weakly) Pareto improving directions. The two special cases studied below present even stronger contractibility assumptions, whereby firms can specify a precise directionality a_i instead of a range. Although this latter degree of contractibility may be difficult in practice due to the challenge of monitoring the other firms' efforts in contract creation, it provides a useful starting point to understand each firms' incentives to expend effort in creating a novel contract. Note that even in this setting, the multiplicative nature of the productivity shock ϵ_j means that firms cannot verify whether any realized contract value is due to the other firm's search intensity or dumb luck. Thus, search intensity r_i is not directly verifiable and therefore is assumed not contractible in this restricted setting.

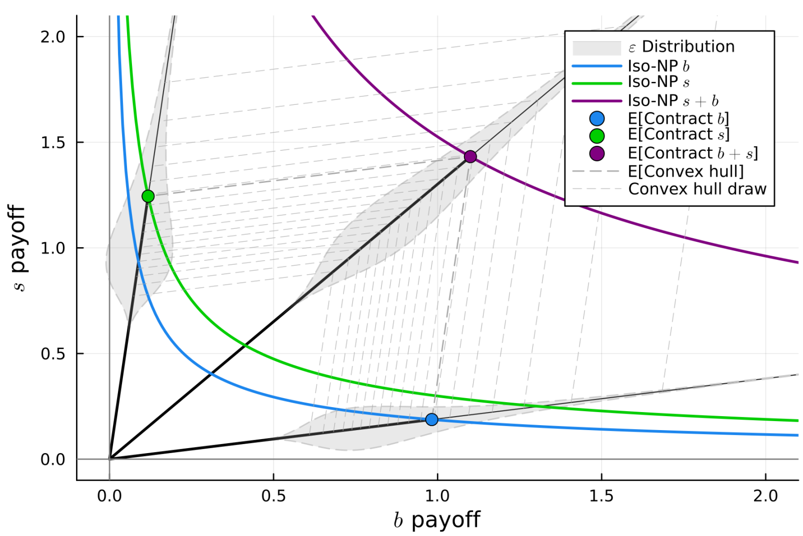
Assumption **A2** imposes an additional restriction that firms' individual and joint term production processes have perfectly correlated shocks. This may occur because of deal-specific challenges, or perhaps due to more willingness to accept non-boilerplate contract terms (for high productivity shocks). While perfect correlation in productivity shocks is admittedly a strong assumption, it significantly simplifies the firms' expectations over term payoffs and emphasizes that the ability to innovate new terms is related to deal-specific factors. (And in any event, we explicitly relax this assumption in the subsequent section.)

Under assumptions **A1** and **A2**, we have the orderings $NP_{bs,PF} > NP_{b,PF}$, $NP_{s,PF}$ and $NP_{bs,PL} > NP_{b,PL}$, $NP_{s,PL}$ for all $\tau \in (0, 1)$, regardless of the choice of radii r_i. Thus implying

> **Lemma 1.** Let **A1** and **A2** hold. Then both b's and s's contract terms are always incorporated into the final contract in both games.

An illustration of this is presented in Figure 2.2 for $r_b = 1$, $r_s = 1.25$, $a_b = -0.19$, $a_s = 0.22$, and $\tau = 0.4$ for the price-first game. The expected contract payoffs are plotted in the direction of search, along with the expected convex hull connecting them. Several other realizations of the convex hull are also plotted in lighter colors; these are all proportional because of assumption **A1**. The curves in Figure 1 represent the iso-Nash product lines, or the set of all contracts with equivalent Nash products. As shown in Lemma 1, the combined term yields a higher Nash product than the individual terms for any realization of ϵ.

Figure 2.2: Firm payoffs in the bartering stage of the price-first game (perfect correlation in ϵ)



Notes: None of these values necessarily represent equilibrium actions. Dashed lines represent possible convex hulls of the choice set of terms, for varying draws of ϵ with associated densities plotted around the ray corresponding to each contract as in Figure 2.1(c).

Lemma 1 has important implications for our analysis. First, it allows us to simplify the expected term-stage payoffs from the previous section as simply the expected term-stage payoffs from the combined term m_{bs}. Second, in both of the cases we next consider, it implies that each firm's choice of r_i is not affected by the other's search costs. Since each firm's optimal search intensities are driven only by bargaining power and their own search costs, the following analysis holds even when each firm's search cost is private information to each firm.

Together with general assumptions on the term cost function, these assumptions imply existence and uniqueness of each firm's investment equilibria in both the price-first and price-last games for a given set of search angles a_i. They also allow for a relative ordering of each firms' search efforts between the price-first and price-last settings.

Proposition 6. Let **A1** and **A2** hold, and let a_i be fixed for $i \in \{b, s\}$. Further assume search costs $c_i(r_i, a_i)$ are increasing and strictly convex in r_i with $c_i(0, a_i) = 0$ and $\frac{\partial c_i(r_i,a_i)}{\partial r_i} |_{r_i=0}= 0$. Then

(i) the equilibrium of the term choice stage exists and is unique for both the price-first and price-last games.

(ii) the search radius for firm b is weakly higher in the price-first game than in the price-last game when $a_b \leq \frac{1}{\pi}\arctan(\frac{1-\tau}{\tau}) - 0.25$.

(iii) the search radius for firm s is weakly higher in the price-first game than in the price-last game when $a_s \geq \frac{1}{\pi}\arctan(\frac{1-\tau}{\tau}) - 0.25$.

(iv) in the price-first game, both firms search strictly less than is socially optimal except when $|a_i| = 0.25$, in which case both firms search at the socially optimal level.

(v) in the price-last game, both firms search strictly less than is socially optimal except for when one firm has all the bargaining power ($\tau \in \{0, 1\}$), in which case only that firm searches at the socially optimal level.

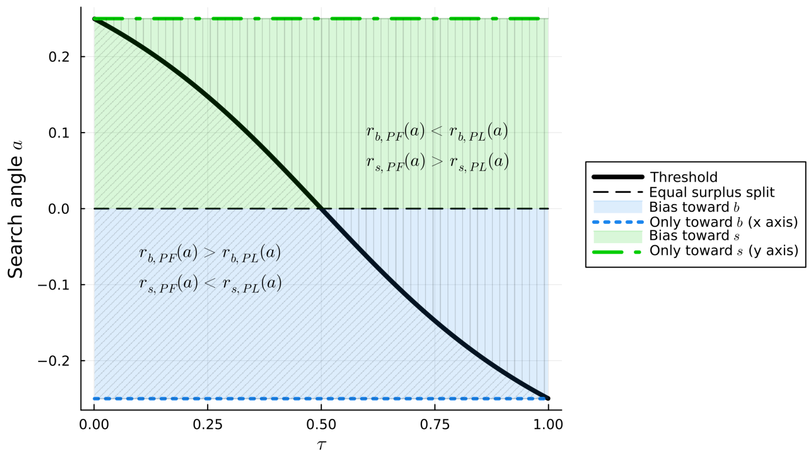
(Proof in Appendix B.2)

These results are straightforward: firms have a unique choice of search intensity (radius) for any prescribed search angle due to convex search costs. However, this result allows for a direct comparison between the outcomes of the two different versions of the contract game. In particular, firms search more intensely in price-first settings where they have relatively less bargaining power and their search angle is more biased in their own favor. Still, neither firm has incentives to search at the socially optimal level unless they can obtain all the surplus generated from their efforts.

Figure 2.3 illustrates heuristically the regions in which firms search more or less in the price-first game relative to the price-last game, as a function of a given directionality a and (buyer) bargaining power τ. In the vertically hatched region, the the seller's search intensity is greatest in the price-first game while the buyer's is greatest in the price last game. In the diagonally hatched regions, these orderings reverse for both the seller and buyer.

To understand the intuitions behind this Figure, consider the "northeast" quadrant where $a \in (0, 0.25)$ and $\tau \in (0.5, 1)$; that is, where the posited search angle is tilted toward the seller and the bargaining power is tilted toward the buyer. In this region, firm s searches more intensely when price is set first, since the fruits of its efforts cannot be expropriated by the powerful buyer later on. In contrast, firm b searches relatively more intensely in the price-last game, since the posited direction favors the seller and thus the buyer can only appropriate the fruits of its efforts through using superior bargaining power to extract price concessions. The "southwest" quadrant where $a \in (-0.25, 0)$ and $\tau \in (0, 0.5)$ is symmetric to the "northeast" quadrant, with the incentives reversed. In the remaining off-diagonal quadrants that contain the thick black curve (which represents the set of points at which each firm chooses the same search intensity in both games), bargaining power is more aligned with the posited search angle, resulting in a larger degree of near indifference (by both players) between the two approaches.

Figure 2.3: Relative search intensities for firms in price-first and price-last games



Notes: This figure represents results (ii) and (iii) of Proposition 6. Each optimal search radius r_i is a posited search angle a_i. The blue and green areas represent search directions that sit on opposite sides of the 45-degree line of an angle. The vertically-hatched region corresponds to the area where b's search intensity is greatest in the price-last game, and s's search intensity is greatest in the price-first game; the opposite is depicted in the diagonally-hatched region.

Although the qualitative characterization from Proposition 1 is interesting, additional insights (including comparative statics) are possible with the addition of an explicit functional form for the firms' search cost functions. The following assumption posits a flexible and intuitive structure:

A3 The investment cost function is $c_i(r_i, a_i) = 0.5\gamma_i r_i^2 \exp(-\gamma_a a_i^2)$ where γ_i is a firm-specific cost parameter for $i \in \{b, s\}$ independent of the search angle a_i, and $\gamma_a \geq 0$ is common to both firms.

Note that under this assumption, it is cheaper to search for more biased terms, i.e. those away from the 45-degree line. While in general we may expect search for more welfare-enhancing terms to be more costly, the limiting case of $\gamma_a = 0$ (where all angles of search are equally costly) provides a useful benchmark for our analysis.

The preceding assumptions are sufficient to deliver closed form characterizations for the firms' equilibrium search decisions for any given search angles. From Lemma 1, the term-search maximization problems of firms b and s in the price-first setting are respectively

$$\begin{aligned}
&\max_{r_b} \quad r_b \cos(\theta(a_b)) + r_s \cos(\theta(a_s)) - 0.5\gamma_b r_b^2 \exp(-\gamma_a a_b^2) \\
&\max_{r_s} \quad r_b \sin(\theta(a_b)) + r_s \sin(\theta(a_s)) - 0.5\gamma_s r_s^2 \exp(-\gamma_a a_s^2)
\end{aligned}$$

while for the price-last setting, the problems are

$$\begin{aligned}
\max_{r_b} \quad & \tau \cdot [r_b \cos(\theta(a_b)) + r_s \cos(\theta(a_s)) + r_b \sin(\theta(a_b)) + r_s \sin(\theta(a_s))] \\
& - 0.5\gamma_b r_b^2 \exp(-\gamma_a a_b^2) \\
\max_{r_s} \quad & (1-\tau) \cdot [r_b \cos(\theta(a_b)) + r_s \cos(\theta(a_s)) + r_b \sin(\theta(a_b)) + r_s \sin(\theta(a_s))] \\
& - 0.5\gamma_s r_s^2 \exp(-\gamma_a a_s^2)
\end{aligned}$$

Evaluating the first-order conditions of the respective maximization problems yields the optimal search radii for any given search angles, as reflected in Lemma 2.

Lemma 2. Let **A1**, **A2**, and **A3** hold. Then for any a_b and a_s, the optimal search radii are as follows in each case:

(i) the price-first game

$$\begin{aligned}
r^*_{b,PF}(a_b) &= \frac{1}{\gamma_b} \cos(\theta(a_b)) \exp(\gamma_a a_b^2) \\
r^*_{s,PF}(a_s) &= \frac{1}{\gamma_s} \sin(\theta(a_b)) \exp(\gamma_a a_s^2)
\end{aligned}$$

(ii) the price-last game

$$\begin{aligned}
r^*_{b,PL}(a_b) &= \frac{\tau}{\gamma_b} [\cos(\theta(a_b)) + \sin(\theta(a_b))] \exp(\gamma_a a_b^2) \\
r^*_{s,PL}(a_s) &= \frac{1-\tau}{\gamma_s} [\cos(\theta(a_s)) + \sin(\theta(a_s))] \exp(\gamma_a a_s^2)
\end{aligned}$$

(iii) the socially-optimal outcome

$$r^*_{b,opt}(a_b) = \frac{1}{\gamma_b}[\cos(\theta(a_b)) + \sin(\theta(a_s))] \exp(\gamma_a a_b^2)$$
$$r^*_{s,opt}(a_s) = \frac{1}{\gamma_s}[\cos(\theta(a_s)) + \sin(\theta(a_s))] \exp(\gamma_a a_s^2)$$

Significantly, note that equilibrium investment intensity in the price-first game does not depend on the bargaining power, while in the price-last game each firm's investment intensity is increasing in its relative bargaining power.

We next turn to a comparative statics analysis under three special cases. In the first two cases, each firm's search angle is exogenously assigned, reducing the search optimization problem of each firm to choosing one variable: search intensity r_i. In the first case we explore the comparative statics in the extreme case of orthogonal (self-interested) search. In the second case we explore the comparative statics in the opposite extreme of aligned search. In the third case, we assume the firms' equilibrium search angles are endogenously chosen, but bounded in absolute value by some exogenous constant $\bar{a} \leq 0.25$.

2.4.2 Orthogonal (self-interested) search

We first consider the special case in which both firms are constrained to search for terms that are value enhancing for themselves but payoff-neutral for their counterparty. Subject to this constraint, we solve for each firms search intensity in both the price first and price last model. In the first comparative statics analysis we assume firm-specific search costs γ_i are the same for both firms. In the second comparative statics analysis we assume search costs differ. In both analyses we find that in the special case of orthogonal search, both firms strictly prefer the price-first game to the price-last game for all but the most extreme values of the bargaining weight τ.

In the context of our model, orthogonal search means we assume the firm search angles are $a_b = -0.25$ and $a_s = 0.25$ (that is, along the x- and y-axes in Figure 2.2). Since we assume search angles are exogenous for this special case, we also fix the directional cost parameter at $\gamma_a = 0$. By Lemma 1 and the functional form assumptions above, we know that the chosen term in game G has the associated expected payoff pair $v_b(m_{bs,G}) = r_{b,G}$ and $v_s(m_{bs,G}) = r_{s,G}$. This yields the following equilibrium investment in the price-first and price-last games.

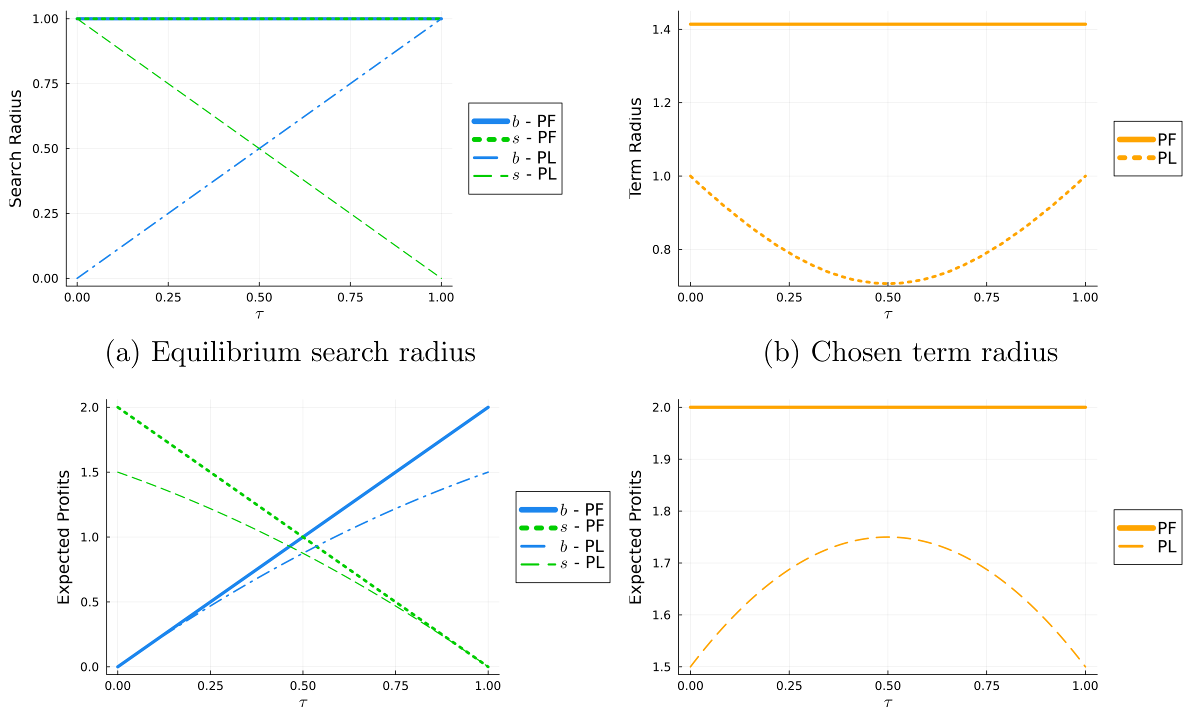
$$r^*_{b,PF} = \frac{1}{\gamma_b} \qquad r^*_{b,PL} = \frac{\tau}{\gamma_b}$$
$$r^*_{s,PF} = \frac{1}{\gamma_s} \qquad r^*_{s,PL} = \frac{1-\tau}{\gamma_s}$$

For all but extreme values of τ (i.e., $\tau = 0$ or $\tau = 1$), both firms under-invest in the price-last game relative to the price-first game. The socially optimal level of investment, given the fixed search angles, is obtained by both firms in the price-first game.

With $r^*_{i,PF}$ and $r^*_{i,PL}$ pinned down as a simple expression of exogenous parameters, we can now compare firm search intensity decisions and the corresponding firm payouts across the price first and price last models. We first compare outcomes when search costs γ_i are the same for both firms across the full range of values for the bargaining weight τ. We then compare outcomes when search costs are different for each firm by pinning down the search cost of the seller and then exploring how outcomes vary across a range of values for γ_b, the search cost of the buyer.

Figure 2.4 presents several comparative statics for both games when setting search costs equal across firms and varying τ. The default parameter values are $\gamma_b = \gamma_s = 1$, $\gamma_a = 0$, $\pi_b = 2$, and $\pi_s = 1$. Panel (a) demonstrates how, except in the extreme case of either $\tau = 0$ or $\tau = 1$, firms in the price-first setting search harder for new terms than in the price-last setting. As shown in panel (b), this yields more value-enhancing tailoring in the price-first game as well, particularly for intermediate levels of bargaining power. Regardless of the bargaining power held by each firm, panel (c) shows that both firms strictly prefer the price-first game to the price-last game for all $\tau \in (0, 1)$, yielding uniformly greater total surplus (as indicated in panel (d)).

Figure 2.4: Comparative statics with respect to τ (orthogonal search)



(a) Equilibrium search radius

(b) Chosen term radius

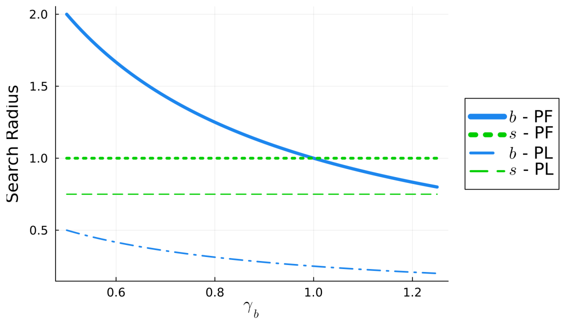
(c) Expected profits

(d) Expected combined profits

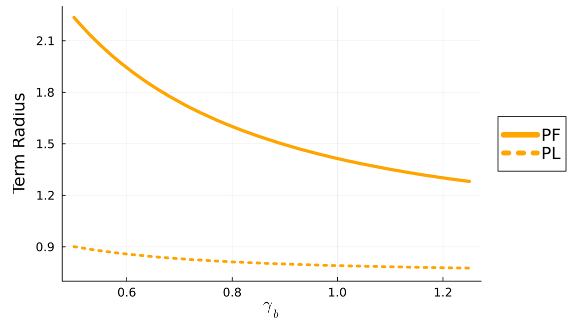
Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\gamma_b = \gamma_s = 1$, $\gamma_a = 0$, $\pi_b = 2$, and $\pi_s = 1$.

Figure 2.5 examines the alternative case where firm b's bargaining power remains fixed ($\tau = 0.25$), but its search costs γ_b vary around the default seller cost parameter $\gamma_s = 1$. As before, we maintain $\pi_b = 2$ and $\pi_s = 1$. In both games, both firms ultimately benefit when the buyer faces lower search costs. As can be seen in panel (c), however, both firms gain more from a reduction in γ_b in the price-first game than in the price-last game: firm b's investment response in the price-last game is muted by its inability to recover the full fruits of the investment in the price-last game. While the relative order of the contract terms and profits is preserved between the two games (as shown in Figure 2.4 for $\tau = 0.25$), we note that the contract terms and price are more sensitive to changes in γ_b in the price-first game relative to the price-last game.

