Essays in Industrial Organization

by

Yuki Ito

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Benjamin Handel, Chair
Professor Matthew Backus
Professor Kei Kawai

Spring 2025

Essays in Industrial Organization

Copyright 2025
by
Yuki Ito

Abstract

Essays in Industrial Organization

by

Yuki Ito

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Benjamin Handel, Chair

This dissertation develops and applies economic frameworks in industrial organization that allow us to empirically study policy-related questions.

Chapter 1 considers how the possibility of credible announcements about firms' entry would alter market outcomes. If a firm can announce its intent to enter a market in a credible manner, it may be able to deter rival firms from entering. However, such an announcement may also provoke more aggressive competition from entrants, compared to remaining silent. In a setup of procurement auctions, we analyze the equilibrium effects of credible entry announcements, while firms face this strategic trade-off. The empirical setting I study is procurement auctions conducted by Montana Department of Transportation, where a designated online Q&A forum serves as an entry disclosure device. I specify and estimate a model of procurement auctions with costly entry, in which firms have the option to disclose entry during a period before the bids are submitted. We find that a disclosure deters entry from other firms, but leads remaining entrants to bid more aggressively. Overall, disclosure is beneficial for a firm if they can disclose at an early period. The availability of disclosure device decreases the auctioneer's payment by 6.3%, while increasing the winner's construction costs by 4.5% and decreasing the total entry costs by 11.1%.

Chapter 2, co-authored with Konan Hara and Paul Koh, studies identification and estimation of dynamic games when the underlying information structure is unknown to the researcher. To tractably characterize the set of Markov perfect equilibrium predictions while maintaining weak assumptions on players' information, we introduce Markov correlated equilibrium, a dynamic analog of Bayes correlated equilibrium. The set of Markov correlated equilibrium predictions coincides with the set of Markov perfect equilibrium predictions that can arise when the players can observe more signals than assumed by the analyst. Using Markov correlated equilibrium as the solution concept, we propose tractable computational strategies for informationally robust estimation, inference, and counterfactual analysis that deal with the non-convexities arising in dynamic environments. By applying this method, we

analyze the dynamic entry game between Starbucks and Dunkin' in the US and the role of informational assumptions.

Chapter 3, co-authored with Konan Hara, investigates the role of racial concordance among doctors, specialists, and patients in doctors' referral decisions. Using a 20% random sample of Medicare beneficiaries from 1999–2019 in the United States, we find that doctors are more likely to refer patients to specialists of the same race, especially when the patient also shares that race. To isolate the effect of racial concordance from differences in expected health outcomes and travel distance, we develop a structural model of referral decisions. The results indicate that Asian, Black, and Hispanic patients are, on average, referred to specialists with worse expected health outcomes than they would be in the absence of racial concordance preferences.

To my family.

Contents

Contents **ii**

List of Figures **iv**

List of Tables **v**

1 Entry Deterrence in Procurement Auctions **1**
1.1 Introduction 1
1.2 Institutional Background and Data 7
1.3 Model 9
1.4 Descriptive Evidence 14
1.5 Identification 18
1.6 Estimation 24
1.7 Estimation Results 29
1.8 Counterfactual Analysis 35
1.9 Conclusion 40

2 Identification and Estimation of Dynamic Games with Unknown Information Structure **42**
2.1 Introduction 42
2.2 Theory of Markov Correlated Equilibrium 45
2.3 Informationally Robust Identification 54
2.4 Estimation 59
2.5 Inference 61
2.6 Empirical Application 63
2.7 Conclusion 69

3 The Role of Racial Concordance in Doctor Referral Decisions **71**
3.1 Introduction 71
3.2 Data and Background 73
3.3 Preliminary Evidence of Racially Concordant Referrals 77
3.4 Model of Doctors' Referral Decisions 80

3.5 Results . . . 83
3.6 Counterfactual Analysis . . . 86
3.7 Conclusion . . . 88

Bibliography **89**

A Appendix to Chapter 1 **98**
A.1 Proof: Equilibrium for Two-agent Example . . . 98
A.2 Step 2 of Identification: Asymmetric Case . . . 101
A.3 Details on the Estimation . . . 103
A.4 Simulating the Status Quo and Model Fit . . . 106
A.5 Appendix Tables . . . 108

B Appendix to Chapter 2 **110**
B.1 Proofs . . . 110
B.2 Interpretation of Counterfactual Analysis . . . 113

C Appendix to Chapter 3 **115**
C.1 Other Summary Statistics . . . 115
C.2 Other Preliminary Evidence . . . 116
C.3 Other Structural Estimates . . . 118
C.4 Other Counterfactual Figures . . . 119

List of Figures

1.1 Screenshot of the Q&A forum 8
1.2 Example: Bid Function 13
1.3 CDF of arrival timing 30
1.4 Value of entry and entry probability 30
1.5 Value of disclosure and disclosure probability 33
1.6 Value of disclosure 34
1.7 Value of entry 35
1.8 *Ex-ante* value of arrival timing 36
1.9 Changes in auction outcomes under different counterfactuals 39

2.1 Average number of active firms over time 53
2.2 Projections of the identified set 58
2.3 Number of Starbucks and Dunkin' stores in the US over time 63
2.4 Growth of Starbucks and Dunkin' 64
2.5 Firm entries in counterfactual equilibria 69

3.1 Number of THA/TKA cases and revisions 74
3.2 Number of THA/TKA referrals 75
3.3 Number of THA/TKA referring doctors and performing specialists by race . . . 75
3.4 Number of referrals by HRR in 2019 78
3.5 Race R specialist referral share against their population share by doctor-patient race group in 2019 79

A.1 Model Fit 107

C.1 Race R orthopedist referral share against their population share by doctor-patient race group in 2019 116
C.2 Race R neurosurgeon referral share against their population share by doctor-patient race group in 2019 117
C.3 Distribution of Expected Health Outcome: Baseline and Counterfactual Scenarios 119
C.4 Distribution of Travel Distance: Baseline and Counterfactual Scenarios 120

List of Tables

1.1 Summary Statistics 9
1.2 Presence of question and entry probability 15
1.3 Timing of questions and best bid from opponent 17
1.4 Estimated parameters of the model 31
1.5 Distribution of Construction Costs: by Construction Types 31
1.6 Distribution of opponents' best bid: Log-normal 32
1.7 Description of the Counterfactuals 36

2.1 Probability of entry in three Markov perfect equilibria 53
2.2 Markov correlated equilibrium identified sets 58
2.3 Summary statistics of 35,686 coffee chain markets from 2003 to 2017 65
2.4 Structural parameter estimates 67

3.1 Summary statistics of THA/TKA patients, doctors, and specialists (1999–2019) 77
3.2 Summary statistics related to referral choice set (1999–2019) 78
3.3 THA/TKA referral share regressions (1999–2019) 81
3.4 THA/TKA outcome regressions 84
3.5 Distribution of specialist-specific outcome effects ξ_k by specialist race group . . . 85
3.6 Estimates of structural parameters for THA/TKA referral decision 85
3.7 Choice probabilities : Pr(Referred to Race R Specialist$|\cdot$) 86
3.8 Change in mean of Expected Revision Rate (%) 87
3.9 Change in mean of Travel Distance (in kilometers) 87

A.1 Secret reserve price 108
A.2 History Transition 109

C.1 Average THA/TKA patient comorbidities (1999–2019) 115
C.2 Coefficients on patient comorbidities for THA/TKA outcome regressions 118

Acknowledgments

First and foremost, I would like to thank my advisors, Ben Handel, Matthew Backus, and Kei Kawai for their unwavering support and guidance throughout my graduate studies. Ben, my primary advisor, has been instrumental at every stage of research. His encouragement, insightful suggestions, and provision of a fruitful environment for research have been invaluable. It has been a privilege to collaborate with Ben on research projects. I have learned a lot from him, especially how to form research questions from empirical facts. I am also deeply grateful to Matt for his numerous comments and invaluable advice. His guidance has always helped me identify what is and could be interesting in my work, and what needs to be improved. His enthusiasm has been a constant source of motivation. I would like to express my sincere appreciation to Kei for being an exceptional mentor, bringing me inspiration for research, and his profound influence on my interest in Industrial Organization, especially topics related to auctions. My approach to formulating research questions is deeply inspired by him. Together, they have been exceptional mentors and role models. I aspire to follow them in both my research and mentorship as I continue my journey as an economist.

I would also like to express my sincere appreciation to faculty members at Berkeley who have offered support and valuable feedback throughout my time here: Nano Barahona, Stefano DellaVigna, Matthew Grennan, Yuichiro Kamada, Jon Kolstad, Sarah Moshary, Olivia Natan, Christine Parlour, Demian Pouzo, Cailin Slattery, Carolyn Stein, and Steven Tadelis. I am especially grateful to Yuichiro for his mentorship and insightful feedback on theory-related work. Our interactions have been both intellectually enriching and personally meaningful. I would also like to thank our graduate advisor, Janene Vernard, for the support.

I am also deeply thankful to the faculty at the University of Tokyo, Yasuki Kobayashi, Hidehiko Ichimura, and Toshiaki Iizuka, for their encouragement and support when I was considering a shift in my academic focus from Public Health to Economics. Their guidance played a pivotal role in that decision, and I would not be where I am today without their support.

I am also thankful for my wonderful friends and colleagues at Berkeley. I am especially grateful to Jun Aoyagi and Akiko Aoyagi for their kindness, which allowed me to get through the pandemic. Special thanks goes to Catherine Che, Bryan Chu, Yikun Jiang, Grady Killeen, Kristy Kim, Louis Luther, Josef Nagelschmidt, Carlos Paramo, Nadav Tadelis, Takanao Tanaka, Dan Tavrov, Kiyotaka Yageta, and Mingduo Zhao for their friendship. Beyond Berkeley, I would like to thank my friend and co-author, Konan Hara. Without his presence and influence, I might never have considered pursuing research in Economics—something that brings me great joy today. I am also grateful to my co-author, Paul Koh, for his brilliance, kindness, and the opportunity to collaborate.

Finally, I extend my deepest gratitude to my family for their encouragement, care and support. I thank my wife, Yui; my sister, Ayaki; my parents, Toshifumi and Makiko; and my grandparents, Taizo, Koji, and Saeko–each of whom has been a role model to me. I am profoundly indebted to my family for the countless opportunities they have offered me.

Chapter 1

Entry Deterrence in Procurement Auctions

1.1 Introduction

When multiple firms contemplate entering a market, there may not be sufficient capacity for the market to profitably accommodate all potential entrants. In such an environment, beliefs about others' entry become crucial. As discussed in Farrell (1987), if a firm can influence the beliefs of other firms regarding its intent to enter, it may compel those other firms to reconsider their own entry decisions. For instance, once all potential entrants believe that a given firm will enter the market, this can benefit the firm since the other firms may then be less inclined to enter. This argument highlights the importance of managing information about entry when designing markets.

In attempting to influence the beliefs of rival firms and deter their entry, it is common for firms to publicly announce their intent to enter the market. For example, a firm may make a pre-announcement on releasing new products for this purpose. In the early 1990s, Microsoft was accused of making product pre-announcements just for the purpose of deterring competitors from entering. The district court judge noted that "Microsoft could unfairly hold onto this [dominant] position with aggressive pre-announcements of new products in the face of the introduction of possibly superior competitive products."[1] Although strategic entry deterrence through disclosure raises concerns from an antitrust perspective, there is a notable lack of empirical research investigating its effects on market outcomes.

In this paper, we investigate how entry disclosure affects auction outcomes by studying procurement auctions conducted by the Montana Department of Transportation (MDOT). Public procurement is an economically important setup, with global spending reaching $11 trillion in 2019, equivalent to 12% of the world's GDP (Bosio et al., 2022). A notable feature of the auctions that we study is that there is a designated online forum on MDOT's website,

[1]The ruling of Judge Stanley Sporkin in Civil Action No. 94-1564 (United States of America v.s. Microsoft Corporation 1995).

where potential bidders can post questions about the project being let. The questions, the identity of the firm asking, the posting time, as well as MDOT's responses, are all publicly accessible information. The unique feature is that the forum gets continuously updated: questions become publicly visible almost immediately after posting. Since posting a question on the forum typically requires a firm to have invested some time in reviewing the project plan, posting a question on the online forum serves as an entry disclosure. Indeed, over 99% of the questions are posted by actual entrants.

To understand how participating firms perceive the Q&A forum, we conducted interviews with the participating firms. Their responses reveal that the firms indeed perceive that questions are posted in a strategic manner and not always intended to gather information about the project:

> "*There is always a strategical consideration to the questions we ask and is not solely determined by us needing the information. It can be gamesmanship with the other bidders.*"

Moreover, their claim indicates that they take the questions as a *credible* signal for a firm entering the auction:

> "*It's safe to assume that contractors would not be asking questions unless they are going to bid the project,*"

which aligns with the fact that over 99% of the questions are posted by the actual entrants. These claims support the idea of considering the Q&A forum as a disclosure device, which forms the foundation of the paper.

In our setup, entry disclosure has two distinct and competing effects from a firm's perspective. First, as noted earlier, a firm can alter opponents' beliefs by disclosing, thereby reducing their expected profits from entry, which consequently leads to less entry from other firms. On the other hand, because the exact set of entrants is unknown at the time of bidding, if a firm discloses its entry, other entrants may bid more aggressively. Thus, entry disclosure can ultimately disadvantage firms that choose to disclose.

The effect of introduction of the Q&A forum on auction outcomes is ambiguous and thus an empirical question. This ambiguity arises because firms may use the forum to coordinate their entry decisions, potentially resulting in fewer auctions with a small number of entrants and fewer auctions with a large number of entrants. From a policy perspective, a recent survey highlights that DOTs vary in their treatment of entry information, where 40% of them provide the identity of the firms who enter before the bidding happens, while others do not (Liscow et al., 2024). This variation in practices across entities raises the question of how information regarding entry should be managed.

To understand the impact of the option to disclose entry, we construct and estimate a model of a procurement auction with costly entry, wherein firms can post questions on a Q&A forum that serves as an entry disclosure device. Our model consists of two stages: (i) entry and disclosure; and (ii) bidding. In the first stage, firms sequentially arrive at the market.

They make entry decisions based on their private entry costs and information available on the Q&A forum – specifically, the firms are aware of who has disclosed entry. Upon entry, firms draw their construction costs and may strategically choose to post a question on the forum, thereby disclosing its entry. Firms also disclose due to non-strategic reasons with a fixed probability. The second stage occurs after all the entry decisions have been finalized. Entrants submit their bids simultaneously, taking into account all disclosures and their own private construction costs. The bidding procedure is a first-price sealed bid auction, and the bidder with the lowest bid wins.

We establish three key descriptive patterns that illustrate the economic forces of the posted questions, i.e., entry disclosure. The following observations allow us to identify the model primitives. First, we document that the presence of a question on the forum is associated with a lower entry probability among bidders. This is a pattern we would see if disclosures indeed deter entry. Second, entrants submit stronger bids when faced with a greater number of questions. This pattern aligns with the hypothesis that firms' bidding behaviors respond to the information presented on the Q&A forum. Third, the firms who disclose early face weaker best rivals' bid compared to those who never disclose and those who disclose late. In light of our model, while the first two patterns are in play, firms benefit from making early entry disclosures because the effect from the first pattern, entry deterrence effect, dominates the effect from the second pattern. One alternative force that could explain the first pattern presented above is variation in quality of the proposals across auctions, where low quality proposals could impose additional costs to the firms. However, this force cannot explain the other two patterns and thus we believe that this factor is unlikely to be the primary driver of our data.

We show that the primitives of the model are non-parametrically identified from firms' entry, entry disclosure, and bidding behaviors. The primitives we aim to recover are the distribution of firms' arrival timing, entry costs, costs associated with entry disclosure, i.e., posting questions, and construction costs. The primary challenge in identification arises from the fact that arrival timing is only observed for the firms who disclose their entry. We start by recovering firm X's beliefs on the evolution of disclosure history, conditional on X entering at a fixed time point. If firm X discloses their entry, this object is directly identified from the observed patterns. However, if X does not disclose, we cannot identify firm X's beliefs directly from the data, since we do not observe their entry timing. To overcome this problem, we construct a mapping from the observed pattern of disclosures to firm's beliefs on the evolution of disclosure histories. The idea is to treat the setup as a survival analysis with competing risks. Here, the event is an entry disclosure, and the possibility of multiple firms disclosing can be seen as competing risks. Our conditional independence assumption between the firms allows us to identify each firm's duration until they make a disclosure, starting from any time point (Tsiatis, 1975). We can then construct firm X's beliefs on evolution of disclosure history, since each firm's duration until disclosure is known under any history.

Given the firms' beliefs, the primitives – specifically, distribution of firms' arrival timing, entry costs, costs of entry disclosure, and construction costs – are identified through the

following six steps. First, we recover the construction costs and their distribution, following the methodology from Guerre et al. (2000). Second, by considering the expected value from the auction stage and the beliefs on the evolution of disclosure history, we are able to determine the value functions with and without disclosure when firms decide whether to disclose. Thus, value of disclosure is identified for each history and construction cost. Third, by exploiting variation in the amount of disclosures and values of disclosure at the same history but under different construction costs, we can identify the distribution of disclosure costs. Fourth, given the knowledge of values with and without disclosures, along with the distribution of disclosure costs, value of entry is identified for each disclosure history. Fifth, by exploiting variation in the amount of disclosures, value of entry, and value of disclosure at the same time under different disclosure histories, we can identify the distribution of entry costs. Finally, we can identify the distribution of arrival timing by comparing the amount of disclosures across different time points.

Given our estimates, we can quantify the value of disclosure. First, we show that disclosure is beneficial for firms at the beginning of the entry period, but becomes detrimental toward the end. For a bidder with median construction costs, the value of disclosure is 1.5% of the estimated project cost at the beginning, whereas it becomes costly by the end. The intuition behind this finding is that if a bidder enters early and discloses, they can deter entry from others, even though remaining entrants may bid more aggressively. In our scenario, the deterrence effect dominates. However, if a bidder enters late and discloses, the deterrence effect diminishes since there are fewer potential entrants remaining on the sideline. As a result, aggressive bidding from other entrants negatively impacts the late-disclosing bidder. Next, stronger bidders who have lower construction costs derive larger values from disclosure. At the beginning of the entry period, the value of disclosure is 2.1% of the estimated project cost for a bidder whose cost is at the 25-th percentile, while the value is 0.7% of the estimated project cost for a bidder at the 75-th percentile. This result indicates that entry disclosure also serves as a signal of a bidder's strength – construction cost.

We also quantify the value of entry, and show how it changes by the presence of disclosures. First, we find that the value of entry increases as we progress to later periods, holding the number of disclosures available at the firms' arrival time fixed. Under the case where there are no disclosures, value of entry is 9.8% of the estimated project cost at the beginning of the entry period and rises to 10.5% by the end. At a fixed time point, an increase in the number of disclosures decreases firm's value of entry, resulting in a reduction of their entry probability by 4–6%. These findings indicate that both arrival timing and the disclosures firms face have a significant impact on entry decisions. Finally, we show that the expected profit from arriving at the end is 7% lower than the the case when a firm arrives at the beginning. Early arrival allows firms to capture greater gains from disclosures. Conversely, firms arriving late can make more efficient entry decisions due to increased information availability. In our setup, the former effect dominates.

In our counterfactual analysis, we compare equilibrium auction outcomes under three alternative scenarios: shutdown of the forum, where the Q&A forum is never made public; last minute disclosure, where the forum becomes public after the entry period but before

the bidding occurs; and the status quo, where the current Q&A forum is available. We use the first scenario, in which the forum is shut down as our benchmark for exposition. Our objective is to understand the effects of entry disclosure, which operate through two channels: entry deterrence and provision of additional information at the bidding stage.

First, we make the Q&A forum public after the entry period ends. In this scenario, entry disclosures affect outcomes solely through the second channel – providing additional information when firms bid. Firms cannot deter entry, as the information only becomes public after firms have made their entry decisions. In this case, we observe a 0.8% increase in auctioneer's payment, a 1.4% increase in winner's construction cost, and a 3.2% increase in total entry costs, compared to the shutdown scenario. Firms still make disclosures, but only due to exogenous reasons. McAfee and McMillan (1987) and Harstad et al. (1990) have pointed out that uncertainty about firm's entry does not affect the auctioneer's payment, when bidders are risk-neutral. In line with this result, we only see a small change in the auctioneer's payment. This small change comes from asymmetry among the bidders (Maskin and Riley, 2000), which also creates efficiency loss in the winner's cost. Consider the following example: there are two entrants, firms X and Y. Firm X discloses its entry, while firm Y remains silent. Firm Y will adopt a more aggressive bidding strategy than firm X since they are sure about facing a competitor. Consequently, firm Y may win some auctions even when firm Y has a higher construction cost than firm X. This inefficiency does not arise when firms are in a symmetric position and employ monotone symmetric strategies. The asymmetry created by disclosures leads to inefficiency. Increase in auctioneer's payment dominates the increase in the winner's cost, and thus total entry increases in equilibrium.

Next, we implement the current Q&A forum, allowing firms to deter others' entry by disclosing their intent to enter. Under the current forum, entry disclosures affect outcomes through both channels – entry deterrence and additional information at the bidding stage. In this case, we observe a more significant impact: relative to the benchmark, there is a 6.3% decrease in auctioneer's payment, a 4.5% increase in winner's construction cost, and an 11.1% decrease in total entry cost. The availability of the forum allows the firms to coordinate on their entry to some extent. Coordination among the firms reduces the number of auctions where there is no competition, i.e., only one entrant. Since the possibility of a firm becoming the only bidder is a strong force that increases the auctioneer's payment, coordination among the firms helps the auctioneer by reducing their payment. Another key observation here is that the stronger bidders with small construction costs are more likely to disclose their entry. While firms can deter entry through disclosure, they forfeit information rents associated with their entry status *and* strength. As a result, we see a decrease in auctioneer's payment. Moreover, this creates a "stronger" asymmetry among the entrants: firms who disclose will be bidding for certain and strong, while firms who stay silent may not be present and weak even if they do enter. Consequently, winner's cost increases.

In summary, the current platform introduces a new dimension regarding firms' types: arrival time. When a firm arrives early, to take advantage of this new dimension, they disclose their entry status to deter other firms' entry even at the cost of sacrificing information rents. Furthermore, such information transmission allows the firms to coordinate their entry be-

havior, eliminating *ex-post* non-competitive auctions. Consequently, this leads to a decrease in auctioneer's payment. Furthermore, disclosures put firms into asymmetric positions, resulting in inefficiency in terms of the winner's cost. Together, we find that the existence of the Q&A forum, which allows the firms to send out some information, has a significant impact on auction outcomes. More broadly, these results suggest that market designers must exercise caution in how information is transmitted before agents take actions.

Related Literature The paper contributes to two strands of literature—the literature on strategic entry deterrence and the literature on costly entry into auctions.

The paper provides an empirical equilibrium analysis to test how strategic entry deterrence can affect market outcomes, taking entry disclosure as a tool to deter entry. A significant amount of theoretical work on strategic entry deterrence has been carried out, e.g., Spence (1977), Dixit (1979), and Milgrom and Roberts (1982a,b).[2] Farrell (1987) illustrates how communication can allow firms to coordinate their entry. However, empirical work on strategic entry deterrence is still limited. Goolsbee and Syverson (2008) and Sweeting et al. (2020) study how limit pricing by the incumbent affects entry behavior in the airline market. Scott Morton (2000) and Ellison and Ellison (2011) studies strategic investment, such as advertisement, to deter entry in the pharmaceutical market. Ely and Hossain (2009) studies the effects of early period bidding in online auctions. Although they find a similar result to our paper that early period bidding deters entry but causes more aggressive bidding from the entrants, there are two important distinctions. First, Ely and Hossain (2009) tests for such effect by experimentally placing bids, while our analysis analyzes the effect of entry disclosure, which arises as an equilibrium behavior. Next, they study a second-price auction setup, while ours is a first-price auction. In second-price auctions, more aggressive bidding due to entry disclosure is not a pattern we would expect under a private-value framework, since bidding their own value would be an undominated strategy for the bidders. In contrast, entry disclosure may cause more aggressive bidding from others under our setup, sealed-bid first-price auctions with private values.

The paper also relates to the literature on costly entry into auctions. Overall, the literature has pointed out the importance of incorporating entry costs in analysis of auctions. Ye (2007) and Quint and Hendricks (2018) theoretically study indicative bidding; Bhattacharya et al. (2014) studies the effect of introducing an entry rights auction; De Silva et al. (2008) studies the effect of releasing information about seller's valuation on bidding in procurement auctions; Krasnokutskaya and Seim (2011) studies how the introduction of bid preference program affects firms entry and bid decisions; and Gentry and Li (2014) studies non-parametric identification of an auction game with selective entry. The paper also studies a setting where entry is costly, but is the first to empirically study how entry disclosure can deter entry from others in first-price auctions with costly entry.

[2]Wilson (1992) provides a review on this topic.

1.2 Institutional Background and Data

Institutional Background

We describe the letting process of procurement auctions conducted by the Montana Department of Transportation (MDOT). MDOT uses sealed-bid first price auctions to award construction projects. The set of firms who participate in bidding will not get disclosed by MDOT until the final auction result is announced.

MDOT advertises projects four weeks prior to the bidding date, providing detailed specifications of each project. On the same day as the advertisement, a Q&A forum is launched on MDOT's website. On this forum, firms can post questions about the project, and MDOT provides answers to the posted questions. The questions become publicly visible when the firms post them, subject to a quick review by the MDOT. Answers from MDOT are provided within two days in most cases. The forum displays the time at which the question got posted, the name of the company, the contact person, the question, and the corresponding answer. Figure 1.1 presents a screenshot of the forum. The forum remains active until three days before the bidding window closes. While other public procurement auctions also accept questions from the firms, the unique feature here is that this forum gets continuously updated along with identity of the firms who posted questions and a timestamp.

The questions posted on the forum vary in value. Some turn out to be valuable, while others offer little additional information. For example, a question pointed out that there is an unnecessary item listed in the contract. MDOT responded by removing the item from the contract and issued an addendum.[3] On the other hand, MDOT responds to some questions by referring firms to existing documents, instructing them to review the relevant sections.[4] This observation is in line with the quote from a firm presented above that some questions are not solely intended to obtain information.

To participate in bidding on a project, firms must prepare the necessary documents and submit them along with their bids. They must also engage in negotiations with subcontractors. These tasks involve significant costs, as they require substantial time and effort. As a result, entry into auctions is inherently costly.[5]

Data

Our data covers projects auctioned between January 2017 and December 2022. For each auction, it includes the project description, location, the engineer's estimate of the total cost of the project, and the submitted bids along with the identity of the bidding firms. Additionally, the dataset contains information from the Q&A forum, including the posted questions, MDOT's responses, identities of the firm who posted the questions, and the timestamps of

[3]For example, see `https://app.mdt.mt.gov/qaf/external/archive/view/493`

[4]For example, see question #1 from `https://app.mdt.mt.gov/qaf/external/archive/view/463`

[5]Costliness of entry into procurement auctions have been pointed out in the literature (e.g., Li and Zheng (2009)).

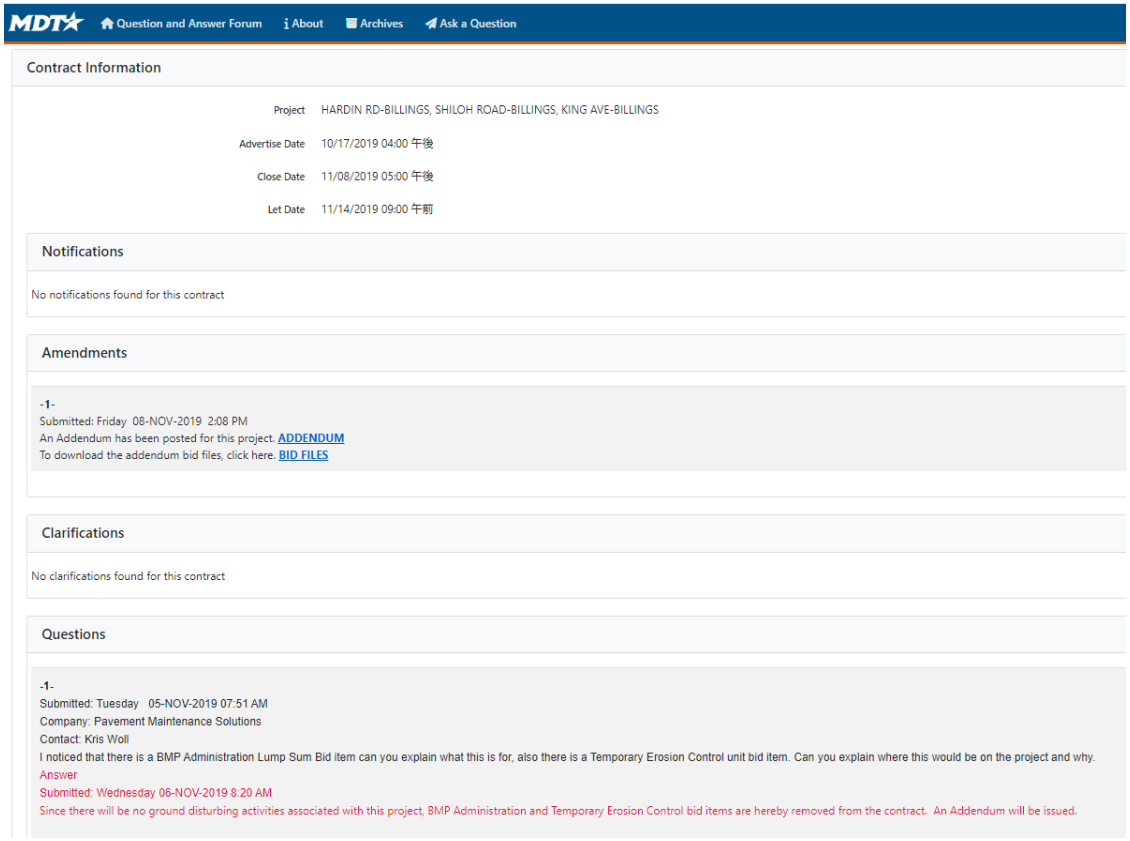


Figure 1.1: Screenshot of the Q&A forum

the posts. During the sample period, 592 projects were advertised, while we focus on 434 projects whose construction reports were available, which allow us to identify the type of construction of the projects.

Table 1.1 presents summary statistics of the auctions. The median engineer's estimate is approximately \$1.30 million, while the median winning bid is around \$1.22 million. For our analysis, we will normalize the bids by the engineer's estimate. The median normalized winning bid is 3.5% below the engineer's estimate. MDOT reserves the rights to reject all the bids, and 16 auctions during the sample period experienced such rejections.[6] On average, we have three entrants. We define potential entrants as firms that entered in at least one auction within the same district $\times$ type of construction pair during the sample period. A typical auction has 12 potential entrants. Regarding the Q&A forum, we observe slightly fewer than one question per auction on average.[7] There is some variety in the types of projects, with overlay projects being the most common (18%).[8] Project distribution is relatively balanced

[6]If bids are rejected, the project may be revised and advertised at a later date.

[7]We only include the questions submitted by the potential entrants. Moreover, there are some cases where a firm posts multiple questions. We only keep the first questions from such firm in our dataset.

[8]We follow the categorization of types of construction provided in the construction reports provided by

Table 1.1: Summary Statistics

	Mean	Standard deviation	10th percentile	Median	90th percentile
Engineer's estimate ($000)	2,949	4,315	144	1,297	8,597
Lowest bid ($000)	3,022	4,702	154	1,225	8,382
Lowest bid / Engineer's estimate	1.021	0.314	0.750	0.965	1.320
#Entrants	2.82	1.50	1	3	5
#Potential entrants	12.44	5.62	4	12	20
#Questions	0.83	0.97	0	1	2
Type of projects	N	percent			
Bridge construction	51	11.8			
Overlay	78	18.0			
Reconstruction	46	10.6			
Safety	67	15.4			
Others	192	44.2			
Districts	N	percent			
Missoula	94	21.7			
Butte	76	17.5			
Great Falls	113	26.0			
Glendive	73	16.8			
Billings	78	18.0			

Note: Total number of projects is 434. There were 5 auctions without an entrant.

across districts, although the Great Falls district accounts for the largest share (26%).[9]

1.3 Model

In this section, we develop a model of a procurement auction with costly entry and the option for firms to disclose their entry. The auctioneer seeks to procure a project through a first-price auction. There are N potential bidders who may choose to participate in bidding. We denote the set of potential bidders as $\mathcal{N} = \{1, \cdots, N\}$. In our empirical application, the questions submitted to the Q&A forum serve as entry disclosures.

The model consists of two stages: (i) entry and disclosure; and (ii) bidding. In the first stage, firms sequentially arrive at the market randomly without knowledge of others' arrival times. When they arrive at the market, firms observe the disclosures that have been made, make decisions on entry, and choose whether to disclose if they enter. Once the first stage concludes, the entrants proceed to the second stage, which involves bidding. In this stage, firms observe the complete history of disclosures, and place bids simultaneously.

MDOT. Some projects fall under multiple categories and if so we assign the project to the more popular type.

[9]We split the state into five districts, following the coverage of five MDOT district offices. See `https://www.mdt.mt.gov/contact/organization/districts.aspx`

First stage An auction is announced and the Q&A forum, serving as the disclosure device, becomes available at $t = 0$. The disclosure device closes at $t = T$, although disclosures will remain observable after its closure. Each potential bidder $i \in \mathcal{N}$ draws $\tau_i \in [0, T]$ from the distribution F_τ, which represents the time at which bidder i decides whether to participate in the auction.[10] At $t = \tau_i$, firm i arrives at the market, observes the disclosure history h^{τ_i}, and draw its entry cost c_i^E from distribution F_E. The disclosure history h^t is public information and records the time at which questions are posted as well as the the identities of the posting firms, up to time t. We denote the set of all time-t histories as H^t. The entry cost, c_i^E encompasses the cost of reviewing the project plan, assessing required materials and labor, negotiating with subcontractors, and arriving at a cost estimate. Firm i may choose to enter the auction by paying the entry cost c_i^E or opt to remain out without incurring any costs. We denote the firm i's entry strategy as: $\chi_{i,\tau}$: $\chi_{i,\tau}(h^{\tau_i}, c_i^E) \mapsto a_i^E \in \{0, 1\}$.

If firm i decides to enter the auction, at the same time $t = \tau_i$, it draws its construction cost c_i, and faces an opportunity to disclose its entry. With probability p^Q, firm i faces a need to disclose and always discloses without paying any additional cost. During the review of the project plan, issues may arise that prevent the firms from making progress in the process. This part reflects such scenario and assume that it happens with probability p^Q. With the other probability $1-p^Q$, firm i may opt for costly disclosures by paying a disclosure cost c_i^Q, which follows the distribution F_Q. This cost can be viewed as the expense associated with formulating an appropriate question, as well as a potential reputation cost. Only firms that have entered the auction are allowed to make disclosures. We denote firm i's disclosure strategy as: $\iota_{i,\tau}(h^{\tau_i}, c_i, c_i^Q) \mapsto a_i^Q \in \{0, 1\}$. If firm i discloses, disclosure history h^τ is updated accordingly.

Second stage After the forum closes at $t = T$, entrants – the firms who have entered – participate in bidding. The auction format is a sealed-bid first price auction. Before placing their bids, the entrants observe the complete disclosure history h^T. Given h^T, the entrants submit their bids b_i simultaneously. We denote the firm i's bidding strategy as: $\beta_i(h^T, c_i) \mapsto b_i \in \mathbb{R}^+$.

Payoff The payoff π_i for firm i is given by:

$$\pi_i = (b_i - c_i)\mathbb{1}\{i \text{ wins}\} - c_i^E a_i^E - c_i^Q a_i^Q,$$

where the first term represents gains from the auction, the second accounts for the entry cost, and the third is the disclosure cost.

Assumption on firms' types As described above, each firm i's type is characterized by the tuple $(\tau_i, c_i^E, c_i^Q, c_i)$, which includes arrival timing, entry cost, disclosure cost, and con-

[10] We do not rule out the possibility that the distributions that appear in the model are common across firms. Although we assume that these distributions are common in our empirical application, identification argument still works with across-firm heterogeneity in these distributions. Here, we omit the index i on these distributions for the sake of exposition.

struction cost. We assume that these four random variables are mutually independent, with draws across firms being independent as well. Furthermore, we assume that the distributions F_τ, F_E, F_Q, and F_c are continuously differentiable.

Equilibrium We consider the Perfect Bayesian Equilibrium of the game presented above. Equilibrium consists of firms' strategy profile $(\chi_{i,\tau}, \iota_{i,\tau}, \beta_i)$ such that:

1. firm i enters ($a_i^E = 1$) if and only if its expected profit from entry exceeds the entry cost c_i^E:
$$\mathbb{E}\left[(b_i - c_i)\mathbb{1}\{i \text{ wins}\} - c_i^Q a_i^Q | h^{\tau_i}, a_i^E = 1\right] > c_i^E$$

2. firm i costly discloses ($a_i^Q = 1$) if and only if its expected gain from entry exceeds the disclosure cost c_i^Q
$$\mathbb{E}\left[(b_i - c_i)\mathbb{1}\{i \text{ wins}\}|h^{\tau_i}, a_i^E = 1, a_i^Q = 1\right] - \mathbb{E}\left[(b_i - c_i)\mathbb{1}\{i \text{ wins}\}|h^{\tau_i}, a_i^E = 1, a_i^Q = 0\right] > c_i^Q$$

3. firm i bids b_i that maximizes its expected profit conditional on the complete disclosure history h^T and construction cost c_i
$$b_i = \underset{b}{\text{argmax}}\ (b - c_i)\Pr(i \text{ wins}|h^T, b).$$

In addition, firms have consistent beliefs given the strategy profile. We assume that an equilibrium exists, and if there are multiple equilibria, we assume that one equilibrium is selected and played.

Discussion of the model In our setup, entry disclosures occur through the posting of questions. Consequently, firms may sometimes need to post questions to address issues that arise during their preparation for bidding, even if their primary motivation is not to disclose their entry. Our model does incorporate this feature. However, it does not account for informational spillovers that may benefit other firms. Our estimates indicate that the strongest bid from the opponents tend to be weaker when a firm discloses its entry, holding others' disclosure activities fixed. If information spillovers had first-order effects that outweighed the effects of entry disclosures, the sign of this effect would be the opposite. Therefore, we believe that this is not a primary concern in our model, though informational spillovers may exist.

Example: Two firms

Here, we provide a simple example with two firms. The purpose of this example is to illustrate the main economic forces that emerge from the option to disclose entry.

There are two firms i and j who are ex-ante symmetric. We assume that the distribution of arrival timings follow $F_\tau \sim U[0,1]$,[11] the distribution of construction costs follows $F_c \sim U[0,1]$, and disclosure can be made at no cost, i.e, $c_i^Q = c_j^Q = 0$. Moreover, firms are never forced to disclose their entry, $p^Q = 0$. We leave the distribution of entry costs F_E to be unspecified at this point. Suppose that there is a reserve price $R = 1$.

For some entry cost distribution F_E, there exists an equilibrium that consists of the following strategies:

1. Second stage: Bidding
   - If h^T includes two disclosures or none: bid $\beta_i(h^T, c_i) = (1 + c_i)/2$.
   - If h^T includes one disclosure and that is from i, firm i bids $\beta_1(h^T, c)$ such that $\beta_1^{-1}(b; h^T)$ equals to

$$1 - \frac{1}{(b - \frac{1}{r})\left(\frac{r^2}{(1-r)^2}\log\left(\frac{1-b}{\frac{1}{r}-b}\right) - \frac{r}{1-r}\frac{1}{b-\frac{1}{r}} - r(1+r) - \frac{2r^2}{(1-r)^2}\log(r) - \frac{r^2(1+r)}{1-r}\right)}$$

     where β_1^{-1} is the inverse bid function, and r is i's belief on j's entry probability. Firm j bids $\beta_2(h^T, c)$ such that $\beta_2^{-1}(b; h^T)$ equals to

$$\frac{1}{r} - \frac{1}{\frac{r}{1-r} + (b-1)\left(\frac{r^2}{(1-r)^2}\log\left(\frac{\frac{1}{r}-b}{1-b}\right) - (1+r) + \frac{2r^2}{(1-r)^2}\log(r) + \frac{1+r}{1-r}\right)}$$

     where β_2^{-1} is the inverse bid function.

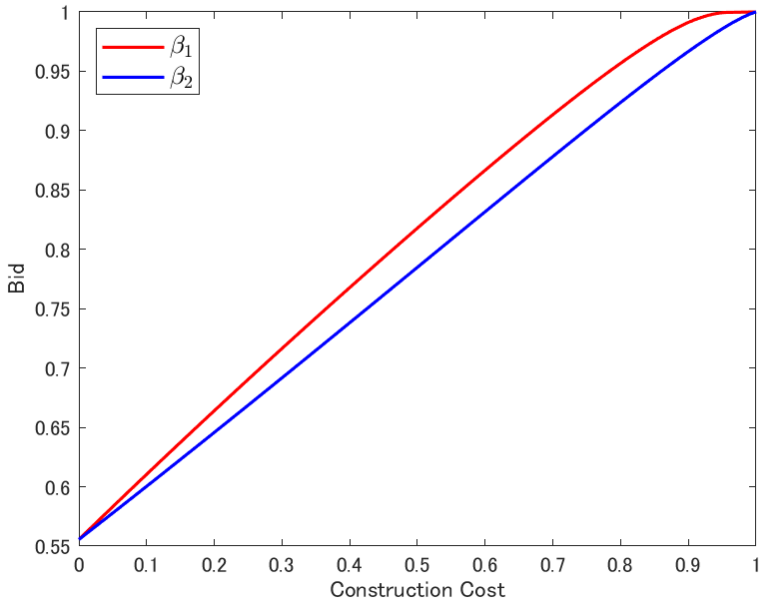
2. First stage: Entry
   - If h^τ does not include any disclosures, firms always enter $\chi_{i,\tau}(h^\tau, c_i^E) = 1$.
   - If h^τ includes one disclosure, firm enters $\chi_{i,\tau}(h^\tau, c_i^E) = 1$ if and only if $\pi_2 > c_i^E$, where π_2 is the expected profit from the bidding stage.

