IN DEFENSE OF THE MIDDLEMAN: QUALITY FAILURES IN THE GENERIC PHARMACEUTICAL MARKET

CATHERINE ISHITANI
Yale University

This paper studies how intermediary buyers discipline quality in the generic pharmaceutical market. Consumers are uninformed about generic quality and implicitly rely on intermediaries, such as wholesalers and pharmacies, to select quality for them. Using the universe of FDA manufacturing quality disclosures from 2000-2022, I first document pervasive quality failures—showing that generics are not interchangeable, as is conventionally assumed. Disclosing quality failures through recalls reduces purchases of low-quality drugs by 60% for over a decade. To isolate the role of intermediaries, I develop a scoring auction model of generic procurement that is identified with incomplete price data. I find that intermediaries are willing to pay a 2% premium for every 10% reduction in recalls. Their preferences, rather than manufacturer responses, drive the quality benefits of disclosure and increase the equilibrium share of high-quality drugs by 27%. Finally, I show that modest subsidies can improve the alignment between intermediary and social objectives, yielding greater welfare gains than interventions that limit the supply of low-quality drugs.

KEYWORDS: Healthcare, quality regulation, disclosure, intermediation.

1. INTRODUCTION

Quality disclosure is supposed to improve choice and welfare (Akerlof, 1970). However, its impact is often limited by frictions in consumer responses. Expansive evidence shows that consumers frequently overlook or ignore disclosures (Ben-Shahar and Schneider, 2017). Instead, in many markets they rely on intermediary buyers to demand quality

Catherine Ishitani: catherine.ishitani@yale.edu

I thank Alex Olssen, Guy David, and Abby Alpert for their invaluable guidance. I also thank Atul Gupta, Rena Conti, Elena Krasnokutskaya, Ashley Swanson, Eric Richert, Natalia Serna, Maria Polyakova, Yunan Ji, John Romley, Jacob Dorn, Yihao Yuan, and seminar participants at Wharton and several institutions for helpful suggestions. I gratefully acknowledge financial support from the Mack Institute and Leonard David Institute.

on their behalf. These firms are pervasive in settings with severe information asymmetry, from healthcare to household goods. For example, the typical consumer trusts that her pharmacy selects authorized COVID tests and that her grocery stocks safety-certified produce. This paper studies how informed intermediaries discipline quality in the market for generic prescription drugs.

The theoretical benefits of informed intermediaries are straightforward. Relative to consumers, intermediaries are often better positioned to observe and react to quality disclosures, given their scale and experience. Downstream, intermediaries can alleviate adverse selection and improve the quality of goods available to consumers. Upstream, they can incentivize suppliers to invest in higher quality. Intermediaries, however, may not fully internalize these social benefits, particularly if they lack reputational or legal incentives. This creates the conditions for a conflict of interest, in which intermediaries select goods that benefit themselves rather than consumers. In this context, intermediaries can enhance *or* distort the benefits of disclosure, with ambiguous implications for equilibrium quality and welfare.

The generic pharmaceutical market is distinguished by an unusual degree of asymmetric information and intermediation, making it an ideal setting to study informed intermediaries.[1] In contrast to branded drugs, generics are conventionally described as almost perfectly substitutable "commodities" (FTC, 2011). Patients and physicians typically cannot differentiate between generics made by alternative manufacturers, much less between differences in their quality. Instead, they implicitly trust intermediaries, such as wholesalers and pharmacies, to select high-quality drugs for them. These firms make purchasing decisions that determine the quality and wholesale prices of nearly all generics in the U.S. market (Fein, 2023). For their part, intermediaries promise to deliver safe and effective drugs, with slogans like, "It's not just a package, it's a patient."

