

NBER WORKING PAPER SERIES

DOES ENTRY REMEDY COLLUSION?
EVIDENCE FROM THE GENERIC PRESCRIPTION DRUG CARTEL

Amanda Starc
Thomas G. Wollmann

Working Paper 29886
<http://www.nber.org/papers/w29886>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2022, Revised April 2023

We thank John Asker, Emily Cuddy, Kathleen Hui, JF Houde, Michi Igami, Louis Kaplow, Casey Mulligan, Rob Porter, Fiona Scott Morton, Mike Sinkinson, Bob Topel, Brett Wendling, as well as seminar participants at the University of Wisconsin, University of Pennsylvania, University of Texas-Austin, Stanford University, US Dept. of Justice, Penn State University, Chicago Booth, UC Berkeley, UCLA Anderson, Federal Trade Commission, Harvard Business School, Congressional Budget Office, Cornerstone Research, Midwest IO Fest, Cowles Conference on Models and Measurement, SITE, Northwestern Antitrust Conference, and ASSA/AEA Annual Meeting for their comments. We benefited immensely from conversations with Doni Bloomfield, with whom we have related work. Paulo Ramos provided not only superb research assistance but also substantive contributions to the model and its estimation. Paloma Avendano also provided excellent research assistance. Wollmann thanks the Becker Friedman Institute's Industrial Organization Initiative for support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Amanda Starc and Thomas G. Wollmann. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Does Entry Remedy Collusion? Evidence from the Generic Prescription Drug Cartel
Amanda Starc and Thomas G. Wollmann
NBER Working Paper No. 29886
March 2022, Revised April 2023
JEL No. L11,L41,L65

ABSTRACT

Entry represents a fundamental threat to cartels. We study the extent and effect of this behavior in the largest price-fixing case in US history, which involves generic drug manufacturing. We link information on the cartel's internal operations to regulatory filings and market data. There is a substantial increase in entry after cartel formation but regulatory approvals delay most entrants by 2-4 years. We then estimate a structural model to simulate counterfactual equilibria. Absent entry, cartel profits would be dramatically higher. Correspondingly, reducing regulatory delays by just 1-2 years equates to consumer compensating variation of \$559 million-\$1.3 billion.

Amanda Starc
Kellogg School of Management
Northwestern University
2211 Campus Drive
Evanston, IL 60208
and NBER
amanda.starc@kellogg.northwestern.edu

Thomas G. Wollmann
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
Thomas.Wollmann@chicagobooth.edu

1 Introduction

Cartels thwart competition, even in modern economies. The US Department of Justice has prosecuted price-fixing by all three major canned tuna brands, 70 auto parts suppliers, 15 global financial institutions trading foreign currency, over 100 real estate investors bidding in foreclosure auctions, and many others in the last few years alone. However, as cartels raise prices and profits, they may also attract uncooperative entrants, whose efforts to gain market share will undercut the incumbents' agreements. Hence, entry can serve as a fundamental safeguard against sustained collusion. In many markets, though, entry is a slow and expensive process, so the likelihood that entry restores competitive prices depends critically on barriers to entry and the length of delays. We study the extent and effect of entry on cartelized markets in the context of the largest price fixing case in US history, which involves generic prescription drug manufacturers.

Historically, most economists and policymakers have thought of generic drugs as a competition success story. When branded drugs lose patent protection an influx of generics typically follows, capturing more than half of the market at less than half the branded equivalent's price (Scott Morton (1999); Wendling et al. (2011)). In recent years, however, the prices of many generics have risen substantially. According to court documents described below, many of these increases can be traced to a single precipitating event: in 2013, Teva Pharmaceuticals, the largest generic firm, hired NP, a marketing executive with especially strong industry relationships, and tasked her with "price increase implementation."¹ Over an 18-month period, industry participants exchanged thousands of calls and texts—alongside countless LinkedIn, Facebook, and WhatsApp messages and face-to-face conversations—with contacts at rival firms to coordinate the increases (Complaint, page 322).² Following this period, prescription drug expenditures by governments, private insurers, and individuals rose sharply by billions of dollars.

In this paper, we measure the effect of price fixing on market entry, estimate a structural model of generic drug competition, and use our estimates to assess counterfactual behavior and policies. We exploit detailed information on the cartel's internal operations, which were revealed when a complaint was filed in May 2019.³ The complaint presents witness testimonies, private communications within and between rival firms, and internal documents, collected in the course of an extensive government investigation. It includes a list of the drugs (i.e., substance-delivery-release-strength combinations) for which NP fixed prices, the dates on which those prices were fixed, and the criteria NP used to select which prices to fix. It also reports NP's own numerical measure of the strength of her relationships with the sales and marketing executives at competing firms.

Information from the complaint is crucial to our research design, which compares cartelized and uncartelized drugs before and after collusion. Such comparisons typically raise concerns about unobservable differences across markets and time that relate to both the likelihood of collusion and outcomes of interest. However, information in the complaint reveals how NP selected markets, so we can observe the variables on which the selection depends. These variables are both stable over time and independent of the outcomes of interest. Also, the vast majority of the price increases were implemented within about a year of NP's first anniversary of joining Teva, mitigating concerns that they were timed to coincide with unobservable

¹We identify individuals using their initials. Their names are discoverable in court documents but irrelevant to our analysis.

²Throughout the paper, we use the word "cartel" or similar terms to characterize the behavior and economic arrangement, not reflect the findings of the court. The case against Teva, its former employees, and several co-conspirators is ongoing at the time of writing, so from a legal perspective, this is an "alleged cartel" with respect to their involvement. However, the facts presented in the complaint, which we take at face value, satisfy any substantive definition of collusion. See Section 2 for details.

³See <https://portal.ct.gov/AG/Press-Releases/2019-Press-Releases/DRUG-PRICE-FIXING-COMPLAINT-UNSEALED>.

environmental changes. Moreover, observable outcomes in cartelized and uncartelized markets track extremely closely with each other leading up to collusion (but clearly diverge sharply afterwards).

Our sample consists of drugs manufactured by Teva in the quarter just prior to NP joining the firm. It includes all firms' versions of those drugs and spans 2008 to 2019.⁴ We measure quantity using prescriptions, which we obtain from IQVIA and Medicaid, and we measure point-of-sale price in dollars per prescription, which we obtain from Medicaid. Before manufacturing a drug, firms must file an Abbreviated New Drug Application (ANDA) and have it approved by the Food and Drug Administration (FDA). ANDA filings closely correspond to entry. We obtain the filings and their respective dates from the FDA.⁵

We find four key patterns in the data. First, prices rise sharply following cartel formation. Compared to uncartelized markets, there are price increases that average about 50% in cartelized markets. Second, cartel formation is followed by significantly more entry, as measured by ANDA filings. Third, regulation introduces delays of 2-4 years. As a result, a firm that filed an ANDA as early as 2013 might not be authorized to enter before, say, 2017. Fourth, the data suggests that entry by nonmembers exerts downward pressure on drug prices.

These facts inform our structural model, which emphasizes two main decisions that generic firms face. In the first stage, firms *without* regulatory approval file ANDAs if the sum of expected discounted profits associated with entering exceeds the sunk costs, given their expectations about future pricing behavior. Cartel members do not coordinate first stage decisions (even though they may cooperate in the second stage).⁶ Entry is not immediate, since the process of obtaining ANDA approval involves significant, stochastic delays.⁷

In the second stage, firms *with* regulatory approval set prices. In uncartelized markets, and in cartelized markets prior to cartel formation, firms set prices that maximize individual profits. A cartel forms when NP reaches an agreement to set prices that maximize the joint profits of the members, in which case nonmembers set prices that best respond. We assume the agreement is supported by trigger strategies, so realized prices depend on the history of play as well as incentive compatibility constraints. The model highlights how cartel formation raises prices, which in turn increases incentives to enter. Moreover, it emphasizes that entry exerts downward pressure on price, given that not all entrants are cartel members.

Next, we take the model to the data. We estimate demand, use these estimates to recover marginal costs, and combine these results to forecast the profits that firms would earn for any hypothetical market structure and conduct. We use these forecasts in conjunction with the empirical distribution of entry delays to compute the value of entry, which equals the sum of discounted expected future profits. Firms enter if and only if the value of entry exceeds the sunk cost of doing so, so observed entry decisions map to the parameters that determine those costs. Since two types of firms (i.e., cartel members and nonmembers) make entry decisions, multiplicity is the rule rather than the exception. Thus, we rely only on the necessary

⁴For example, Teva manufactures pravastatin 100mg immediate release tablets in 2013 Q1, so this drug is in our sample.

⁵All approval dates are reported by the agency, but only a subset of filing dates are. However, ANDA numbers are issued approximately sequentially, so filing dates can be inferred without meaningful error. See the Online Appendix A for a detailed description of the process.

⁶Fershtman and Muller (1986) call this "semicollusive" behavior: long-run choices are competitively chosen but short-run choices may be cooperatively decided. As for why NP could not restrict entry by members, our conversations with industry participants suggest it entailed too much risk. At some firms, entry and pricing decisions are approved by different managers, meaning that restricting entry would require many more individuals be complicit in cartelization. For an illuminating case of strict market segmentation, see Clark et al. (2018). Note, though, that the asphalt cartel they study resorted to violence, which is presumably infeasible in our setting.

⁷See Section 2 for more detail.

conditions implied by Nash equilibrium, bounding the parameters of interest using moment *inequalities* (Tamer, 2003; Pakes et al., 2015). These methods yield parameter estimates that imply accurate profit margins and sensible sunk costs of entry when compared to outside sources, such as audited financial statements or remarks by former FDA officials.

Our estimates reflect assumptions about conduct. To ensure the assumptions are consistent with the individual incentives of the cartel members, we introduce a dynamic games framework. Following Igami and Sugaya (2022), we measure each member's incentives to comply with the cartel agreement, plot the lower envelope over time across market structures and draws of the disturbances, and find that trigger strategies form an equilibrium. We then test whether the observed prices are consistent with our conduct assumptions, following Backus, Conlon, and Sinkinson (2021). Most notably, we find that collusive prices persist long after the government's investigation and ongoing prosecution, presumably supported by tacit agreements.⁸ The data also rejects a model in which nonmembers price cooperatively in favor of a model in which they best respond to cartel behavior, which means entrants can exert downward pressure on prices.

With estimates of the parameters governing demand, prices, and entry in hand, we simulate equilibria under alternative assumptions and policies. First, to assess the importance of observed entry on cartelized markets, we counterfactually prohibit firms from entering markets in which cartels have formed and recompute second stage outcomes. Our most striking finding is that observed entry reduces profits earned by incumbent members by three-quarters, reflecting lower prices as well as smaller market shares.

Second, to inform policymakers, we counterfactually reduce the regulatory costs and delays by up to \$600,000 and two years, respectively. These alternative policies are logical and important for several reasons. First, generic production requires low and expensive drug-specific government approvals. Firms must prepare and file an ANDA with the FDA, and the approval process can take years and cost millions of dollars. Second, these factors are within the government's control: the FDA has varied fees considerably over the past decade and experimented with various expedited approval programs. Former commissioner Scott Gottlieb has lamented both the cost and time associated with entry (Gottlieb (2016)). Third and perhaps most important, regulatory costs and delays are basic features of many large industries. Examples range from air travel and power generation to defense contracting and retail banking. Hence, our broad conclusions are likely to be relevant beyond prescription drugs. We find that reducing entry costs leads to significant entry but relatively modest consumer benefits, which are roughly offset by lost profits. In contrast, we find that reducing delays has a comparatively large effect on consumer surplus—the consumer compensating variation exceeds \$1.3 billion.⁹

Our findings show that entry can play a key role in disciplining cartels, with a first order impact on consumer welfare. Nonetheless, the discipline it exerts may be incomplete and slow, due to high costs and long delays, many of which can be attributed to regulation.¹⁰ We contribute to the growing body of empirical work on cartels by incorporating equilibrium entry behavior and post-entry pricing. While

⁸Ongoing tacit collusion is consistent with statements by NP that price fixing in the generic drug industry is hard to initiate but easy to sustain.

⁹These gains can be compared to potential costs of relaxing regulations. For generic drugs, we note that the active ingredients have already been deemed "safe and effective" by the FDA. Indeed, policymakers have moved to streamline the approval process with little concern about safety. However, for additional analysis of these tradeoffs, see Grennan and Town (2020).

¹⁰We do not in any way mean to imply that the time and expense associated with generic drug evaluations are wasteful, only that limiting fees and hastening approvals generates significant surplus for buyers when incumbent firms are playing cooperatively. Lower fees may draw resources away from other oversight activities, while quicker approvals may require additional staff (or, again, result in lax enforcement). Forecasting the costs associated with these changes is far beyond the scope of the present paper, which focuses on the relationships between cartel formation, entry, pricing, and purchases in equilibrium.

foundational work by Stigler (1964) names both cheating and entry as threats to collusive agreements, most early and influential formalizations of the cartel problem (e.g., Green and Porter (1984)) study the former threat but rule the latter one out. Theoretically, Harrington (1984) illustrates how the profitability of entry depends on post-entry pricing behavior.¹¹ Fershtman and Pakes (2000) allow for entry by way of simulations, numerically solving Markov perfect equilibria that permit entry, exit, and quality differences across firms. The authors report a strong industry relationship between collusion and subsequent entry—a relationship they state "has largely been ignored in the literature and has important implications for the welfare analysis of collusive behavior."¹²

Empirically, Levenstein and Suslow (2006) note that entry can "undermine the best-laid collusive plans."¹³ Collusion is often difficult to sustain, absent barriers to entry (Levenstein, 1995; Symeonidis, 2003; Vasconcelos, 2004), cartels may take steps to limit entry (Levenstein and Suslow, 2011; Harrington et al., 2018), and yet may need or choose to accommodate some entrants (Scott Morton, 1997; Podolny and Scott Morton, 1999). In many recent empirical papers, authors have carefully chosen markets where it is reasonable to assume that entry is either unlikely to occur in the near-term (Miller and Weinberg (2017)) or that entry is determined by factors that are "outside" the model (Porter (1983), Byrne and de Roos (2019), Igami and Sugaya (2022)).¹⁴ We model endogenous entry and subsequent pricing, which allows us to estimate the sunk costs and perform counterfactuals. To do so, we draw on work by Borenstein (1989), Bresnahan and Reiss (1991), Mazzeo (2002), Seim (2006), and the large, growing literature that followed it.

This paper also improves our understanding of prescription drug pricing (see, for example, Berndt et al. (2018)). Sco (2000) shows that larger revenue markets, markets with more hospital sales, and products that treat chronic conditions attract more entry. At the same time, Ganapati and McKibbin (2021) show that entry—a traditional mechanism for reducing market power—has been limited in the United States in recent years. Mulligan (2021) argues the problem is burdensome regulation. We expand this literature by exploring the strategic incentives faced by generic manufacturers and the impact of these incentives on consumers. Cuddy (2020) models this market as a series of simultaneous procurement auctions, in order to study the equilibrium relationship between competition and prices. She estimates large damages from cartel behavior using a counterfactual auction model of competitive behavior. Similarly, Clark et al. (2021) estimate large damages using a reduced form approach. We focus on the market forces that may serve to alleviate these harms.

The paper is organized as follows. Sections 2 and 3 describe the institutional setting and data,

¹¹While incumbent firms may be unable to maintain collusion under the threat of entry when the anticipated response is discontinuing cooperation or accommodating new firms, if potential entrants expect the incumbent firms to respond aggressively, incumbents may be able to sustain collusion even absent high entry costs (Harrington, 1989b).

¹²In particular, Fershtman and Pakes (2000) report, "The contrast between the entry states in the model that allows for collusion and the entry states in the model that does not is quite striking. The [relationships between entry and the state variables] simply *disappear* when collusion is not allowed" (emphasis in original). De Roos (2006) notes that the dynamics of entry and that characterize the lysine cartel are not adequately addressed in the literature. Fershtman and Pakes (2000) emphasize the universally positive effect collusion has on entry; however their statement contains a slight but especially interesting caveat due to work by Asker (2010), who studies bidding rings among stamp deals and shows how their arrangement induces *overbidding*, which deters entry.

¹³Entry can undermine a cartel by, e.g., reducing the amount by which collusion increases profit, which threatens stability.

¹⁴For example, Igami and Sugaya (2022) study vitamin cartels in the 1990s. In their setting, competition unexpectedly arose from technology change, which benefited fringe suppliers. Igami (2015) also considers the impact of fringe Vietnamese coffee producers. We build on their approach by modeling the sunk cost of entry. Elsewhere, Miller and Weinberg (2017) study the effect of the Miller-Coors joint venture on the US beer market. In their setting, the emergence of another major brewer is very unlikely, at least in the near future, given that the leading brands have accumulated brand equity over decades of advertising which insulates them from startup competition. Moreover, while the "craft" segment has grown over the past three decades, their Table I shows the top two firms have nearly two-thirds of nationwide share. Entry barriers are not necessary, however, to restrict entry. For instance, Harrington (1989a) shows that sufficiently patient incumbents can credibly threaten entrants.

respectively. Section 4 reports patterns in the data that motivate our structural model. Sections 5 and 6 describe and estimate the structural model. Section 7 measures incentives to collude and tests the assumptions we maintain. Section 8 simulates counterfactual outcomes under alternative assumptions and policies, and Section 9 concludes.

2 Industry setting and cartel operations

2.1 The US generic prescription drug industry

Generic drugs are a competition success story. When a branded drug loses exclusivity, generic entry drives prices down towards marginal cost. Generally, the first generic entrant will price its product slightly lower than the branded drug, and the second generic entrant will reduce the price to approximately 50% of the branded drug price. Conditional on having a large number of entrants, prices fall to around 20% of the price of the branded counterpart (Scott Morton (1999)). For this reason, the market is the "most dynamic and cost-effective in the world" (Scott Morton and Boller (2017)). Historically, barriers to entry have been relatively low. Under the Hatch-Waxman Act, firms can enter by filing an ANDA that shows that the active ingredient, delivery mechanism, strength, and dose of the generic drug are the same as the branded drug. The generic drug must be "bioequivalent" to the branded drug.

Generic drug manufacturers compete with each other to sell their products to wholesalers, distributors, and in some cases, directly to retail pharmacy chains, mail-order and specialty pharmacies, hospital chains, and some health plans. Due to complex "cost-plus" reimbursement rules, higher wholesale prices may weakly benefit these market participants.¹⁵ In recent years, wholesale prices for some generic molecules have increased substantially. The increases have been attributed to both supply shortages and anti-competitive behavior (Cuddy, 2020).

2.2 The cartel

The cartel we study traces back to a single change in industry leadership: on April 22, 2013, NP joined Teva Pharmaceuticals.¹⁶ This event was special for several reasons. First, in the years leading up to the cartel's formation, NP forged uniquely strong relationships. She worked as the Director of Global Generic Sourcing for Amerisource Bergen (ABC), one of the three major US drug distributors. The role led to "routine interaction with representatives from every major generic drug manufacturer" (Complaint, page 158). Second, Teva was—and still is—the world's largest generic manufacturer. By early 2013, for example, it produced about one out of every three generic tablets and capsules.¹⁷ Third, NP's role as Director of Strategic Customer Marketing involved, in her own words, "price increase implementation" (Complaint, page 158). The significance of this move was not lost on other industry leaders. At Taro Pharmaceuticals,

¹⁵McKesson 2014 10-K, Cardinal 2014 10-K, ABC 2014 Annual Summary all explicitly state that their profits are positively affected by manufacturer price increases (due to cost-plus arrangements).

¹⁶Note that while the complaint focuses on drugs affected by NP, it alleges other segments of the industry were not entirely immune to coordinated behavior. These allegations do not affect our analysis. The complaint describes antitrust violations in the other segments that are qualitatively and comparatively unimportant in our setting.

¹⁷Most generic drugs are manufactured in a pill/capsule form, which is the delivery mechanism we study. Injectable drug markets differ in a host of ways—manufacturing processes are very different and customers are typically hospitals rather than retail pharmacies.

another leading generic firm, the Vice President of Sales and Marketing emailed the COO just days after NP left ABC to say, "[NP] Going to Teva—Hush Hush for now" (Complaint, page 159). Our sample includes only drugs that Teva historically produced. As the complaint states, Teva was uncooperative before NP joined.

The complaint states that just 8 days after joining Teva, NP began identifying target markets. The process was highly structured. She started by assigning each generic firm an individual "quality" rating, which ranged from -3 to 3, and reflected the strength of her relationships with their sales and marketing executives. Next, she combined these ratings into a drug-specific score, which was the most important element in the selection process. Finally, she factored in the number of firms in the market and other minor considerations. In an internally distributed 2014 document, she summarizes "Candidate Identification" as "Target 2-4 total players, where the quality of competitor is high" and "[where] Teva has majority share and quality of competitors is high" (Complaint, page 217).¹⁸

The complaint states that the outcome of the process was a spreadsheet titled "Immediate PI File" (where "PI" stands for "Price Increase"). It contained a list of drugs for which NP expected to cooperatively raise prices. NP forwarded the list to supervisors on May 24, and Teva changed prices on July 3, preceded or followed closely by other firms. The magnitudes of the increases are especially noteworthy. Many commonly prescribed medications—drugs treating cancer, bacterial infections, arthritis pain, and high blood pressure, to name a few—doubled or more in price.

Fixing prices required considerable coordination, which in turn required frequent communication. Leading up to the increases, phone records reveal thousands of calls and messages between NP and her counterparts at firms producing the drugs for which she sought to raise prices. For instance, the day before an increase, NP exchanged 15 calls with 3 individuals.¹⁹ Interestingly, since perfectly synchronizing price changes is impossible, sharp increases by one firm often prompted its customers to approach other producers, who were forced to decline the business. In some cases, firms outright declined, citing fictitious supply problems (Complaint, pages 239 and 265), while in others they quoted inflated prices—a tactic NP called "fluff pricing" (Complaint, page 146). This frustrated buyers such as Schnuck's, a Midwest grocery chain, which felt "so insulted" by one of the egregious quotes (Complaint, pages 146-7). Ultimately, NP coordinated five main rounds of price increases between July 2013 and January 2015 (Complaint, page 3).²⁰ Notably, the rollout would have been even more compressed were NP not to have taken maternity leave in the middle of it, which resulted in a six-month "hiatus" (Complaint, page 212).

According to the Connecticut Attorney General (AG), "suspicious price increases of certain generic

¹⁸The complaint reports NP's quality ratings and we observe market structure, so we can confirm NP used precisely these criteria. Firm quality is presented in Table A.2 and average quality across markets is presented in Figure A.1. She also mentions targeting "Exclusive items" (i.e., those where Teva is the only supplier). Our focus is collusion, so we exclude these cases from our sample. They are rare, so this restriction is without any meaningful loss of generality.

¹⁹It does not affect our analysis but is interesting to note that most conspirators knew their behavior was illegal. For instance, when NP described communications with rival firms during a 2013 internal meeting, MP, another Teva executive, "smiled, put her hands over her ears, and pretended that she could not hear what was being said" (Complaint, page 337). In other instances, executives deliberately avoided written communications. When a senior executive at Taro asked about the arrangement, a fellow executive replied, "No emails please. Phone call. [Redacted] let's discuss" (Complaint, pages 49-50). In yet other instances, executives used personal email accounts to transfer illegally obtained information and deleted illegal communications when they learned of the government investigation (Complaint, page 341).

²⁰After the first round, opportunities for subsequent increases emerged for various reasons. One source is leadership changes. If an employee with close ties to NP moved from Firm A to Firm B, then the likelihood of Firm B complying with the cartel agreement increased. For example, when Zydus, which was initially rated -3, hired NP's colleague, KG, the firm's score increased to 2. Subsequent phone and text records indicate that the two communicated extensively (Complaint, page 272-273). A second source of opportunities was supply disruptions. If a firm without close ties to NP lost access to the active pharmaceutical ingredient and left the market, then the likelihood of cartel formation increased.

drugs" prompted the state in July 2014 to begin an investigation, which is still underway.²¹ On May 10, 2019, 43 US states and territories filed a complaint, which was unsealed the following month. It alleges a "horizontal conspiracy" to fix prices for multiple generic drugs in violation of Section 1 of the Sherman Act. Alongside the states' civil suits, the US Department of Justice brought criminal charges against several firms and individuals.²²

The complaint states that NP's supervisor was "aware of the government investigations that had been commenced" by early 2015. She "told [NP] that the government was showing up on people's doorsteps," and "warned [NP] to be careful about communicating with competitors" (Complaint, page 341). Again according to the complaint, news of the investigation ended the price increases (Complaint, page 338). However, the news does not appear to have undermined existing agreements, which the firms may have maintained tacitly. This is consistent with NP's assertions that generic drug price fixing is hard to initiate but easy to sustain (Complaint, page 160) as well as internal documents referencing cooperative arrangements years after the government launched its investigation (Complaint, page 50). We address this issue directly in Section 7.