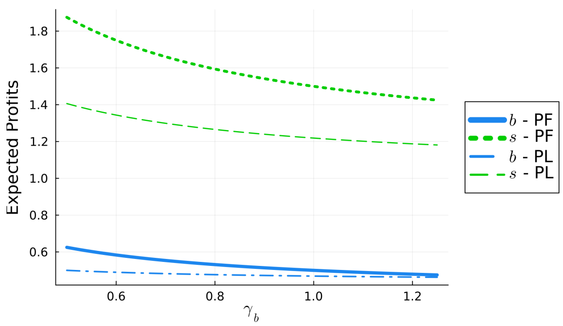
Figure 2.5: Comparative statics with respect to γ_b (orthogonal search)



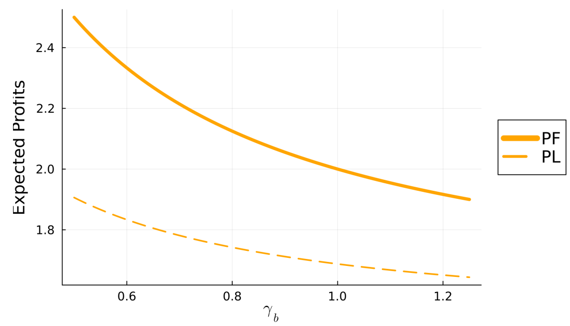
(a) Equilibrium search radius



(b) Chosen term radius



(c) Expected profits



(d) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\tau = 0.25$, $\gamma_s = 1$, $\gamma_a = 0$, $\pi_b = 2$, and $\pi_s = 1$.

2.4.3 Aligned (surplus-maximizing) search

We now explore the special case where firms are constrained to search in an aligned fashion, so that their search angles are equal at $a_b = a_s = 0$ (i.e., the 45-degree line, which is the expected-surplus-maximizing angle for any fixed search radius). As before, we fix $\gamma_a = 0$ and consider the firms' search intensities and expected payoffs with both identical and heterogeneous search costs. Applying Lemma 1, we observe the expected payoffs in game G

of $v_b(m_{bs,G}) = \sqrt{0.5}[r_{b,G} + r_{s,G}]$ and $v_s(m_{bs,G}) = \sqrt{0.5}[r_{b,G} + r_{s,G}]$.

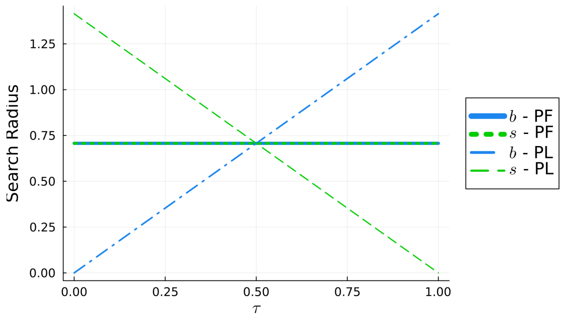
From the equilibrium search intensities and the posited search angle, the firms choose the following equilibrium search radii in the price-first and price-last games.

$$r^*_{b,PF} = \frac{\sqrt{0.5}}{\gamma_b} \qquad r^*_{b,PL} = \frac{\tau\sqrt{2}}{\gamma_b}$$
$$r^*_{s,PF} = \frac{\sqrt{0.5}}{\gamma_s} \qquad r^*_{s,PL} = \frac{(1-\tau)\sqrt{2}}{\gamma_s}$$

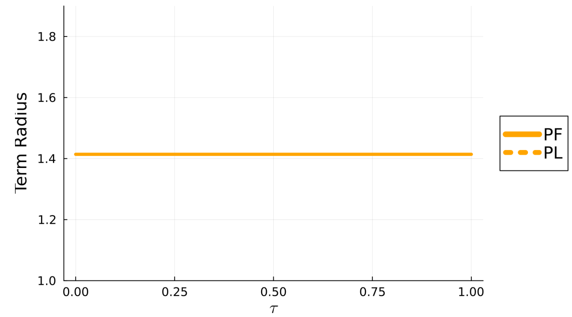
As before, the firms' respective search intensities do not turn on bargaining power in the price-first game. However, in the price-last game, the firm with more (less) bargaining power will over- (under-) invest in search relative to the price-first model; the two coincide for $\tau = 0.5$. As shown in Proposition 6, the socially optimal level of investment is only attained in the price-last game when one firm has all the bargaining power.

Figure 2.6 presents comparative statics in the aligned search case for varying values of τ in both games. The default values are $\gamma_b = \gamma_s = 1$, $\gamma_a = 0$, $\pi_b = 2$, and $\pi_s = 1$. Panel (a) shows that the firm with more bargaining power will choose a larger search radius in the price-last game than in the price-first game, though as indicated by panel (b) the chosen term is in expectation identical in both games. Panels (c) and (d) in turn plot the individual and combined profits for both firms in both the price-first and price-last games. In contrast to the orthogonal search setting (where the price-first game strictly dominates the price-last game), the two firms are indifferent between the two games when $\tau = 0.5$. However, their preferences diverge with unequal bargaining power. Here, the more powerful bargainer generally prefers the price-first game, while the less powerful bargainer generally leans the other way. As panel (d) shows, however, the preferences are not zero sum, and the price-first game once again outperforms the price-last game in terms of total payoff (for all but the case of $\tau = 0.5$, where they produce equivalent payoffs).

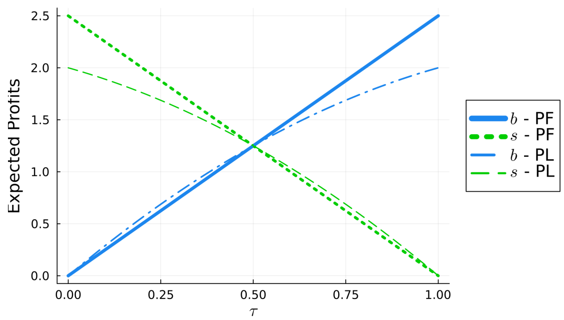
Figure 2.6: Comparative statics with respect to τ (aligned search)



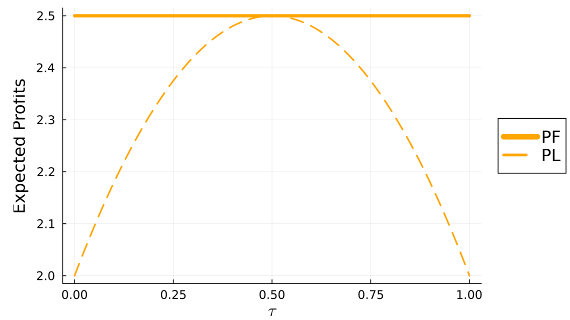
(a) Equilibrium search radius



(b) Chosen term radius



(c) Expected profits

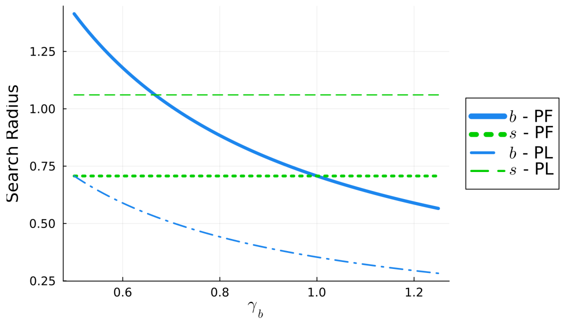


(d) Expected combined profits

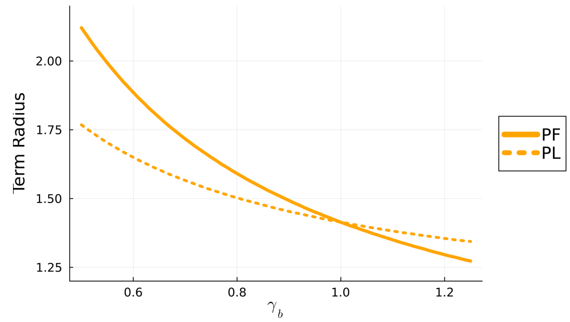
Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\gamma_b = \gamma_s = 1$, $\gamma_a = 5$, $\pi_b = 2$, and $\pi_s = 1$.

As above, we can also illustrate comparative statics in varying heterogeneous search costs. Figure 2.7 again shows where firm b has low bargaining power ($\tau = 0.25$) while varying γ_b around the default seller cost parameter $\gamma_s = 1$. As before, we maintain $\pi_b = 2$ and $\pi_s = 1$. When compared to Figure 2.5(a), Figure 2.7 (a) shows that firm b's response as search costs γ_b vary are in some ways similar to the orthogonal case. At the same time, while both firms gain from the lower search costs across games, the benefit is more drastic for firm b in the price-first game. In fact, for low enough search costs, firm b (the weaker bargainer) no longer prefers the price-last game; this differs from the symmetric-cost bargaining setting illustrated in Figure 2.6(c). Panel (b) shows that the price-first game yields more investment in new terms when the weaker bargainer is the stronger searcher, while overall innovation is less sensitive to changes in firm b's cost parameter in the price-last game. Since both firms search along the 45-degree line in this setting, the resulting contract terms always provide equal value to both parties.

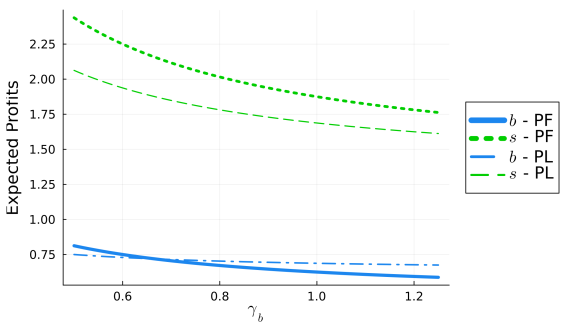
Figure 2.7: Comparative statics with respect to γ_b (aligned search)



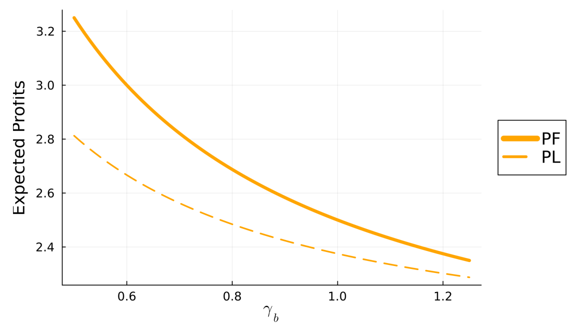
(a) Equilibrium search radius



(b) Chosen term radius



(c) Expected profits



(d) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\tau = 0.25$, $\gamma_s = 1$, $\gamma_a = 5$, $\pi_b = 2$, and $\pi_s = 1$.

Collectively, these comparative statics analyses under both orthogonal and aligned search demonstrate that, when one accounts for the value of term innovation in the contracting process, setting price first is either weakly Pareto optimal or Kaldor Hicks optimal relative to setting price last across the full range of exogenous parameter values. This prediction, although consistent with industry practice in high stakes M&A deals, stands in stark contrast to the standard intuition in contract design that welfare is maximized when parties barter over terms first and set price last.

2.4.4 Partially contractible term search

We now consider the case in which firms choose both their search intensity r^* and their search angle a_i^* subject to the constraint that $|a_i^*| \leq \bar{a}$. This case weakens the assumptions of the previous two cases, in which the angles are exogenous, but still restricts the angle of search to fall within some weak subset of the first quadrant. We assume this restriction arises from some combination of professional norms or the technology by which new terms are produced, ensuring that all terms must be at least weakly value-improving for both parties.

We begin by presenting a second proposition that follows from assumptions A1, A2, and A3 when we free up the search angle to be endogenous, though still constrained to fall within the first quadrant. To build intuition, we then consider two special cases of the value of the angle cost parameter: $\gamma_a = 0$ and $\gamma_a = 5$. Building on these special cases, we then present comparative statics for the full range of values for the bargaining weight τ and a wide range of values for γ_a. These comparative statics demonstrate that the price first game induces more aggressive term innovation relative to the price last game for all values of τ and γ_a and that the price first game is Pareto and/or Kaldor-Hicks dominant relative to the price last game for a wide range of the parameter space.

From the previous assumptions and Proposition 6, we obtain the unique optimal search radii as a function of the firms' search angles a_i. From the firms' optimal strategies for search intensity, we then solve for the optimal search angles under the constraint $|a_i| \leq \bar{a}$. We characterize this equilibrium in the following proposition.

Proposition 7. Let **A1**, **A2**, and **A3** hold. An equilibrium exists for both the price-first and price-last games when firms choose both the search radius r_i and the search angle a_i. Further

(i) in the price-first game, the unique optimal search angles are $a_{b,PF}^* = -\bar{a}$ and $a_{s,PF}^* = \bar{a}$.

(ii) in the price-last game for $\gamma_a \in [0, \pi^2]$, there is a unique optimal search angle

$a^*_{i,PL} = 0$.

(iii) in the price-last game for $\gamma_a > \pi^2$, each firm has two optimal search angles that are unique up to their sign. These angles coincide with the constraint, i.e. $|a^*_{i,PL}| = \bar{a}$, for all $\gamma_a \geq \frac{\pi}{\bar{a}} \tan(\pi \bar{a})$.

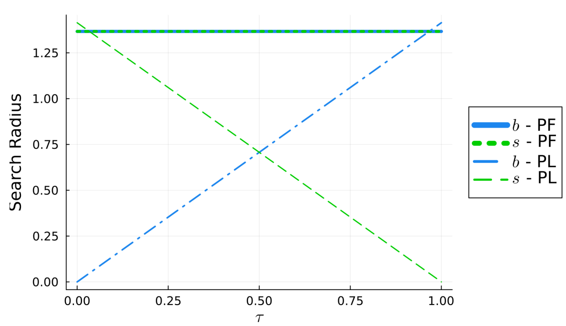
(Proof in Appendix B.2)

The symmetric equilibrium search angles exist in the price-last game because all surplus is evenly split between both firms, so only the amount (and not the original allotments) of total surplus matters. In this case, both firms are equally compensated for their efforts when searching for terms that improve either their or their counterparts' payoffs. For clarity, and to facilitate comparisons with the previous two sections, we restrict attention to equilibria where $a_b \leq 0$ and $a_s \geq 0$ (that is, each firm searches on its own side of the 45-degree line).

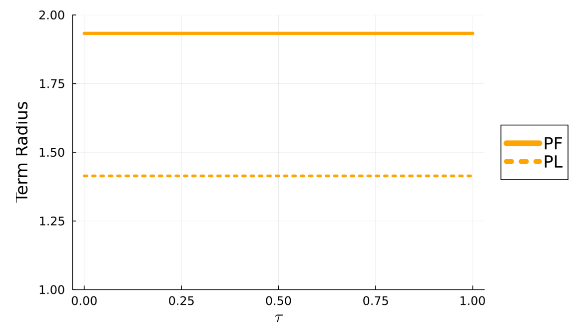
To fix ideas, we first examine the case where $\gamma_a = 0$ and $\bar{a} = 0.25$, i.e. when there is no penalty to searching along the 45-degree line and the entire first quadrant can be searched. In this case, the price-last game incentivizes the firms to search in the surplus-maximizing direction, since they will ultimately earn a share of the total surplus they generate. This coincides with the aligned search game considered above. In contrast, the price-first game incentivizes the firms to search in the most efficient direction to maximize their own payoff. Since the firms' search is limited only to the first quadrant, this coincides with the orthogonal search case. Evaluating firms' strategies and outcomes for various values of τ reveals that both the price-first and price-last games yield the same radius and angle for the resulting contract term, regardless of the value of τ. Further, both firms prefer the price-last game to the price-first game for any $\tau \in (0, 1)$ (see Figure B.2 in the appendix for more details).

To compare firms' search decisions and outcomes in this more relaxed setting, we now present several comparative statics with respect to the key parameter values in this model: the bargaining weight τ and the angle-specific cost parameter γ_a. Figure 2.8 illustrates the case where $\gamma_a = 5$ and $\bar{a} = 0.25$. Panel (a) shows that, with the exception of the stronger bargaining firm for extreme values of τ, the price-first game incentivizes larger firm search radii than the price-last game; these imply the equilibrium radius of the chosen term is larger in the price-first game than the price-last game (see panel (b)). Panels (c) and (d) together illustrate how the higher cost to searching along the 45-degree line yields higher surplus in settings where one firm is a particularly stronger bargainer. This holds even when the cost parameter γ_a is not sufficiently large to deter firms from searching in the surplus-maximizing direction in the price-last game.

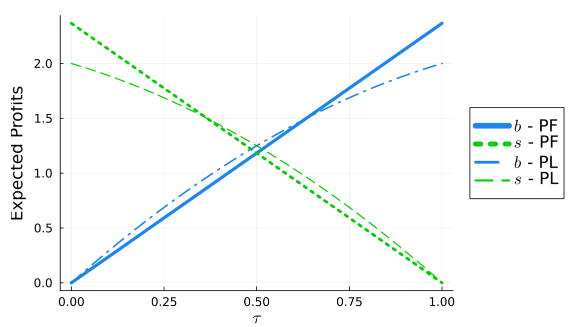
Figure 2.8: Comparative statics with respect to τ (endogenous angle search)



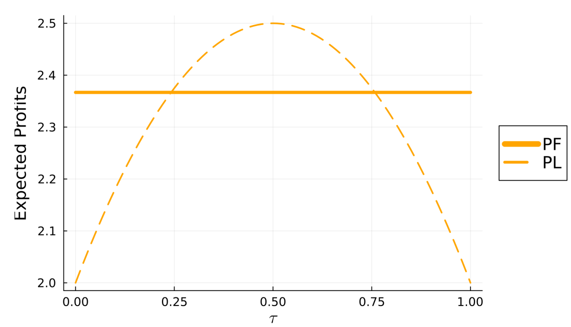
(a) Equilibrium search radius



(b) Chosen term radius



(c) Expected profits

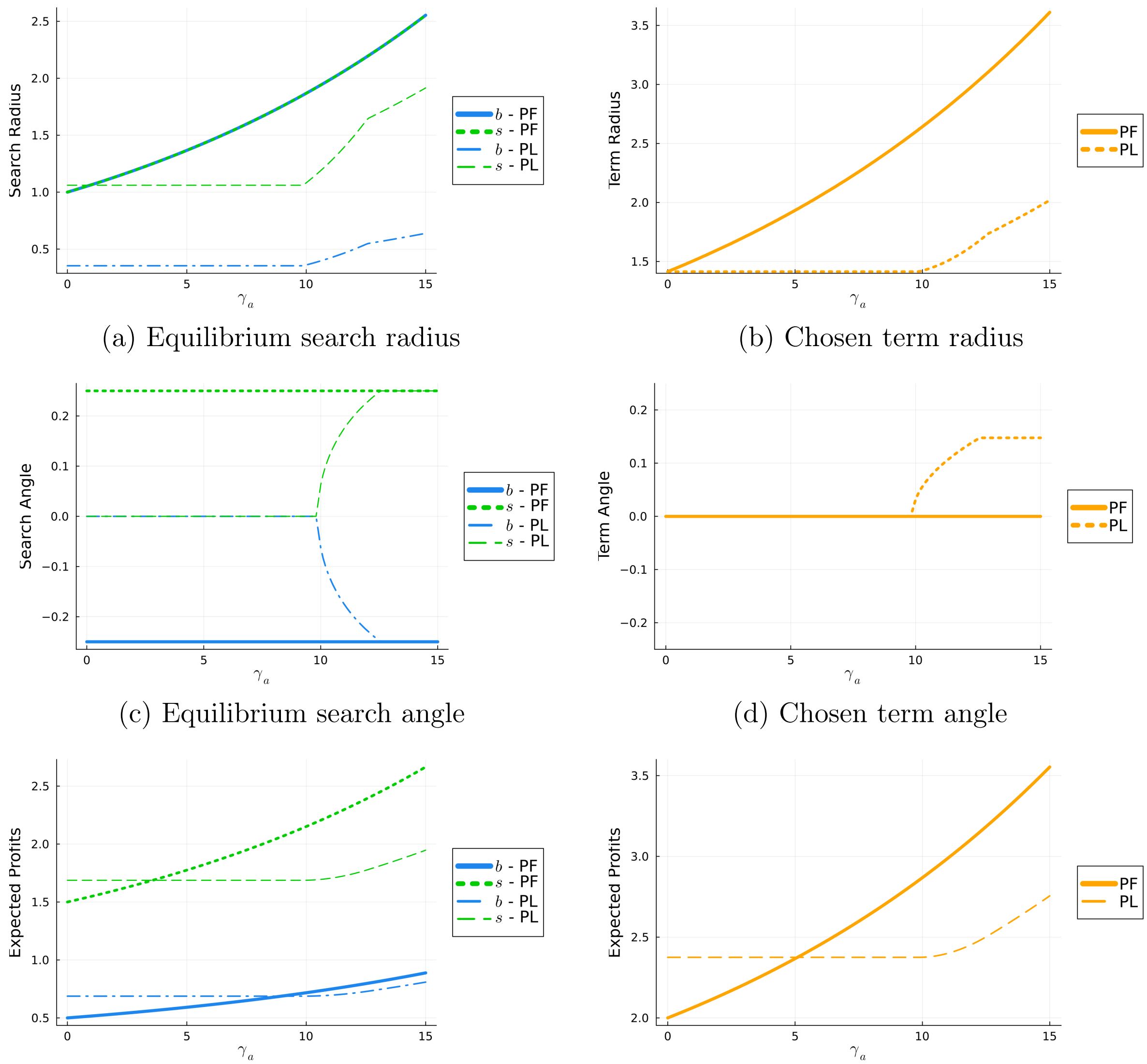


(d) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\gamma_b = \gamma_s = 1$, $\gamma_a = 5$, $\bar{a} = 0.25$, $\pi_b = 2$, and $\pi_s = 1$.