Most patients perceive generics to be tightly regulated, commensurate with their high stakes for health. Although generics must meet a minimum clinical standard at entry, their *manufacturing quality* can deteriorate in the decades after. This type of quality reflects a

[1] A generic drug is a copy of a branded drug whose primary patents have expired; two-thirds of Americans consume these drugs each year (Cohen and Mykyta, 2023). A *drug* is a molecule-formulation made by a specific manufacturer (e.g., atorvastatin-tablet-Teva Pharmaceuticals).

drug's purity, strength, and ability to perform its approved therapeutic purpose—capturing, for example, whether it is contaminated with bacteria or has insufficient active ingredient.[2] The Food and Drug Administration (FDA) primarily maintains manufacturing quality through recall and inspection failure announcements. These disclose that a drug poses a "reasonable probability" of serious harm but do not permanently remove it from the market (FDA, 2024b). Recalls also require manufacturers to retrieve defective batches, which can temporarily disrupt supply.

I begin by providing evidence of pervasive quality failures in the U.S. generics market. Next, I show that disclosing these failures sharply reduces purchases of low-quality drugs. Finally, to isolate the role of intermediaries, I develop a structural model of generic procurement. I show that intermediaries enhance the effects of disclosure and substantially improve equilibrium quality, and then evaluate leading proposals to improve drug quality and supply reliability.

First, I construct a new dataset that provides the most comprehensive information on drug manufacturing quality to date. The data include the universe of FDA disclosure actions from 2000-2022, obtained through Freedom of Information Act (FOAI) requests. I use drug labels, import manifests, and administrative records to identify drug manufacturing locations—historically considered trade secrets—and to link drugs to quality disclosures (Conti et al., 2020). These data reveal that quality failures are pervasive and persistent: 57% of facilities fail inspections, and 12% of drugs are recalled during my study period, demonstrating that quality issues are not limited to a few bad actors. Failing drugs are much more likely to fail again, relative to the average drug, suggesting that quality resembles a fixed type. Taken together, these results contradict the conventional wisdom: generic drugs are not perfect substitutes.

Next, I explore whether disclosing quality failures improves equilibrium outcomes. This is not obvious *a priori*. Since patients face strong barriers to observing and using disclosures, the FDA's policy might be expected to have limited impact. I quantify the causal effects of disclosure using commercial claims data and a generalized difference-in-differences design that exploits variation in disclosure timing. I find that recall announcements reduce intermediary purchases of low-quality, recalled drugs by 60%. This effect

[2]These are representative examples of top quality failures, which are summarized in Table V.

persists for at least ten years—long after recalls are typically resolved—suggesting it is driven by long-run changes in intermediary preferences or manufacturer costs, rather than short-run supply disruptions. In contrast, disclosing even the most serious inspection failures has little effect on intermediary purchases. Manufacturers thus appear to be penalized more stringently for poor supply reliability than for low quality alone. Finally, I study other determinants of patient welfare, including drug prices and accessibility. While recalls have limited impact on prices, they increase the probability of a national drug shortage by 33% and reduce patient medication adherence by 5%. Thus, even as recalls improve quality, they also reduce patient welfare by disrupting drug supply.

The quasi-experimental evidence demonstrates that disclosure improves drug quality, but it cannot disentangle supply- and demand-side mechanisms. For example, disclosure might alter manufacturer *costs* and *prices*, as they reoptimize capacity and positioning. On the demand-side, it might shift *intermediary preferences*, as they update their expectations about supply reliability. To separate these mechanisms, I develop a structural model of intermediary purchasing through scoring auctions. The model is based on intermediary reference manuals, which specify their criteria for scoring manufacturers in auctions, such as historical recalls and fill rates (CardinalHealth, 2022, AmerisourceBergen, 2019). I capture this process using a scoring auction framework, in which intermediaries weigh both price and quality, and estimate the model using data from 70,000 auctions.