The complaint further suggests that cartel members—firms participating in the collusive scheme in any market—understood the value of sticking to the agreement. It is also worth noting what type of conduct is absent from the complaint and related filings. Despite thousands of pages of court documents recounting almost countless conversations about prices, no statements suggest that cartel members coordinated entry. In other words, it seems very unlikely that members reached agreements not to enter each others' markets. Consistent with that view, when entry by a cartel member into a cartelized market did occur, incumbent producers responded by accommodating rather than retaliating.²³ We test for this behavior explicitly in Section 7. We further use our conduct tests to validate assumptions about expectations of post-entry pricing behavior.²⁴

3 Data

3.1 Sources

The data come from several sources. The National Drug Code (NDC) Directory, which is published by the FDA, provides a current list of all prescription pharmaceutical products manufactured for sale in the US. Each product has a unique NDC code, which identifies the substance-delivery-release-strength combination of the product, the firm that produces it, and the ANDA that authorizes that production. The FDA updates the list daily. Using the Internet Archive we take annual snapshots of the directory, which are the starting point for our panel dataset.

²¹To be precise, the investigation traces back to a *New York Times* article titled "Rapid Price Increases for Some Generic Drugs Catch Users by Surprise" (Rosenthal, 2014). A supervisor in the Connecticut AG office's unit of antitrust and fraud read and forwarded the article to a staff attorney, who subsequently sought subpoenas (Pazniokas, 2019).

²²Teva, e.g., was charged in August 2020. The charges are serious. For instance, Taro paid over \$200 million to settle its case. Moreover, if found guilty, executives at these firms could face federal prison sentences.

²³ANDA launches by Aurobindo, Lupin, and Actavis are examples. See Complaint, pages 74, 81-82, and 103-104, respectively.

²⁴Further evidence comes from Civica Rx, a startup generic manufacturer owned by a consortium of hospitals. In a *New York Times* article titled "Fed Up With Drug Companies, Hospitals Decide to Start Their Own" (Abelson and Thomas, 2018), the firm stated that when it enters the market it expected incumbents to respond by "quickly dropping the price of the drugs in question." Further, the firm stated that were it to exit subsequently, it expected incumbents to respond by "raising them again later."

Medicaid State Drug Utilization Data, which is published by the Centers for Medicare & Medicaid Services, is the primary source of price and quantity information. Each quarter, all states and the District of Columbia report the number of prescriptions filled by Medicaid enrollees and the corresponding expenditure. We construct prices by dividing expenditure by quantity. We download national aggregate statistics, and we merge to the NDC Directory data, described above, at the NDC-quarter level. Since NDCs are specific to drug packaging (e.g., 1000-count bottles, 14-count blister packages, etc.), firms are associated with multiple NDCs per drug. Thus, we sum over NDCs to get drug-firm-quarter observations.

We also acquired quantity data from IQVIA, a private provider whose National Prescription Audit covers 92% of US pharmacies. The IQVIA data provides detailed information on aggregate quantities at the drug-firm-quarter level but no information on price. We first use IQVIA data to scale up the quantities reported in State Drug Utilization Data, which covers only Medicaid, to US totals. We do this by computing the ratio of total IQVIA prescriptions to total Medicaid prescriptions, and we scale up by that factor. Second, we use IQVIA to check carefully whether patterns in Medicaid are representative of industry-wide purchasing behavior. Total purchases track extremely closely with one another over time, and their responses to cartel-induced price changes are almost identical (see Appendix Figure A.2).

The complaint, which is described in the previous section, provides information about cartel formation. The complaint lists the drugs for which NP fixed prices, the dates on which those prices were fixed, and NP's relationships. From this information, we construct an indicator variable for whether a cartel was ever formed in a particular drug market (i.e., market for a particular substance-delivery-release-strength). Within each drug market where a cartel is formed, we also construct an indicator for whether a particular period precedes or follows cartel formation. Finally, we construct an indicator variable for cartel membership.

We then merge in ANDA filing and approval dates. To obtain filing dates, we scrape the FDA website for all available ANDA approval letters, most of which state the date the application was filed. ANDA numbers are largely assigned in chronological order. We were able to parse filing dates for 8,185 ANDAs from PDFs. We infer the rest using the FDA's numbering system (e.g., if ANDA 70001 was filed on January 1 and ANDA 70015 was filed on January 10, we infer that ANDAs 70002–700014 were filed between January 1 and January 10). To fill in missing filing dates, we regress ANDA numbers on filing dates and then assess fit. To minimize measurement error, we aggregate to the quarterly level. See Appendix Figure A.3 and the associated discussion, which shows that we measure ANDA filing dates without meaningful error. We obtain approval dates from the Orange Book, which is published by the FDA.

Our sample covers 2008 to 2019 and consists of all drugs manufactured by Teva in the quarter prior to NP joining the firm, regardless of manufacturer, subject to some exceptions. We omit drugs that lose exclusivity during the sample period.²⁵ We also omit 11 substance-delivery-release combinations for which NP raised prices twice. Last, we drop 38 injectable drugs: injectables occupy a very different segment of the industry, require radically different production methods, have short shelf lives, are mostly manufactured by different producers than those we study, and most importantly are never mentioned in the complaint. The sample construction is described in detail in Appendix Table A.1.

²⁵They exhibit drastic price declines that ruin comparability across units and reflect idiosyncrasies of patent challenges and expiration, which are unrelated to cartel formation, including, for example, Paragraph IV challenges.

3.2 Summary Statistics

Table I summarizes the data at the drug-year level. There are 4,992 observations spread over 416 drugs and 12 years. The average number of prescriptions filled is 1.9 million, and the average price is \$31. Average yearly expenditures are \$31 million, or about \$250 billion over the full panel. The mean number of manufacturers in the market is 4.15, while the mean number of ANDA filings is 0.21, highlighting that entry into generic drug markets is typically rare.²⁶ Most ANDA filings can be attributed to cartel "nonmembers" (i.e., firms that are not named in the Complaint). The group accounts for about 60% of entry over the full sample and closer to 67% in the periods after NP joined Teva, regardless of whether we consider all drug markets or just those in which cartels are eventually formed (not shown). Cartels are formed in the markets for 113 drugs, or 27% of the sample.²⁷

Table I: *Summary statistics*

	Count	Mean	Std. dev.	Minimum	Maximum
Price	4,992	31.66	40.42	2.32	628.53
Quantity (in thousands)	4,992	1,904.20	3,618.82	0.15	29,571.61
Expenditure (in millions)	4,992	31.09	110.51	0.01	4,188.54
Number of firms	4,992	4.15	1.93	1.00	12.00
Number of ANDA filings	4,992	0.21	0.54	0.00	5.00
Cartelized drug	4,992	0.27	0.44	0.00	1.00
Cartelized drug \times Post cartel formation	4,992	0.14	0.34	0.00	1.00

The unit of observation is a drug-quarter. Price refers to dollars per prescription. Quantity is measured in thousands of prescriptions. Expenditure is measured in millions of dollars. "Cartelized drug" is an indicator for whether a cartel formed in the drug market at any point in the sample. "Cartel drug \times Post cartel formation" is the interaction of "Cartelized drug" and an indicator for periods following cartel formation (i.e., the indicator is "on" if the cartel for a particular drug formed in April 2013 and the observation summarizes Q3 2013 or later).

4 Descriptive analysis

In this section, we describe the patterns in the data that motivate the structural analysis.

4.1 Price changes following cartel formation

We first ask whether and to what extent prices increase following cartel formation. To answer it, we plot the average log prices of cartelized and uncartelized drugs against calendar time in quarters. We normalize each price series to zero in the period just prior to the one in which NP joins Teva, and we mark that period with a vertical red line.

Figure I reports the result. Four aspects of the graph stand out. First, the prices of cartelized and uncartelized drugs exhibit very similar trends in the periods leading up to Teva hiring NP. Second, within

²⁶The exception is the period immediately following the loss of exclusivity of the reference branded drug. Using data from an earlier period, Scott Morton (1999) shows that 80% of entry happens within six months after the initial generic approval.

²⁷NP selects markets where she has close relationships with substantively all of the other firms. The probability this occurs falls with the number of firms in the market, so it also falls with market size. Thus, cartelized markets should be somewhat smaller and have fewer firms. Appendix Table A.3 confirms this prediction.

the pre-event trends, and quarter-to-quarter innovations in the price series are correlated. Third, the price of uncartelized drugs trend smoothly through Teva’s hiring of NP, suggesting that the generic markets were not affected by any major industry-wide shocks (aside from cartel formation). Fourth, the price of cartelized drugs rise sharply following NP joining Teva—they increase about 50% by the end of 2014.



Figure I: Prices rise sharply following cartel formation

This figure plots the average log price of cartelized and uncartelized drugs on the y-axis against calendar quarter on the x-axis. The vertical red line corresponds to the first quarter of 2013—the period in which NP joined Teva. Prices are normalized to zero in that quarter.

We report supplementary results in Online Appendix B. In Appendix Figure B.3, we show that the relationship between the "quality" scores assigned by NP and the likelihood of cartel formation is so tight that we can produce a graph similar to Figure I by comparing "high quality" and "low quality" markets (rather than cartelized and uncartelized ones).²⁸ For a more precise sense of timing and magnitude, we report event study estimates that compare prices before and after cartel formation, conditional on drug and year fixed effects in Appendix Figure B.1. Finally, to ensure our findings are not different by variation in the timing of cartel formation, we follow Sun and Abraham (2020), whose approach accounts for potential contamination of leading and lagging coefficients, and obtain nearly identical estimates in Appendix Figure B.2.

4.2 Entry following cartel formation

We then examine entry. Before a firm launches a generic product, it must file an ANDA and then obtain FDA approval. We begin by studying the first step in the process captured in our data, ANDA filings. Since ANDAs apply to all strengths within a substance-delivery-release combination, we measure entry at that

²⁸In contrast to these results, very different patterns emerge when we restrict attention to markets in which Teva is a monopolist (see Appendix Figure B.4).

level both here and in the structural model. We plot the average number of ANDA filings in cartelized and uncartelized markets against calendar time in years.²⁹ We normalize each series to zero in 2013, the year NP joins Teva, and mark that period with a vertical red line.

Panel A of Figure II reports the result. Measured by ANDA filings, entry into cartelized and uncartelized drug markets track closely with one another prior to NP joining Teva, and entry into uncartelized markets trends smoothly through this event. Yet, as with prices, ANDA filings rise sharply in cartelized markets in the periods immediately following NP's hiring.³⁰

ANDA filings do not permit immediate entry. The FDA must review and approve the ANDA, which involves delays. To capture effective entry, we repeat the process used to produce Panel A but replace ANDA filings with ANDA approvals. Panel B of Figure II reports the result. Whereas ANDA filings in cartelized markets increase in 2014, ANDA approvals do not rise until 2017, consistent with regulatory delays of roughly three years.³¹

For a more complete portrait of delays, we plot the full distribution of delays measured in years. Appendix Figure B.8 reports the result. The median time between ANDA filing and approval is 3 years, consistent with Figure II. Notably, the distribution of total delays—the time between the formation of the cartel and the start of production—is shifted right by 1.5 years, suggesting ANDA preparation and/or factory setup also delay entry.

4.3 Post-entry price changes

The negative relationship between entry and prices is better established in US drug markets than perhaps any other industry (Wendling et al., 2011)), but the situation may be complicated by collusion in our setting. Here we present additional, suggestive data patterns that shed light on the entry-price relationship.

First, to compare markets where entry is more or less likely, we exploit variation in market size, following a large body of established work (see, e.g., Berry and Waldfoel (1999) and the work cited therein). Market size shifts the probability of entry but is less likely to be correlated with unobserved factors that influence prices.³² In Appendix Figure B.9, we confirm that, following cartel formation, larger markets attract more entrants, echoing cross-sectional results reported in Scott Morton (1999). Then, in Panel A of Figure III, we compare prices in large and small cartelized markets over time, using prices of uncartelized drugs to control for any time-varying factors affecting the entire industry. To produce it, we estimate

$$y_{dt} = \sum_{\tau=-21}^{24} \beta^{\tau} x_{dt}^{\tau} + a_d + b_t + e_{dt}, \quad (1)$$

²⁹Aggregating up to the year level improves legibility.

³⁰Infrequently, producers re-enter markets using dormant ANDAs (i.e., approved filings that are no longer associated with production but once were). This might be especially common in markets where cartels have recently formed, since they drive price rises that could encourage not only *de novo* entry in the form of ANDA filings but also induce inactive firms to become active once again. To evaluate this possibility we re-estimate equation 26, replacing Y with the number of "re-entries" per drug-year. Due to the relatively small number of these occurrences, we report the result in Appendix Figure B.7. Re-entry spikes in the year immediately following collusion. This pattern is consistent with the cartel turning some unprofitable markets profitable once again, and it squares with the fact that already-approved applications can enter without delay.

³¹For completeness, we also plot entry in event time relative to cartel formation (see Appendix Figure B.6).

³²Cross-sectional differences in entry are plainly endogenous. Nonetheless, entry is associated with lower prices in cartelized markets. That is, prices in cartelized markets with and without entry rise sharply after cartel formation and peak around 18 months later, but they follow qualitatively different paths thereafter. Measured from their peaks to their values at the end of the sample, average prices in markets with entry fall 39 percent, whereas average prices in markets without entry are stable (i.e., they plateau and do not move more than about one percent in either direction).

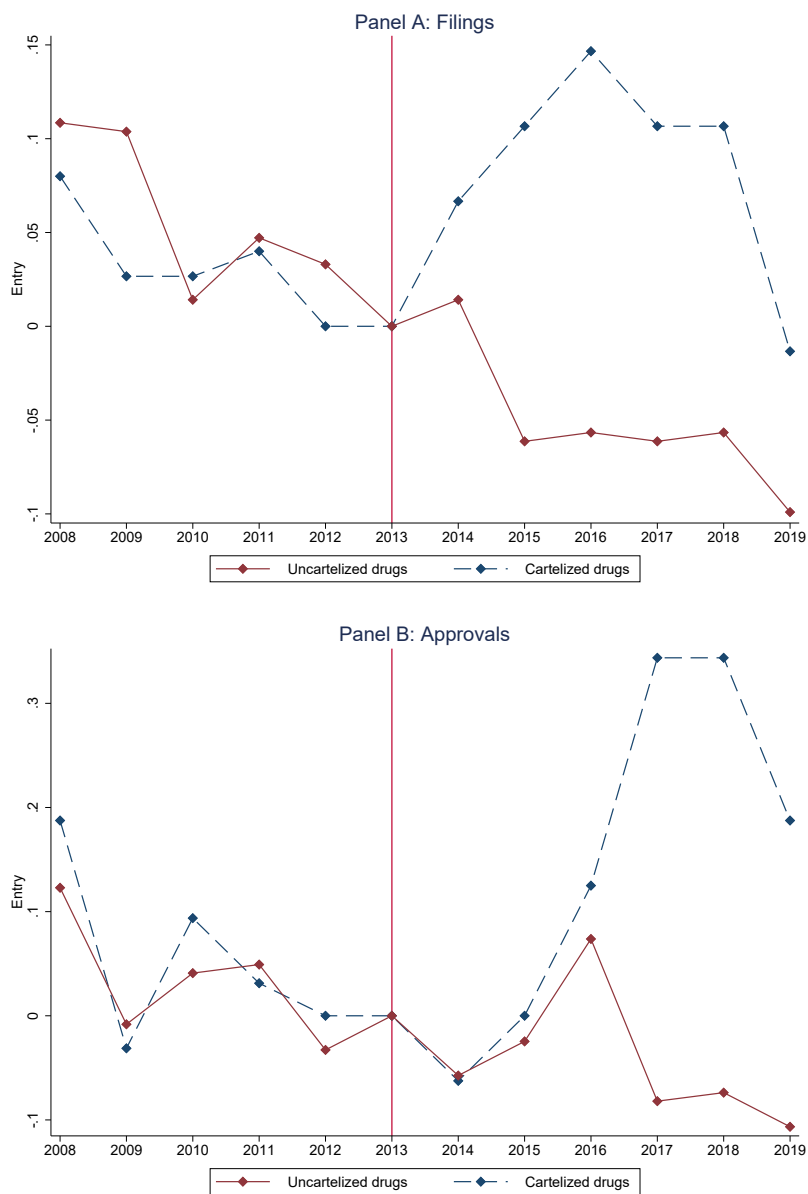


Figure II: ANDA filings rise following cartel formation, but ANDA approvals are delayed.

This figure plots average number of entrants per substance-delivery-release-combination on the y-axis against calendar year on the x-axis. Panel A measures entry using ANDA filings, whereas Panel B measures it using ANDA approvals. The vertical red line corresponds to the first quarter of 2013—the period in which NP joined Teva. The number of entrants are normalized to zero in that quarter.

where y_{dt} denotes the average log price of drug d in quarter t , a_d and b_t denote drug and quarter fixed effects, respectively, and x_{dt}^τ denotes an indicator variable that equals one if and only if d is a cartelized drug and t is τ periods from cartel formation. Given the timing of cartel formation in relation to our sample period, we restrict attention to $\tau \in [-21, 24]$, and we set $\beta^{-1} = 0$ to facilitate comparisons to the period just prior to cartel formation. We first estimate equation 1 on uncartelized and large cartelized markets and plot the results along the dashed, blue line. We then re-estimate equation 1 on uncartelized and small cartelized markets and plot those coefficients along the solid, red line.

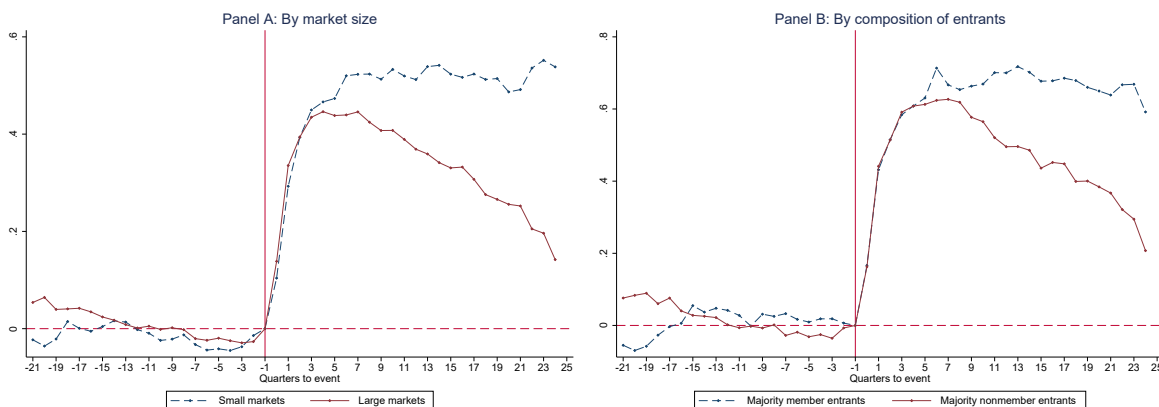


Figure III: Prices after cartel formation

In this figure, we plot estimates of β^τ , which are derived from equation 1, against x^τ , which represents event time in quarters. For Panel A, we proxy for market size by computing the number of prescriptions filled in the period just prior to NP joining Teva, and we distinguish between large and small cartelized markets based on how this figure relates to the median value. For Panel B, we calculate the fraction of entrants that arrive as a result of ANDAs filed after NP joins Teva and identify markets in which a majority of those firms are members and nonmembers, respectively.

Prices in large cartelized markets exhibit very similar pre-event trends as those in small cartelized markets. After cartel formation, both series rise sharply and reach similar peaks. After that point, the series diverge. Prices in large markets decline throughout the remainder of the panel, while prices in small markets remain completely stable. Readers should obviously interpret this result cautiously: we cannot, for example, rule out the possibility that this divergence reflects, in part, an effort to reduce antitrust risk.³³

Second, we expect that nonmembers drive the entry-price relationship. According to the Complaint, NP repeatedly coordinated pricing decisions by member entrants to "avoid competition" (see, e.g., Complaint (page 107)). Yet, nowhere in the over-500-page document does the government allege NP coordinated prices set by nonmember entrants (or nonmember incumbents, for that matter). To make comparisons based on the cartel affiliation of the entrants, we identify firms whose ANDAs were filed after NP arrives at Teva and compute the fraction that are members in each cartelized market. In Panel B of Figure III, we compare the prices of cartelized drugs for which the majority of entrants are members with the prices of cartelized drugs for which the majority are nonmembers, using the prices of uncartelized drugs as controls.

³³On the one hand, high prices for popular drugs may generate headlines, in which case members may reduce prices in large markets in an attempt to avoid prosecution. On the other hand, when we trace the investigation back to its origin, obtain the list of drugs whose price changes triggered the inquiry, and compare them to all products manufactured in our sample, their sizes are similarly distributed (see Appendix Figure B.10 for details.) Moreover, despite carefully describing how members allegedly responded to the government's investigation (e.g., by deleting text messages), no statements come close to suggesting that they adjusted prices in response.

We then follow exactly the same procedure that we use to produce Panel A.

Again, we find that the two price series exhibit similar pre-event trends and reach similar peak prices about one year after cartel formation but diverge afterwards. Cartelized markets in which most entrants are nonmembers experience declining prices, whereas ones in which most entrants are members experience very stable prices. These patterns are consistent with claims that member entrants comply with existing cartel agreements while nonmembers do not.³⁴ Ultimately, incentives to enter depend on post-entry conduct. To address these issues, we turn to a structural model of entry, pricing, and demand.

5 Model

In this section, we present a model of generic prescription drug supply and demand. For each drug in each period, there are two stages. In the first stage, firms that do not have regulatory approval and would like to enter file an ANDA. The FDA approves these filings after a delay of uncertain duration. In the second stage, firms that have regulatory approval manufacture and sell the drug to buyers. To ease exposition, we describe the model in reverse chronological order, starting with demand.

5.1 Stage I: Demand

Patients are the final consumers of generic prescription medications but are typically not the ones to decide which firm's product to purchase. Instead, buyers are intermediaries such as wholesalers (e.g., Cardinal Health), group purchasing organizations (i.e., cooperatives formed to source drugs), and large retail chains (e.g., CVS). The buying process resembles a procurement auction. For every drug in every period, each buyer submits a request for proposals (RFP), to which any firm with an approved ANDA may respond. Next, the buyer assigns each proposal a score, which is the payoff associated with selecting that proposal's product. Buyers maximize payoffs by selecting the highest-scoring proposal or an outside option.³⁵

Scores depend on price and non-price attributes. For instance, Cardinal Health states that it values not only low prices but also short lead times, frequent product availability reports, infrequent recalls, accurate invoicing, high historical fill rates, and high quality, on-time deliveries (Cardinal Health, 2020).³⁶ Since non-price attributes mostly reflect large, historic investments or past operating performance, we assume that these aspects of the proposal are fixed when firms submit proposals.

To formalize the process, we index firms by f , buyers by i , drugs by d , and periods by t . The payoff associated with buyer i selecting the product made by f is given by

$$v_{idft} = \lambda_d + \lambda_t + \zeta_{dft} + \zeta_{idt} + (1 - \sigma)\epsilon_{idft} - \alpha_d p_{idft}. \quad (2)$$

λ_d and λ_t capture differences in the prevalence of medical conditions and the popularity of treatment

³⁴In Appendix Figure B.11, we show that cartel formation affected entry among members and nonmembers similarly.

³⁵In our setting, the outside option entails drawing down existing stock, aggressively optimizing inventory across warehouses and stores, and intermittently stocking out at certain locations. Note that stockouts of generic drugs are common during the sample period. See, e.g., the FDA's 2020 report for a summary of the problem, which is available at <https://www.fda.gov/media/150409/download>.

³⁶See also the RFP template published by the Minnesota Multistate Contract Alliance for Pharmacy (MMCAP), which is available at <https://www.bidnet.com/bneattachments?/489130294.pdf>.

regiments that vary over drugs and periods, respectively. ζ_{dft} , ϵ_{idft} , and ζ_{idt} are i.i.d. disturbances.³⁷ We assume that ϵ_{idft} is distributed Type 1 extreme value and that ζ_{idt} is distributed such that $[\zeta_{idt} + (1 - \sigma)\epsilon_{idft}]$ is also distributed Type 1 extreme value. As is common in the literature, we assume firms know ζ but not ζ or ϵ when they choose prices. α_d measures the sensitivity of the payoff to price, represented by p_{idft} , and may vary by drug. Since buyers rank alternatives, the mean payoff assigned to the outside option is normalized to zero without loss of generality, so $v_{id0t} = \epsilon_{id0t}$.

We assume there are a large number of similarly sized buyers, which is a realistic approximation to the actual "downstream" market structure.³⁸ This assumption makes the problem tractable and, as we will show, produces sensible estimates.³⁹ Under these conditions we can treat the collection of procurement scoring auctions as a nested logit demand system (Einav, 2003; Miller, 2014; Miller and Sheu, 2020). See Online Appendix D for the proof.