We now examine comparative statics with respect to γ_a in Figure 2.9. For these figures, we set $\tau = 0.25$ and $\gamma_b = \gamma_s = 1$. As shown in panel (a), increasing γ_a makes searching near the axes (i.e., for high $|a_i|$) cheaper, incentivizing both firms to increase their search radii whenever $a_i \neq 0$. Firms' additional efforts in increasing the search radius implies strictly more innovation in the chosen contract term in the price-first game relative to the price-last game, as shown in panel (b). Panel (c) plots the equilibrium search angles as indicated by Proposition 7, with the additional restriction that $a_b \leq 0$ and $a_s \geq 0$. When the cost of searching for surplus-maximizing (i.e., low $|a_i|$) terms is sufficiently high, the price-last game will be biased toward the firm with more bargaining power (panel (d)). Together, this implies that sufficiently high γ_a implies the price-first game generates more total surplus (see panel (f)). In fact, panel (e) indicates that the price-first game may even be strictly preferred by both firms for high enough values of γ_a.

Figure 2.9: Comparative statics with respect to γ_a (endogenous angle search)



(a) Equilibrium search radius

(b) Chosen term radius

(c) Equilibrium search angle

(d) Chosen term angle

(e) Expected profits

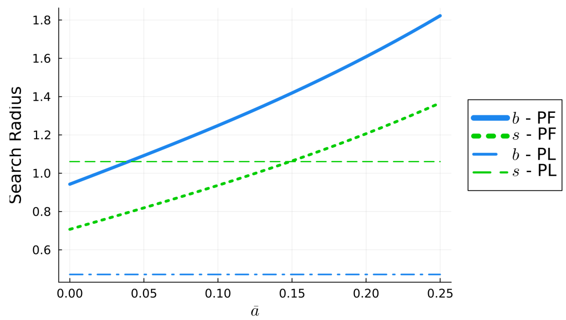
(f) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\tau = 0.25$, $\gamma_b = \gamma_s = 1$, $\pi_b = 2$, and $\pi_s = 1$.

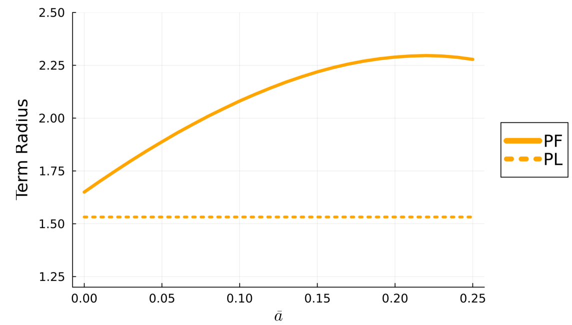
Now we present comparative statics with respect to the search angle constraint $\bar{a}$ in Figure 2.10. Since γ_a is sufficiently low, the price-last game does not incentivize firms to do anything other than search along the 45-degree line. This contrasts with the price-first game, which as shown in Proposition 2(i) incentivizes firms to search at their boundary (either $-\bar{a}$

for firm b or $\bar{a}$ for firm s). This implies that both the resulting term radius and the expected combined profits are non-monotonic in $\bar{a}$. Despite the two firms monotonically increasing their search efforts in panel (a), their increasingly selfish search implies that this yields less total innovation than for smaller values of $\bar{a}$. Interestingly, panel (f) shows that the joint profit-maximizing constraint for the price-first game is approximately $\bar{a} = 0.12$. While full contractibility on the firms' search angles may not be possible, some partial restrictions may in fact make the price-first game generate more surplus.

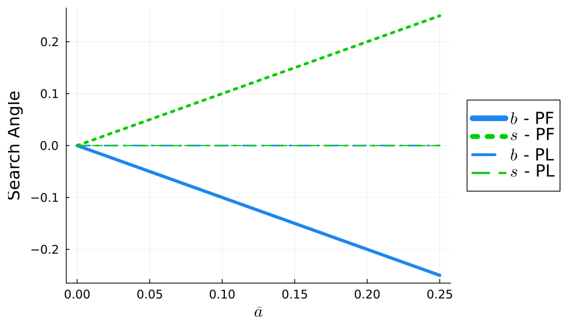
Figure 2.10: Comparative statics with respect to $\bar{a}$ (endogenous angle search)



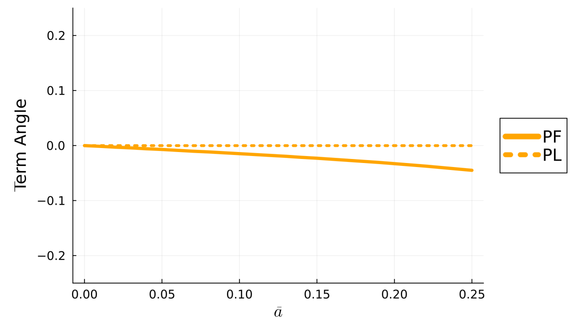
(a) Equilibrium search radius



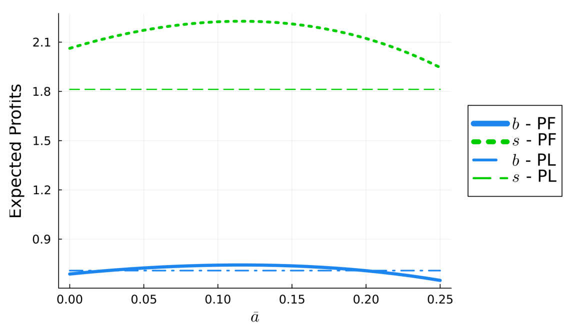
(b) Chosen term radius



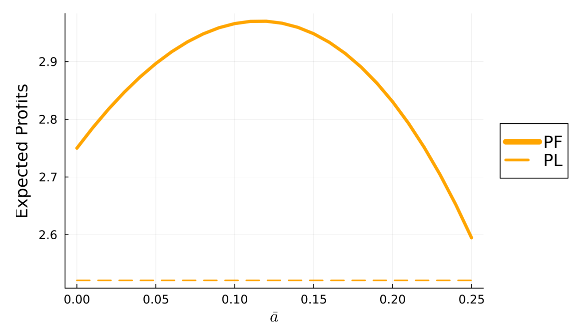
(c) Equilibrium search angle



(d) Chosen term angle



(e) Expected profits



(f) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\tau = 0.25$, $\gamma_b = \gamma_s = 1$, $\gamma_a = 5$, $\pi_b = 2$, and $\pi_s = 1$.

2.4.5 Equilibrium contract under alternative assumptions

We now briefly examine another specification of the model, which allows for endogenous search angles and radius under various alternative assumptions. Unlike in the previous section, firms can now search for any type of contract term—even those that may be actively harmful to their counterpart. We also allow for term-specific productivity shocks to vary across terms, implying that any individual firm's proposed contract term may be preferred to the combined contract term in a specific setting. Thus, firms have full flexibility in using their search decisions to determine the expected payoff from their proposed term, as well as the probability it is selected. The central question is now how timing affects the contract creation process in the absence of any restrictions on firms' search process.

In order to understand this question, we continue to make some simplifying assumptions for tractability. Instead of assuming the shocks ϵ_j are perfectly correlated, we instead assume they are independent and make a functional form assumption that yields closed-form choice probabilities. This means there is an option value to variety even if the terms have the same expected value for both parties. Thus, firms' investment decisions are shaped by the knowledge that their own term may be chosen instead of the combined term, since the combined term may be ill-suited for a deal relative to either of the simpler individual contract terms.

Modifying these assumptions shows how firms' incentives differ when their individual terms may be chosen. Since productivity shocks are independent, and firms have a nonzero chance of only their term being chosen in the bartering process, they offer more neutral terms (lower $|a_i|$) in the price-first game, and they increase their search radii when they have more bargaining power. Firms behave similarly in the price-last game as under the previous assumptions, generally searching in the surplus-maximizing direction because all surplus will be redistributed later and a larger "pie" is beneficial. We study this setting in detail now.

Alternative assumptions and implications for the expected contract

We replace assumptions **A1** and **A2** with the following two assumptions:

B1 The direction of search may be any angle within the entire unit circle, i.e. $|a_i| \leq 1.0$.

B2 ϵ_j is i.i.d. Frechet (inverse Weibull) with shape parameter $\alpha > 1$ and scale parameter $\sigma = \Gamma(1 - 1/\alpha)^{-1}$, implying $\mathbb{E}[\epsilon_j] = 1$ for all j.

The first assumption expands the set of possible contract terms to include those that may be value-destroying for one of the two parties.[19] The second assumption helps achieve tractability by yielding functional forms for conditional choice probabilities and conditional expected values, as in Eaton and Kortum (2002) and more broadly in the empirical discrete choice literature (see e.g. S. T. Berry and Haile 2021).[20]

Assumption **B2** implies several closed forms for important quantities in both the price first and price last game. Recall that in the bartering for terms stage of game G, firms choose whichever term of $\mathcal{M}^* = \{m_s^*, m_b^*, m_{bs}^*\}$ yields the greatest Nash product:

$$m_G^* = \underset{m_j \in \mathcal{M}^*}{\text{argmax}} \quad NP_{j,G}$$

where $NP_{j,PF} = \delta_{j,PF} \cdot \epsilon_j$ and $\delta_{j,PF} = v_b(m_j)^\tau \cdot v_s(m_j)^{1-\tau}$, while $NP_{j,PL} = \delta_{j,PL} \cdot \epsilon_j$ and $\delta_{j,PL} = v_b(m_j) + v_s(m_j)$. [21]

Then the equilibrium choice probability for any term j is

$$\lambda_{j,G}^* \equiv \mathbb{P}[j \text{ is chosen in game } G] = \frac{(\delta_{j,G}^*)^\alpha}{\sum_k (\delta_{k,G}^*)^\alpha}$$

Thus, any term that offers strictly positive surplus to both firms in the price-first game has a strictly positive probability of being selected through Nash bartering. Using the expectation

19. These bounds are only imposed to avoid coterminal angles, i.e. those that differ by some multiple of 2π.

20. Importantly, $\epsilon_{bs} \perp\!\!\!\perp \epsilon_b$, ϵ_s. While this is a strong assumption, it will be helpful in both theoretical and empirical applications; such benefits will be highlighted below. More complex correlation structures may also be helpful; see e.g. the nested Frechet model in Lashkaripour and Lugovskyy, 2023 as well as the larger literature using variations of the nested logit model for demand estimation. We use the extreme case of full independence in contrast with the other extreme, perfect correlation, which we consider above.

21. For ease of reference, recall these derivations are found in Subsection 2.3.2 (price-first) and Subsection 2.3.3 (price-last).

of the maximum of Frechet random variables,[22] we have

$$\mathbb{E}[\epsilon^*_{j,G}] = \frac{1}{\delta^*_{j,G}} \cdot \Big(\sum_k (\delta^*_{k,G})^\alpha\Big)^{1/\alpha} = (\lambda^*_{j,G})^{-1/\alpha}$$

This expression illustrates the option value of choosing between multiple contracts - even though all ϵ_j have mean 1, conditioning on that term being chosen means the Nash product (and the associated payoffs) will be higher on average than the unconditional average payoffs. Finally, by the law of total expectation, the expected value of firm payoffs from the term stage, $U_{i,G}$, is written as

$$U^*_{i,G} = \sum_{j \in \mathcal{M}} v_i(m^*_{j,G}) \cdot (\lambda^*_{j,G})^{1-1/\alpha}$$

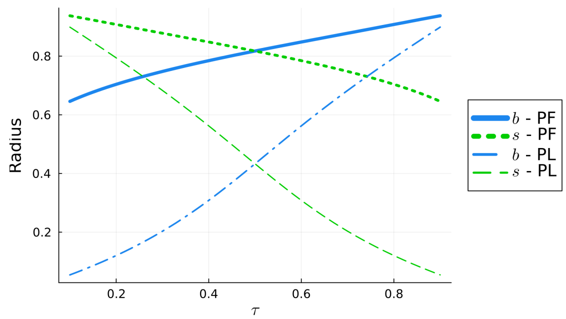
That is, the expected value from each individual contract is weighted by the probability it is chosen and its expected value conditional on being chosen in game G.

Comparative statics

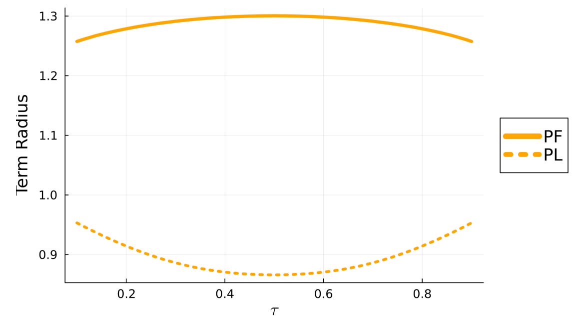
Figure 2.11 plots the outcomes for the contracting process for different values of the bargaining weight τ. In this case we see that both games induce greater search efforts by the stronger bargainer, as represented by the larger search radius. Note that relative to the game in Figure 2.8, the search angles in the price-first game are shifted toward the center, particularly for the firm with less bargaining power. This is because the bartering process rewards terms that offer a higher Nash product, particularly by placing non-zero probability on choosing any individual term proposed by a single firm. However, the firms in the price-last game still search at $a_i = 0$ since it is surplus-maximizing.

22. In particular, $\mathbb{E}[\epsilon^*_{j,G}] = \frac{1}{\delta_{j,G}}\mathbb{E}[NP_{j,G} \mid j \text{ is chosen in } G]$ and under **B2** it holds that $\mathbb{E}[NP_{j,G} \mid j \text{ is chosen in } G] = (\sum_k (\delta^*_{k,G})^\alpha)^{1/\alpha}$.

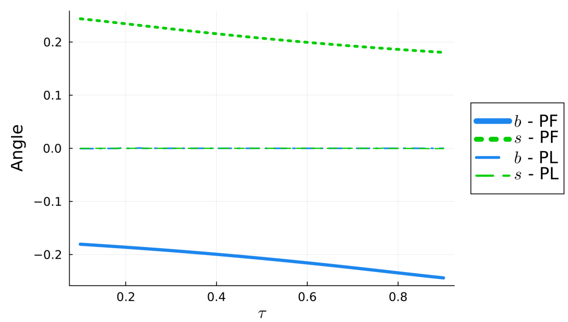
Figure 2.11: Comparative statics with respect to τ (unrestricted term search)



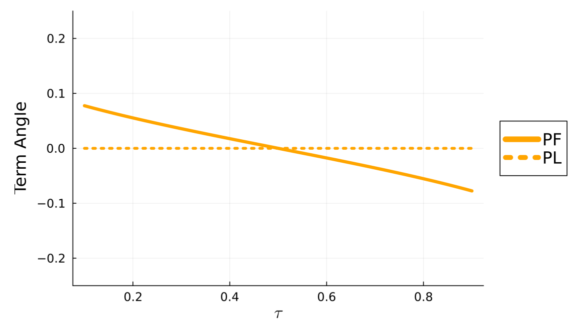
(a) Equilibrium search radius



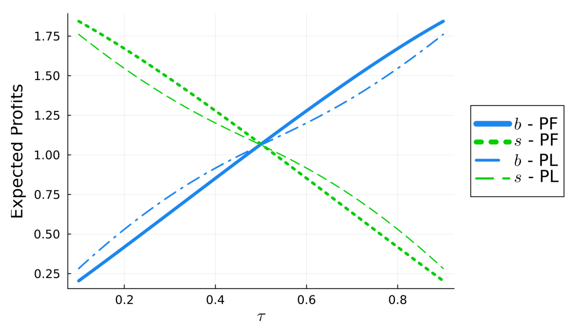
(b) Chosen term radius



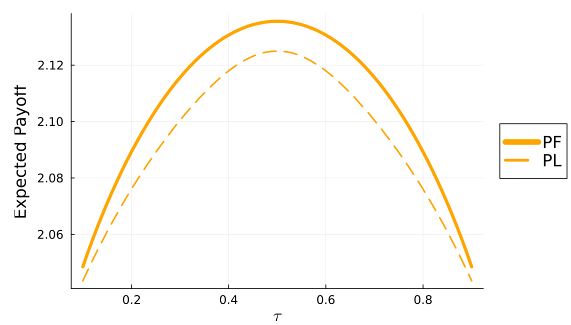
(c) Equilibrium search angle



(d) Chosen term angle



(e) Expected profits



(f) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\alpha = 2$, $\gamma_b = \gamma_s = 1$, $\gamma_a = 5$, $\pi_b = 2$, and $\pi_s = 1$.

As with the endogenous search angle game in 2.4.4, the stronger bargainer generally prefers the price-last game. However, the combination of the firms' angle and radius choices in this setting imply that the expected joint profits are nearly equal across the two games (with the price-first game generating slightly higher surplus). This results in greater investment in the joint term in the price-first game than for the price-last game, particularly for intermediate values of τ, and a term angle that is biased toward the stronger bargainer in the price-first game.

Note that the outcomes for the contracting process for different values of γ_a in the unrestricted model are almost identical to those in the partially contractible term search model presented in Figure 2.9. The main exception is panel (c) in which the equilibrium search angle is constant for both the buyer and the seller in the price-last game for all values of γ_a under this particular set of exogenous parameter values. This in turn implies that the firms are not incentivized to search more intensely as γ_a increases in the price-last game. Because of these similarities, these graphs are presented in Figure B.3 in Appendix B.3.

2.5 Implications

The theoretical framework analyzed in the previous sections yields several surprising results and intuitions relevant to both contract theory and transactional practice.

First, our framework demonstrates that when efficient contract structures are not obvious *a priori*, contract design protocols can play a critical role to incentivize parties to "discover" such terms. Moreover, because bargaining power is not directly contractible, explicitly rewarding a party who discovers such terms through direct price concessions is typically infeasible, especially when that party anticipates being expropriated by a party with appreciable bargaining power. Rather, contract designers must fashion indirect means to encourage search for value enhancing terms. Our framework demonstrates that a seemingly inflexible protocol of cementing price first and then "bartering" non-price terms can be an incentive compatible means for doing so across a dense space of contracting environments. Viewed in this light, the sequential inversion of canonical contract theory is not only intuitive, but it also offers a parsimonious answer to one of the long-standing puzzles from commercial contracting: why it is that the most sophisticated, highest-stakes business contracts so frequently adopt a seemingly-backward "negotiate price first, and other terms later" approach?

In a similar vein, our framework suggests why the "price-first" approach is more typically observed in high-stakes contracts (such as M&A deals and large financings). The dynamics of our model operate only when the payoffs to contract innovation are sufficiently high to justify the parties' search costs. In lower-stakes contracts, by contrast, there are fewer economies of scale to efficient contract design, and accordingly the benefits of incentivizing contractual tailoring are more modest.[23]

Our framework does not merely offer a solution to this longstanding puzzle, however; it also provides insights about other phenomena that observers struggle to understand. One such phenomenon is the incidence of deal failure. Although most M&A practitioners heavily prioritize the certainty of closing, between five and ten percent of publicly announced deals nonetheless fail to close.[24] The failure rate is no doubt higher for preliminary deals that have signed up a term sheet but have yet to reach a definitive agreement (thought to be in the range of 20-40 percent range).[25] While deal failure no doubt has many root causes, our framework suggests an intriguing one: That a signed, price-first deal may ultimately tank because the deal was (mildly) value-destroying *from the very beginning*—and the parties had been relying on subsequent search efforts to tailor the contract language and bridge the valuation gap. In our framework, however, reliance on later search efforts is not a sure thing, even if it is a rational strategy in expectation. Accordingly, deal failure can be an

23. By way of comparison, Gabaix and Landier (2008) make a similar argument to predict that the highest-quality executives will sort into the largest firms, because even modest skill advantages translate into appreciable payoff differences when deployed at scale (and are reflected in higher equilibrium compensation packages as well).

24. *See* Ricks and Lin (2024) (reporting between 4 and 6 percent); Dariush Bahreini et al., "Done deal? Why many large transactions fail to cross the finish line," McKinsey & Co. (2019) (reporting 10 percent).

25. Because term sheets are not publicly disclosed, it is difficult to empirically measure deal failure before a definitive agreement is announced. The 20-40% figure, however, comports with common practitioner estimates.

equilibrium phenomenon.[26]

Relatedly, our framework helps provide insights about why courts have increasingly become attentive to the *pre-contractual* conduct of the parties. Traditionally, an aspiring contractual party enjoyed no legal rights against their counter-party unless and until a fully spelled out contract (a "Type I" agreement, in the parlance of U.S. contract law) had emerged from negotiations.[27] Until that magic moment arrived, both parties were free to walk away from (or even sabotage) the incipient deal. Over the last five decades, however, courts have progressively warmed to the theory that, even when only a preliminary agreement is in place with price and only a few central terms (a "Type II" agreement), the parties begin to bear at least some exposure should they walk away.[28] In particular, a party who fails to negotiate in "good faith" may be found to have breached a preliminary agreement, and thereby subjected to damages claims. Our analysis suggests an economic rationale for this form of liability: to the extent that the parties' endogenous search terms are at least partially contractible, their incentives to search for (and produce) payoff-enhancing terms may be further augmented.[29]

Our results also may bear directly on the long-simmering debate about the value transactional lawyers contribute to deals. Some commentators have suggested that deal lawyers represent little more than transactional deadwood, "churning" out contractual provisions that do little more than lard up billable hours.[30] As evidence, they note that the disclosure of

26. On this note, our framework may also provide intuitions about the use of termination fees within *preliminary* (as opposed to definitive) agreements. Our model predicts that such fees can play a helpful role in incentivizing the discovery of value-enhancing non-price by the lowest cost searcher.