My estimation approach extends a recent method from Laffont et al. (2020) and introduces a new technique that non-parametrically identifies the model primitives when only winning prices are observed. While scoring auctions are common in practice, current estimation methods require complete data on prices, which is infeasible in most empirical applications. Identification with incomplete prices is non-standard because the auction model incorporates two sources of asymmetry: bidder characteristics and unobserved buyer preferences. My approach draws on the empirical auction literature to identify manufacturer costs (Guerre et al., 2000), and then imposes standard restrictions on bidders' cost distributions to identify intermediary preferences (Athey and Haile, 2002). I provide empirical support for these restrictions by exploiting variation in recall timing and showing balance in uncontrolled drug characteristics. Using Monte Carlo simulations, I show that my algorithm recovers unbiased estimates of the model primitives. Finally, I generalize my ap-

proach to multiple scoring auction formats, including linear, optimal, and quality-to-price ratio scores.

The estimated model shows that intermediary responses are central to the equilibrium effects of disclosure. Intermediaries prefer high-quality suppliers and are willing to pay a 13% premium for them (roughly 2% for each 10% reduction in recall risk). Their preferences account for over 100% of quality improvements observed after recalls, and increase the share of high-quality drugs by 18%. Manufacturer cost and pricing responses offset 60% of these gains, lowering quality by 11 percentage points (pp). On net, disclosure benefits patients but reduces producer surplus by 14%, due to increased price competition and failure-to-supply penalties. This loss helps explain why—counter to economic theory—manufacturers do not voluntarily disclose their quality types, and regulation is required in the first place (Grossman, 1981, Milgrom, 1981).

Next, I quantify the equilibrium role of intermediaries. Simulating a counterfactual environment without quality preferences shows that intermediaries increase the share of high-quality drugs by 27%. In contrast, they have limited effects on prices and producer surplus (which each increase by less than 3%). Overall, intermediaries improve short-run quality and welfare, and thus partially correct market failures from patient asymmetric information. By encouraging competition on quality, and not just price, they may also strengthen manufacturer incentives to invest in quality in the long-run.

In my last counterfactuals, I study the welfare implications of policies that seek to better align intermediary and social incentives, or to alter the composition of manufacturers in the U.S. market. During a period of historic pandemics and drug shortages, the federal government has proposed four such policies: expanding quality disclosures, subsidizing purchases of high-quality drugs, raising minimum quality standards, and reshoring drug production to the U.S. (HHS, 2024, GAO, 2024, The White House, 2025). The quasi-experimental evidence indicates that intermediaries already respond strongly to quality signals, suggesting that more salient disclosures may yield limited gains. In contrast, the structural model shows that modest subsidies could increase consumer welfare by reinforcing intermediary incentives to select high-quality drugs. Improving quality, however, is increasingly expensive, and reaching an 80% high-quality drug supply would require price increases of more than 30%. In the long-run, large subsidies are also likely to drive low-quality manufacturers from the market, heightening supply risks in markets already prone to shortage.

Raising quality standards could improve average quality by approximately 60%, but at the cost of reduced competition and nearly 1.25 times greater prices. Reshoring production produces the least efficient quality gains, increasing prices by a factor of 2.8 for only 50% higher quality. These outcomes reflect more expensive domestic production (with 9% higher marginal costs in the U.S.), lower competition (with 50% fewer bidders), and poor targeting (as baseline U.S. quality is only slightly higher). Finally, to evaluate the welfare consequences of these proposals, I apply laboratory evidence on the value of quality. The FDA estimates that 145 recently recalled drugs increased the incidence of cancer by approximately 1 in 8000, implying that higher-quality versions may be worth at least 20% more (at standard values of Quality-Adjusted Life Years (QALYs)). Overall, this analysis suggests that moderate subsidies would improve welfare, and are likely to yield substantially greater gains than proposals that raise quality standards or reshore manufacturing.