The market share of f in t at d is given by

$$s_{dft} = \frac{e^{\lambda_d + \alpha_d p_{dft} + \zeta_{dft}}}{\Lambda^\sigma (1 + \Lambda)^{1 - \sigma}}, \quad (3)$$

where $\Lambda = \sum_{f' \in \mathcal{F}_d} e^{\lambda_d + \alpha_d p_{df'} + \zeta_{df'}}$. The market share of the outside good, s_{df0} , is obtained by replacing the numerator in equation 3 with one and the denominator with $(1 + \Lambda)^{1 - \sigma}$.

5.2 Stage II: Pricing

Profit is given by

$$\pi_{dft} = (p_{dft} - mc_{dft}) s_{dft} A_d, \quad (4)$$

where mc_{dft} denotes marginal cost and A_d denotes market size.⁴⁰ We parameterize marginal cost such that

$$\ln(mc_{dft}) = \gamma_d + \gamma_t + \omega_{dft}, \quad (5)$$

where ω_{dft} is an i.i.d. disturbance.

In uncartelized markets, and in cartelized markets prior to cartel formation, we assume that all firms set *competitive* prices, which maximize individual profits. We denote competitive prices by p_{dft}^B and corresponding profits by π_{dft}^B . Alternatively, when a cartel forms, we assume that NP reaches an agreement

³⁷Within a given drug-period, ζ captures product-specific shocks such as manufacturing defects that temporarily reduce quality. For instance, certain lots of Aurobindo's valsartan were contaminated with the carcinogen N-Nitrosodiethylamine (NDEA). The lots were recalled, but news of the incident temporarily reduced demand for the firm's version of the drug. ϵ captures idiosyncratic buyer-product-specific shocks. For example, distributors including Cardinal value the ability to "backhaul" shipments, which ultimately depends on where its distribution facilities are located in relation to where the firms warehouse their product. ζ allows the score of "inside" products to be correlated with one another.

³⁸State attorneys name 35 different buyers in the complaint: Amerisource Bergen, Cardinal, HD Smith, McKesson, Morris and Dickson, Viziant, Premier, Intalere, MMCAP, Econdisc, OptiSource, Humana, NC Mutual, PVA Health, Cigna, OptumRx, Prime, Kaiser, Armada, ANDA, Omnicare, Key Source, CVS, Walgreens, Rite Aid, Publix, Walmart, Target, Giant Eagle, Schnucks, Ahold, Hannaford, Kinney Drug, and H-E-B. However, these are merely names that arise in anecdotes describing cartel behavior, so the true number of intermediaries with whom firms contract is even larger (e.g., Safeway does not appear in this list). To be transparent, in reality, some buyers are larger than others. For instance, Teva's 2013 annual report states its largest buyer accounts for 17% of sales.

³⁹We verify estimates (e.g., levels and changes in profit margins) using outside sources and forecast damages that are quantitatively similar to Cuddy (2020), who models the auctions directly.

⁴⁰We use IQVIA data to compute the total number of prescriptions per year per drug. We compute the maximum value for each drug over the sample and set market size equal to fifty percent above this figure.

with other members to set *collusive* prices, which maximize cartel members' joint profit. Based on facts presented in the complaint and patterns in the data, we also assume NP does not approach nonmembers to discuss the agreement, so each member best responds to the cartel. When members maximize their joint profit and nonmembers best respond, we denote each member's price by $p_{dft}^{C,M}$ and each nonmember's price by $p_{dft}^{C,N}$, and we denote the corresponding profits they earn by $\pi_{dft}^{C,M}$ and $\pi_{dft}^{C,N}$, respectively.

Merely reaching an agreement does not guarantee compliance. Absent other considerations, not complying is always profitable for members, so the agreement must stipulate punishment. Given the effort required to coordinate behavior, the legal risk involved with price fixing, anger that would arise from betrayal, and subsequent loss of trust, it is likely that a single intentional deviation would not only undermine the cartel but also preclude its members from forming a new one. Thus, we assume cooperation among cartel members is supported by trigger strategies. Conceptually, all members agree that in every future period, each member will set collusive prices so long as every member has set collusive prices since cartel formation, and they also agree that if any member does not comply, then all members will set competitive prices forever after. In this sense, noncompliance "triggers" reversion to Nash-Bertrand.

Under trigger strategies, prices depend on the beliefs of firms and the history of play. In Section 7, we formalize the dynamic game, measure the incentives to comply with the cartel agreement at every point in time, report that trigger strategies form an equilibrium, and explicitly evaluate conduct. For the remainder of this section and the entirety of the next, we (temporarily) assume that once a cartel forms, it is in each member's individual self-interest to set collusive prices. Doing so greatly eases exposition and economizes on notation.

5.3 Stage I: Entry

In the first stage of the period in which a cartel forms, we assume that firms without an approved ANDA have a one-time opportunity to file one. Filing an ANDA allows a firm to enter after a delay of uncertain duration, denoted by D . For tractability, we assume that other market structure changes are determined outside the model. Since entry into mature generic drug markets is otherwise rare, and since exit is usually precipitated by supply disruptions outside a firm's control, this assumption greatly reduces the computational burden but is unlikely to have a material impact on our conclusions.

We assume that firms know the distributions of ξ , ω , and D , denoted by F_ξ , F_ω , and F_D , respectively, and they have rational expectations about future realizations of these random variables. Experienced firms learn about these objects over time and nascent firms can easily estimate them from data similar to ours. Also, we assume that firms know how λ_t , γ_t , and the number of member and nonmember incumbents will evolve over time, and we assume that these objects are fixed at their end-of-sample-values. In mature generic drugs, market-wide demand and cost shocks are rare, exit is infrequent, and re-entry is equally infrequent (provided at least four years have passed since cartel formation).

Since ANDAs apply to all strengths within a substance-delivery-release combination, our analysis of entry occurs at that level. We index substance-delivery-release combinations by j and denote the set of d corresponding to j by \mathcal{D}_j . The value of entering j equals the value of entering each drug market associated with it and depends on the number of member and nonmember entrants, denoted by χ_M and

χ_N , respectively. For a member, the value of entering j is given by

$$VE_j^M(\chi^M, \chi^N) = \sum_{d \in \mathcal{D}_j} VE_d^M(\chi^M, \chi^N). \quad (6)$$

where

$$VE_d^M(\chi^M, \chi^N) = \sum_{t=1}^{\infty} \delta^t F_D(t) \sum_{e_M=0}^{\chi_M-1} \sum_{e_N=0}^{\chi_N} \left[\rho(\chi_M - 1, e_M, t; F_D) \rho(\chi_N, e_N, t) \right. \\ \left. \times \int_{\xi} \int_{\omega} \pi_{d,f,t_d+t}^{C,M}(M_{d,t_d+t} + e_m + 1, N_{d,t_d+t} + e_n, \xi_{d,t_d+t}, \omega_{d,t_d+t}) dF_{\xi} dF_{\omega} \right]. \quad (7)$$

δ denotes the discount factor. $M_{d,t}$ and $N_{d,t}$ denote the number of member and nonmember incumbents in d at t , respectively. t_d denotes the period in which the cartel is formed. $\rho(\chi_M - 1, e_M, t; F_D)$ is the probability that e_M other cartel member entrants are active t periods after the cartel is formed, while $\rho(\chi_N, e_N, t; F_D)$ is an analogous probability for nonmember entrants. Both are binomial expansions given by

$$\rho(a, b, t) = \frac{a!}{(a-b)! b!} F_D(t)^b [1 - F_D(t)]^{a-b}. \quad (8)$$

The final term in equation 7, $\int_{\xi} \int_{\omega} \pi(\cdot) dF_{\xi} dF_{\omega}$, represents the expected variable profits earned by the potential entrant, assuming it is active at $t_d + t$. For a nonmember, the value of entry is denoted by $VE_j^N(\chi^M, \chi^N)$ and given by an expression analogous to the one that appears on the right-hand side of equation 6. The payoff to a potential entrant that does not file an ANDA is normalized to zero regardless of cartel membership.

Potential entrants weigh the value of entry against expected sunk costs. We denote the sunk cost for f in j by θ_{jf} , its information set by \mathcal{I}_{jf} , and its expectation operator as \mathcal{E} . We assume firms know the vector of θ_{jf} when they make entry decisions. Cartel member f files an ANDA for j if and only if

$$VE_j^M - \mathcal{E}[\theta_{jf} | \mathcal{I}_{jf}] \geq 0. \quad (9)$$

Nonmember f' follows an analogous rule, which depends on VE_j^N instead of VE_j^M and $\mathcal{E}[\theta_{jf'} | \mathcal{I}_{jf'}]$ instead of $\mathcal{E}[\theta_{jf} | \mathcal{I}_{jf}]$. Entry decisions form a simultaneous move Nash equilibrium. More than one may exist.⁴¹

6 Estimation

In this section, we describe the methods we employ to estimate the structural parameters. We also report the resulting estimates.

⁴¹Multiplicity does not pose a problem for estimation, since our strategy relies only on the necessary conditions for equilibrium (see the following section). However, when we recompute equilibrium under counterfactual policies, we must select among them (see the Section 8).

6.1 Stage II: Demand

To estimate the demand system, we rely on the market share inversion proposed by Berry (1994) and estimate

$$\ln(s_{dft}) - \ln(s_{d0t}) = \lambda_d + \lambda_t + \alpha_d p_{dft} + \sigma \ln(s_{dft|dgt}) + \xi_{dft}, \quad (10)$$

where $s_{dft|dgt}$ denotes f 's share of the inside good, which equals $s_{dft}/(1 - s_{d0t})$. The left-hand side of equation 10 is a straightforward transformation of the data, and the right-hand side is linear in parameters to be estimated— λ_d , λ_t , σ , and α_d —and an error term. In our initial specification, we restrict $\alpha_d \equiv \bar{\alpha}$, but we partially relax this restriction below.

Since firms know ξ when they set p , prices are endogenously determined in equation 10, so OLS estimates of $\bar{\alpha}$ will be biased. We estimate the parameter using price variation induced by the cartel. Our price instrument equals the product of a dummy for cartelized drug markets and a dummy for $t \geq t_d$. The intuition behind the identification strategy is transparent. Following cartel formation, the price and consumption of cartelized drugs rise and fall, respectively, compared to uncartelized drugs. The relative changes reflect sensitivity to price, which maps to the $\bar{\alpha}$. Given the origin of the cartel, the selection of drugs, and the timing, it is reasonable to assume that values of the instrument are uncorrelated with drug-firm-year specific demand shocks.⁴² In nested logit models, an additional right-hand side term— $s_{dft|dgt}$ —is also endogenously determined. Our instrument counts the number of products in the market, so we estimate σ using share changes following entry and exit.⁴³

Table II reports the demand estimates. Column 1 corresponds exactly to equation 10. We find that preferences are correlated within the "nest" of inside goods, with $\sigma = 0.57$. As expected, we also find buyers dislike higher prices, with $\bar{\alpha} = -0.11$. Standard errors are clustered at the drug level, and both coefficients are significant at the 1% level. In Column 2, we assess the importance of accounting for "authorized generic" products. Authorized generics are generic versions of drugs that are manufactured by the patent owner, which are often introduced around the point at which it loses exclusivity. Throughout our analysis, we do not distinguish these products from "ordinary" generics. To ensure this distinction is empirically unimportant, we append equation 10 to include a dummy for authorized generics. Reassuringly, estimates of σ and α are entirely unchanged, and the coefficient on the authorized generic dummy is small and not significant.

We observe that two drug classes experience larger-than-expected price changes at the onset of collusion (as shown in Appendix Figure C.1) and witness more subsequent entry than one would predict on the basis of market size alone. Both patterns are consistent with less elastic buyers. Thus, to accurately represent demand and correctly forecast entry incentives, we allow α_d to vary based on whether the drug is an β -blocker, anticonvulsant, or other generic. That is, we replace $\bar{\alpha}$ on the right-hand side of equation 10 with $\alpha_0 + \alpha_1 \mathbb{1}\{\beta\text{-blocker}_d\} + \alpha_2 \mathbb{1}\{\text{antiepileptic}_d\}$, where $\mathbb{1}\{\cdot\}$ denotes an indicator variable.⁴⁴

Column 3 reports the result. We find that $\alpha_1 = 0.16$ and $\alpha_2 = 0.20$, consistent with less elastic buyers in the separately named drug classes. As expected, we find a higher baseline price coefficient (in absolute

⁴²To focus on periods proximate to cartel formation, we restrict attention to a five year window around cartel formation, and to avoid complications arising from mid-year cartel formation, we omit $t = t_d$.

⁴³The instrument equals $M_{dt} + N_{dt}$. In the presence of drug specific dummy variables, which all specifications include, the instrument captures period-to-period changes in the number of active firms. Though the parameters are jointly determined, one can build intuition around identification by considering exit, which reduces the value of the instrument. If the existing firm cedes a large share to other drugs instead of the outside option, this implies σ is relatively large.

⁴⁴Allowing α to vary more flexibly (e.g., by drug class) may be desirable but generates noisy estimates.

Table II: Demand estimates

VARIABLES	(1) IV	(2) IV	(3) IV	(4) IV
Log of inside share	0.56*** (0.16)	0.56*** (0.16)	0.73** (0.31)	0.73** (0.30)
Price	-0.11*** (0.031)	-0.11*** (0.031)	-0.21** (0.10)	-0.21** (0.10)
Price X β -blocker			0.16* (0.089)	0.16* (0.088)
Price X Anticonvulsant			0.20** (0.094)	0.20** (0.094)
Indicator for authorized generic		-0.043 (0.17)		-0.067 (0.31)
Observations	19,299	19,299	19,299	19,299
Drug FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Number of drugs	415	415	415	415

, **, and * denote significance at the 10%, 5%, and 1% levels, respectively. The unit of observation is a drug-year-firm. Standard errors are clustered at the drug level.*

value terms), with $\alpha_0 = -0.21$. We also estimate a stronger correlation of products within the nest of inside goods. All estimates are significant at the 5% level except α_1 , which is significant at the 10% level.⁴⁵

6.2 Stage II: Marginal costs

To recover marginal cost, we restrict attention to these unambiguously competitive markets. Specifically, we isolate observations from (a) markets that were never cartelized and (b) years prior to cartelization in markets that were cartelized. For each observation, we take the derivative of profit with respect to price and solve for marginal cost, which yields

$$mc_{dft} = p_{dft} - \left(\frac{\partial \mathcal{J}_{dft}}{\partial p_{dft}} \right)^{-1} \mathcal{J}_{dft}. \quad (11)$$

We then replace $\partial \mathcal{J} / \partial p$ with our estimate of its value to obtain $\hat{m}c$, and plug the result into the empirical analog of equation 5 to recover estimates of γ_d , γ_t , and F_ω .

We find that the average marginal cost of a drug in our sample is \$10.45. Heterogeneity exists across drugs and over time. For a sense of dispersion, we compute the expected log marginal cost of each drug, $\hat{\gamma}_d + \hat{\gamma}_t$, and plot the resulting densities for cartelized and uncartelized drugs separately. Appendix Figure C.2 reports the result, which shows that the mean, variance, and "shape" of the two distributions are similar.

We use estimates of γ_d and γ_t to predict how marginal costs would have evolved in cartelized markets

⁴⁵Column 4—included mostly for completeness—shows that column 3's results are equally as insensitive as column 1's to distinguish between authorized and ordinary generics. Since accounting for heterogeneity in price sensitivity is important, while separating out authorized generics is not, we rely on estimates from column 3 for the remainder of the paper.

following cartel formation. In Online Appendix C, we provide two pieces of evidence to support our extrapolation. First, Appendix Figure C.3 shows goodness-of-fit for a leave-out sample. Second, we show that our estimated margins match the firm financial statements. We further show that our estimates generate reasonable implied damages between 2013 and 2015.

6.3 Stage I: Sunk costs

6.3.1 Setup

The necessary conditions of a simultaneous move Nash equilibrium imply bounds on sunk costs. If we observe cartel member f enter j , then

$$\sum_{d \in \mathcal{J}} VE_d^M(\chi^M, \chi^N) \geq \mathcal{E}[\theta_{jf} | \mathcal{I}_{jf}], \quad (12)$$

and if not, then

$$\sum_{d \in \mathcal{J}} VE_d^M(\chi^M + 1, \chi^N) < \mathcal{E}[\theta_{jf} | \mathcal{I}_{jf}]. \quad (13)$$

If we observe nonmember f enter j , then

$$\sum_{d \in \mathcal{J}} VE_d^N(\chi^M, \chi^N) \geq \mathcal{E}[\theta_{jf} | \mathcal{I}_{jf}] \quad (14)$$

and if not, then

$$\sum_{d \in \mathcal{J}} VE_d^N(\chi^M, \chi^N + 1) < \mathcal{E}[\theta_{jf} | \mathcal{I}_{jf}]. \quad (15)$$

Inequalities 12-15 are the basis for estimation.

We allow sunk costs to vary with two important observable factors. First, sunk costs depend on the number of strengths associated with the substance-delivery-release combination (e.g., Teva produces atorvastatin in four strengths: 10, 20, 40, and 80 milligrams).⁴⁶ When preparing an ANDA for filing, firms must demonstrate "bioequivalence" to the innovator drug at each strength level. The process is costly, as it involves measuring the time it takes for a given amount of the substance to reach the bloodstream in healthy volunteers. Separately, once approval is granted, most delivery methods require firms to design distinct packaging and install (or repurpose) separate equipment for each strength. Second, sunk costs vary with the delivery method. Compared to tablets and capsules, which are by far the most common dosage forms among orally administered medications, irregular delivery methods such as syrups, solutions, and chewables involve specialized equipment and more complicated packaging. Moreover, these methods are much more susceptible to bacteria growth, which necessitates sterile packing conditions. Finally, sunk costs depend on a symmetric i.i.d. disturbance that is unobserved by the econometrician but known by the firms when they decide whether to enter. We assume it is independent of the substance-delivery-release

⁴⁶Production requires know-how that some firms may lack. To accurately analyze alternative policies that induce more entry, we require that a sufficient number of potential entrants possess this know-how. The requirement is easily met in our setting, which studies relatively "simple" drugs like orally administered solids (i.e., tablets and capsules), but it could bind in other settings.

combination characteristics (i.e., the number of strengths and the delivery method).

Formally, we parameterize sunk costs such that

$$\theta_{jf} = \theta_0 + \theta_1 r_j + \theta_2 \ell_j + \eta_j, \quad (16)$$

where r denotes the number of additional strengths associated with the substance-delivery-release combination, ℓ is an indicator variable that equals one for drugs with an irregular delivery method. η_j represents the disturbance term known to the firms when they contemplate entry. This term permits the sunk costs of entering one market (e.g., warfarin tablets) to differ from the sunk costs of entering another (e.g., cefdinir capsules) in ways that influence behavior.

If the predictions of the model and the actions of the firms differ beyond the flexibility provided for in equation 16, then the differences are rationalized by mean-zero expectation errors, which we denote by ν_{jf} . Their presence reconciles a small number of cases where entry by a nonmember implies that a member should enter but does not (and vice versa). We assume that all potential member entrants realize the same expectational error and that the same is true for potential nonmember entrants. Given this assumption and the fact that these disturbances rationalize differences in the realized payoffs of cartel members relative to nonmembers, we can ascribe them to either group without loss of generality. We arbitrarily pin them to cartel nonmembers, defining $\theta_{jM} = \mathcal{E}[\theta_{jM} | \mathcal{I}_{jM}]$ and $\theta_{jN} = \mathcal{E}[\theta_{jN} | \mathcal{I}_{jN}] + \nu_j$.⁴⁷

Finally, we assume that $\hat{F}_D = F_D$ and that

$$\int_{\xi'} \int_{\omega'} \hat{\pi}_{d,t}^k(M, N, \xi, \omega) d\hat{F}_{\xi} d\hat{F}_{\omega} = \int_{\xi'} \int_{\omega'} \pi_{d,t}^k(M, N, \xi, \omega) dF_{\xi} dF_{\omega}, \quad (17)$$

where $k \in \{(C, M), (C, N), B\}$. \hat{F}_D , \hat{F}_{ξ} , and \hat{F}_{ω} denote the empirical distributions of delays, demand shocks, and marginal cost shocks, and $\hat{\pi}_{d,t}^k(M, N, \xi, \omega)$ denotes $\pi_{d,t}^k(M, N, \xi, \omega)$ evaluated at our estimates of λ_d , λ_t , α_d , σ , γ_d , and γ_t rather than the true parameters. The assumption that expected profits are measured without error is commonly employed in the literature, since it greatly reduces the computational burden of calculating the bounds and is a reasonable approximation to reality in our setting as well.

6.3.2 Moments and inference

To estimate sunk costs, we must address selection introduced by η , which is observed by firms prior to entry but unobserved by the econometrician. To see the problem clearly, suppose we ignore it, naively replace expected sunk costs with $\theta_0 + \theta_1 r_j + \theta_2 \ell_j$, and take inequalities 12-15 directly to the data. Since entry is especially common in markets where sunk cost shocks are small, this approach over-samples negative η when we compute lower bounds, resulting in upward bias. By the same logic, it over-samples positive η when we compute upper bounds, resulting in downward bias.

For lower bounds, the solution to the selection problem introduced by the structural errors lies in the fact that although the conditional expectation of η varies with observed entry, its unconditional expectation is nonetheless mean zero (Ishii, 2005; Ho, 2009; Pakes et al., 2015). To see this conceptually, suppose for the sake of illustration that at least one cartel member enters each substance-delivery-release combination.

⁴⁷For readers more familiar with Pakes et al. (2015), our ν and η correspond to their $\nu_{1,-}$ and $\nu_{2,-}$ -type errors, respectively.

Further, suppose that we construct precisely one instance of inequality 12 for each substance-delivery-release combination, substitute measured objects for true values, ignore the error terms, pool the instances together, and calculate their mean. This process collects one η_j from each j , yielding an unselected set of disturbances whose expected value is zero.⁴⁸

We compute

$$\frac{1}{J} \sum_j \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \frac{1}{2} \sum_{k \in \{M, N\}} \left[\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] h_j^\iota < 0. \quad (18)$$

Moments indexed by ι are formed by interacting the bracketed term with a weight function, denoted by h_j^ι . The weight function includes a constant, an indicator for standard delivery method drugs, an indicator for nonstandard delivery method drugs, and indicators for substance-delivery-release combinations available in one, two, and more than two strengths. J denotes the number of unique substance-delivery-release combinations, and μ_j denotes the number of drugs associated with each substance-delivery-release combination. $\widehat{VE}_d^M(\cdot)$ denotes $VE_d^M(\cdot)$ evaluated at our estimates of $\pi(\cdot)$, F_{ξ} , F_{ω} , and F_D rather than the true values. $\widehat{VE}_d^N(\cdot)$ is defined analogously.

For upper bounds, we take a slightly different approach, since not every substance-delivery-release combination experiences entry. To solve the selection problem, we exploit the symmetry of the distribution of η (Powell, 1986; Pakes et al., 2015). This approach requires additional notation. Let L be the set of j with at least one entrant, J_L be the size of that set, and w^i be a positive valued function of r_j and ℓ_j . Also, define

$$VE_j^+ = \frac{1}{2} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} \left[\widehat{VE}_d^k(\chi^M, \chi^N + \mathbb{1}\{k = N\}) \right], \quad (19)$$

which represents the average of the cartel members' and nonmembers' entry values in j . Finally, for each moment i , order j by their values of $w_j^i VE_j^+$, and let Ψ_{wVE} denote the set of j that corresponds to the $J - J_L$ smallest values. The second set of moments is then given by

$$\begin{aligned} & \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} \widehat{VE}_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M = 0, \chi^N = 0\}} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{wVE}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] < 0. \quad (20) \end{aligned}$$

In Online Appendix D, we prove that inequalities 18 and 20 produce consistent bounds.

Our moments provide one-sided restrictions on the parameters of interest, the most informative of which are the greatest lower bound and the least upper bound. Since our inference procedure is based on maxima and minima, respectively, rather than averages, we cannot rely on the central limit theorem. To

⁴⁸This approach exploits the "ordered choice" nature of the problem. Ishii (2005) illustrates the approach most clearly. We differ from her approach by having two types of entrants—cartel members and nonmembers—and permitting expectational error to reconcile differences in the sunk costs implied by their decisions. See Section III.C of Wollmann (2018) for a general discussion of ways to relax this assumption. To name one, the econometrician could specify the shape of the structural error and take the "probability inequality" approach, proposed by Tamer (2003), though this approach is computationally infeasible in our setting.

obtain 95% confidence intervals around the true parameters, we follow Andrews and Soares (2010).⁴⁹ We search over a three 3-dimensional grid of candidate vectors for $[\theta_0 \theta_1 \theta_2]$, and we invert an Anderson-Rubin type test at each one (Chernozhukov et al., 2007).