3. First stage: Disclosure
   - If h^τ does not include any disclosures, firms always disclose $\iota_{i,\tau}(h^\tau, c_i) = 1$.
   - If h^τ includes one disclosure, firms never disclose $\iota_{i,\tau}(h^\tau, c_i) = 0$,

and consistent beliefs, given these strategies. The construction for the bidding part follows Kaplan and Zamir (2012). The proof is given in Appendix A.1.

The required conditions for the entry cost distribution F_E are:

[11] This assumption is made for simplicity. It can be any continuous distribution without a mass.



Notes: This figure shows the bidding functions employed in our two firm example, for the case $r = 0.8$.

- firms always enter under no disclosure:

$$F_E(\pi_1) = 1$$

  where π_1 is the expected profit from the bidding stage for the case where you have disclosed but your opponent has not.

- firms enter with probability r when there is one disclosure:

$$F_E(\pi_2) = r$$

  where π_2 is the expected profit from the bidding stage for the case where your opponent has disclosed but you have not.

This equilibrium presents multiple key observations that may arise from an environment with entry disclosures. First, entry disclosure deters subsequent firms from entering the auction after a disclosure has been made. The firm that arrives first always discloses, which decreases the entry value for the next arriving firm, thereby reducing its probability of entry.

Second, entry disclosure compels the other firm to bid more aggressively. To illustrate this, we can compare the scenario in which the first firm discloses with the scenario in which it does not. In the former case, the second firm employs strategy β_2, while in the latter case, it employs strategy β_1. As shown in Figure 1.2, β_2 represents a more aggressive strategy.[12]

[12]This claim holds for any entry probability of the second firm, r.

Therefore, conditional on the other firm entering, the other firm bids more aggressively when faced with a disclosure.

Third, early disclosures are valuable, while late disclosures are detrimental. We observe that only the firm who arrives first chooses to disclose, as disclosure is beneficial for that firm. In contrast, the firm that arrives second finds that disclosing only prompts the first firm to bid more aggressively, leading it to refrain from making a disclosure.

Here, through a simple example, we have shown three key observations that are expected to arise in an environment where entry disclosure is an available option. In the following section, we provide evidence that our data aligns with these key observations.

1.4 Descriptive Evidence

We use the data to establish three empirical facts that highlight the trade-offs firms face when considering disclosures: disclosure may deter entry from other firms, and it may also compel other entrants to bid more aggressively. First, we examine the relationship between the presence of a question on the forum and firms' entry probability. Second, we show how bids are related to the number of questions that a firm encounters. Finally, we analyze how the timing of questions and bids from opponents are related. Our findings align with the patterns illustrated in the two-firm example presented in Section 1.3.

Fact 1: Presence of a question and entry probability

As we have seen in Section 1.3, entry disclosures may deter entry from others. If a question serves as an entry disclosure and deter entry from others, presence of a question on the forum would reduce the probability of entry from other firms. To assess the relationship between the presence of a question and entry probability, we run the following regression:

$$\mathbb{1}\{\text{firm enters}\}_{ia} = \beta_0 + \beta_1 \mathbb{1}\{\text{question is posted from an opponent}\}_{ia} + \beta^X X_a + \varepsilon_{ia} \quad (1.1)$$

where i denotes the firm and a denotes the auction. Auction-level characteristics X_a include the number of potential bidders, type of construction, and district where the project is located. Our primary interest is in the sign of β_1. The first column from table 1.2 presents the results from this regression. We observe that presence of a question from opponents is associated with a 3.4 percentage point decrease in entry probability, and this association is significant at the 5% level. However, concerns may arise regarding within-auction variation influencing this result. Specifically, firms that post questions always enter, which mechanically results in fewer questions being visible on the forum within an auction. To mitigate such concern, we restrict our sample to firms that do not post a question. The second column of Table 1.2 shows the results for this restricted sample. We find that the result is mostly unchanged.

Table 1.2: Presence of question and entry probability

	Dependent variable: Entry			
	(1)	(2)	(3)	(4)
Sample	all	only not posted	all	only not posted
Q from opponent is present	-0.034	- 0.041		
	(0.014)	(0.012)		
Number of Qs from opponent				
1			-0.030	-0.039
			(0.017)	(0.013)
≥ 2			-0.042	-0.044
			(0.020)	(0.015)
Auction-level characteristics	Yes	Yes	Yes	Yes
N	5,397	5,042	5,397	5,042

Note: Results in columns (1) and (3) is based on the entire sample of potential entrants. Results in columns (2) and (4) is based on the sample of potential entrants who have not posted a question. Standard errors are clustered at the auction level.

We also investigate how this relationship varies with the number of questions firms observe on the forum. To do this, we run the following regression:

$$\begin{aligned}\mathbb{1}\{\text{firm enters}\}_{ia} = \beta_0 &+ \beta_1 \mathbb{1}\{\text{One question is posted from an opponent}\}_{ia} \\ &+ \beta_2 \mathbb{1}\{\text{Two or more questions are posted from an opponent}\}_{ia} \\ &+ \beta^X X_a + \varepsilon_{ia}.\end{aligned}$$

Our interest lies in the signs and the relative magnitudes of β_1 and β_2. We again observe a negative relation between presence of opponents' questions and entry probability. Although the strength of this relationship by the number of questions is not statistically different (β_1 and β_2), our point estimate for the coefficient on seeing two or more questions is larger than the coefficient on seeing one question.

Together, the relationship between presence of opponents' questions and firms' entry probability in our data aligns with the hypothesis that questions deter entry from others.

Fact 2: Questions from opponents and bid

If firms incorporate the disclosures into their bids, they are likely to be placing stronger bids if they observe disclosures from opponents. To assess the relationship between presence of questions from opponents and a firm's bid, we run the following regression:

$$b_{i,a} = \beta_0 + \beta_1 \mathbb{1}\{\#\text{ questions from opponents}\}_{ia} + \beta^X X_a + \varepsilon_{ia},$$

where we control for auction-level characteristics X_a.

We find that $\hat{\beta}_1 = -0.039$ (S.E. = 0.013), which is consistent with our hypothesis.[13] Firms facing more questions from opponents tend to submit stronger bids compared to those who facing fewer questions; in fact, seeing one additional question corresponds to placing a stronger bid by 3.9% of the engineer's estimate. This pattern reinforces our hypothesis that questions function as entry disclosures and that firms incorporate this information into their bidding behavior.

Fact 3: Timing of questions and bids from opponents

Disclosures made in early periods may have strong deterrence effects, while other entrants may bid more aggressively compared to the case where the firm had stayed silent. On the other hand, late disclosures may be detrimental because they may lack deterrence effects, while other entrants still bid aggressively.

Now, note that from one entrant's point of view, their profit depends on the best bid among their opponents in a first price auction. If the deterrence effect from early disclosures is strong enough, we would expect to observe weaker best bid from opponents compared to scenarios where the firm does not disclose or disclose late. To evaluate the relationship between timing of question postings and best bid from opponents, we run the regression:

$$\wedge b_{-i,a} = \beta_0 + \beta_1 \mathbb{1}\{\text{posted a question}\}_{ia} + \beta_2 \mathbb{1}\{\text{posted a question}\}_{ia} \times \tau_{ia} + \beta^X X_{ia} + \varepsilon_{ia}.$$

where $\wedge b_{-i,a}$ is the best bid among opponents and $\tau_{ia} \in [0, 1]$ denotes the timing of the question posting.[14] The control variables X_{ia} include auction-level characteristics and number of questions posted by opponents. We normalize the period at which firms can post questions to the interval $[0, 1]$. We expect $\beta_1 > 0$, $\beta_2 < 0$, and $\beta_1 + \beta_2 < 0$ for our estimates to be consistent with the hypothesis presented above.

Table 1.3 reports the results from this regression. First, we observe that firms who post questions at $t = 0$ face a weaker best bid from opponents compared to those who never post by 8.0% of the engineer's estimate ($\beta_1 > 0$). Second, firms that post questions at $t = 0$ face a weaker best bid from opponents compared to those who post at $t = 1$ by 9.1% of the engineer's estimate ($\beta_2 < 0$). Third, our point estimates suggest that firms who post questions at $t = 1$ face a stronger best bid from opponents than those who never post by 1.1% of the engineer's estimate ($\beta_1 + \beta_2 < 0$), although the relation is not statistically significant.

These results are consistent with the hypothesis that: early disclosures are beneficial because there is a strong enough deterrent effect; and late disclosures are detrimental as they lead other entrants bid more aggressively while the deterrent effect becomes weak.

[13] Adding in a dummy variable for firm i posting a question and/or firm i's timing of question posting does not change our estimate in a meaningful way.

[14] The sample for this regression is the entrants who had at least one competing entrant.

Table 1.3: Timing of questions and best bid from opponent

Dependent variable: Best bid from opponent		
	Coefficient	Standard Error
Posted a Q	0.080	0.027
Timing of Q: τ	-0.091	0.040
Controls	Yes	
N	1,144	

Note: Estimation is based on the sample: entrants who had at least one opponent. Standard errors are clustered at the auction level.

Discussion

One may consider an alternative hypothesis: if there is unobserved heterogeneity in the quality (or uncertainty) of the government's proposal across projects and the presence/number of questions acts as a proxy for such quality, we may observe the same pattern as Fact 1. Now, suppose that the presence/number of questions do act as a proxy for quality of the proposals. Then, we would see weaker bids in auctions where questions are posted. However, our findings indicate the opposite trend. While Fact 3 partly addresses this concern, we run the following regression:

$$b_{i,a} = \beta_0 + \beta_1 \mathbb{1}\{\#\text{total questions}\}_a + \beta^X X_a + \varepsilon_{ia}.$$

Our estimate is $\hat{\beta}_1 = -0.025$ (S.E. $= 0.011$), suggesting that the level of bids are stronger in auctions with greater number of questions. This suggests that unobserved heterogeneity in quality of proposals does not seem to be a primary concern in our analysis.

Another alternative hypothesis may posit that if firms who have lower costs tend to arrive earlier, we may observe the same pattern to that described as Fact 2. However, it is important to note that this hypothesis alone cannot explain Fact 1. To further explore this point, although with parametric assumptions, we can estimate the construction costs of each firm without imposing structure on how decisions on entry and question posting are made. We consider two scenarios when we estimate the construction costs for this practice: (i) firms condition their bidding strategies on the question postings; and (ii) firms do not condition their strategies on any question postings.We run the following regression:

$$\hat{c}_{ia} = \beta_0 + \beta_1 \tau_{ia} + \beta^X X_a + \varepsilon_{ia}.$$

where the sample consists of firms that posted a question, $\hat{c}_{ia}$ is the estimated cost, and τ_{ia} is the timing of question posted. If this hypothesis significantly influences the data, we would expect to see $\beta_1 < 0$. However, our estimates under both scenarios do not support this pattern, where $\hat{\beta}_1 = 0.017$ (S.E. $= 0.081$) for scenario (i) and $\hat{\beta}_1 = 0.016$ (S.E. $= 0.079$) for scenario (ii). Therefore, we do not consider this hypothesis to be a primary driver of our data.

1.5 Identification

In this section, we provide a discussion on identification of the model. We identify the model primitives in a sequential manner through the six steps outlined below. The primitives we aim to identify include: the distribution of arrival timing F_τ, entry costs F_E, disclosure costs F_Q, construction costs F_c, each entrant's construction cost c_i and the probability of forced disclosure p^Q. Construction costs are identified from the bidding stage, while the other primitives are identified from variation in entry and disclosure behaviors. It is worth emphasizing that entry timing τ_i is observed for the firms that disclose, but not for the firms that do not disclose.

While the discussion below holds even when conditioning on auction-level characteristics, we omit such expressions for the sake of exposition. Moreover, all the distributions we aim to identify are identified at the firm level. In what follows, we may omit the firm-level index as well.

Step 1. Construction costs c_i and its distribution F_{c_i} First, we aim to identify the construction costs c_i of each bidder and the distribution F_{c_i}. The argument follows the strategy established by Guerre et al. (2000). We impose the following assumption on bidders' strategies:

Assumption 1. *Firm i's bidding strategy is strictly increasing in their construction costs c_i, conditional on the complete disclosure history h^T.*

For each public history h^T, the bidder i's problem at the bidding stage is to maximize their expected value $V_i(h^T, c_i)$:

$$V_i(h^T, c_i) = \max_b (b - c_i) \left(1 - G_{-i}(b|h^T)\right), \tag{1.2}$$

where $G_{-i}(b|h^T)$ denotes the distribution of the lowest rival bid, $\wedge b_{-i}$, conditional on h^T. Note that $G_{-i}(\cdot|h^T)$ is nonparametrically identified, and thus c_i for each entrant and $F_{c_i}(\cdot|h^T)$ are identified by exploiting the first order condition of this problem. Moreover, $F_{c_i}(c)$ is identified by pooling across all realizations of h^T:

$$F_{c_i}(c) = \Pr(c_i \leq c) = \int F_{c_i}(c|h^T) dF_{H_i^T}(h^T). \tag{1.3}$$

Note that the right-hand side of (1.3) is the probability that c_i is less than c without conditioning on h^T (but conditional on i bidding in the auction). The distribution, $F_{H_i^T}$, is the distribution of time-T history h^T in H_i^T, which is the set of all possible time-T histories with i entering.

Step 2. Beliefs on history evolution $h^{\tau_i} \to h^T$ Next, we identify the beliefs of firm i on time-T history h^T conditional on the history at entry timing h^{τ_i} and disclosure action a_i^Q. We denote such beliefs as $\mu_i(h^T|h^\tau, \tau_i = \tau, a_i^Q)$.

When firm i discloses, their beliefs $\mu_i(h^T|h^\tau, \tau_i = \tau, a_i^Q = 1)$ are directly identified from the data. However, when firm i does not disclose, their beliefs $\mu_i(h^T|h^\tau, \tau_i = \tau, a_i^Q = 0)$ cannot be directly identified from data because there is a selection issue due to the fact that their entry timing is not observed. In this case, the key idea for identification is to think of this setup as a survival analysis, where the event is a disclosure from a firm. Additionally, the possibility of multiple firms disclosing can be treated as competing risks. Following this idea, we can establish a simple mapping from the observed evolution of disclosure histories to firm's beliefs about the evolution. It has been shown that the hazard function for each risk can be identified, if potential survival times of each risk are mutually independent (Tsiatis, 1975). In our context, this corresponds to identifying each firm's hazard function for disclosing, given any history. We can then identify the beliefs of a firm on how the disclosure history would evolve.

For the sake of simplicity, we provide an argument for the symmetric case. The proof that allows for asymmetry is given in Appendix Section A.2. First, consider the following probability $p^{noQ}(\tau^1, \tau^2|h^{\tau^1})$:

$$p^{noQ}(\tau^1, \tau^2|h^{\tau^1}) \equiv \Pr(\text{no disclosure between time } \tau^1 \text{ and } \tau^2|h^{\tau^1})$$

where $\tau^1 < \tau^2$. Let the number of firms who have not disclosed under h^{τ^1} be M. Moreover, let the firms who have disclosed by τ^1 be $j_1, \ldots, j_J$, with their corresponding timings $\tilde{\tau}_1, \ldots, \tilde{\tau}_J$. We can express this probability as follows:

$$\begin{aligned}
p^{noQ}(\tau^1, \tau^2|h^{\tau^1}) &= \frac{\Pr(h^{\tau^2})}{\Pr(h^{\tau^1})} \\
&= \frac{\prod_{m=1}^{J} \int_0^\infty f_\tau(\tilde{\tau}_m) A^{\tilde{\tau}_m}(h^{\tau_m}(\tilde{\tau}_1, \ldots, \tilde{\tau}_{m-1}), c) f_c(c) dc}{\prod_{m=1}^{J} \int_0^\infty f_\tau(\tilde{\tau}_m) A^{\tilde{\tau}_m}(h^{\tau_m}(\tilde{\tau}_1, \ldots, \tilde{\tau}_{m-1}), c) f_c(c) dc} \\
&\quad \times \frac{\left\{1 - F_\tau(\tau^2) + \int_0^\infty \int_0^{\tau^2} f_\tau(t) \left(1 - A^t(h^{\tau_2}, c)\right) f_c(c) dt dc\right\}^M}{\left\{1 - F_\tau(\tau^1) + \int_0^\infty \int_0^{\tau^1} f_\tau(t) \left(1 - A^t(h^{\tau^1}, c)\right) f_c(c) dt dc\right\}^M} \\
&= \frac{\left\{1 - F_\tau(\tau^2) + \int_0^\infty \int_0^{\tau^2} f_\tau(t) \left(1 - A^t(h^{\tau_2}, c)\right) f_c(c) dt dc\right\}^M}{\left\{1 - F_\tau(\tau^1) + \int_0^\infty \int_0^{\tau^1} f_\tau(t) \left(1 - A^t(h^{\tau^1}, c)\right) f_c(c) dt dc\right\}^M}
\end{aligned}$$

where h^{τ_2} is a history that includes the same set of disclosures as h^{τ_1} and has no disclosures between τ_1 and τ_2. The probability of entering and then disclosing under history h and construction cost c, $A^t(h, c)$ is:

$$A^t(h, c) \equiv F_E(V^t(h)) \left(p^Q + (1 - p^Q) F_Q(\Delta v_i^t(h, c)\right)$$

where $V^t(h)$ represents the value of entry at time t conditional on history h, and $\Delta v_i^t(h, c)$ is the value of disclosure at time t conditional on history h and construction cost c.[15]

[15]Formally, if history h contains information up to time t' ($t' > t$), we are conditioning on the restricted history $h^t|h^{t'}$ that contains the same information as $h^{t'}$ but only those that happen before t.

Now, let us consider the beliefs of firm i on the same object when the firm enters but does not disclose at τ ($\tau \leq \tau_1 < \tau_2$). We denote such beliefs as $\mu_i^{noQ}(\tau_1, \tau_2 | h^{\tau_1}, \tau_i = \tau, a_i^Q = 0)$. Then,

$$\begin{aligned}
\mu_i^{noQ}(\tau_1, \tau_2 | h^{\tau_1}, \tau_i = \tau, a_i^Q = 0) &= \frac{\Pr(h^{\tau^2} \cap \{\tau_i = \tau, a_i^Q = 0\})}{\Pr(h^{\tau^1} \cap \{\tau_i = \tau, a_i^Q = 0\})} \\
&= \frac{\int_0^\infty F_E(V^\tau(h^\tau)) \left(1 - \left(p^Q + (1-p^Q) F_Q(\Delta v(h^\tau, c))\right)\right) f_c(c) dc}{\int_0^\infty F_E(V^\tau(h^\tau)) \left(1 - (p^Q + (1-p^Q) F_Q(\Delta v(h^\tau, c))\right) f_c(c) dc} \\
&\times \frac{\prod_{m=1}^J \int_0^\infty f_\tau(\tilde{\tau}_m) A^{\tilde{\tau}_m}(h^{\tau_m}(\tilde{\tau}_1, \ldots, \tilde{\tau}_{m-1}), c) f_c(c) dc}{\prod_{m=1}^J \int_0^\infty f_\tau(\tilde{\tau}_m) A^{\tilde{\tau}_m}(h^{\tau_m}(\tilde{\tau}_1, \ldots, \tilde{\tau}_{m-1}), c) f_c(c) dc} \\
&\times \frac{\left\{1 - F_\tau(\tau^2) + \int_0^\infty \int_0^{\tau^2} f_\tau(t) \left(1 - A^t(h^{\tau_2}, c)\right) f_c(c) dt dc\right\}^{M-1}}{\left\{1 - F_\tau(\tau^1) + \int_0^\infty \int_0^{\tau^1} f_\tau(t) \left(1 - A^t(h^{\tau^1}, c)\right) f_c(c) dt dc\right\}^{M-1}} \\
&= \frac{\left\{1 - F_\tau(\tau^2) + \int_0^\infty \int_0^{\tau^2} f_\tau(t) \left(1 - A^t(h^{\tau_2}, c)\right) f_c(c) dt dc\right\}^{M-1}}{\left\{1 - F_\tau(\tau^1) + \int_0^\infty \int_0^{\tau^1} f_\tau(t) \left(1 - A^t(h^{\tau^1}, c)\right) f_c(c) dt dc\right\}^{M-1}}
\end{aligned}$$

holds. Note that firm i's knowledge of its own construction cost does not affect this belief.

From the above, we can see that there is a simple relationship between the observed probability $p^{noQ}(\tau^1, \tau^2 | h^{\tau^1})$ and belief $\mu_i^{noQ}(\tau_1, \tau_2 | h^{\tau_1}, \tau_i = \tau, a_i^Q = 0)$. The former can be expressed as q^M, while the latter is expressed as q^{M-1}, where q is some number.[16] Thus,

$$\mu_i^{noQ}(\tau_1, \tau_2 | h^{\tau_1}, \tau_i = \tau, a_i^Q = 0) = p^{noQ}(\tau^1, \tau^2 | h^{\tau^1})^{\frac{M-1}{M}} \tag{1.4}$$

holds. Since the right-hand side is directly identified from data, i's beliefs on having no disclosures between two time points are identified. Therefore, i's belief on time-T history for the case where firm i does not disclose is identified as well.

Step 3. Value of disclosure We argue that the value of disclosure is identified. The expected value, $v_i^{1,\tau}(h^\tau, c_i)$, from disclosing at time τ under history h^τ when i's construction cost is c_i is simply

$$v_i^{1,\tau}(h^\tau, c_i) = \int V_i(h^T, c_i)\ \mu_i(h^T | h^\tau, \tau_i = \tau, a_i^Q = 1)\ dh^T, \tag{1.5}$$

where $V_i(h^T, c_i)$ is the value from the bidding stage and is given by expression (1.2). For each time-T history h^T, the expected value of bidder i with cost realization c_i is $V_i(h^T, c_i)$. By integrating $V_i(h^T, c_i)$ with respect to $\mu_i(h^T | h^\tau, \tau_i = \tau, a_i^Q)$, the beliefs on distribution

[16] Explicitly, $q = \frac{1 - F_\tau(\tau^2) + \int_0^\infty \int_0^{\tau^2} f_\tau(t)\left(1 - A^t(h^{\tau_2}, c)\right) f_c(c) dt dc}{1 - F_\tau(\tau^1) + \int_0^\infty \int_0^{\tau^1} f_\tau(t)\left(1 - A^t(h^{\tau^1}, c)\right) f_c(c) dt dc}$.

over possible time-T histories, we obtain the expected value. Similarly, the expected value, $v_i^{0,\tau}(h^\tau, c_i)$, from *not* disclosing at time τ under history h^τ when construction cost is c_i is

$$v_i^{0,\tau}(h^\tau, c_i) = \int V_i(h^T, c_i)\ \mu_i(h^T | h^\tau, \tau_i = \tau, a_i^Q = 0)\ dh^T. \tag{1.6}$$

Note that $V_i(h^T, c_i)$ is identified for all time-T histories h^T and construction costs c_i in Step 1, and μ_i was identified in Step 2. Therefore, the terms on the right-hand sides of equations (1.5) and (1.6) are all identified, establishing that $v_i^{1,\tau}(h^\tau, c_i)$ and $v_i^{0,\tau}(h^\tau, c_i)$ are both identified. We denote the value of disclosure as $\Delta v_i^\tau(h^\tau, c_i) \equiv v_i^{1,\tau}(h^\tau, c_i) - v_i^{0,\tau}(h^\tau, c_i)$. We can easily see that the value of disclosure is identified.

Step 4. Probability of forced disclosure p^Q and distribution of disclosure costs F_Q First, we introduce the following assumption:

Assumption 2. *Support of disclosure values is $[\underline{v^Q}, \overline{v^Q}]$. Firms always disclose at the upper bound of the disclosure value: $F_Q(\overline{v^Q}) = 1$.*

The key variation we leverage here is the difference in disclosure values across firms with varying construction costs but those who are facing the same disclosure history. When a firm is not forced to disclose, the decision to disclose is given by the following expression:

$$\begin{cases} \iota_{i,\tau}(h^\tau, c_i, c_i^Q) = 1 & \text{if } \Delta v_i^\tau(h^\tau, c_i) \geq c_i^Q \\ \iota_{i,\tau}(h^\tau, c_i, c_i^Q) = 0 & \text{otherwise} \end{cases} \tag{1.7}$$

The expected value of the auction $\tilde{v}_i^\tau(h^\tau, c_i)$ at time τ under history h^τ and construction costs c_i is:

$$\tilde{v}_i^\tau(h^\tau, c_i) = p^Q v_i^1(h^\tau, c_i) + (1 - p^Q)\mathbb{E}_{F^Q}\left[\max\left\{v_i^0(h^\tau, c_i), v_i^1(h^\tau, c_i) - c_i^Q\right\}\right]. \tag{1.8}$$

The first term corresponds to the case where firm i is forced to disclose. Inside the expectation bracket, the first term represents the expected value from not disclosing, and the second term represents the expected value from costly disclosing. The expected value of entry, $v_i(h^\tau)$, is then

$$v_i^\tau(h^\tau) = \mathbb{E}_{F_{c_i}}\left[\tilde{v}_i^\tau(h^\tau, c_i)\right]. \tag{1.9}$$

The decision to enter at time τ under history h^τ is given by the following expression:

$$\begin{cases} \chi_{i,\tau}(h^\tau, c_i^E) = 1 & \text{if } v_i^\tau(h^\tau) \geq c_i^E \\ \chi_{i,\tau}(h^\tau, c_i^E) = 0 & \text{if otherwise} \end{cases} \tag{1.10}$$

Fix values $v', v'' \in \mathbb{R}$. Recall that the expected gain from disclosure, $\Delta v_i^\tau(h^\tau, c_i)$, is identified for all τ, h^τ and c_i. Now let us take c_i' and c_i'' appropriately so that $\Delta v(h^\tau, c_i') = v'$

and $\Delta v(h^\tau, c_i'') = v''$ for some h^τ. The density that a firm with type c_i' discloses at h^τ is given by:

$$\begin{aligned} & f\left(a_i^Q = 1, \tau_i = \tau, h^\tau, c_i'\right) \\ = \; & f_{H^\tau}(h^\tau | c_i', \tau_i = \tau)\; f_\tau(\tau)\; f_{c_i}(c_i')\; F_E(v_i^\tau(h^\tau))\; \left(p^Q + (1-p^Q)F_Q(v')\right) \\ = \; & f_{H^\tau}(h^\tau | \tau_i = \tau)\; f_\tau(\tau)\; f_{c_i}(c_i')\; F_E(v_i^\tau(h^\tau))\; \tilde{F}_Q(v') \end{aligned} \tag{1.11}$$

where f_{H^τ}, f_τ and f_c are the densities of F_{H^τ}, F_τ and F_c, respectively. Also, we denote $\tilde{F}_Q(v) = \left(p^Q + (1-p^Q)F_Q(v)\right)$. The first term on the right-hand side of (1.11) is the probability that event h^τ occurs conditional on arrival timing τ_i being equal to τ. The second term is the probability of τ_i being equal to τ, and the third term gives the probability that the cost draw is c_i'. The fourth term corresponds to the entry probability. Finally, the last term represents the probability of disclosure ($a_i^Q = 1$) conditional on time-τ history h^τ and construction cost c_i'. This final probability accounts for (i) forced disclosure and (ii) costly disclosure if the disclosure cost, c_i^Q, is less than v', i.e., $c_i^Q \leq \Delta v_i^\tau(h^\tau, c_i')\ (= v')$.

Similarly, the density that a firm with type c_i'' discloses under history h^τ is as follows:

$$\Pr\left(a_i^Q = 1, \tau_i = \tau, h^\tau, c_i''\right) = f_{H^\tau}(h^\tau | \tau_i = \tau)\; f_\tau(\tau)\; f_{c_i}(c_i'')\; F_E(v_i(h^\tau))\; \tilde{F}_Q(v''). \tag{1.12}$$

Since construction costs c_i are identified for all entrants from Step 1, the left-hand sides of expressions (1.11) and (1.12) are both identified. Moreover, $f_c(c_i')$ and $f_c(c_i'')$ are both identified because $F_{c_i}(c)$ is identified. Hence, from the ratio of expressions (1.11) and (1.12), we identify $\tilde{F}_Q(v')/\tilde{F}_Q(v'')$. Because F_Q is a distribution, p^Q and F_Q are identified.[17][18]

Step 5. Value of entry As demonstrated in (1.9), value of entry $v_i^\tau(h^\tau)$ can be expressed as:

$$\begin{aligned} v_i^\tau(h^\tau) &= \mathbb{E}_{F_{c_i}}\left[\tilde{v}_i^\tau(h^\tau, c_i)\right] \\ &= \int \tilde{v}_i^\tau(h^\tau, c_i) dF_{c_i}(c_i) \\ &= \int p^Q v_i^{1,\tau}(h^\tau, c_i) + (1-p^Q)\mathbb{E}_{F^Q}\left[\max\left\{v_i^{0,\tau}(h^\tau, c_i), v_i^{1,\tau}(h^\tau, c_i) - c_i^Q\right\}\right] dF_{c_i} \\ &= \iint p^Q v_i^{1,\tau}(h^\tau, c_i) + (1-p^Q)\max\left\{v_i^{0,\tau}(h^\tau, c_i), v_i^{1,\tau}(h^\tau, c_i) - c_i^Q\right\} dF_Q\; dF_{c_i} \end{aligned}$$

[17] If $\tilde{F}_Q(v')/\tilde{F}_Q(v'')$ is identified for all $v', v'' \in \mathbb{R}$, it implies that $\tilde{F}_Q(v)$ is identified up to a constant, say, $\tilde{F}_Q(0)$. This is because we can express $\tilde{F}_Q(v)$ as follows: $\tilde{F}_Q(v) = \tilde{F}_Q(0)\left(\tilde{F}_Q(v)/\tilde{F}_Q(0)\right)$, where the ratio $\left(\tilde{F}_Q(v)/\tilde{F}_Q(0)\right)$ is identified. There is a unique value of $\tilde{F}_Q(0)$ such that $\lim_{v \to \overline{v^D}} F_Q(v) = 1$.

[18] Strictly speaking, we need an assumption for this argument to be valid. We assume that: For all $(v', v'') \in [\underline{v^Q}, \overline{v^Q}]^2$, there exist a sequence of histories $(h_1^{\tau_1}, \ldots, h_H^{\tau_H})$ and a sequence of numbers $(v_1, \ldots, v_{H+1})$ such that (i) $v_1 = v'$; (ii) $v_{H+1} = v''$: and (iii) there exist some $c_{k,1}, c_{k,2} \in \mathbb{R}^+$ that $\Delta v_i^{\tau_k}(h^{\tau_k}, c_{k,1}) = v_k$ and $\Delta v_i^{\tau_k}(h^{\tau_k}, c_{k,2}) = v_{k+1}$ for all k $(1 \leq k \leq H)$.

Since all the objects that appear in this expression are identified objects, value of entry $v_i^\tau(h^\tau)$ is also identified.

Step 6. Distribution of entry costs F_E and arrival timing F_τ In this final step, our goal is to identify the remaining two distributions: the entry cost distribution and arrival timing distribution. To achieve this, the idea is to exploit variation in value of entry and value of disclosure across firms facing different disclosure histories.

Suppose that under time-τ history h^τ, where bidders $j_1, \ldots, j_J$ have each disclosed at $\tau_{j_1}, \ldots, \tau_{j_J}$ ($\tau_{j_1} < \cdots < \tau_{j_J}$), and the remaining bidders i and $k_1, \ldots, k_K$ have not disclosed. As in Step 2, let

$$A_i^t(h, c_i) \equiv F_E^i(V_i^t(h)) F_Q(\Delta v_i^t(h, c_i)) \tag{1.13}$$

First, consider the following density P:[19]

$$\begin{aligned}
P &= \Pr\left(j_m \text{ signals at } \tau_{j_m}\ \forall m, i \text{ signals at } \tau, k_n \text{ does not signal before } \tau\ \forall n,\ \vec{c}_j, c_i\right) \\
&= \prod_m f_\tau^{j_m}(\tau_{j_m}) A_{j_m}^{\tau_{j_m}}(h^{\tau_{j_m}}(\tau_{j_1}, \ldots, \tau_{j_J}), c_{j_m}) f_{c_{j_m}}(c_{j_m}) \\
&\times \prod_n \left\{ 1 - F_\tau^{k_n}(\tau) + \int_0^\infty \int_0^\tau f_\tau^{k_n}(t) \left(1 - A_{k_n}^t(h^t(\tau_{j_1}, \ldots, \tau_{j_J}), c_{k_n})\right) f_{c_{k_n}}(c_{k_n}) dt dc_{k_n} \right\} \\
&\times f_\tau^i(\tau) A_i^\tau(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_{c_i}(c_i)
\end{aligned} \tag{1.14}$$

Next, We consider the following density Q:

$$\begin{aligned}
Q &= \Pr\left(j_m \text{ signals at } \tau_{j_m}\ \forall m, i \text{ does not signal before } \tau, k_n \text{ does not signal before } \tau\ \forall n,\ \vec{c}_j\right) \\
&= \prod_m f_\tau^{j_m}(\tau_{j_m}) A_{j_m}^{\tau_{j_m}}(h^{\tau_{j_m}}(\tau_{j_1}, \ldots, \tau_{j_J}), c_{j_m}) f_{c_{j_m}}(c_{j_m}) \\
&\times \prod_n \left\{ 1 - F_\tau^{k_n}(\tau) + \int_0^\infty \int_0^\tau f_\tau^{k_n}(t) \left(1 - A_{k_n}^t(h_t(\tau_{j_1}, \ldots, \tau_{j_J}), c_{k_n})\right) f_{c_{k_n}}(c_{k_n}) dt dc_{k_n} \right\} \\
&\times \left\{ 1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h^t(\tau_{j_1}, \ldots, \tau_{j_J}), c_i)\right) f_{c_i}(c_i) dt dc_i \right\}.
\end{aligned} \tag{1.15}$$

Taking the ratio between these two densities gives us:

$$P/Q = \frac{f_\tau^i(\tau) A_i^\tau(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_{c_i}(c_i)}{\left\{1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h^t(\tau_{j_1}, \ldots, \tau_{j_J}), c_i)\right) f_{c_i}(c_i) dt dc_i\right\}} \tag{1.16}$$

Exploiting the relation that

$$\begin{aligned}
&\frac{\partial(1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h^t, c_i)\right) f_c(c_i) dt dc_i)}{\partial \tau} \\
&= -\int_0^\infty f_\tau^i(\tau) A_i^\tau(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_c(c_i) dc_i,
\end{aligned}$$

[19] When we write $h^t(\tau_{j_1}, \ldots, \tau_{j_J})$, this is a time-$t$ history such that disclosures are made at times that are in the set $\{\tau_{j_1}, \ldots, \tau_{j_J}\}$ and before t. Note that we are also suppressing the expression on which firm disclosed at which timing. Moreover, we use the notation Pr to represent densities for simplicity in exposition.

the function

$$\Gamma_i(\tau; h^\tau = (\tau_{j_1}, \ldots, \tau_{j_J})) = 1 - F^i_\tau(\tau) + \int_0^\infty \int_0^\tau f^i_\tau(t) \left(1 - A^t_i(h^t, c_i)\right) f_c(c_i) dt dc_i$$
$$= 1 - \int_0^\infty \int_0^\tau f^i_\tau(t) A^t_i(h^t, c_i) f_c(c_i) dt dc_i \tag{1.17}$$

is identified up to scale for all $\tau \in [\tau_{j_J}, T]$.[20] And thus

$$-\frac{\partial \Gamma_i}{\partial \tau}(\tau; (\tau_{j_1}, \ldots, \tau_{j_J})) = \int_0^\infty f^i_\tau(\tau) A^\tau_i(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_c(c_i) dc_i$$
$$= f^i_\tau(\tau) F^i_E(V^\tau_i(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J})))$$
$$\times \int_0^\infty F_Q(\Delta v^\tau_i(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i)) f_c(c_i) dc_i$$

is identified up to scale. Since $\Gamma_i(0; h^0 = \phi)$=1 holds, $\Gamma_i(\tau; h^\tau = \phi)$ is identified for all $\tau \in [0, T]$.[21] Therefore, $\frac{\partial \Gamma_i}{\partial \tau}(\tau; h^\tau = \phi)$ is identified for all $\tau \in [0, T]$. Furthermore, since F_Q is identified, $f^i_\tau(\tau) F^i_E(V^\tau_i(h^\tau = \phi))$ is identified for all $\tau \in [0, T]$.

Now, given that $f^i_\tau(\tau) F^i_E(V^\tau_i(\phi))$ is identified, $f^i_\tau(t) A^t_i(h^t = \phi, c_i)$ is identified for all $t \in [0, T]$. Consequently, as Γ_i is expressed as (1.17), $\Gamma_i(\tau; h^\tau = \tau)$ is identified. Since Γ_i is identified up to scale and now that $\Gamma_i(\tau; h^\tau = \tau)$ is identified, $\Gamma_i(\tau'; h^{\tau'} = \tau)$ is identified for all $\tau' \in [\tau, T]$. As a result, $f^i_\tau(t)\ F^i_E(V^t_i(h^t))$ such that h^t includes one disclosure is identified. By induction on the number of disclosures made, repeating this argument will allow us to identify $f^i_\tau(\tau)\ F^i_E(V^\tau_i(h^\tau))$ for all histories h^τ.

To proceed with the identification of f_τ, fix $\tau', \tau'' \in [0, T]$ and take $h^{\tau'}$ and $h^{\tau''}$ appropriately so that $V^{\tau'}_i(h^{\tau'}) = V^{\tau''}_i(h^{\tau''}) = v$ for some constant $v \in \mathbb{R}$. We identify the ratio, $f_\tau(\tau')/f_\tau(\tau'')$ from the ratio of $f^i_\tau(\tau') F^i_E(V^{\tau'}_i(h^{\tau'}))$ and $f^i_\tau(\tau'') F^i_E(V^{\tau''}_i(h^{\tau''}))$. Since F_τ is a distribution, F_τ is identified.[22]

Finally, since f_τ and $f^i_\tau(\tau) F^i_E(V^\tau_i(h^\tau))$ are identified, F^E is also identified.

1.6 Estimation

In this section, we provide an outline of the estimation procedure, which closely follows the identification argument.

[20] Note that $\frac{dF(x)}{dx}/F(x) = \frac{d(\log F(x))}{dx}$ holds. If the left-hand side object is identified, $\log F(x)$ is identified up to a constant. Therefore, $F(x)$ is identified up to scale.

[21] When we write $h^t = \phi$, we mean that time-t history h^t does not include any disclosures.

[22] Strictly speaking, we need an assumption for this argument to be valid. We assume that: For all $(\tau', \tau'') \in [0, T]^2$, there exist a sequence of timings $(\tau^1, \ldots, \tau^H)$ and a sequence of numbers $(v_1, \ldots, v_H)$ such that (i) $\tau^1 = \tau'$; (ii) $\tau^H = \tau''$: and (iii) there exist some histories h^{τ_k} and $\tilde{h}^{\tau_{k+1}}$ such that $V^{\tau_k}_i(h^{\tau_k}) = V^{\tau_{k+1}}_i(h^{\tau_{k+1}})$ for all k $(1 \leq k \leq H-1)$.

Parametric assumptions

To apply our model to data, we introduce parametric assumptions, despite having established non-parametric identification. First, we assume that firms are *ex-ante* symmetric, conditional on auction-level characteristics: all the firms share the same distribution for arrival timing, entry costs, disclosure costs, and construction costs if the auction is the same construction type and from the same district.

In what follows, we set $T = 1$. We specify parametric forms for the distributions of entry timing, entry costs, and disclosure costs. We assume that the distribution of arrival timing follows a Beta distribution, which lies within the interval $[0, 1]$, with two shape parameters $\theta^\tau \equiv (\alpha_\tau, \beta_\tau)$. Next, we make the following assumption on entry costs. With probability p^E, each firm gets a chance to consider whether they would enter an auction, while a firm always stays out with the other probability $1 - p^E$. This reflects the fact that firms may face various constraints, such as other ongoing projects. If a firm considers entry, they draw an entry cost c_i^E from a truncated normal distribution on $[0, \infty)$ with parameters $\theta^E \equiv (\mu_E, \sigma_E)$. Here, we parameterize $\mu_E = X_a\beta^E + \alpha^E$, where X_a is the logarithm of number of potential entrants. Finally, we assume that the distribution of disclosure costs F_Q follows a truncated normal distribution on $[0, \infty)$ with parameters $\theta^Q \equiv (\mu_Q, \sigma_Q)$. Note that we have also assumed that firms are in a position where they must disclose with probability p^Q, reflecting the fact that disclosures are done through posting questions in our setup.

Estimation procedure

We estimate our parameters in four steps. In the first step, we start by estimating the construction costs for each entrant and the distribution of such costs, exploiting the bidding results. Next, we estimate firms' beliefs on how disclosure history evolves over time, conditional on their disclosure actions. With the estimates from the bidding stage and estimated beliefs on disclosure history, we turn to the estimation of the value of disclosure and entry. Finally, using the obtained estimates, we estimate the remaining model primitives via maximum likelihood estimation.

Step 1. Construction costs c_i To account for the fact that some bids are ultimately rejected, we assume the presence of a secret reserve price p^r, which follows a log-normal distribution. To estimate the construction costs c_i, we exploit the optimality of the bids as in Guerre et al. (2000). Specifically, construction cost c_i when the bid is b_i is estimated exploiting the first order condition for bidding:

$$c_i = b_i - \frac{1 - G_{-i}(b)}{g_{-i}(b)} \tag{1.18}$$

where G_{-i} is the CDF of the lowest bid among the opponents, and g_{-i} is the corresponding pdf.[23]

[23]This will be the minimum of the opponents' bid and the secret reserve price.

We assume that G_{-i} follows a log-normal distribution log-$\mathcal{N}$ (μ_b, σ_b) with:

$$\mu_b = \mathbf{X_i^{\mu_b}}\beta^{\mu_b}, \qquad \sigma_b = \mathbf{X_i^{\sigma_b}}\beta^{\sigma_b},$$

where $X_i^{\mu_b}$ includes a dummy indicating whether i disclosed, the time at which i disclosed, the number of disclosures made by others, the number of potential bidders, construction type dummies, and district dummies. For the variance, $X_i^{\sigma_b}$ includes the number of others' disclosures, the number of potential bidders, construction type dummies, and district dummies. We estimate the parameters $(\beta^{\mu_b}, \sigma^{\mu_b})$ via maximum likelihood. Once we obtain the estimates for the distribution G_{-i}, we exploit (1.18) and estimate construction costs for each entrant. While we allow the distribution of construction costs F_c to be fully nonparametric, we assume that the distribution depends on the type of construction and district at which the project is located.