27. *See Teachers' Insurance*, *supra* note 6.

28. *Id.*

29. *See* SIGA v. PharmaThene, 67 A.3d 330 (Del. 2013) (awarding expectation damages for breaching a Type II agreement. In a related vein, even prior to cases like *SIGA*, the emergent "promissory estoppel" doctrine may have also served to incentivize efficiency-enhancing contract design in preliminary negotiations. Hoffman v. Red Owl Stores, Inc., 26 Wis. 2d 683 (Wis. 1965).

30. *See, e.g.*, Anderson (2020) and Anderson and Manns (2013, 2017a, 2017b). Other work finds evidence of perverse bargaining outcomes arising from poorly set incentives, misaligned in ways other than the opportunistic churning of billable hours. *See, e.g.*, Clayton (2023) and Gulati and Scott (2012a).

the definitive merger agreement's non-price terms do not have significant effects on markets in comparison to the initial announcement of the merger some days earlier.[31] Our analysis provides an accounting not only for the value of good lawyering,[32] but also why it wouldn't be manifest in announcement-day returns: Equilibrium expectations. In our model, all parties (including the investing public) will know that a merger announcement was the product of a multi-stage equilibrium, whereby efficient terms were discovered and embraced (even if not fully disclosed alongside the bare-bones pricing terms). Since markets can price those expectations in immediately on announcement, we should not expect systematic directional returns when traders "update" their knowledge set by seeing additional granular details. To the contrary, our model *does* predict that skilled transactional attorneys bring considerable value to a transaction, a trend that is evidently born out in market data: Evidence from recent years shows that top M&A lawyers now routinely out-earn top bankers and other financial professionals in hourly compensation rates.[33]

Finally, our framework may provide important insights about the differences between the majority of deals (where pricing is fixed up front) and the small minority of "auction" deals where pricing occurs last. Very few merger deals take place through a full-blown auction with multiple bidders; indeed, even for seller-initiated processes that invite interested prospective bidders early on, it is overwhelmingly common for a seller to "go exclusive" with a preferred bidder and negotiate one on one.[34] Nevertheless, auctions do occur from time to time, and our framework offers insights for assessing them against negotiated deals. Most notably, auctions mechanically vest the bargaining power with the seller, since buyer competition

31. *See, e.g.*, Anderson and Manns (2017b).

32. In this respect, our project is consistent with the side of the debate finding evidence that transactional lawyers add value. *See, e.g.*, A. Badawi et al. (2023), Coates (2016a), Gilson (1984), and Jennejohn et al. (2022)

33. *See, e.g., Wall Street Journal*, Rock-Star Law Firms Are Billing Up to $2,500 per Hour. Clients Are Indignant (Oct. 4, 2024); *Wall Street Journal*, On Wall Street, Lawyers Make More Than Bankers Now (June 22, 2023)

34. *See* Liu et al. (2022).

typically means the target will reap most of the deal surplus. Moreover, when competing bids are collected, bidders are heavily encouraged to bid on a deal that already has the non-price terms cemented, so that the seller can make apples-to-apples comparisons among buyers. Thus, auction deals are one of the few circumstances in mergers and acquisitions practice where pricing is set last. Our framework suggests that, ceteris paribus, auctions will tend to be preferred in situations where the seller (and not the bidders) is best situated to search for and discover non-price terms. In such a situation bargaining power and search skill are aligned, and our framework predicts optimal contrating would ensue.

2.6 Conclusion

In this article, we have presented an analytic framework that combines a bargaining model and a search game over innovative contractual provisions to reconcile a longstanding puzzle in contract design: the counterintuitive practice in complex transactions of cementing core price terms before negotiating other (non-price) terms. Our framework delivers a robust and tractable set of intuitions about when fixing price before other terms optimally incentivizes non-contractible investments by the contracting parties in contract design.

We are optimistic that our efforts here will serve as a metaphorical "term sheet" upon which future researchers might build to investigate how contractual design, process, and structure can efficiently interact. By modeling firms' investment decisions in the contract construction process, we allow for extensions to the case where firms can exit the negotiation process after discovering new terms. Given the empirical tractability of our model, this enables researchers to evaluate the impact of reliance and expectation measures of damages. More broadly, this model can be used to more accurately estimate the value of contract terms in real-world contracts even in the absence of explicit price renegotiation.

REFERENCES

Adelson, P., Jennejohn, M., Nyarko, J., & Talley, E. L. (2024). Introducing a new corpus of definitive m&a agreements, 2000-2020. *Columbia Law and Economics Working Paper Forthcoming, European Corporate Governance Institute-Law Working Paper Forthcoming.*

Aguirregabiria, V., & Magesan, A. (2020). Identification and estimation of dynamic games when players' beliefs are not in equilibrium. *The Review of Economic Studies*, *87*(2), 582–625.

Anderson, R. (2020). Path dependence, information, and contracting in business law and economics. *Wis. L. Rev.*

Anderson, R., & Manns, J. (2013). The merger agreement myth. *Cornell L. Rev.*

Anderson, R., & Manns, J. (2017a). Engineering greater efficiency in mergers and acquisitions. *The Business Lawyer.*

Anderson, R., & Manns, J. (2017b). The inefficient evolution of merger agreements. *Geo. Wash. L. Rev.*

Backus, M., Blake, T., Larsen, B., & Tadelis, S. (2020). Sequential bargaining in the field: Evidence from millions of online bargaining interactions. *The Quarterly Journal of Economics*, *135*(3), 1319–1361.

Badawi, A., de Fontenay, E., & Nyarko, J. (2023). The value of m&a drafting. *Duke Law School Public Law & Legal Theory Series No. 2023-14.*

Badawi, A. B., & de Fontenay, E. (2019). Is there a first-drafter advantage in m&a? *California Law Review*, *107*(4), 1119–1172.

Bajari, P., Cen, Z., Chernozhukov, V., Manukonda, M., Vijaykumar, S., Wang, J., Huerta, R., Li, J., Leng, L., Monokroussos, G., et al. (2023). Hedonic prices and quality adjusted price indices powered by ai. *arXiv preprint arXiv:2305.00044.*

Basak, D., & Khan, U. (2024). Aspiration and bargaining. *Available at SSRN 4799699.*

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841–890.

Berry, S. T., & Haile, P. A. (2021). Foundations of demand estimation. In *Handbook of industrial organization* (pp. 1–62, Vol. 4). Elsevier.

Bolton, P., & Dewatripont, M. (2004). *Contract theory.* MIT press.

Brewster, B. (1862). Portfolio. *Yale Literary Journal*, *416*, 202.

Bronnenberg, B. J., Kim, J. B., & Mela, C. F. (2016). Zooming in on choice: How do consumers search for cameras online? *Marketing Science*, *35*(5), 693–712.

Bulow, J., & Klemperer, P. (1996). Auctions versus negotiations. *American Economic Review*, *86*(1), 180–94.

Bulow, J., & Roberts, J. (1989). The simple economics of optimal auctions. *Journal of political economy*, *97*(5), 1060–1090.

Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., & Robins, J. M. (2022). Locally robust semiparametric estimation. *Econometrica*, *90*(4), 1501–1535.

Cho, S., & Rust, J. (2010). The flat rental puzzle. *The Review of Economic Studies*, *77*(2), 560–594.

Choi, A., & Triantis, G. (2008). Completing contracts in the shadow of costly verification. *The Journal of Legal Studies*, *37*(2), 503–534.

Choi, A., & Triantis, G. (2012). The effect of bargaining power on contract design. *Virginia Law Review*, 1665–1743.

Choi, S. J., Gulati, M., Jennejohn, M., & Scott, R. E. (2022). Contract production in m&a markets. *U. Pa. L. Rev.*, *171*, 1881.

Choi, S. J., Gulati, M., & Scott, R. E. (2021). Investigating the contract production process. *Capital Markets Law Journal*, *16*(4), 414–431.

Clayton, W. (2023). High-end bargaining problems. *Vand. L. Rev.*

Coates, J. (2016a). Why have m&a contracts grown? evidence from twenty years of deals. *Harvard Law School John M. Olin Center Discussion Paper No. 889*.

Coates, J. (2016b). Why have m&a contracts grown? evidence from twenty years of deals. *Harvard Law School John M. Olin Center Discussion Paper No. 889*.

Cogley, T., Colacito, R., & Sargent, T. J. (2007). Benefits from us monetary policy experimentation in the days of samuelson and solow and lucas. *Journal of Money, Credit and Banking*, *39*, 67–99.

Compiani, G., Lewis, G., Peng, S., & Wang, P. (2022). *Online search and product rankings: A double logit approach* (tech. rep.). Working paper.

Compiani, G., Morozov, I., & Seiler, S. (2023). Demand estimation with text and image data.

Crawford, V. P. (1982). A theory of disagreement in bargaining. *Econometrica: Journal of the Econometric Society*, 607–637.

Davis, A. M., Katok, E., & Kwasnica, A. M. (2011). Do auctioneers pick optimal reserve prices? *Management Science*, *57*(1), 177–192.

de Fontenay, E. (2015). Law firm selection and the value of transactional lawyering. *Journal of Corporation Law*, *41*, 393.

de Fontenay, E. (2017). Market information and the elite law firm. *Duke Law School Public Law & Legal Theory Series*, (2017-32).

Doraszelski, U., Lewis, G., & Pakes, A. (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review*, *108*(3), 565–615.

Dunn, A., Gottlieb, J. D., Shapiro, A. H., Sonnenstuhl, D. J., & Tebaldi, P. (2024). A denial a day keeps the doctor away. *The Quarterly Journal of Economics*, *139*(1), 187–233.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, *70*(5), 1741–1779.

Engelbrecht-Wiggans, R. (1987). On optimal reservation prices in auctions. *Management Science*, *33*(6), 763–770.

Erdem, T., & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, *15*(1), 1–20.

Farrell, M. H., Liang, T., & Misra, S. (2020). Deep learning for individual heterogeneity: An automatic inference framework. *arXiv preprint arXiv:2010.14694*.

Fong, J., Manchanda, P., & Song, Y. (2023). How effective is suggested pricing?: Experimental evidence from an e-commerce platform. *Experimental Evidence from an E-Commerce Platform (November 10, 2023)*.

Foroughifar, M. (2023). *The challenges of deploying an algorithmic pricing tool: Evidence from airbnb* [Doctoral dissertation, University of Toronto (Canada)].

Foster, L., Haltiwanger, J., & Syverson, C. (2016). The slow growth of new plants: Learning about demand? *Economica*, *83*(329), 91–129.

Fox, J. T., il Kim, K., Ryan, S. P., & Bajari, P. (2012). The random coefficients logit model is identified. *Journal of Econometrics*, *166*(2), 204–212.

Freund, J. (1975). *Anatomy of the merger: Strategies and techniques for negotiating corporate acquisitions*. Law Journal Press.

Freyberger, J., & Larsen, B. J. (2022). Identification in ascending auctions, with an application to digital rights management. *Quantitative Economics*, *13*(2), 505–543.

Gabaix, X., & Landier, A. (2008). Why has executive pay increased so much. *The Quarterly Journal of Economics*, 49–100.

Gilson, R. (1984). Value creation by business lawyers: Legal skills and asset pricing. *Yale L. J.*

Gomes, R. (2014). Optimal auction design in two-sided markets. *The RAND Journal of Economics*, *45*(2), 248–272.

Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of political economy*, *94*(4), 691–719.

Gulati, M., & Scott, R. (2012a). *The three and a half minute transaction: Boilerplate and the limits of contract design*. University of Chicago Press.

Gulati, M., & Scott, R. (2012b). *The three and a half minute transaction: Boilerplate and the limits of contract design*. Univ. of Chicago Press.

Haggag, K., McManus, B., & Paci, G. (2017). Learning by driving: Productivity improvements by new york city taxi drivers. *American Economic Journal: Applied Economics*, *9*(1), 70–95.

Hasker, K., & Sickles, R. (2010). Ebay in the economic literature: Analysis of an auction marketplace. *Review of Industrial Organization*, *37*(1), 3–42.

Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica: journal of the econometric society*, 679–694.

Herrero, M. J. (1989). The nash program: Non-convex bargaining problems. *Journal of Economic Theory*, *49*(2), 266–277.

Hitsch, G. J. (2006). An empirical model of optimal dynamic product launch and exit under demand uncertainty. *Marketing Science*, *25*(1), 25–50.

Hodgson, C., & Lewis, G. (2023). *You can lead a horse to water: Spatial learning and path dependence in consumer search* (tech. rep.). National Bureau of Economic Research.

Huang, Y., Ellickson, P. B., & Lovett, M. J. (2020). Learning to set prices. *Available at SSRN 3267701*.

Hwang, C. (2016). Unbundled bargains: Multi-agreement dealmaking in complex mergers and acquisitions. *University of Pennsylvania Law Review*.

Hwang, C. (2018). Deal momentum. *UCLA Law Review*.

Hwang, C., & Jennejohn, M. (2018). Deal structure. *Northwestern University Law Review*.

Ichimura, H., & Newey, W. K. (2022). The influence function of semiparametric estimators. *Quantitative Economics*, *13*(1), 29–61.

Jennejohn, M. (2018). The architecture of contract innovation. *Boston College Law Review*.

Jennejohn, M. (2020). Transformation cost engineering. *Wisconsin Law Review*.

Jennejohn, M., Nyarko, J., & Talley, E. (2022). Contractual evolution. *U. Chi. L. Rev.*, *89*, 901.

Jullien, B., Pavan, A., & Rysman, M. (2021). Two-sided markets, pricing, and network effects. In *Handbook of industrial organization* (pp. 485–592, Vol. 4). Elsevier.

Kafka, F. (1915). *Metomorphosis*. Die weißen Blätter.

Katkar, R., & Reiley, D. H. (2007). Public versus secret reserve prices in ebay auctions: Results from a pokémon field experiment. *The BE Journal of Economic Analysis & Policy*, *6*(2), 0000102202153806371442.

Keller, G., & Rady, S. (1999). Optimal experimentation in a changing environment. *The Review of Economic Studies*, *66*(3), 475–507.

Kim, Y. (2020). *Customer retention under imperfect information* [Doctoral dissertation, The University of Chicago].

Klein, B., Crawford, R. G., & Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The journal of Law and Economics*, *21*(2), 297–326.

Klein, B., Lerner, A. V., Murphy, K. M., & Plache, L. L. (2005). Competition in two-sided markets: The antitrust economics of payment card interchange fees. *Antitrust LJ*, *73*, 571.

Lashkaripour, A., & Lugovskyy, V. (2023). Profits, scale economies, and the gains from trade and industrial policy. *American Economic Review*, *113*(10), 2759–2808.

Levin, D., & Smith, J. L. (1996). Optimal reservation prices in auctions. *The Economic Journal*, *106*(438), 1271–1283.

Li, X. (2005). A note on expected rent in auction theory. *Operations Research Letters*, *33*(5), 531–534.

List, J. A. (2003). Does market experience eliminate market anomalies? *The Quarterly Journal of Economics*, *118*(1), 41–71.

List, J. A. (2004). Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica*, *72*(2), 615–625.

List, J. A. (2011). Does market experience eliminate market anomalies? the case of exogenous market experience. *American Economic Review*, *101*(3), 313–317.

Liu, T., Officer, M. S., & Tu, D. (2022). Negotiation, auction, or negotiauction?! evidence from the field. *Evidence from the Field (October 1, 2022)*.

Lu, J. (2019). Bayesian identification: A theory for state-dependent utilities. *American Economic Review*, *109*(9), 3192–3228.

Marra, M. (2019). *Pricing and fees in auction platforms with two-sided entry* (tech. rep.). Sciences Po Departement of Economics.

McFadden, D. (1972). Conditional logit analysis of qualitative choice behavior.

Mela, C. F., Roos, J. M., & Sousa, T. (2023). Advertiser learning in direct advertising markets. *arXiv preprint arXiv:2307.07015*.

Muthoo, A. (1992). Revocable commitment and sequential bargaining. *The Economic Journal*, *102*(411), 378–387.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, *6*(1), 58–73.

Nash, J. (1950). The bargaining problem. *Econometrica: Journal of the econometric society*, 155–162.

Nash, J. (1953). Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, 128–140.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science*, *31*(3), 521–543.

O'Donoghue, T., & Sprenger, C. (2018). Reference-dependent preferences. In *Handbook of behavioral economics: Applications and foundations 1* (pp. 1–77, Vol. 1). Elsevier.

Ostrovsky, M., & Schwarz, M. (2016). Reserve prices in internet advertising auctions: A field experiment. typescript.

Platt, B. C. (2017). Inferring ascending auction participation from observed bidders. *International Journal of Industrial Organization*, *54*, 65–88.

Quint, D. (2017). Common values and low reserve prices. *The Journal of Industrial Economics*, *65*(2), 363–396.

Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. In *The economics of the internet and e-commerce*. Emerald Group Publishing Limited.

Ricks, M., & Lin, D. (2024). How deals die. *Available at SSRN*.

Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, *1*(4), 990–1029.

Rothschild, M. (1974). A two-armed bandit theory of market pricing. *Journal of Economic Theory*, *9*(2), 185–202.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, 97–109.

Schelling, T. C. (1956). An essay on bargaining. *The American Economic Review*, *46*(3), 281–306.

Serrano, R., & Shimomura, K.-I. (1998). Beyond nash bargaining theory: The nash set. *journal of economic theory*, *83*(2), 286–307.

Simonsohn, U. (2010). Ebay's crowded evenings: Competition neglect in market entry decisions. *Management Science*, *56*(7), 1060–1073.

Strulov-Shlain, A. (2021). More than a penny's worth: Left-digit bias and firm pricing. *Chicago Booth Research Paper*, *19–22*.

Tadelis, S., Hooton, C., Manjeer, U., Deisenroth, D., Wernerfelt, N., Dadson, N., & Greenbaum, L. (2023). *Learning, sophistication, and the returns to advertising: Implications for differences in firm performance* (tech. rep.). National Bureau of Economic Research.

Talley, E. (2009). On uncertainty, ambiguity, and contractual conditions. *Del. J. Corp. Law*, *34*, 755–812.

Tong, L. C., Ye, K. J., Asai, K., Ertac, S., List, J. A., Nusbaum, H. C., & Hortaçsu, A. (2016). Trading experience modulates anterior insula to reduce the endowment effect. *Proceedings of the National Academy of Sciences*, *113*(33), 9238–9243.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, *16*(1), 8–37.

Wang, Z., Chang, Y., Yang, N., & Garcia, A. (2024). Retail investment under aggregate fluctuations. *Available at SSRN*.

Wang, Z., & Yang, N. (2024). Identification of structural learning models. *Available at SSRN*.

Williamson, O. E. (1983). Credible commitments: Using hostages to support exchange. *The American economic review*, *73*(4), 519–540.

Wu, R., Huang, Y., & Li, N. (2023). Platform information design and competitive price targeting. *Available at SSRN*.

Xu, B., Deng, Y., & Mela, C. (2022). A scalable recommendation engine for new users and items. *arXiv preprint arXiv:2209.06128*.

Zhou, L. (1997). The nash bargaining theory with non-convex problems. *Econometrica: Journal of the Econometric Society*, 681–685.

APPENDIX A

APPENDIX FOR CHAPTER 1

A.1 Summary statistics and additional descriptive evidence

Table A.1 shows selected summary statistics for the sample used in the data. To limit the effect of prediction error in estimated item values (described in more detail in 1.5.1), I drop all items with a standardized reserve price and standardized revenue greater than the 99th quantile of the respective variables.

Table A.1: Summary statistics

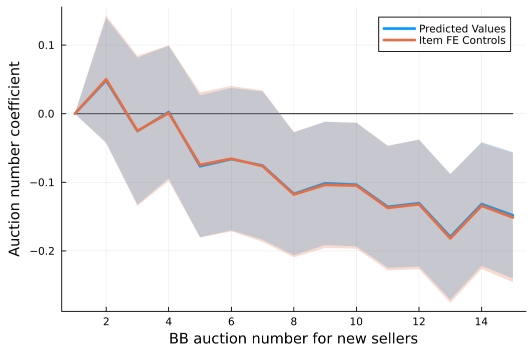
Variable	Mean	Std. Dev.	Minimum	Maximum
Minimum Bid	13.85	48.43	0.01	10,000
Reserve Price	15.39	139.4	0	68,000
Revenue	15.66	83.5	0	68,000
# Bidders	2.62	2.93	0	36
Sell	0.55	0.5	0	1
Fees	1.05	1.97	0	867.12

Notes: These statistics are from the 1,038,383 items included in the analysis data. The top 1% of the items (by standardized effective reserve price) have been removed from the analysis data.

Figure A.1 estimates the same equation as Figure 1.2, but the dependent variables are the inverse hyperbolic sine (IHS) of effective reserve price and net revenue. I also restrict the sample to items where the seller has listed at least one other item with the same description, and run the regression with predicted item values and item fixed effects to compare the resulting estimates. The trends are quite similar whether using predicted item values or fixed effects, which suggests the predicted item values capture economically meaningful information. They are also similar to the trends in Figure 1.2, though with the caveat that the greater magnitude of the coefficients in panel (b) may be in part driven by re-listed items that were not sold the first time.[1]

1. One limitation of the dataset is that I cannot observe seller inventories, so I cannot see how many of the identical items are true duplicates as opposed to relisting.