6.3.3 Parameter estimates

Table III reports sunk cost estimates. The final column provides 95% confidence intervals, all of which exclude zero. The coefficient on the constant term, θ_0 , is bounded between \$475,000 and \$2.375 million. Reflecting the fact that sunk costs scale with the number of strengths per substance-delivery-release combinations, we find that θ_1 is bounded between \$1.0 million and \$2.225 million. Consistent with irregular delivery methods involving substantially higher sunk costs, we find that θ_2 is between \$1.4 MILLION and \$6.8 million. Calculated at the midpoint of each identified set, entry costs range between \$1.425 million and \$14.325 million, depending on the characteristics of the substance-delivery-release combination. The average across drugs, weighting each equally, is \$3.78 million.

Table III: *Sunk cost estimates*

Variable	Parameter	95% confid. interval
Constant	ϕ_0	[0.475,2.375]
Number of strengths	ϕ_1	[1.0,2.225]
Indicator for irregular delivery	ϕ_2	[1.4,6.8]
Observations		220
Moments		10
Minimum sunk cost of entry		1.425
Average sunk cost of entry		3.78
Maximum sunk cost of entry		14.325

Sunk cost estimates are reported in millions of dollars. The bounds reported in the rightmost column are intervals in which the true parameters lie 95% of the time. To calculate the minimum, mean, and maximum sunk costs across the substance-delivery-release combinations, we set θ_0 , θ_1 , and θ_2 equal to their respective midpoints.

Our entry cost estimates align with statements made by agency officials and medical researchers. For example, in 2014, former FDA Commissioner Gottlieb stated, "Filing a generic application requires an average of about \$5 million and can cost as much as \$15 million"—very close to the \$3.8 million mean and \$14.3 million maximum values we report in Table III Gottlieb (2016).⁵⁰ Similarly, our estimates fall

⁴⁹Tests based on inequalities depend on the degree to which the moments are binding. One consequence is that including uninformative moments (i.e., ones that are satisfied for a very wide range of parameters) typically widens confidence intervals. The main innovation of Andrews and Soares (2010) is a procedure for, loosely speaking, deciding which moments are sufficiently uninformative to be discarded. The authors have found moment selection to be incredibly important in other settings. We deemphasize it here only because it does not have a big effect on our results.

⁵⁰Given his November 20, 2014 Senate testimony, we take this to mean "the cost of bringing a drug to market." Note that while our estimates are lower than his, we expect this. Our figures pertain mainly to "solid dose" drugs, whereas his figures also cover injectables. Injectable drugs require higher setup costs because they must be specially formulated and packaged to inhibit bacteria

squarely within the bounds reported by Scott Morton (1999), who surveyed FDA officials around 1999 and found that sunk costs range from \$382,000 to \$31 million (in 2019 dollars, i.e., \$250,000 to \$20 million in the original text). Even more to the point, Baker-Smith et al. (2008) ran six bioequivalence studies between 1997 and 2004 and documented their per-study expenditures.⁵¹ The authors spent between \$807,000 and \$1.25 million per study (in 2019 constant dollars), which aligns with our estimate of per-strength entry costs.⁵²

7 Testing Assumptions

Throughout the last two sections, we have assumed that once a cartel forms, members set collusive prices and nonmembers best respond in all subsequent periods.⁵³ In this section, we test our assumptions. We first formally quantify the incentives of cartel members to collude, relying on definitions and methods proposed by Igami and Sugaya (2022). Following Backus et al. (2021), we then test whether the data is consistent with our conduct assumptions.

7.1 Incentives to collude

To accurately characterize incentives to comply, we must pay special attention to timing. Throughout the estimation of the model and simulation of counterfactual outcomes, we assume each period lasts one year. This level of aggregation is common in the literature and reflects limitations imposed by data and computational costs. However, in our setting, cartel members can detect and, potentially retaliate against, noncompliance *very* quickly.⁵⁴ For instance, in 2013, Glenmark undercut Teva in one of the cartelized markets, apparently due to miscommunication. The response was almost immediate. A confused Teva employee sent an email whose only contents were "???" to NP, and exactly five minutes later, NP replied, "I know...made the call already." The next day NP spoke to her counterpart at Glenmark, who rescinded the offer (Complaint, page 134).⁵⁵

To reflect the speed at which punishment would occur without complicating notation throughout the rest of the paper, we maintain that all changes in market structure and environment occur annually, but we let cartel members punish noncompliance within two weeks.⁵⁶ We again note that detection is described as nearly instantaneous in the complaint and hypothetical retaliation could quickly follow. Formally, we assume that first change decisions are made instantaneously and that the second stage of each period has S sub-stages of equal duration, which are indexed by s . In this subsection only, we also assume that in every

growth.

⁵¹Putting aside factory setup, pharmacological studies are the main reason entry costs depend on the number of strengths, so they are an especially good benchmark for our estimates of the coefficient on r .

⁵²Note that while the range Baker-Smith et al. (2008) gives is centered around the lower bound of our confidence interval, our figure also accounts for the equipment required to produce the drug.

⁵³Recall firms expect incumbents to respond "rationally" to changes in market structure. For instance, the complaint describes firms "playing nice in the sandbox" as "responsible" or "rational" competitors (Complaint, page 42). By contrast, the complaint does not explicitly discuss punishments. We assume the simplest model of Nash reversion.

⁵⁴Decentralized letting of contracts allows for aggressive bidding for subsequent customers. For existing customers, two dynamics are important. First, requests for bids are often erratic. Second, any unexpected changes in the market, ranging from pricing changes to a new entrant, can prompt a request for new bids.

⁵⁵This event also speaks to how gravely members treated noncompliance. By withdrawing an attractive price, Glenmark upset the buyer, who was a very large customer.

⁵⁶As we show below, incentive compatibility constraints are met in our data. Nearly all exhibit significant slack. When the detection/punishment period is as long as two months, a small number of constraints are violated for extreme draws of ζ and ω .

sub-stage, firms set prices while buyers draw new demand shocks and make purchase decisions.⁵⁷ We emphasize that because all determinants of market structure, demand, and cost are fixed throughout the year, this assumption does not affect any other analysis in the paper.

We assume that prices charged in one sub-period are observed by the time the next one starts. The history of play at period t , stage 2, and sub-stage s is then given by

$$h^{ts} = \left(\left(p_{df\tau\psi}, p_{df\tau\psi}^{C,M} \right)_{f \in \mathcal{M}_{d\tau}, t_d \leq \tau < t, 1 \leq \psi \leq S'} \left(p_{df\tau\psi}, p_{df\tau\psi}^{C,M} \right)_{f \in \mathcal{M}_{d\tau}, 1 \leq \psi \leq s-1} \right), \quad (21)$$

If there exists a $\tau \leq t-1$ or there exists $\tau = t$ and $\psi \leq s-1$ such that $p_{df\tau\psi} \neq p_{df\tau\psi}^{C,M}$, then we call the cartel "noncompliant." Otherwise, we call it "compliant."

In any period $t \geq t_d$ and sub-stage s when the cartel is stable, for any number of members and nonmembers that might have entered by that time and any vectors of demand and marginal cost shocks, the value to member f of complying with the cartel agreement is equal to cartel profits today plus the continuation value associated with the cartel existing in the next period:

$$\begin{aligned} V_{df\tau s}^C(E_m, E_n, \zeta, \omega) &= \frac{1}{S} \sum_{\psi=0}^{S-s} \tilde{\delta}^\psi \pi_{d,f,t}^{C,M}(M_{d,t}+E_m, N_{d,t}+E_n, \zeta, \omega) \\ &+ \tilde{\delta}^s \sum_{\tau=1}^{\infty} \delta^\tau F_D(\tau) \sum_{e_M=0}^{\chi_{M_d}-E_m} \sum_{e_N=0}^{\chi_{N_d}-E_n} \rho(\chi_{M_d}-E_m, e_M, \chi_{N_d}-E_n, e_N, \tau) \\ &\times \int_{\zeta', \omega'} \pi_{d,f,t+\tau}^{C,M}(M_{d,t+\tau}+E_m+e_m, N_{d,t+\tau}+E_n+e_n, \zeta', \omega') dF_{\zeta'} dF_{\omega'}, \end{aligned} \quad (22)$$

where $\tilde{\delta} \equiv \delta^{1/S}$. The value of noncompliance is the value of not complying today plus the expectation of profits given Nash reversion:

$$\begin{aligned} V_{df\tau s}^D(E_m, E_n, \zeta, \omega) &= \frac{1}{S} \left[\pi_{d,f,t}^D(M_{d,t}+E_m, N_{d,t}+E_n, \zeta, \omega) + \sum_{\psi=2}^{S-s} \tilde{\delta}^\psi \pi_{d,f,t}^B(M_{d,t}+E_m, N_{d,t}+E_n, \zeta, \omega) \right] \\ &+ \tilde{\delta}^s \sum_{\tau=1}^{\infty} \delta^\tau F_D(\tau) \sum_{e_M=0}^{\chi_{M_d}-E_m} \sum_{e_N=0}^{\chi_{N_d}-E_n} \rho(\chi_{M_d}-E_m, e_M, \chi_{N_d}-E_n, e_N, \tau) \\ &\times \int_{\zeta', \omega'} \pi_{d,f,t+\tau}^B(M_{d,t+\tau}+E_m+e_m, N_{d,t+\tau}+E_n+e_n, \zeta', \omega') dF_{\zeta'} dF_{\omega'}. \end{aligned} \quad (23)$$

The cartel is stable at (t, s) if

$$\min_{(\tau, \psi): \tau > t | (\tau=t, \psi \geq s)} \left\{ \min_{E_M, E_N} \left\{ \min_{\zeta, \omega} \left\{ \min_{f \in \mathcal{M}_{dt}} \left\{ V_{df\tau\psi}^C(E_m, E_n, \zeta, \omega) - V_{df\tau\psi}^D(E_m, E_n, \zeta, \omega) \right\} \right\} \right\} \right\} \geq 0. \quad (24)$$

Inequality 24 ensures that all members have an incentive to comply at the current time and at any time in the future, regardless of what delays, demand shocks, and marginal cost shocks are realized.

We follow Igami and Sugaya (2022) and evaluate cartel stability in the data using the following procedure.

⁵⁷In the model, buyers are sufficiently small that integrating out over redrawn (ϵ, ζ) does not affect our results.

For each period τ , sub-stage ψ , and cartelized drug d , we enumerate every possible combination of member and nonmember entrants that might have entered by that time, denoted by E_m and E_n , respectively.⁵⁸ Then, for each combination, we draw a large number (1,000) of vectors of demand and marginal cost disturbances, denoted by ζ and ω , respectively. Next, we calculate the lower envelope of the difference between $V_{df\tau\psi}^C(E_m, E_n, \zeta, \omega)$ and $V_{df\tau\psi}^D(E_m, E_n, \zeta, \omega)$ over members, draws of the disturbances, and combinations of entrants. If the value is positive for all $\tau > t_d$ and for all $\psi \geq s$ given $\tau = t$, then the cartel is stable—trigger strategies form an equilibrium. Finally, we plot the distribution of the lower envelope across drugs over time.

Figure IV reports the result following cartel formation. Since the least-slack ICC is satisfied, all cartels are stable after being formed. As expected, incentives to collude fall 3 years after cartel formation, when entrants begin to arrive. Notice the decline is abrupt at $t = 2$. The pattern over time reflects three features of the calculation: the plot reflects minimum values over possible combinations of member and nonmember entrants, the least-slack ICC always corresponds to all entrants arriving as soon as possible, and the cumulative mass function of delays is exactly zero for the first two years following cartel formation.

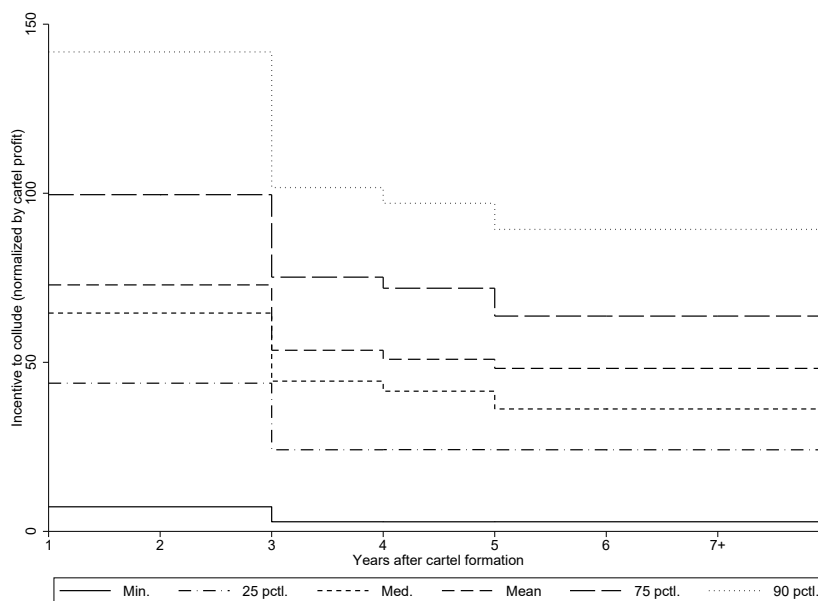


Figure IV: *Generic drug cartels are stable*

In this figure, we plot the distribution of incentives to collude across drugs along the y-axis and the number of years since cartel formation on the x-axis. To compute the incentives, we calculate the difference between the value of complying and not complying with the cartel agreement, divide by cartel profit, and obtain the lower envelope over all possible market structures and draws of the demand and marginal cost shocks (as described in the text).

Note that while Figure IV suffices to show stability, ICCs are likely satisfied by even greater margins. For example, we abstract away from multi-market contact, i.e., trigger strategies in which a deviation in *any market* is punished by reversion to Nash-Bertrand in *all markets*. Yet, members' statements published in the complaint can easily be interpreted as concerns over cross-market punishment, which increase

⁵⁸A trivial number of delays are shorter than 3 years, so we set $F_D(2) \equiv 0$.

incentives to collude.⁵⁹⁶⁰ Also, note that while the alternative policies that we study in Section 8, which reduce regulatory costs and delays, result in more entry, the least-slack ICC is still satisfied under the counterfactual market structures generated by those policies.

7.2 Conduct testing

We now test whether the conduct we observe in the data is consistent with the assumptions of our model. To do so, we rely on an intuitive test proposed by Backus et al. (2021) which adapts the non-nesting framework proposed by Rivers and Vuong (2002) to exclusion restrictions suggested by Berry and Haile (2014). The test provides pairwise comparisons but does not require that either behavioral model is correctly specified. It is based on the idea that if conduct is correctly specified, then variables that determine markups but do not influence marginal costs will be uncorrelated with marginal cost disturbances implied by that model.

Recall our maintained assumptions: all firms set competitive prices in uncartelized markets; all firms set competitive prices in cartelized markets before cartel formation; and after cartel formation, members set prices that maximize their joint profits while nonmembers best respond. We conduct four tests. In each, we arbitrarily number our "baseline" model of conduct "1" and the comparison model "2."

In Tests A and B, we assess whether, following cartel formation, members set prices that *maximize joint profit*. Model 2 is identical to Model 1 except that, after a cartel forms, members set prices that *maximize individual profits* in Model 2. To abstract away from the effect of the government's investigation, Test A omits any observations after 2015. To study post-investigation behavior, Test B omits the period between 2013 and 2015. In Test C, we assess whether nonmembers comply with cartel agreements (i.e., they are "folded into" the cartel). Model 2 is identical to Model 1 except that, after a cartel forms, all firms set prices that maximize the joint profit of all firms (rather than only members setting price that maximizes their joint profit). In Test D, we assess whether member entrants comply with the cartel agreement. That is, Model 2 is identical to Model 1 except that only member incumbents set prices that maximize their joint profit (as opposed to all members setting prices that maximize their joint profit).

Each test involves the following seven steps. First, for Model 1, compute equilibrium markups, denoted by \widehat{mk}_{dft}^1 , subtract them from observed prices to obtain \widehat{mc}_{dft}^1 , and repeat this process for Model 2 to obtain \widehat{mk}_{dft}^2 and \widehat{mc}_{dft}^2 . Second, for Model 1, estimate the empirical analog of equation 11 to obtain $\widehat{\omega}_{dft}^1$, and repeat this calculation for Model 2 to obtain $\widehat{\omega}_{dft}^2$. Third, let z_{dft} denote a vector that includes drug fixed effects, year fixed effects, and an indicator variable that takes a value of one if and only if the complaint alleges the price of d was cartelized in or before year t . Then, regress $\widehat{mk}_{dft}^1 - \widehat{mk}_{dft}^2$ on z_{dft} and obtain predicted values, denoted by $\widehat{g}(z_{dft})$. Fourth, let \tilde{Q}_1 denote the square of the average value of $\widehat{\omega}_{dft}^1 \widehat{g}(z_{dft})$, let \tilde{Q}_2 denote an analogous value for Model 2, and compute \tilde{Q}_1 and \tilde{Q}_2 . Fifth, repeat steps 1-4 on a large number (500) of bootstrapped samples and compute the standard error of the difference between \tilde{Q}_1 and \tilde{Q}_2 , denoted by $\hat{\zeta}$. Sixth, compute the test statistic, which equals $\hat{\zeta}^{-1}(\tilde{Q}_1 - \tilde{Q}_2)$ and is distributed standard normal. The null hypothesis is that both models describe the data equally well. A negative test statistic indicates that the data favors our Model 1, while a positive one indicates that the data favors Model 2.

Table IV reports the results. To interpret them, recall that the test statistic is distributed standard normal and that smaller values imply that the data fits Model 1 better than Model 2. Each test strongly favors

⁵⁹Least-slack constraints reflect asymmetries among members within markets. Cross-market pooling creates ICC slack.

⁶⁰Noncompliance, i.e., cheating, may also involve significant shame and guilt, which increases incentives to comply.

Model 1 in favor of Model 2. Test A supports the government’s claim that defendants conspired to fix prices through at least 2015, while Test B implies collusion continued for the remainder of the sample, albeit perhaps in a tacit form. Test C indicates that nonmembers are not folded into the cartel. The finding is consistent with nonmembers undercutting members; we further note that NP formed cartels in markets where almost all other firms were members. (The test further demonstrates that potential entrants were attracted to supernormal *equilibrium* profits, not simply high prices; consistent with Caoui (2017), entrants price less aggressively when facing the cartel.) Test D implies that members entrants complied with cartel agreements, which is consistent with how carefully NP coordinated the entry of a member entrant into cartelized markets.

Table IV: Results of conduct tests

	$\tilde{Q} \times 100$		Test statistic	p-value
	Baseline	Alternative		
Test A. Model vs. competition (pre-investigation)	.16247	.20688	-3.03	.001
Test B. Model vs. competition (post-investigation)	.20286	.34235	-4.63	< .0001
Test C. Model vs. nonmembers comply	1.3979	4.6533	-5.57	< .0001
Test D. Model vs. member entrants do not comply	.00008	.00008	-3.06	.001

This table reports the results of the testing procedure proposed by Backus et al. (2021) for pairwise comparisons described in the text. The test statistic is distributed standard normal. The standard error of the difference between \tilde{Q}_1 and \tilde{Q}_2 is obtained via bootstrapping.

8 Counterfactual analysis

In this section, we measure the effect of equilibrium entry on the cartelized markets. We then evaluate alternative FDA policies that attract additional entrants and/or permit them to enter faster.

8.1 Effects of equilibrium entry

To assess how important entrants are in cartelized drug markets, we compare the predictions of a model that allows for entry following cartel formation with the predictions of a model that prohibits it. To arrive at the first set of predictions, we let the market structure evolve according to equilibrium entry. For each cartelized market d and year t following cartel formation, we draw a large number (1,000) of vectors of demand and marginal cost shocks, and for each draw, we recompute the equilibrium of the second stage. For the second set of predictions, we replicate this process but remove all products associated with ANDAs filed after cartel formation. This eliminates 207 entrants spread over 69 markets. To summarize our findings, we compute the sum of profits by year and plot the difference between the first and second model.

Panel A of Figure V reports the result in event time, with year zero corresponding to cartel formation. Entrants arrive in markets as early as three years after cartel formation and exert almost immediately

downward pressure on profit, reducing it by an average of \$60 million or about 10%. More arrive over the following four years so that within seven years of cartel formation, entry reduces profit by an average of \$140 million or 23%. Measured at year zero and under the assumption that cartel members continue to comply with the agreement, albeit perhaps tacitly, entry reduces the present value of firms' profits by just under \$1 billion. Moreover, consumer surplus rises more than second stage profit falls, as many buyers switch away from the outside option as prices fall.

The counterfactual exercise also speaks to potential deterrence. Deciding whether to fix prices requires weighing supra-normal profits against coordination costs and legal risks. Entry reduces the supra-normal profits without reducing the risks and coordination costs, so it may deter cartel formation altogether in some settings, even though it did not do so here. To understand the effect it can have on incentives to reach a collusive agreement, we replicate Panel A of Figure V but plot the profit earned by *incumbent cartel members* profit (rather than by *all firms*).

Panel B of Figure V reports this result (note that the scale of the axes has changed). Since the profit measured in Panel A includes profit earned by entrants, which are irrelevant to incumbent cartel members, the profit measured in Panel B declines even more steeply as entrants arrive. Within seven years of cartel formation, entry reduces the amount cartel members earn by nearly half. In light of this sharp drop, it is easy to see how entry might deter some cartels from forming in the first place, even if it was insufficient deterrence in this instance.

8.2 Effects of alternative policies

Our estimates imply that even when entrants face large sunk costs and long regulatory delays, they can still significantly affect equilibrium outcomes. This finding naturally raises the question of whether society would benefit from policies that encourage more entrants and hasten their arrival. We examine two sets of policy changes. First, we reduce the cost of entry by between \$150,000 and \$600,000, a 4-16% decline for the average substance-delivery-release combination. Second, we reduce regulatory delays by between 0.5 and 2 years.

It is within the FDA's control to implement these changes. For instance, the agency could dramatically reduce the sunk cost of entry by cutting fees, or with other actions that maintain the revenue income but still reduce the costs of entry.⁶¹ Recent proposals include publishing "best practices for manufacturers in submitting an ANDA" and "enhanc[ing] communication between manufacturers and FDA" (Waxman et al., 2018). The agency could also reallocate staff to prioritize timely reviews. Notably, the agency has experimented with expedited reviews, proving that fast approval is feasible (presumably without compromising the integrity of the process). For instance, the "Prioritization of the Review of Original ANDAs, Amendments, and Supplement" policy, introduced in 2020, targets 4-8 month approval for certain filings.

These policy alternatives map to changes in θ_0 or F_D . For each of them, we adjust the primitives, simulate market outcomes, and report how these differ from the status quo. We recompute equilibrium entry in the first stage, equilibrium prices in the second stage, and integrate out over the distributions of demand and marginal cost disturbances and delays. Multiple equilibria may arise in the first stage, so our

⁶¹In 2022, manufacturers must pay a base rate of \$1.5 million per year, \$200,000 per factory per year, \$225,712 to file an ANDA, and another \$75,000 per drug to ensure the details of production remain confidential.