Step 2. Belief on the evolution of disclosure history Closely following the identification argument, we start by estimating the observed evolution of disclosure histories. Let us note here again that the observed evolution of disclosure histories and the beliefs on the histories are not identical. We parameterize the distribution of time intervals between the n-th and $(n+1)$-th disclosure as follows (time interval between $t = 0$ and the first disclosure will be also included as case $n = 0$):

$$\Pr(\tau^{n+1} - \tau^n \leq x) = \frac{\Phi((x - \mu_t)/\sigma_t) - \Phi(-\mu_t/\sigma_t)}{1 - \Phi(-\mu_t/\sigma_t)}$$

if the $(n+1)$-th disclosure exists and

$$\Pr\left((n+1)\text{-th disclosure does not exist}\right) = \frac{1 - \Phi\left(((1 - \tau^n) - \mu_t)/\sigma_t\right)}{1 - \Phi(-\mu_t/\sigma_t)}$$

where timing of the n-th disclosure is given by τ^n.[24][25] The parameters (μ_t, σ_t) are characterized as:

$$\mu_t = \mathbf{X}^{\mu_t}\beta^{\mu_t}, \qquad \sigma_t = \mathbf{X}^{\sigma_t}\beta^{\sigma_t},$$

where X^{μ_t} includes the n-th disclosure timing τ^n, the number of disclosures n, the log of (number of firms who have not disclosed yet $+1$), construction type dummies, and district dummies. For the variance, X^{σ_t} includes the number of disclosures n, and the log of (number of firms who have not disclosed yet $+1$). We estimate $(\beta^{\mu_t}, \sigma^{\mu_t})$ via maximum likelihood.

Next, we turn to the estimation of the beliefs of the firms. First, we estimate the belief of a firm when the firm discloses at some time τ facing history h^τ. To estimate this belief, we simulate the evolution of the disclosure history, using the distribution of time intervals

[24] We are implicitly assuming that the decision to disclose depends on the number of disclosures that have occurred up to that time and the timing of the most recent disclosure.

[25] If n is equal to the number of total potential bidders N, then the next disclosure does not exist with probability one.

between disclosures we have estimated above. Suppose that i's disclosure was the n-th one. Then, the next disclosure timing τ^{n+1}'s distribution is characterized by the above distribution we have estimated. By repeating this process, until we draw the probability of next disclosure not existing, we can obtain the distribution of time-T histories. We denote the estimated belief as: $\hat{\mu}_i(h^T|h^\tau, \tau_i = \tau, a_i^Q = 1)$.

Next, we estimate a firm's belief when it enters but does not disclose at time τ facing history h^τ. Let τ^n represent the most recent disclosure before τ, with $\tau^n = 0$ if none exists. We exploit the following relationship, which we have shown as equation 1.4:

$$\begin{aligned}&\Pr(\tau^{n+1} > t \text{ or no } (n+1)\text{-th disclosure})^{(M-1)/M}\\&= \mu_i(\tau^{n+1} > t \text{ or no } (n+1)\text{-th disclosure})\end{aligned}$$

where M is the number of potential entrants who have not disclosed at τ, including firm i. Moreover, from equation 1.4,

$$\begin{aligned}&\Pr(\tau^{n+k} > t \text{ or no } (n+k)\text{-th disclosure}|(n+k-1)\text{-th disclosure is at } \tau^{n+k-1})^{(M-k)/(M-k+1)}\\&= \mu_i(\tau^{n+k} > t \text{ or no } (n+k)\text{-th disclosure}|(n+k-1)\text{-th disclosure is at } \tau^{n+k-1})\end{aligned}$$

holds for all k ($1 \leq k \leq M-1$) and all t ($\tau^{n+k-1} < t < 1$). These relationships allow us to back out the distribution of time intervals between disclosures from firm i's perspective, since the left hand-side objects are what we had estimated at the beginning of Step 2. Now, we can simulate the evolution of disclosure history using i's belief on the timing of disclosures. We denote the estimated belief as: $\hat{\mu}_i(h^T|h^\tau, \tau_i = \tau, a_i^Q = 0)$.

Step 3. Value of disclosure In this step, we aim to obtain an estimate for the value of disclosures. First, we start by estimating the value from the bidding stage $V_i(h^T, c)$. For clarity, each objects are estimated for every construction type and district pair, although we omit these specific dependencies here for simplicity. We estimate the value from bidding conditional on time-T history $V_i(h^T, c)$ by:

$$\hat{V}_i(h^T, c) = \max_b (b-c)(1 - \hat{G}_{-i}(b|h^T)),$$

where $\hat{G}_{-i}$ is the estimated CDF of the lowest bid from opponents.

Next, we estimate the values with and without disclosure, $v^{1,\tau}(h^\tau, c)$ and $v^{0,\tau}(h^\tau, c)$. This value is estimated by:

$$\hat{v}^{j,\tau}(h^\tau, c) = \int \hat{V}_i(h^T, c)\ \hat{\mu}_i(h^T|h^\tau, \tau_i = \tau, a_i^Q = j)\ dh^T$$

for $j = 0, 1$. Note that although we have expressed this as an integration, due to our estimation procedure, $\hat{\mu}_i$ is a discrete distribution. Therefore, this turns out to be a summation, in practice. Then, value of disclosure $\Delta v^\tau(h^\tau, c)$ can be estimated as:

$$\widehat{\Delta v}^\tau(h^\tau, c) = \hat{v}^{1,\tau}(h^\tau, c) - \hat{v}^{0,\tau}(h^\tau, c).$$

Step 4. Model primitives In our final step, we estimate our model primitives–specifically, the distribution of entry timing F_τ, entry costs F_E and p^E, and disclosure costs F_Q and p^Q– using observed entry and disclosure data. Note that if we have estimates $\hat{p}^Q$ and $\hat{F}^Q$ for p^Q and F_Q, value of entry $v(h^\tau)$ can be estimated as:

$$\begin{aligned}\hat{v}^\tau(h^\tau) &= \sum_c \int \hat{p}^Q \hat{v}^{1,\tau}(h^\tau, c) + (1-\hat{p}^Q)\mathbb{E}_{\hat{F}^Q}\left[\max\left\{\hat{v}^{0,\tau}(h^\tau,c), \hat{v}^{1,\tau}(h^\tau,c) - c^Q\right\}\right] d\hat{F}_Q \\ &= \sum_c \Bigg[\left(\hat{p}^Q + (1-\hat{p}^Q)\hat{F}_Q(\max\{\hat{v}^{1,\tau}(h^\tau,c) - \hat{v}^{0,\tau}(h^\tau,c), 0\})\right)\hat{v}^{1,\tau}(h^\tau,c) \\ &\quad + (1-\hat{p}^Q)\left(1 - \hat{F}_Q(\max\{\hat{v}^{1,\tau}(h^\tau,c) - \hat{v}^{0,\tau}(h^\tau,c), 0\})\right)\hat{v}^{0,\tau}(h^\tau,c) \\ &\quad - (1-\hat{p}^Q)\int_0^{\max\{\hat{v}^{1,\tau}(h^\tau,c) - \hat{v}^{0,\tau}(h^\tau,c), 0\}} c^Q d\hat{F}_Q \Bigg]. \end{aligned} \tag{1.19}$$

where summation is taken over the estimated costs c.

Using this expression, we estimate the model primitives via maximum likelihood. Suppose firms $i_1, \ldots, i_I$ do not enter, firms $j_1, \ldots, j_J$ enter but do not disclose, and firms $k_1, \ldots, k_K$ enter and disclose at time $\tau_{k_1}, \ldots, \tau_{k_K}$. Also, let each entrant i's cost be c_i. The likelihood function for observing this history h^T and set of entrants, say $\mathcal{N}^E$, is:

$$\begin{aligned}\mathcal{L}(h^T, \mathcal{N}^E) = &\prod_l \int f_c(c_{i_l})dc \int f_\tau(t)(1 - \tilde{F}_E(v(h^t)))dt \\ &\times \prod_m f_c(c_{j_m}) \int f_\tau(t)\tilde{F}_E(v(h^t))(1 - \tilde{F}_Q(\Delta v(h^t, c_{j_m})))dt \\ &\times \prod_n f_c(c_{k_n}) f(\tau_{k_n})\tilde{F}_E(v(h^{\tau_{k_n}}))\tilde{F}_Q(\Delta v(h^{\tau_{k_n}}, c_{k_n})),\end{aligned}$$

where h^t and $h^{\tau_{k_n}}$ are time-t and time-τ_{k_n} histories that arise as a natural restriction of h^T. Furthermore, $\tilde{F}_E$ is defined as: $\tilde{F}_E = p^E F_E$. The parameters we aim to estimate here are: $\Theta \equiv (\theta^\tau, \theta^Q, p^Q, \theta^E, p^E)$. We solve:

$$\max_\Theta \sum_{a=1}^A \log \hat{\mathcal{L}}(h_a^T, \mathcal{N}_a^E),$$

where $a = 1, \ldots, A$ represents each auction, and $\hat{\mathcal{L}}$ is the likelihood function with previous estimates plugged in:

$$\begin{aligned}\hat{\mathcal{L}} = &\prod_l \int f_\tau(t;\theta^\tau)(1 - \tilde{F}_E(\hat{v}(h^t;(\theta^Q,p^Q));(\theta^E,p^E)))dt \\ &\times \prod_m \hat{f}_c(\hat{c}_{j_m}) \int f_\tau(t;\theta^\tau)\tilde{F}_E(\hat{v}(h^t;(\theta^Q,p^Q));(\theta^E,p^E))(1 - \tilde{F}_Q(\widehat{\Delta v}^t(h^t, c_{j_m});(\theta^Q,p^Q)))dt \\ &\times \prod_n \hat{f}_c(\hat{c}_{k_n}) f_\tau(\tau_{k_n};\theta^\tau)\tilde{F}_E(\hat{v}(h^{\tau_{k_n}};(\theta^Q,p^Q));(\theta^E,p^E))\tilde{F}_Q(\widehat{\Delta v}^{\tau_{k_n}}(h^{\tau_{k_n}}, c_{k_n});(\theta^Q,p^Q))\end{aligned}$$

with $\hat{v}^\tau(h^\tau)$ as described in Equation 1.19.

1.7 Estimation Results

This section discusses the results from the estimation of the parameters in the model.

Parameter Estimates

Table 1.4 presents the estimation results for the model parameters. Figure 1.3 illustrates the CDF of firms' arrival timing. Our estimates suggest that firms are more likely to arrive during the latter half of the entry period, with 70% of the firms arriving in this time frame. The median arrival time is 0.71, approximately one week before the forum's closure.

Figure 1.4 shows the relationship between values of entry and entry probabilities in a scenario with 12 potential bidders, which is the median size of the bidders' pool. Firms consider entering with 28% probability when there are 12 potential bidders. This probability decreases with the number of potential bidders. Additionally, the median entry cost is estimated to be 3.4% of the engineer's estimate. Our estimate of entry costs is comparable with the numbers reported in the literature (Bajari et al., 2010; Krasnokutskaya and Seim, 2011).

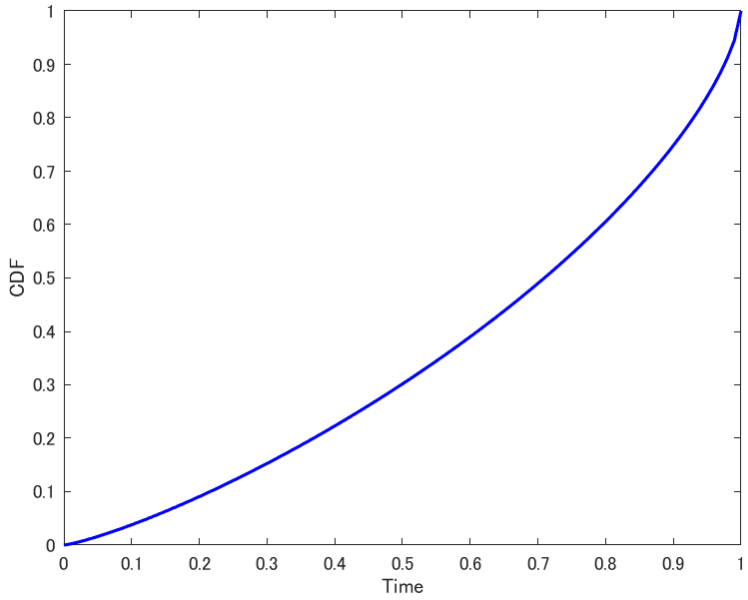
Figure 1.5 shows the relationship between values of disclosure and disclosure probabilities. Firms get in need for posting a question with a probability of 27%, which reflects the likelihood of disclosure even when it may be disadvantageous to them. As value of disclosure increases, disclosure probability also rises. For instance, when the value of disclosure is 1% of the estimated cost, firms disclose with a probability of 31%.

We summarize the estimated distributions of construction costs in Table 1.5. The table reports the median, 25th and 75th percentiles as fractions of the engineer's estimate for each construction type. The median construction cost is estimated to range from 74% to 89% of the engineer's estimate. Overall, projects related to overlay and reconstruction have higher construction costs compared to the other project types.

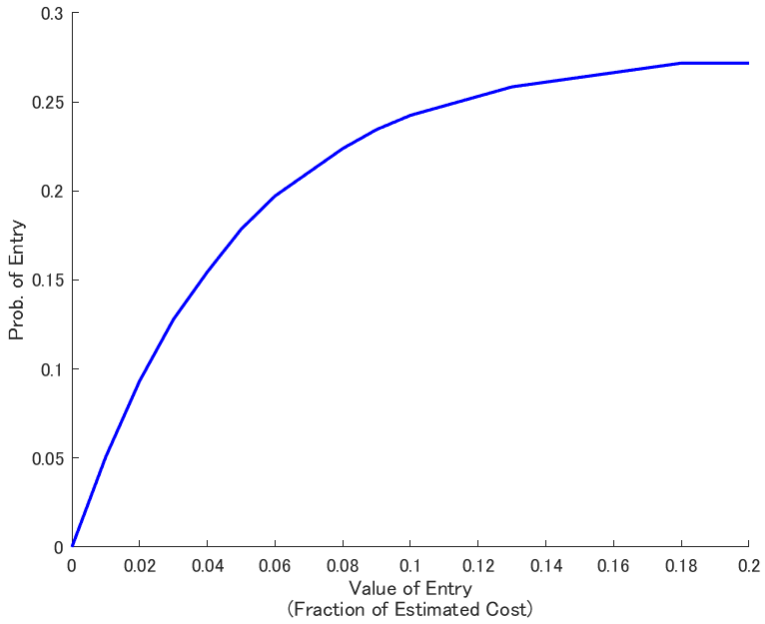
Table 1.6 presents the estimation results for the distribution of opponents' best bids.[26] In a first-price auction, it is the opponents' best bid that ultimately determines an individual firm's profit. First, opponents' best bid becomes more aggressive as a firm faces more disclosures (questions). This trend reflects the fact that disclosures are made by the *actual* entrants.

Next, the results suggest that making disclosures at earlier periods makes opponents' best bid less aggressive, if we hold others' disclosure behaviors fixed. To illustrate this impact, we compare two scenarios: (i) Firm X discloses at $t = 0$ while no other firm discloses; and (ii) No firms disclose at all. If firm X places the median bid ($b_i = 1.03$), its probability of winning probability increase by 11.4 p.p. in scenario (i) compared to scenario (ii), for an auction on an overlay project in the Missoula district. Conversely, the estimates imply that a

[26]Estimates for the distribution of the secret reserve price is presented in Appendix Table A.1.



Notes: This figure shows the CDF of arrival timing.



Notes: This figure shows the relationship between values of entry and entry probability for the case where we have 12 potential bidders, which is the median size of entrants' pool.

Table 1.4: Estimated parameters of the model

Distribution	c_E: Truncated Normal on $[0, \infty)$ c_Q: Truncated Normal on $[0, \infty)$ τ: Beta	
	Estimate	S.E.
Entry		
Prob. of considering entry: p^E		
Const.	0.851	0.021
ln(# Pot bidder)	-0.231	0.009
μ_E	-2.926	0.015
σ_E	0.383	0.033
Disclosure		
Prob. of always disclosing: p^Q	0.268	0.019
μ_Q	-2.416	0.251
σ_Q	0.642	1.213
Timing		
α_τ	1.227	0.099
β_τ	0.661	0.052

Note: Table presents estimates of the model parameters. Standard errors are calculated using 100 bootstrap draws, with sampling at the auction level.

Table 1.5: Distribution of Construction Costs: by Construction Types

Construction Types	25-th percentile	Median	75-th percentile
Bridge	0.38	0.75	1.02
Overlay	0.74	0.88	1.01
Reconstruction	0.72	0.89	1.11
Safety	0.50	0.74	0.99
Others	0.56	0.74	0.97

Note: Total number of projects is 434. There were 5 auctions without an entrant. The presented numbers are fractions of the engineer's estimate.

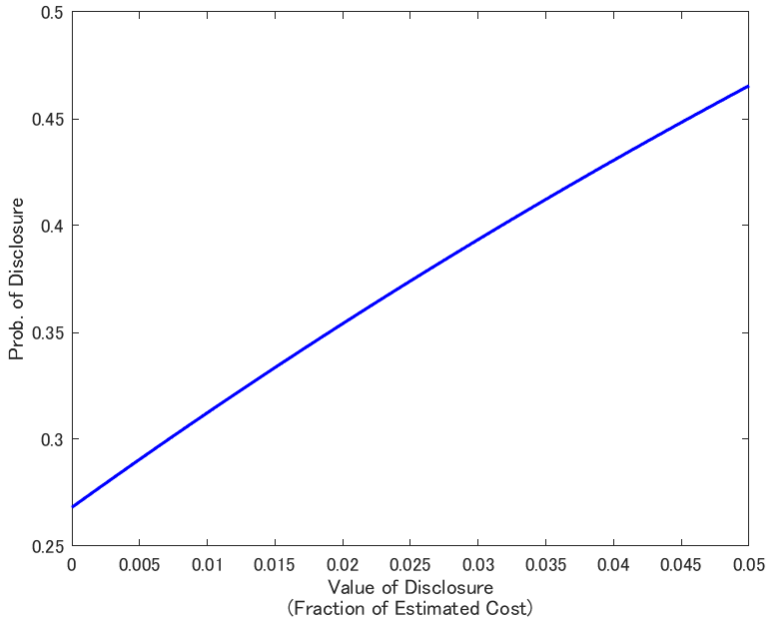
last minute disclosure would make opponents' best bid more aggressive, although this effect is not statistically significant.

These results align with our earlier discussion on the trade-off associated with entry disclosures. As previously noted, while disclosures can reduce the number of entrants, they may simultaneously provoke more aggressive bids from remaining entrants. Our results suggest that the former effect dominates for disclosures made during earlier periods, while

Table 1.6: Distribution of opponents' best bid: Log-normal

Variables	Estimate	S.E.
μ		
Constant	-0.023	0.067
Asked	0.044	0.027
Asked $\times\tau$	-0.045	0.033
ln(# Pot. Bidders)	-0.017	0.034
# Q from others	-0.037	0.010
Type:		
Overlay	0.066	0.030
Safety	0.018	0.050
Bridge	0.207	0.054
Recons	0.102	0.044
Others	Reference	
District:		
Missoula	Reference	
Butte	-0.027	0.028
Great Falls	0.050	0.029
Glendive	0.035	0.032
Billings	-0.022	0.043
$\log\sigma$		
Constant	-1.232	0.144
ln(# Pot. Bidders)	-0.018	0.082
# Q from others	-0.207	0.039
Type:		
Overlay	-0.596	0.100
Safety	0.043	0.134
Bridge	0.274	0.140
Recons	-0.240	0.139
Others	Reference	
District:		
Missoula	Reference	
Butte	-0.148	0.092
Great Falls	-0.149	0.091
Glendive	-0.028	0.094
Billings	0.320	0.123

Note: This table presents estimated parameters of the distribution of opponents' best bid. The opponents' best bid is defined as the minimum of the opponents' bid and the secret reserve price.



Notes: This figure shows the relationship between value of disclosure and disclosure probability.

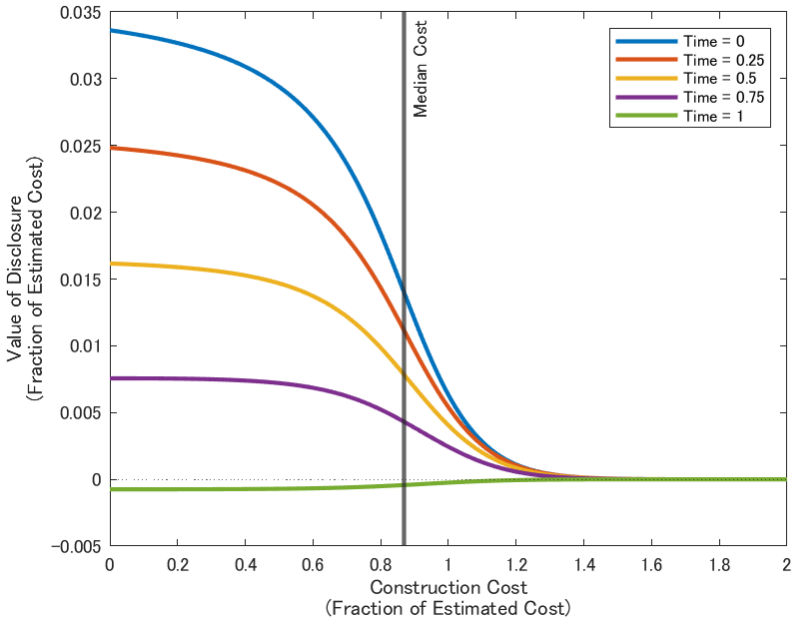
the latter effect dominates when disclosures are made at the last minute, at a time close to $t = 1$.

It is important to note that these results do not account for the potential influence of one firm's disclosures on the disclosure behaviors of others. We now shift our focus to the value of disclosures, entry, and arrival timing, incorporating the evolution of disclosure history and allowing firms to optimize their bidding strategies. In what follows, the numbers and figures we present in this section will be based on the auctions on overlay projects from the Missoula district, which has the mode for the number of auctions across (type of construction, district)-pairs and has the median number of potential bidders totaling 12.

Value of Disclosure, Entry, and Arrival

Figure 1.6 shows how the value of disclosure evolves over time and varies across different levels of firms' construction costs. The values shown in the figure are for the case where no disclosure have been made up to the corresponding time. The results indicate that the value of disclosure decreases as time progresses. For example, a bidder with median construction cost ($c = 0.86$) would find the value of disclosure to be 1.5% of the engineer's estimate at $t = 0$. However, this value declines to 0.8% by $t = 0.5$, and ultimately becomes a loss of 0.04% at $t = 1$. This downward trend is consistent across all levels of construction costs.

At a given point in time, the value of disclosure tends to be lower for firms with higher construction costs, except when the timing approaches the end of the period at $t = 1$. For example, at $t = 0$, firms with construction costs at the 25th percentile ($c = 0.75$) would value



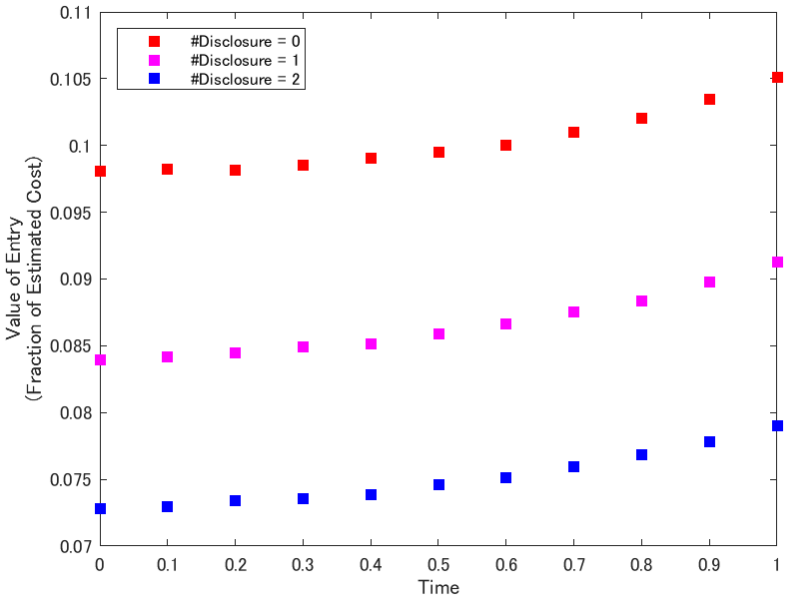
Notes: This figure shows the relationship between values of disclosure and construction costs. The results are based on auctions on overlay projects from the Missoula district. The values are for the case where there is no disclosure up to the corresponding time.

disclosure at 2.1% of the engineer's estimate, compared to 1.5% for those at the median cost ($c = 0.86$) and 0.7% for firms at the 75th percentile ($c = 0.98$). This pattern suggests that the stronger entrants– those with low construction costs– are more likely to disclose entry. As a result, the disclosures reveal not only information about firms' entry but also signal the strength of those bidders.

Finally, the value of disclosure at $t = 1$ turns out to be negative. When disclosures are made at the very end of the entry period, they fail to deter entry from others and instead prompt remaining entrants to bid more aggressively. Thus, firms disclosing at the last minute would incur a loss.

Next, we examine the value of entry for the firms over time, considering the number of disclosures available on the forum. Figure 1.7 illustrates the value of entry by time and number of disclosures. The values shown in the figure are for the case where the most recent disclosure is made at $t = 0$ if there is one. As time progresses, the value of entry increases, provided that the number of disclosures remains constant. For example, with no disclosures, the value of entry starts at 9.8% of the engineer's estimate at $t = 0$, and rises to 10.5% by $t = 1$. As we would expect, this increase reflects that the absence of disclosures becomes more informative and valuable to firms arriving later in the period.

At any given point in time, an increase in the number of disclosures reduces the value of entry for firms. For example, having one more disclosure decreases a firm's entry value by 1.3–1.4 % of the engineer's estimate, compared to the case without disclosures. This



Notes: This figure shows the relationship between values of entry and entry timing, across different numbers of disclosures. The results are based on auctions on overlay projects from the Missoula district. The values are for the case where the most recent disclosure is made at $t = 0$ if there is one.

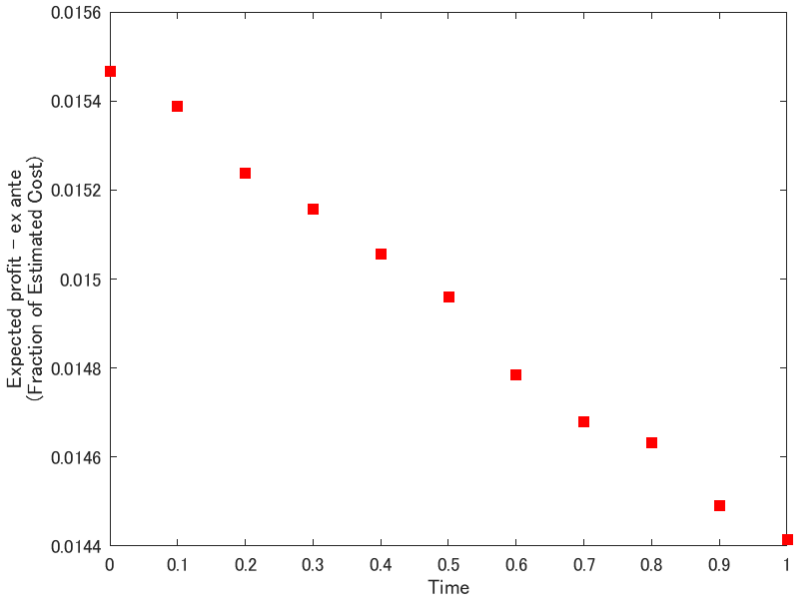
corresponds to a 4–5% drop in entry probability. A second disclosure further lowers entry value by 1.1–1.2% of the engineer's estimate, leading to an additional 5–6% reduction in entry probability.

Finally, we analyze the value of arrival timing from an *ex-ante* perspective. Figure 1.8 presents values of entry based on arrival timing, regardless of whether firms decide to enter. Arriving at $t = 0$ yields an *ex-ante* value of 1.55% of the engineer's estimate, while arriving at $t = 1$ has a 7% lower value, 1.44% of the engineer's estimate.

Two opposing effects influence the value of arrival timing. Early arrival enables a firm to enter the auction and potentially deter others from entering by making a disclosure. However, early arrivals also encounter higher uncertainty about the number of entrants. In contrast, firms arriving later can benefit from available disclosures, which provide information about whether entry is advantageous, potentially avoiding inefficient auctions. In our setup, the benefit of early arrival is more substantial, suggesting that the potential for entry deterrence through disclosure outweighs the informational advantages of arriving later.

1.8 Counterfactual Analysis

In this section, we use our model and estimates to evaluate the performance of the current platform design, the Q&A forum, in comparison to alternative platform designs. We sim-



Notes: This figure shows the values of arrival time, from an *ex-ante* perspective. The results are based on auctions on overlay projects from the Missoula district.

Table 1.7: Description of the Counterfactuals

Counterfactual	Description	Entry Deterrence	Additional Info at Bid
(0) Shutdown	Q&A never becomes public		
(1) Last minute disclosure	Q&A revealed publicly at $t = 1$		✓
(2) Status quo	Current Q&A forum	✓	✓

ulate how auction outcomes–the auctioneer's payment (conditional on the project getting allocated), the winner's construction cost, and entry behavior– would change under different designs for handling questions, or more broadly, how entry information is treated. To better understand the role of entry disclosure, we run three counterfactuals, summarized in Table 1.7.

It is important to emphasize that entry disclosure impacts the auction outcomes through two main channels: (i) entry deterrence – entry disclosure reduces the value of entry for firms arriving after the disclosure, potentially deterring their entry; and (ii) additional information at the bidding stage – a firm's disclosure may lead other entrants to bid more aggressively.

The first counterfactual, **(0) Shutdown**, corresponds to the case where we shut down the Q&A forum. In this case, the firms would communicate privately with the auctioneer

if they have any questions about the project, and those questions would not be publicly visible. Consequently, entry disclosure would not be an option, removing both the potential for entry deterrence and the added information at the bidding stage.

The second counterfactual, **(1) Last minute disclosure**, considers a case where the Q&A forum becomes public only after $t = T(= 1)$ but before the bidding window closes. Here, firms can still submit questions but their entry would only be disclosed after the entry period ends. As a result, entry disclosures would not deter other firms from entering the auction, though they would still provide additional information about each firm's entry status at the bidding stage.

Our final counterfactual, **(2) Status quo**, represents the current setup with the current Q&A forum as implemented by MDOT. Disclosures are made public immediately upon posting, potentially deterring entry from other firms. Furthermore, these disclosures provide information to entrants during the bidding stage. According to our empirical estimates, firms with lower construction costs—typically stronger bidders—are more likely to disclose, meaning that disclosures signal both entry and entrant strength.

In the following sections, we simulate outcomes for auctions on overlay projects, which constitute the most common type of auction held by MDOT, to analyze how these alternative platform designs impact key metrics such as auctioneer payments, winner's construction costs, and entry behavior. We will use scenario (0) as our benchmark for comparison.

Last Minute Disclosure

First, we discuss the auction outcomes under counterfactual (1), where the Q&A forum becomes public after $t = 1$. In the equilibrium we have estimated, firms do not engage in costly disclosures, which are made for strategic purposes. Any disclosure made are from firms requiring information for exogenous reasons. This results in a disclosure rate of 27% among the entrants.

Figure 1.9 provides estimated changes in auction outcomes compared to our benchmark case, (0) Shutdown. Moving from a no-disclosure environment (0) to counterfactual (1), auctioneer's payment increases by 0.8%, which translates to an increase by $10,000 for a median-sized project. Moreover, we observe a loss in efficiency in terms of the winner's construction cost, increasing by 1.4%. In terms of entry, we see a increase in the number of entrants by 0.6% and 3.2% increase in the total entry cost.

Through disclosures, firms are giving up their information about their own entry, which is originally private information for them. In this counterfactual, some entrants are forced to give up such information since they are in need for asking a question through the forum. However, McAfee and McMillan (1987) and Harstad et al. (1990) have shown that whether or not bidders know the set of bidders essentially has no impact on the expected payment for the auctioneer. Therefore, overall level of information about entry does not directly change the auction outcomes.

Another factor that plays an important role here is asymmetry among the bidders. Note that in our benchmark case (0), entrants are in a symmetric position since there is no addi-

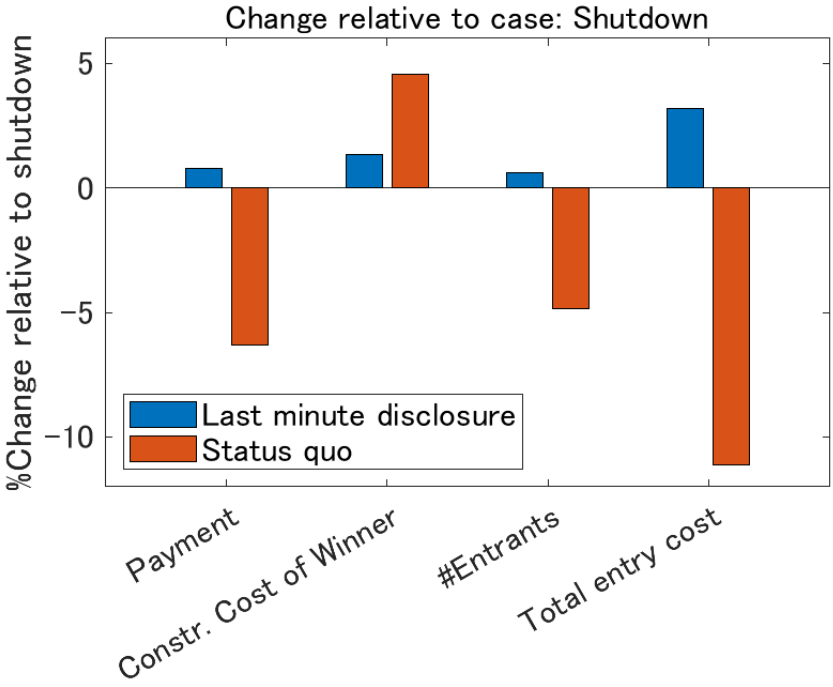
tional information for them. As long as the firms employ monotone and symmetric strategies in such scenario, the winner is the firm with the lowest construction cost, maintaining efficiency in allocation. This observation does not hold in our counterfactual case (1). Although firms are *ex-ante* symmetric, the entrants find themselves in asymmetric positions based on whether they have disclosed their entry or not. For example, let us consider a case where there are two entrants X and Y: firm X have disclosed entry, while Y have not, thereby an asymmetry in knowledge arises. Firm Y knows that firm X is participating, but firm X remains uncertain about Firm Y's entry status. This difference in information leads Y to adopt a more aggressive bidding strategy than X, ultimately introducing inefficiencies. Consequently, there are cases where Y may win the auction despite having a higher construction cost than X, resulting in inefficiency in terms of winner selection. Furthermore, auctioneer's payment increases due to this asymmetry in beliefs about entrants. While the effect of asymmetry on auctioneer can go either way (Maskin and Riley, 2000), auctioneer's payment increases in our case.

We observe a increase in number of entrants and total entry cost. As the rise in auctioneer's payment outweighs the increase in the winner's construction cost, value of entry increases in equilibrium. And as a result, number of entrants increases. While this increase in the number of entrants would counteract against the increase in payments, this force is not large enough to flip the sign. Finally, note that total entry cost increases by a larger fraction than the number of entrants because the firms who are marginal here are the firms who have the largest entry costs among the entrants.

Status quo

Next, we describe the auction outcomes under counterfactual (2), where we are in the status quo with the current Q&A forum. As we have seen in Section 1.7, the value of disclosure is more pronounced in the earlier stages of the entry period. Approximately one-third of the firms entering during the first half of the entry period choose to disclose entry, with about one-fifth of these disclosures made strategically, involving a payment of positive disclosure costs. Moreover, we have shown that firms opting to disclose tend to have lower construction costs. We find that, overall, the firms who disclose have 1.5% lower construction cost than the firms who do not disclose. Therefore, disclosures also act as a signal for strength of the firms.

The estimated changes in auction outcomes for counterfactual (2), the status quo, relative to our benchmark case, (0) Shutdown, are presented in Figure 1.9. By introducing the Q&A forum, corresponding to moving from our benchmark (0) to counterfactual (2), auctioneer's payment decreases by 6.3%, which is a substantial reduction. This corresponds to a decrease of $82,000 for a median-sized project. As in counterfactual (1), we see a loss in efficiency regarding the winner's construction cost compared to our benchmark scenario (0). Winner's construction cost increases by 4.5%, corresponding to a $38,000 increase for a median-sized project. This change is more significant than the change in counterfactual (1). In addition,



Notes: Figure shows auction outcomes under different counterfactuals. The first two bars show the change in auctioneer's payment from scenario (0) to counterfactuals (1) and (2), respectively. The next two bars show changes in winner's construction cost. The rest of the bars show changes in the number of entrants and total entry costs.

we see a decrease in the number of entrants by 4.9% (equivalent to 0.15 entrants) and 11.1% decrease in total entry costs.

In this counterfactual disclosure conveys information in two dimensions: entry status and firms' strength. Stronger firms are more likely to disclose their entry status, which means that, in their efforts to deter entry from competitors, they relinquish not only their information rents related to their private entry status but also those pertaining to their construction costs. Unlike the scenario where information pertains solely to entry, this additional information about construction costs reduces the auctioneer's payment.

Another factor contributing to the decrease in the auctioneer's payment is the coordination of entry among bidders. A situation that the auctioneer would want to avoid is where bidders assign high probabilities to being the only entrant. With the availability of the forum, firms can gain information about others' entry intentions, resulting in a smaller fraction of auctions featuring only one bidder compared to our benchmark case. We observe that this fraction decreases by 3% due to the presence of the forum, in contrast to scenarios where entry decisions are made independently.

Entrants would be placed into an asymmetric position in this counterfactual as well. Let us consider the same example, there are two entrants X and Y: firm X has disclosed entry,

while Y has not. From X's perspective, uncertainty remains regarding Y's entry status. Moreover, X recognizes that Y is likely a relatively weaker firm, as stronger firms are more prone to disclose their entry. On the other hand, Y is certain of X's entry and holds a belief that X is a relatively stronger firm. As a result, X is inclined to submit a weaker bid than it would in a situation without any information, while Y is tempted to submit a stronger bid. Therefore, this environment creates a larger gap in these firms' bidding strategies, which gives the firm with larger construction cost a greater chance of winning the auction. Consequently, we observe a larger efficiency loss in the winner's construction cost compared to counterfactual (1).

In terms of entry, we observe a significant decrease in number of entrants and total entry cost. With a decline in the auctioneer's payment and an increase in the winner's construction cost, the overall value of entry diminishes in equilibrium. The upward pressure on payments resulting from the decrease in the number of entrants is insufficient to reverse this trend.

In summary, the availability of this entry disclosure device–the Q&A forum–compels firms to engage in deterring others' entry through their disclosures. A key aspect of this setup is that firms differ along a new type dimension: arrival time. When assigned a strong type in this new dimension—meaning they arrive early—firms can effectively deter entry from competitors through their disclosures. Consequently, we observe a loss in efficiency regarding the winner's construction cost, as factors beyond construction costs now influence firms' strategies at the bidding stage. Additionally, as firms relinquish their information rents through disclosures to capitalize on their advantageous arrival time, auctioneer's payment decreases. Bernheim (1984) noted that the possibility of strategic entry deterrence has ambiguous effects on market concentration in scenarios where firms sequentially arrive at the market. In our setup, the entry disclosure device reduces the firms' expected profit from entry on average, leading to fewer entrants overall.

1.9 Conclusion

In this paper, we study how the option to disclose entry affects market outcomes by analyzing procurement auctions conducted by MDOT. We develop and estimate a model of a procurement auction with costly entry, where firms sequentially arrive at the market and make decisions on entry and disclosure. We present evidence that entry disclosure has two competing effects: it enables a firm to deter entry from others; and remaining entrants bid more aggressively in response to disclosures. Our analysis reveals that firms who disclose in early periods benefit from disclosures, as the deterrence effect of disclosure dominates the impact of more aggressive bidding from other entrants. On the other hand, late disclosures are detrimental for the firms, as aggressive bidding by opponents becomes more pronounced. We also document that early arrivals are relatively more valuable in our setting since the firms can enjoy the gains from disclosures, even though firms have informational advantage when they arrive late.

We then use our model to compare alternative platform designs. Compared to a scenario

where the Q&A forum is shut down– thus eliminating entry disclosures–the auctioneer's payment is lower under the current Q&A forum. Entry behavior becomes more efficient, as fewer firms choose to enter overall. However, this comes at the cost of efficiency in terms of the winner's construction cost. Thus, the auctioneer must carefully weigh this trade-off when deciding whether to implement such a platform. Two key factors drive these outcomes. First, the forum encourages early-arriving firms to disclose their entry status, which also signals their strength. As a result, firms forgo some information rents, leading to a reduction in the auctioneer's payment. Next, the forum allows the firms to coordinate their entry behavior, leading to a smaller fraction of auctions with few entrants. This coordination further contributes to the reduction in the auctioneer's payment.

Our analysis shows how transmission of information can alter market outcomes. We demonstrate that even a simple Q&A forum can serve as a tool for agents to transmit information strategically. Designing such a platform requires careful consideration, as it can significantly impact the market designer's objectives, such as welfare.