Figure A.1: Time trends of variables of interest with auction experience



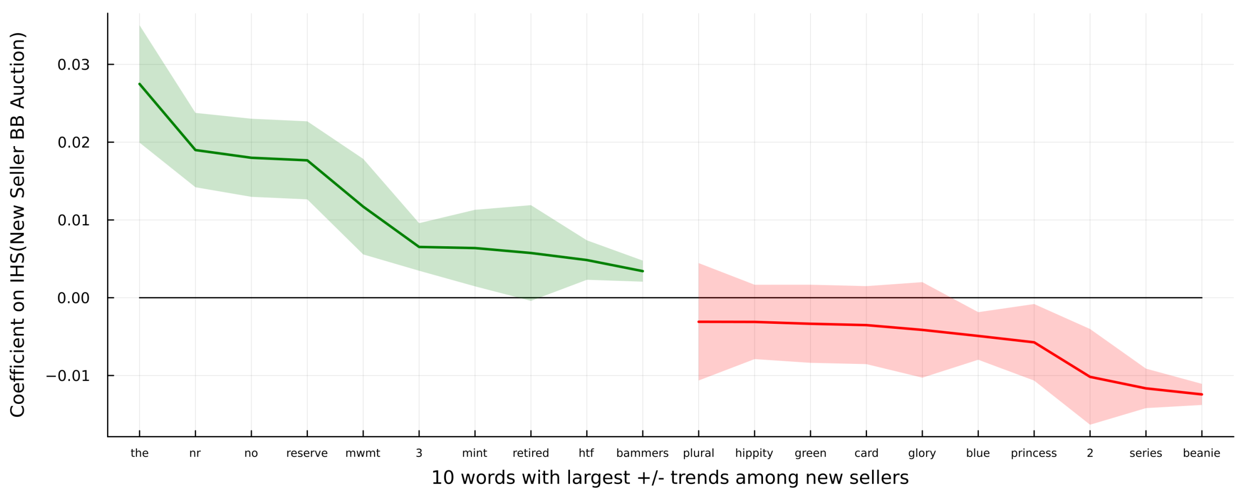
(a) IHS(Effective reserve)

(b) IHS(Net revenue)

Notes: These regressions pool 1,639 new sellers' first 15 auctions with all auctions by 5,165 experienced sellers (defined as those with $\geq$47 auctions at the start of the data, which is the 75th percentile of initial experience). The sample is limited to sellers with at least 15 auctions in the data. The results are similar when using different values of T_{New}.

Figure A.2 shows the words that most increased and decreased in their usage by new sellers in their first 15 auctions. While the trends are largely small, the words that most increased in frequency include "nr" (short for "no reserve") and "no" "reserve". This is consistent with sellers becoming more aware of the possible effect of their pricing decisions on bidder entry.

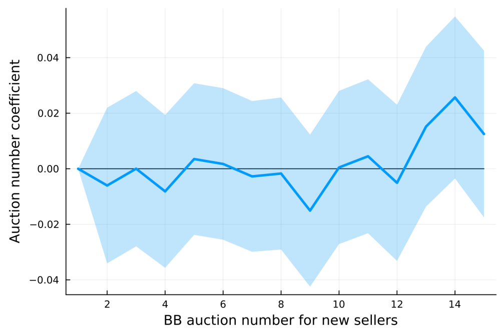
Figure A.2: Trends in the frequency of words in item descriptions



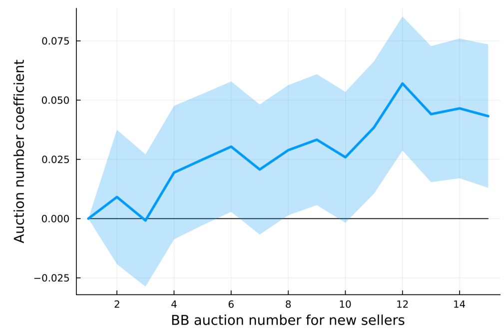
Notes: These are the 10 most positive and 10 most negative (by absolute value) coefficients when regressing $\mathbb{1}$[item contains word] on the inverse hyperbolic sine (IHS) of new sellers' auction number (among the first 15 auctions of sellers who have at least 15 auctions in the data or sellers with >75th percentile of experience at the start of the data), along with predicted item value, IHS(feedback count), feedback percentage, and seller and month fixed effects.

Figure A.3 shows additional trends in non-price variables among sellers with at least 15 auctions. New sellers show some trends in the timing of an auction (panels (b) and (c)), where they favor shorter auctions that end on weekends. As shown in panel (d), new sellers also become less likely to feature items. To ensure that estimation is computationally tractable, and since these trends are generally smaller in magnitude relative to the baseline averages of each variable, I focus attention on the choice of reserve prices.

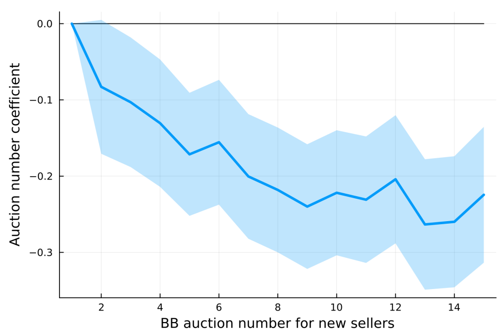
Figure A.3: Trends in non-price variables among new sellers



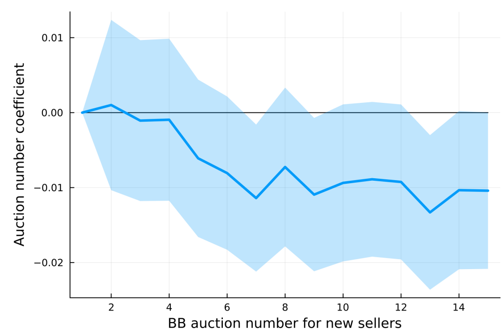
(a) End time in evening



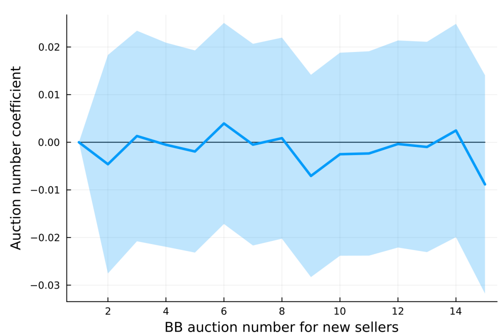
(b) End time on weekend



(c) Auction length (days)



(d) Featured item

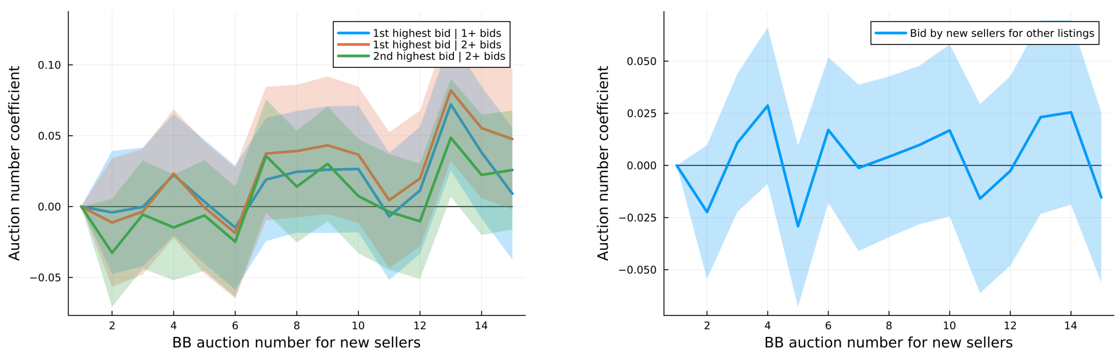


(e) Has secret reserve

Notes: These figures display coefficients of the first 15 auctions of new sellers in a regression of the various outcomes on predicted item values, seller feedback scores, inverse hyperbolic sine (IHS) of auction experience, and seller and month fixed effects. The regression is on new sellers with at least 15 auctions in the data and experienced sellers (defined as sellers with >75th percentile of experience at the start of the data).

Figure A.4 estimates similar regressions for bids within auctions, though with the number of observed bids in each auction additional control. Panel (a) shows the trends in first and second highest bids faced by new sellers, since these are the only bids known to reflect the first and second highest values in an ascending IPV setting. Panel (b) shows the trend in bids placed by new sellers on other listings on or before their kth listing. The trend lines for both figures are relatively noisy, with a small but statistically insignificant trend upward in panel (a). Note that panel (a) conditions on the number of observed bidders, which is a noisy measure of the true number of bidders that is unobserved due to eBay's increasing minimum bid rule.

Figure A.4: Time trends of bids with auction experience



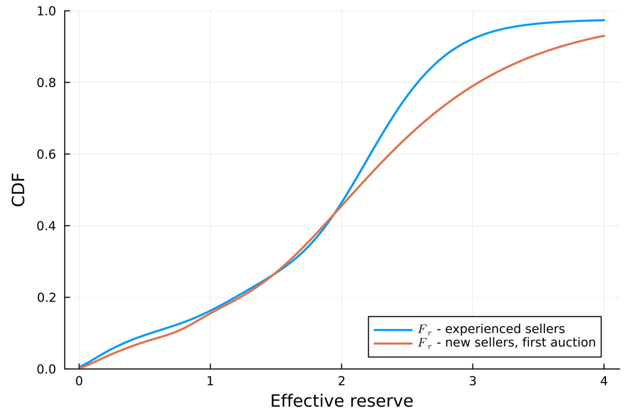
(a) Trends in bids faced by new sellers

(b) Trends in bids placed by new sellers

Notes: These regressions pool bids from 1,639 new sellers' first 15 auctions with bids from all auctions by 5,165 experienced sellers (defined as those with $\geq$47 auctions at the start of the data, which is the 75th percentile of initial experience). The sample is limited to sellers with at least 15 auctions in the data. Panel (a) restricts the sample to all auctions with at least 1 or 2 bids (as specified in the legend), and panel (b) examines bids placed by new sellers before they list their kth item. Both panels control for seller fixed effects as in section 1.3.2, as well as fixed effects for the number of other observed bids in each auction.

Figure A.5 shows the estimated CDFs of the reserve prices among experienced vs. inexperienced sellers in their first auction, for all inexperienced sellers that list at least 15 items. The distributions largely align for approximately 50% of listings, but diverge for higher-reserve listings.

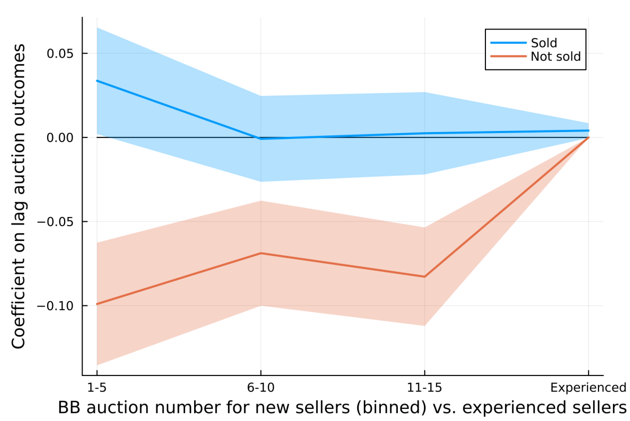
Figure A.5: Distributions of reserve prices among new sellers in their first auction vs. experienced sellers



Notes: These are the estimated CDFs of the standardized reserve prices for new sellers in their first auction (among new sellers who have at least 15 auctions in the data) and experienced sellers (those with >75th percentile of experience at the start of the data). I use 5-component log-Normal mixture models to smooth the estimated CDFs in both cases.

Figure A.6 plots coefficients from a regression of standardized effective reserve prices on interactions between seller experience indicators and lag auction outcome indicators. Note that only the first coefficient for the "lag sale" indicator is positive and statistically significant (though it is close), while new sellers failing to sell an item in the previous auction is more strongly (and negatively) correlated with effective reserve prices.

Figure A.6: Coefficients for regressing lag auction sale outcomes on current effective reserve price, by experience bin



Notes: The plotted coefficients are from regressing standardized effective reserve prices on seller experience indicators interacted with lag auction sale indicators, along with 95% confidence intervals for the estimated coefficients. This regression controls for lag effective reserve price, bidder feedback, feedback score, and month fixed effects.

A.2 Testing for common values

One potential concern in this setting is that Beanie Babies have a common value component that is unobservable to the bidders, opening the possibility of a winner's curse that affects bidding strategies. I regress log standardized bids on the number of bidders observed in the auction, which is a noisy but biased measure of the total number of bidders who have entered the auction. I examine both the total number of bidders and the number of serious bidders, or those whose bids exceed the reserve price, since some bids may be mechanically lower due to a low minimum bid allowing low-value bidders to submit bids that they would not ordinarily be able to submit. I also control for the standardized minimum bid and other covariates that may influence bidder arrival, namely the log item value, the inverse hyperbolic sine of the feedback count, and sellers' feedback percentage score, along with month fixed effects. These regressions are not causal, as the number of bidders is likely correlated with other drivers of bidder values.

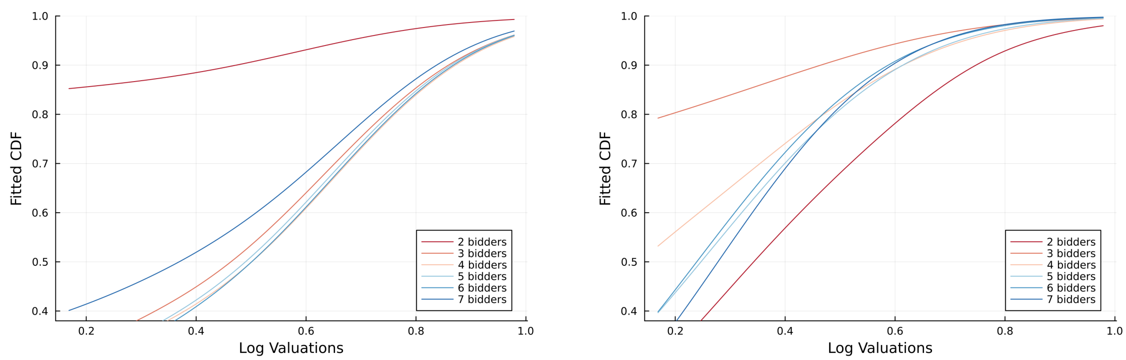
Table A.2: Regression of log standardized bid values on number of bidders

Variable	(a)	(b)
Number of bidders	-0.008	-
	(2e-4)	-
Number of serious bidders	-	0.004
	-	(1e-4)
Adjusted R^2	0.159	0.159

Notes: These regressions also control for IHS(feedback), feedback percentage, log mean item value, standardized minimum bid, month fixed effects, though these coefficients are omitted from the table. Specifications (a) and (b) differ by whether the number of bidders or the number of serious bidders (defined as bidders whose bid is equal to or greater than the reserve price) is used.

Table A.2 shows that while there is some statistically significant correlation between the number of bidders and the log standardized bid values, its magnitude is small. I also use the two highest bids, binning by the number of either observed or serious bidders, and estimate the corresponding distributions; under common values, these distributions should be stochastically ordered so higher-bidder auctions have lower value distributions. Figure A.7 presents the estimated distributions where I bin distributions for auctions by number of serious bidders (those with a bid above the secret reserve price, to discard unusually low bids where there is a low minimum bid) and the total number of observed bidders. In panel (a), the distributions are quite similar for all but the auctions with only 2 serious bidders. In panel (b), the distribution estimated on auctions with only 2 bidders is to the right of the remaining distributions, but the others have no clear stochastic ordering that is indicative of common values. While not conclusive, this is suggestive evidence that independent private values is a useful modeling assumption in this setting.

Figure A.7: Estimated distributions of bidder valuations by the number of observed bidders per auction



(a) Binned by number of serious bidders (b) Binned by number of observed bidders

Notes: I use 5-component Gaussian mixture models to estimate the distributions for each binned number of bidders. Serious bidders are defined as those with bids above the reserve price. These distributions are plotted between the 5th and 95th percentile of the second-highest bids in the data.

A.3 Proofs

Proof of Proposition 1

This closely follows the corollary in Marra (2019), though with the caveat that r is observed in this setting and thus directly impacts the Poisson mean Λ. First, denote

$$\pi_B(r \mid n, c) = \frac{1}{n} \cdot \mathbb{E}\Big[v_{n:n} - (1 + c_B^R) \max\{v_{(n-1):n}, r^*\} \;\Big|\; v_{n:n} \geq r^B\Big] \cdot (1 - F_B(r^B)^n)$$

Since F_B satisfies the strict monotone hazard rate property, Li (2005) implies the inequality $\mathbb{E}[v_{(n+1):(n+1)} - v_{n:(n+1)}] < \mathbb{E}[v_{n:n} - v_{(n-1):n}]$; this holds when conditioning on r since r is set before the auction and does not vary with the number of bidders n that arrive. Additionally, $\frac{1}{n}(1 - F_B(r)^n) \geq \frac{1}{n+1}(1 - F_B(r)^{n+1})$,[2] so $\pi_B(r \mid n)$ is decreasing in n.

We temporarily abuse notation to write the probability mass at n given Λ as $p_n(\Lambda)$. Since arrival is Poisson, $p_n(\Lambda')$ first-order stochastically dominates $p_n(\Lambda)$ for $\Lambda' > \Lambda$. Increasing Λ therefore decreases $\sum_{n=1}^{\mathbf{N}_B-1} \pi_B(r \mid n, c)p_n(\Lambda)$ since $\pi_B(r \mid n)$ is monotonically decreasing in n. Thus, there exists a unique Λ that solves the zero profit condition.

Similarly, $\pi_B(r \mid n, c)$ is decreasing in r, since the measure of the set $\{v_{n:n} \geq r^B\}$, the probability of winning $1 - F_B(r^B)^n$, and the winner's expected surplus are all all decreasing in r. Since higher r corresponds to strictly lower surplus, Λ is strictly decreasing in r.

Proof of Proposition 3

We differentiate the first-order condition with respect to v_0:

$$\frac{\partial^2 \Pi(v_0, r \mid b, c)}{\partial r \partial v_0} = (1 - c_S^R)R_{rr}(r^*(v_0))\frac{\partial r^*(v_0)}{\partial v_0} + K_r(r^*(v_0) \mid b) + v_0 K_{rr}(r^*(v_0))\frac{\partial r^*(v_0)}{\partial v_0}$$

Rearranging, we have

$$\frac{\partial r^*(v_0)}{\partial v_0} = \frac{K_r(r^*(v_0) \mid b)}{-[(1 - c_S^R)R_{rr}(r^*(v_0)) + v_0 K_{rr}(r^*(v_0))]}$$

The denominator is positive at the interior optimum because it is the negative second-order condition of the seller's profit maximization problem. Thus, $\frac{\partial r^*}{\partial v_0}$ when $K_r(\cdot \mid b) > 0$. Further, since the inverse of an increasing function is itself increasing, the virtual type function $\psi(\cdot \mid b, c)$ is monotonic increasing.

2. To see this, first note that $F_B(r) \in [0, 1]$, and the expression is equivalent to showing $(1 + \frac{1}{n})\frac{1-F_B(r)^n}{1-F_B(r)^{n+1}} \geq 1$. Let $x \in [0, 1]$, and note that $1 - x^n = (1 - x)b_n$, where $b_n \equiv \sum_{k=0}^{n-1} x^k$. Then $\frac{1-x^n}{1-x^{n+1}} = \frac{b_n}{b_n+x^n} = \frac{1}{1+\frac{x^n}{b_n}}$, and $\frac{1}{n} \geq \frac{x^n}{b_n}$, since $\sum_{k=0}^{n-1} x^k \geq \sum_{k=0}^{n-1} x^n$.

Lemma 3. Let $p_n(x \mid \delta) = \frac{1}{n!}\exp(-\Lambda(x \mid \delta))\Lambda(x \mid \delta)^n$, where the Poisson mean is $\Lambda(x \mid \delta) = \exp(\delta_1 + \delta_2\rho(x))$ and and $\rho(x^*) = 0$. Then for each $k = 1, 2, ...,$ $\frac{\partial^k}{\partial x^k} p_n(x^* \mid \delta)$ has the form

$$\sum_{\ell=0}^{k} \delta_2^{\ell} \cdot h_{\ell,k}\left(n, \delta_1, \left\{\frac{\partial^t}{\partial x^t}\rho(x^*)\right\}_{t=1}^{k-\ell+1}\right)$$

where each $h_{\ell,k}$ is known.