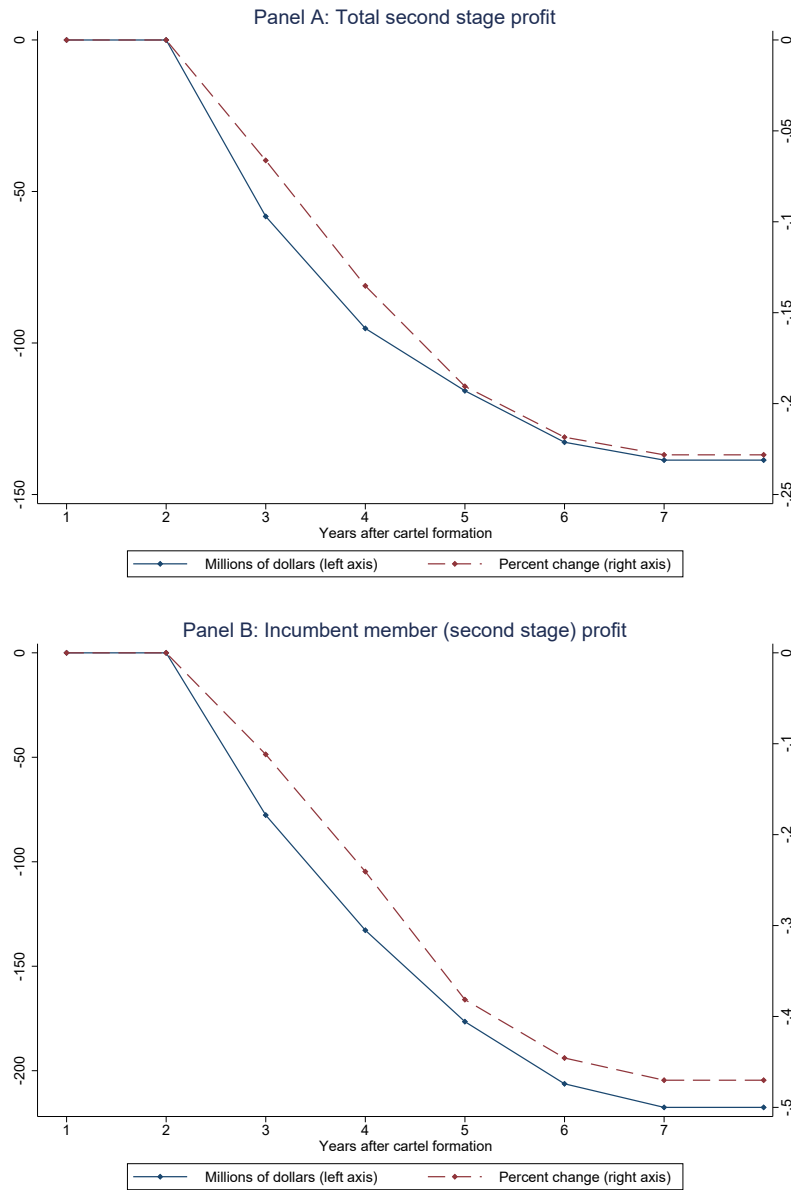


Figure V: Effects of equilibrium entry on firms' profit

This figure reports the difference in profit due to observed entry on the y-axes against the number of years since cartel formation on the x-axes. Primary y-axes measure dollar-denominated differences, while secondary y-axes measure percentage point differences (relative to profit earned under the no-entry model). Panel A reflects profits earned by all firms, whereas Panel B reflects profits earned by incumbent cartel members.

simulations must select among them. We obtain a unique Nash equilibrium by assuming that the ratio of member to nonmember entrants is as close as possible to the ratio observed in the data following cartel formation.

Table V reports the result. Reducing the sunk costs of entry attracts as many as 90 new entrants.

Compensating variation rises by at most \$179 million and profits decline by at most \$297 million. Assuming fees account for one-half of entry cost reductions, FDA revenue falls by between around \$25-140 million. For instance, with a \$600,000 cost reduction, additional entry raises its revenue by \$22 million, but fee cuts reduce FDA revenue by \$160 million.

Table V: *Effects of reducing entry costs and delays*

	Sunk costs reduced by				Delays shortened by			
	\$150,000	\$300,000	\$450,000	\$600,000	0.5 years	1 year	1.5 years	2 years
Mean sunk cost	4.19	4.04	3.89	3.74	4.34	4.34	4.34	4.34
Mean delay	5.24	5.24	5.24	5.24	4.74	4.24	3.74	3.24
Drugs with additional entry	14	15	24	31	14	18	20	23
Additional entrants	13	47	70	90	17	36	52	60
Change in agency revenue (I)	-39.9	-79.8	-119.7	-159.6	-	-	-	-
Change in agency revenue (II)	5.72	26.27	30.49	21.58	10.5	32.07	47.71	55.2
Change in total profits	-28.28	-165.03	-248.79	-296.94	-157.6	-339.24	-571.78	-718.45
Consumer comp. variation	12.73	77.87	134.78	178.55	270.88	558.73	1019.39	1338.75

Columns correspond to distinct counterfactual policy experiments. Columns 1-4 correspond to reducing the sunk entry costs; columns 4-8 report the result of shortening entry delays. "Change in agency revenue (I)" represents lost revenue due to fee reductions, which we assume account for one-half of entry cost reductions. "Change in agency revenue (II)" represents revenue gained from additional filing fees. We assume that the FDA collects one-quarter of entry costs, less fee reductions.

Bringing products to market faster has much greater overall effects. Even a half-year reduction in delays produces roughly the same changes in profits and consumer compensating variation as a \$600,000 reduction in sunk costs. From the perspective of patients and payers, two mechanisms merit consideration. First, when nonmembers arrive, they undercut the cartel in an effort to steal share, and the cartel responds by also lowering prices, so short delays mean fewer periods with monopoly prices. Second, shorter delays allow entrants to recoup their sunk investments in entry faster, which encourages additional entry.

Equally striking, the magnitudes of the changes scale roughly linearly with the duration of the reduction. Bringing products to market two years faster cuts industry profit by over \$718 million and equates to \$1.3 billion in consumer compensating variation. The extent of additional entry is on par with a \$600,000 cost reduction, highlighting the relative importance of the first mechanism: it is essential that monopoly-like prices charged by the cartel are disciplined as quickly as possible.

9 Conclusion

We investigated the likelihood and impact of entry in the context of the largest price-fixing case in US history, which involved US generic drugs. Our data suggest that cartel formation successfully raised prices sharply but also attracted significant entry, even though entrants faced regulatory delays of several years. The facts of the case and patterns in the data indicate that cartel members were folded into existing cartels while nonmembers were not.

To accurately forecast how market participants would behave under alternative behavioral assumptions

and regulatory policies, we estimated a structural model that allows firms to endogenously enter and set prices, and we used the resulting estimates to simulate equilibrium outcomes. We found that cartel members have strong incentives to comply with the collusive agreements, which may partly explain why we also found that price fixing persisted long after the government launched its investigation. Our estimates confirm that entry is not just slow but also expensive, with sunk entry costs ranging from \$1.4 million to \$14.3 million per substance-delivery-release combination.

We simulated counterfactual outcomes absent entry and found that cartel profits would have been 27% higher over our sample period absent entry. This means that entry protected consumers from facing even higher prices, even if it was incomplete and delayed. Since policymakers often have considerable discretion over costs and delays, we also simulated counterfactual outcomes under alternative regulatory environments. Our results indicate that reducing either entry costs or delays can improve consumer surplus, although the latter is especially effective.

References

- Barriers to entry, brand advertising, and generic entry in the us pharmaceutical industry. International Journal of Industrial Organization, 18(7):1085 – 1104, 2000.
- R. Abelson and K. Thomas. Fed up with drug companies, hospitals decide to start their own. The New York Times, 18, 2018.
- D. W. Andrews and G. Soares. Inference for parameters defined by moment inequalities using generalized moment selection. Econometrica, 78(1):119–157, 2010.
- J. Asker. A study of the internal organization of a bidding cartel. American Economic Review, 100(3): 724–62, June 2010.
- M. Backus, C. Conlon, and M. Sinkinson. Common ownership and competition in the ready-to-eat cereal industry. Technical report, National Bureau of Economic Research, 2021.
- C. M. Baker-Smith, D. K. Benjamin Jr, H. G. Grabowski, E. D. Reid, B. Mangum, J. V. Goldsmith, M. D. Murphy, R. Edwards, E. L. Eisenstein, J. Sun, et al. The economic returns of pediatric clinical trials of antihypertensive drugs. American Heart Journal, 156(4):682–688, 2008.
- E. R. Berndt, R. M. Conti, and S. J. Murphy. The generic drug user fee amendments: an economic perspective. Journal of Law and the Biosciences, 5(1):103–141, 2018.
- S. T. Berry. Estimating discrete-choice models of product differentiation. The RAND Journal of Economics, pages 242–262, 1994.
- S. T. Berry and P. A. Haile. Identification in differentiated products markets using market level data. Econometrica, 82(5):1749–1797, 2014.
- S. T. Berry and J. Waldfogel. Public radio in the united states: does it correct market failure or cannibalize commercial stations? Journal of Public Economics, 71(2):189–211, 1999.

- S. Borenstein. Hubs and high fares: dominance and market power in the us airline industry. The RAND Journal of Economics, pages 344–365, 1989.
- T. F. Bresnahan and P. C. Reiss. Entry and competition in concentrated markets. Journal of political economy, 99(5):977–1009, 1991.
- D. P. Byrne and N. de Roos. Learning to coordinate: A study in retail gasoline. American Economic Review, 109(2):591–619, February 2019.
- E. Caoui. A study of umbrella damages from bid-rigging. Working Paper, 2017.
- Cardinal Health. Manufacturer reference manual: For manufacturers to the pharmaceutical segment distribution supply chain. Technical report, 2020 2020. URL <https://www.cardinalhealth.com/content/dam/corp/web/documents/Manual/cardinalhealth-pharma-supplier-guidebook.pdf>.
- V. Chernozhukov, H. Hong, and E. Tamer. Estimation and confidence regions for parameter sets in econometric models 1. Econometrica, 75(5):1243–1284, 2007.
- R. Clark, D. Coviello, J.-F. Gauthier, and C. U. Shneyerov. Bidding rigging and entry deterrence in public procurement: Evidence from an investigation into collusion and corruption in quebec. The Journal of Law, Economics, and Organization, 34(3), 2018.
- R. Clark, C. A. Fabiilli, L. Lasio, et al. Collusion in the US generic drug industry. Technical report, 2021.
- E. Cuddy. Competition and collusion in the generic drug market. Technical report, 2020.
- N. De Roos. Examining models of collusion: The market for lysine. International Journal of Industrial Organization, 24(6):1083–1107, 2006.
- L. Einav. A note on the analogies between empirical models of auctions and of differentiated product markets. Technical report, August 2003.
- C. Fershtman and E. Muller. Capital investments and price agreements in semicollusive markets. The Rand Journal of Economics, pages 214–226, 1986.
- C. Fershtman and A. Pakes. A dynamic oligopoly with collusion and price wars. The RAND Journal of Economics, 31(2):207–236, 2000.
- S. Ganapati and R. McKibbin. Markups and fixed costs in generic and off-patent pharmaceutical markets. Technical report, National Bureau of Economic Research, 2021.
- S. Gottlieb. How Obama’s FDA Keeps Generic Drugs Off the Market. Missouri Medicine, 113(6):444–445, 2016.
- E. J. Green and R. H. Porter. Noncooperative collusion under imperfect price information. Econometrica, 52(1):87–100, 1984.
- M. Grennan and R. J. Town. Regulating innovation with uncertain quality: information, risk, and access in medical devices. American Economic Review, 110(1):120–61, 2020.

- J. E. Harrington. Noncooperative behavior by a cartel as an entry-deterring signal. The RAND Journal of Economics, pages 426–433, 1984.
- J. E. Harrington. Collusion among asymmetric firms: The case of different discount factors. International Journal of Industrial Organization, 7(2):289 – 307, 1989a.
- J. E. Harrington. Collusion and predation under (almost) free entry. International Journal of Industrial Organization, 7(3):381–401, 1989b.
- J. E. Harrington, K. Hüschelrath, and U. Laitenberger. Rent sharing to control noncartel supply in the german cement market. Journal of Economics & Management Strategy, 27(1):149–166, 2018.
- K. Ho. Insurer-provider networks in the medical care market. American Economic Review, 99(1):393–430, 2009.
- M. Igami. Market power in international commodity trade: The case of coffee. The Journal of Industrial Economics, 63(2):225–248, 2015.
- M. Igami and T. Sugaya. Measuring the incentive to collude: The vitamin cartels, 1990–99. The Review of Economic Studies, 89(3):1460–1494, 2022.
- J. Ishii. Compatibility, competition, and investment in network industries: ATM networks in the banking industry. Unpublished working paper, 2005.
- M. Levenstein. Mass production conquers the pool: firm organization and the nature of competition in the nineteenth century. The Journal of Economic History, 55(3):575–611, 1995.
- M. C. Levenstein and V. Y. Suslow. What determines cartel success? Journal of economic literature, 44(1): 43–95, 2006.
- M. C. Levenstein and V. Y. Suslow. Breaking up is hard to do: Determinants of cartel duration. The Journal of Law and Economics, 54(2):455–492, 2011.
- M. J. Mazzeo. Product choice and oligopoly market structure. RAND Journal of Economics, pages 221–242, 2002.
- N. Miller and G. Sheu. Quantitative methods for evaluating the unilateral effects of mergers. 2020.
- N. H. Miller. Modeling the effects of mergers in procurement. International Journal of Industrial Organization, 37:201–208, 2014.
- N. H. Miller and M. C. Weinberg. Understanding the Price Effects of the MillerCoors Joint Venture. Econometrica, 85(6):1763–1791, 2017.
- C. Mulligan. Peltzman revisited: Quantifying 21st century opportunity costs of fda regulation. Technical report, National Bureau of Economic Research, 2021.
- A. Pakes, J. Porter, K. Ho, and J. Ishii. Moment inequalities and their application. Econometrica, 83(1): 315–334, 2015.

- M. Pazniokas. Drug price-fixing lawsuit pushes CT probe into national spotlight. CT Post, May 2019.
- J. M. Podolny and F. M. Scott Morton. Social status, entry and predation: The case of british shipping cartels 1879–1929. The Journal of Industrial Economics, 47(1):41–67, 1999.
- R. H. Porter. A study of cartel stability: The joint executive committee, 1880-1886. The Bell Journal of Economics, 14(2):301–314, 1983.
- J. L. Powell. Symmetrically trimmed least squares estimation for tobit models. Econometrica, 54(6): 1435–1460, 1986.
- D. Rivers and Q. Vuong. Model selection tests for nonlinear dynamic models. The Econometrics Journal, 5 (1):1–39, 2002.
- E. Rosenthal. Rapid price increases for some generic drugs catch users by surprise. New York Times, July 2014.
- F. Scott Morton. Entry and predation: British shipping cartels 1879–1929. Journal of Economics & Management Strategy, 6(4):679–724, 1997.
- F. Scott Morton. Entry decisions in the generic pharmaceutical industry. The RAND Journal of Economics, 30(3):421–440, 1999.
- F. Scott Morton and L. Boller. Enabling competition in pharmaceutical markets. Hutchins Center Working Paper, (30), 2017.
- K. Seim. An empirical model of firm entry with endogenous product-type choices. The RAND Journal of Economics, 37(3):619–640, 2006.
- G. J. Stigler. A theory of oligopoly. The Journal of Political Economy, 72(1):44–61, 1964.
- L. Sun and S. Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics, 2020.
- G. Symeonidis. In which industries is collusion more likely? evidence from the uk. The journal of industrial economics, 51(1):45–74, 2003.
- E. Tamer. Incomplete simultaneous discrete response model with multiple equilibria. The Review of Economic Studies, 70(1):147–165, 2003.
- H. Vasconcelos. Entry effects on cartel stability and the joint executive committee. Review of Industrial Organization, 24(3):219–241, 2004.
- H. Waxman, B. Corr, K. Martin, and S. Duong. What commissioner gottlieb’s fda is doing to lower prescription drug prices and steps congress can take to help. Technical report, The Commonwealth Fund, 2018.
- B. W. Wendling, D. R. Schmidt, M. Hertzendorf, L. M. Olson, H. Hadeishi, K. Cagatay, G. Kuczura, E. Connelly, J. Towey, K. A. Goldman, W. E. Cohen, K. A. Grimm, C. N. Bryan, and R. A. Nelson Jr. Authorized generic drugs: Short-term and long-term impact. Technical report, Federal Trade Commission, 2011.

T. G. Wollmann. Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. American Economic Review, 108(6):1364–1406, 2018.

A [For Online Publication] Data appendix

A.1 Sample construction

Table A.1 describes the construction of our sample. Table A.2 lists cartel members and their "quality," as assigned by NP. Figure A.1 presents a histogram of assigned qualities. Table A.3 summarizes the differences between cartelized and uncartelized markets.

Table A.1: *Sample construction*

Total number of generic drugs approved prior to 2008:	3722
<i>Teva does not participate in the market in 2013q1.</i>	-2627
<i>Teva does not participate in the market just prior to hiring of NP (i.e., in 2013q1).</i>	-339
<i>First generic launches during sample period.</i>	-138
<i>Complaint indicates KG and/or DR may have affected prior to NP joining Teva.</i>	-62
<i>Injectible, dental, shampoo, suppository, or aerosol.</i>	-60
<i>Complaint is ambiguous with respect to alleged conduct.</i>	-28
<i>Exceptionally high price due to ongoing/potential litigation.</i>	-28
<i>Complaint alleges two price increases.</i>	-23
<i>Particular strength not sold in any meaningful quantity.</i>	-1
Total number of generic drug markets in the sample:	= 416

The drugs with "exceptionally high price due to ongoing/potential litigation" include (a) tretinoin/isotretinoin, which are Vitamin A derivatives including Accutane that were facing litigation due to certain birth defects, (b) methotrexate, and (c) immune system suppressants such as cyclosporine.)

Table A.2: Firm quality and cartel membership.

Firm	Original quality (as of 2013q2)	Updated quality (as of 2014q2)	Is a cartel member (as of 2014q2)?
Actavis/Watson	3	3	Yes
Amneal	1	2	Yes
Apotex	-3	2	Yes
Breckenridge	1	2	Yes
Glenmark	3	2	Yes
Greenstone	0	1	Yes
Heritage	0	3	Yes
Lupin	2	3	Yes
Mylan	3	3	Yes
Par	1	2	Yes
Rising	1	2	No
Sandoz	3	3	Yes
Taro	3	3	Yes
Upsher Smith	2	2	Yes
Versapharm	-2	-2	No
Zydus	-3	2	Yes

The table is restricted to firms explicitly assigned a score greater than -3 by NP. The other firms include Accord, Acella, Aci Healthcare, Afaxys, Ajanta, Akorn, Akron, Alembic, Alkem, Almus, Alvogen, American Antibiotics, Amici, Ani, Apace Ky, Apnar, Arbor, Ascend, Aurobindo, Avpak, Banner, Bausch, Bayshore, Belcher, Biocon, Biomes, Bionpharma, Blu, Bristol, Burel, Cadista, Camber, Cambridge Therapeutic Tech, Cameron, Carlsbad Tech, Cipla, Citron, Clay Park Lab, Corepharma, Cosette, Daiichi, Dash, Dava, Edenbridge, Endo, Epic, Ethex, Excellium, Exelan, Eywa, Gavis, Gen Source Rx, Genbiopro, Glaxosmithkline, Granules, Granules India, Gw, Hi Tech, Hikma, Impax, Ingenus, Int Medication Systems, Invagen, Inwood Lab, Ipca Lab, Johnson Johnson, Jsj, Kremers Urban, Kvk Tech, Lannet, Lannett, Larken, Laurus Lab, Leading, Lineage Therapeutic, Macleods, Major, Mallinckrodt, Marathon, Mayne, Medimetriks, Medstone, Megalith, Metcure, Method, Micro Lab, New Horizon Rx Group, Nivagen, Nostrum, Novitium, Orchid, Osmotica, Pack, Patriot, Pernix Therapeutic, Perrigo, Pharm Assoc, Pharmacist, Polygen, Prasco, Precision Dose, Princeton, Puracap, Puracap Lab, Purdue, Quagen, Quinn, Reddy, Roxane, Sagent, Sanofi, Sciegen, Sigmapharm Lab, Silarx, Sky Packaging, Solco, Sti, Strides, Sun, Sunrise, Tagi, Time Cap Lab, Torrent, Tris, Trupharma, Unichem, Valeant, Vensun, Vertical, Vertical Trigen, Viona, Virtus, Vista, Warner Chilcott, Westminster, Wilshire, Winthrop, Wockhardt, Woodward Services, X Gen, Xiromed, and Yiling.

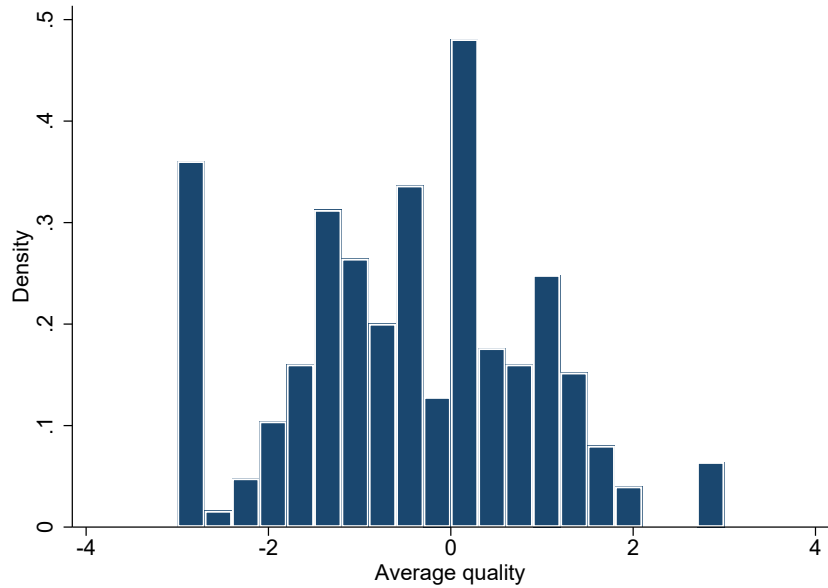


Figure A.1: Histogram of average quality. The unit of observation is a drug market. (For instance, a "3" means NP was friends with all other rivals.)

Table A.3: Balance Table

	Cartelized	Uncartelized	Difference	
Price	27.38 (1.15)	33.86 (0.91)	-6.49 (0.74)	1.00
Quantity (in thousands)	870.78 (58.10)	1572.71 (73.26)	-701.93 (56.06)	1.00
Expenditure (in millions)	16.09 (1.24)	26.96 (1.36)	-10.87 (1.05)	1.00
Number of ANDA filings	0.18 (0.02)	0.29 (0.02)	-0.11 (0.01)	1.00
Number of firms	3.76 (0.07)	4.36 (0.06)	-0.60 (0.05)	1.00

The unit of observation is drug-year across all observations between 2008 and 2012, inclusive. There are 113x5 observations for the cartelized drugs and 303x5 observations for the uncartelized drugs. Standard errors of the means are given in parentheses.

The set of cartelized drugs include Amoxicillin Clavulanate Potassium tablet chewable oral in 2 strengths,

Azithromycin for suspension oral in 2 strengths, Baclofen tablet oral in 2 strengths, Bethanechol Chloride tablet oral in 4 strengths, Bumetanide tablet oral in 3 strengths, Carbamazepine tablet chewable oral in one strength, Carbamazepine tablet oral in one strength, Cefdinir capsule oral in one strength, Cefdinir for suspension oral in 2 strengths, Cefprozil tablet oral in 2 strengths, Cephalexin for suspension oral in 2 strengths, Ciprofloxacin Hydrochloride tablet oral in 3 strengths, Clarithromycin ER tablet oral in one strength, Clotrimazole solution topical in one strength, Cyproheptadine Hydrochloride tablet oral in one strength, Desmopressin Acetate tablet oral in 2 strengths, Dicloxacillin Sodium capsule oral in 2 strengths, Diflunisal tablet oral in one strength, Disopyramide Phosphate capsule oral in 2 strengths, Doxazosin Mesylate tablet oral in 4 strengths, Estazolam tablet oral in 2 strengths, Estradiol tablet oral in 3 strengths, Ethosuximide capsule oral in one strength, Ethosuximide syrup oral in one strength, Etodolac tablet oral in 2 strengths, Etodolac ER tablet oral in 3 strengths, Fluconazole tablet oral in 4 strengths, Fluoxetine Hydrochloride tablet oral in one strength, Flurbiprofen tablet oral in one strength, Flutamide capsule oral in one strength, Glimepiride tablet oral in 3 strengths, Griseofulvin Microsize suspension oral in one strength, Hydroxyurea capsule oral in one strength, Hydroxyzine Pamoate capsule oral in 2 strengths, Isoniazid tablet oral in 2 strengths, Ketoconazole cream topical in one strength, Ketoconazole tablet oral in one strength, Loperamide Hydrochloride capsule oral in one strength, Medroxyprogesterone Acetate tablet oral in 3 strengths, Moexipril Hydrochloride tablet oral in 2 strengths, Nabumetone tablet oral in 2 strengths, Nadolol tablet oral in 3 strengths, Nortriptyline Hydrochloride capsule oral in 4 strengths, Nystatin tablet oral in one strength, Oxybutynin Chloride tablet oral in one strength, Penicillin V Potassium tablet oral in 2 strengths, Pentoxifylline ER tablet oral in one strength, Pravastatin Sodium tablet oral in 4 strengths, Prochlorperazine Maleate tablet oral in 2 strengths, Propranolol Hydrochloride tablet oral in 2 strengths, Ranitidine Hydrochloride tablet oral in 2 strengths, Sotalol Hydrochloride tablet oral in 3 strengths, Tamoxifen Citrate tablet oral in one strength, Theophylline ER tablet oral in one strength, and Warfarin Sodium tablet oral in 9 strengths .