Chapter 2

Identification and Estimation of Dynamic Games with Unknown Information Structure

2.1 Introduction

This paper develops a computationally tractable econometric framework for estimating empirical dynamic discrete game models with weak assumptions on players' information. Let

$$f : (\Theta, \mathcal{S}) \rightrightarrows \mathcal{D} \tag{2.1}$$

describe the data-generating process under Markov perfect equilibrium, where Θ is the set of structural parameters, $\mathcal{S}$ is the set of information structures, and $\mathcal{D}$ is the data (e.g., conditional choice probabilities and state transition probabilities); the correspondence in (2.1) reflects the possibility of multiple equilibria. The standard econometric approach obtains the identified set by inverting (2.1) after assuming that the true information structure takes a particular form, namely that players observe their payoff shocks but not their opponents' (Rust, 1994; Aguirregabiria and Mira, 2007; Bajari, Benkard and Levin, 2007; Pakes, Ostrovsky and Berry, 2007; Pesendorfer and Schmidt-Dengler, 2008; Bugni and Bunting, 2021; Dearing and Blevins, 2024).[1]

However, assuming the true information structure is *known* to the researcher is far from innocuous. While such an assumption has facilitated computationally tractable estimation of dynamic games, researchers often have little evidence to defend the assumptions. Recent works have documented that misspecification of players' information may create substantial bias in both static and dynamic settings (Fershtman and Pakes, 2012; Grieco, 2014; Pakes et al., 2015; Aguirregabiria and Magesan, 2020; Syrgkanis, Tamer and Ziani, 2021; Gualdani and Sinha, 2024; Magnolfi and Roncoroni, 2023; Koh, 2023). The information structure misspecification problem appears especially serious in dynamic game settings because fewer

[1] See Aguirregabiria, Collard-Wexler and Ryan (2021) for a recent survey of the literature.

informational assumptions have been entertained in econometric works; we are unaware of an econometric framework that considers an alternative information assumption other than Ericson and Pakes (1995). Unfortunately, attempts to invert (2.1) directly are likely futile since the set of information structures is large and infinite-dimensional.

In this paper, we develop a tractable empirical framework for analyzing dynamic discrete games while treating the underlying information structure as *unknown* to the analyst. We assume that the researcher only knows minimal information available to the players but remains agnostic about the possibility that the players might observe more signals.[2] The framework typically yields a set of partially identified parameters robust to misspecification on the information structure.

The main innovation of this paper is the development of *Markov correlated equilibrium*, a dynamic (Markovian) analog of Bayes correlated equilibrium (Bergemann and Morris, 2016). We show that Markov correlated equilibrium can capture the set of Markov perfect equilibrium predictions (the joint distribution of actions and states) that can arise when the players might observe more signals than assumed by the analyst. We establish the informational robustness property of Markov correlated equilibrium by formulating Markov perfect equilibrium as the Bayes Nash equilibrium of the associated reduced one-shot game via the one-shot deviation principle and invoking Theorem 1 of Bergemann and Morris (2016).

We build on the theory of Markov correlated equilibrium and propose computationally tractable approaches for estimation and inference. Roughly speaking, we prove

$$\Theta_I^{MCE} = \bigcup_{\tilde{S} \in \mathcal{S}} \Theta_I^{MPE}\left(\tilde{S}\right), \tag{2.2}$$

where Θ_I^{MCE} is the identified set of parameters after assuming that a Markov correlated equilibrium generated the data, and $\Theta_I^{MPE}(\tilde{S})$ is the Markov perfect equilibrium identified set when the information structure is assumed to be $\tilde{S}$. Directly inverting (2.1) requires constructing the right-hand side of (2.2), which is computationally infeasible when $\mathcal{S}$ is large. However, we show that constructing the left-hand side of (2.2) is computationally tractable. In sum, treating the data *as if* they were generated by a Markov correlated equilibrium enables an informationally robust econometric analysis. Identified set (2.2) is usually non-singleton. We propose a computationally attractive strategy for estimation and inference by formulating the estimation problem as mathematical programs with equilibrium constraints (MPEC) (Su and Judd, 2012) and applying the inference strategy of Horowitz and Lee (2023) by nesting the confidence set for the conditional choice probabilities into the MPEC problem.[3]

[2]Note that this assumption restricts the set of possible information structure $\mathcal{S}$ in (2.1). Our framework allows $\mathcal{S}$ to span the entire set of information structures, but the corresponding identified set tends to be large and uninformative. We focus on more restrictive sets in our empirical application.

[3]The standard estimation strategies for dynamic games are inapplicable (e.g., the nested fixed-point algorithm of Rust (1987) or the conditional choice probability inversion of Hotz and Miller (1993)).

We apply our framework to study the dynamic entry and exit decisions of Starbucks and Dunkin', the two largest coffee chains in the US. We estimate a tight identified set by leveraging the large support idea of Tamer (2003). Specifically, we estimate the firms' profit parameters by identifying markets where they operate as single players and applying the standard dynamic discrete choice estimation methods. Next, we estimate the competitive effects parameters under a Markov correlated equilibrium assumption using the data from the markets where the two firms compete. The single-agent profit parameters are point-estimated, but the competitive effects parameters are set-estimated. We also estimate the parameters under the standard incomplete information Markov perfect equilibrium assumption.

Comparison between the estimates from the standard incomplete information Markov perfect equilibrium assumption and our informationally robust Markov correlated equilibrium assumption suggests that the standard incomplete information assumption may be misspecified. The estimates from the standard approach suggest Dunkin' likes Starbucks' presence, while Starbucks is not affected by Dunkin's presence. However, the Markov correlated equilibrium identified set shows that these conclusions are not robust to the misspecification of information structures. The Markov correlated equilibrium identified set does not rule out both positive and negative spillover effects.

We also conduct a counterfactual experiment in which we cut the firms' entry costs and simulate the evolution of the number of active firms in our sample period. We consider two parameter values that capture the opposite extremes of the Markov correlated equilibrium identified set estimates. The two cases capture scenarios in which firm entry decisions are strategic substitutes and complements. We find that the counterfactual outcomes can vary widely depending on the true value of the competitive effects parameters (which we cannot determine under the Markov correlated equilibrium assumption in the parameter estimation). However, we find that for a fixed parameter value, informationally robust bounds on counterfactual outcomes are tight, indicating that an information designer cannot affect the outcomes much. These tight counterfactual outcome bounds occur because the competitive effects parameters are small in absolute value, leaving the information designer with little room to affect the firms' decisions relative to the case where they operate as single agents who do not account for the opponents' entry decisions strategically.

Our work advances the literature on the econometric analysis of game-theoretic models in several ways. First, we contribute to the dynamic games literature by developing a new solution concept with attractive computational properties. Markov correlated equilibrium is a natural extension of Bayes correlated equilibrium (Bergemann and Morris, 2013, 2016) to a class of dynamic games that build on Maskin and Tirole (2001). While various versions of correlated equilibrium in dynamic games have been studied (e.g., see Nowak and Raghavan (1992), Duffie et al. (1994), Solan (2001), Solan and Vieille (2002), Mailath and Samuelson (2006), Von Stengel and Forges (2008), He and Sun (2017), Forges (2020) and the works cited therein), our version is new as we extend Bayes correlated equilibrium to a class of dynamic stochastic Markov games with incomplete information. Doval and Ely (2020) and Makris and Renou (2023) also extend Bayes correlated equilibrium to dynamic environments, but

our setting is different as we focus on a class of dynamic discrete games common in empirical works.

This paper also contributes to the literature on misspecification-robust econometric analysis of game-theoretic models. Our work is most closely related to the econometric frameworks that pursue robustness to misspecification on information structure (Grieco, 2014; Pakes et al., 2015; Syrgkanis, Tamer and Ziani, 2021; Magnolfi and Roncoroni, 2023; Koh, 2023; Gualdani and Sinha, 2024; Han, Kaido and Magnolfi, 2024). We contribute by developing a novel framework for estimating a class of dynamic games, building on the aforementioned works on the econometric analysis of dynamic Markovian games. Our work also relates to works that develop econometric approaches for estimating incomplete models with multiple equilibria (e.g., Tamer (2003), Ciliberto and Tamer (2009), Beresteanu, Molchanov and Molinari (2011), and Galichon and Henry (2011)) since our approach does not require an assumption on equilibrium selection.

The rest of the paper is organized as follows. In Section 2.2, we define Markov correlated equilibrium and establish a connection between Markov correlated equilibria and Markov perfect equilibria in a class of dynamic games and discuss its properties. In Section 2.3, we introduce an econometric model and discuss informationally robust identification. Section 2.4 proposes strategies for computing the identified set. Section 2.5 discusses our approach to inference. Section 2.6 applies our framework to the dynamic entry/exit game by Starbucks and Dunkin' in the US. Section 2.7 concludes. All proofs are in Appendix B.1.

2.2 Theory of Markov Correlated Equilibrium

We consider a class of dynamic Markovian games that has been used as a standard framework for empirical works. We introduce the concept of Markov correlated equilibrium for informationally robust analysis and discuss its properties.

Setup

Let $t = 1, 2, ..., \infty$ denote discrete time. A *stationary dynamic Markov game of incomplete information* is defined by a pair of basic game G and information structure S. A basic game $G = \langle \mathcal{I}, (\mathcal{A}_i, u_i)_{i\in\mathcal{I}}, \mathcal{X}, \mathcal{E}, \psi, f, \delta \rangle$ specifies the payoff-relevant primitives of the model: $i \in \mathcal{I} = \{1, 2, ..., I\}$ indexes the players; $a_{it} \in \mathcal{A}_i$ denotes player i's action at time t, where $\mathcal{A}_i$ is a finite set of actions available to player i; $a_t = (a_{1t}, ..., a_{It}) \in \mathcal{A} \equiv \times_{i\in\mathcal{I}}\mathcal{A}_i$ denotes an action profile; $x_t \in \mathcal{X}$ denotes a state variable that is publicly observed by the players, and $\mathcal{X}$ is assumed to be finite; $\varepsilon_t \in \mathcal{E}$ denotes a latent state variable that is not directly observed by the players, and $\mathcal{E}$ is assumed to be finite[4]; $\psi_{\varepsilon_t|x_t}$ represents the players' prior on

[4]The finite state space assumption is non-material and simplifies the notation. The assumption also facilitates the transition from theory to econometrics since the state space will be discretized for feasible estimation. There is no conceptual difficulty in extending the theory to a more general space; see Syrgkanis, Tamer and Ziani (2021) and Magnolfi and Roncoroni (2023) for the case of Bayes correlated equilibrium.

the probability of ε_t conditional on x_t; $f_{x_{t+1}|a_t,x_t,\varepsilon_t}$ specifies the probability of transitioning to state x_{t+1} conditional on $(a_t, x_t, \varepsilon_t)$; $u_i : \mathcal{A} \times \mathcal{X} \times \mathcal{E} \to \mathbb{R}$ is the flow payoff function of player i; $\delta \in [0, 1)$ is the players' discount factor.

An information structure $S = ((\mathcal{T}_i)_{i\in\mathcal{I}}, \pi)$ specifies the information-relevant primitives; $\tau_{it} \in \mathcal{T}_i$ denotes player i's private signal at period t, where $\mathcal{T}_i$ represents a finite set of private signals; $\tau_t = (\tau_{1t}, ..., \tau_{It}) \in \mathcal{T} \equiv \times_{i\in\mathcal{I}}\mathcal{T}_i$ is a signal profile; $\pi : \mathcal{X} \times \mathcal{E} \to \Delta(\mathcal{T})$ is a signal distribution that maps the state of the world to a signal profile. Note that players' signals are allowed to be correlated. The interpretation is that each player i does not directly observe the latent state ε_t but receives a private signal τ_{it} whose informativeness about ε_t depends on π.

The game (G, S) is common knowledge to the players. The model proceeds as follows. At the beginning of period t, the common knowledge state $x_t \in \mathcal{X}$ is given and publicly observed. Conditional on x_t, the latent state $\varepsilon_t \in \mathcal{E}$ is drawn from the probability distribution $\psi_{\varepsilon_t|x_t}$. Next, a profile of private signals $\tau_t = (\tau_{1t}, ..., \tau_{It}) \in \mathcal{T}$ is drawn from the signal distribution $\pi_{\tau_t|x_t,\varepsilon_t}$. At this point, each player i observes (x_t, τ_{it}) and uses Bayes' rule to infer ε_t. Then, the players simultaneously choose actions a_{it}, $i = 1, ..., I$, and each player i receives period t payoff $u_i(a_t, x_t, \varepsilon_t) \in \mathbb{R}$. Finally, the observable state x_t transitions to x_{t+1} via the probability kernel $f_{x_{t+1}|a_t,x_t,\varepsilon_t}$. Given x_{t+1}, period $t + 1$ begins.

Note that the primitives imply that the transition probability of state variables factors as $\Pr(x_{t+1}, \varepsilon_{t+1}|a_t, x_t, \varepsilon_t) = \psi_{\varepsilon_{t+1}|x_{t+1}} f_{x_{t+1}|a_t,x_t,\varepsilon_t}$. Assuming that the latent variables are generated independently of the previous-period states is standard in the empirical literature.[5] The assumption also plays a vital role in this paper as it simplifies the characterization of correlated equilibria in dynamic settings.

The players are rational and forward-looking. In period t, each player i chooses action plans to maximize the expected discounted sum of intertemporal payoffs

$$\mathbb{E}\bigg[\sum_{s=t}^{\infty} \delta^{s-t} u_i(a_s, x_s, \varepsilon_s)|x_t, \tau_{it}\bigg].$$

The conditioning on (x_t, τ_{it}) reflects the assumption that i observes the public state x_t and her signal τ_{it} before choosing period t action. Since the environment is stationary, we suppress time subscripts unless necessary.

Example 1 (Two-player dynamic entry game)**.** As a running example, we consider the two-player dynamic entry/exit model of Pesendorfer and Schmidt-Dengler (2008), which has served as a representative example in the literature (Egesdal, Lai and Su, 2015; Aguirregabiria et al., 2021; Dearing and Blevins, 2024). We also adopt a similar model in our empirical application. There are two players, $i = 1, 2$. In each period, firms simultaneously

[5]When the latent variables are correlated over time so that information asymmetries become persistent, computing the perfect Bayesian equilibria easily becomes intractable because the players need to keep track of the entire history and apply Bayes rule for each contingent. See Fershtman and Pakes (2012) for a discussion.

decide whether to be active ($a_{it} = 1$) or inactive ($a_{it} = 0$). The per-period payoff is

$$u_i(a_{it}, a_{jt}, z_{it}, \varepsilon_{it}) = \begin{cases} (1 - a_{jt})\pi^m + a_{jt}\pi^d + (1 - z_{it})c + \varepsilon_{it} & \text{if } a_{it} = 1 \\ z_{it}\kappa & \text{if } a_{it} = 0 \end{cases}. \tag{2.3}$$

Here, π^m, π^d, c, and κ represent monopoly profit, duopoly profit, entry cost, and scrap value, respectively. The public state variable is $x_t = (z_{1t}, z_{2t})$, where z_{it} represents whether firm i is an incumbent ($z_{it} = 1$) or not ($z_{it} = 0$). The firms' previous actions determine the incumbency states as $z_{it} = a_{i,t-1}$. In empirical applications, x_t usually contains additional observable market characteristics. The latent state variable is $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})$, where $\varepsilon_{it} \in \mathcal{E}_i \subseteq \mathbb{R}$ only enters player i's payoff. Each ε_{it} independently follows the standard normal distribution.

Standard empirical dynamic discrete game models assume that each player observes ε_{it} but not ε_{jt} for $j \neq i$. Formally, $\mathcal{T}_i = \mathcal{E}_i$, and the private signal is $\tau_{it} = \varepsilon_{it}$ with probability one. Alternatively, if the information is complete, players publicly observe ε_t, so $\mathcal{T}_i = \mathcal{E}$, and $\tau_{it} = \varepsilon_t$ with probability one.

Markov Perfect Equilibrium

A Markov strategy of player i is a mapping $\beta_i : \mathcal{X} \times \mathcal{T}_i \to \Delta(\mathcal{A}_i)$ that specifies a probability distribution over actions at each realization of observable state and private signal.[6] A strategy profile $\beta = (\beta_1, ..., \beta_I)$ is a *Markov perfect equilibrium* if it is a subgame perfect equilibrium.

To leverage the results in static games to a dynamic environment, it is useful to characterize the Markov perfect equilibrium conditions using the one-shot deviation principle.[7] The one-shot deviation principle formulation has two advantages. First, we can leverage results established in static environments as it translates a Markov perfect equilibrium as a Bayes-Nash equilibrium of the associated normal-form game. Second, it facilitates equilibrium computation during econometric analysis.

[6]The strategies are Markov in the sense that they are memoryless and depend only on the current state. Although the standard Markov strategy assumption only allows players to condition on payoff-relevant state (Maskin and Tirole, 2001), our definition allows players' strategy to depend on signals that may be unrelated to the fundamentals (e.g., a public correlation device under complete information), which is a consequence of considering arbitrary information structures. However, our assumptions align with the spirit of Maskin and Tirole (2001) because we rule out arbitrary strategies by requiring actions to depend only on signals of a memoryless latent state.

[7]One-shot deviation principle states that a strategy profile is subgame-perfect if and only if no player has an information set at which a one-shot deviation is profitable. A well-known sufficient condition for the one-shot deviation principle for infinite-horizon games is the "continuity at infinity" condition, which states that events in the distant future are relatively unimportant (Fudenberg and Tirole, 1991). Standard dynamic discrete game frameworks satisfy the condition because the overall payoffs are a discounted sum of per-period payoffs, and the expected stage payoffs that arbitrary strategies can attain are uniformly bounded (even though the payoffs themselves might have unbounded support due to payoff shocks with unbounded support).

Suppose a strategy profile $\beta = (\beta_1, ..., \beta_I)$ is given. Define the *ex-ante value function* $V_i^\beta(x) \in \mathbb{R}$ as the expected payoff to player i at state x prior to the realization of ε when all players follow the prescriptions in β. For each $i \in \mathcal{I}$, $V_i^\beta \in \mathbb{R}^{|\mathcal{X}|}$ is the unique solution to

$$V_i^\beta(x) = \sum_{\varepsilon \in \mathcal{E}, \tau \in \mathcal{T}, a \in \mathcal{A}} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \beta_{a|x,\tau} \left\{ u_i(a, x, \varepsilon) + \delta \sum_{x' \in \mathcal{X}} V_i^\beta(x') f_{x'|a,x,\varepsilon} \right\}, \quad \forall x \in \mathcal{X} \tag{2.4}$$

where $\beta_{a|x,\tau} \equiv \prod_{i=1}^I \beta_i(a_i|x, \tau_i)$ denotes the conditional distribution over action profiles induced by the strategy profile β. Next, define the *outcome-specific payoff function* given strategy profile β as

$$v_i^\beta(a, x, \varepsilon) \equiv u_i(a, x, \varepsilon) + \delta \sum_{x' \in \mathcal{X}} V_i^\beta(x') f_{x'|a,x,\varepsilon},$$

which represents the continuation payoff to player i when (a, x, ε) is realized today, and all players follow the prescriptions in β from tomorrow onward.

A strategy profile β in game (G, S) induces a reduced normal-form basic game $G^\beta = \langle \mathcal{I}, (\mathcal{A}_i, v_i^\beta)_{i \in \mathcal{I}}, \mathcal{X}, \mathcal{E}, \psi \rangle$. Then (G^β, S) describes a "static" game in which player i's "static" payoff function is given by $v_i^\beta : \mathcal{A} \times \mathcal{X} \times \mathcal{E} \to \mathbb{R}$. The following lemma allows us to import the results in static environments into dynamic environments.

Lemma 1 (One-shot deviation principle formulation of Markov perfect equilibrium)**.** *A strategy profile β is a Markov perfect equilibrium of (G, S) if and only if β is a Bayes Nash equilibrium of (G^β, S).*

Markov Correlated Equilibrium

We introduce a dynamic analog of Bayes correlated equilibrium for dynamic discrete games. A *stationary Markov decision rule* in (G, S) is a mapping

$$\sigma : \mathcal{X} \times \mathcal{E} \times \mathcal{T} \to \Delta(\mathcal{A}),$$

which specifies a probability distribution over action profiles at each realization of state and signals. It is instructive to think of σ as a recommendation strategy of an omniscient mediator who observes (x, ε, τ) and privately recommends action to each player. Suppose the mediator commits to a recommendation strategy σ and announces it to the players. After the state and players' signal (x, ε, τ) are realized, an action profile $a = (a_1, ..., a_I)$ is drawn from the probability distribution $\sigma_{a|x,\varepsilon,\tau}$, and each a_i is privately recommended to each player i. Each player i, having observed (x, τ_i, a_i), decides whether to obey (play a_i) or not (deviate to $a_i' \neq a_i$). If the decision rule σ is such that the players are always obedient, we call σ a Markov correlated equilibrium of (G, S).

We formalize the equilibrium condition as follows. Let $V_i^\sigma \in \mathbb{R}^{|\mathcal{X}|}$ denote the vector of ex-ante value functions induced by σ, obtained as a unique solution to

$$V_i^\sigma(x) = \sum_{\varepsilon \in \mathcal{E}, \tau \in \mathcal{T}, a \in \mathcal{A}} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \sigma_{a|x,\varepsilon,\tau} \left\{ u_i(a, x, \varepsilon) + \delta \sum_{x' \in \mathcal{X}} V_i^\sigma(x') f_{x'|a,x,\varepsilon} \right\}, \quad \forall x \in \mathcal{X}. \tag{2.5}$$

Define the outcome-specific value function associated with σ as

$$v_i^\sigma(a,x,\varepsilon) \equiv u_i(a,x,\varepsilon) + \delta \sum_{x' \in \mathcal{X}} V_i^\sigma(x') f_{x'|a,x,\varepsilon},$$

which represents the payoff to player i if (a,x,ε) is realized today and σ determines the players' actions in the future. Let $\mathbb{E}^\sigma[v_i^\sigma(a_i',a_{-i},x,\varepsilon)|x,\tau_i,a_i]$ be the expected payoff to player i from choosing a_i' when i observes (x,τ_i) and receives recommendation a_i. The following definition states that σ is a Markov correlated equilibrium of (G,S) if the players do not deviate from the mediator's recommendations.

Definition 1. *A decision rule σ is a Markov correlated equilibrium of (G,S) if for each $i \in \mathcal{I}$, $x \in \mathcal{X}$, $\tau_i \in \mathcal{T}_i$, and $a_i \in \mathcal{A}_i$, we have*

$$\mathbb{E}^\sigma[v_i^\sigma(a_i,a_{-i},x,\varepsilon)|x,\tau_i,a_i] \geq \mathbb{E}^\sigma[v_i^\sigma(a_i',a_{-i},x,\varepsilon)|x,\tau_i,a_i] \tag{2.6}$$

for each $a_i' \neq a_i$ whenever $\Pr^\sigma(x,\tau_i,a_i) > 0$.

Since

$$\begin{aligned}\mathbb{E}^\sigma[v_i(\tilde{a}_i,a_{-i},x,\varepsilon)|x,\tau_i,a_i] &= \sum_{\varepsilon,a_{-i}} v_i^\sigma(\tilde{a}_i,a_{-i},x,\varepsilon)\Pr^\sigma(\varepsilon,a_{-i}|x,\tau_i,a_i) \\ &= \sum_{\varepsilon,a_{-i}} v_i^\sigma(\tilde{a}_i,a_{-i},x,\varepsilon)\left(\frac{\sum_{\tau_{-i}}\psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}}{\sum_{\varepsilon,\tau_{-i},a_{-i}}\psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}}\right),\end{aligned}$$

canceling out the denominator, which is constant across all possible realizations of $(\varepsilon,\tau_{-i},a_{-i})$—lets us rewrite the obedience condition (2.6) as, for each i,

$$\sum_{\varepsilon,\tau_{-i},a_{-i}} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}v_i^\sigma(a,x,\varepsilon) \geq \sum_{\varepsilon,\tau_{-i},a_{-i}} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}v_i^\sigma(a_i',a_{-i},x,\varepsilon), \quad \forall x,\tau_i,a_i,a_i'. \tag{2.7}$$

Equation (2.7) serves as the operational definition of Markov correlated equilibrium in our econometric analysis.

A Markov correlated equilibrium always exists. Nowak and Raghavan (1992) establishes the existence of a stationary correlated equilibrium in a discounted stochastic game with complete information under standard regularity conditions. We will show in the following sections that the set of Markov correlated equilibria under arbitrary information structure always includes a complete information stationary correlated equilibrium. Thus, the set of Markov correlated equilibria is always non-empty.

Relationship to Bayes Correlated Equilibrium

As in Section 2.2, a decision rule σ in game (G,S) induces a reduced normal-form basic game $G^\sigma = \langle \mathcal{I}, (\mathcal{A}_i, v_i^\sigma)_{i\in\mathcal{I}}, \mathcal{X}, \mathcal{E}, \psi\rangle$. Then (2.7) shows that if σ is a Markov correlated equilibrium of (G,S), it is a Bayes correlated equilibrium of (G^σ,S).

Lemma 2 (One-shot deviation principle characterization of Markov correlated equilibrium)**.** *A decision rule σ is a Markov correlated equilibrium of (G, S) if and only if it is a Bayes correlated equilibrium of (G^σ, S).*

When players fully discount the future ($\delta = 0$), a Markov correlated equilibrium collapses to a Bayes correlated equilibrium of the stage game: for each i,

$$\sum_{\varepsilon,\tau_{-i},a_{-i}} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}u_i(a,x,\varepsilon) \geq \sum_{\varepsilon,\tau_{-i},a_{-i}} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}u_i(a_i',a_{-i},x,\varepsilon), \quad \forall x,\tau_i,a_i,a_i'.$$

Thus, Markov correlated equilibrium is a proper dynamic analog of Bayes correlated equilibrium.

Unfortunately, contrary to its static analog, the Markov correlated equilibrium conditions are not linear with respect to the decision rule, which creates two challenges that are absent in the static case. First, Markov correlated equilibrium is more difficult to compute. While computing a Bayes correlated equilibrium amounts to solving a linear program, computing a Markov correlated equilibrium generally requires solving a non-convex program. Second, the set of Bayes correlated equilibria is convex, but the set of Markov correlated equilibria is generally non-convex. The convexity of the set of Bayes correlated equilibria implies that a mixture of multiple equilibria is observationally equivalent to a single equilibrium; Syrgkanis, Tamer and Ziani (2021) and Magnolfi and Roncoroni (2023) use this property to stay agnostic about the true equilibrium selection rule in the data generating process. However, since Markov correlated equilibrium does not share this property, assuming a single equilibrium selection rule is not without loss.

Informational Robustness

We establish the informational robustness property of Markov correlated equilibrium. Let an equilibrium prediction refer to the joint distribution on actions, states, and signals that are implied by either an equilibrium strategy profile or decision rule. We prove the following claim.

Claim 1 (Markov correlated equilibrium is informationally robust)**.** *The set of Markov correlated equilibrium predictions is equal to the set of Markov perfect equilibrium predictions that can arise when players can observe more signals that are unknown to the analyst.*

In the empirical application, we consider a scenario where the analyst knows that each firm observes its firm-specific payoff shock but cannot rule out the possibility that they receive some signals about opponents' shocks. Claim 1 says that the analyst can capture the set of all predictions that can arise in such a scenario by resorting to Markov correlated equilibrium. Thus, as in Bayes correlated equilibrium, Markov correlated equilibrium attains a form of informational robustness that allows the analyst to be agnostic about the players observing additional signals after specifying players' minimal information. We formalize

Claim 1 by following Bergemann and Morris (2016)'s framework and connecting Markov correlated equilibrium to Bayes correlated equilibrium via the one-shot deviation principle.

First, we introduce a partial order on the set of information structures using the notion of *expansion* as in Bergemann and Morris (2016). Let $\omega \in \Omega \equiv \mathcal{X} \times \mathcal{E}$ denote the state of the world.

Definition 2 (Expansion)**.** *Let $S = (\mathcal{T}, \pi)$ be an information structure. Information structure $S^* = (\mathcal{T}^*, \pi^*)$ is an expansion of S, or $S^* \succsim_E S$, if there exists $(\tilde{\mathcal{T}}_i)_{i\in\mathcal{I}}$ and $\lambda : \Omega \times \mathcal{T} \to \Delta(\tilde{\mathcal{T}})$ such that $\mathcal{T}_i^* = \mathcal{T}_i \times \tilde{\mathcal{T}}_i$ for all $i \in \mathcal{I}$ and $\pi^*_{\tau,\tilde{\tau}|\omega} = \pi_{\tau|\omega}\lambda_{\tilde{\tau}|\omega,\tau}$.*

Intuitively, if $S^* \succsim_E S$, the players receive *extra* signals in S^* than they do in S. Specifically, complete information (players observe the latent state ε with probability one) is an expansion of arbitrary information structure.

Next, we claim that if a strategy profile induces the same joint distribution over states and actions as a decision rule, the two induce the same reduced normal-form games. Let S^* be an expansion of S. A strategy profile β in (G, S^*) *induces* a decision rule σ for (G, S) if

$$\sigma_{a|x,\varepsilon,\tau} = \sum_{\tilde{\tau}\in\tilde{\mathcal{T}}} \lambda_{\tilde{\tau}|x,\varepsilon,\tau} \prod_{i=1}^{I} \beta_i(a_i|x, \tau_i, \tilde{\tau}_i), \quad \forall a, x, \varepsilon, \tau.$$

The strategy profile β in a game with more signals induces the same joint distribution over actions, states, and signals as the decision rule σ in a game with fewer signals.

Lemma 3 (Identical reduced normal-form basic games)**.** *If β induces σ, then the associated reduced normal-form basic games, G^β and G^σ, are identical.*

Lemma 3 is intuitive: if β induces σ, then β and σ induce identical distributions over action profiles at each state, so the associated ex-ante value functions must be identical to each other, i.e., $V_i^\beta = V_i^\sigma$. Then, the outcome-specific payoff functions are also identical to each other, i.e., $v_i^\beta = v_i^\sigma$, making all primitives of the basic games identical to each other.

Finally, a solution concept generates a *prediction* defined as a probability distribution over actions at each state and signal. Let $\mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G, S)$ represent the set of predictions that can be induced by a Markov perfect equilibrium in (G, S).[8] Let $\mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G, S)$ be defined similarly for a Markov correlated equilibrium.

We are ready to state our main theorem. Suppose the analyst knows that the players observe signals specified by S but does not know whether the players have access to extra signals. What is the set of feasible Markov perfect equilibrium predictions? The following theorem says that the set of Markov correlated equilibria of (G, S) captures all Markov perfect equilibrium predictions when the players can observe more signals than in S.

[8]We slightly abuse the notation and use $\mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G, S)$ to denote a set of conditional distributions over $\mathcal{A}$ defined at each (x, ε, τ).

Theorem 1 (Markov correlated equilibrium is informationally robust)**.** *For any basic game G and information structure S,*

$$\mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G,S) = \bigcup_{S^* \succsim_E S} \mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G,S^*).$$

Theorem 1 is attractive because $\mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G,S)$ is far easier to characterize and compute than $\bigcup_{S^* \succsim_E S} \mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}$; the former does not require searching over the set of information structures. Note that when $\delta = 0$, Theorem 1 reduces to

$$\mathcal{P}^{BCE}_{a|x,\varepsilon,\tau}(G,S) = \bigcup_{S^* \succsim_E S} \mathcal{P}^{BNE}_{a|x,\varepsilon,\tau}(G,S^*),$$

which is analogous to Bergemann and Morris (2016) Theorem 1.

In empirical works, it is common to assume that the analyst observes the *conditional choice probabilities* that represent the probability of each action profile at each state that is observable to the econometrician. Thus, it is useful to characterize the implications of Theorem 1 in terms of conditional choice probabilities. Let $\mathcal{P}^{MPE}_{a|x}(G,S)$ denote the set of feasible conditional choice probabilities that can arise under a Markov perfect equilibrium of (G,S). Let $\mathcal{P}^{MCE}_{a|x}(G,S)$ be defined similarly.

Corollary 1. *For any basic game G and information structure S,*

$$\mathcal{P}^{MCE}_{a|x}(G,S) = \bigcup_{S^* \succsim_E S} \mathcal{P}^{MPE}_{a|x}(G,S^*).$$

As in the case of Bayes correlated equilibrium, if players observe more signals, the set of Markov correlated equilibria becomes smaller. In other words, more information implies a smaller set of Markov correlated equilibria.

Theorem 2. *Let G be an arbitrary basic game. If $S^1 \succsim_E S^2$, then $\mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G,S^1) \subseteq \mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G,S^2)$.*

Numerical Example

We continue Example 1 to illustrate the Markov correlated equilibrium identified set. The true parameters in (2.3) are set to $(\pi^m, \pi^d, c, \kappa) = (1.2, -1.2, -0.2, 0.1)$, and we use the discount factor $\delta = 0.9$. We generate data by finding a Markov perfect equilibrium under the standard assumption that ε_i is private information to player i. Pesendorfer and Schmidt-Dengler (2008) finds five equilibria, two of which are symmetric to other equilibria. We report the conditional choice probabilities of equilibria (i), (ii), and (iii) of Pesendorfer and Schmidt-Dengler (2008) in Table 2.1.[9]

[9]We use Dearing and Blevins (2024)'s MATLAB code to generate the equilibrium conditional choice probabilities.

Table 2.1: Probability of entry in three Markov perfect equilibria

	(i)		(ii)		(iii)	
	P1	P2	P1	P2	P1	P2
Out/Out	0.733	0.276	0.615	0.528	0.576	0.576
Out/In	0.613	0.420	0.312	0.840	0.305	0.842
In/Out	0.800	0.223	0.831	0.303	0.842	0.305
In/In	0.752	0.294	0.606	0.578	0.595	0.595

Note: Each value represents the entry probability of each player at each state. P1 and P2 represent player 1 and player 2.

To characterize the bounds on the game outcomes that can arise in a Markov correlated equilibrium, we compute the Markov correlated equilibria that maximize and minimize the expected number of active firms.[10] Using the associated conditional choice probabilities, we simulate the number of active firms over $T = 6$ periods, assuming that the initial state has no incumbent, i.e., $x_1 = (0, 0)$.

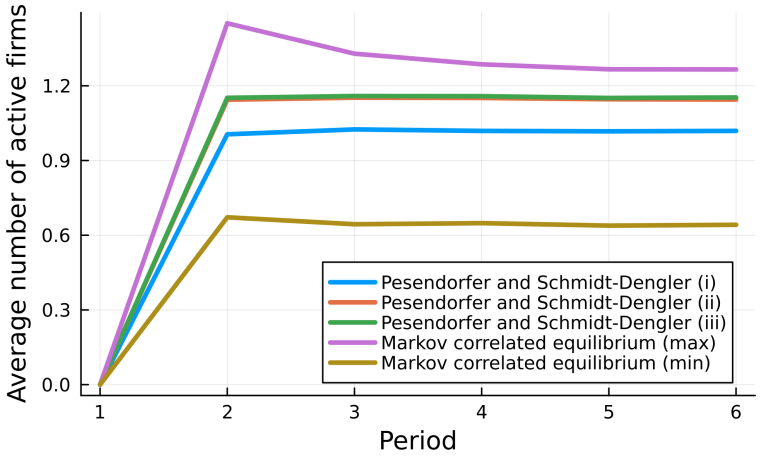


Figure 2.1: Average number of active firms over time

Figure 2.1 plots the simulated average number of active firms for $T = 6$ periods. Note that the outcomes of Pesendorfer and Schmidt-Dengler (2008)'s three Markov perfect equilibria are bounded by those of the Markov correlated equilibria. The Markov correlated equilibrium bounds characterize the sharp bounds on the Markov perfect equilibrium outcomes that can arise when players can receive information about the opponent's payoff shock.[11]

[10] Specifically, let $\phi^\sigma(a_1, a_2|x)$ be the conditional choice probability implied by decision rule σ. The expected number of active firms at state x is $0 \cdot \phi^\sigma(0,0|x) + 1 \cdot (\phi^\sigma(0,1|x) + \phi^\sigma(1,0|x)) + 2 \cdot \phi^\sigma(1,1|x)$. We find the Markov correlated equilibria that maximize and minimize the unweighted average over x.

[11] An alternative interpretation is that the Markov correlated equilibrium bounds characterize the limits to which an omniscient information designer who observes players' information can manipulate the outcome.

2.3 Informationally Robust Identification

In this section, we specialize the model for econometric analysis and discuss informationally robust identification in dynamic discrete games using Markov correlated equilibrium.

Setup

Econometric Assumptions

To facilitate econometric analyses, we impose the following assumptions on the basic game primitives. As commonly assumed in empirical works, we let $\varepsilon \equiv (\varepsilon_1, \varepsilon_2, ..., \varepsilon_I)$ where $\varepsilon_i \in \mathcal{E}_i$ only enters player i's payoff and $f_{x'|a,x,\varepsilon} = f_{x'|a,x}$, i.e., the transition probability of public state x is independent of latent state ε. We assume that the payoff functions and the prior distribution are parameterized by a finite-dimensional vector θ so that $u_i = u_i^\theta$ and $\psi = \psi^\theta$.

The econometrician observes data

$$\{a_{m,t}, x_{m,t} : \ m = 1, 2, ..., M, \ t = 1, 2, ..., T\},$$

which have players' actions $a_{m,t}$ and common knowledge state variables $x_{m,t}$ across M independent markets over T periods. We assume that M is large and that the conditional choice probabilities $\phi : \mathcal{X} \to \Delta(\mathcal{A})$ and the transitional probability function $f : \mathcal{A} \times \mathcal{X} \to \Delta(\mathcal{X})$ can be non-parametrically identified from the data. Finally, we assume that the common discount factor $\delta \in [0, 1)$ is known to the econometrician. In sum, we treat the basic game $G^\theta = \langle \mathcal{I}, (\mathcal{A}_i, u_i^\theta)_{i \in \mathcal{I}}, \mathcal{X}, \mathcal{E}, \psi^\theta, f, \delta \rangle$ as known to the econometrician up to θ.

We summarize the baseline assumptions for econometric analysis as follows.

Assumption 3 (Baseline assumptions for econometric analysis)**.**

1. *The set of covariates $\mathcal{X}$ and the set of latent states $\mathcal{E}$ are finite.*
2. *The prior distribution ψ^θ and the payoff functions u_i^θ are known up to a finite-dimensional parameter θ.*
3. *The latent state is a vector of player-specific payoff shocks, i.e., $\varepsilon = (\varepsilon_1, ..., \varepsilon_I)$ and $u_i^\theta(a, x, \varepsilon) = u_i^\theta(a, x, \varepsilon_i)$, and does not affect the transition probability of public states, i.e., $f_{x'|a,x,\varepsilon} = f_{x'|a,x}$.*
4. *The conditional choice probabilities $\phi : \mathcal{X} \to \Delta(\mathcal{A})$ and the transition probability function $f : \mathcal{A} \times \mathcal{X} \to \Delta(\mathcal{A})$ are identified from the data.*

The only "unconventional" assumption is that $\mathcal{E}$ is finite. Although unnecessary for the identification arguments, the discrete support assumption is necessary for feasible estimation and makes the connection to the previous section straightforward. Other econometric papers that use Bayes correlated equilibrium also discretize the state space for feasible estimation (Syrgkanis, Tamer and Ziani, 2021; Magnolfi and Roncoroni, 2023; Gualdani and Sinha,

2024). The player-specific payoff shock assumption, which is a conventional assumption, makes the role of the latent state clear, but it is also not required for our identification arguments.

Identified Set

Given a game (G^θ, S), a solution concept determines a set of feasible predictions, which in turn induces a set of feasible conditional choice probabilities. A standard econometric approach is to assume that the information structure S is known and identify the model parameters by inverting the mapping from the model parameters to the observed conditional choice probability vector ϕ.

Definition 3 (Identified set)**.** *Given Assumption 3, a solution concept SC, and an information structure S, the identified set of parameters is defined as:*

$$\Theta_I^{SC}(S) \equiv \left\{\theta \in \Theta : \ \phi \in \mathcal{P}_{a|x}^{SC}(G^\theta, S)\right\}.$$

In words, a candidate parameter θ enters the identified set if the observed conditional choice probabilities ϕ can be rationalized by some equilibrium (defined by the solution concept) of the model.

Informationally Robust Identified Set

We consider a scenario where a Markov perfect equilibrium generates the data, but the true information structure is unknown to the analyst; misspecifying the information structure leads to biased parameter estimates. We assume that the analyst only knows that the players minimally observe signals in S but may observe more. For instance, as assumed in Magnolfi and Roncoroni (2023), the analyst might know that each player i observes at least ε_i but does not know whether they can receive extra information about opponents' payoff shocks.

Assumption 4 (Identification under Markov perfect equilibrium)**.** *The data are generated by a Markov perfect equilibrium of (G^{θ_0}, S^0), where S^0 is an expansion of S.*

It is well-understood that working directly with Assumption 4 is computationally infeasible; when the analyst knows S but does not know the true information structure S^0, the analyst has to consider the set of all information structures that are expansions of S, but the set is too large. We can make the identification problem computationally feasible by replacing Assumption 4 with the following.

Assumption 5 (Identification under Markov correlated equilibrium)**.** *The data are generated by a Markov correlated equilibrium of $\left(G^{\theta_0}, S\right)$.*

Theorem 3 (Equivalence of identified sets)**.** *The identified set under Assumptions 3 and 4 is equal to the identified set under Assumptions 3 and 5.*

Theorem 3 says that an econometrician who does not know the true information structure underlying the data-generating process can proceed by treating the data *as if* they were generated by a Markov correlated equilibrium. Syrgkanis, Tamer and Ziani (2021), Magnolfi and Roncoroni (2023), Koh (2023), and Gualdani and Sinha (2024) use similar results in static settings. Theorem 3 extends the results to a class of dynamic games. Note that a Markov correlated equilibrium identified set always includes the parameters that can be obtained under the complete information Markov perfect equilibrium assumption with mixed strategies (Ericson and Pakes, 1995; Doraszelski and Satterthwaite, 2010).

Properties of Informationally Robust Identified Set

Identified Set Shrinks with Stronger Assumptions on Players' Information

The Markov correlated equilibrium identified set depends on the baseline information structure set by the analyst, which specifies what players minimally observe. Theorem 2 has shown that a stronger assumption on players' information leads to a tighter set of predictions. The following translates this intuition to identified sets.

Theorem 4. *If* $S^1 \succsim_E S^2$*, then* $\Theta_I^{MCE}(S^1) \subseteq \Theta_I^{MCE}(S^2)$.

Theorem 4 implies a trade-off between informational robustness and the tightness of the identified set. Using weaker assumptions on information is more robust but increases the size of the identified set. The econometrician obtains the tightest identified set when players' baseline information is complete; the econometrician obtains the loosest identified set when players minimally observe nothing about ε. Magnolfi and Roncoroni (2023) and Koh (2023) have shown that fully robust identified sets (i.e., ones with no assumption on the players' information) can be quite large and empirically uninformative.