Proof

We first show that for every $k = 1, 2, ...,$ $\frac{\partial^k}{\partial x^k} p_n(x \mid \delta)$ has the form

$$\sum_{\ell=0}^{k} \delta_2^{\ell} \cdot \tilde{h}_{\ell,k}\left(n, \Lambda(x \mid \delta), \left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)$$

for known $\tilde{h}_{\ell,k}$. Beginning with $k = 1$, note that

$$\begin{aligned}\frac{\partial^1}{\partial x^1} p_n(x \mid \delta) &= \delta_2 \cdot p_n(x \mid \delta) \cdot [n - \Lambda(x \mid \delta)] \cdot \rho'(x) \\ &\equiv \delta_2 \cdot \tilde{h}_{1,1}(n, \Lambda(x \mid \delta), \rho'(x))\end{aligned}$$

since p_n is a known function of Λ. Now suppose $\frac{\partial^k}{\partial x^k} p_n(x \mid \delta)$ has the form above. Then using the shorthand $\tilde{h}_{j,\ell,k}$ to denote the first derivative with respect to the jth argument of

$\tilde{h}_{\ell,k}$, note that by the chain rule,

$$\begin{aligned}
&\frac{\partial}{\partial x}\sum_{\ell=0}^{k}\delta_2^{\ell}\cdot\tilde{h}_{\ell,k}\left(n,\Lambda(x\mid\delta),\left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)\\
&= \qquad \sum_{\ell=0}^{k}\delta_2^{\ell}\cdot\left[\tilde{h}_{2,\ell,k}\left(n,\Lambda(x\mid\delta),\left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)\cdot\Lambda(x\mid\delta)\cdot\rho'(x)\cdot\delta_2\right.\\
&\qquad\qquad \left.+\,\tilde{h}_{3,\ell,k}\left(n,\Lambda(x\mid\delta),\left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)^{\intercal}\left\{\frac{\partial^{t+1}}{\partial x^{t+1}}\rho(x)\right\}_{t=1}^{k-\ell+2}\right]\\
&\equiv \sum_{\ell=0}^{k+1}\delta_2^{\ell}\cdot\tilde{h}_{\ell,k+1}\left(n,\Lambda(x\mid\delta),\left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)
\end{aligned}$$

Evaluating the expression above at x^*, where $\rho(x^*=0)$ and therefore $\Lambda(x^*\mid\delta)=\exp(\delta_1)$, yields the desired result.

Proof of Proposition 5

By assumption (i), restricting attention to all sellers with the same history $\mathcal{H}$ is equivalent to restricting attention to sellers with identical beliefs b.

By Proposition 3, $\psi(\cdot\mid b,c)$ is increasing. Using the seller first-order condition, a change-of-variables can be applied to the reserve price distribution to write it in terms of the virtual type function and the known seller value distribution:

$$\mathbb{P}[r\le x]=\mathbb{P}[v_0\le\psi(x\mid b,c)]=F_S(\psi(x\mid b,c))$$

Further, the selection rule for seller entry implies $\bar{v}(b,c)\ge v_0$. Taken together, the probability that a reserve price is less than or equal to x, conditional on beliefs b and cost vector c, is $F_S(\psi(x\mid b,c))/F_S(\bar{v}(b,c))$. Since by assumption (ii) the entry threshold $\bar{v}(b,c)$ is known, inverting the empirical reserve price distribution $\phi(x\mid\mathcal{H})$ of sellers with history $\mathcal{H}$ yields the virtual type function:

$$\psi(x \mid b, c) = F_S^{-1}\Big[\phi(x \mid \mathcal{H}) \cdot F_S(\bar{v}(b, c))\Big]$$

for all x such that $\psi(x \mid b, c) < \bar{v}(b, c)$. By assumption (iv) this support includes a positive-measure interval including some value x^* for which $\rho(x^*) = 0$; in what follows we restrict attention to this interval.

The virtual type function $\psi(x \mid b, c)$ is proportional to the ratio of $R_r(x \mid b)$ and $K_r(x \mid b)$, and its derivatives are

$$\begin{aligned}\frac{\partial^k}{\partial x^k}\psi(x \mid b, c) &= -(1 - c_S^R)\frac{\partial^k}{\partial x^k}\frac{R_r(x \mid b)}{K_r(x \mid b)} \\ &= -(1 - c_S^R)\sum_{\ell=0}^{k}\binom{k}{\ell}\Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}}R_r(x \mid b)\Big) \cdot \Big(\frac{\partial^\ell}{\partial x^\ell}\big(K_r(x \mid b)\big)^{-1}\Big)\end{aligned}$$

where by Faà di Bruno's formula

$$\frac{\partial^\ell}{\partial x^\ell}\big(K_r(x \mid b)\big)^{-1} = \sum_{t=0}^{\ell}\frac{(-1)^t \cdot t!}{(K_r(x \mid b)^{t+1})}B_{\ell,t}\Big(\frac{\partial}{\partial x}K_r(x \mid b),\ \dots,\ \frac{\partial^{\ell-t+1}}{\partial x^{\ell-t+1}}K_r(x \mid b)\Big)$$

in which $B_{\ell,t}$ are Bell polynomials. In turn, for both functions $G \in \{R, K\}$, we have

$$\frac{\partial^k}{\partial x^k}G(x \mid b) = \sum_{\ell=0}^{k}\binom{k}{\ell}\left(\sum_{n=0}^{\infty}\Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}}G_n(x)\Big) \cdot \Big(\frac{\partial^\ell}{\partial x^\ell}p_n(x \mid b)\Big)\right)$$

Thus, the kth derivative of the virtual type function is a known function of the 0,...,kth derivatives of R_n, K_n, and p_n.

We now turn our attention to the Poisson mass function and its derivatives, which are the only arguments of ξ_k that depend on beliefs b. Expanding $\frac{\partial^k}{\partial x^k}p_n(x \mid b)$ and imposing independence between the marginal beliefs about the two parameters yields

$$\frac{\partial^k}{\partial x^k} p_n(x \mid b) = \int \int \left[\frac{\partial^k}{\partial x^k} p_n(x \mid \delta) \right] b_{\delta_1}(\delta_1) b_{\delta_2}(\delta_2) d\delta_1 d\delta_2$$

Evaluating this at x^* and applying Lemma 3, this can be expanded to yield

$$\sum_{\ell=0}^{k} \mathbb{E}_{b_{\delta_2}}[\delta_2^\ell] \cdot \hat{h}_{\ell,k}(n)$$

where $\hat{h}_{\ell,k}(n) \equiv \int h_{\ell,k}(n, \delta_1, \{\frac{\partial^t}{\partial x^t} \rho(x^*)\}_{t=1}^{k-\ell+1}) b_{\delta_1}(\delta_1) d\delta_1$ is known under the assumption that b_{δ_1} is known. Note that $\mathbb{E}_{b_{\delta_2}}[\delta_2^\ell]$ does not depend on n, so this term can be pulled out of all the infinite sums in which it appears, i.e. for $G \in \{R, K\}$

$$\begin{aligned}
\frac{\partial^k}{\partial x^k} G(x^* \mid b) &= \sum_{\ell=0}^{k} \binom{k}{\ell} \left(\sum_{n=0}^{\infty} \Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}} G_n(x^*) \Big) \cdot \Big(\sum_{t=0}^{\ell} \mathbb{E}_{b_{\delta_2}}[\delta_2^t] \cdot \hat{h}_{t,\ell}(n) \Big) \right) \\
&= \mathbb{E}_{b_{\delta_2}}[\delta_2^k] \left(\sum_{n=0}^{\infty} G_n(x^*) \hat{h}_{k,k}(n) \right) \\
&\qquad + \sum_{\ell=0}^{k-1} \binom{k}{\ell} \left(\sum_{n=0}^{\infty} \Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}} G_n(x^*) \Big) \cdot \Big(\sum_{t=0}^{\ell} \mathbb{E}_{b_{\delta_2}}[\delta_2^t] \cdot \hat{h}_{t,\ell}(n) \Big) \right)
\end{aligned}$$

Note the $k-1$th derivative of the virtual type function is a function of the kth raw moment of b_{δ_2}. By assumption, the $k-1$th derivative of $\psi(x^* \mid b, c)$ is invertible in the coefficient of this first term, yielding identification of the kth raw moment from the $k-1$th derivative of $\psi(x^* \mid b, c)$ and knowledge of lower-order moments. Since b_{δ_2} satisfies the Carleman condition, its moments uniquely characterize the distribution, and b_{δ_2} is identified up to the $\bar{k}$th moment.

A.4 Demand estimation details

I use text data from item descriptions to estimate the average value for each item. Since item descriptions are seller-provided, there is significant variation in how words are spelled, which poses a challenge for tractably estimating item values. To address this, I manually created a crosswalk of individual words to their apparent intended word to decrease the dimensionality of the space item descriptions (e.g., replacing "beaneis" and "babys" with "beanies" and "babies"). I then constructed a dictionary of the 5,368 words that appear at least 10 times in the cleaned item descriptions. I also include indicators for each month in the dataset.

I tested multiple neural network architectures for γ via out-of-sample validation and with built-in dropout layers to find the architecture that achieved the lowest out-of-sample loss using the likelihood derived in A.6. In particular, I use 80% of the sample to train the model, 10% of the sample to test out-of-sample loss during training, and the remaining 10% of the sample for out-of-sample validation after training. I used an early stopping rule to determine the number of epochs with which to train the full model: I use the smallest number of epochs after which the testing loss fails to improve for 10 consecutive epochs. I then select the architecture with the lowest validation loss. The resulting architecture has 9,111,233 parameters; additional information on the various architectures is presented in Table A.3. This table shows that the nonparametric specifications outperform the parametric model in the first line of the table. Larger numbers of parameters generally improve validation loss, though there are diminishing returns to increased complexity. The empirical results of the paper are similar using different architectures.

Table A.3: Neural network architectures and performance

Model	# 1st-Layer Nodes	# Parameters	Train Loss	Test Loss	Val Loss	Epochs	R^2
1	-	5,374	2.6013	2.6165	2.6104	46	0.691
2	512	2,900,321	1.0885	1.2647	1.2680	34	0.984
2	1,024	5,782,881	0.9965	1.2264	1.2374	38	0.992
2	1,536	8,665,441	1.0006	1.2316	1.2752	30	0.989
3	512	3,083,969	1.0509	1.2638	1.2791	47	0.987
3	1,024	6,097,601	0.9144	1.1934	1.2007	43	0.993
3	1,536	9,111,233	0.8590	1.1774	1.1933	46	-

Notes: Model 1 is fully parametric, with a Gaussian distribution for log-values and no hidden nodes (i.e., log-values are modeled as a linear combination of word-specific fixed effects). Models 2 and 3 both use 5-component Gaussian mixture models, with a varying number of nodes in the first hidden layer; both have 4 hidden layers with corresponding dropout of 50%, 40%, 30%, and 20% for each. The second through fourth hidden layers contain 256, 64, and 16 nodes for model 2 and 512, 128, and 32 nodes for model 3. The number of parameters is the total number of trained parameters in each specification. The train, test, and validation loss columns denote the loss of each of the 80%, 10% and 10% samples used in comparing each of the models. The number of epochs is chosen via early stopping, since subsequent training after the listed number of epochs yields no test loss improvement for at least 10 epochs. The R^2 is taken from regressing the fitted item values from each architecture on the architecture with the lowest validation loss.

A.5 Orthogonalization of the likelihood function

This section derives a method to estimate the true parameter ϑ_0 without bias due to estimation error for the nonparametric component γ_0. I denote the log-likelihood as ℓ; its derivation is shown in the following section. All relevant data for this demand-side likelihood is abbreviated as $\mathbf{D}_d$ to differentiate it from the data $\mathbf{D}$ that is used by sellers in updating their beliefs.

Denote the score function for the structural parameters ϑ_0 as

$$g(\vartheta \mid \gamma, \mathbf{D}_d) = \frac{\partial \ell(\vartheta, \gamma \mid \mathbf{D}_d)}{\partial \vartheta}$$

and note that $\mathbb{E}[g(\vartheta_0 \mid \gamma_0, \mathbf{D}_d)] = 0$. To derive a Neyman orthogonal score $g^*(\vartheta \mid \gamma, \mathbf{D}_d)$ for the average score $\mathbb{E}[g(\vartheta_0 \mid \gamma_0, \mathbf{D}_d)]$, Ichimura and Newey (2022) provide a method for finding a candidate first-stage influence function that will be added to the original score. I follow the steps in their Proposition 1 to show how this applies to a setting with both low and high dimensional parameters, where we orthogonalize with respect to the high dimensional parameter.

By way of notation, γ_0 as the true high-dimensional parameter under the true distribution function, and γ_τ is the perturbation in the direction of some alternative $\tilde{\gamma}$ (i.e., $\gamma_\tau = (1-\tau)\gamma_0 + \tau\tilde{\gamma}$). The Gateaux derivative $\frac{\partial}{\partial\tau}$ is the derivative with respect to τ from above evaluated at zero ($\tau \downarrow 0$). I assume that $\mathbb{E}[\frac{\partial}{\partial a}\frac{\partial}{\partial a}\ell(\vartheta, \gamma(X) + a \mid \mathbf{D}_d) \mid X = x] = 0$, which implies $\mathbb{E}[b(X)\ell(\vartheta, \gamma(X) + a \mid \mathbf{D}_d) \mid X = x] = 0$ for all b.

In this setting, Assumptions 1 and 2 of Ichimura and Newey (2022) are that there exist $\alpha_1(\vartheta \mid x)$ and $\alpha_2(\vartheta \mid x)$ with finite variance (and where α_2 is bounded away from zero) such that

$$\frac{\partial}{\partial\tau}\mathbb{E}\left[\frac{\partial}{\partial\vartheta}\ell(\mathbf{D}_d, \vartheta, \gamma_\tau(X))\right] = \frac{\partial}{\partial\tau}\mathbb{E}\left[\alpha_1(\vartheta \mid X)\gamma_\tau(X)\right]$$
$$\frac{\partial}{\partial\tau}\mathbb{E}\left[b(X)\frac{\partial}{\partial\gamma}\ell(\mathbf{D}_d, \vartheta, \gamma_\tau(X))\right] = \frac{\partial}{\partial\tau}\mathbb{E}\left[b(X)\alpha_2(\vartheta \mid X)\gamma_\tau(X)\right]$$

By the chain rule and iterated expectations on the score above, we have

$$\alpha_1(\vartheta \mid x) = \mathbb{E}\left[\frac{\partial g(\vartheta \mid a, \mathbf{D})}{\partial a}\Big|_{a=\gamma(X)}\middle| X = x\right]$$
$$\alpha_2(\vartheta \mid x) = \mathbb{E}\left[\frac{\partial^2 \ell(\vartheta \mid a, \mathbf{D})}{\partial a^2}\Big|_{a=\gamma(X)}\middle| X = x\right]$$

Writing the derivative of the likelihood with respect to the scalar output of γ as

$$\tilde{g}(\vartheta \mid \gamma, \mathbf{D}) = \frac{\partial \ell(\vartheta \mid a, \mathbf{D})}{\partial a}\Big|_{a=\gamma(X)}$$

we can combine these terms to form the orthogonal score

$$g^*(\vartheta \mid \gamma, \mathbf{D}_d) = g(\vartheta \mid \gamma, \mathbf{D}_d) - \alpha_1(\vartheta \mid x) \cdot \alpha_2(\vartheta \mid x)^{-1} \cdot \tilde{g}(\vartheta \mid \gamma, \mathbf{D}) \tag{A.1}$$

This orthogonal score may then be used to estimate θ while removing bias due to the plug-in estimator γ_0.

The nuisance parameters α_1 and α_2 are projections of second derivatives of ℓ onto the space of covariates X that enter γ. Unlike regression settings, each depends on the structural parameters ϑ; this is similar to Example 3 of Chernozhukov et al. (2022). As in that setting, initial estimators $\hat{\gamma}$ and $\hat{\theta}$ can be constructed using sample splitting, and then plugged into α_1 and α_2 to get predicted values and estimate the conditional expectations $\hat{\alpha}$ using nonparametric regression on X. These "plugin" estimators form the nuisance parameter $\hat{\alpha}(x) = \hat{\alpha}_1(\hat{\vartheta} \mid x) \cdot \hat{\alpha}_2(\hat{\vartheta} \mid x)^{-1}$ that yields a version of equation (A.1) that will be used in estimation (I omit the multiple indices used in sample splitting for ease of exposition):

$$g^*(\vartheta \mid \gamma, \mathbf{D}_d) = g(\vartheta \mid \gamma, \mathbf{D}_d) - \hat{\alpha}(x) \cdot \tilde{g}(\vartheta \mid \gamma, \mathbf{D}_d)$$

This orthogonal moment can be used as in standard GMM both to estimate ϑ without bias and construct the asymptotic variance matrix of the structural parameters. I follow the steps in Chernozhukov et al. (2022) with threefold sample splitting.

A.6 Likelihood derivation: demand side

I now derive the demand side likelihood as presented in equation (1.10). Subscripts for j and t will be omitted where possible (since this focuses on the likelihood contribution for any given auction), as well as the dependence on parameters ϑ_d, to streamline notation.

The likelihood that no bids are observed is simply the likelihood that the highest bid (integrated over the arrival distribution) is smaller than the minimum bid. Thus, the likelihood contribution for observing no bids is

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{\Lambda^n e^{-\Lambda}}{n!} \underbrace{F_B(m)^n}_{\mathbb{P}[\text{all bids below } m]} &= \\
&= e^{-\Lambda[1-F_B(m)]} \sum_{n=0}^{\infty} \frac{[\Lambda F_B(m)]^n e^{-[\Lambda F_B(m)]}}{n!} \\
&= e^{-\Lambda[1-F_B(m)]}
\end{aligned}$$

where the third equality holds since the sum is the integral of a Poisson density with mean $\Lambda F_B(m)$.

The likelihood contribution from auctions with one observed bidder uses the fact that one bid is not censored, but the other $n-1$ are. For any n bidders that arrive,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\Lambda^n e^{-\Lambda}}{n!} \underbrace{n F_B(m)^{n-1} f_B(v^{(1)})}_{\mathbb{P}[\text{only 1 bid above } m]} &= \\
&= \frac{f_B(v^{(1)})}{F_B(m)} e^{-\Lambda[1-F_B(m)]} \sum_{n=0}^{\infty} \frac{[\Lambda F_B(m)]^n e^{-[\Lambda F_B(m)]}}{n!} n \\
&= f_B(v^{(1)}) \Lambda e^{-\Lambda[1-F_B(m)]}
\end{aligned}$$

where the second equality holds since the summand is 0 for $n = 0$, and the last equality holds since the sum is the first moment of the Poisson distribution with mean $\Lambda F_B(m)$.

The last case, where $N_j \geq 2$, combines all possible arrival orders with at least 2 bidders. The precise values of other bidders are not necessary to construct the partial likelihood; the two highest bids provide enough information about the arrival process.

$$\begin{aligned}
&\sum_{n=2}^{\infty}\frac{\Lambda^n e^{-\Lambda}}{n!}\underbrace{f_B(v^{(1)} \mid v^{(2)})f_B(v^{(2)} \mid n \text{ bids})}_{\mathbb{P}[2 \text{ highest bids}]} = \\
&\quad = f_B(v^{(1)} \mid v^{(2)})\sum_{n=2}^{\infty}\frac{\Lambda^n e^{-\Lambda}}{n!}n(n-1)(1-F_B(v^{(2)}))F_B(v^{(2)})^{n-2}f_B(v^{(2)}) \\
&\quad = f_B(v^{(1)} \mid v^{(2)})\frac{(1-F_B(v^{(2)}))}{F_B(v^{(2)})}\frac{f_B(v^{(2)})}{F_B(v^{(2)})}e^{-\Lambda[1-F_B(v^{(2)})]} \\
&\qquad\quad \cdot \sum_{n=0}^{\infty}\frac{[\Lambda F_B(v^{(2)})]^n e^{-[\Lambda F_B(v^{(2)})]}}{n!}(n^2-n) \\
&\quad = f_B(v^{(1)})f_B(v^{(2)})\Lambda^2 e^{-\Lambda[1-F_B(v^{(2)})]}
\end{aligned}$$

where the third equality follows since $n^2 - n = 0$ for $n = 0, 1$, the fourth equality comes from the difference of the first and second raw moments of the Poisson distribution and further simplification. Combining the likelihood component of each case ($N = 0$, $N = 1$, or $N \geq 2$) with the density of the reserve price, we obtain equation (1.10).

Test with simulated data

For the simulations, I use a modified version of the data-generating process various architectures for the item value index γ. In the first architecture, I assume item j's log-value γ_j is known (i.e. $\gamma(\gamma_j) = \gamma_j$). In the following, I assume item values are a function of 50 indicator variables, each of which is randomly generated with average probability 0.1. I allow item values to be generated from a dense neural network (mapping from 50 indicator variables to two hidden layers of 10 nodes each before outputting to a scalar).

The number of bidders is Poisson distributed with mean $\Lambda_j = \exp(\delta_{0,1} + \delta_{0,2}\gamma_j)$, so in these simulations bidder arrival does not depend on the reserve price. I also set $m = r$ in the simulations, so the minimum bid and reserve price are the same. I model $r_j \sim \mathcal{N}(\mu_r, \sigma_r^2)$ and $v_{ij} \sim \mathcal{N}(0, \sigma_B^2)$. I report the results for this set of simulations in Table A.4.

Table A.4: Simulations for maximum likelihood estimation (demand)

	Regression: γ_0 on $\hat{\gamma}$		Structural Parameters				
	Intercept	Coef	σ_B^2	μ_r	σ_r^2	$\delta_{0,1}$	$\delta_{0,2}$
True values	0.0	1.0	0.5	0.25	0.5	1.0	0.5
Known γ_0							
$N = 2{,}000$	-	-	0.498	0.251	0.499	1.019	0.469
	-	-	(0.009)	(0.012)	(0.008)	(0.031)	(0.047)
$N = 10{,}000$	-	-	0.5	0.25	0.5	1.022	0.473
	-	-	(0.004)	(0.005)	(0.004)	(0.012)	(0.02)
Estimated γ_0 (uncorrected)							
$N = 2{,}000$	0.002	0.976	0.501	0.261	0.491	1.095	0.748
	(0.112)	(0.091)	(0.06)	(0.112)	(0.02)	(0.155)	(0.164)
$N = 10{,}000$	0.031	1.099	0.55	0.283	0.528	1.049	0.749
	(0.055)	(0.04)	(0.05)	(0.087)	(0.014)	(0.088)	(0.108)
Estimated γ_0 (orthogonalized)							
$N = 2{,}000$	-0.022	0.91	0.661	0.235	0.625	0.809	0.163
	(0.134)	(0.119)	(0.098)	(0.136)	(0.042)	(0.243)	(0.296)
$N = 10{,}000$	0.035	1.069	0.59	0.273	0.556	0.985	0.554
	(0.07)	(0.054)	(0.062)	(0.101)	(0.021)	(0.105)	(0.153)

Notes: Average (standard deviation) parameters are from 100 simulations for each case. Starting values were chosen randomly using Julia's Flux package initializations.