The set of uncarterized drugs include Acetaminophen Codeine Phosphate tablet oral in 3 strengths, Acyclovir capsule oral in one strength, Acyclovir tablet oral in 2 strengths, Albuterol Sulfate syrup oral in one strength, Alendronate Sodium tablet oral in 4 strengths, Amiodarone Hydrochloride tablet oral in one strength, Amlodipine Besylate tablet oral in 3 strengths, Amoxicillin capsule oral in 2 strengths, Amoxicillin for suspension oral in 4 strengths, Amoxicillin tablet chewable oral in 2 strengths, Amoxicillin tablet oral in 2 strengths, Amoxicillin Clavulanate Potassium for suspension oral in one strength, Amoxicillin Clavulanate Potassium tablet oral in 2 strengths, Mixed Amphetamine Salt (long name) tablet oral in 7 strengths, Anagrelide Hydrochloride capsule oral in 2 strengths, Atenolol tablet oral in 3 strengths, Azithromycin tablet oral in 3 strengths, Benazepril Hydrochloride tablet oral in 4 strengths, Benzonatate capsule oral in one strength, Benztrapine Mesylate tablet oral in 3 strengths, Bisoprolol Fumarate tablet oral in 2 strengths, Bupropion Hydrochloride ER tablet oral in one strength, Cabergoline tablet oral in one strength, Calcitriol capsule oral in 2 strengths, Carbidopa Levodopa tablet oral in 3 strengths, Carvedilol tablet oral in 4 strengths, Cefadroxil Cefadroxil Hemihydrate capsule oral in one strength, Cefadroxil Cefadroxil Hemihydrate tablet oral in one strength, Cephalexin capsule oral in 2 strengths, Cephalexin tablet oral in 2 strengths, Chlordiazepoxide Hydrochloride capsule oral in 3 strengths, Chlorzoxazone tablet oral in one strength, Ciclopirox solution topical in one strength, Cilostazol tablet oral in 2 strengths, Citalopram Hydrobromide tablet oral in 3 strengths, Clarithromycin tablet oral in 2 strengths, Clindamycin Hydrochloride capsule oral in 2 strengths, Clomiphene Citrate tablet oral in one strength, Clonazepam

tablet oral in 3 strengths, Clonazepam tablet orally disintegrating oral in 5 strengths, Clozapine tablet oral in 4 strengths, Cromolyn Sodium solution drops ophthalmic in one strength, Cromolyn Sodium solution inhalation in one strength, Cyclobenzaprine Hydrochloride tablet oral in one strength, Danazol capsule oral in 2 strengths, Dexmethylphenidate Hydrochloride tablet oral in 3 strengths, Dextroamphetamine Sulfate ER capsule oral in 3 strengths, Dextroamphetamine Sulfate tablet oral in 2 strengths, Diazepam tablet oral in 3 strengths, Diclofenac Sodium ER tablet oral in one strength, Dipyridamole tablet oral in 3 strengths, Disulfiram tablet oral in 2 strengths, Doxepin Hydrochloride concentrate oral in one strength, Doxycycline Hyclate tablet oral in one strength, Enalapril Maleate Hydrochlorothiazide tablet oral in 2 strengths, Ergocalciferol capsule oral in one strength, Estradiol Norgestimate tablet oral in one strength, Ethambutol Hydrochloride tablet oral in one strength, Ethinyl Estradiol Levonorgestrel tablet oral in 3 strengths, Ethinyl Estradiol Norgestimate tablet oral in 2 strengths, Etodolac capsule oral in one strength, Famotidine tablet oral in 2 strengths, Finasteride tablet oral in one strength, Fluconazole for suspension oral in one strength, Fludrocortisone Acetate tablet oral in one strength, Fluoxetine Hydrochloride capsule oral in 3 strengths, Fluoxetine Hydrochloride solution oral in one strength, Fluvoxamine Maleate tablet oral in 2 strengths, Fosinopril Sodium tablet oral in 3 strengths, Gabapentin tablet oral in 2 strengths, Gemfibrozil tablet oral in one strength, Glipizide tablet oral in 2 strengths, Glipizide Metformin Hydrochloride tablet oral in 3 strengths, Glyburide tablet oral in 6 strengths, Glyburide Metformin Hydrochloride tablet oral in 3 strengths, Haloperidol Lactate concentrate oral in one strength, Hydralazine Hydrochloride tablet oral in 4 strengths, Hydrochlorothiazide tablet oral in 2 strengths, Hydrochlorothiazide Lisinopril tablet oral in 3 strengths, Hydrocodone Bitartrate Ibuprofen tablet oral in one strength, Hydroxyzine Hydrochloride tablet oral in 3 strengths, Indomethacin capsule oral in 2 strengths, Lamotrigine tablet chewable oral in 2 strengths, Leflunomide tablet oral in 2 strengths, Lidocaine Hydrochloride jelly topical in one strength, Lisinopril tablet oral in 6 strengths, Lovastatin tablet oral in 3 strengths, Mefloquine Hydrochloride tablet oral in one strength, Megestrol Acetate tablet oral in 2 strengths, Meloxicam tablet oral in 2 strengths, Metformin Hydrochloride tablet oral in 3 strengths, Metformin Hydrochloride ER tablet oral in 2 strengths, Methyl dopa tablet oral in 2 strengths, Metoclopramide Hydrochloride tablet oral in 2 strengths, Metoprolol Tartrate tablet oral in 2 strengths, Metronidazole capsule oral in one strength, Metronidazole cream topical in one strength, Metronidazole tablet oral in 2 strengths, Mexiletine Hydrochloride capsule oral in 3 strengths, Minocycline Hydrochloride capsule oral in 3 strengths, Mirtazapine tablet oral in 3 strengths, Mirtazapine tablet orally disintegrating oral in 3 strengths, Misoprostol tablet oral in one strength, Mometasone Furoate cream topical in one strength, Mometasone Furoate lotion topical in one strength, Mometasone Furoate ointment topical in one strength, Mupirocin ointment topical in one strength, Naltrexone Hydrochloride tablet oral in one strength, Naproxen tablet oral in 3 strengths, Naproxen DR tablet oral in 2 strengths, Naproxen Sodium tablet oral in one strength, Nefazodone Hydrochloride tablet oral in 5 strengths, Neomycin Sulfate tablet oral in one strength, Nifedipine ER tablet oral in 3 strengths, Oxaprozin tablet oral in one strength, Oxazepam capsule oral in 3 strengths, Oxybutynin Chloride ER tablet oral in 3 strengths, Pantoprazole Sodium DR tablet oral in 2 strengths, Paroxetine Hydrochloride tablet oral in 4 strengths, Penicillin V Potassium for solution oral in 2 strengths, Piroxicam capsule oral in 2 strengths, Prednisolone syrup oral in one strength, Protriptyline Hydrochloride tablet oral in one strength, Ramipril capsule oral in 3 strengths, Simvastatin tablet oral in 5 strengths, Sucralfate tablet oral in one strength, Terazosin Hydrochloride capsule oral in 4 strengths, Terbinafine Hydrochloride tablet oral in one strength, Tetracycline Hydrochloride capsule oral in one strength, Toremide tablet oral in 4 strengths, Tramadol Hydrochloride tablet oral in one strength, Trandolapril tablet oral in 3 strengths, Trazodone Hydrochloride tablet oral in 4 strengths,

Ursodiol capsule oral in one strength, Valproic Acid capsule oral in one strength, Venlafaxine Hydrochloride tablet oral in 5 strengths, Verapamil Hydrochloride ER tablet oral in 2 strengths, and Zolpidem Tartrate tablet oral in 2 strengths .

A.2 Comparing data sources

The dataset we obtained from IQVIA reports the number of dispensed prescriptions nationally at the drug-month-year level but is subject to two minor limitations. First, it covers the first quarter of 2011 through the fourth quarter of 2017, inclusive. In other words, we do not span the period studied in the body of the main text (i.e., 2008-2019 inclusive). We were limited by cost as well as historical availability, so we focused on acquiring data around cartel formation. Second, the data aggregates tablet and capsule purchases. In a small number of cases, substances are delivered in both forms, so we are forced to drop those drugs. The data does not distinguish between immediate and extended-release versions of theophylline and etodolac. Again, we drop those drugs. Observing these distinctions would require more granular data, which was *much* more expensive, and since the omissions reflect random features of the sample, they will not affect the comparisons we derive from them.

In terms of model predictions, the most influential feature of the quantity data is the mean change around cartel formation in cartelized markets relative to uncartelized ones. As a result, we compare the Medicaid and IQVIA datasets along this dimension. Specifically, we denote the log of the number of prescriptions of drug d consumed in quarter t as y_{dt} . We then estimate

$$y_{dt} = \sum_{\tau=-13}^{13} \beta^{\tau} x_{dt}^{\tau} + a_d + b_t + e_{dt}, \quad (25)$$

where a_d and b_t denote drug and quarter fixed effects, respectively, and x_{dt}^{τ} denotes an indicator variable that equals one if and only if d is a cartelized drug and t is τ periods from cartel formation. We $\beta^{-1} = 0$ to facilitate comparisons to the period just prior to cartel formation.

Figure A.2 plots β^{τ} estimated on each dataset. Despite how differently the underlying observations are collected, the sources present very similar graphs. We observe (a) a very slightly positive pre-event trend one to three years prior to cartel formation, (b) no appreciable pre-event trend just prior to cartel formation, and (c) a clear decline in quantity thereafter, (d) culminating in a statistically significant decrease of about 15%. By way of this comparison to the "gold standard" represented by IQVIA, we conclude that Medicaid utilization data accurately measures changes in prescription drug quantities and is well-suited for demand estimation.

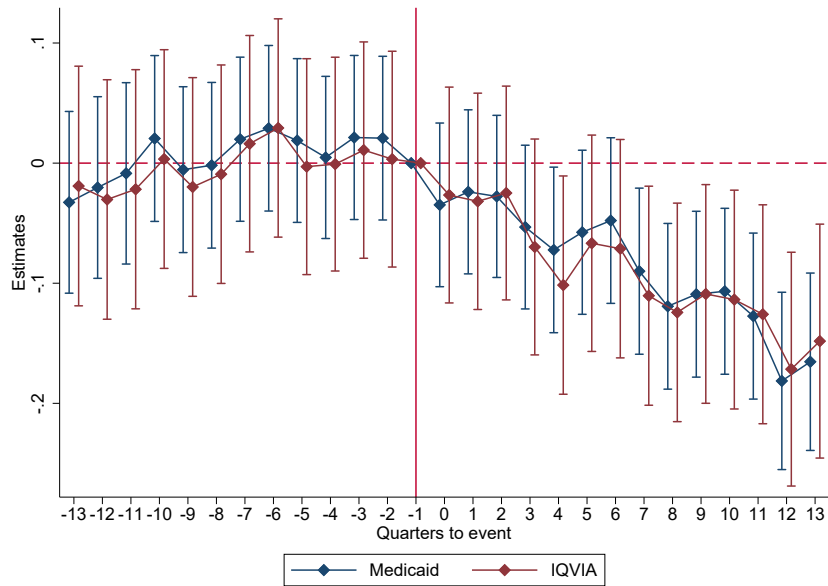


Figure A.2: *The quantity responses reported in IQVIA and Medicaid are very similar.*

This figure plots coefficients obtained by estimating equation 26 on the y-axis against event time on the x-axis. Log prescriptions are the outcome of interest, and the unit of observation is a drug-year. The vertical red line at event time -1 corresponds to the year just prior to cartel formation. Vertical bars around the point estimates show 95 percent confidence intervals, based upon standard errors that are clustered by drug. Notice that while the quantity decline in the Medicaid data following cartel formation lags the one evidenced by IQVIA data, the delayed response does not have a meaningful effect on our results; when we estimate demand, we omit observations from the year in which each cartel is formed.

A.3 Inferring filing dates

Figure A.3 shows that while there is no comprehensive correspondence between ANDA numbers and filing dates, the latter can be inferred without meaningful error.

To obtain filing dates, we downloaded all available approval letters from the FDA website and parsed out filing dates from the PDFs. Since 2000, the agency has issued three "waves" of ANDA numbers. Within each wave, numbers are assigned in chronological order. Specifically, the agency issued numbers in the 70,000s from 2000 to 2008, in the 90,000s from 2008 to 2010, and in the 200,000s thereafter.

In Panel A, we plot filing dates on the x-axis against ANDA numbers on the y-axis. The graph reflects 8,185 ANDAs for which we were able to obtain filing dates from parsed PDFs. The three waves are clearly visible. Putting aside a small number of very early ANDAs, which are filed years before our sample starts, there are only two obvious parsing errors.

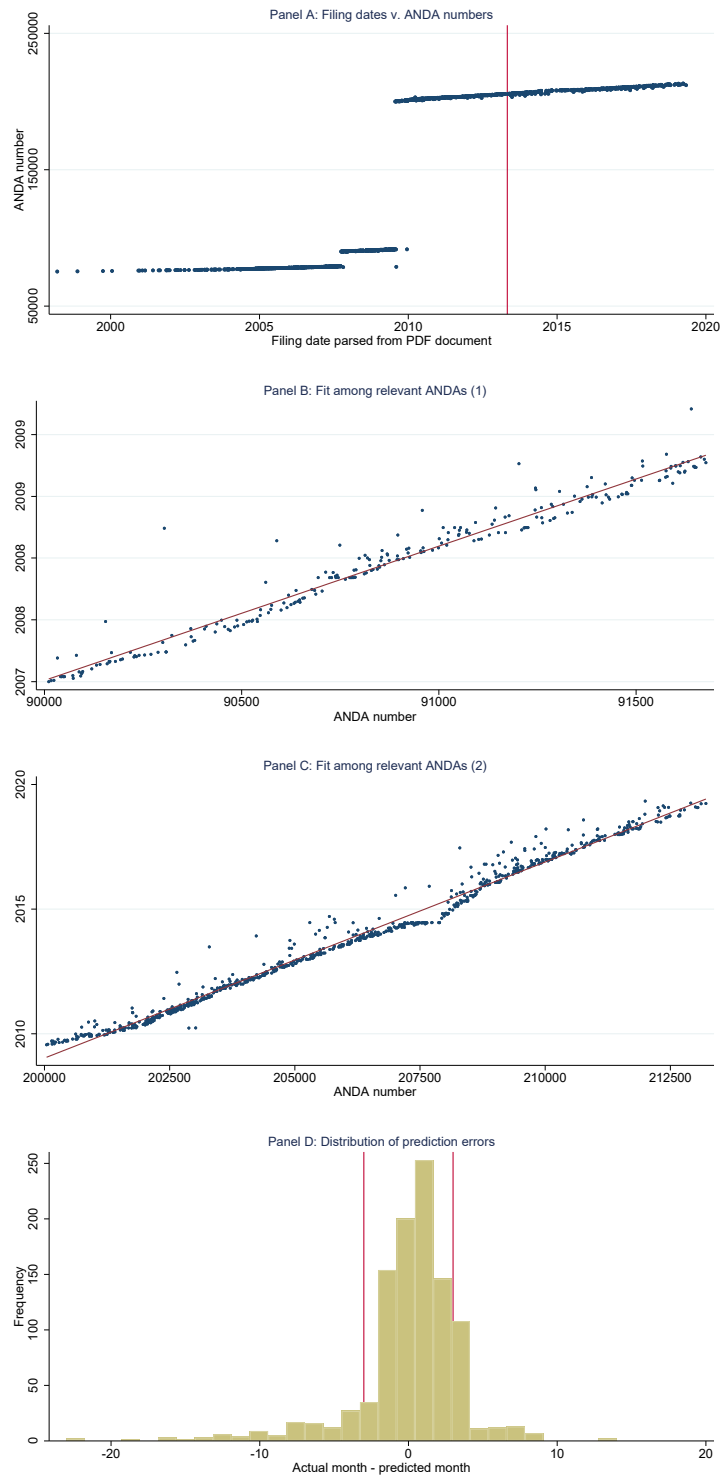


Figure A.3: Filing dates are inferred without meaningful error.

Panel A plots parsed ANDA filing dates on the x-axis and ANDA numbers on the y-axis. Panels B and C plot ANDA numbers on the x-axis and parsed filing dates on the y-axis for the relevant "waves" of ANDAs. Panel D plots the difference between the actual and predicted filing months.

Panel A shows that the sample period corresponds to the second and third wave of ANDA numbers. Thus, in Panels B and C, we isolate the waves separately. Panel B depicts ANDA numbers in the 90,000s, while Panel C depicts them in the 200,000s. In each panel, we plot ANDA numbers on the x-axis and filing dates on the y-axis. Both graphs illustrate the linearity of this relationship.

Given the aforementioned linearity, we regress ANDA numbers on filing dates within each wave and then predict filing dates for the remaining ANDA numbers. To assess overall fit, we compute the difference between actual and predicted dates measured in months and plot the density in Panel D. Substantially all of the parsed dates fall within 3 months of the predicted dates. Especially given that we aggregate ANDAs to the annual level, we conclude that ANDAs filing dates are measured accurately.

B [For Online Publication] Supplement to descriptive analyses

B.1 Tables and figures related to Section 4.1

We compare the price of cartelized drugs before and after cartel formation, using uncartelized drug prices to control for any time-varying factors affecting all generic markets. The unit of analysis is a drug-quarter, and the estimating equation is given by

$$y_{dt} = \sum_{\tau=-21}^{24} \beta^{\tau} x_{dt}^{\tau} + a_d + b_t + e_{dt}. \quad (26)$$

y_{dt} represents the log of average price of drug d in calendar-quarter t . a_d and b_t represent drug- and quarter-specific fixed effects, while x_{dt}^{τ} is an indicator variable that equals one if and only if d is a cartelized drug and t is τ periods from cartel formation. Given the timing of the price increases and the time period covered by our data, τ can take a value between -21 and 24. We set $\beta^{-1} = 0$, so the coefficients on x^{τ} terms represent differences in prices relative to the period immediately preceding cartel formation. Figure B.1 reports estimates of β^{τ} .

Figure B.1 replicates Figure I in the body of the main text with an exception: instead of employing two-way fixed effects, we follow Sun and Abraham (2020), whose approach accounts for potential contamination of leading and lagging coefficients. (See their paper for details.) We obtain similar coefficients.

Figure B.3 replicates Figure I in the body of the main text with an exception: rather than distinguish between cartelized and uncartelized markets, we distinguish between *markets we predict are cartelized and uncartelized*. Our predictions are based on scores assigned by NP, which reflect the strength of her personal relationships.

Figure B.4 replicates Figure I in the body of the main text with an exception. We add a third price series, which tracks the average log price of markets in which Teva is a monopolist (as of the first quarter of 2013). Figure B.4 reports this result.

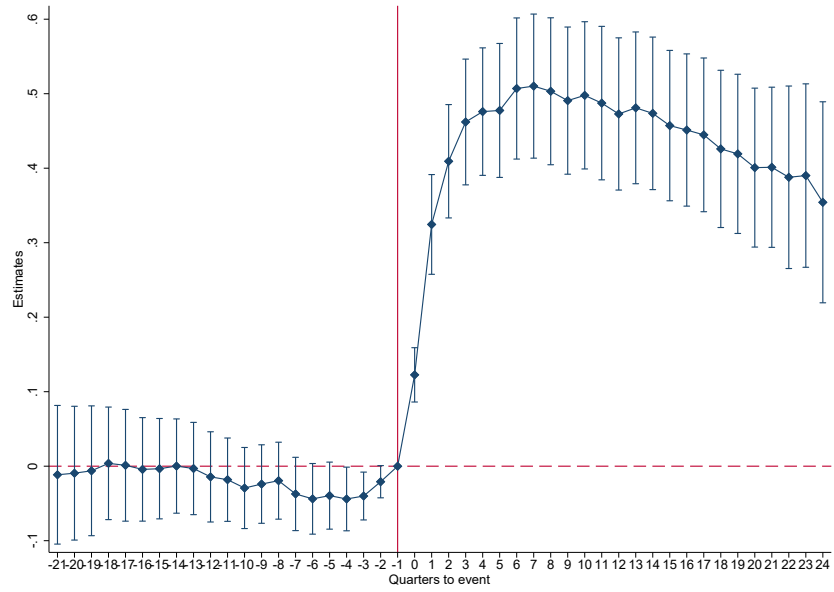


Figure B.1: Prices in event time (two-way fixed effects)

This figure plots β^t , which is obtained by estimating equation 26, on the y-axis against event time on the x-axis. The unit of observation is a drug-year. The outcome variable is low average price. The vertical red line at event time -1 corresponds to the year the cartel is formed. Vertical bars around the point estimates show 95 percent confidence intervals for those coefficients. Standard errors are clustered by substance-delivery-release.

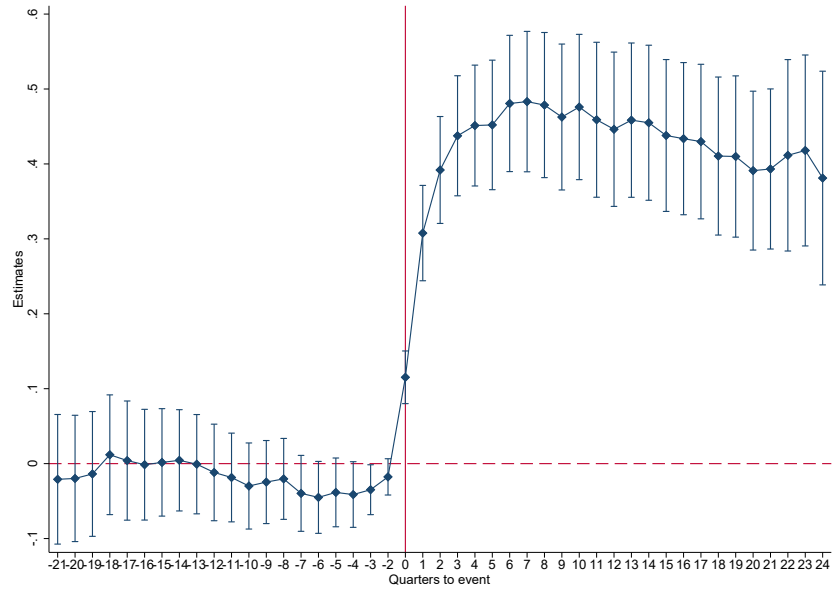


Figure B.2: Prices in event time (Sun and Abraham (2020) approach)

This figure plots β^t , which is obtained by estimating equation 26, on the y-axis against event time on the x-axis. The unit of observation is a drug-quarter. Vertical bars around the point estimates show 95 percent confidence intervals for those coefficients. Standard errors are clustered by drug.

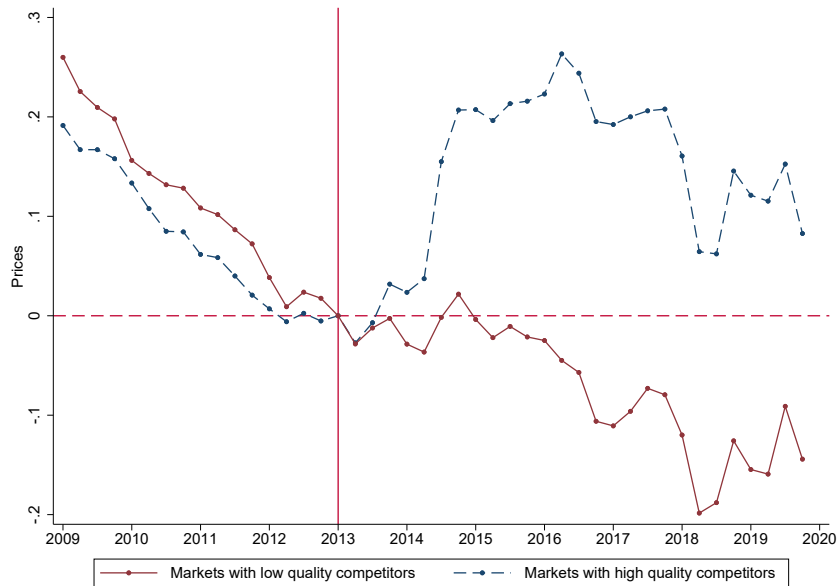


Figure B.3: Prices in markets where NP does/does not have strong personal relationships

This figure replicates Figure I in the body of the main text with one exception: rather than compare cartelized and uncartelized markets, we compare markets where NP does and does not have especially strong relationships. To do so, we obtain the "quality" of each Teva competitor from the Complaint (i.e., a score assigned by NP to reflect the strength of her personal relationships with the key sales and marketing persons at each firm). To approximate the process that NP actually used to determine what markets to cartelize, we compute each drug's weighted average competitor quality. We define "high quality" markets as ones in which the average competitor quality is greater than or equal to two, and we call the remaining markets "low quality."