Markov Correlated Equilibrium is Not Robust to Multiple Equilibria Issue

In the static case, assuming that a single Bayes correlated equilibrium generates the data is innocuous because any mixture of Bayes correlated equilibria is a Bayes correlated equilibrium. However, this property does not carry over to the dynamic case because a mixture of Markov correlated equilibria need not be a Markov correlated equilibrium. This occurs because the Bayes correlated equilibrium obedience condition is linear in the decision rule while the Markov correlated equilibrium obedience condition is nonlinear (non-convex) in the decision rule. Thus, although we remain agnostic about the equilibrium selection rule, we assume that the data are generated by a single equilibrium. Identification with data generated by multiple equilibria is possible (e.g., by using Beresteanu, Molchanov and Molinari (2011)) but far more computationally challenging.

Excluded Variables with Large Support Ensure Point-Identification

While the tightness of the identified set is largely an empirical question, the identified set can shrink substantially when excluded variables have large support. Magnolfi and Roncoroni (2023) prove that players' payoff parameters are point-identified under Bayes correlated equilibrium when there are player-specific excluded variables with large support.[12] Consider, for example, a two-player entry game. Driving a variable that enters only player j's payoff to either positive or negative infinity ensures that the player always stays in or stays out, rendering the other player's decision problem a single-agent binary choice problem. The econometrician can use the single-agent data to identify the payoff parameters, including the competition effect parameters (Tamer, 2003). The same intuition holds in the dynamic case: in markets where one player always stays in or out, the other player solves a single-agent dynamic discrete choice problem while taking the other player's action as fixed. The players' payoff parameters can be estimated using the single-agent dynamic discrete choice data.

In our empirical application, we leverage the intuition from the large support strategy. We can identify many local markets where one of two firms is highly unlikely to enter so that we can exclude the firm as a potential entrant. Then, the other firm is solving a single-agent dynamic discrete choice problem, which is not subject to the identification problem associated with the econometrician not knowing the true information structure underlying the strategic interactions.[13] Thus, we can apply the standard single-agent dynamic discrete choice estimation methods to identify the parameters that govern the single-agent dynamic decision problem.[14] After estimating the "single-agent" parameters, we can use the Markov correlated equilibrium to identify the competitive effects parameters. This estimation strategy that mimics the large support strategy helps us shrink the identified set dramatically.

Numerical Example

Following Pesendorfer and Schmidt-Dengler (2008), we assume that the econometrician knows the values of the scrap value $\kappa = 0.1$ and the discount factor $\delta = 0.9$. However, instead of assuming that the researcher knows the information structure exactly, we assume that the researcher only knows that i minimally observes ε_i and leaves open the possibility that i receives extra signals about ε_{-i}.

Table 2.2 reports the projection intervals of the Markov correlated equilibrium identified sets obtained under each equilibrium. In all cases, the Markov correlated equilibrium identified set contains the true parameter vector, as expected. However, the size of the identified set can vary depending on which equilibrium generated data. We find that the Markov correlated equilibrium identified set obtained from equilibrium (i)—which is the only nested

[12] Parameter governing the correlation of players' payoff types may not be point-identified.

[13] Specifically, we assume that each firm observes its ε_i but allow for the firms to observe arbitrary signals about ε_j. However, if firm i knows j always stays out, then ε_j is payoff-irrelevant to firm i under our econometric assumption.

[14] We cannot apply this idea to identify the competitive effects parameters since we cannot identify markets where one of two firms is always present.

pseudo-likelihood-stable equilibrium and generates highly asymmetric choice probabilities across players—is the tightest.[15] The identified sets under equilibria (ii) and (iii) are indistinguishable because the equilibrium conditional choice probabilities are similar to each other.

Table 2.2: Markov correlated equilibrium identified sets

	True	(i)	(ii)	(iii)
π^m	1.2	$[0.93, 1.83]$	$[0.88, 2.52]$	$[0.88, 2.52]$
π^d	-1.2	$[-1.87, -1.01]$	$[-3.16, -0.65]$	$[-3.16, -0.65]$
c	-0.2	$[-0.51, 0.37]$	$[-0.67, 1.02]$	$[-0.67, 1.02]$

Note: The table reports the projections of the Markov correlated equilibrium identified set assuming that the data were generated from Pesendorfer and Schmidt-Dengler (2008)'s Markov perfect equilibrium (i), (ii), or (iii).

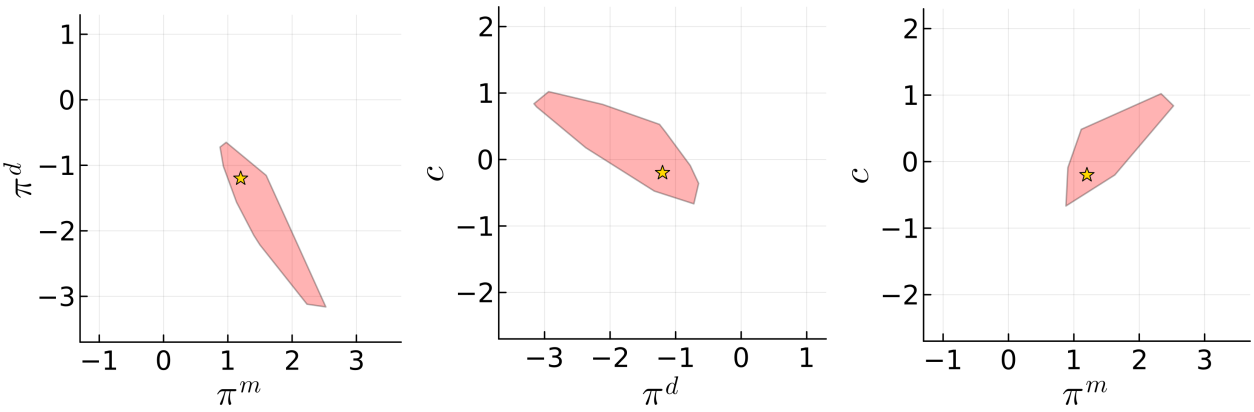


Figure 2.2: Projections of the identified set

Figure 2.2 visualizes the results by plotting the convex hull of the sharp identified set corresponding to equilibrium (ii). In each subfigure, the shaded area represents the convex hull of the projected identified sets, and the star represents the location of the true parameters. As expected, the identified set is non-singleton and contains the true parameter.

Our numerical example shows that a strong assumption on information is far from innocuous: removing arbitrary assumptions on who observes what may result in a significant loss of identifying power. Robustness to misspecification on players' information structure expands the identified set, but the degree to which the identified set expands can vary depending

[15]The nested pseudo-likelihood-stability has been the subject of investigation in the dynamic games estimation literature (see Dearing and Blevins (2024) for a review). Equilibrium (i) is nested pseudo-likelihood-stable while equilibria (ii) and (iii) are not. Our Markov correlated equilibrium estimator is not sensitive to the nested pseudo-likelihood-stability of the underlying equilibrium.

on which equilibrium is selected in the data generating process. Our framework provides a computationally feasible approach to calculating the sharp identified set associated with a weak assumption on players' information.

2.4 Estimation

We propose computationally feasible estimation strategies by formulating a mathematical program with equilibrium constraints à la Su and Judd (2012). We assume the econometrician knows the true conditional choice probability vector ϕ (we relax this assumption in the next section). We also assume that the baseline information structure S is fixed and denote the sharp identified set as $\Theta_I \equiv \Theta_I^{MCE}(S)$.

Mathematical Program

For a candidate parameter θ, we have $\theta \in \Theta_I$ if and only if there exists a decision rule σ that satisfies the Markov correlated equilibrium obedience conditions and induces the observed conditional choice probabilities. We characterize the Markov correlated equilibrium as a set of constraints that the econometrician can plug into optimization software.

Theorem 5 (Sharp Identified Set)**.** *Suppose Assumption 3 holds. Then $\theta \in \Theta_I$ if and only if there exists a pair (σ, V) that satisfies*

$$\sigma_{a|x,\varepsilon,\tau} \geq 0 \text{ for each } a \text{ and } \sum_a \sigma_{a|x,\varepsilon,\tau} = 1, \qquad \forall x, \varepsilon, \tau \tag{2.8}$$

$$\phi_{a|x} = \sum_{\varepsilon,\tau} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \sigma_{a|x,\varepsilon,\tau}, \qquad \forall a, x \tag{2.9}$$

$$\sum_{\varepsilon,\tau_{-i},a_{-i}} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \sigma_{a|x,\varepsilon,\tau} \partial v_i^\theta(a_i', a, x, \varepsilon_i) \leq 0, \qquad \forall i, x, \tau_i, a_i, a_i' \tag{2.10}$$

$$V_{i,x} = \sum_{\varepsilon,\tau,a} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \sigma_{a|x,\varepsilon,\tau} u_i^\theta(a, x, \varepsilon_i) + \delta \sum_{a,x'} \phi_{a|x} V_{i,x'} f_{x'|a,x}, \qquad \forall i, x \tag{2.11}$$

where $v_i^\theta(a, x, \varepsilon_i) \equiv u_i^\theta(a, x, \varepsilon_i) + \delta \sum_{x'} V_{i,x'} f_{x'|a,x}$ is the outcome-specific value function, and $\partial v_i^\theta(a_i', a, x, \varepsilon_i) \equiv v_i^\theta(a_i', a_{-i}, x, \varepsilon_i) - v_i^\theta(a_i, a_{-i}, x, \varepsilon_i)$ represents the deviation payoff from a_i to a_i'.

Constraints (2.8) mean that σ is a proper conditional probability distribution, as required by the definition of a decision rule; (2.9) requires that σ induces the observed conditional choice probabilities ϕ; (2.10) is the Markov correlated equilibrium obedience condition; (2.11) is the ex-ante value function equation described in (2.5) but one that makes the equation

linear in (σ, V) by imposing (2.9).[16] Note that ψ, π, ϕ, δ, and f are treated as known objects in the mathematical program.

Theorem 5, which casts the identification problem as a nonlinear feasibility program, provides a basis for estimating Θ_I. First, one may approximate Θ_I by checking feasibility for a large number of candidate points that approximate the parameter space. This approach is straightforward but may be computationally intensive.[17] Second, one may find a projection interval of Θ_I by minimizing $p^\top \theta$ subject to the constraints above, where $p \in \mathbb{R}^{\dim(\theta)}$ is a direction vector (see, e.g., Kaido, Molinari and Stoye (2019)).[18] The projection approach alleviates the need for a grid search over the entire parameter space but can be numerically difficult.

Fully Robust Identified Set

In the special case where the analyst sets the baseline information structure to players having no minimal signals, S^{null}, the corresponding identified set—"fully robust identified set"—allows for the underlying information structure to be completely arbitrary. It turns out that determining whether a candidate parameter enters the fully robust identified set is computationally easier because the associated program is linear. Since $\Theta_I^{MCE}(S) \subseteq \Theta_I^{MCE}(S^{null})$ for any information structure S, the analyst can take advantage of the computational tractability to restrict the search area.

Theorem 6 (Fully Robust Identified Set)**.** *Suppose 3 holds. $\theta \in \Theta_I^{MCE}(S^{null})$ if and only if there exists (σ, V) such that*

$$\sigma_{a|x,\varepsilon} \geq 0 \text{ for each } a \text{ and } \sum_a \sigma_{a|x,\varepsilon} = 1, \qquad \forall x, \varepsilon$$

$$\phi_{a|x} = \sum_\varepsilon \psi_{\varepsilon|x} \sigma_{a|x,\varepsilon}, \qquad \forall a, x \tag{2.12}$$

$$\sum_{\varepsilon, a_{-i}} \psi_{\varepsilon|x} \sigma_{a|x,\varepsilon} \partial u_i^\theta(a_i', a, x, \varepsilon_i) + \delta \sum_{a_{-i}, x'} \phi_{a|x} V_{i,x'} \partial f_{x'|a_i', a, x} \leq 0, \qquad \forall i, x, a_i, a_i' \tag{2.13}$$

$$V_{i,x} = \sum_{\varepsilon, a} \psi_{\varepsilon|x} \sigma_{a|x,\varepsilon} u_i^\theta(a, x, \varepsilon_i) + \delta \sum_{a, x'} \phi_{a|x} V_{i,x'} f_{x'|a,x}, \qquad \forall i, x \tag{2.14}$$

where $\partial u_i^\theta(a_i', a, x, \varepsilon_i) \equiv u_i^\theta(a_i', a_{-i}, x, \varepsilon_i) - u_i^\theta(a, x, \varepsilon_i)$ represents the deviation flow payoff from a_i to a_i', and $\partial f_{x'|a_i', a, x} \equiv f_{x'|a_i', a_{-i}, x} - f_{x'|a,x}$ is the change of the transition probability from a_i to a_i'.

[16] Note (2.11) also allows us to express each V_i as a closed-form function of σ. This is similar to expressing the ex-ante value function as a closed-form function of conditional choice probabilities in Aguirregabiria and Mira (2007).

[17] In practice, it is useful to transform the feasibility program—which only provides a binary conclusion—into one that measures the degree of constraint violations, e.g., a criterion function approach of Chernozhukov, Hong and Tamer (2007).

[18] For example, $p = (1, 0, ..., 0)$ and $(-1, 0, ..., 0)$ give the endpoints of Θ_I's projection interval in the first component of θ.

The expressions above reflect the assumption that the players receive no signals about the latent state ε. They only rely on their common prior to update their beliefs after observing private recommendations by the information designer. Since the constraints in Theorem 6 are linear in (σ, V), checking whether a candidate θ enters the fully robust identified set amounts to solving a linear program.

2.5 Inference

The inference problem in this paper is characterized by a moment inequality model with a large number of moment inequalities and a high-dimensional nuisance parameter, making it difficult to apply many inference methods for partially identified models.[19] Inference methods used by Syrgkanis, Tamer and Ziani (2021), Magnolfi and Roncoroni (2023), and Gualdani and Sinha (2024) (e.g., Chernozhukov, Hong and Tamer (2007)'s subsampling approach) are computationally challenging to implement in our setting because the Markov correlated equilibrium constraints are non-convex.

We propose a computationally attractive strategy for inference based on the key insights from Horowitz and Lee (2023).[20] Our approach is intended to minimize the need for repeatedly solving non-convex optimization problems. For simplicity, let us assume that the transition probability function is known to the researcher. However, an extension to the case where f needs to be estimated is straightforward. Our approach is also applicable to Bayes correlated equilibrium.

In our setting, the sampling errors are only associated with the reduced-form parameters, namely the conditional choice probabilities (CCPs) ϕ. If ϕ is known, the identified set can be estimated without statistical uncertainty. Thus, we propose controlling for the sampling uncertainty using a confidence set for the CCPs.[21]

Confidence Set

Suppose Φ_α is a confidence set for the CCPs ϕ with the property that $\phi \in \Phi_\alpha$ with probability at least $1-\alpha$ as the sample size n goes to infinity. Leading examples of Φ_α are box constraints or ellipsoids. Let $\Theta_I(\phi)$ be the identified set when the CCPs are ϕ. Define the confidence set as

$$\widehat{\Theta}_I^\alpha \equiv \bigcup_{\phi \in \Phi_\alpha} \Theta_I(\phi).$$

[19] See Canay and Shaikh (2017) and Canay, Illanes and Velez (2023) for recent surveys of the inference methods for partially identified models.

[20] Koh (2023) uses a similar approach in static discrete games settings.

[21] Horowitz and Lee (2023) proposes methods for carrying out (non-asymptotic) inference when the partially identified parameters are characterized by constraints that involve unknown population means that can be estimated from data. Our approach follows their insights in treating the CCPs as unknown parameters that determine the partially identified set. A similar insight has been employed by Kline and Tamer (2016), who propose a Bayesian approach.

If Φ_α covers the true ϕ with high probability, then $\widehat{\Theta}_I^\alpha$ covers the true Θ_I with high probability. Furthermore, checking whether $\theta \in \widehat{\Theta}_I^\alpha$ amounts to solving a nonlinear program in Theorem 5, but with ϕ as part of the optimization variables subject to the constraint $\phi \in \Phi_\alpha$.

Theorem 7. *Let Φ_α be a confidence region for ϕ such that*

$$\liminf_{n\to\infty} \Pr\left(\phi \in \Phi_\alpha\right) \geq 1-\alpha.$$

Then

$$\liminf_{n\to\infty} \Pr\left(\Theta_I \subseteq \widehat{\Theta}_I^\alpha\right) \geq 1-\alpha.$$

Moreover, $\theta \in \widehat{\Theta}_I^\alpha$ if and only if there exists (σ, V, ϕ) subject to $\phi \in \Phi_\alpha$, (2.8), (2.10), and (2.11).

Implementation

We briefly discuss possible approaches for constructing Φ_α. While there are many possible approaches available, it is helpful to construct Φ_α as linear constraints because they are computationally easy to handle. For example, one can construct simultaneous confidence intervals of the form

$$\Phi \equiv \left\{\phi : \hat{L}_{a|x} \leq \phi_{a|x} \leq \hat{U}_{a|x}, \quad \forall a, x\right\}$$

where the vectors $\hat{L}, \hat{U} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{X}|}$ are determined by data. Examples include simultaneous confidence intervals for multinomial probabilities proposed by Fitzpatrick and Scott (1987) or sup-t band by Montiel Olea and Plagborg-Møller (2019); the former is easier to compute since the endpoints of the intervals can be computed in closed-form, but the latter can be tighter.

An alternative is to take a parametric approach.[22] Suppose $\phi = \phi(\gamma)$, where γ is the parameter vector that governs the CCPs. One may estimate γ and construct simultaneous confidence intervals of the form

$$\Phi \equiv \left\{\gamma : \hat{L}_k \leq \gamma_k \leq \hat{U}_k, \quad \forall k\right\}$$

where $\hat{L}, \hat{U} \in \mathbb{R}^{\dim(\gamma)}$ can be computed using the asymptotic distribution of $\hat{\gamma}$. One may also consider replacing ϕ with its linear approximation $\phi(\gamma) \approx \phi(\hat{\gamma}) + \nabla\phi(\hat{\gamma})^\top(\gamma - \hat{\gamma})$ to further simplify the computation. In the identified set search, we also add the constraints defining ϕ as proper conditional distributions, i.e., $\phi_{a|x} \geq 0$, $\forall a, x$, and $\sum_a \phi_{a|x} = 1$, $\forall x$.

[22] When the number of observations is small relative to the dimension of observable covariates, estimating CCPs nonparametrically can be challenging. For example, there might be a covariate bin with no data points. Employing a parametric approach can alleviate this challenge. In particular, a flexibly-specified multinomial logit model can be useful. While such an estimator is generally inconsistent because of model misspecification, it has a lower variance than nonparametric frequency estimates and, thus, may perform better in small samples. Aguirregabiria and Mira (2007) test a similar approach for implementing the two-step pseudo maximum likelihood estimation and report that it works quite well.

2.6 Empirical Application

We study a dynamic entry game by Starbucks and Dunkin' in the US and illustrate the usefulness of our framework.

Background

The US coffee chain industry has grown significantly in recent decades due to the growing demand for specialty coffee and the rise of a culture emphasizing quality, convenience, and socialization. The coffee chain industry is highly competitive, with companies competing for market share through pricing, product differentiation, marketing strategies, and in-store service. In particular, Starbucks and Dunkin' have grown to be the two largest players in the industry, significantly outpacing other competitors by a wide margin. In 2019, Starbucks and Dunkin' accounted for 65% and 28% (resp. 55% and 35%) of total sales (resp. outlets) from the top 8 coffee chains in the US (Technomic, 2019).[23] Figure 2.3 shows that the number of Starbucks and Dunkin' stores in the US increased dramatically from 1997 to 2023. Figure 2.4 illustrates how the firms dramatically expanded their geographic footprint over time.

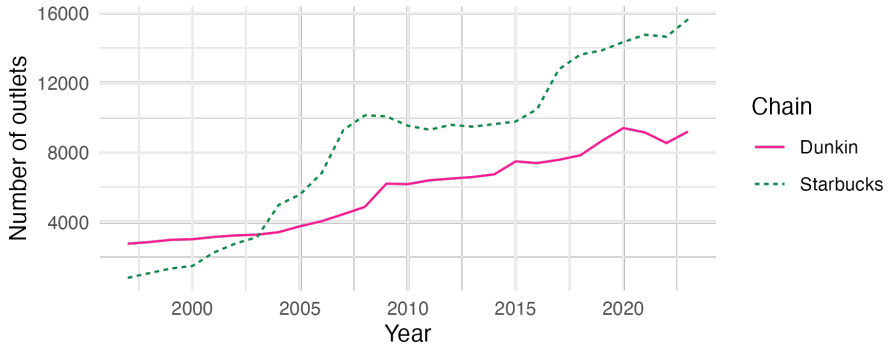
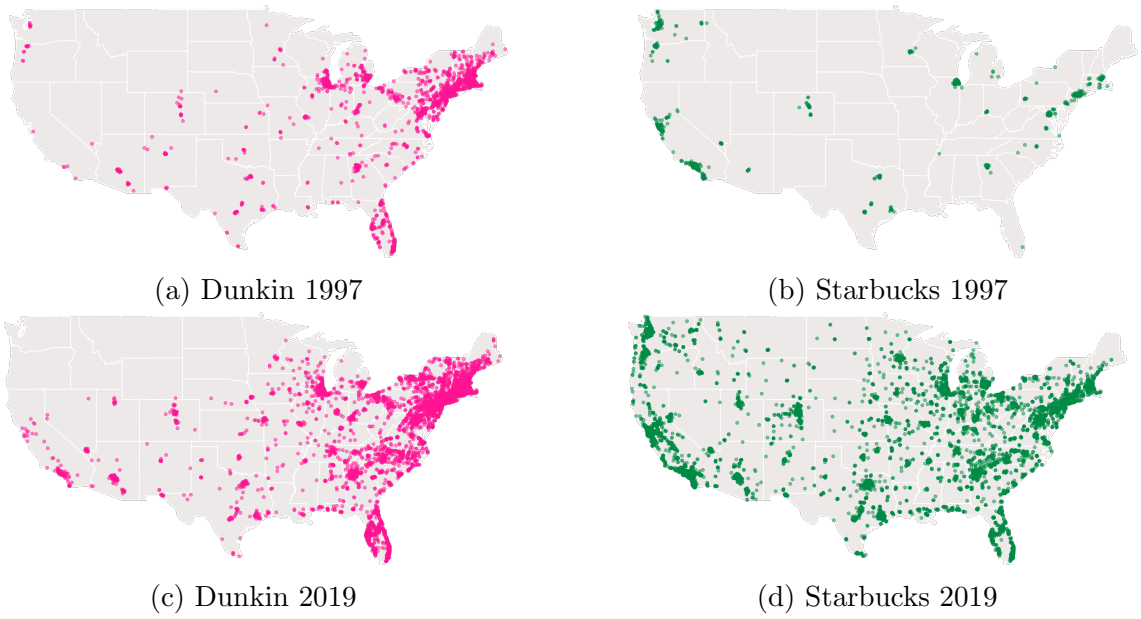


Figure 2.3: Number of Starbucks and Dunkin' stores in the US over time

Little is publicly known about how coffee chains predict opponents' strategic entry decisions. It is reasonable to expect that firms have good information about their stores' (potential) profitability. Still, less is known about how firms learn about their opponents' profitability. Resorting to the standard "incomplete" information assumption—firms observe their payoff shocks but observe nothing about their opponents' shocks—will likely result in a model misspecification. Whether this misspecification is an empirical concern depends on how much it may change the structural estimates and counterfactual outcomes. We apply our Markov correlated equilibrium framework to estimate the parameters and conduct a counterfactual experiment with weaker assumptions on players' information.

[23]Starbucks and Dunkin' outperform other chains by a significant margin. The third and fourth chains were Tim Hortons and Dutch Bros., which accounted for 2.2% and 1.7% (resp. 2.6% and 1.3%) of the sales (resp. outlets) out of the top 8 chains.



(a) Dunkin 1997 (b) Starbucks 1997

(c) Dunkin 2019 (d) Starbucks 2019

Figure 2.4: Growth of Starbucks and Dunkin'

Data

Our primary dataset is Data Axle's U.S. Historic Business Database, which has annual location information of all business establishments in the U.S. from 1997 to 2023. We define markets as four-digit 2010 census tracts.[24] Our geographic market definition assumes that coffee chains compete in small geographic areas.[25] We supplement the firm-location data with the demographic/geographic data from the Longitudinal Track Data Base (Logan, Xu and Stults, 2014) and the neighborhood characteristics data from the National Neighborhood Data Archive. We limit our sample period to 2003–2017 to match the sample period of the neighborhood characteristics data. We define a firm as a potential entrant in each market if it ever operated an outlet in the market or had at least one outlet in the corresponding county and adjacent markets during the sample period.

Table 2.3 reports the summary statistics. Our final sample comprises a balanced panel of $M = 35,686$ tracts observed over $T = 15$ years. Starbucks and Dunkin' are defined as potential entrants in 83% and 71% of the markets. The firms are active in 15% and 12% of the observations, respectively. Entries and exits are rare (around 1% for entry and 0.2% for exit), indicating a high correlation in the incumbency status. The number of eating places is the key common market characteristic that controls for the overall market attractiveness (customer traffic increases with the number of food establishments in the neighborhood). Finally, we

[24] Four-digit census tracts are census-defined geographic units that are contiguous and strike a balance between county (which is too large) and six-digit census tracts (which is too small).

[25] Guler (2018) shows that demand for Starbucks is highly localized. Thomadsen (2005) shows evidence that only fast-food chains within approximately 0.5 miles—which amounts to 10 minutes walking distance—compete as close substitutes in California; it is likely that coffee consumers travel even less.

use own-chain outlets in the county and adjacent markets as firm-specific excluded variables that capture the network effects associated with economies of density (Guler, 2018).

Table 2.3: Summary statistics of 35,686 coffee chain markets from 2003 to 2017

Variable	Mean	Median	Std. Dev.	Min	Max
Potential Entrant					
SB is player	0.831	1	0.375	0	1
DK is player	0.708	1	0.455	0	1
Entry/Exit Decisions					
SB is active	0.151	0	0.358	0	1
DK is active	0.122	0	0.328	0	1
SB entered	0.012	0	0.107	0	1
DK entered	0.007	0	0.086	0	1
SB exited	0.002	0	0.043	0	1
DK exited	0.002	0	0.043	0	1
Market Characteristics					
Number of eating places	16.508	9	26.667	0	910
SB outlets in county	1021.233	252	2127.375	0	10923
DK outlets in county	668.877	107	1346.954	0	7050
SB outlets in adjacent markets	1.733	1	3.413	0	108
DK outlets in adjacent markets	1.109	0	2.142	0	41

Note: SB and DK stand for Starbucks and Dunkin'.

Model

Our model follows the standard dynamic oligopoly entry/exit model with heterogeneous firms (Aguirregabiria and Mira, 2007; Pakes, Ostrovsky and Berry, 2007; Pesendorfer and Schmidt-Dengler, 2008). In each market m and time period t, firm i decides whether to operate in the market ($a_{imt} = 1$) or not ($a_{imt} = 0$). An active firm's flow profit is

$$u_i^\theta(a_{imt} = 1, a_{jmt}, x_{mt}, \varepsilon_{imt}) = \theta_{i,w}^\top w_{mt} + \theta_{i,ec}(1 - z_{imt}) + \theta_{i,ce} a_{jmt} + \varepsilon_{imt}.$$

Here, public state x_{mt} includes two components, observable market characteristics w_{mt} and incumbency status $z_{imt} \equiv a_{im,t-1}$. w_{mt} includes a constant, (log) number of eating places, (log) own-chain outlets in the county, and (log) own-chain outlets in adjacent markets. Parameters $\theta_{i,w}$, $\theta_{i,ec}$, and $\theta_{i,ce}$ represent coefficients on market characteristics, entry cost, and competitive effects, respectively. Finally, ε_{imt} is the idiosyncratic per-period payoff shock assumed to be independently drawn from the standard logistic distribution. We set the flow payoff from being inactive to zero.[26] We assume the agents take the market characteristics as

[26]Normalizing the flow profit from the outside option to zero is not without loss of generality for counterfactual analysis (Aguirregabiria and Suzuki, 2014; Kalouptsidi, Scott and Souza-Rodrigues, 2017, 2021;

exogenous and fixed over time. The state transition is deterministic since the only dynamic state variable is the incumbency status, determined by the firms' actions in the previous period.

Estimation Strategy

Baseline Information Structure. We estimate the Markov correlated equilibrium identified set assuming that the players observe at least their idiosyncratic shocks ε_{imt}. Our identified set is robust to the assumption that the players may observe more information about opponents' payoff shocks before making decisions.

Single-Agent Dynamic Discrete Choice Estimation. We leverage the intuition from the literature that having excluded variables with large support can substantially tighten the identified set (Tamer, 2003; Magnolfi and Roncoroni, 2023). Recall that our definition of potential entrant requires that a firm has had at least one outlet in the county during the sample period. In markets where a firm did not operate even a single outlet in the relevant county, the firm is highly unlikely to enter.[27] In markets where Dunkin' (resp. Starbucks) is not a potential entrant, Starbucks (resp. Dunkin') solves a single-agent dynamic entry problem. The econometrician can then apply single-agent discrete choice estimation methods to identify the payoff parameters except the competition effects parameters. We apply the two-step pseudo-likelihood method.

Markov Correlated Equilibrium Estimation. We then use data from markets in which both firms are players to estimate the competitive effects parameters via our Markov correlated equilibrium framework outlined in Section 2.5.[28] We estimate the first-stage CCPs using flexible logit.[29] We construct simultaneous confidence intervals for CCPs with a plug-in sup-t implementation (Montiel Olea and Plagborg-Møller, 2019).[30] We estimate a criterion function $Q(\theta_{ce})$ that measures the degree of violation in the Markov correlated equilibrium obedience condition (2.10) for given competitive effects parameters θ_{ce}. We construct identified set estimates by collecting θ_{ce} that satisfies the estimated $Q(\theta_{ce}) \leq c$ for a small threshold $c > 0$. We use non-zero c to accommodate misspecifications in the structural and first-stage CCP logit models. The identified set estimates are more conservative as we use a larger c. We do not incorporate sampling error associated with the first-stage single-agent dynamic

Aguirregabiria, Collard-Wexler and Ryan, 2021). We show that our counterfactual experiment of reducing a proportion of entry cost is equivalent to providing an entry subsidy equal to the same proportion of the entry cost net of scrap value without this normalization in Online Appendix B.2.

[27] Formally, let $\tilde{x}_{jmt}$ represent the log of the number of firm j's outlets in the county. Suppose that $\tilde{x}_{jmt}$ is excluded from i's payoff functions and that j's payoff is increasing in $\tilde{x}_{jmt}$. Then markets in which firm j operates no outlet in the county has $\tilde{x}_{jmt} = \log 0 = -\infty$, naturally satisfying the large support assumption.

[28] Aradillas-Lopez and Rosen (2022) also takes a similar two-step approach in static ordered-response games setting. The authors point-identify the non-competitive effects parameters and use the estimates to set-identify the competitive effects parameters.

[29] Nonparametric estimation is infeasible since observations do not exist in some covariate bins.

[30] We use the covariance matrix for CCPs estimated by bootstrapping at the market level in the plug-in sup-t implementation.

discrete choice estimation, so a larger c may prevent us from making identified set estimates too tight.

We also estimate the parameters under the standard incomplete information Markov perfect equilibrium assumption for comparison. We set the discount factor to $\delta = 0.9$. We discretize each continuous variable into two bins by grouping the observations above and below the median and replacing the values with within-bin means.

Estimation Results

Table 2.4 reports our estimation results. Column MPE reports the estimates from the standard incomplete information Markov perfect equilibrium assumption. The competitive effects parameter estimates suggest the presence of Starbucks has a positive spillover effect on Dunkin', while the presence of Dunkin' does not affect Starbucks' profitability. The market characteristics coefficient estimates show positive effects of attractive markets on profitability and economies of density. The entry cost estimates are quite large compared to the other parameter estimates, which rationalizes the low entry probability.

Table 2.4: Structural parameter estimates

	MPE		MCE	
	Estimate	95% CI	Estimate	95% CI
Starbucks				
Intercept	-0.010	$[-0.032, 0.018]$	-0.014	$[-0.045, 0.027]$
Log eating places	0.128	$[0.122, 0.135]$	0.130	$[0.123, 0.138]$
Log outlets in county	0.004	$[0.002, 0.007]$	0.008	$[0.005, 0.013]$
Log adjacent outlets	0.060	$[0.054, 0.067]$	0.053	$[0.041, 0.063]$
Entry cost	-8.064	$[-8.155, -7.976]$	-8.115	$[-8.263, -7.984]$
Competitive effect	0.011	$[-0.001, 0.024]$	-	$[-0.056, 0.096]$
Dunkin'				
Intercept	0.011	$[-0.010, 0.035]$	0.149	$[0.105, 0.183]$
Log eating places	0.071	$[0.065, 0.077]$	0.076	$[0.066, 0.086]$
Log outlets in county	0.018	$[0.016, 0.020]$	0.016	$[0.012, 0.023]$
Log adjacent outlets	0.083	$[0.075, 0.092]$	0.001	$[-0.015, 0.016]$
Entry cost	-8.106	$[-8.194, -8.012]$	-8.543	$[-8.692, -8.395]$
Competitive effect	0.084	$[0.071, 0.098]$	-	$[-0.098, 0.132]$

Note: Parameters excluding the competitive effects parameters in column MCE are obtained using single-agent discrete choice estimation using data where the firms operate as single players. The Markov correlated equilibrium identified set for the competitive effects parameters does not account for sampling error associated with first-stage single-agent dynamic discrete choice estimations. Confidence intervals (CIs) are bootstrapped with resampling at the market level.

Column MCE reports the estimates from our Markov correlated equilibrium framework. The Markov correlated equilibrium identified set estimates of the competitive effects subsume their confidence intervals under the standard incomplete information assumption in column MPE and do not allow us to determine the signs for both firms. The parameter estimates from the single-agent dynamic discrete choice estimation are similar to those reported in column MPE, but the confidence intervals do not overlap for some parameters. The discrepancy of these parameter estimates can be attributed to either misspecification in the information structure of the standard incomplete information assumption or the differences in the parameters underlying separate segments of the markets.[31]

Counterfactual

As a counterfactual experiment, we reduce each firm's entry cost by 40% and study how the number of active firms evolves beginning from 2003. We consider two competitive effects parameter values: lower and upper bounds of their estimated identified set in column MCE of Table 2.4. The lower bounds represent the most extreme case of strategic substitutes because both firms lose their profits most due to the competitor's presence. The upper bounds are the other extreme of strategic complements, as both firms benefit most from the competitor's presence. By considering the two extremes, we can investigate the bounds of counterfactual outcomes that can arise. We fix the parameters other than the competitive effects at their two-step pseudo maximum likelihood estimates in column MPE of Table 2.4.

At each candidate parameter, we find Markov correlated equilibria that maximize and minimize the expected number of active firms. The expected number of active firms at state x is $0 \times \phi(0,0|x) + 1 \times (\phi(0,1|x) + \phi(1,0|x)) + 2 \times \phi(1,1|x)$, where $\phi(a_1, a_2|x)$ is the equilibrium CCPs. We take an unweighted average over x to obtain the (unconditional) expected number of active firms. We use the equilibrium CCPs to simulate the evolution of active firms. We allow the information structure in the counterfactual equilibrium to vary from the equilibrium in the estimation.[32]

Figure 2.5 plots the counterfactual results. The black curve shows the evolution of active firms in the markets using the CCPs estimated from the data. The blue curves represent the bounds on counterfactual outcomes at the most strategic substitute case. The red curves represent the bounds on counterfactual outcomes at the most strategic complement case. As expected, the strategic complement case has a higher number of active firms than the strategic substitute case because the presence of opponents boosts firms' profits when entries are strategic complements.

Figure 2.5 reveals that the counterfactual outcomes are not sensitive to assumptions on information structure in the counterfactual equilibrium. The differences between the maximum and minimum expected number of active firms are tight for a given set of parameter

[31] We do not provide a formal test of misspecification in the information structure. See Han, Kaido and Magnolfi (2024) for a related test in the Bayes correlated equilibrium context.

[32] We do not apply the methodology of Bergemann, Brooks and Morris (2022), which shows how to conduct counterfactual analysis with information structure fixed at the pre-counterfactual equilibrium.

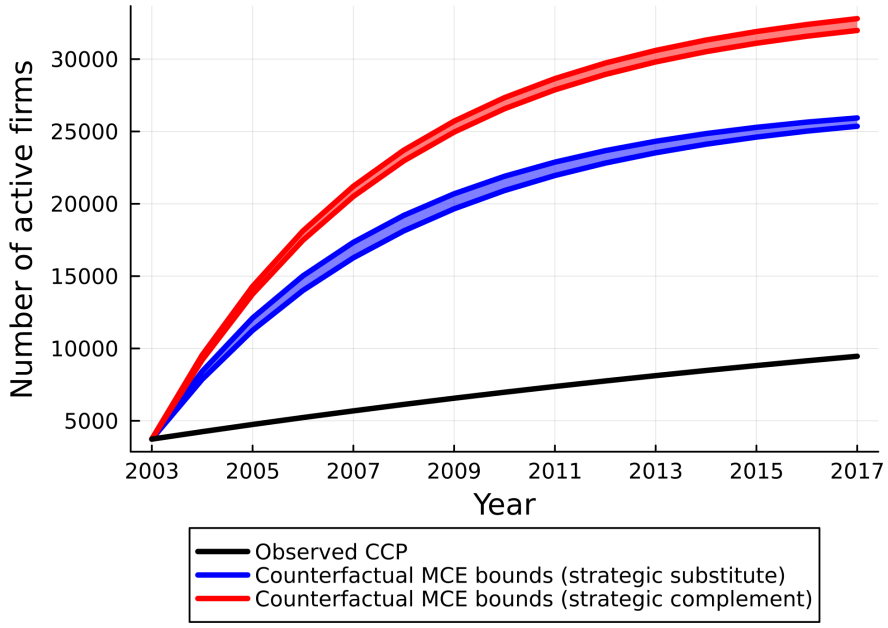


Figure 2.5: Firm entries in counterfactual equilibria

values. The region between the maximum and minimum includes the expected numbers of active firms arising from any Markov perfect equilibrium when the players minimally observe their payoff shocks and may observe more information about opponents' payoff shocks. Thus, the tight counterfactual outcome bounds indicate that the ability to obtain more information about opponents' payoff shocks has little impact on the outcome. The small competitive effects parameters (in absolute values) can explain the tight counterfactual bounds. If the competitive effects are null, then the firms do not care about the opponents' behavior, and there is no room for an information designer to affect the outcomes. Here, the competitive effects parameters range between -0.10 and 0.14, which is small considering the standard deviation of the idiosyncratic payoff shock (the standard deviation of the standard logistic random variable is $\sqrt{\pi^2/3} \approx 1.81$). Thus, the degree to which the outcomes can vary in counterfactual Markov correlated equilibrium depends largely on the competitive effects parameter values.

2.7 Conclusion

In this paper, we have studied the identification, estimation, and counterfactual analyses of a common class of dynamic games for empirical analysis, assuming the researcher knows some minimal information available to players. To facilitate informationally robust econometric analysis of dynamic games, we have defined *Markov correlated equilibrium*, a dynamic analog of Bayes correlated equilibrium, and studied its properties. We have discussed new challenges that arise in dynamic environments and proposed practical solutions. We have applied our

proposed method to study the dynamic entry and exit decisions of the two largest coffee chains in the US.

We conclude this paper with a discussion of future research directions. First, we think further refinement of Markov correlated equilibrium (and thus Bayes correlated equilibrium) will attract larger interest from empiricists. The type of informational robustness offered by Markov/Bayes correlated equilibrium admits too many possibilities. Markov correlated equilibrium identified sets and counterfactual outcomes can span a wide range, which may not be empirically informative; Magnolfi and Roncoroni (2023) find similar results in their empirical application. Numerical examples from Koh (2023) and Magnolfi and Roncoroni (2023) show that even complete information correlated equilibrium identified set can be large, indicating that correlation in actions can also weaken the identifying power substantially. In some settings, the researcher may be able to reasonably exclude some information structures from consideration. It will be fruitful to find a tractable approach that can accommodate further assumptions on the underlying information structure to narrow down the possibilities.

Second, estimation with Markov correlated equilibrium remains computationally challenging. A similar problem exists for Bayes correlated equilibrium since the analyst needs to run a grid search. Studying how to conduct computationally tractable estimation and inference in this class of partially identified models will be important to broaden the scope of applications.

Chapter 3

The Role of Racial Concordance in Doctor Referral Decisions

3.1 Introduction

Customers often request advice from an expert to reach out to a service provider. Examples include doctor referrals, referral-based hiring (Miller and Schmutte, 2021), and real estate agent recommendations (Christensen and Timmins, 2022). Experts may exhibit ingroup bias toward providers or stereotype customers, and they may make referral decisions based on their characteristics on top of the customer's benefit. In this paper, we study the role of racial concordance among doctors, specialists, and patients in doctors' referral decisions.[1]

Racially concordant referrals may have conflicting implications for patient care. On one hand, racially concordant referrals may steer patients toward low-quality care. For instance, black patients tend to visit black doctors, and black specialists tend to work in low-quality hospitals (Bach et al., 2004; Chandra et al., 2024). On the other hand, racially concordant referrals may also improve patients' outcomes because racial concordance between patients and physicians can enhance the quality of care and improve health outcomes (Alsan et al., 2019; Greenwood et al., 2020). Thus, seemingly racially concordant referrals may reflect doctors' intentions to improve patient outcomes (statistical discrimination). In contrast, we consider racially concordant referrals that persist after controlling for the expected patient outcome and travel distance as evidence of ingroup bias. Such effect may arise from taste-based discrimination (Becker, 1957), stereotyping (Bordalo et al., 2016), or biased beliefs (Bohren et al., 2019). Prior work documents the presence of ingroup bias in healthcare markets (Hoffman et al., 2016; Singh and Venkataramani, 2022). Understanding the effects of racially concordant referrals on health outcomes can help us evaluate the potential inefficiencies and equity concerns in the doctor referral system.

[1] We include both race and ethnicity when we use the term "race" in this paper. We call a referral-sending physician a doctor and a referral-receiving physician a specialist regardless of their specialties, following Zeltzer (2020).

To that end, we propose a structural framework to identify the racial concordance effect, net of benefits from outcome and travel distance on doctors' referral decisions. We build a doctor referral decision model and a patient outcome model. The referral decision model considers racial concordance, expected outcome, and travel distance. The outcome model incorporates the racial concordance effect on the outcome and the specialist-specific outcome effect. We use a 20% random sample of Medicare beneficiaries from 1999–2019 in the United States to estimate the model, where we predict physicians' race from their names and locations (Imai and Khanna, 2016). We first estimate the outcome model and then plug in the outcome model estimates to gauge the expected outcome that feeds into the referral decision model. Since revision is a relatively rare event, our estimates for the expected outcome may suffer from measurement error. We use specialist exits from the local market as an instrument to deal with measurement error of the specialist-specific outcome effect.