A.7 Likelihood approach: supply side belief estimation

Several functions (e.g. R, K, and their derivatives with respect to r, and expectations with respect to belief densities or bidder values) involve multiple integrals and/or summations, and are therefore infeasible to compute repeatedly for all possible parameters ϑ_s (the parameters of the seller value and cost distributions) and ϑ_b (the parameters of the seller prior beliefs). I use dense neural networks to approximate several functions used in the estimation procedure.

Each neural network maps from $\mathbb{R}^M$ to $\mathbb{R}$, and each is composed of one input layer, 9 hidden layers, and one output layer. The activation function for each hidden layer is leakyrelu, and the number of nodes from input layer to output layer for each neural network is: M, 50, 100, 100, 200, 300, 3000, 300, 200, 100, 100, 50, 1. The activation function for the output layer is listed with the associated function below.

To construct each approximation, I generate datasets on which to train each neural network for various parameter values. The bounds of each variable used in the approximations are chosen to cover the empirical support of the corresponding variables where they are observed (e.g., $\delta_{j,0,1}$) and sufficiently large support where they are unobserved (e.g. seller prior parameters). I use 99% of each dataset for training and 1% for holdout validation. I train each neural network on the respective training datasets in batches of 50 for 100 epochs before training the network on the full training dataset for 50 epochs; I exit training early if the mean square error of the holdout sample is less than 1e-5. The approximations (in bold) are constructed in the following order, with additional details listed for each approximation and the construction of the associated datasets. Each function is fit by minimizing mean square prediction error, though in some cases transformations are applied to improve relative accuracy for some parameter values.

1. *Functions with δ known.*
   (a) *Evaluate revenue and keep probabilities.* Using the estimated bidder value distribution and arrival parameters, I first evaluate R_n and K_n for $n = 0, 1, ..., 150$. I then construct 316 Chebyshev nodes in each dimension for $r \in [0.01, 6.25]$ and $\Lambda \in [-1.5, \ln(150)]$ and evaluate R and K, respectively, by taking their dot product with $\{p_n(r, \Lambda)\}_{n=0}^{150}$ evaluated at each node (I chose 316 because Floor(100,000$^{0.5}$) = 316).
      i. **Expected revenue** R (softplus activation). Inputs: r and $\delta_{j,0,1}$.
      ii. **Keep probability** K (sigmoid activation). Inputs: r and $\delta_{j,0,1}$.
   (b) *Search for optimal reserve price.* I construct 316 Chebyshev nodes in each dimension for $v_0 \in [-1.25, 6.25]$ and $\delta_{j,0,1} \in [-1.5, \ln(150)]$ and search for the optimal reserve price r^* in 0.01, 0.02, ..., 6.25 along with the expected profit and seller surplus (profit minus outside option value) at the optimum.

i. **Virtual type** ψ (identity activation). Inputs: r^* and $\delta_{j,0,1}$.

ii. **Optimal reserve price** ψ^{-1} (identity activation). Inputs: v_0 and $\delta_{j,0,1}$.

iii. **Expected surplus** $\Pi^* - v_0$ (exponential activation) Inputs: v_0 and $\delta_{j,0,1}$. Since expected surplus is positive when entry costs are zero, I minimize the mean square prediction error of the *log* expected surplus. This increases the relative accuracy of predicted expected surplus where it is small, which is important for precisely approximating the entry threshold in the next step.

(c) *Entry threshold.* I then construct 316 Chebyshev nodes in each dimension for $c_E \in [0, 0.5]$ and $\delta_{j,0,1} \in [-5, \ln(150)]$ and search for the maximum $v_0 \in [0, 6.25]$ such that expected surplus is weakly positive. Since expected surplus is monotonic in v_0, I use a binary search algorithm (i.e., evaluating expected surplus at the midpoint of $[0, 6.25]$, determining whether $\bar{v}$ lies above or below the midpoint, and iterating with additional intervals) until the difference in successive iterations is less than 0.01.

i. **Entry threshold** $\bar{v}$ (identity activation). Inputs: c_E and $\delta_{j,0,1}$.

2. *Functions with δ unknown.* I approximate the following functions for the full model with 2-dimensional unknown parameter δ_0 and the arrival coefficient model where only $\delta_{0,2}$ is unknown. The prior parameters are the mean $\{\mu_{0,1}, \mu_{0,2}\}$, standard deviations $\sigma_{0,1}$ and $\sigma_{0,2}$, and correlation $\tilde{\rho}_0$. Due to the higher dimensionality due to the belief parameters, I sample 100,000 input values for each step rather than constructing a grid of Chebyshev nodes. Unless otherwise specified, each input is sampled uniformly on the stated support.

(a) *Search for optimal reserve price.* I sample $v_0 \sim F_S$ (bounded on $[0.0, 6.25]$), $\delta_{j,0,1} \sim U[-1.5, \ln(150)]$, $\mu_{0,2} \sim U[-0.75, 0.75]$, $\sigma_{0,i} \sim U[0.01, 1.0]$ for $i = 1, 2$, and $\tilde{\rho}_0 \sim U[-0.9, 0.9]$ ($\sigma_{0,2}$ and $\tilde{\rho}_0$ are only sampled for the full model). I then search for the optimal reserve price r^* in 0.01, 0.02, ..., 6.25 along with the expected profit and seller surplus (profit minus outside option value) at the optimum, integrating over sellers' belief distributions to do so using Gauss-Hermite quadrature with 5 points in each dimension.[3]

i. **Virtual type** ψ (identity activation). Inputs: r^*, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$.

ii. **Optimal reserve price** ψ^{-1} (identity activation). Inputs: v_0, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$.

iii. **Expected surplus** $\Pi - v_0$ (exponential activation). Inputs: v_0, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. As with step 2(c), I minimize mean square prediction error using the log expected surplus.

(b) *Entry threshold.* I sample the parameters $c_E \sim U[0,1]$ (bounded on $[1.25, 6.25]$), $\delta_{j,0,1} \sim U[-1.5, \ln(150)]$, $\mu_{0,2} \sim U[-0.75, 0.75]$, $\sigma_{0,i} \sim U[0.01, \sqrt{0.5}]$ for $i = 1, 2$, and $\tilde{\rho}_0 \sim U[-0.95, 0.95]$ ($\sigma_{0,2}$ and $\tilde{\rho}_0$ are only sampled for the full model). I use the same binary search algorithm as in 1(c) to find $\bar{v}$.

i. **Entry threshold** $\bar{v}$ (identity activation). Inputs: c_E, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$.

(c) *Updating process.* I sample 100,000 draws of $v_0 \sim F_S$ (truncated on $[0, 6.25]$), $Z'\lambda \sim \mathcal{N}(2,1)$ (truncated on $[-1.5, \ln(150)]$), $\mu_{0,i} \sim U[-0.75, 0.75]$ for $i = 1, 2$, $\sigma_{0,i} \sim U[0.01, 1.0]$ for $i = 1, 2$, and $\tilde{\rho}_0 \sim U[-0.9, 0.9]$ ($\sigma_{0,2}$ and $\tilde{\rho}_0$ are only sampled for the full model). I sample $r^* \sim 0.2\mathcal{N}(0.25, 0.5^2) + 0.7\mathcal{N}(1.25, 0.5^2) + 0.1\mathcal{N}(3.0, 1.5^2)$ (a Gaussian mixture), truncated on $[0, 6.25]$.[4] I also simulate profit signals ϵ as in equation (1.6) from $\mathcal{N}(0, \sigma_\Pi^2(r))$, where $\sigma_\Pi^2(r) = \exp(\tilde{\sigma}_{0,\Pi} + \tilde{\sigma}_{1,\Pi} r)^2$ is estimated using the empirical difference between profit signals and expected

3. This procedure yields candidate quadrature nodes $\ln(\Lambda)$ at which I evaluate the expected revenue and keep probaiblities. For $\ln(\Lambda) \in [-5, \ln(150)]$ I use the evaluated R_n and K_n as in step 1(a) above, and for $\ln(\Lambda) \notin [-5, \ln(150)]$ I extrapolate using the fitted values of R and K.

4. I chose this after experimenting with various distributions from which to sample; this density oversamples for relatively low draws of the reserve price where R can vary significantly with r^* depending on the value of $\delta_{0,2}$.

profit among experienced sellers.

I compute the updated beliefs via a Laplace approximation to the posterior for each observation. I first use a resilient backpropogation algorithm to search for the new maximum a posteriori estimates $\delta^*_{j,0,1}$ and $\mu^*_{0,2}$ given prior parameters, r^*, and ϵ. I then evaluate the updated covariance parameters ($\sigma^*_{0,1}$, $\sigma^*_{0,2}$, and $\tilde{\rho}^*_0$) by evaluating the Hessian of the posterior evaluated at the maximum a posteriori estimate. I drop all evaluations for which this algorithm returns either posterior parameters outside the simulation bounds or an invalid covariance matrix, and evaluate the neural networks using the resulting posterior parameters.

i. **Updating means $\mu_{0,1}$ and $\mu_{0,2}$** (identity activation). Inputs: ϵ, $\delta_{j,0,1}$, $\rho(r^*)$, $\mu_{0,2} \cdot \rho(r^*)$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. Instead of approximating each updated parameter directly, I minimize mean square prediction error for the standardized difference $(\mu^*_{0,i} - \mu_{0,i})/\sigma_{0,i}$ for $i = 1, 2$; this ensures more accurate update steps for the sellers' mean parameters when beliefs are more highly concentrated.

ii. **Updating standard deviations $\sigma_{0,1}$ and $\sigma_{0,2}$** (identity activation). Inputs: ϵ, $\delta_{j,0,1}$, $\rho(r^*)$, $\mu_{0,2} \cdot \rho(r^*)$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. Instead of approximating each updated parameter directly, I minimize mean square prediction error for the ratio $(\sigma^*_{0,1})/\sigma_{0,i}$ for $i = 1, 2$; this ensures more accurate update steps for the sellers' covariance parameters when beliefs are more highly concentrated.

iii. **Updating posterior correlation $\tilde{\rho}$** (identity activation). Inputs: ϵ, $\delta_{j,0,1}$, $\rho(r^*)$, $\mu_{0,2} \cdot \rho(r^*)$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. Instead of approximating the updated parameter directly, I minimize mean square prediction error for the difference $\tilde{\rho}^*_0 - \tilde{\rho}_0$ for $i = 1, 2$.

I chose both $\rho(r^*)$ and $\mu_{0,2} \cdot \rho(r^*)$ as inputs after experimenting with various architectures.

APPENDIX B

APPENDIX FOR CHAPTER 2

B.1 Stochastic Nash bartering

Traditional Nash bargaining assumes a convex choice set, though other work weakens this assumption (Herrero, 1989; Serrano & Shimomura, 1998; Zhou, 1997). In our setting, we have a discrete set of possible terms that can be chosen. We wish to represent the term bartering game as an instant decision that relies on firm bargaining power in the same way as prices are set via Nash bargaining. We illustrate how the stochastic Nash bargaining solution is in expectation equal to the classic Nash bargaining solution over the convex hull of *ex ante* expected payoffs from possible choices of contract terms.

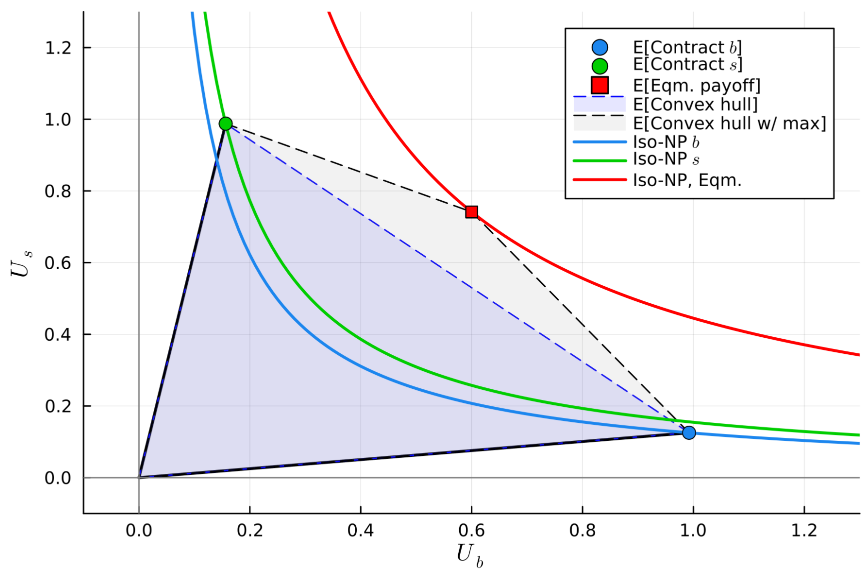
The stochastic Nash bartering solution we propose also coincides with the Nash bargaining solution on an *ex ante* convexification of the expected payoff set. We define the expected payoff set $PS_E = \{\mathbb{E}[v(m_j;\epsilon_j)] \mid m_j \in \mathcal{M}\} \cup \{\mathbb{E}[v(m_j;\epsilon_j) \mid NP_j \geq NP_k \ \forall k \neq j]\}$. That is, PS_E is the set of achievable expected payoffs from any individual term and the expected payoff from selecting the term with the highest realized Nash product (under whichever game is being considered). Let PS_∞ be the convex hull of the expected payoff set PS_E. Consider any randomization protocol on PS_E, so any feasible payoff in the convex hull PS_∞ can be represented by the lottery on PS_E. Then the solution of the Nash bargaining program over the convex set PS_∞ is equal to the solution of choosing $\mathbb{E}[v(m_j;\epsilon_j) \mid NP_j \geq NP_k \ \forall k \neq j]$ from the restricted set PS_E.[1] The difference $\mathbb{E}[v(m_j;\epsilon_j) \mid NP_j \geq NP_k \ \forall k \neq j] - \max_{m_j \in \mathcal{M}}\{\mathbb{E}[v(m_j;\epsilon_j)]\}$ is strictly positive when there are at least two non-default terms in $\mathcal{M}$, and it represents the expected option value

1. In other words, any distinct randomization protocol over the terms in the set PS_E (that places positive probability on any term in $\mathcal{M}$, regardless of the realization of ϵ) will in expectation do worse (in the Nash program sense) than choosing whichever term has the highest Nash product after ϵ is drawn.
In the simple case where there are only 2 terms other than the default term (and terms cannot be added together), the expected payoff is also a weighted average of the expected payoffs from each term conditional on that term being selected; see Figure B.4(a) for an illustration.

from having multiple potential contracts to choose from.

Figure B.1 illustrates the expected convex hull PS_∞ of the set of expected payoffs PS_E, represented by the grey and blue shaded areas. The blue shaded area represents the convex hull only over the proposed contracts m_0, m_b, and m_s (i.e., $PS_E \setminus \{\mathbb{E}[v(m_j;\epsilon_j) \mid NP_j \geq NP_k\ \forall k \neq j]\}$). If the firms commit *ex ante* to select the term with highest Nash product once all uncertainty is resolved, the expected payoff exceeds that of any other contract selection rules in the convex hull.

Figure B.1: Firm payoffs in the bartering stage of the price-first game (unrestricted model)



0

Notes: The plots assume $r_b = 1$, $r_s = 1.25$, $a_b = -0.19$, $a_s = 0.22$, $\alpha = 4$, and $\tau = 0.4$. None of these values necessarily represent equilibrium actions. For clarity in illustrating the expected convex hull, we present a modified case where the combined term is not considered by the firms.

B.2 Proofs and derivations

Proof of Proposition 6

The first-order conditions for b and s in PF and PL are

$$
\begin{aligned}
&\{b, PF\} && \frac{\partial}{\partial r_b} c_b(r_b, a_b) = \cos(\theta(a_b)) \\
&\{s, PF\} && \frac{\partial}{\partial r_s} c_s(r_s, a_s) = \sin(\theta(a_s)) \\
&\{b, PL\} && \frac{\partial}{\partial r_b} c_b(r_b, a_b) = \tau[\cos(\theta(a_b)) + \sin(\theta(a_b))] \\
&\{s, PL\} && \frac{\partial}{\partial r_s} c_s(r_s, a_s) = (1-\tau)[\cos(\theta(a_s)) + \sin(\theta(a_s))]
\end{aligned}
$$

For fixed angles a_i in the first quadrant and fixed bargaining parameter $\tau \in [0,1]$, each of the right-hand side expressions are weakly positive constants, while the left-hand side expressions are increasing in r. Since $c_i(\cdot, a_i) = 0$ and $\frac{\partial c_i(r_i,a_i)}{\partial r_i}|_{r_i=0} = 0$, this is the unique value that satisfies the first-order conditions. Further, since the right-hand sides of the above expressions do not depend on r_i and costs are convex in r_i, the second-order conditions hold. Thus, the equilibrium exists and is unique for both the price-first and price-last games.

Further note that since c_i is increasing in the search radius r_i and search angles a_i are fixed across the two games, search intensity is higher in whichever game results in higher investment costs. For firms b and s respectively, this implies that search intensity is higher when

$$
\begin{aligned}
&\{b\} && \cos(\theta(a_b)) \geq \tau[\cos(\theta(a_b)) + \sin(\theta(a_b))] \\
&\{s\} && \sin(\theta(a_s)) \geq \tau[\cos(\theta(a_s)) + \sin(\theta(a_s))]
\end{aligned}
$$

or equivalently,

$$
\begin{aligned}
&\{b\} \qquad\qquad a_b \leq \frac{1}{\pi}\arctan\left(\frac{1-\tau}{\tau}\right) - 0.25 \\
&\{s\} \qquad\qquad a_s \geq \frac{1}{\pi}\arctan\left(\frac{1-\tau}{\tau}\right) - 0.25
\end{aligned}
$$

Note that the socially optimal level of investment maximizes the total surplus from producing new terms, net of search costs. That is, a social planner solves

$$
\max_{r_b, r_s} \quad [r_b\cos(\theta(a_b)) + r_b\sin(\theta(a_b)) + r_s\cos(\theta(a_s)) + r_s\sin(\theta(a_s))] - c_b(r_b, a_b) - c_s(r_s, a_s)
$$

Taking first-order conditions yields

$$
\begin{aligned}
\frac{\partial}{\partial r_b}c_b(r_b, a_b) &= \cos(\theta(a_b)) + \sin(\theta(a_b)) \\
\frac{\partial}{\partial r_s}c_s(r_s, a_s) &= \sin(\theta(a_s)) + \cos(\theta(a_s))
\end{aligned}
$$

The right-hand sides weakly exceed the corresponding right-hand side expressions for the first-order conditions of both firms in the both the price-first and price-last games. Since costs are convex, this implies the socially optimal level of investment is weakly higher than the investment by either firm in either game. This inequality is strict except where $\tau \in \{0, 1\}$ in the price-last game (in which case exactly one of the two firms invests at the socially optimal level) and where $|a_i| = 0.25$ in the price-first game (in which case both firms invest at the socially optimal level).

Proof of Proposition 7

Lemma 1 provides the optimal search radii in each game as functions of the search angle. Plugging in these strategies yields the following maximization problems in the price-first setting

$$\max_{a_b\colon |a_b|\leq \bar{a}} \quad \left[\frac{1}{\gamma_b}\cos(\theta(a_b))^2\exp(\gamma_a a_b^2) + \frac{1}{\gamma_s}\cos(\theta(a_s))\sin(\theta(a_s))\exp(\gamma_a a_s^2)\right]$$
$$- 0.5\frac{1}{\gamma_b}\cos(\theta(a_b))^2\exp(\gamma_a a_b^2)$$
$$\max_{a_s\colon |a_s|\leq \bar{a}} \quad \left[\frac{1}{\gamma_b}\cos(\theta(a_b))\sin(\theta(a_b))\exp(\gamma_a a_b^2) + \frac{1}{\gamma_s}\sin(\theta(a_s))^2\exp(\gamma_a a_s^2)\right]$$
$$- 0.5\frac{1}{\gamma_s}\sin(\theta(a_s))^2\exp(\gamma_a a_s^2)$$

and the price last setting

$$\max_{a_b\colon |a_b|\leq \bar{a}} \quad \tau\cdot\left[\frac{\tau}{\gamma_b}(1+\sin(2\theta(a_b)))\exp(\gamma_a a_b^2) + \frac{1-\tau}{\gamma_s}(1+\sin(2\theta(a_s)))\exp(\gamma_a a_s^2)\right]$$
$$- 0.5\frac{\tau^2}{\gamma_b}[1+\sin(2\theta(a_b))]\exp(\gamma_a a_b^2)$$
$$\max_{a_s\colon |a_s|\leq \bar{a}} \quad (1-\tau)\cdot\left[\frac{\tau}{\gamma_b}(1+\sin(2\theta(a_b)))\exp(\gamma_a a_b^2) + \frac{1-\tau}{\gamma_s}(1+\sin(2\theta(a_s)))\exp(\gamma_a a_s^2)\right]$$
$$- 0.5\frac{(1-\tau)^2}{\gamma_s}[1+\sin(2\theta(a_s))]\exp(\gamma_a a_s^2)$$

Proof of (i). We now focus on the price-first setting. First define $f_{i,PF}(a_1)$ for $i \in \{b, s\}$ as firm i's expected payoff when choosing angle a minus their expected payoff from choosing $-a$, for any fixed angle from firm i's counterpart $-i$. That is,

$$f_{b,PF}(a) = \frac{1}{2\gamma_b}\exp(\gamma_a a_1^2)\left[\cos(\theta(a))^2 - \cos(\theta(-a))^2\right]$$
$$f_{s,PF}(a) = \frac{1}{2\gamma_s}\exp(\gamma_a a_2^2)\left[\sin(\theta(a))^2 - \sin(\theta(-a))^2\right]$$

For $a > 0$, we have $f_{b,PF}(a) < 0$ ($-a$ dominates a) and $f_{s,PF}(a) > 0$ (a dominates $-a$). Thus b always chooses $a^*_{b,PF} \in [-\bar{a}, 0]$ (the "lower half" of the first quadrant) and s always chooses $a^*_{s,PF} \in [0, \bar{a}]$ (the "upper half" of the first quadrant).