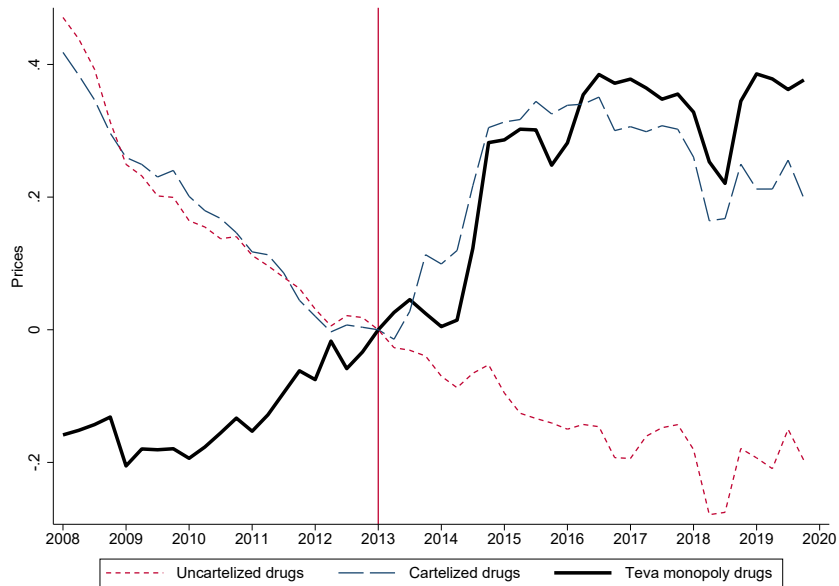


Figure B.4: Prices in markets where Teva has a monopoly

Unlike cartelized drugs, prices in markets where Teva was a monopolist did not rise discontinuously when NP joined; although there is a slight jump late in 2014, the fact is that they were increasing and continued increasing at roughly the same pace.

B.2 Tables and figures related to Section 4.2

Firms with dormant ANDAs—ones that were once associated with positive production but no longer are—might re-enter the market when cartels form. To study this possibility, we plot re-entry in calendar time for cartelized and uncartelized drugs. Figure B.5 reports the result. (The format of the graph is identical to the format of Figure II in the body of the main text except that the y-axis measures the average number of (re-)entrants per drug and year rather than per substance-delivery-release combination and year. Cartel formation clearly induces re-entry.

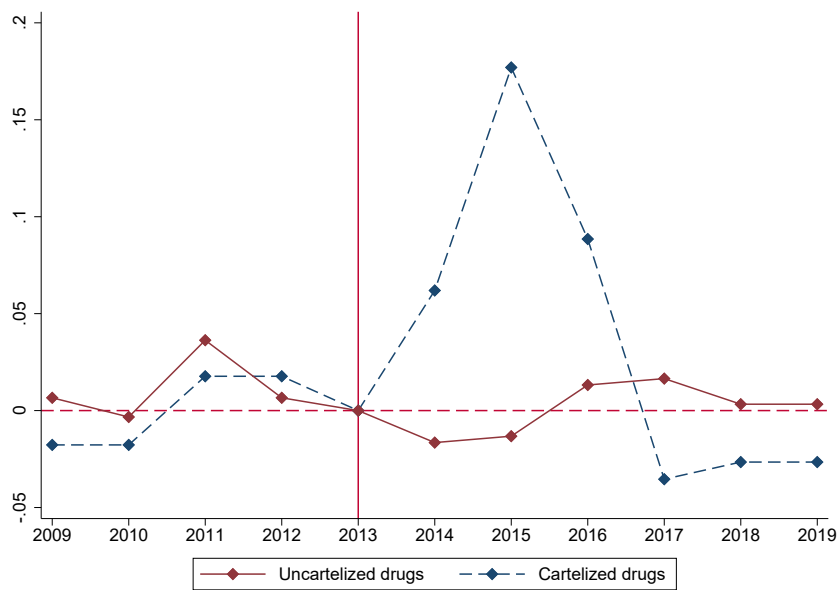


Figure B.5: *Reentry in event time*

This figure plots average number of re-entrants per drug on the y-axis against calendar year on the x-axis. The vertical red line corresponds to the first quarter of 2013—the period in which NP joined Teva. The number of entrants are normalized to zero in that quarter.

We also compare entry into cartelized markets before and after cartel formation, using uncartelized drug prices to control for any time-varying factors affecting all generic markets. We denote the number of ANDA filings for a substance-delivery-release combination j in year t by y_{jt} , and we estimate

$$y_{jt} = \sum_{\tau=-5}^5 \beta^{\tau} x_{jt}^{\tau} + a_j + b_t + e_{jt}. \quad (27)$$

a_j and b_t represent drug- and year-specific fixed effects, respectively, while x_{jt}^{τ} is an indicator variable that equals one if and only if j is a cartelized substance-delivery-release combination and t is τ periods from cartel formation. We set $\beta^{-1} = 0$, so the coefficients on x^{τ} terms represent differences in prices relative to the period immediately preceding cartel formation. We plot estimates of β^{τ} in event-time. Panel A of Figure B.6 reports this result. We then replicate this procedure but replace ANDA filings with ANDA approvals, again, plot estimates of β^{τ} in event time. Panel B of Figure B.6 reports this result. We also replicate this procedure but replace ANDA filings or approvals with re-entry, and we plot estimates of β^{τ} in event time. Figure B.6 reports this result.

Figure B.8 plots the distribution of total and regulatory-specific delays.

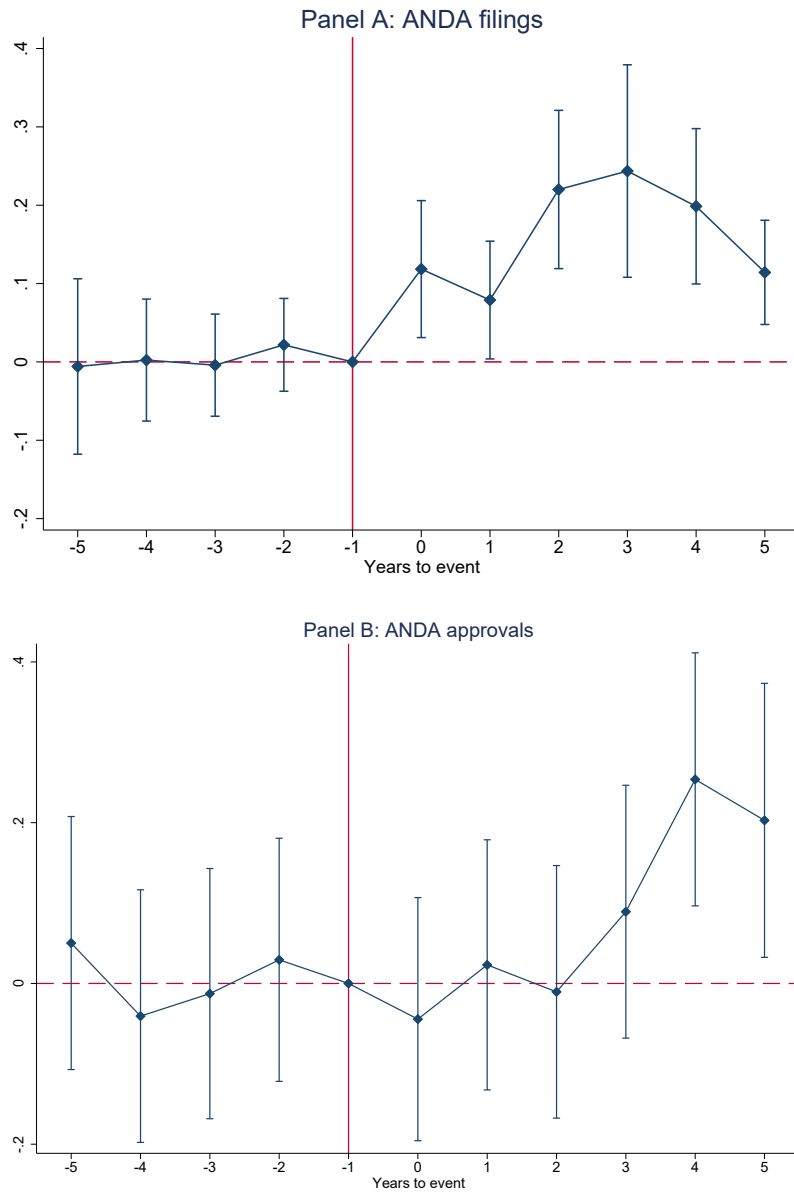


Figure B.6: Entry in event time

This figure plots β^T , which is obtained by estimating equation 26, on the y-axis against event time on the x-axis. The unit of observation is a substance-delivery-release-year. In Panel A, ANDA filings are the outcome of interest. In Panel B, ANDA launches are the outcome of interest. The vertical red line at event time -1 corresponds to the year immediately prior to cartel formation, as described in the complaint. Vertical bars around the point estimates show 95 percent confidence intervals for those coefficients. Standard errors are clustered by drug.

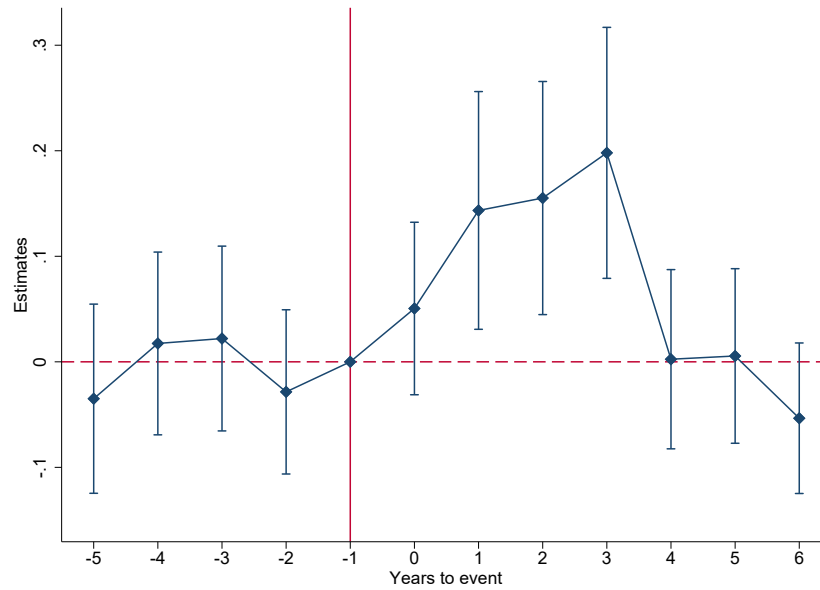


Figure B.7: Reentry in event time

In the body of the main text, we find that cartel formation attracts new entrants (i.e., ANDA filings). However, re-entry is also possible. Firms with dormant ANDAs—ones that were once associated with positive production but no longer are—might re-enter the market when cartels form. To study this possibility, we plot re-entry in event time for cartelized and uncartelized drugs. There is an economically and statistically increase in entry, and that re-entry occurs very soon after collusion begins. The stark contrast of this figure and the one that reports ANDA launches highlights the effect of approval delays. This figure plots coefficients obtained by estimating equation 26 on the y-axis against event time on the x-axis. The unit of observation is a drug-year, and the outcome of interest is a ANDA re-entry. Re-entering ANDAs are those where the ANDA was associated with some output, then went at least one year without being associated with output, and then re-entered (i.e., was once again associated with output). The vertical red line at event time zero corresponds to the year in which Teva hired NP. Vertical bars around the point estimates show 95 percent confidence intervals for those coefficients, based upon standard errors that are clustered by drug.

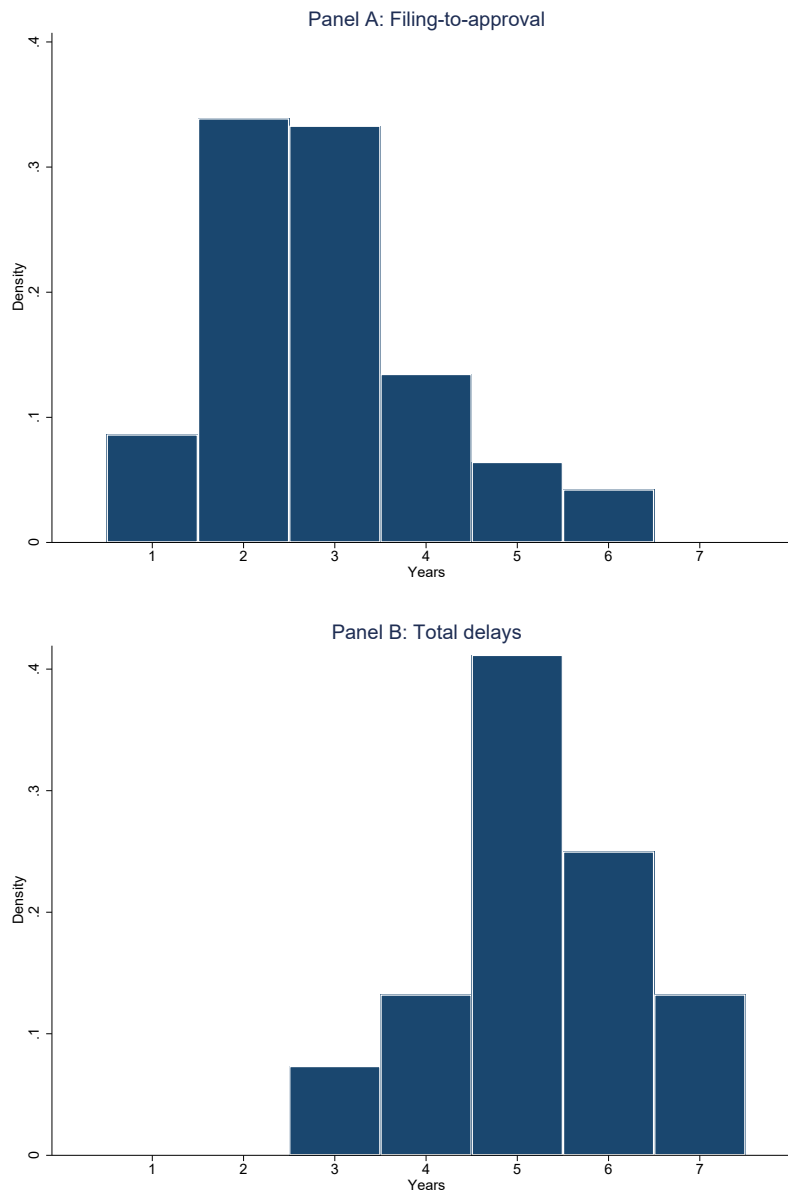


Figure B.8: *Distributions of delays*

This figure plots the distribution of delays. The unit of observation is an ANDA. Panel A reports the distribution of regulatory delays (i.e., the time between ANDA filing and approval), whereas Panel A reports the distribution of regulatory delays (i.e., the time between cartel formation and sales).

B.3 Tables and figures related to Section 4.3

Figure B.9 examines the relationship between market size and entry in cartelized markets. For the purposes of this figure, we proxy for market size using each drug's total revenue in the quarter just prior to NP joining Teva. In Panel A, we measure entry as the probability that each cartelized market experiences entry following cartel formation. In Panel B, we measure entry as the average number of firms entering each

cartelized market following cartel formation. The graph shows that regardless of how we measure it, entry is closely related to market size. Only about 8% of drugs with around \$1 million in revenue attract any entry at all. However, drugs with over \$1 billion in revenue almost always attract entry, with an average of three firms filing ANDAs following cartel formation.

As we state in the body of the main text, comparing the paths of prices in large and small markets is complicated in this setting by antitrust risk. Large markets may be more "visible" than small ones, and the earliest entry events roughly coincide with the government's investigation, so cartel members may have reduced prices in large markets in an effort to reduce scrutiny of their behavior. In other words, prices in large cartelized markets might have fallen regardless of entry. To investigate this issue, we trace the investigation back to its origin, obtain the list of drugs whose price changes triggered the inquiry, and plot the size distribution of these drugs to the size distribution of the full sample in Figure B.10. The array of lawsuits faced by the firms all trace back to the Connecticut AG's office, which launched its initial inquiry based on July 8, 2014, *New York Times* article, which in turn described price changes first reported by Adam Fein of Pembroke Consulting and the Drug Channels Institute. Fein's report, which appears as a November 19, 2013 article titled "Retail Generic Drug Costs Go Up, Up, and Away," simply orders the drugs in terms of year-over-year percentage increases in NADAC prices. The report cites increases in the price of doxycycline, clomipramine, albuterol sulfate, captopril, tetracycline, digoxin, and benazepril. Figure B.10 shows that the distribution of market size for the drugs that prompted the government's investigation is very similar to the distribution of market size for the full sample.

One-third of entrants into cartelized markets are cartel members. Of the remaining entrants, which are nonmembers, the vast majority (68%) were in existence prior to cartel formation. Figure B.11 plots entry into cartelized and uncartelized drug markets over time separately for members and nonmembers, and we observe similar patterns across the two groups. The distribution of member and nonmember entrants into cartelized markets does not change around the time NP is hired by Teva.

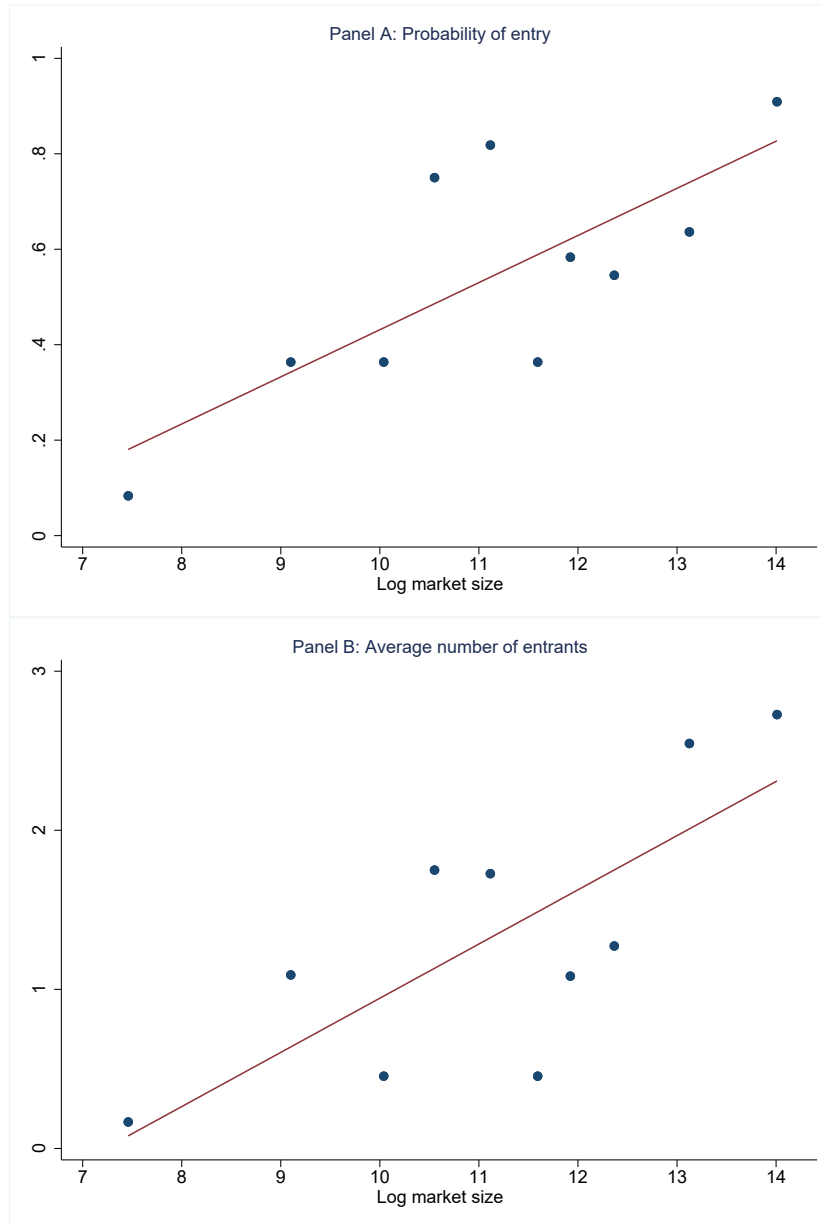


Figure B.9: Larger markets attract more entry

This figure plots log market size on the x-axis against the number of entrants in the post-collusion period. The unit of observation is a drug. Market size is measured in total revenue in 2012, the year immediately prior to NP joining Teva. Data are binned according to x-axis values, so averages within the bin are plotted (i.e., the graph represents a "binscatter").

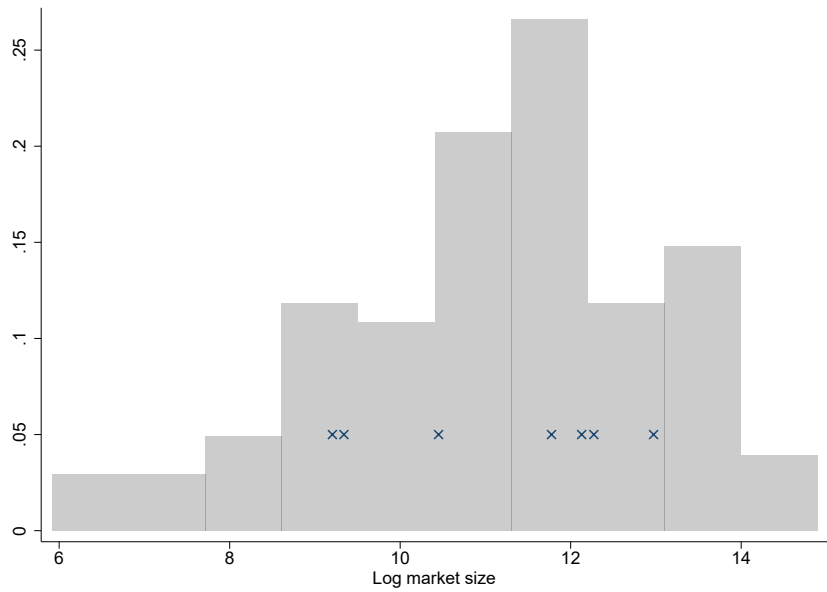


Figure B.10: Size distribution of drugs that triggered the government's investigation

This figure reports the distribution of (logged) size, measured by 2012 revenue. The drugs identified by Fein are marked with a blue "x."

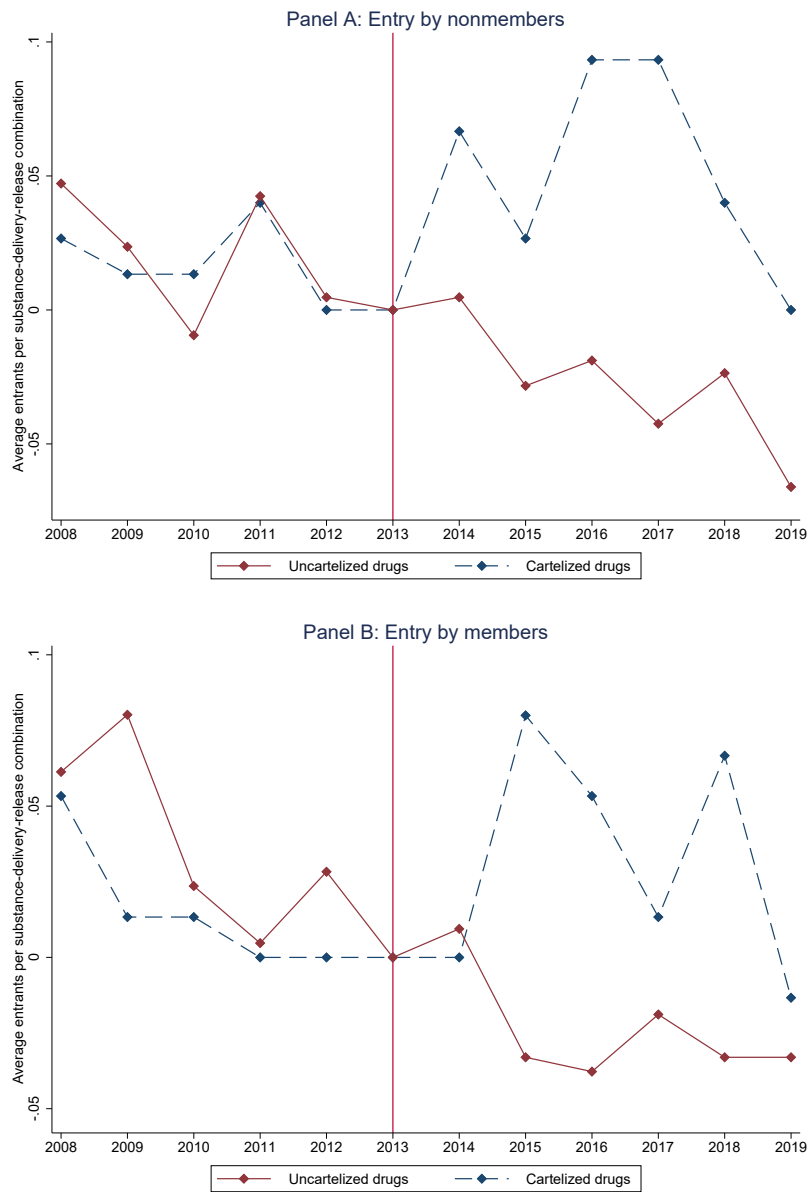


Figure B.11: Entry in calendar time by cartel membership

In the body of the main text, we find that cartel formation attracts new entrants (i.e., ANDA filings). Here, we plot the relationship separately for cartel nonmembers (left panel) and cartel members (right panel).