We apply our model to patients who underwent Total Hip Arthroplasty (THA) or Total Knee Arthroplasty (TKA). We focus on these surgeries performed in orthopedics for several reasons. First, doctors would likely face a limited specialist choice set when referring patients with acute conditions, such as acute myocardial infarction and stroke. Thus, THA/TKA—primarily due to chronic conditions, osteoarthritis and rheumatoid arthritis—are better suited to analyze doctors' referral decisions. Second, THA/TKA patients have a risk of revision surgery in the short term, which has severe consequences on outcomes and costs (Nichols and Vose, 2016; Seidelman et al., 2023). Third, revision surgeries are unlikely to be affected by the racial concordance.[2] The irrelevance of racial concordance on outcomes makes it easier to clearly show racially concordant referrals even if there are no outcome improvements because of the racial concordance.

We find that doctors make racially concordant referrals even after accounting for outcome benefits and travel distance for THA/TKA cases. Within a local market, Asian doctors refer non-Asian patients to Asian specialists 1.8 percentage points (pp) more often than non-Asian doctors. Similarly, Asian patients are referred by non-Asian doctors to Asian specialists 3.4 pp more than non-Asian patients. We then quantify the trade-off between the racial concordance and expected health outcome. We estimate that doctors are willing to trade off 0.03–0.12 percentage points of revision risk to refer patients to a specialist of their own race. Additionally, they are willing to accept 0.07–0.54 pp of additional revision risk to match the patient with a specialist of the same race. Furthermore, we find that there is a trade-off between health outcomes and patient's travel distance to the specialist. Our estimates suggest that an improvement of 1 pp in the revision rate is equivalently valued as 79.6–82.4 kilometers of travel to the specialist.

In our counterfactual analysis, we compare outcomes under two scenarios: (i) a baseline scenario in which racial concordance influences referral decisions, and (ii) a counterfactual scenario in which the effect of racial concordance is removed. Through this analysis, we aim to understand how patient outcomes are affected by the existence of racially concordant

[2]We confirm that revision surgeries are not affected by the racial concordance in the paper.

referrals. We find that the probability of referring a patient to an Asian, Black, or Hispanic specialist significantly declines in the absence of racial concordance effects, with the decline most pronounced when the doctor and patient share the same race. For instance, when both the doctor and patient are Asian, the probability of referral to an Asian specialist drops from 36.0% in the baseline to 14.6% in the counterfactual scenario. Additionally, for doctor–patient pairs of the same race (Asian, Black, or Hispanic), the patient's expected revision rate improves by 0.03 pp on average—a 4% relative improvement—under the counterfactual. This improvement comes at the cost of a modest increase in travel distance, ranging from 0.4 to 0.9 kilometers.

Our paper builds on multiple strands of literature. First, this paper is related to the literature on heterogeneity in physician practice patterns (Abaluck et al., 2016; Chandra and Staiger, 2020; Chan et al., 2022). We specifically examine heterogeneity arising from racial differences among agents in the healthcare market. Prior research has highlighted the impact of racial concordance in healthcare (Bach et al., 2004; Alsan et al., 2019; Greenwood et al., 2020). We study the effect of racial concordance in doctor referral decisions. Second, Sarsons (2022) and Zeltzer (2020) have shown sex homophily in doctor referrals. We extend their ideas to racially concordant referrals. Importantly, we incorporate patient outcomes and travel distance into our model to disentangle the racial concordance effect from other factors.

Furthermore, this paper contributes to the literature on disparity in health. Existing work has emphasized the role of access, social determinants, and doctor-patient relationships in generating disparities (Alsan et al., 2019; Chandra et al., 2024). We focus on a novel channel, doctor referrals. We document that, due to the existence of racial concordant referrals, some racial groups experience worse health outcomes than other groups.

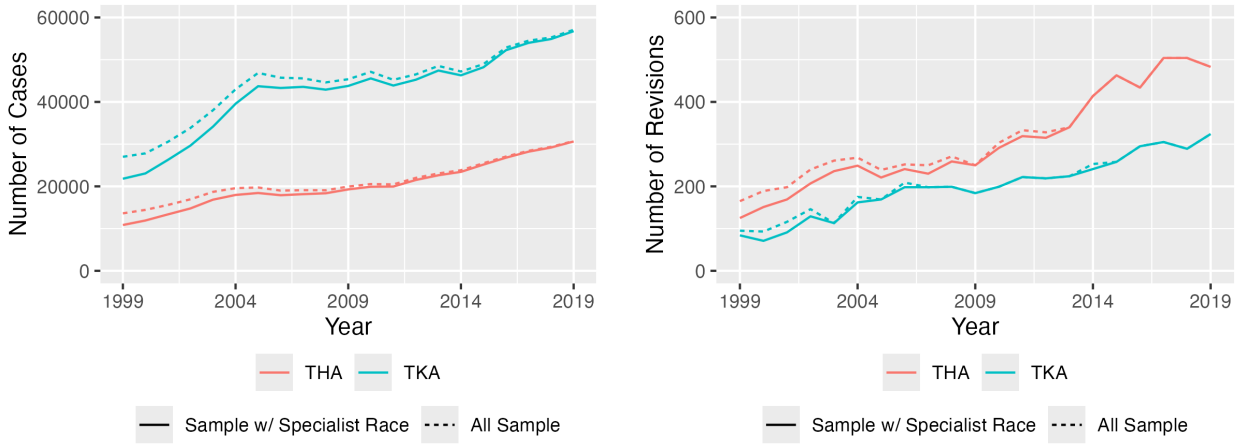
3.2 Data and Background

The primary data source for our study is the U.S. Medicare fee-for-service claims data, maintained by the Centers for Medicare & Medicaid Services (CMS). We use a 20% random sample of Medicare beneficiaries from 1999–2019, drawn from the CMS Research Identifiable Files (RIFs). We combine two RIFs to construct our dataset: the Carrier File and the Master Beneficiary Summary File. We collect the date of service, Healthcare Common Procedure Coding System (HCPCS) codes, identity and specialty of the performing physician (doctor), and identity of the referring physician (specialist) from the Carrier File. Physician identifiers are linked to names using the CMS National Plan and Provider Enumeration System (NPPES).[3] We infer physicians' race based on their names and geographic location using the

[3]In 2008, physician identifiers transitioned from the Unique Physician Identification Number (UPIN) to the National Provider Identifier (NPI). To consistently track physicians across this change, we follow the methodology proposed by Parsons et al. (2017) to link records and identify the same physician before and after the transition.

wru package in R (Imai and Khanna, 2016).[4] From the Master Beneficiary Summary File, we obtain beneficiaries' demographic characteristics—including sex, age, race, and zip code of residence—as well as information on comorbidities. In accordance with the CMS cell size suppression policy, all reported statistics are based on cells with more than 10 beneficiaries, to protect individual privacy.

Our main analysis focuses on patients who underwent Total Hip Arthroplasty (THA) or Total Knee Arthroplasty (TKA). We use the THA/TKA revision within 90 days of the initial surgery to evaluate the outcome of THA/TKA referrals. The leading causes of these early revisions are surgical site infection (SSI) and dislocation for THA (Heckmann et al., 2021) and SSI for TKA (Postler et al., 2018; Urish et al., 2020). These early revisions are associated with substantially worse clinical outcomes and increased healthcare costs (Nichols and Vose, 2016; Seidelman et al., 2023). The prevalence of SSI is estimated at 1–2% (Li et al., 2022), while the prevalence of dislocation within 3 months for THA is 1% (Meek et al., 2006). Importantly, the prevalence of SSI and dislocation varies across specialists (Hedlundh et al., 1996; Li et al., 2022; Seidelman et al., 2023).



(a) Total (b) Revision

Figure 3.1: Number of THA/TKA cases and revisions

Figure 3.1 presents the number of THA/TKA cases and revisions in our dataset. The number of THA/TKA cases have been increasing over time. In 2019, we see 30,000 THA cases and 58,000 TKA cases. Revision within 90 days occurs in 1–2% of THA cases and 0.5% of TKA cases. For each THA/TKA patient and performing specialist, we identify the THA/TKA referring doctor by looking back at most 180 days from the surgery date of the patient's visits to the performing specialist.[5] Figure 3.2 depicts the number of THA/TKA

[4]While this package was originally created in the field of political science, the package has been used in Public Health studies (e.g., Studdert et al., 2020).

[5]If multiple doctors meet this criterion, we retain the one whose observation is closest to the surgery date.

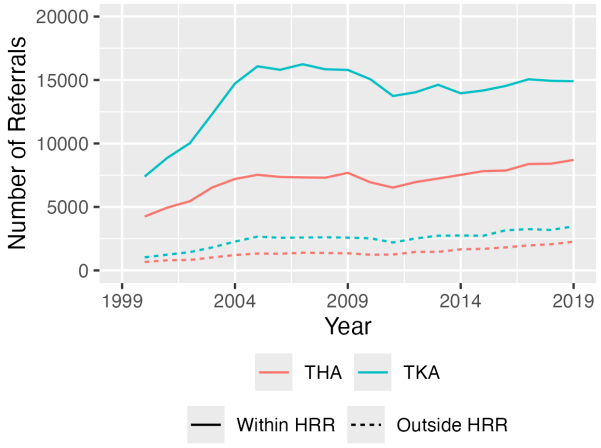
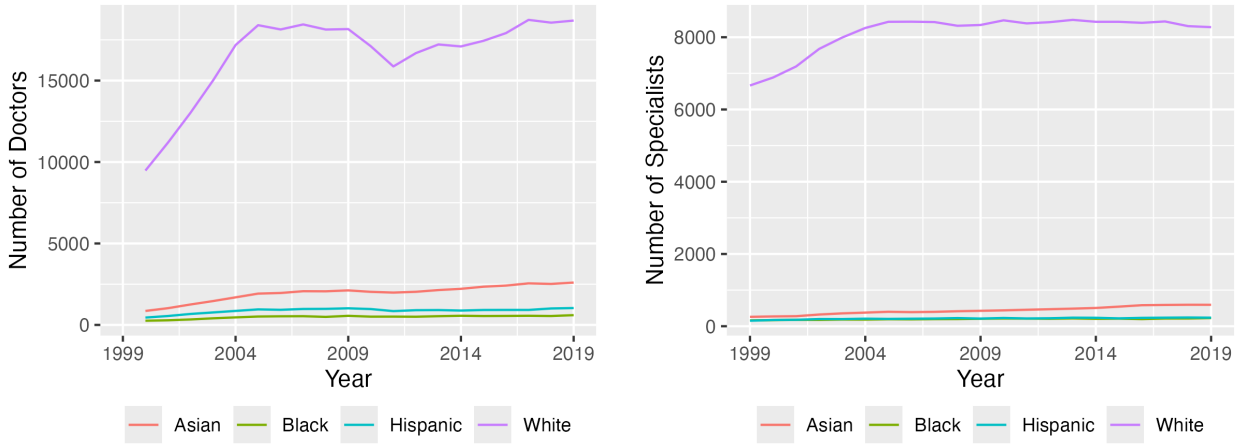


Figure 3.2: Number of THA/TKA referrals



(a) Doctor

(b) Specialist

Figure 3.3: Number of THA/TKA referring doctors and performing specialists by race

referrals over time. Approximately 80% of referrals occur within the same Hospital Referral Region (HRR) as the referring doctor.

Figure 3.3 shows the number of THA/TKA referring doctors and performing specialists by race. Compared to the summary statistics of physicians' self-reported race reported by the Association of American Medical Colleges (AAMC), Black doctors/specialists represent a higher share in our dataset. According to the AAMC's 2018 report, the number of Asian primary care physicians (PCPs) is 15,323; Black is 6,581; Hispanic is 7,647; and White is 66,898. The numbers for orthopedists are 1,255 for Asian, 513 for Black, 523 for Hispanic, and 13,443 for White. The shares of Asian, Black, and Hispanic PCPs and specialists reported by the AAMC exceed those observed in our dataset. This suggests that our methodology slightly undercounts Asian, Black, and Hispanic physicians and specialists overall.

We exclude specialists with fewer than 10 observed THA/TKA cases in our main analysis to estimate the specialist-specific outcome effects. Specialists with 10 or more observed THA/TKA cases, whom we refer to as *experienced specialists*, are expected to have performed more than 50 cases, given that our dataset represents only a 20% random sample of Medicare beneficiaries and does not include non-Medicare patients. The experienced specialists account for 94.1% of all THA/TKA cases in our dataset.

Table 3.1 summarizes the characteristics of THA/TKA patients, doctors, and specialists, both before and after the exclusion of low-volume specialists. Column 1 tabulates the summary statistics for the full sample, and column 2 tabulates them for the restricted sample with experienced specialists. The patient characteristics are largely similar across the full and restricted samples (see patient comorbidities in Table C.1 in Appendix C.1). The population shares by race are Asian 5.9%, Black 12.1%, Hispanic 18.7%, and White 57.8% in the 2020 U.S. Census. Thus, compared to the population shares, White patients represent a higher share of THA/TKA patients, whereas Asian, Black, and Hispanic patients represent lower shares. On average, we observe 100 THA/TKA cases over 14 years for experienced specialists. These specialists are expected to have performed more than 500 cases. Thus, the orthopedists in our sample perform THA/TKA once a week. We observe fewer than one revision per experienced specialist.

We model doctors' referral decisions, restricting each doctor's referral choice set to specialists located within the same HRR, as doctors mostly refer to specialists within the same HRR, which is a pattern we have seen in Figure 3.2.[6] We estimate the structural parameters of the referral decision contrasting two mutually exclusive race groups, e.g., Asian and non-Asian. To estimate the structural parameters contrasting race $R \in \{\text{Asian, Black, Hispanic, White}\}$ or non-R groups, we exclude HRRs that have had at least one year without either race R or non-R specialists. We also exclude HRR-years without race R or non-R referrals for any doctor-patient race group combinations.

Table 3.2 presents summary statistics related to doctors' referral choice sets. On average, each doctor has 39 specialists in their referral choice set, and patients travel an average of 22.5 kilometers to reach their referred specialist. Among patient-specialist pairs located within the same HRR, the average distance is 56.4 kilometers, suggesting that travel distance plays an important role in referral decisions.

Figure 3.4 illustrates the number of referrals for each HRR in year 2019.[7] The comparison between Asian and non-Asian referrals excludes much of the Northwestern and Southern areas. In contrast, the Hispanic versus non-Hispanic comparison is more geographically concentrated near the southern border.

[6]HRRs are designed to contain the referral hospitals most often used by region residents (Wennberg and Cooper, 1996). There are 306 HRRs in total. HRRs include a city where patients can have major cardiovascular surgical procedures and neurosurgery, and more than 120,000 residents populate them. More than 85% of referrals are within the same HRR (Zeltzer, 2020).

[7]Areas with less than 10 referrals are excluded from the analysis.

Table 3.1: Summary statistics of THA/TKA patients, doctors, and specialists (1999–2019)

	Full Sample (1)		Experienced Specialist (2)	
	Mean	S.D.	Mean	S.D.
Patient	(N = 1,315,864)		(N = 1,238,534)	
Age	74.38	6.17	74.35	6.16
Male	0.3598	-	0.3600	-
Asian	0.0075	-	0.0075	-
Black	0.0500	-	0.0493	-
Hispanic	0.0102	-	0.0098	-
White	0.9141	-	0.9152	-
THA/TKA Revision	0.0081	-	0.0080	-
Doctor	(N = 143,290)		(N = 141,427)	
Observation Length (year)	5.19	5.44	5.17	5.42
Number of Referrals	2.91	4.39	2.90	4.39
Specialist	(N = 17,496)		(N = 12,798)	
Observation Length (year)	11.76	6.79	14.58	5.25
Number of THA/TKA Cases	71.8	111.3	96.8	120.7
Number of THA/TKA Revisions	0.58	1.24	0.78	1.39
Proportion of THA/TKA Revisions (%)	1.11	4.96	0.95	1.86

Notes: Table only shows the mean for binary variables. The experienced specialists have 10 or more observed THA/TKA cases.

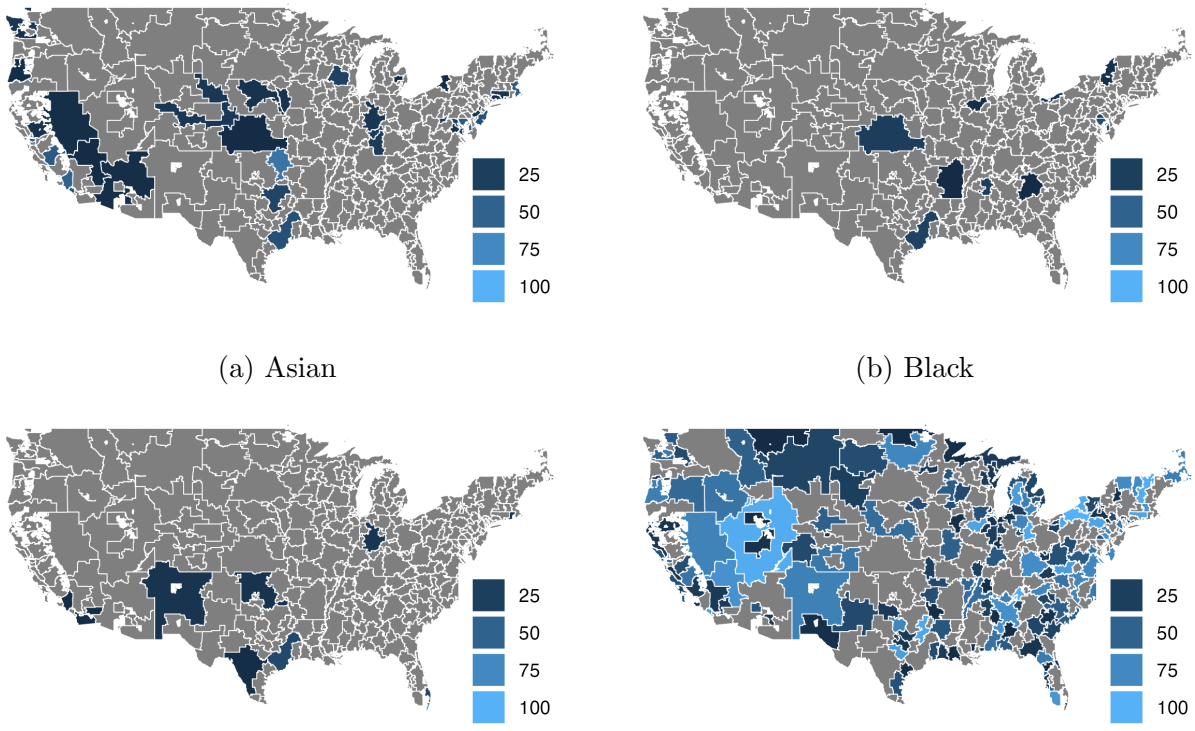
3.3 Preliminary Evidence of Racially Concordant Referrals

We show that racially concordant referrals appear within HRRs. In this section, we broaden our focus to include all types of referrals, rather than limiting our analysis to THA/TKA-related referrals. We first plot race $R \in \{\text{Asian,Black,Hispanic,White}\}$ specialist referral share against their corresponding specialist population share, stratified by doctor-patient race group (race R or not) combinations for each HRR in 2019. In this analysis, we define a referral as a unique doctor–patient–specialist triple, following Zeltzer (2020).[8]

[8]We exclude triples with the same physician as the "doctor" and "specialist." These cases are often lab tests in which the ordering physician is reported as the referring physician.

Table 3.2: Summary statistics related to referral choice set (1999–2019)

	Mean	S.D.
Referral		
Number of Specialists in Same HRR	39.20	28.78
Distance to *Chosen* Specialist (km)	22.51	29.48
N	394,527	
Patient-Specialist Pair (Both Chosen & not Chosen)		
Distance to Specialist in Same HRR (km)	56.39	55.08
N	15,467,299	



(a) Asian (b) Black

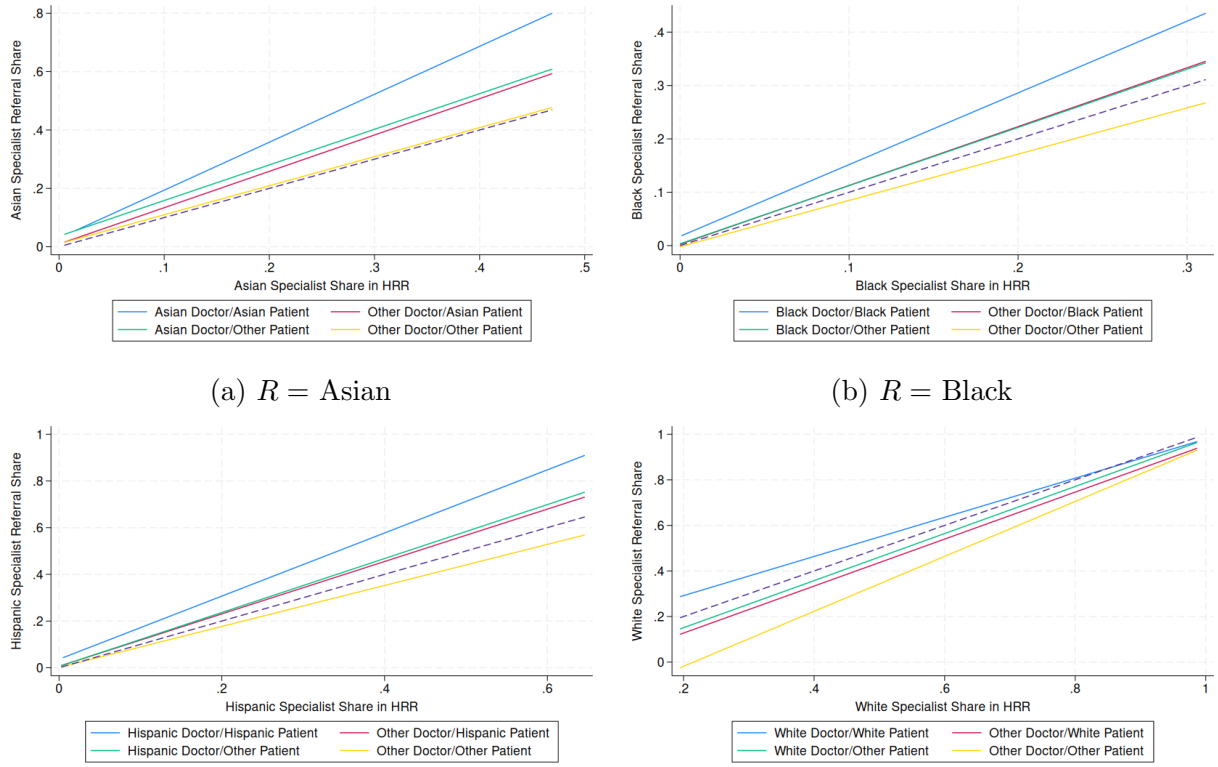
(c) Hispanic (d) White

Figure 3.4: Number of referrals by HRR in 2019

Figure 3.5 depicts the fitted lines for each doctor-patient race group combination. As expected, race R specialist referral shares are higher in HRRs with higher proportion of race R specialists. More interestingly, we see patterns that suggest the existence of racially

concordant referrals within HRRs. For instance, in regions where 20% of specialists are Asian, the share of non-Asian doctors referring non-Asian patients to Asian specialists is 20% (yellow), non-Asian doctors referring Asian patients to Asian specialists is around 25% (red), Asian doctors referring non-Asian patients to Asian specialists is around 25% (green), and Asian doctors referring Asian patients to Asian specialists is around 35% (blue).

Race R specialist referral shares for race R doctor–race R patient pairs (blue line) are higher than other doctor-patient race group combinations within HRRs. Additionally, race R specialist referral shares are the lowest for race non-R doctor–race non-R patient pairs (yellow line). Putting into words, doctors are more likely to refer patients to specialists of the same race, particularly when the patient shares the same race as the referring doctor. Similarly, patients are more frequently referred to specialists of their own race, especially when the referring doctor is also of the same race. We obtain similar patterns for orthopedists and neurosurgeons (Figures C.1 and C.2 in Appendix C.2).[9]



(a) R = Asian

(b) R = Black

(c) R = Hispanic

(d) R = White

Figure 3.5: Race R specialist referral share against their population share by doctor-patient race group in 2019

Next, we quantify the racial concordance effect on referral patterns within HRRs using

[9]Inconsistent results for Hispanic neurosurgeons are likely due to the small sample size.

the THA/TKA dataset from 1999–2019. We calculate race R specialist referral share by doctor-patient race group combination for each HRR year. Denote the referral share for race $r \in (R, \text{non-}R)$ doctor and race $\rho \in (R, \text{non-}R)$ patient pair in HRR l and year t as $\text{RefShare}_{r\rho lt}$. We then run the following regression:

$$\text{RefShare}_{r\rho lt} = \beta_1 \mathbb{1}(r = R) + \beta_2 \mathbb{1}(\rho = R) + \beta_3 \mathbb{1}(r = R) \cdot \mathbb{1}(\rho = R) + X_{lt}\beta_4 + \varepsilon_{r\rho lt},$$

where X_{lt} includes the covariates for HRR and year.

Table 3.3 presents the regression results. Column 1 only controls for year fixed effects. Columns 2 and 3 additionally control for the race R specialist population share and HRR fixed effects, respectively. A positive coefficient on the race R doctor (β_1) indicates doctors of race R are more likely to refer patients to race-R specialists than non-R doctors, when referring non-R patients. Similarly, a positive coefficient on the race R patient (β_2) indicates that race R patients are referred to race R specialists more than non-R patients when the patient obtains a referral from a non-R doctor. Including the race R specialist population share or HRR fixed effects attenuates the magnitude of these coefficients but they remain positively significant. For example, controlling for the Asian specialist population share ($R =$ Asian, column 2), Asian doctors refer to Asian specialists 1.8 pp more often than non-Asian doctors for non-Asian patients. In addition, Asian patients are referred to Asian specialists 3.4 pp more frequently than non-Asian patients when referred by non-Asian doctors.

A positive sum of coefficients on the race R patient and the interaction term (i.e., $\beta_2+\beta_3 > 0$) indicates that doctors are more likely to refer patients to same race specialists more when the patient is also of the same race, relative to other patients. Similarly, a positive sum of coefficients on the race R doctor and the interaction term (i.e., $\beta_1 + \beta_3 > 0$) indicates that patients are more likely to be referred to same race specialists when the doctor is also of the same race, relative to other doctors. For instance, controlling for the Asian specialist population share ($R =$ Asian, column 2), Asian doctors refer Asian patients to Asian specialists 13.9 pp more than they refer non-Asian patients. Also, Asian patients are referred to an Asian specialist 12.3 pp more frequently by Asian doctors than non-Asian doctors.

The race R specialist population share affects the race R specialist referral share in a large magnitude. Specifically, A 10 pp increase in the race R specialist population share is associated with an increase in the race R specialist referral share by 10.1–11.4 pp. Replacing this covariate with HRR fixed effects yields similar estimates for the racial concordance effect, suggesting that the observed referral patterns are not solely driven by regional variation in preference for racial factors.

3.4 Model of Doctors' Referral Decisions

To understand doctors' referral behavior while accounting for specialist quality and travel distance, we develop a model of doctors' referral decisions. Doctors choose a specialist from within their HRR, considering the patient's travel distance, expected health outcome and

Table 3.3: THA/TKA referral share regressions (1999–2019)

	Race R Specialist Referral Share					
	R = Asian			R = Black		
	(1)	(2)	(3)	(1)	(2)	(3)
Race R Doctor	0.0260***	0.0183***	0.0186***	0.0343***	0.0155***	0.0163***
	(0.0036)	(0.0030)	(0.0033)	(0.0037)	(0.0033)	(0.0032)
Race R Patient	0.0708***	0.0344***	0.0303***	0.0245***	0.0127***	0.0135***
	(0.0109)	(0.0068)	(0.0070)	(0.0037)	(0.0033)	(0.0032)
Race R Doctor × Race R Patient	0.1288***	0.1049***	0.1060***	0.0604***	0.0493***	0.0471***
	(0.0200)	(0.0193)	(0.0198)	(0.0130)	(0.0118)	(0.0114)
Race R Specialist Population Share		1.018***			1.007***	
		(0.0691)			(0.0660)	
Year Fixed Effects	X	X	X	X	X	X
HRR Fixed Effects			X			X
Number of HRR-year	12,353	12,353	12,353	13,259	13,259	13,259
R^2	0.098	0.372	0.340	0.048	0.337	0.300
	R = Hispanic			R = White		
	(1)	(2)	(3)	(1)	(2)	(3)
Race R Doctor	0.0206***	0.0093**	0.0109***	0.0963***	0.0652***	0.0660***
	(0.0056)	(0.0035)	(0.0037)	(0.0093)	(0.0074)	(0.0077)
Race R Patient	0.0494***	0.0164**	0.0177**	0.0938***	0.0600***	0.0605***
	(0.0145)	(0.0065)	(0.0068)	(0.0094)	(0.0074)	(0.0077)
Race R Doctor × Race R Patient	0.1128***	0.0623***	0.0615***	-0.0644***	-0.0406***	-0.0408***
	(0.0251)	(0.0175)	(0.0168)	(0.0074)	(0.0063)	(0.0061)
Race R Specialist Population Share		1.140***			1.012***	
		(0.0595)			(0.0301)	
Year Fixed Effects	X	X	X	X	X	X
HRR Fixed Effects			X			X
Number of HRR-year	11,556	11,556	11,556	18,697	18,697	18,697
R^2	0.072	0.591	0.576	0.037	0.443	0.391

Notes: Standard errors are clustered at the HRR and year level. * p-value < .1, ** p-value < .05, *** p-value < .01.

racial concordance preferences that are unrelated to the outcome. The outcome depends on the racial concordance between the patient and specialist, the specialist-specific outcome effect, and exogenous covariates, such as patient comorbidities. We contrast two mutually exclusive race groups, 1 and 0, e.g., Asian and non-Asian, in the model.[10]

Doctor i with $\text{Race}_i \in \{0, 1\}$ makes a referral decision for patient j with $\text{Race}_j \in \{0, 1\}$ in HRR l at year-quarter t. We assume that doctors' choice set consists of specialists who have performed THA/TKA in HRR l at least once. Furthermore, we assume that there is no outside option because the patient needs the referral based on her health conditions that

[10]Extending the model to account for three or more race groups is straightforward. However, empirical applications will become challenging due to limited sample sizes in less prevalent race group combinations.

require specialist treatments.

Doctor i's utility of referring patient j to specialist k is

$$u_{ijlt}(k) = g(\text{Race}_i, \text{Race}_j, \text{Race}_k) + \beta \mathbb{E}[\text{Outcome}_{jklt} | \mathcal{I}_{ijklt}] + \gamma \text{Distance}_{jk} + \varepsilon_{ijklt}. \quad (3.1)$$

Function g represents the racial concordance effect net of the outcome benefit. Outcome_{jklt} is the outcome when patient j is referred to specialist k. $\mathcal{I}_{ijklt}$ is doctor i's information set when they make the referral decision. Distance_{jk} is the distance between patient j's residential area and the location where specialist k provides service. ε_{ijklt} is an i.i.d. mean-zero preference shock, and is independent of $\text{Race}_i, \text{Race}_j, \text{Race}_k, \mathbb{E}[\text{Outcome}_{jklt} | \mathcal{I}_{ijklt}]$, and Distance_{jk}. We are assuming travel distance to be a numeraire in the absence of prices, as commonly assumed in the empirical market design literature (Abdulkadiroğlu et al., 2017; Agarwal and Somaini, 2018; Huitfeldt et al., 2024).

Patient j's health outcome when treated by specialist k is:

$$\text{Outcome}_{jklt} = \gamma_1 \text{Race}_j + \gamma_2 \text{Race}_j \cdot \text{Race}_k + \xi_k + W_{jlt}\gamma_3 + e_{jklt}, \quad (3.2)$$

where ξ_k is the specialist-specific outcome effect and W_{jlt} is exogenous covariates.[11] We assume that outcome shock e_{jklt} satisfies:

$$\mathbb{E}[e_{jklt} | \text{Race}_j, \text{Race}_k, \xi_k, W_{jlt}] = 0.$$

The coefficient γ_2 is the racial concordance effect on the outcome. We include patients' sex, age, and comorbidities, and year-quarter fixed effects in W_{jlt}. We assume that doctor i's information set is:

$$\mathcal{I}_{ijklt} = \{\text{Race}_j, \text{Race}_k, \xi_k, W_{jlt}\}.$$

This assumption implies that doctor i knows patient j's conditions and the specialist k's ability ξ_k, while i does not have any information on the realization of the shock e_{jklt}.

Estimation

The estimation procedure consists of two steps: (i) estimating the parameters of the outcome model, and (ii) estimating the parameters of the referral decision model.

Step 1. Outcome In the first step, we estimate parameters of the outcome model. We estimate equation (3.2) by specialist fixed effects regression, where the term ξ_k is the fixed effect. We take THA/TKA revisions as the outcome of interest, i.e., Outcome = 1 if and only if the patient experienced a revision. This regression allows us to obtain parameter estimates, $\hat{\gamma}_1, \hat{\gamma}_2$, and $\hat{\gamma}_3$. Then, we recover the estimates for the fixed effects as the mean of

[11] In principle, we can extend the outcome model to include terms for the doctor's race and make the specialist-specific outcome effect depend on the doctor-patient race group combination. However, it will require the specialists to have a sufficient number of referrals for each doctor-patient race group combination for estimation.

the residuals. Furthermore, we correct for the estimation error of the fixed effects ξ_k through an empirical Bayes approach, following Kane and Staiger (2008).

Step 2. Referral Decision In the second step, we estimate parameters of the referral decision model. One concern when estimating this model is that the expected outcome, $\mathbb{E}[\text{Outcome}_{jklt}|\mathcal{I}_{jklt}]$, is estimated with error even after applying an empirical Bayes approach, since the number of observations per specialist is limited and revisions are relatively rare. To mitigate such concern, we take the control function approach (Petrin and Train, 2010). We assume the following first-stage equation:

$$\hat{\mathbb{E}}[\text{Outcome}_{jklt}|\mathcal{I}_{jklt}] = h(\text{Race}_i, \text{Race}_j, \text{Race}_k) + \delta_1 \text{Distance}_{jk} + \delta_2 Z_{ijklt} + \eta_{ijklt} \tag{3.3}$$

where h is a function that maps the tuple of races to some value. Z_{ijklt} are instruments, which are the inverse of distance between the exiting specialist and the patient, and the inverse of distance between the exiting specialist and the doctor, while exit has to occur before t.[12] We assume that Z_{ijklt} is independent of ε_{ijklt} and η_{ijklt}. The underlying assumption is that the exiting behavior of a specialist is independent of the unobserved taste for specialists and the unobserved component of the health outcomes. Moreover, exit of a specialist affects health outcomes in the same HRR, while this pattern can be explained through a competition effect and a volume effect (Lau et al., 2012). We further assume that:

$$\varepsilon_{ijklt} = \lambda \eta_{ijklt} + \tilde{\varepsilon}_{ijklt}.$$

Let $\zeta_{ijklt} = \tilde{\varepsilon}_{ijklt} - \beta \left(\hat{\mathbb{E}}[\text{Outcome}_{jklt}|\mathcal{I}_{jklt}] - \mathbb{E}[\text{Outcome}_{jklt}|\mathcal{I}_{jklt}] \right)$. We assume that the conditional distribution, $\zeta_{ijklt}|\{\text{Race}_i, \text{Race}_j, \text{Race}_k, \text{Distance}_{jk}, Z_{ijklt}, \eta_{ijklt}\}$, follows a type I extreme value distribution.

First, we estimate equation (3.3) and obtain the residuals $\hat{\eta}_{ijklt}$. Then, we estimate the parameters based on the following equation:

$$u_{ijlt}(k) = g(\text{Race}_i, \text{Race}_j, \text{Race}_k) + \beta \hat{\mathbb{E}}[\text{Outcome}_{jklt}|\mathcal{I}_{ijklt}] + \gamma \text{Distance}_{jk} + \lambda \hat{\eta}_{ijklt} + \zeta_{ijklt}.$$

For $\hat{\mathbb{E}}[\text{Outcome}_{jklt}|\mathcal{I}_{ijklt}]$ and $\hat{\eta}_{ijklt}$, we use the estimates from the previous steps.

3.5 Results

First, we present results for our outcome model. Table 3.4 shows coefficient estimates for the outcome model (equation (3.2)). The racial concordance effect on the outcome (γ_2) is not statistically significant, except for the case $R =$ Hispanic where the significance is relatively weak.[13] Thus, racial concordance between the patient and specialist is not likely to change the likelihood of THA/TKA revision. THAs are 0.8 pp more likely to result in a revision than TKAs. Appendix Table C.2 displays coefficient estimates for patient comorbidities. More comorbidities result in a higher likelihood of THA/TKA revision.

[12] Explicitly, we set $Z = 1/(1+d)$ where d is the distance. If there was no exit, we set $Z = 0$.

[13] Note that the sign of γ_2 for Hispanic is even in the direction of a negative effect of racial concordance.

Table 3.4: THA/TKA outcome regressions

	Outcome: THA/TKA Revision (%)			
	R = Asian	R = Black	R = Hispanic	R = White
	(1)	(2)	(3)	(4)
Race				
Race R Patient, γ_1	-0.0319	-0.1678***	-0.3389***	0.1856**
	(0.1094)	(0.0450)	(0.0784)	(0.0722)
R Patient $\cdot$ R Specialist, γ_2	-0.1309	-0.0297	0.3899*	-0.0276
	(0.1973)	(0.1984)	(0.2232)	(0.0813)
Patient Characteristics				
THA	0.8052***	0.8051***	0.8038***	0.8028***
	(0.0266)	(0.0266)	(0.0266)	(0.0266)
Male	0.1292***	0.1254***	0.1291***	0.1264***
	(0.0179)	(0.0179)	(0.0178)	(0.0180)
Patient Age	X	X	X	X
Patient Comorbidities	X	X	X	X
Specialist Fixed Effects	X	X	X	X
Year-quarter Fixed Effects	X	X	X	X
Number of Observations	1,239,535	1,239,535	1,239,535	1,239,535
R^2	0.02796	0.02797	0.02797	0.02798

Notes: Standard errors are clustered at the HRR and year-quarter level. Comorbidities are coded by the Chronic Condition Warehouse. * p-value $< .1$, ** p-value $< .05$, *** p-value $< .01$.

Table 3.5 compares the distribution of specialist-specific outcome effects across specialist race groups. On average, White specialists have slightly better outcomes than non-White specialists. White specialists are expected to have a 0.11 pp lower probability of revisions, but the standard deviation of the distribution is 0.64 pp. For other race groups, we see the opposite pattern. The overall negative average indicates that specialists who have performed more THA/TKA cases have worse specialist-specific outcome effects. The empirical Bayes shrinkage procedure can be one of the reasons for this since the procedure shrinks less experienced specialists' outcome effect more toward zero.

Table 3.6 shows coefficient estimates for the referral decision model (equation (3.1)). We find racially concordant referrals, conditional on patient benefits for THA/TKA, which consists of expected health outcome and distance to the specialist. To illustrate the structural parameters, we interpret the estimated coefficients in terms of travel distance equivalents. An improvement of 1 pp in the revision rate is equivalently valued as 79.6–82.4 kilometers of travel to the specialist. Patients are more likely to be referred to Asian specialists. This

Table 3.5: Distribution of specialist-specific outcome effects ξ_k by specialist race group

	Race R Specialist			non-R Specialist		
	N	Mean	S.D.	N	Mean	S.D.
R = Asian	680	-0.6417	0.5891	12,118	-0.7583	0.6182
R = Black	285	-0.6420	0.5868	12,513	-0.7538	0.6176
R = Hispanic	305	-0.7030	0.6340	12,493	-0.7530	0.6169
R = White	11,388	-0.8044	0.6416	1,410	-0.6939	0.6281

tendency is valued as 8.8 kilometers of travel. Patients are unlikely to be sent to White specialists, where the effect is 20.5 kilometers of travel. Overall, we observe a positive effect of racial concordance between patients, doctors, and specialists for referral decisions. For example, For an Asian patient, Asian doctors value Asian specialists as 28.9 kilometers of travel, relative to non-Asian specialist. Even non-Asian doctors demonstrate a preference for racial concordance, valuing Asian specialists 14.6 kilometers of travel more than non-Asian specialists for Asian patients.

Table 3.6: Estimates of structural parameters for THA/TKA referral decision

	Referral Decision to a Specialist			
	R = Asian (1)	R = Black (2)	R = Hispanic (3)	R = White (4)
Distance between Patient and Specialist, γ	-5.821	-5.822	-5.827	-5.822
(100 km)	(0.048)	(0.048)	(0.047)	(0.048)
Expected Outcome (% Revision), β	-4.751	-4.797	-4.753	-4.636
	(1.185)	(1.226)	(1.154)	(1.188)
R Specialist	0.515	0.563	0.060	-1.193
	(0.275)	(0.382)	(0.254)	(0.239)
R Doctor/R Specialist	0.555	0.133	0.352	0.434
	(0.094)	(0.248)	(0.263)	(0.059)
R Patient/R Specialist	0.332	0.403	2.577	0.360
	(1.103)	(0.930)	(1.077)	(0.457)
R Patient/R Doctor/R Specialist	0.279	0.536	0.125	-0.128
	(0.181)	(0.236)	(0.351)	(0.049)
Residual($\hat{\eta}$), λ	4.008	4.054	4.011	3.915
	(1.189)	(1.228)	(1.156)	(1.189)

Notes: Standard errors are calculated using 100 cluster-bootstrap replications. Cluster is at the HRR-year level.

3.6 Counterfactual Analysis

We now use our structural estimates to predict the referral decisions when the effect of racial concordance is absent.[14] Our aim is to understand how racially concordant referrals influence patient outcomes and travel distances. Gaining insight into these effects allows us to assess whether such referrals contribute to disparities in health.