Taking derivatives of the firms' profit functions with respect to their choice variables yields the following expressions

$$
\begin{aligned}
&\{b\} \qquad \frac{1}{\gamma_b}\exp(\gamma_a a_b^2)\cdot\cos(\theta(a_b))\cdot\Big[\gamma_a a_b\cos(\theta(a_b))-\pi\sin(\theta(a_b))\Big] \\
&\{s\} \qquad \frac{1}{\gamma_s}\exp(\gamma_a a_s^2)\cdot\sin(\theta(a_s))\cdot\Big[\gamma_a a_s\sin(\theta(a_s))+\pi\cos(\theta(a_s))\Big]
\end{aligned}
$$

For $|a| \leq 0.25$ (i.e., a within the first quadrant), $\cos(\theta(a))$ and $\sin(\theta(a))$ at least weakly positive, which implies that the terms preceding the brackets are positive. Note that $sign(\gamma_a a_b\cos(\theta(a_b))) = sign(\gamma_a a_s\sin(\theta(a_s))) = sign(a_s)$ for a_b, a_s within the first quadrant. This implies that the derivative for firm b is weakly negative when $a_b \in [-\bar{a}, 0]$ and the derivative for firm s is weakly positive for $a_s \in [0, \bar{a}]$ (these are strict for either $\gamma_a > 0$ or $|a_i| \neq 0.25$). Therefore, for all $\gamma_a \geq 0$, it holds that the unique optimal search angles in the price-first game are $a^*_{b,PF} = -\bar{a}$ and $a^*_{s,PF} = \bar{a}$.

Proof of (ii). We now turn to the price-last setting. We have the following derivatives of the firms' maximization problems with respect to their own search angles, after applying trigonometric identities:

$$
\begin{aligned}
&\{b\} \qquad \frac{1}{\gamma_b}\exp(\gamma_a a_b^2)\cdot\tau^2\cdot 2\pi\cos(\pi a_b)\Big[\frac{\gamma_a}{\pi}a_b\cos(\pi a_b)-\sin(\pi a_b)\Big] \\
&\{s\} \qquad \frac{1}{\gamma_s}\exp(\gamma_a a_s^2)\cdot(1-\tau)^2\cdot 2\pi\cos(\pi a_s)\Big[\frac{\gamma_a}{\pi}a_s\cos(\pi a_s)-\sin(\pi a_s)\Big]
\end{aligned}
$$

The terms preceding the brackets are weakly positive for all angles within the first quadrant (strictly so for $\tau \in (0,1)$). Therefore, the signs and zeros of these derivatives are determined solely by the signs and zeros of the bracketed terms. We now restrict attention to only the bracketed terms, which have the same functional form for both firms.

Denote $f_1(a) = \frac{a}{\pi}\cos(\pi a)$ and $f_2(a) = \sin(\pi a)$, and define $f(a, \gamma_a) \equiv \gamma_a f_1(a) + f_2(a)$. Since $f(0, \gamma_a) = 0$, the angle $a_i = 0$ always satisfies the interior first-order condition. Differentiating $f(a, \gamma_a)$ with respect to a, we have

$$\frac{1}{\pi}[\gamma_a - \pi^2]\cos(\pi a) - \gamma_a a \sin(\pi a)$$

Note that for $|a| \leq 0.25$, it holds that $\cos(\pi a) > 0$ and $a\sin(\pi a) \geq 0$ (with strict inequality for $a \neq 0$). Assume that $\gamma_a < \pi^2$, which implies that $\frac{\partial}{\partial a} f(a, \gamma_a) < 0$ for $|a| \leq \bar{a}$. Since $f(0, \gamma_a) = 0$, monotonicity of f in a implies

$$f(a, \gamma_a) \begin{cases} > 0 & \text{if } a \in [-\bar{a}, 0) \\ < 0 & \text{if } a \in (0, \bar{a}] \end{cases}$$

That is, firms' profits are increasing in a for $a < 0$ and decreasing in a for $a > 0$. This implies that the unique optimal choice of a is $a^* = 0$ for $|a| \leq 0.25$ for $\gamma_a < \pi^2$.

Proof of (iii). We now consider the case where $\gamma_a > \pi^2$. Denoting $g_{i,PL}(a, \gamma_a)$ as the derivative of firm i's maximization problem with respect to its own choice angle, note that $g_{i,PL}(a, \gamma_a) = -g_{i,PL}(-a, \gamma_a)$, implying $g'_{i,PL}(a, \gamma_a) = g'_{i,PL}(-a, \gamma_a)$. Thus, for any angle a that is optimally chosen by firm i, the angle $-a$ also satisfies both the first- and second-order conditions. We therefore restrict attention (without loss of generality) to $a \in [0, \bar{a}]$.

We first consider $a = 0$. Note that $f(0, \gamma_a) = 0$ and $\frac{\partial f(a, \gamma_a)}{a}|_{a=0} > 0$, implying by the product rule that the second derivative of firms' maximization problem is positive at $a = 0$. Thus, $a = 0$ is not optimal when $\gamma_a > \pi^2$.

We now consider $a \in (0, \bar{a}]$, and look for solutions of the first-order condition. Dividing the equation $f(a, \gamma_a) = 0$ on both sides by $\frac{a}{\pi}\cos(\pi a)$ (which is strictly positive for $0 < |a| \leq 0.25$) yields

$$0 = \gamma_a - \frac{\pi}{a}\tan(\pi a)$$

We examine the second term, $\frac{\pi}{a}\tan(\pi a)$, to determine the behavior of this transformed first-order condition as a varies. Note that at the lower limit of this interval, we have

$$\lim_{a\downarrow 0}\frac{\pi}{a}\tan(\pi a) = \lim_{a\downarrow 0}\pi^2\sec^2(\pi a) = \pi^2$$

We examine how this function varies with a for $a > 0$. By applying trigonometric identities, we obtain

$$\begin{aligned}\frac{\partial}{\partial a}\left(\frac{\pi}{a}\tan(\pi a)\right) &= \frac{\pi}{a^2}\left[\pi a\sec^2(\pi a) - \tan(\pi a)\right]\\ &= \frac{\pi}{a^2\cos^2(\pi a)}\left[\pi a - 0.5\sin(2\pi a)\right]\end{aligned}$$

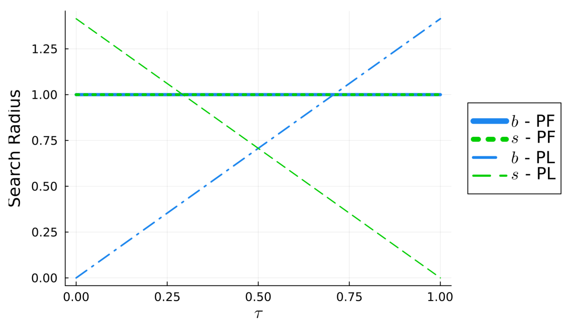
The term outside the brackets is strictly positive for $a > 0$. Defining the bracketed term in the second line as $h(a) \equiv \pi a - 0.5\sin(2\pi a)$, we have $h'(a) = \pi(1-\cos(2\pi a))$. Since $h'(a)$ is strictly positive for $a \in (0, 0.25)$ and $h(0) = 0$, we have that $h(a) > 0$ for $a \in (0, 0.25]$. Thus $\frac{\partial \frac{\pi}{a}\tan(\pi a)}{\partial a} > 0$ over the same interval. In turn, this implies that $f(\cdot, \gamma_a)$ has one zero in $(0, \bar{a}]$ if $\gamma_a \in (\pi^2, \frac{\pi}{\bar{a}}\tan(\pi\bar{a})]$.

We now prove that this zero is in fact optimal. By a similar argument as in (ii), for the angle $a^* \in (0, \bar{a}]$ such that $\gamma_a = \frac{\pi}{a^*}\tan(\pi a^*)$, the function $f(a, \gamma_a)$ is positive (and therefore the firms' profits are increasing) for any $a < a^*$ and it is negative (implying firms' profits are decreasing) for $a > a^*$. Thus a^* and $-a^*$ are optimal for the firms.

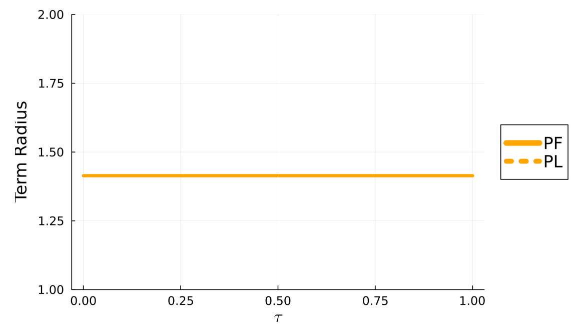
Finally, for $\gamma_a > \frac{\pi}{\bar{a}}\tan(\pi\bar{a})$, the firms' first-order condition is positive (and therefore profits are increasing in a) for all $a \in (0, 0.25]$. This implies that both the upper bound $\bar{a}$ and lower bound $-\bar{a}$ are optimal for sufficiently large γ_a.

B.3 Additional figures

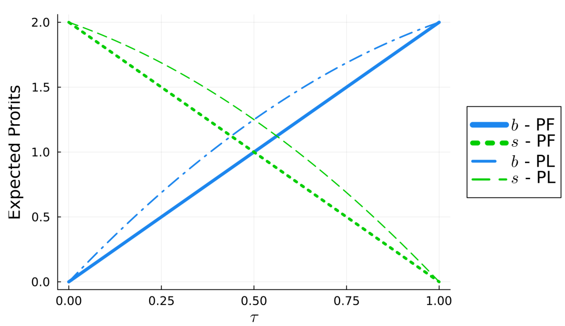
Figure B.2: Comparative statics with respect to τ (endogenous angle search)



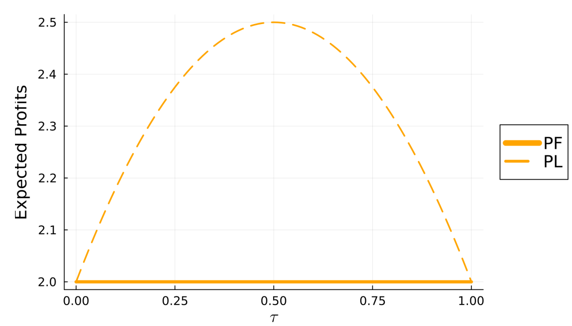
(a) Equilibrium search radius



(b) Chosen term radius



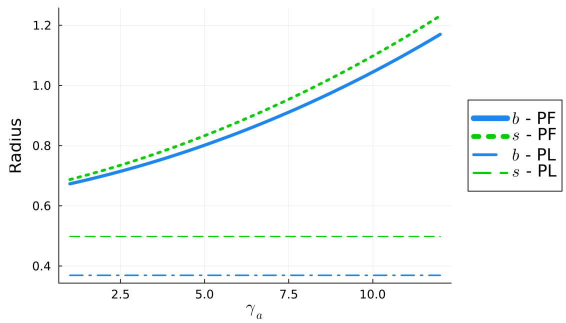
(c) Expected profits



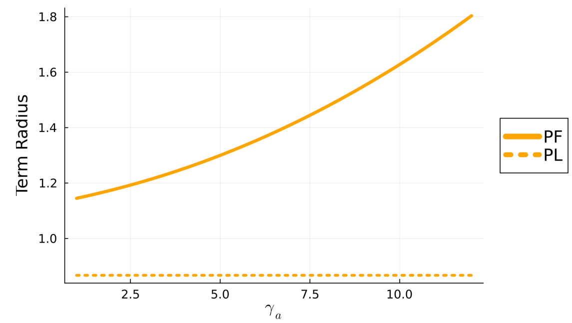
(d) Expected combined profits

Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\alpha = 2$, $\gamma_b = \gamma_s = 1$, $\gamma_a = 0$, $\pi_b = 2$, and $\pi_s = 1$. Comparative statics with respect to the search and term angles are omitted, as they are constant for all values of τ.

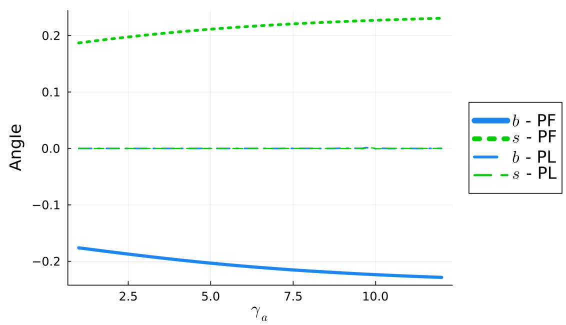
Figure B.3: Comparative statics with respect to γ_a (independent term shocks)



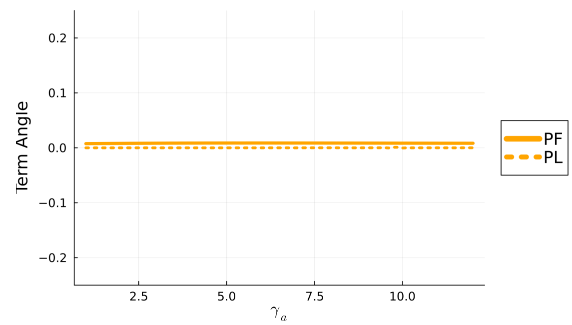
(a) Equilibrium search radius



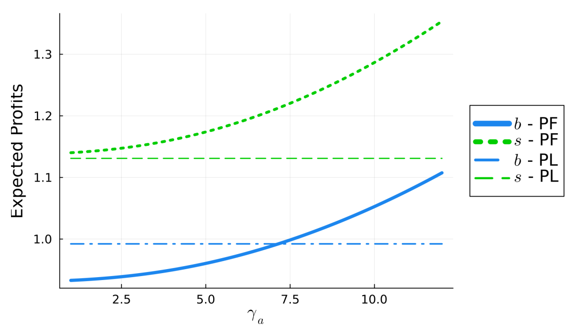
(b) Chosen term radius



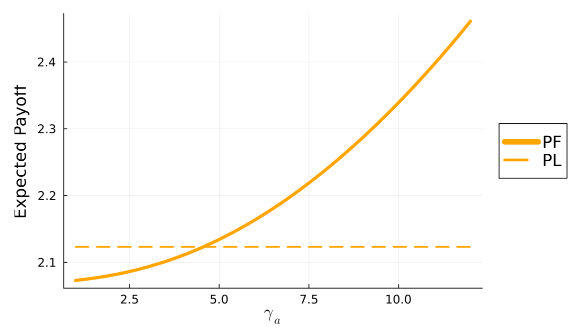
(c) Equilibirum search angle



(d) Chosen term angle



(e) Expected profits



(f) Expected combined profits

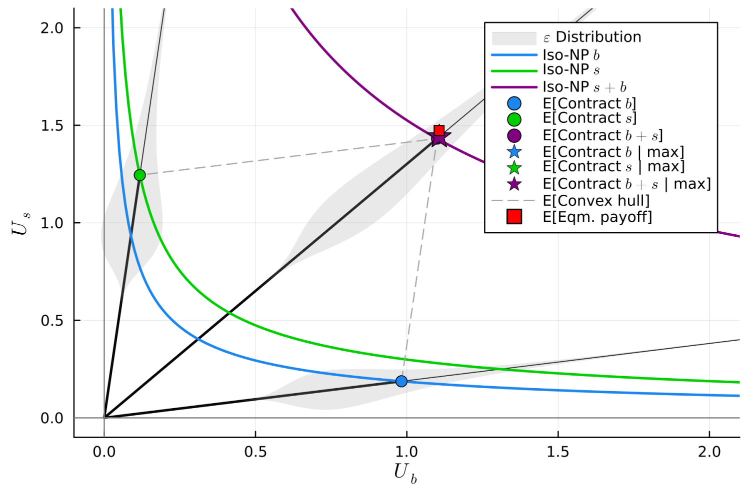
Notes: The panels on the left-hand side plot the variable of interest for both firms b and s in the price-first ("PF") and price-first ("PL") games. The panels on the right-hand side depict the corresponding outcomes of the contract in both games, for the combined firms. The plots assume $\tau = 0.45$, $\alpha = 2$, $\gamma_b = \gamma_s = 1$, $\pi_b = 2$, and $\pi_s = 1$.

B.4 Equilibrium contract with unrestricted term search and independent productivity shocks

Figure B.4 illustrates the bargaining game for the full model. In contrast to Figure 2.2, where ϵ's components are perfectly correlated, independent realizations of term-specific shocks lead to a different "shape" for the convex hull of every bargaining set; we therefore represent the variation due to ϵ with densities graphed along the radius of each term.

Figure B.4 also shows the option value from variety: the red square representing the expected equilibrium payoff is closer to the top-right than any of the individual terms. This is also illustrated by the expected payoff to contract terms, represented by stars in the figure below: individual terms must have a favorable ϵ draw to be chosen, so their expected value to the firms is greater when conditioning on the terms being chosen. Note that the stars for the individual terms do not appear in the figure, since the value of the shock must be so extreme as to push the value of the joint payout to a higher iso-curve relative to the iso-curve of the combined term. For the combined term m_{bs}, this difference is only slight since it is a low-probability event that either individual term will be chosen over the combined term.

Figure B.4: Firm payoffs in the bartering stage of the price-first game (unrestricted model)



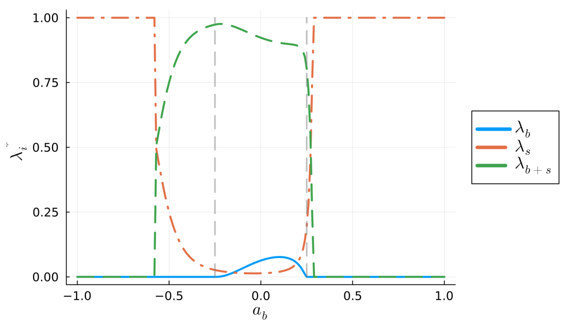
Notes: The plots assume $r_b = 1$, $r_s = 1.25$, $a_b = -0.19$, $a_s = 0.22$, $\alpha = 4$, and $\tau = 0.4$. These values are chosen for clarity of exposition and do not necessarily represent equilibrium actions. The blue and green stars are not shown on the figure, but instead lie far along their respective rays.

Figure B.5 similarly provides intuition for how $\lambda_{j,G}$ varies with firm search decisions. Firm b and s both have a positive probability that their term will be chosen when searching within the first quadrant ($a_b \in [-0.25, 0.25]$, marked with dashed grey lines in the figures). This highlights how Lemma 1 does not hold in this setting since all terms have a positive probability of being chosen.[2] However, the combined term is preferred in expectation except when firm b searches in a sufficiently value-destroying direction and makes term s relatively more favorable.[3] While the term choice probabilities look broadly similar in panels (a) and (b), the price-last game has a nonzero probability of choosing term b even when a_b is outside the first quadrant; this is due to the redistribution of term payoffs in the price-last game.

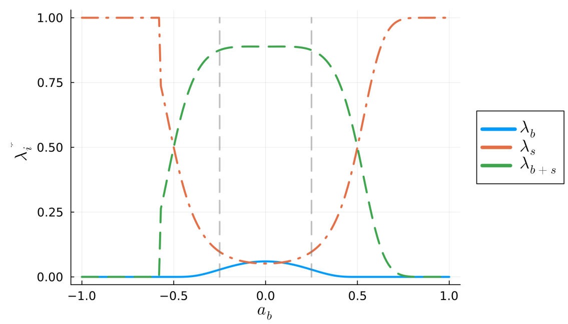
2. Note that taking the limit as $\alpha \to \infty$ (thereby decreasing the variance of ϵ) pushes $\lambda_{bs} \to 1$.

3. Searching outside the first quadrant ($a_b \notin [-0.25, 0.25]$) results in some value destruction, but as shown in panel (b), this may still result in a term that is preferred to an individual term in expectation.

Figure B.5: Term choice probabilities $\lambda_{j,G}$



(a) Price-first game



(b) Price-last game

Notes: The plots assume $r_b = 1$, $r_s = 1.25$, $a_s = 0.22$, $\alpha = 4$, and $\tau = 0.4$, and a_b varies to show the resulting probabilities that each term is chosen. These values are chosen for clarity of exposition and do not necessarily represent equilibrium actions.