C [For Online Publication] Supplement to structural analyses

C.1 Price elasticity of demand

We better understand heterogeneity in the price elasticity of demand, we plot inverse price coefficients (i.e., $1/\alpha$) against corresponding drug classes. Figure C.1 reports the result. Buyers of two classes of

drugs—other antiepileptics and β -blockers—are especially inelastic.

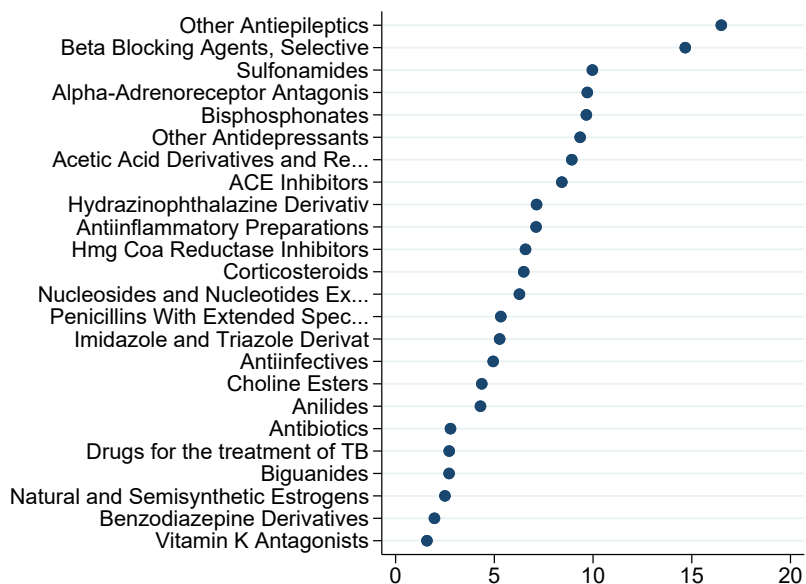


Figure C.1: Patients consuming antiepileptics and β -blockers are inelastic.

This figure plots inverse price coefficients (i.e., $1/\alpha$) on the x-axis against corresponding drug classes on the y-axis. Buyers of two classes of drugs—other antiepileptics and β -blockers—are especially inelastic, so we incorporate this heterogeneity in the demand system. See Section 6 for more details.

C.2 Marginal cost

In Appendix Figure C.2, we report the distribution of marginal costs. Appendix Figure C.3 shows that the model fits well out of sample. The estimates imply markups that very closely align with figures reported by Teva in their financial statements. To obtain values implied by our model, we set the first order condition of the profit function with respect to price equal to zero, solve for $p_{dft} - mc_{dft}$, divide the resulting markups by prices, and average over drugs manufactured by Teva, weighting by revenue. To obtain analogous figures from Teva's annual reports, we extract segment-specific income statements and compute the ratio of operating profits to total revenue for their generic division.⁶² Our model assumes competitive pricing and implies that profit margins average 19.7%, while Teva's financial statements imply 20.0% in the two years prior to NP joining Teva. In the two years after NP joins Teva, our model assumes NP has cartelized many drug markets and implies that profit margins average 39.6%, while Teva's financial statements imply 39.9%. In other words, forecasts from the model not only match profit rates in levels but also changes around cartel formation.⁶³

⁶²Operating margin is the right choice, given how Teva reports its income. Operating profit reduces total revenue by cost of goods sold and selling/marketing expenses, which are mostly variable, but not general/administrative expenses, (e.g., executive compensation, headquarters operations, etc.), which are mostly fixed/sunk.

⁶³Although careful demand estimation contributed to this result, we believe that *such a close correspondence* between the model's predictions and the financial statement analysis is, at least in part, coincidental. The goal of this exercise was to see if the model was in the neighborhood of the annual reports—not whether it was a close match.

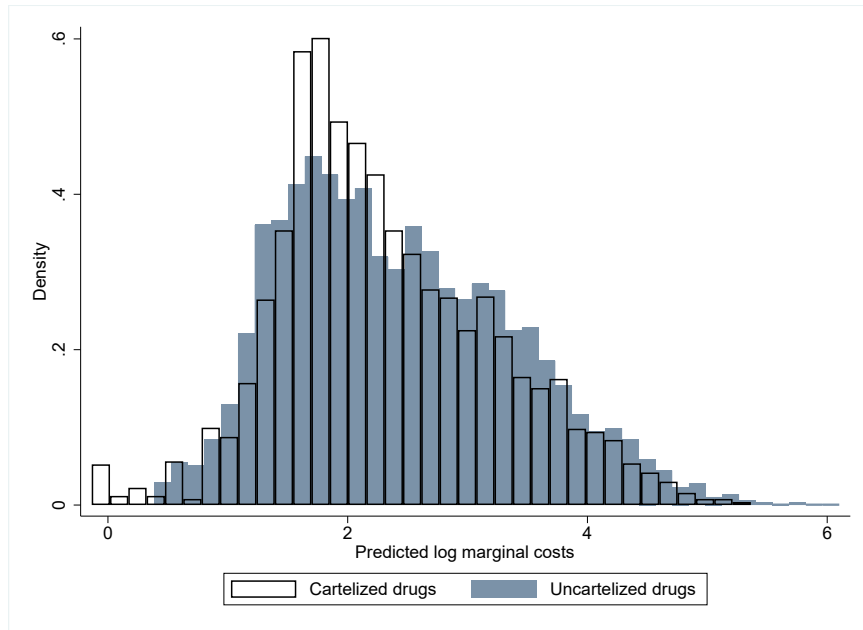


Figure C.2: *Distribution of average predicted log marginal cost*

This figure plots the density of predicted log marginal costs, $\hat{\gamma}_d + \hat{\gamma}_t$, separately for cartelized and uncartelized drugs. It comprises all drug-year observations. However, to avoid taking a stand on conduct prior to formalizing testing it, we estimate the parameters using unambiguously competitive drug-year observations (i.e., using (a) drugs whose prices were never fixed and (b) periods prior to cartelization for drugs whose prices were fixed).

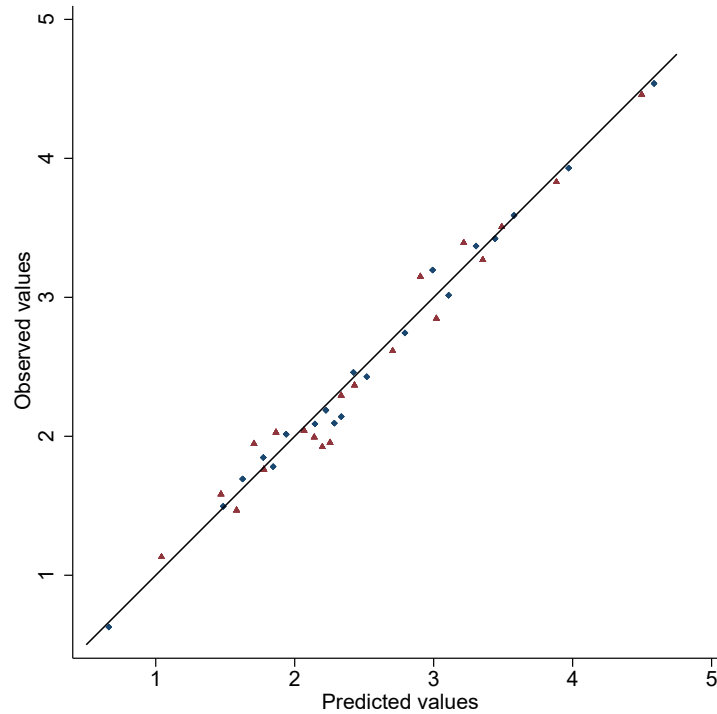


Figure C.3: *Out-of-sample evaluation of log marginal cost predictions*

This figure is constructed as follows: we restrict attention to 2008-2012 and “back out” marginal cost estimates assuming firms set Bertrand-Nash prices. Then, we estimate a regression using uncartelized drugs from 2008-2012 and cartelized drugs from 2008-2010. Next, we predict marginal costs for our “leave-out” sample, which comprises cartelized drugs from 2011-2012. Last, we plot (“binscatter”) predicted marginal costs for the leave-out sample against our marginal cost estimates from those drugs and periods.

C.3 Damage assessment

Using our model and demand estimates, we can compute damages to consumers. For each product in a cartelized drug market, we compute equilibrium prices under competition and collusion, and we multiply the difference by the number of observed prescriptions. The median price differences are \$6.1, \$6.4, and \$5.5 per prescription for 2013, 2014, and 2015, respectively. Mean price differences are slightly higher at \$10.0, \$7.3, and \$7.7, respectively. Damages total \$787 million, \$1.4 billion, and \$1.5 billion, respectively. That is, damages total \$3.8 billion over the three-year period, which averages out to \$18.2 million per drug per year.

Our figures are very similar to those reported in two other recent studies. Cuddy (2020) finds that collusion induced nationwide annual damages of \$49.5 million per substance-delivery-release combination.⁶⁴ Even though our data and models differ, we arrive at a similar figure, \$40.5 million. Clark et al. (2021) study six substance-delivery-release combinations that were affected by price fixing, estimating damages

⁶⁴To arrive at \$49.5 million, we start with average annual damages for the insurer she studies (\$1.3755 billion, per her Table 8), scale up to nationwide damages (by a factor of 5.8, per her Section 6.4), and divide by the number of substance-delivery-release combinations in her sample (161, per her Appendix Table B.1).

using a carefully constructed difference-in-difference research design. Again, we reach similar estimates.⁶⁵ Since the source of our quantity data is the same as theirs, we predict nearly identical damages for the substance-delivery-release combinations for which we overlap.

D [For Online Publication] Technical appendix

D.1 Moment Inequalities

THEOREM I. Moments indexed by i and given by

$$\frac{1}{J} \sum_j \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \frac{1}{2} \sum_{k \in \{M, N\}} h_j^i \left[\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] < 0 \quad (28)$$

produce consistent upper bounds.

PROOF.

For each moment indexed by i , we have

$$\begin{aligned} & \frac{1}{J} \sum_j \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \frac{1}{2} \sum_{k \in \{M, N\}} h_j^i \left[\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ &= \frac{1}{J} \sum_j \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \frac{1}{2} \sum_{k \in \{M, N\}} h_j^i \left[VE_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ &= \frac{1}{J} \sum_j \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \frac{1}{2} \sum_{k \in \{M, N\}} h_j^i \left[VE_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \theta_{jk} + \eta_j \right] \\ &= \frac{1}{J} \sum_j \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \frac{1}{2} \sum_{k \in \{M, N\}} h_j^i \left[VE_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \mathcal{E}(\theta_{jk} | \mathcal{I}_{jk}) - v_{jN} + \eta_j \right] \\ &< \frac{1}{J} \sum_j \left[h_j^i \eta_j \right] - \frac{1}{J} \sum_j \left[h_j^i v_{jN} \right] \xrightarrow{p} \mathbb{E}[h^i \eta] - \mathbb{E}[h^i v_{jN}] = \mathbb{E}[h^i] \mathbb{E}[\eta] - \mathbb{E}[h^i] \mathbb{E}[v_{jN}] = 0. \quad (29) \end{aligned}$$

The first equality results from replacing $\widehat{VE}_d^k(\cdot)$ with $VE_d^k(\cdot)$. $\widehat{VE}_d^k(\cdot)$ is a function of $\widehat{\pi}_{d,t}^k(\cdot)$, \widehat{F}_{ξ} , \widehat{F}_{ω} , and \widehat{F}_D , and $VE_d^k(\cdot)$ is a function of $\pi_{d,t}^k(\cdot)$, F_{ξ} , F_{ω} , and F_D . Since $\pi_{d,t}^k(M, N, \xi_{dt}, \omega_{dt})$, F_{ξ} , F_{ω} , and F_D are measured without error, and since the value entry depends on only those objects, entry values are measured without error. The second equality results from replacing $\theta_0 + \theta_1 r_j + \theta_2 \ell_j$ with $\theta_{jk} - \eta_j$, which follows directly from equation 16. The third equality results from replacing θ_{jk} with $\mathcal{E}[\theta_{jk} | \mathcal{I}_{jk}] + v_{jN}$, which follows directly from the definition of an expectational error. The inequality follows from the necessary conditions of a simultaneous move Nash equilibrium, which require $VE_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \mathcal{E}(\theta_{jk} | \mathcal{I}_{jk}) < 0$. (If this condition were false, then another firm would have expected to profitably enter.) The next step follows from the law of large numbers. Since h^i depends on r and ℓ , η is independent of r and ℓ , v is independent of r and ℓ , the fourth inequality holds. To arrive at the final inequality, notice that η and v are both unconditionally mean zero (i.e., $\mathbb{E}[\eta] = 0$ and $\mathbb{E}[v] = 0$). ■

⁶⁵Whereas they estimate 44.2% and 13.5% price increases for nystatin and theophylline, respectively, our structural model predicts 41.4% and 20.8% changes. To arrive at these figures, we divide the estimated damages per defined daily dose by pre-collision prices, both of which are reported by the authors in their Table 7. Specifically, we define \$0.21 by \$1.561 and \$0.155 by \$0.350.

THEOREM II. Moments indexed by i and given by

$$\begin{aligned} & \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} \widehat{VE}_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M=0, \chi^N=0\}} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{wVE}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \geq 0 \quad (30) \end{aligned}$$

produce consistent lower bounds.

PROOF.

For each moment i , order j by their value of $w_j^i \eta_j$, let $L_{w\eta}$ denote the set of j that correspond to the smallest J values, and let $\Psi_{w\eta}$ correspond to the smallest $J - J_L$ values. For each moment indexed by i , we have

$$\begin{aligned} & \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} \widehat{VE}_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M=0, \chi^N=0\}} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & \geq \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} \widehat{VE}_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M=0, \chi^N=0\}} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\widehat{VE}_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & = \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} VE_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M=0, \chi^N=0\}} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{VE_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \theta_0 - \theta_1 r_j - \theta_2 \ell_j \right] \\ & = \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} VE_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M=0, \chi^N=0\}} - \theta_{jk} + \eta_j \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{VE_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \mathbb{E}(\theta_{jk} | \mathcal{J}_{jk}) + \eta_j \right] \\ & = \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{\mathbb{1}\{\chi^k \geq 1\} VE_d^k(\chi^M + \mathbb{1}, \chi^N)}{\mathbb{1}\{\chi^M \geq 1\} + \mathbb{1}\{\chi^N \geq 1\} + \mathbb{1}\{\chi^M=0, \chi^N=0\}} - \theta_{jk} + \eta_j - \nu_{jN} \right] \\ & - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i \left[\frac{VE_d^k(\chi^M + \mathbb{1}\{k=M\}, \chi^N + \mathbb{1}\{k=N\})}{2} - \mathbb{E}(\theta_{jk} | \mathcal{J}_{jk}) + \eta_j - \nu_{jN} \right] \\ & \geq \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i [\eta_j - \nu_{jN}] - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i [\eta_j - \nu_{jN}]. \quad (31) \end{aligned}$$

The first inequality follows from the construction of set Ψ_{wVE} . The first equality results from replacing $\widehat{VE}_d^k(\cdot)$ with $VE_d^k(\cdot)$. $\widehat{VE}_d^k(\cdot)$ is a function of $\hat{\pi}_{d,t}^k(\cdot)$, \hat{F}_ξ , \hat{F}_ω , and \hat{F}_D , and $VE_d^k(\cdot)$ is a function of $\pi_{d,t}^k(\cdot)$, F_ξ , F_ω , and F_D . Since $\pi_{d,t}^k(M, N, \xi_{dt}, \omega_{dt})$, F_ξ , F_ω , and F_D are measured without error, and since the value

entry depends on only those objects, entry values are measured without error. The second equality results from replacing $\theta_0 + \theta_1 r_j + \theta_2 \ell_j$ with $\theta_{jk} - \eta_j$, which follows directly from equation 16. The third equality results from replacing θ_{jk} with $\mathcal{E}[\theta_{jk} | \mathcal{S}_{jk}] + \nu_{jN}$, which follows directly from the definition of an expectational error. The second inequality follows from the necessary conditions of a simultaneous move Nash equilibrium. That is, for $k \in \{M, N\}$, these conditions require $VE_d^k(\chi^M, \chi^N) - \mathcal{E}(\theta_{jk} | \mathcal{S}_{jk}) \geq 0$ as well as $VE_d^k(\chi^M + \mathbb{1}\{k = M\}, \chi^N + \mathbb{1}\{k = N\}) - \mathcal{E}(\theta_{jk} | \mathcal{S}_{jk}) < 0$.

Assume that $J_L/J \xrightarrow{p} q$ by the law of large numbers. We then have

$$\begin{aligned}
& \frac{1}{J} \sum_{j \in L} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i [\eta_j - \nu_{jN}] - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} \frac{1}{\mu_j} \sum_{d \in \mathcal{J}} \sum_{k \in \{M, N\}} w_j^i [\eta_j - \nu_{jN}] \\
&= \frac{1}{J} \sum_{j \in L} [w_j^i \eta_j] - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} [w_j^i \eta_j] - \left(\frac{1}{J} \sum_{j \in L} [w_j^i \nu_{jN}] - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} [w_j^i \nu_{jN}] \right) \\
&\geq \frac{1}{J} \sum_{j \in L_{w\eta}} [w_j^i \eta_j] - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} [w_j^i \eta_j] - \left(\frac{1}{J} \sum_{j \in L} [w_j^i \nu_{jN}] - \frac{1}{J} \sum_{j \in \Psi_{w\eta}} [w_j^i \nu_{jN}] \right) \\
&\xrightarrow{p} \mathbb{E}[\eta | \eta < F^{-1}(q), w^i = 1] - \mathbb{E}[\eta | \eta < F^{-1}(1 - q), w^i = 1] - (\mathbb{E}[\nu | \eta < F^{-1}(q), w^i = 1] - \mathbb{E}[\nu | \eta < F^{-1}(1 - q), w^i = 1]) \\
&= \mathbb{E}[\eta | \eta < F^{-1}(q)] - \mathbb{E}[\eta | \eta < F^{-1}(1 - q)] - (\mathbb{E}[\nu] - \mathbb{E}[\nu]) = 0. \quad (32)
\end{aligned}$$

The first equality follows from the fact that w^i and η_j , and ν_{jN} do not depend on d or k . The first inequality follows from the construction of $L_{w\eta}$. The third step follows from the law of large numbers. The second equality follows from the fact that η and ν are independent of r and ℓ , on which w^i depends. Thus, for example, $\mathbb{E}[\eta | \eta < F^{-1}(q), w^i = 1] = \mathbb{E}[\eta | \eta < F^{-1}(q)]$. To arrive at the final step, notice that $\mathbb{E}[\eta | \eta < F^{-1}(q)]$ and $\mathbb{E}[\eta | \eta < F^{-1}(1 - q)]$ are values that are equidistant from zero, so their difference is zero. Also, notice that ν is independent of η and is unconditionally mean zero. ■

D.2 Procurement scoring auctions vs. nested logit demand

D.2.1 Purchase probabilities and market shares

Consider the procurement scoring auction described in the body of the main text, and omit drug and time subscripts to simplify notation. Let \mathcal{F} denote the set of firms with products in this market, $\mathcal{P}(\cdot)$ denote the probability that the event in the parentheses occurs, and $\mathcal{P}_i(\cdot)$ denote the probability that i buys the drug from f . \mathcal{P}_{if} can be written as

$$\begin{aligned}
& \mathcal{P}(\lambda + \xi_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p_{if} \geq \lambda + \xi_{f'} + \zeta_i + (1 - \sigma)\epsilon_{if'} - \alpha p_{if'} \quad \forall f' \in \mathcal{F}) \\
& \quad \times \mathcal{P}(\lambda + \xi_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p_{if} \geq \epsilon_{i0}). \quad (33)
\end{aligned}$$

Since ϵ and ζ are not known by the firms when they submit proposals and are i.i.d. with respect to i (and other indices), f sets the same price to all i . Thus, \mathcal{P}_{if} can be rewritten as

$$\begin{aligned}
& \mathcal{P}(\lambda + \xi_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p_f \geq \lambda + \xi_{f'} + \zeta_i + (1 - \sigma)\epsilon_{if'} - \alpha p_{f'} \quad \forall f' \in \mathcal{F}) \\
& \quad \times \mathcal{P}(\lambda + \xi_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p_f \geq \epsilon_{i0}). \quad (34)
\end{aligned}$$

Since ϵ and $[\zeta + (1 - \sigma)\epsilon]$ are distributed Type 1 extreme value, \mathcal{P}_{if} can further be rewritten as

$$\frac{e^{\lambda + \alpha p_f + \zeta_f}}{\Lambda^\sigma (1 + \Lambda)^{1 - \sigma}}, \quad (35)$$

where $\Lambda = \sum_{f' \in \mathcal{F}} e^{\lambda + \alpha p_{f'} + \zeta_{f'}}$.

Let w_i denote the size of buyer i and s_f denote the market share of f . s_f equals the weighted average probability of winning the procurement scoring auctions, which is given by $\frac{\sum_i w_i \mathcal{P}_{if}}{\sum_i w_i}$. If the buyers are approximately symmetric, then $\frac{\sum_i w_i \mathcal{P}_{if}}{\sum_i w_i} \approx \frac{\sum_i w \mathcal{P}_{if}}{\sum_i w} = \frac{1}{n_i} \sum_i \mathcal{P}_{if}$, where n_i denotes the number of buyers. If the number of buyers is very large, then $\frac{1}{n_i} \sum_i \mathcal{P}_{if} \approx \mathbb{E}[\mathcal{P}_{if}]$, which is the same as the individual probability that i purchases the drug from f , since \mathcal{P}_{if} does not depend on i . Hence, $s_f = \frac{e^{\lambda + \alpha p_f + \zeta_f}}{\Lambda^\sigma (1 + \Lambda)^{1 - \sigma}}$.

Thus, under the collection of procurement scoring auctions described in the body of the main text, each firm's market share is approximately equal to the market share obtained under a nested logit demand system.

D.2.2 Bids and prices

Consider the procurement scoring auction described in the body of the main text, and omit drug and time subscripts to simplify notation. Let $\mathcal{P}(\cdot)$ denote the probability that the event in the parentheses occurs. Each firm f wishes to maximize its profit, so it solves

$$\max_p \left\{ \mathcal{P}(\lambda + \zeta_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p \geq \lambda + \zeta_{f'} + \zeta_i + (1 - \sigma)\epsilon_{if'} - \alpha p_{f'}) \forall f' \in \mathcal{F} \right. \\ \left. \times \mathcal{P}(\lambda + \zeta_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p \geq \epsilon_{i0})(p - mc_f) \right\}. \quad (36)$$

As in the explanation above, since ϵ_{if} and ζ_i are not known by the firms when they submit proposals and are i.i.d. with respect to i (and other indices), each f' sets the same price to all buyer (i.e., $p_{if'} \equiv p_{f'}$). Thus, the preceding optimization problem can be rewritten as

$$\max_p \left\{ \mathcal{P}(\lambda + \zeta_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p \geq \lambda + \zeta_{f'} + \zeta_i + (1 - \sigma)\epsilon_{if'} - \alpha p_{f'}) \forall f' \in \mathcal{F} \right. \\ \left. \times \mathcal{P}(\lambda + \zeta_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p \geq \epsilon_{i0})(p - mc_f) \right\}. \quad (37)$$

Since ϵ_{if} and $[\zeta_i + (1 - \sigma)\epsilon]_{if}$ are distributed Type 1 extreme value, the preceding optimization problem can further be rewritten as

$$\max_p \left\{ \frac{e^{\lambda + \alpha p + \zeta_f}}{\Lambda^\sigma (1 + \Lambda)^{1 - \sigma}} (p - mc_f) \right\}, \quad (38)$$

where $\Lambda = \sum_{f' \in \mathcal{F}} e^{\lambda + \alpha p_{f'} + \zeta_{f'}}$.

Separately, consider a nested logit demand model. Let the buyer's "utility" takes the same form as its "payoff" in the procurement auction, so that $u_{if} = \lambda + \zeta_f + \zeta_i + (1 - \sigma)\epsilon_{if} - \alpha p$, and let ζ_f , ζ_i , and ϵ_{if} have the same distributional assumptions as in the procurement auction case. Firm f multiplies the product of

market share and per-unit profit margin, i.e., it solves

$$\max_p \left\{ \frac{e^{\lambda + \alpha p + \xi_f}}{\Lambda^\sigma (1 + \Lambda)^{1 - \sigma}} (p - mc_f) \right\}. \quad (39)$$

where $\Lambda = \sum_{f' \in \mathcal{F}} e^{\lambda + \alpha p_{f'} + \xi_{f'}}$.

Thus, bid setting under the procurement scoring auction described in the body of the main text is isomorphic to price setting in a nested logit demand system.