In specific, we consider the case where doctor i's utility of referring patient j to specialist k is:

$$\bar{u}_{ijlt}(k) = \beta \mathbb{E}[\text{Outcome}_{jklt}|\mathcal{I}_{ijklt}] + \gamma \text{Distance}_{jk} + \varepsilon_{ijklt}. \tag{3.4}$$

Note that we have eliminated the racial concordance term $g(\text{Race}_i, \text{Race}_j, \text{Race}_k)$ from equation (3.1). We compare the outcomes between two scenarios: (i) the baseline scenario, where the effect of racial concordance is present; and (ii) the counterfactual scenario where racial concordance effect is absent. We assume that the expected health outcome and agents' locations are unchanged between the two scenarios.

Table 3.7: Choice probabilities : Pr(Referred to Race R Specialist$|\cdot$)

Race R	Asian	Black	Hispanic	White
Baseline				
Race R Doctor · Race R Patient	0.360	0.149	0.400	0.935
Race R Doctor · non-Race R Patient	0.087	0.071	0.097	0.873
non-Race R Doctor · Race R Patient	0.144	0.055	0.101	0.860
non-Race R Doctor · non-Race R Patient	0.036	0.013	0.012	0.701
Overall Average	0.042	0.017	0.017	0.914
Counterfactual				
Race R Doctor · Race R Patient	0.146	0.044	0.151	0.957
Race R Doctor · non-Race R Patient	0.038	0.041	0.082	0.926
non-Race R Doctor · Race R Patient	0.091	0.026	0.027	0.922
non-Race R Doctor · non-Race R Patient	0.023	0.008	0.011	0.839
Overall Average	0.025	0.010	0.015	0.947

Table 3.7 presents the choice probabilities for doctors referring patients to a race R specialist, comparing the baseline scenario with the counterfactual scenario. We see a decrease in the choice probability of race R specialist for the cases where R is Asian, Black, and Hispanic. The level of decrease is stronger in the cases where the doctor and the specialist share the same race R (= {Asian/Black/Hispanic}). For example, when the patient and

[14] Here, we include the effect of a specialist being a specific race as part of the effect of racial concordance.

the doctor are both Asian, we see a significant reduction in the probability of referring the patient to Asian specialists (36.0% to 14.6%). Even for the case where the doctor is Asian and the patient is non-Asian, we see a reduction (8.7% to 3.8%). For the case where the doctor is non-Asian and the patient is Asian, we see a reduction as well (14.4% to 9.1%).

Table 3.8: Change in mean of Expected Revision Rate (%)

Race R	Asian	Black	Hispanic	White
Race R Doctor · Race R Patient	-0.032	-0.030	-0.031	-0.002
Race R Doctor · non-Race R Patient	-0.004	-0.004	-0.018	-0.008
non-Race R Doctor · Race R Patient	-0.006	-0.007	0.001	-0.005
non-Race R Doctor · non-Race R Patient	-0.001	-0.001	0.000	-0.018
Overall Average	-0.001	-0.002	0.000	-0.004

Table 3.9: Change in mean of Travel Distance (in kilometers)

Race R	Asian	Black	Hispanic	White
Race R Doctor · Race R Patient	0.405	0.419	0.880	0.011
Race R Doctor · non-Race R Patient	0.000	0.022	0.067	0.095
non-Race R Doctor · Race R Patient	0.086	0.016	0.028	0.010
non-Race R Doctor · non-Race R Patient	0.005	-0.007	0.000	0.199
Overall Average	0.013	-0.002	0.000	0.029

Table 3.8 presents the change in the mean of expected probability of receiving a revision between two scenarios.[15] When the doctor and the patient share the same race, we see an improvement in the expected revision rate. When R = Asian/Black/Hispanic, this improvement is about 0.03 pp reduction in the expected revision rate, which corresponds to a 4% relative improvement. Moreover, while the effects are smaller than these cases, we see an improvement in patients' health outcome in most of the pairs.

Table 3.9 reports the change in the mean of travel distance between two scenarios. When the doctor and the patient share the same race, we observe that patients experience a longer travel. When R = Asian/Black/Hispanic, the change is about 0.41–0.88 kilometers of travel. Instead of getting to referred to a same race specialist nearby to the patients' residential area, patients travel more to seek care from a specialist with better ability. Distributions of expected revision rates and travel distance are shown in the Appendix Figures C.3 and C.4.

[15]While we are aware of potential measurement error in the expected probability of revision, the results presented here are based on the estimated expected probability of revision without further adjustment.

3.7 Conclusion

We propose a structural framework to identify the racial concordance effect net of outcome benefits and travel cost in doctors' referral decisions. Applying the model to THA/TKA cases, we find evidence of racially concordant referrals conditional on outcome benefits and travel distance. Our results suggest that racially concordant referrals have a negative impact on health outcomes for some race groups, such as Asian, Black, or Hispanic, while we do not find outcome benefits from racial concordance. The negative effect is most pronounced when the referring doctor and patient share the same race. In the absence of racially concordant referrals, Asian, Black, and Hispanic patients are more frequently referred to White specialists. Since White specialists are associated with better outcomes in our setting, these patients experience improved health outcomes under this counterfactual scenario.

A key imitation of our study is that we do not disentangle the underlying motivations for racially concordant referrals. Such referrals may arise from multiple factors. For example, patients may prefer seeing a specialist with the same race, conditional on the specialist's quality and travel costs, since they may experience improved communication (Alsan et al., 2019). Next, doctors and specialists with the same race may have stronger professional or personal ties, which may reduce uncertainty about the specialist's quality. Our current approach does not isolate these channels, which may have an impact when we consider welfare.

We conclude by outlining directions for future research. Although we do not find evidence that racial concordance affects health outcomes, previous studies have shown that such effects can exist in other contexts (Bach et al., 2004; Alsan et al., 2019; Greenwood et al., 2020). Exploring the implications of racially concordant referrals in these alternative settings remains an important area for future research. We also would like to emphasize that our framework allows us to consider other types of concordance, such as shared medical school training or professional networks.

Furthermore, we do not account for potential capacity constraints that may become a problem in the absence of racially concordant referrals. For example, if the highest-performing specialists attract a disproportionate number of patients, they may be assigned a number of patients beyond their capacity. As a result, the change in outcomes we estimate may be overstated. Accounting for specialist capacity constraints is an important direction for future research.

Bibliography

Abaluck, Jason, Leila Agha, Chris Kabrhel, Ali Raja, and Arjun Venkatesh, "The determinants of productivity in medical testing: Intensity and allocation of care," American Economic Review, 2016, 106 (12), 3730–3764.

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak, "The welfare effects of coordinated assignment: Evidence from the New York City high school match," American Economic Review, 2017, 107 (12), 3635–3689.

Agarwal, Nikhil and Paulo Somaini, "Demand analysis using strategic reports: An application to a school choice mechanism," Econometrica, 2018, 86 (2), 391–444.

Aguirregabiria, Victor, Allan Collard-Wexler, and Stephen P Ryan, "Dynamic Games in Empirical Industrial Organization," in "Handbook of Industrial Organization," Vol. 4 2021, pp. 225–343.

_ and Arvind Magesan, "Identification and Estimation of Dynamic Games When Players' Beliefs Are Not in Equilibrium," Review of Economic Studies, 2020, 87 (2), 582–625.

_ and Junichi Suzuki, "Identification and Counterfactuals in Dynamic Models of Market Entry and Exit," Quantitative Marketing and Economics, 2014, 12, 267–304.

_ and Pedro Mira, "Sequential Estimation of Dynamic Discrete Games," Econometrica, 2007, 75 (1), 1–53.

Alsan, Marcella, Owen Garrick, and Grant Graziani, "Does Diversity Matter for Health? Experimental Evidence from Oakland," American Economic Review, 2019, 109 (12), 4071–4111.

Aradillas-Lopez, Andres and Adam M Rosen, "Inference in Ordered Response Games With Complete Information," Journal of Econometrics, 2022, 226 (2), 451–476.

Bach, Peter B., Hoangmai H. Pham, Deborah Schrag, Ramsey C. Tate, and J. Lee Hargraves, "Primary Care Physicians Who Treat Blacks and Whites," New England Journal of Medicine, 2004, 351 (6), 575–584.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin, "Estimating Dynamic Models of Imperfect Competition," Econometrica, 2007, 75 (5), 1331–1370.

_ , Han Hong, and Stephen P Ryan, "Identification and estimation of a discrete game of complete information," Econometrica, 2010, 78 (5), 1529–1568.

Becker, Gary S, The Economics of Discrimination, University of Chicago press, 1957.

Beresteanu, Arie, Ilya Molchanov, and Francesca Molinari, "Sharp Identification Regions in Models with Convex Moment Predictions," Econometrica, 2011, 79 (6), 1785–1821.

Bergemann, Dirk and Stephen Morris, "Robust Predictions in Games With Incomplete Information," Econometrica, 2013, 81 (4), 1251–1308.

_ and _ , "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," Theoretical Economics, 2016, 11 (2), 487–522.

_ , Benjamin Brooks, and Stephen Morris, "Counterfactuals with Latent Information," American Economic Review, 2022, 112 (1), 343–368.

Bernheim, B Douglas, "Strategic deterrence of sequential entry into an industry," The Rand Journal of Economics, 1984, pp. 1–11.

Bhattacharya, Vivek, James W Roberts, and Andrew Sweeting, "Regulating bidder participation in auctions," The RAND Journal of Economics, 2014, 45 (4), 675–704.

Bohren, J Aislinn, Alex Imas, and Michael Rosenberg, "The dynamics of discrimination: Theory and evidence," American Economic Review, 2019, 109 (10), 3395–3436.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," The Quarterly Journal of Economics, 2016, 131 (4), 1753–1794.

Bosio, Erica, Simeon Djankov, Edward Glaeser, and Andrei Shleifer, "Public procurement in law and practice," American Economic Review, 2022, 112 (4), 1091–1117.

Bugni, Federico A and Jackson Bunting, "On the Iterated Estimation of Dynamic Discrete Choice Games," Review of Economic Studies, 2021, 88 (3), 1031–1073.

Canay, Ivan A and Azeem M Shaikh, "Practical and Theoretical Advances in Inference for Partially Identified Models," Advances in Economics and Econometrics, 2017, 2, 271–306.

_ , Gastón Illanes, and Amilcar Velez, "A user's guide for inference in models defined by moment inequalities," Journal of Econometrics, 2023, p. 105558.

Chan, David C, Matthew Gentzkow, and Chuan Yu, "Selection with variation in diagnostic skill: Evidence from radiologists," The Quarterly Journal of Economics, 2022, 137 (2), 729–783.

Chandra, Amitabh and Douglas O Staiger, "Identifying sources of inefficiency in healthcare," The Quarterly Journal of Economics, 2020, 135 (2), 785–843.

_ , Pragya Kakani, and Adam Sacarny, "Hospital allocation and racial disparities in health care," Review of Economics and Statistics, 2024, 106 (4), 924–937.

Chernozhukov, Victor, Han Hong, and Elie Tamer, "Estimation and Confidence Regions for Parameter Sets in Econometric Models," Econometrica, 2007, 75 (5), 1243–1284.

Christensen, Peter and Christopher Timmins, "Sorting or steering: The effects of housing discrimination on neighborhood choice," Journal of Political Economy, 2022, 130 (8), 2110–2163.

Ciliberto, Federico and Elie Tamer, "Market Structure and Multiple Equilibria in Airline Markets," Econometrica, 2009, 77 (6), 1791–1828.

Dearing, Adam and Jason R Blevins, "Efficient and Convergent Sequential Pseudo-Likelihood Estimation of Dynamic Discrete Games," Review of Economic Studies, 2024, p. rdae050.

Dixit, Avinash, "A model of duopoly suggesting a theory of entry barriers," The Bell Journal of Economics, 1979, pp. 20–32.

Doraszelski, Ulrich and Mark Satterthwaite, "Computable Markov-Perfect Industry Dynamics," RAND Journal of Economics, 2010, 41 (2), 215–243.

Doval, Laura and Jeffrey C Ely, "Sequential Information Design," Econometrica, 2020, 88 (6), 2575–2608.

Duffie, Darrell, John Geanakoplos, Andreu Mas-Colell, and Andrew McLennan, "Stationary Markov Equilibria," Econometrica, 1994, 62 (4), 745–81.

Egesdal, Michael, Zhenyu Lai, and Che-Lin Su, "Estimating Dynamic Discrete-Choice Games of Incomplete Information," Quantitative Economics, 2015, 6 (3), 567–597.

Ellison, Glenn and Sara Fisher Ellison, "Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration," American Economic Journal: Microeconomics, 2011, 3 (1), 1–36.

Ely, Jeffrey C and Tanjim Hossain, "Sniping and squatting in auction markets," American Economic Journal: Microeconomics, 2009, 1 (2), 68–94.

Ericson, Richard and Ariel Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," Review of Economic Studies, 1995, 62 (1), 53.

Farrell, Joseph, "Cheap talk, coordination, and entry," The RAND Journal of Economics, 1987, pp. 34–39.

Fershtman, Chaim and Ariel Pakes, "Dynamic Games with Asymmetric Information: A Framework for Empirical Work," The Quarterly Journal of Economics, 2012, 127 (4), 1611–1661.

Fitzpatrick, Simon and Alastair Scott, "Quick Simultaneous Confidence Intervals for Multinomial Proportions," Journal of the American Statistical Association, 1987, 82 (399), 875–878.

Forges, Françoise, "Correlated Equilibria and Communication in Games," Complex Social and Behavioral Systems: Game Theory and Agent-Based Models, 2020, pp. 107–118.

Fudenberg, Drew and Jean Tirole, Game Theory, MIT press, 1991.

Galichon, Alfred and Marc Henry, "Set identification in models with multiple equilibria," Review of Economic Studies, 2011, 78 (4), 1264–1298.

Gentry, Matthew and Tong Li, "Identification in auctions with selective entry," Econometrica, 2014, 82 (1), 315–344.

Goolsbee, Austan and Chad Syverson, "How do incumbents respond to the threat of entry? Evidence from the major airlines," The Quarterly Journal of Economics, 2008, 123 (4), 1611–1633.

Greenwood, Brad N., Rachel R. Hardeman, Laura Huang, and Aaron Sojourner, "Physician–Patient Racial Concordance and Disparities in Birthing Mortality for Newborns," Proceedings of the National Academy of Sciences, 2020, 117 (35), 21194–21200.

Grieco, Paul LE, "Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets," RAND Journal of Economics, 2014, 45 (2), 303–340.

Gualdani, Cristina and Shruti Sinha, "Identification in discrete choice models with imperfect information," Journal of Econometrics, 2024, 244 (1), 105854.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong, "Optimal nonparametric estimation of first-price auctions," Econometrica, 2000, 68 (3), 525–574.

Guler, Ali Umut, "Inferring the Economics of Store Density from Closures: The Starbucks Case," Marketing Science, 2018, 37 (4), 611–630.

Han, Sukjin, Hiroaki Kaido, and Lorenzo Magnolfi, "Testing Information Ordering for Strategic Agents," arXiv preprint arXiv:2402.19425, 2024.

Harstad, Ronald M, John H Kagel, and Dan Levin, "Equilibrium bid functions for auctions with an uncertain number of bidders," Economics Letters, 1990, 33 (1), 35–40.

He, Wei and Yeneng Sun, "Stationary Markov Perfect Equilibria in Discounted Stochastic Games," Journal of Economic Theory, 2017, 169, 35–61.

Heckmann, Nathanael D., JaeWon Yang, Kevin L. Ong, Edmund C. Lau, Brian C. Fuller, Daniel D. Bohl, and Craig J. Della Valle, "Revision Surgery for Instability After Total Hip Arthroplasty: Does Timing Matter?," The Journal of Arthroplasty, 2021, 36 (5), 1779–1783.e2.

Hedlundh, Urban, Lennart Ahnfelt, C.-H. Hybbinette, Johan Weckström, and Hans Fredin, "Surgical Experience Related to Dislocations after Total Hip Arthroplasty," The Journal of Bone and Joint Surgery. British volume, 1996, 78-B (2), 206–209.

Hoffman, Kelly M, Sophie Trawalter, Jordan R Axt, and M Norman Oliver, "Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites," Proceedings of the National Academy of Sciences, 2016, 113 (16), 4296–4301.

Horowitz, Joel L and Sokbae Lee, "Inference in a class of optimization problems: Confidence regions and finite sample bounds on errors in coverage probabilities," Journal of Business & Economic Statistics, 2023, 41 (3), 927–938.

Hotz, V Joseph and Robert A Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," Review of Economic Studies, 1993, 60 (3), 497–529.

Huitfeldt, Ingrid, Victoria Marone, and Daniel C Waldinger, "Designing Dynamic Reassignment Mechanisms: Evidence from GP Allocation," Technical Report, National Bureau of Economic Research 2024.

Imai, Kosuke and Kabir Khanna, "Improving ecological inference by predicting individual ethnicity from voter registration records," Political Analysis, 2016, 24 (2), 263–272.

Kaido, Hiroaki, Francesca Molinari, and Jörg Stoye, "Confidence Intervals for Projections of Partially Identified Parameters," Econometrica, 2019, 87 (4), 1397–1432.

Kalouptsidi, Myrto, Paul T Scott, and Eduardo Souza-Rodrigues, "On the Non-identification of Counterfactuals in Dynamic Discrete Games," International Journal of Industrial Organization, 2017, 50, 362–371.

_ , _ , and _ , "Identification of Counterfactuals in Dynamic Discrete Choice Models," Quantitative Economics, 2021, 12 (2), 351–403.

Kane, Thomas J and Douglas O Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," NBER Working Paper No. 14607 2008.

Kaplan, Todd R and Shmuel Zamir, "Asymmetric first-price auctions with uniform distributions: analytic solutions to the general case," Economic Theory, 2012, 50 (2), 269–302.

Kline, Brendan and Elie Tamer, "Bayesian Inference in a Class of Partially Identified Models," Quantitative Economics, 2016, 7 (2), 329–366.

Koh, Paul S, "Stable Outcomes and Information in Games: An Empirical Framework," Journal of Econometrics, 2023, 237 (1), 105499.

Krasnokutskaya, Elena and Katja Seim, "Bid preference programs and participation in highway procurement auctions," American Economic Review, 2011, 101 (6), 2653–2686.

Lau, Rick L, Anthony V Perruccio, Rajiv Gandhi, and Nizar N Mahomed, "The role of surgeon volume on patient outcome in total knee arthroplasty: a systematic review of the literature," BMC musculoskeletal disorders, 2012, 13, 1–10.

Li, Tao, Haining Zhang, Ping Keung Chan, Wing Chiu Fung, Henry Fu, and Kwong Yuen Chiu, "Risk Factors Associated with Surgical Site Infections Following Joint Replacement Surgery: A Narrative Review," Arthroplasty, 2022, 4 (1), 11.

Li, Tong and Xiaoyong Zheng, "Entry and competition effects in first-price auctions: Theory and evidence from procurement auctions," The Review of Economic Studies, 2009, 76 (4), 1397–1429.

Liscow, Zachary, William Nober, and Cailin Slattery, "Procurement and Infrastructure Costs," Working Paper 2024.

Logan, John R, Zengwang Xu, and Brian J Stults, "Interpolating Us Decennial Census Tract Data From as Early as 1970 to 2010: A Longitudinal Tract Database," Professional Geographer, 2014, 66 (3), 412–420.

Magnolfi, Lorenzo and Camilla Roncoroni, "Estimation of Discrete Games with Weak Assumptions on Information," The Review of Economic Studies, 2023, 90 (4), 2006–2041.

Mailath, George J and Larry Samuelson, Repeated Games and Reputations: Long-Run Relationships, Oxford University Press, 2006.

Makris, Miltiadis and Ludovic Renou, "Information Design in Multistage Games," Theoretical Economics, 2023, 18 (4), 1475–1509.

Maskin, Eric and Jean Tirole, "Markov perfect equilibrium: I. Observable actions," Journal of Economic Theory, 2001, 100 (2), 191–219.

_ and John Riley, "Asymmetric auctions," The Review of Economic Studies, 2000, 67 (3), 413–438.

McAfee, R Preston and John McMillan, "Auctions with a stochastic number of bidders," Journal of Economic Theory, 1987, 43 (1), 1–19.

Meek, R M. D, D. B. Allan, G. McPhillips, L. Kerr, and C. R. Howie, "Epidemiology of Dislocation after Total Hip Arthroplasty," Clinical Orthopaedics & Related Research, 2006, 447 (447), 9–18.

Milgrom, Paul and John Roberts, "Limit pricing and entry under incomplete information: An equilibrium analysis," Econometrica: Journal of the Econometric Society, 1982, pp. 443–459.

_ and _ , "Predation, reputation, and entry deterrence," Journal of Economic Theory, 1982, 27 (2), 280–312.

Miller, Conrad and Ian M. Schmutte, "The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil," NBER Working Paper No. 29246 2021.

Morton, Fiona M Scott, "Barriers to entry, brand advertising, and generic entry in the US pharmaceutical industry," International Journal of Industrial Organization, 2000, 18 (7), 1085–1104.

Nichols, Christine I. and Joshua G. Vose, "Clinical Outcomes and Costs within 90 Days of Primary or Revision Total Joint Arthroplasty," The Journal of Arthroplasty, 2016, 31 (7), 1400–1406.e3.

Nowak, A.S. and T.E.S. Raghavan, "Existence of Stationary Correlated Equilibria with Symmetric Information for Discounted Stochastic Games," Mathematics of Operations Research, 1992, 17 (3), 519–526.

Olea, José Luis Montiel and Mikkel Plagborg-Møller, "Simultaneous Confidence Bands: Theory, Implementation, and an Application to SVARs," Journal of Applied Econometrics, 2019, 34 (1), 1–17.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii, "Moment Inequalities and Their Application," Econometrica, 2015, 83 (1), 315–334.

_ , Michael Ostrovsky, and Steven Berry, "Simple Estimators for the Parameters of Discrete Dynamic Games (With Entry/Exit Examples)," RAND Journal of Economics, 2007, 38 (2), 373–399.

Parsons, Helen M, Lindsey R Enewold, Robert Banks, Michael J Barrett, and Joan L Warren, "Creating a National Provider Identifier (NPI) to unique physician identification number (UPIN) crosswalk for Medicare data," Medical care, 2017, 55 (12), e113–e119.

Pesendorfer, Martin and Philipp Schmidt-Dengler, "Asymptotic Least Squares Estimators for Dynamic Games," Review of Economic Studies, 2008, 75 (3), 901–928.

Petrin, Amil and Kenneth Train, "A control function approach to endogeneity in consumer choice models," Journal of Marketing Research, 2010, 47 (1), 3–13.

Postler, Anne, Cornelia Lützner, Franziska Beyer, Eric Tille, and Jörg Lützner, "Analysis of Total Knee Arthroplasty Revision Causes," BMC Musculoskeletal Disorders, 2018, 19 (1), 55.

Quint, Daniel and Kenneth Hendricks, "A theory of indicative bidding," American Economic Journal: Microeconomics, 2018, 10 (2), 118–151.

Rust, John, "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica: Journal of the Econometric Society, 1987, pp. 999–1033.

_ , "Structural Estimation of Markov Decision Processes," in "Handbook of Econometrics," Vol. 4 1994, chapter 51, pp. 3081–3143.

Sarsons, Heather, "Interpreting Signals in the Labor Market: Evidence from Medical Referrals," Working Paper 2022.

Seidelman, Jessica L., Christopher R. Mantyh, and Deverick J. Anderson, "Surgical Site Infection Prevention," JAMA, 2023, 329 (3), 244.

Silva, Dakshina G De, Timothy Dunne, Anuruddha Kankanamge, and Georgia Kosmopoulou, "The impact of public information on bidding in highway procurement auctions," European Economic Review, 2008, 52 (1), 150–181.

Singh, Manasvini and Atheendar Venkataramani, "Rationing by Race," Technical Report, National Bureau of Economic Research 2022.

Solan, Eilon, "Characterization of Correlated Equilibria in Stochastic Games," International Journal of Game Theory, 2001, 30 (2), 259–277.

_ and Nicolas Vieille, "Correlated Equilibrium in Stochastic Games," Games and Economic Behavior, 2002, 38 (2), 362–399.

Spence, A Michael, "Entry, capacity, investment and oligopolistic pricing," The Bell Journal of Economics, 1977, pp. 534–544.

Stengel, Bernhard Von and Françoise Forges, "Extensive-Form Correlated Equilibrium: Definition and Computational Complexity," Mathematics of Operations Research, 2008, 33 (4), 1002–1022.

Studdert, David M, Yifan Zhang, Sonja A Swanson, Lea Prince, Jonathan A Rodden, Erin E Holsinger, Matthew J Spittal, Garen J Wintemute, and Matthew Miller, "Handgun ownership and suicide in California," New England journal of medicine, 2020, 382 (23), 2220–2229.

Su, Che-Lin and Kenneth L Judd, "Constrained Optimization Approaches to Estimation of Structural Models," Econometrica, 2012, 80 (5), 2213–2230.

Sweeting, Andrew, James W Roberts, and Chris Gedge, "A model of dynamic limit pricing with an application to the airline industry," Journal of Political Economy, 2020, 128 (3), 1148–1193.

Syrgkanis, Vasilis, Elie Tamer, and Juba Ziani, "Inference on Auctions with Weak Assumptions on Information," Working Paper 2021.

Tamer, Elie, "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," Review of Economic Studies, 2003, 70 (1), 147–165.

Technomic, "Top 500 Chain Restaurant Report," 2019. Accessed via `https://www.restaurantbusinessonline.com/financing/largest-coffee-cafe-chains-us`.

Thomadsen, Raphael, "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," RAND Journal of Economics, 2005, pp. 908–929.

Tsiatis, Anastasios, "A nonidentifiability aspect of the problem of competing risks.," Proceedings of the National Academy of Sciences, 1975, 72 (1), 20–22.

Urish, Kenneth L., Yongmei Qin, Bassel Salka, Benjamin Y. Li, Tudor Borza, Michael Sessine, Peter Kirk, Brent K. Hollenbeck, Jonathan E. Helm, Mariel S. Lavieri, Ted A. Skolarus, and Bruce L. Jacobs, "Comparison of Readmission and Early Revision Rates as a Quality Metric in Total Knee Arthroplasty Using the Nationwide Readmission Database," Annals of Translational Medicine, 2020, 8 (11), 687–687.

Wennberg, John E. and Megan McAndrew Cooper, The Dartmouth Atlas of Health Care in the United States, Chicago, Illinois: American Hospital Publishing, Inc., 1996.

Wilson, Robert, "Strategic models of entry deterrence," Handbook of game theory with economic applications, 1992, 1, 305–329.

Ye, Lixin, "Indicative bidding and a theory of two-stage auctions," Games and Economic Behavior, 2007, 58 (1), 181–207.

Zeltzer, Dan, "Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap," American Economic Journal: Applied Economics, 2020, 12 (2), 169–197.

Appendix A

Appendix to Chapter 1

A.1 Proof: Equilibrium for Two-agent Example

Let agent i's inverse bid function be $\phi_i = \beta_i^{-1}$. Cost is distributed $\sim U[0,1]$. Player 1 enters with probability 1. Player 1's belief on Player 2's entry is with probability r. Let the reserve price be $R = 1$.

Derivation and proof for β_1 and β_2 being the equilibrium bidding strategies when Player 1 enters and discloses while Player 2 enters but does not disclose, follows Kaplan and Zamir (2012).

Here, we focus on the proof of optimality of the disclosure strategies taken in our example. First, let us consider the required condition that forces the first entrant to disclose. To consider this condition, let us first prove the following claim:

Claim: Fix r. $\phi_1(b) > r\phi_2(b)$ for all $b \in \left(\frac{1}{1+r}, 1\right)$. Note that support of the bids for each player is $\left[\frac{1}{1+r}, 1\right]$.

Proof: Let $X = \frac{r^2}{(1-r)^2}\log\frac{1-b}{\frac{1}{r}-b} - r(1+r) - \frac{2r^2}{(1-r)^2}\log(r) - \frac{r^2(1+r)}{1-r}$. Then,

$$\phi_1 = 1 - \frac{1}{(b-\frac{1}{r})X - \frac{r}{1-r}},\ \phi_2 = \frac{1}{r} - \frac{1}{-(b-1)X + \frac{r}{1-r}}.$$

To show $\phi_1 > r\phi_2$, it is equivalent to show that:

$$-r(1+r)/(1-r) > (2-br-b)X$$

holds. Let $h(b) = (2-br-b)X$. Note that $h(\frac{1}{1+r}) = -r(1+r)/(1-r)$. Observe that:

$$h'(b) = -(1+r)X + (2-br-b)\frac{r^2}{(1-r)^2}\left(\frac{1}{b-1} - \frac{1}{b-\frac{1}{r}}\right)$$

Again, note that $h'(\frac{1}{1+r}) = 0$. Next,

$$h''(b) = -2(1+r)\frac{r^2}{(1-r)^2}\left(\frac{1}{b-1} - \frac{1}{b-\frac{1}{r}}\right) + (2-br-b)\frac{r^2}{(1-r)^2}\left(-\frac{1}{(b-1)^2} + \frac{1}{(b-\frac{1}{r})^2}\right)$$
$$= \frac{1}{(1-r)^2}(r-1)^3 b < 0$$

Thus, $h'(b) < 0$ for all $b \in (\frac{1}{1+r}, 1)$. Therefore, $h(b) < -r(1+r)/(1-r)$ holds for all $b \in (\frac{1}{1+r}, 1)$. This is what we wanted to show. ■

Now, we are ready to prove the following claim.

Claim: Fix r. $EU_1(c) > EU_2(c)$ for all $c \in (0,1)$. $EU_i(c)$ is the expected utility for Player i conditional on entering and cost c, when players take strategies β_i and β_j.

Proof: Observe that:

$$\begin{aligned} EU_1(c) &= (\beta_1(c) - c)(1-r) + (\beta_1(c) - c)r(1 - \phi_2(\beta_1(c))) \\ &= (b_1(c) - c)(1 - r\phi_2(\beta_1(c))) \end{aligned}$$

and

$$EU_2(c) = (b_2(c) - c)(1 - \phi_1(b_2(c))).$$

By the previous claim, we know that $r\phi_2(b_2(c)) < \phi_1(b_2(c))$. Now, the following holds:

$$\begin{aligned} (b_1(c) - c)(1 - r\phi_2(b_1(c))) &\geq (b_2(c) - c)(1 - r\phi_2(b_2(c))) \\ &> (b_2(c) - c)(1 - \phi_1(b_2(c))) \end{aligned}$$

Hence, $EU_1(c) > EU_2(c)$ holds for all $c \in (0,1)$. ■

This claim implies that if a Player is entering, they prefer being in the position of Player 1. Therefore, if you arrive and enter first, it is optimal for the player to disclose their entry.

The next required condition is that if you are the second entrant facing one disclosure, you do not disclose. To show that this condition holds, we prove the following claim:

Claim: $2b - 1 > \phi_1(b)$ for all $b \in (\frac{1}{1+r}, 1)$ when $r < 1$.

Proof: We will prove $\frac{1}{2(1-b)} > \frac{1}{1-\phi_1}$. Let

$$H(b) = \frac{1}{2(1-b)} - (b - \frac{1}{r})X + \frac{r}{1-r},$$

where $X = \frac{r^2}{(1-r)^2}\log\frac{1-b}{\frac{1}{r}-b} - r(1+r) - \frac{2r^2}{(1-r)^2}\log(r) - \frac{r^2(1+r)}{1-r}$. We shall show $H(b) > 0$ for all $b \in (\frac{1}{1+r}, 1)$. Observe that:

$$H'(b) = \frac{1}{2(1-b)^2} - \left(b - \frac{1}{r}\right)\frac{r^2}{(1-r)^2}\left(\frac{1}{b-1} - \frac{1}{b-\frac{1}{r}}\right) - X$$

and

$$H'\left(\frac{1}{1+r}\right) = \frac{1}{2r^2}(1-r^2)(1+2r) > 0$$

Next,

$$\begin{aligned}
H''(b) &= \frac{1}{(1-b)^3} - \frac{2r^2}{(1-r)^2}\left(\frac{1}{b-1} - \frac{1}{b-\frac{1}{r}}\right) - \left(b-\frac{1}{r}\right)\frac{r^2}{(1-r)^2}\left(-\frac{1}{(b-1)^2} + \frac{1}{(b-\frac{1}{r})^2}\right) \\
&= \frac{1}{(b-1)^3(b-\frac{1}{r})}\frac{r^2}{(1-r)^2}\frac{1-r}{r}\left[-\frac{1-r}{r}(b-\frac{1}{r}) + 2(b-1)^2 + (b-1)(-2b+1+\frac{1}{r})\right] \\
&= \frac{1}{(b-1)^3(b-\frac{1}{r})}\frac{r}{(1-r)}\left[-\frac{1-r}{r}(b-\frac{1}{r}) + 2(b-1)^2 + (b-1)(-2b+1+\frac{1}{r})\right] \\
&= \frac{1}{(b-1)^3(b-\frac{1}{r})}\frac{r}{(1-r)}\left(\frac{1}{r}-1\right)^2 > 0
\end{aligned}$$

Thus, $H'(b) > 0$ holds for all $b \in (\frac{1}{1+r}, 1)$. Moreover, since $H(\frac{1}{1+r}) > 0$, we get $H(b) > 0$ for all $b \in (\frac{1}{1+r}, 1)$. This completes our proof. ■

Now, to complete our discussion, we show that it is optimal for the second entrant to not disclose.

Claim: Let the expected player's profit be $V(c)$ when both bidders take strategy $b(c) = \frac{c+1}{2}$. Then, $EU_2(c) > V(c)$ holds.

Proof: Observe that:

$$\begin{aligned}
V(c) &= (b(c)-c)(1-(2b(c)-1)) \\
&< (b(c)-c)(1-\phi_1(b(c))) \\
&\leq (\beta_2(c)-c)(1-\phi_1(\beta_2(c))) \\
&= EU_2(c).
\end{aligned}$$

First inequality holds since $2b-1 > \phi_1(b)$ holds by the previous claim. The second inequality holds by the optimality of β_2 when the other bidder employs β_1. ■

Combining the results from these claims, we have shown that the disclosure behavior presented in our example satisfies the equilibrium requirements.

A.2 Step 2 of Identification: Asymmetric Case

Suppose that under h^τ, bidders $j_1, \ldots, j_J$ signals before i each at $\tau_{j_1}, \ldots, \tau_{j_J}$ $(\tau_{j_1} < \cdots < \tau_{j_J})$, and for the rest of bidders $k_1, \ldots, k_K$, their signals are yet to be observed. Take τ_k such that $\tau_{j_m} < \tau_k < \tau$ holds for all m. Let

$$A_i^t(h, c_i) \equiv F_E^i(V_i^t(h)) F_Q(\Delta v_i(h, c_i)) \tag{A.1}$$

We consider the following density P:

$$\begin{aligned} P &= \Pr\left(j_m \text{ signals at } \tau_{j_m}\ \forall m, i \text{ signals at } \tau, k_n \text{ does not signal before } \tau\ \forall n,\ \vec{c}_j, c_i\right) \\ &= \prod_m f_\tau^{j_m}(\tau_{j_m}) A_{j_m}^{\tau_{j_m}}(h^{\tau_{j_m}}(\tau_{j_1}, \ldots, \tau_{j_J}), c_{j_m}) f_{c_{j_m}}(c_{j_m}) \\ &\times \prod_n \left\{ 1 - F_\tau^{k_n}(\tau) + \int_0^\infty \int_0^\tau f_\tau^{k_n}(t) \left(1 - A_{k_n}^t(h_t(\tau_{j_1}, \ldots, \tau_{j_J}), c_{k_n})\right) f_{c_{k_n}}(c_{k_n}) dt dc_{k_n} \right\} \\ &\times f_\tau^i(\tau) A_i^\tau(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_{c_i}(c_i) \end{aligned} \tag{A.2}$$

We consider the following density Q:

$$\begin{aligned} Q &= \Pr\left(j_m \text{ signals at } \tau_{j_m}\ \forall m, i \text{ does not signal before } \tau, k_n \text{ does not signal before } \tau\ \forall n,\ \vec{c}_j\right) \\ &= \prod_m f_\tau^{j_m}(\tau_{j_m}) A_{j_m}^{\tau_{j_m}}(h^{\tau_{j_m}}(\tau_{j_1}, \ldots, \tau_{j_J}), c_{j_m}) f_{c_{j_m}}(c_{j_m}) \\ &\times \prod_n \left\{ 1 - F_\tau^{k_n}(\tau) + \int_0^\infty \int_0^\tau f_\tau^{k_n}(t) \left(1 - A_{k_n}^t(h_t(\tau_{j_1}, \ldots, \tau_{j_J}), c_{k_n})\right) f_{c_{k_n}}(c_{k_n}) dt dc_{k_n} \right\} \\ &\times \left\{ 1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h_t(\tau_{j_1}, \ldots, \tau_{j_J}), c_i)\right) f_{c_i}(c_i) dt dc_i \right\} \end{aligned} \tag{A.3}$$

Taking the ratio between these two densities gives us:

$$P/Q = \frac{f_\tau^i(\tau) A_i^\tau(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_{c_i}(c_i)}{\left\{1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h_t(\tau_{j_1}, \ldots, \tau_{j_J}), c_i)\right) f_{c_i}(c_i) dt dc_i\right\}} \tag{A.4}$$

Exploiting the relation that

$$\begin{aligned} &\frac{\partial(1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h^\tau, c_i)\right) f_c(c_i) dt dc_i)}{\partial \tau} \\ &= \int_0^\infty f_\tau^i(\tau) A_i^\tau(h^\tau(\tau_{j_1}, \ldots, \tau_{j_J}), c_i) f_c(c_i) dc_i, \end{aligned}$$

the function

$$\begin{aligned} \Gamma_i(\tau; h^\tau = (\tau_{j_1}, \ldots, \tau_{j_J})) &= 1 - F_\tau^i(\tau) + \int_0^\infty \int_0^\tau f_\tau^i(t) \left(1 - A_i^t(h^\tau, c_i)\right) f_c(c_i) dt dc_i \\ &= 1 - \int_0^\infty \int_0^\tau f_\tau^i(t) A_i^t(h^\tau, c_i) f_c(c_i) dt dc_i \end{aligned}$$

is identified up to scale for all $\tau \in [\tau_{j_J}, T]$. Since $\Gamma_i(0; h^\tau = \phi)$=1 holds, $\Gamma_i(\tau; h^\tau = \phi)$ is identified. Therefore, $f^i_\tau(\tau)A^\tau_i(h^\tau = \phi, c_i)f_c(c_i)$ is identified for all $\tau \in [0, T]$.

Now, given that $f^i_\tau(\tau)A^\tau_i(h^\tau = \phi, c_i)f_c(c_i)$ is identified, $\Gamma_i(\tau; h^\tau = \tau)$ is identified. As a result, $\Gamma_i(t; h^t = \tau)$ $(t \geq \tau)$ such that h^t includes one disclosure at τ is identified. By induction on the number of disclosures made, repeating this argument will allow us to identify $\Gamma_i(t; h^\tau)$ for all histories h^τ. Note that $f^i_\tau(\tau)A^\tau_i(h^\tau, c_i)f_c(c_i)$ is also identified for all h^τ.

Let

$$\begin{aligned} R = \quad & \prod_m f^{j_m}_\tau(\tau_{j_m}) A^{\tau_{j_m}}_{j_m}(h^{\tau_{j_m}}(\tau_{j_1}, \ldots, \tau_{j_{m-1}}), c_{j_m}) f_{c_{j_m}}(c_{j_m}) \\ & \times \prod_p f^{l_p}_\tau(\tau_{l_p}) A^{\tau_{l_p}}_{l_p}(h^{\tau_{l_p}}(\tau_{j_1}, \ldots, \tau_{j_J}, \tau_{l_1}, \ldots, \tau_{l_{p-1}}), c_{l_p}) f_{c_{l_p}}(c_{l_p}) \\ & \times \prod_n \Gamma_{k_n}(T; h^\tau = (\tau_{j_1}, \ldots, \tau_{j_J}, \tau_{l_1}, \ldots, \tau_{l_L})) \end{aligned}$$

and

$$\begin{aligned} S = \quad & \prod_m f^{j_m}_\tau(\tau_{j_m}) A^{\tau_{j_m}}_{j_m}(h^{\tau_{j_m}}(\tau_{j_1}, \ldots, \tau_{j_{m-1}}), c_{j_m}) f_{c_{j_m}}(c_{j_m}) \\ & \times \prod_p \Gamma_{l_p}(\tau; h^\tau = (\tau_{j_1}, \ldots, \tau_{j_J})) \\ & \times \prod_n \Gamma_{k_n}(\tau; h^\tau = (\tau_{j_1}, \ldots, \tau_{j_J})). \end{aligned}$$

Since all the objects that appear in R and S are identified, R and S are identified.

Belief of i can be written as:

$$\Pr_i(h^T | h^\tau, \tau_i = \tau, A^Q_i = 0) = R/S$$

and since R and S are identified, this object is also identified.

A.3 Details on the Estimation

We discuss the details on the estimation that we omit from the main text.

Step 1. Construction costs c_i We estimate G^{-i}, the CDF of the minimum of the bids among the opponents and the secret reserve price. We assume that G^{-i} and the secret reserve price both follows a log-normal distribution. Parametric specification of G^{-i} is given in the main text. We assume that secret reserve price follows log-$\mathcal{N}(\mu^r, \sigma^r)$.

We split our observations into four cases, and show the corresponding likelihood for each case below:

1. When winner was accepted but the bidder lost the auction:

$$g_{-i}(b)$$

   where b is the winner's bid.

2. When you are the winner but rejected:

$$\frac{\int_0^b g_{-i}(\tilde{b}) f_r(\tilde{b}) d\tilde{b}}{F_r(b)}$$

3. When you are the winner and accepted, there is another bidder:

$$\frac{g_{-i}(b^2)(1 - F_r(b^2)) + \int_b^{b^2} g_{-i}(\tilde{b}) f_r(\tilde{b}) d\tilde{b}}{1 - F_r(b)}$$

   where b is the winner's bid and b^2 is the highest opponent's bid.

4. When you are the winner and accepted, there is no other bidder:

$$\frac{\int_b^{+\infty} g_{-i}(\tilde{b}) f_r(\tilde{b}) d\tilde{b}}{1 - F_r(b)}$$

   where b is the winner's bid and b^2 is the highest opponent's bid.

Using this likelihood, we estimate parameters for G^{-i} and the distribution of secret reserve price via maximum likelihood. Integration is computed by evaluating values on a grid.

Given our estimate $\hat{G}^{-i}$, we estimate construction costs for each entrant by:

$$\hat{c}_i = b_i - \frac{1 - \hat{G}_{-i}(b)}{\hat{g}_{-i}(b)}.$$

Step 2. Belief on the evolution of disclosure history Parameters regarding the distribution of time intervals between the n-th and $(n+1)$-th disclosure, $(\beta^{\mu_t}, \beta^{\sigma_t})$ are estimated via maximum likelihood as described in the main text.

Given the estimates on time intervals, we estimate the belief of a firm when the firm discloses at some time τ facing history h^τ. To obtain this estimate, we simulate the evolution of disclosure histories. The simulation procedure proceeds as follows:

1. Suppose that a firm makes the n-th disclosure. We fix the firm's entry and disclosure timing at $\tau^n \in \{0, 0.05, 0.1, \ldots, 0.95\}$.

2. Draw the timing of the next disclosure τ^{n+1}. If the next disclosure does not exist, terminate. Otherwise, repeat this step until the next disclosure does not exist or there are no potential bidders left.

We take 10,000 draws for each τ^n. This gives us the estimated belief $\hat{\mu}_i$ for cases where i discloses.

Next, we estimate the belief of a firm when the firm enters but does not disclose. To obtain this estimate, we again simulate the evolution of disclosure histories. The simulation procedure proceeds as follows:

1. Suppose that a firm observes n disclosures. We fix the firm's entry timing at $\tau \in \{0, 0.05, 0.1, \ldots, 0.95\}$. Also, fix the timing of the latest question $\tau^n \in \{0, \ldots, \tau\}$ ($\tau^0 = 0$ if $n = 0$).

2. Draw the timing of the next disclosure τ^{n+1}. To make a draw, we exploit the following relationship, which we have shown as equation 1.4:

$$\begin{aligned}&\Pr(\tau^{n+1} > t \text{ or no } (n+1)\text{-th disclosure})^{(M-1)/M} \\ &= \mu_i(\tau^{n+1} > t \text{ or no } (n+1)\text{-th disclosure})\end{aligned}$$

   where M is the number of potential entrants who have not disclosed at τ, including firm i. If there are no disclosures beyond, terminate.

3. We draw the next disclosure timing until the draw suggests that next disclosure does not exist. Here , we exploit:

$$\begin{aligned}&\Pr(\tau^{n+k} > t \text{ or no } (n+k)\text{-th discl.} \mid (n+k-1)\text{-th discl. is at } \tau^{n+k-1})^{(M-k)/(M-k+1)} \\ &= \mu_i(\tau^{n+k} > t \text{ or no } (n+k)\text{-th discl.} \mid (n+k-1)\text{-th discl. is at } \tau^{n+k-1}).\end{aligned}$$

We take 10,000 draws for each tuple (τ, n, τ^n). This gives us the estimated belief $\hat{\mu}_i$ for cases where i does not disclose.

When we need to evaluate $\hat{\mu}_i$ when timings are not on the grid $\{0, 0.05, \ldots, 1\}$, we make a linear interpolation.

Step 3. Value of disclosure In this step, we aim to obtain an estimate for the value of disclosures. First, we estimate the value from bidding conditional on time-T history $V_i(h^T, c)$ by:

$$\hat{V}_i(h^T, c) = \max_b (b - c)(1 - \hat{G}_{-i}(b|h^T)),$$

where $\hat{G}_{-i}$ is the estimated CDF of the lowest bid from opponents. We estimate this object for all h^T and for all values of estimated construction costs. For h^T, we include the number of disclosures from opponents, a dummy for whether i disclosed, and i's disclosure timing. we evaluate $\hat{V}_i$ on a grid of disclosure timings: $\{0, 0.05, \ldots, 1\}$. A linear interpolation is made when necessary.

Next, we estimate the values with and without disclosure, $v^{1,\tau}(h^\tau, c)$ and $v^{0,\tau}(h^\tau, c)$. This value is estimated by:

$$\hat{v}^{j,\tau}(h^\tau, c) = \int \hat{V}_i(h^T, c)\ \hat{\mu}_i(h^T | h^\tau, \tau_i = \tau, a_i^Q = j)\ dh^T$$

for $j = 0, 1$. Note that although we have expressed this as an integration, due to our estimation procedure, $\hat{\mu}_i$ is a discrete distribution. Therefore, this calculation turns out to be a summation, in practice. We estimate these objects for all h^τ and for all values of estimated construction costs. For h^τ, we include the number of disclosures from opponents and the most recent disclosure timing. The values are evaluated on a grid of time τ: $\{0, 0.05, \ldots, 1\}$. A linear interpolation is made when necessary.

Then, value of disclosure $\Delta v^\tau(h^\tau, c)$ can be estimated as:

$$\widehat{\Delta v}^\tau(h^\tau, c) = \hat{v}^{1,\tau}(h^\tau, c) - \hat{v}^{0,\tau}(h^\tau, c).$$

Step 4. Model primitives The model primitives are estimated via maximum likelihood as described in the main text. We calculate standard errors via bootstrapping. Our final procedure bootstraps over the entire estimation procedure to incorporate estimation error in earlier steps.

A.4 Simulating the Status Quo and Model Fit

We simulate the auction outcome under status quo using our estimates to consider counterfactuals and evaluate the model fit. The procedure is as follows:

1. For each potential bidder i, we draw a tuple $(\tau_i, c_i^E, c_i^Q, c_i)$ – arrival timing, entry cost, disclosure cost, and construction cost–. Moreover, we make draws for whether i considers entry and whether i is forced to disclose.

2. Starting from the firm with the earliest arrival timing, we determine their actions on entry and disclosure. For entry, we use the estimated value of entry $\hat{v}^{\tau_i}(h^{\tau_i})$ to assign the entry action. For disclosures, we use the estimated value of disclosure $\widehat{\Delta v}^{\tau_i}(h^{\tau_i}, c_i)$.

3. Repeat the previous step until all the actions are determined for the firm with the latest arrival timing.

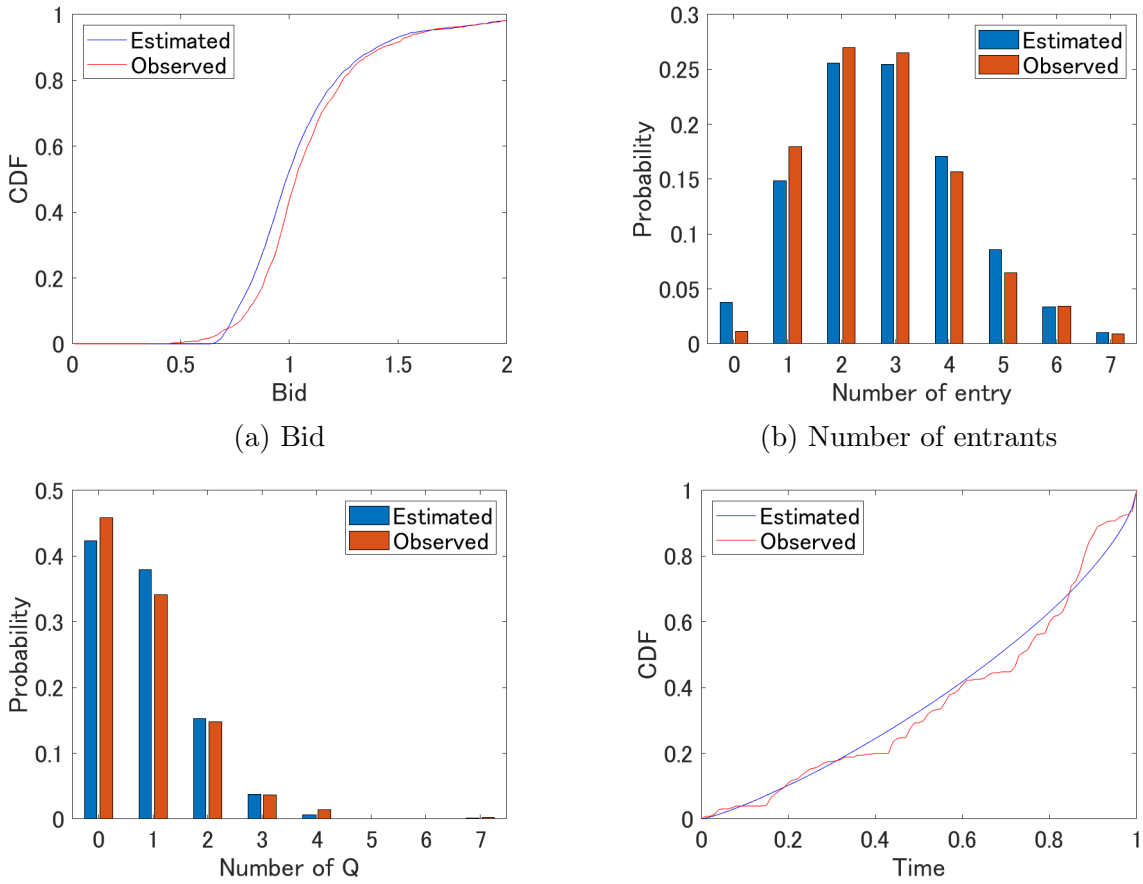
4. Now that the time-T history h^T is determined, we solve:

$$\underset{b}{\text{argmax}}(b - c_i)(1 - \hat{G}_{-i}(b|h^T))$$

   and obtain b_i for each entrant.

We repeat this procedure for 10,000 times.

Figure A.1 shows the observed and simulated outcomes from the auction. The model does relatively well fitting the overall shape of the actual bid distribution, number of entrants, number of disclosures, and timing of disclosures.



(a) Bid

(b) Number of entrants

(c) Number of disclosures

(d) Timing of disclosures

Figure A.1: Model Fit

A.5 Appendix Tables

Table A.1: Secret reserve price

Distribution: Log-normal		
Parameters	Estimate	S.E.
μ	0.738	0.079
σ	0.261	0.057

Note: Table shows parameters for the distribution of the secret reserve price.

Table A.2: History Transition

Variables	Estimate	S.E.
μ_t		
Constant	0.720	0.358
# Q posted (n)	-0.214	0.141
ln(# Pot. Bidders not posted Q yet +1)	0.099	0.136
Time of previous Q (τ^n)	-0.873	0.164
Type:		
Overlay	-0.222	0.081
Safety	-0.054	0.099
Bridge	-0.291	0.100
Recons	-0.655	0.135
Others	Reference	
District:		
Missoula	Reference	
Butte	0.060	0.071
Great Falls	0.133	0.068
Glendive	0.079	0.082
Billings	0.130	0.089
log σ_t		
Constant	0.447	0.248
# Q posted (n)	0.237	0.433
ln(# Pot. Bidders not posted Q yet +1)	-0.407	0.131

Note: This table presents estimated parameters $(\beta^{\mu_t}, \sigma^{\sigma_t})$ that relates to the distribution of $(n+1)$-th disclosure, conditional on the n-th disclosure. The specification is described in Step 2 from Section 1.6.

Appendix B

Appendix to Chapter 2

B.1 Proofs

We use BNE, BCE, MPE, and MCE for Bayes Nash equilibrium, Bayes correlated equilibrium, Markov perfect equilibrium, and Markov correlated equilibrium, respectively, for brevity in the following proofs.

Proof of Lemma 1

Proof. ($\Rightarrow$) Let β be a MPE of (G, S). β induces (G^β, S). The BNE best-response condition is directly implied by the MPE condition of (G, S).

($\Leftarrow$) Let β be a BNE of (G^β, S). The BNE best-response condition implies that there is no profitable one-shot deviation given that each player expects all players to follow the prescription of β in the future. The absence of profitable one-shot deviation implies that β is a subgame perfect equilibrium of (G, S) by the one-shot deviation principle. Therefore, β is a MPE of (G, S). □

Proof of Lemma 2

Proof. The statement follows by applying the one-shot deviation principle. □

Proof of Lemma 3

Proof. It is enough to show that $V_i^\beta(x) = V_i^\sigma(x)$ for all $i \in \mathcal{I}$ and $x \in \mathcal{X}$ because then it follows that $v_i^\beta(a, x, \varepsilon) = v_i^\sigma(a, x, \varepsilon)$ for all $i \in \mathcal{I}$, $a \in \mathcal{A}$, $x \in \mathcal{X}$, and $\varepsilon \in \mathcal{E}$.

Recall that V_i^β and V_i^σ satisfy (2.4) for (G, S^*) and (2.5) for (G, S), respectively. Since β induces σ, (2.4) for (G, S^*) is equivalent to

$$\begin{aligned} V_i^\beta(x) &= \sum_{\varepsilon,\tau,\tilde{\tau},a} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\lambda_{\tilde{\tau}|x,\varepsilon,\tau}\beta_{a|x,\tau,\tilde{\tau}}\left(u_i(a,x,\varepsilon) + \delta\sum_{x'} V_i^\beta(x') f_{x'|a,x,\varepsilon}\right) \\ &= \sum_{\varepsilon,\tau,a} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}\sigma_{a|x,\varepsilon,\tau}\left(u_i(a,x,\varepsilon) + \delta\sum_{x'} V_i^\beta(x') f_{x'|a,x,\varepsilon}\right), \end{aligned}$$

where $\beta_{a|x,\tau,\tilde{\tau}} \equiv \prod_{i=1}^I \beta_i(a_i|x,\tau_i,\tilde{\tau}_i)$. Comparison of the above equation for V_i^β and (2.5) for (G, S) that defines V_i^σ implies that $V_i^\beta = V_i^\sigma$. □

Proof of Theorem 1

Proof. To prove the statement of the theorem, we make use of Theorem 1 of Bergemann and Morris (2016) (BM), which we restate below.

Lemma 4 (BM Theorem 1)**.** *Let (G, S) be a static game of incomplete information. A decision rule σ is a BCE of (G, S) if and only if, for some expansion S^* of S, there is a BNE of (G, S^*) that induces σ.*

($\subseteq$) Suppose σ is a MCE of (G, S). We want to show that there exists an expansion S^* and a strategy profile β in (G, S^*) such that β is a MPE of (G, S^*) and β induces σ. Since σ is a MCE of (G, S), it is a BCE of (G^σ, S). By BM Theorem 1, there exists an expansion S^* of S and a strategy profile β of (G^σ, S^*) such that β is a BNE of (G^σ, S^*) and β induces σ. But since $G^\sigma = G^\beta$, β is also a BNE of $\left(G^\beta, S^*\right)$, which in turn implies that β is a MPE of (G, S^*).

($\supseteq$) Suppose β is a MPE of (G, S^*). We want to show that if β induces σ in (G, S), then σ is a MCE of (G, S). Since β is a MPE of (G, S^*), it is a BNE of $\left(G^\beta, S^*\right)$. By BM Theorem 1, if σ is induced by β, then σ is a BCE of $\left(G^\beta, S\right)$. But since $G^\beta = G^\sigma$, σ is also a BCE of (G^σ, S), which in turn implies that σ is a MCE of (G, S). □

Proof of Corollary 1

Proof. ($\subseteq$) Take $\phi \in \mathcal{P}_{a|x}^{MCE}(G, S)$. By definition, there exists $q \in \mathcal{P}_{a|x,\varepsilon,\tau}^{MCE}(G, S)$ such that q induces ϕ, i.e.,$\phi_{a|x} = \sum_{\varepsilon\in\mathcal{E},\tau\in\mathcal{T}} \psi_{\varepsilon|x}\pi_{\tau|x,\varepsilon}q_{a|x,\varepsilon,\tau}$, $\forall a \in \mathcal{A}, x \in \mathcal{X}$. Then there exists S^* such that $S^* \succsim_E S$ and $q \in \mathcal{P}_{a|x,\varepsilon,\tau}^{MPE}(G, S^*)$, implying that $\phi \in \mathcal{P}_{a|x}^{MPE}(G, S^*)$.

($\supseteq$) Take $\phi \in \bigcup_{S^*\succsim_E S} \mathcal{P}_{a|x}^{MPE}(G, S^*)$. Then there exists S^* such that $S^* \succsim_E S$ and $\phi \in \mathcal{P}_{a|x}^{MPE}(G, S^*)$, which implies that there exists $q \in \mathcal{P}_{a|x,\varepsilon,\tau}^{MPE}(G, S^*)$ such that q induces ϕ. Then since $q \in \mathcal{P}_{a|x,\varepsilon,\tau}^{MCE}(G, S)$, we have $q \in \mathcal{P}_{a|x}^{MCE}(G, S)$. □

Proof of Theorem 2

Proof. Let $\phi \in \mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G, S^1)$. We want to show $\phi \in \mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G, S^2)$. From our Theorem 1, we have

$$\mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G, S^1) = \bigcup_{S^* \succsim_E S^1} \mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G, S^*) \text{ and } \mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G, S^2) = \bigcup_{S^{**} \succsim_E S^2} \mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G, S^{**}).$$

If $\phi \in \mathcal{P}^{MCE}_{a|x,\varepsilon,\tau}(G, S^1)$, then there exists S^* such that $\phi \in \mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G, S^*)$. But since $S^* \succsim_E S^1$ and $S^1 \succsim_E S^2$, we have $S^* \succsim_E S^2$, which implies $\phi \in \bigcup_{S^{**} \succsim_E S^2} \mathcal{P}^{MPE}_{a|x,\varepsilon,\tau}(G, S^{**})$, which is what we wanted to show. □

Proof of Theorem 3

Proof. We want to show that $\Theta^{MCE}_I(S) = \bigcup_{\tilde{S} \succsim_E S} \Theta^{MPE}_I(\tilde{S})$. Take $\theta \in \Theta^{MCE}_I(S)$. By definition, $\phi \in \mathcal{P}^{MCE}_{a|x}(G^\theta, S)$. By Corollary 1, there exists some $S^* \succsim_E S$ such that $\phi \in \mathcal{P}^{MPE}_{a|x}(G, S^*)$. But this immediately implies $\theta \in \Theta^{MPE}_I(S^*)$, and thus $\theta \in \bigcup_{\tilde{S} \succsim_E S} \Theta^{MPE}_I(\tilde{S})$. The proof of converse follows the above steps in reverse. □

Proof of Theorem 4

Proof. The statement follows from Theorem 2. □

Proof of Theorem 5

Proof. The statement directly follows from the fact that the constraints (2.8), (2.9), (2.10), and (2.11) characterize the MCE identification conditions as implied by the definition of identified sets. Note that the ex-ante value function equation (2.11) is equivalent to (2.5) because

$$\begin{aligned}\sum_{\varepsilon,\tau,a} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \sigma_{a|x,\varepsilon,\tau} \sum_{x'} V_{i,x'} f_{x'|a,x} &= \sum_{a,x'} \left(\sum_{\varepsilon,\tau} \psi_{\varepsilon|x} \pi_{\tau|x,\varepsilon} \sigma_{a|x,\varepsilon,\tau} \right) V_{i,x'} f_{x'|a,x} \\ &= \sum_{a,x'} \phi_{a|x} V_{i,x'} f_{x'|a,x}.\end{aligned}$$

□

Proof of Theorem 6

Proof. The ex-ante value function (2.14) can be derived by expanding (2.5) and imposing (2.12) similarly to the proof of Theorem 5 in Appendix B.1. The obedience condition (2.13)

is derived as follows. The obedience condition in the fully robust case is

$$\sum_{\varepsilon,a_{-i}} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon}\partial v_i^\theta(a_i', a_i, x, \varepsilon_i) \leq 0. \tag{B.1}$$

But since

$$\begin{aligned}
&\sum_{\varepsilon,a_{-i}} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon} v_i^\theta(a_i', a_{-i}, x, \varepsilon_i) \\
&= \sum_{\varepsilon,a_{-i}} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon} \left(u_i^\theta(a_i', a_{-i}, x, \varepsilon_i) + \delta \sum_{x'} V_{i,x'} f_{x'|a_i',a_{-i},x} \right) \\
&= \sum_{\varepsilon,a_{-i}} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon} u_i^\theta(a_i', a_{-i}, x, \varepsilon_i) + \sum_{a_{-i}} \left(\sum_{\varepsilon} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon} \right) \left(\delta \sum_{x'} V_{i,x'} f_{x'|a_i',a_{-i},x} \right) \\
&= \sum_{\varepsilon,a_{-i}} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon} u_i^\theta(a_i', a_{-i}, x, \varepsilon_i) + \delta \sum_{a_{-i},x'} \phi_{a|x} V_{i,x'} f_{x'|a_i',a_{-i},x},
\end{aligned}$$

where the last equation is obtained by imposing (2.12), the obedience condition (B.1) can be replaced by

$$\sum_{\varepsilon,a_{-i}} \psi_{\varepsilon|x}\sigma_{a|x,\varepsilon}\partial u_i^\theta(a_i', a, x, \varepsilon_i) + \delta \sum_{a_{-i},x'} \phi_{a|x} V_{i,x'} \partial f_{x'|a_i',a,x} \leq 0, \quad \forall i, x, a_i, a_i'$$

where $\partial u_i^\theta(a_i', a, x, \varepsilon_i) \equiv u_i^\theta(a_i', a_{-i}, x, \varepsilon_i) - u_i^\theta(a, x, \varepsilon_i)$ and $\partial f_{x'|a_i',a,x} \equiv f_{x'|a_i',a_{-i},x} - f_{x'|a,x}$. □

Proof of Theorem 7

Proof. By construction, $\Pr\left(\Theta_I(\phi) \subseteq \bigcup_{\tilde{\phi}\in\Phi_\alpha} \Theta_I(\tilde{\phi})\right) \geq \Pr\left(\phi \in \Phi_\alpha\right)$. By taking the limit inferior on both sides, we obtain the desired result. The remaining statement holds by the definitions of $\Theta_I(\phi)$ and $\widehat{\Theta}_I^\alpha$. □

B.2 Interpretation of Counterfactual Analysis

As noted in a series of papers (Aguirregabiria and Suzuki, 2014; Kalouptsidi, Scott and Souza-Rodrigues, 2017, 2021; Aguirregabiria, Collard-Wexler and Ryan, 2021), normalizing the flow payoff from the outside option to zero is not without loss of generality for counterfactual analysis. However, we provide a way to interpret our normalization and counterfactual analysis under the assumption that a parametric specification is correct.

Suppose that the flow payoff is given by

$$u_i^\theta(a_{it}, a_{jt}, x_t, \varepsilon_{it}) = \begin{cases} \theta_{i,0}^* + \theta_{i,w}^* w_t + \theta_{i,ec}^*(1 - z_{it}) + \theta_{i,ce}^* a_{jt} + \varepsilon_{it} & \text{if } a_{it} = 1 \\ \theta_{i,sv}^* z_{it} & \text{if } a_{it} = 0 \end{cases}, \tag{B.2}$$

where $\theta^*_{i,ec} < 0$ is the true entry cost and $\theta^*_{i,sv} > 0$ is the true scrap value. A critical assumption in (B.2) is that the scrap value is independent of the market characteristics. Payoff (B.2) implies that the payoff difference is

$$u_i^\theta(1, a_{jt}, x_t, \varepsilon_{it}) - u_i^\theta(0, a_{jt}, x_t, \varepsilon_{it}) = \\ \left(\theta^*_{i,0} - \theta^*_{i,sv}\right) + \theta^*_{i,w} w_t + \left(\theta^*_{i,ec} + \theta^*_{i,sv}\right)(1 - z_{it}) + \theta^*_{i,ce} a_{jt} + \varepsilon_{it}.$$

Thus, the estimated intercept captures $(\theta^*_{i,0} - \theta^*_{i,sv})$ (i.e., fixed cost net of scrap value), and the estimated entry cost captures $(\theta^*_{i,ec} + \theta^*_{i,sv})$ (i.e., entry cost net of scrap value). This is consistent with Aguirregabiria and Suzuki (2014) that the fixed cost, entry cost, and scrap value are not separately identified.

Now consider the counterfactual experiment of reducing the entry cost as in Section 2.6. One can show that this is equivalent to holding $\theta^*_{i,sv}$ fixed and reducing $\theta^*_{i,ec}$ (making it less negative) to $\theta^{**}_{i,ec}$ by

$$\theta^{**}_{i,ec} = \theta^*_{i,ec} - \lambda \underbrace{(\theta^*_{i,ec} + \theta^*_{i,sv})}_{<0},$$

where λ is the proportion of entry cost intended to be cut. Thus, our counterfactual experiment is equivalent to providing an entry subsidy equal to λ of the entry cost net of scrap value.

Appendix C

Appendix to Chapter 3

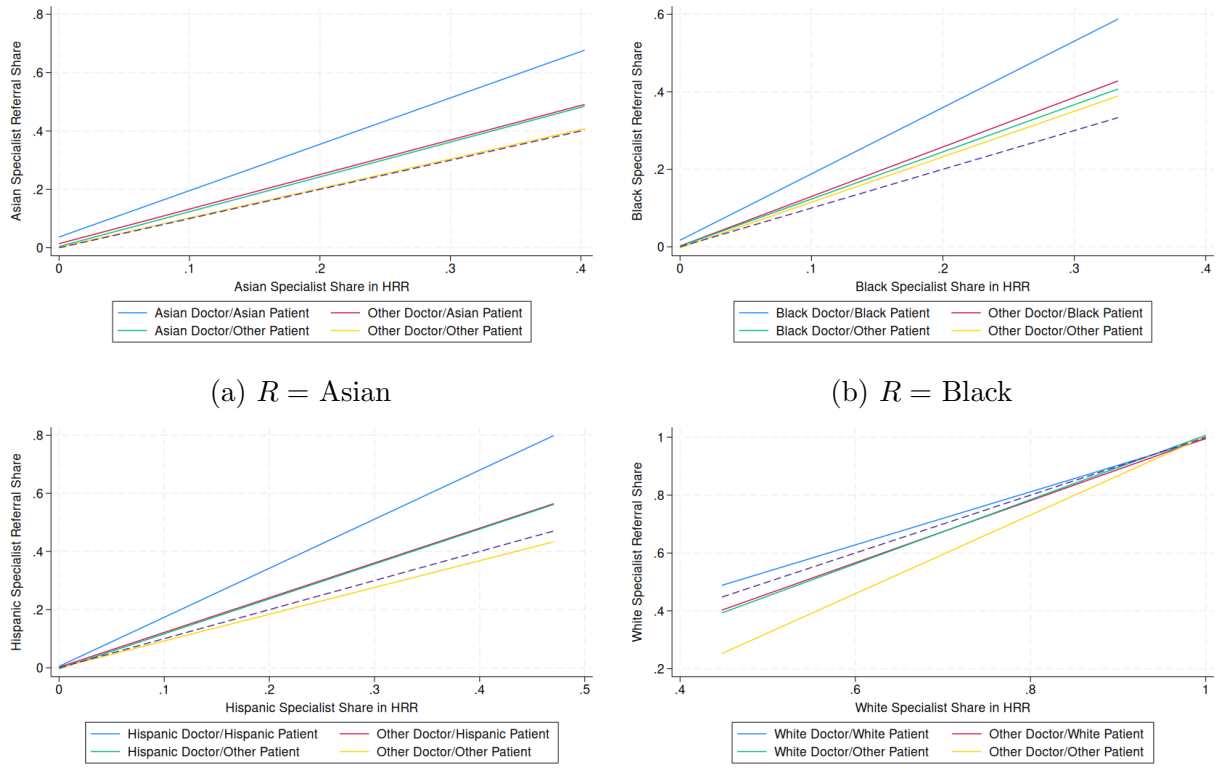
C.1 Other Summary Statistics

Table C.1: Average THA/TKA patient comorbidities (1999–2019)

	Full Sample	Experienced Specialist
N	1,315,864	1,238,534
Alzheimer's Disease	0.017	0.017
Acute Myocardial Infarction	0.007	0.007
Anemia	0.580	0.579
Asthma	0.102	0.102
Atrial Fibrillation	0.111	0.112
Heart Failure	0.150	0.149
Chronic Kidney Disease	0.161	0.163
Cancer	0.127	0.127
Chronic Obstructive Pulmonary Disease	0.132	0.131
Depression	0.203	0.204
Diabetes	0.266	0.266
Hip/Pelvic Fracture	0.025	0.024
Hyperlipidemia	0.627	0.632
Hypertension	0.812	0.812
Ischemic Heart Disease	0.363	0.361
Osteoporosis	0.131	0.131
Rheumatoid Arthritis/Osteoarthritis	0.994	0.994
Stroke/Transient Ischemic Attack	0.034	0.034

Notes: Comorbidities are coded by the Chronic Condition Warehouse (CCW). Cancer combines all site-specific cancers in the CCW (Endometrial, Breast, Colorectal, Lung, and Prostate Cancer).

C.2 Other Preliminary Evidence



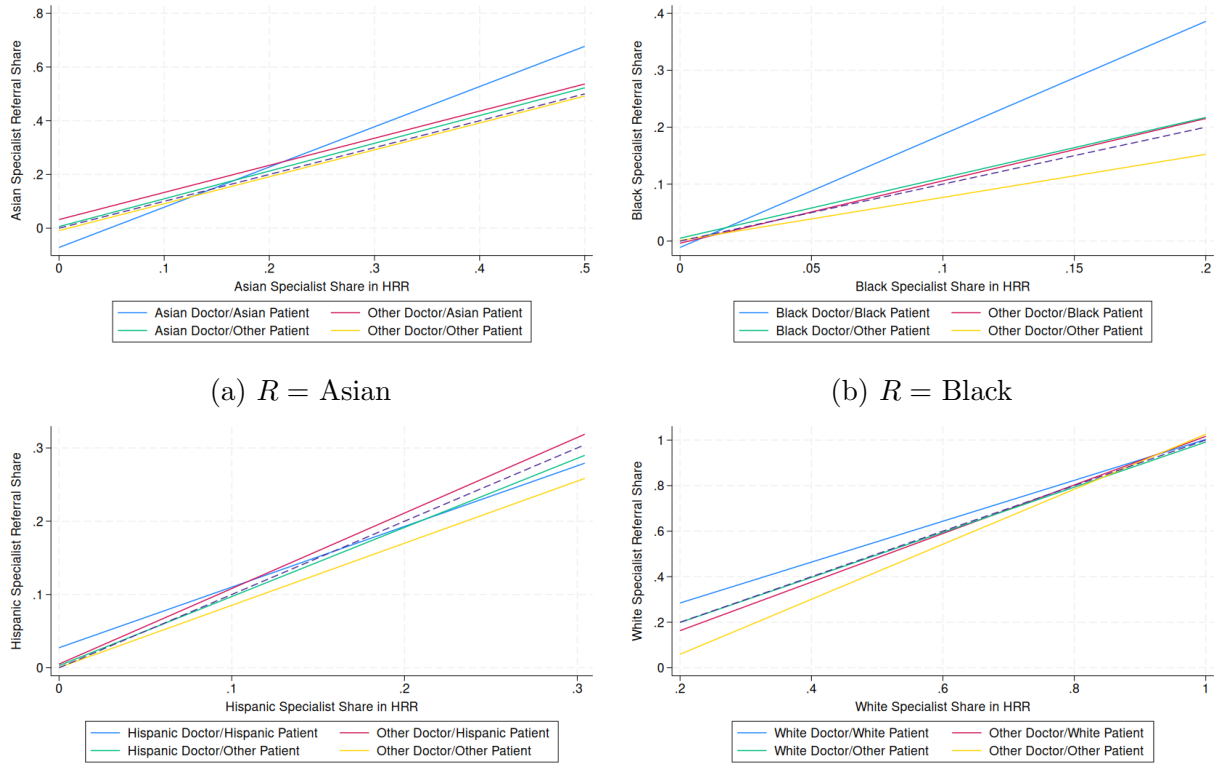
(a) $R =$ Asian

(b) $R =$ Black

(c) $R =$ Hispanic

(d) $R =$ White

Figure C.1: Race R orthopedist referral share against their population share by doctor-patient race group in 2019



(a) R = Asian

(b) R = Black

(c) R = Hispanic

(d) R = White

Figure C.2: Race R neurosurgeon referral share against their population share by doctor-patient race group in 2019

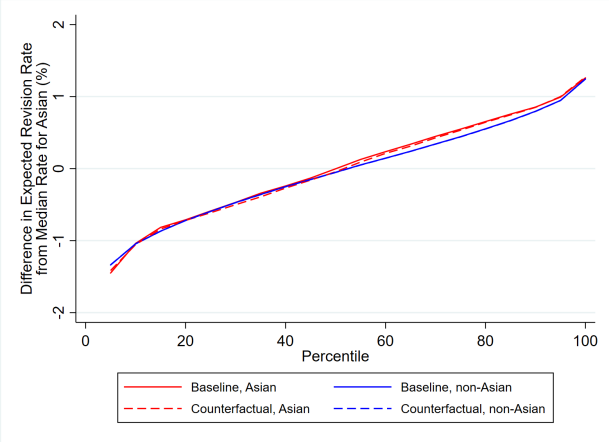
C.3 Other Structural Estimates

Table C.2: Coefficients on patient comorbidities for THA/TKA outcome regressions

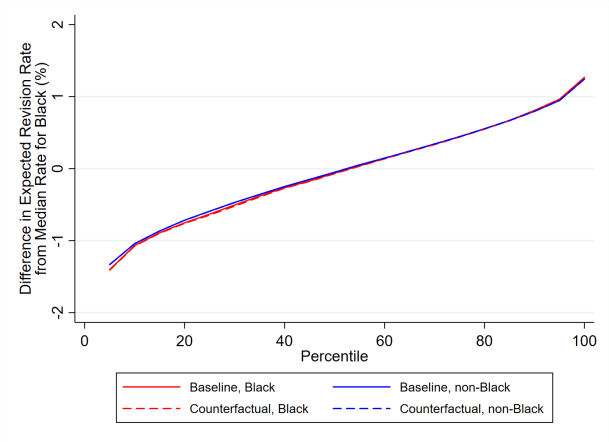
	Outcome: THA/TKA Revision(%)			
	R = Asian (1)	R = Black (2)	R = Hispanic (3)	R = White (4)
Alzheimer's Disease	0.0864	0.0886	0.0885	0.0896
	(0.0895)	(0.0895)	(0.0896)	(0.0895)
Acute Myocardial Infarction	-0.0976	-0.0993	-0.0975	-0.0991
	(0.1150)	(0.1150)	(0.1150)	(0.1151)
Anemia	0.6349***	0.6375***	0.6348***	0.6380***
	(0.0361)	(0.0362)	(0.0360)	(0.0362)
Asthma	0.1232***	0.1257***	0.1232***	0.1255***
	(0.0266)	(0.0266)	(0.0266)	(0.0265)
Atrial Fibrillation	0.1845***	0.1801***	0.1835***	0.1784***
	(0.0348)	(0.0349)	(0.0348)	(0.0348)
Heart Failure	0.2716***	0.2743***	0.2720***	0.2746***
	(0.0355)	(0.0355)	(0.0355)	(0.0355)
Chronic Kidney Disease	0.3629***	0.3660***	0.3626***	0.3653***
	(0.0318)	(0.0316)	(0.0318)	(0.0316)
Cancer	-0.0451*	-0.0442*	-0.0461*	-0.0458*
	(0.0246)	(0.0246)	(0.0246)	(0.0246)
COPD	0.1800***	0.1789***	0.1796***	0.1778***
	(0.0314)	(0.0314)	(0.0315)	(0.0314)
Depression	0.6015***	0.5973***	0.6017***	0.5962***
	(0.0254)	(0.0257)	(0.0255)	(0.0255)
Diabetes	0.0052	0.0091	0.0067	0.0114
	(0.0228)	(0.0226)	(0.0228)	(0.0227)
Hip/Pelvic Fracture	4.472***	4.469***	4.471***	4.468***
	(0.2093)	(0.2091)	(0.2093)	(0.2092)
Hyperlipidemia	-0.0654***	-0.0677***	-0.0659***	-0.0680***
	(0.0159)	(0.0159)	(0.0160)	(0.0160)
Hypertension	0.1680***	0.1725***	0.1681***	0.1723***
	(0.0175)	(0.0177)	(0.0175)	(0.0176)
Ischemic Heart Disease	0.0292	0.0286	0.0291	0.0286
	(0.0227)	(0.0227)	(0.0227)	(0.0227)
Osteoporosis	0.2730***	0.2694***	0.2731***	0.2712***
	(0.0340)	(0.0340)	(0.0340)	(0.0340)
RA/Osteoarthritis	2.678***	2.678***	2.678***	2.678***
	(0.2344)	(0.2343)	(0.2344)	(0.2343)
Stroke/TIA	0.1598**	0.1606**	0.1596**	0.1603**
	(0.0621)	(0.0620)	(0.0620)	(0.0620)

Notes: Standard errors are clustered at the HRR and year level. Comorbidities are coded by the Chronic Condition Warehouse (CCW). Cancer combines all site-specific cancers in the CCW (Endometrial, Breast, Colorectal, Lung, and Prostate Cancer). Abbreviations are COPD for Chronic Obstructive Pulmonary Disease, RA for Rheumatoid Arthritis, and TIA for Transient Ischemic Attack.

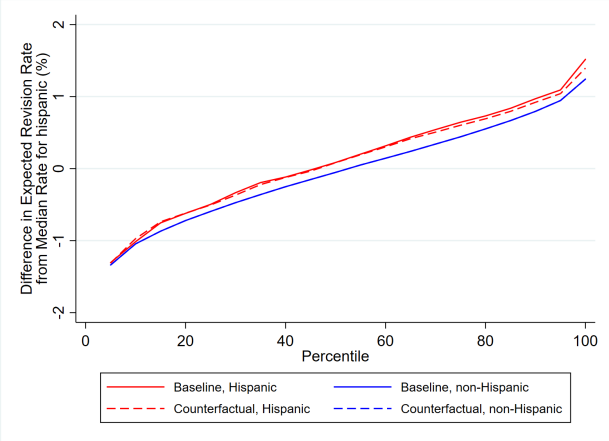
C.4 Other Counterfactual Figures



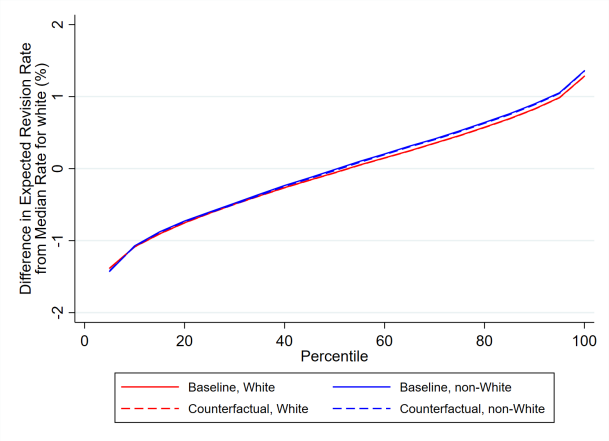
(a) $R =$ Asian



(b) $R =$ Black

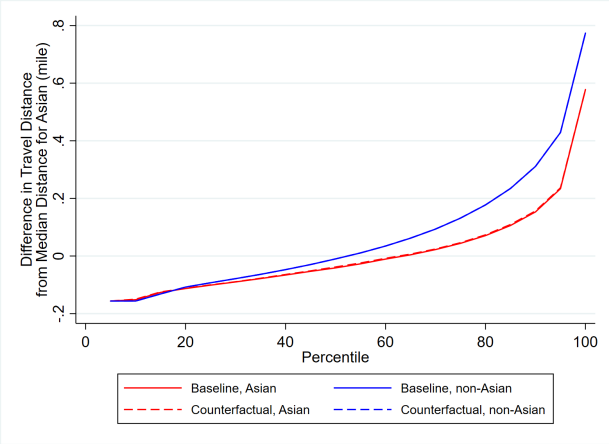


(c) $R =$ Hispanic

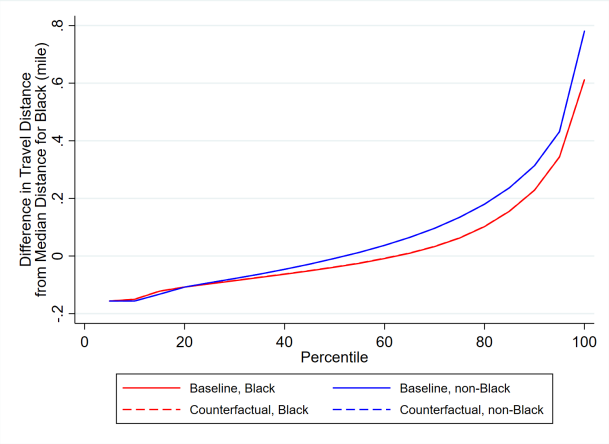


(d) $R =$ White

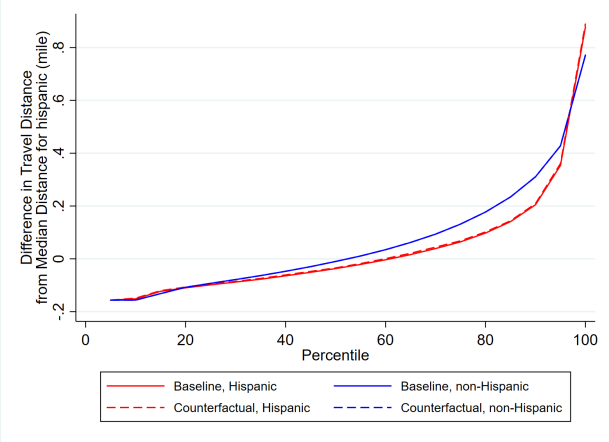
Figure C.3: Distribution of Expected Health Outcome: Baseline and Counterfactual Scenarios



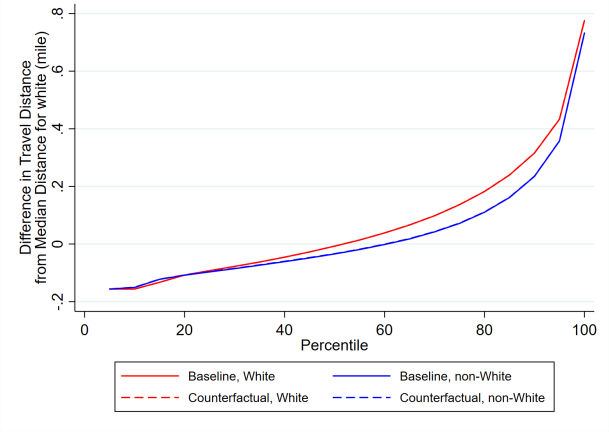
(a) R = Asian



(b) R = Black



(c) R = Hispanic



(d) R = White

Figure C.4: Distribution of Travel Distance: Baseline and Counterfactual Scenarios