Although this paper's focus on drug manufacturing quality is new, research on quality disclosure is well-established (Akerlof, 1970). Prior work documents its effects in settings such as restaurants, packaged foods, toys, and insurance (Dranove and Jin, 2010, Freedman et al., 2012, Barahona et al., 2023, Vatter, 2022). In general, these studies find that disclosure can improve consumer choice and quality, but its impact is often attenuated by consumer responses. Motivated by this result, I study how *pharmacies and wholesalers* respond to quality information, connecting evidence on disclosure to work on expert intermediaries (Grennan et al., 2024, Gavazza and Lizzeri, 2021). I show that intermediaries partially resolve adverse selection problems between buyers and sellers and substantially enhance the equilibrium effects of quality disclosure.

My paper complements a sizable literature on pharmaceutical quality, providing the first economic evidence on quality *after approval* and its welfare implications (Peltzman, 1973, Atal et al., 2022). It also adds to prior research on generic drugs, which has addressed shortages (Yurukoglu et al., 2017, Dubois et al., 2023, Galdin, 2024), pricing (Scott Morton, 2000, Ganapati and McKibbin, 2023), and collusion (Cuddy, 2020, Starc and Wollmann, 2022). I build on this work by introducing data on manufacturing quality, which I show has first order implications for supply reliability and competition. Finally, it contributes methodologically to the literature on empirical scoring auctions. This research has largely focused on settings with complete bid data and known scoring rules (Marra et al., 2024, Allen et al., 2024). Krasnokutskaya et al. (2020) is most similar, in that the authors recover

an unknown scoring rule using incomplete data on quality (but complete prices) in Internet auctions. This paper extends the canonical estimation methods of Guerre et al. (2000) to a broader set of empirical applications with incomplete price and scoring data.

The paper proceeds as follows. Section 2 describes the empirical setting and data. Section 3 presents the quasi-experimental evidence, and Section 4 introduces the structural model. Section 5 discusses the counterfactual analyses, and Section 6 concludes.

2. INSTITUTIONAL SETTING AND DATA

2.1. *Generic Pharmaceuticals and Quality Disclosure*

Generic interchangeability underpins the basic organization of the U.S. pharmaceutical market, where it generates intense price competition. As a result, U.S. generics are priced 80% lower than their branded counterparts and capture 91% of prescription volumes (Cohen and Mykyta, 2023). However, growing evidence suggests that generic drugs may not be perfect substitutes. Generic quality failures lead to approximately 1,000 recalled or banned drugs each year and are the "foremost" cause of drug shortages (Woodcock and Wosinska, 2013, Jewett, 2023). For example, the discovery of a "cascade of failure" in quality control at Intas Pharmaceuticals led to widespread shortages of anti-epileptics and chemotherapies, delaying the care of up to 10% of U.S. cancer patients in 2023 (FDA, 2023, American Cancer Society Cancer Action Network, 2023). While regulators describe pharmaceutical supply as a matter of "national security," little research has examined the determinants and welfare implications of manufacturing quality, key inputs to evaluating policy proposals (The White House, 2024).

Internal FDA data indicate that quality failures are both pervasive and persistent. Over 2000-2022, 57% of manufacturing facilities failed inspections, and 12% of drugs were recalled for health risks.[3] These failure rates have also broadly increased over time, as shown in Figure 1 panel (a). Distributions of quality failures (panels (b)–(c)) show that they affect most facilities, manufacturers, and countries of origin, and thus cannot be attributed to a few bad actors. Nor can failures be explained by random shocks: a drug that fails once is 3.3 times more likely to be recalled over the next decade (relative to one that does not),

[3] A full 28% of facilities received the worst grade (OAI) over the same period. Results using only OAI grades are presented in the Appendix.

suggesting that generics are characterized by persistent quality types (Figure 7). Together, these patterns align with the FDA's narrative of gradually declining manufacturing quality, stemming from deteriorating physical capital with only "limited upgrades" over decades (Woodcock and Wosinska, 2013).

After a drug's approval, the FDA primarily regulates its quality through public disclosures, such as inspection reports and recall announcements. Inspection reports document violations of good manufacturing practices observed during routine inspections, while recalls are initiated when a drug violates the safety standards of the Food, Drug, and Cosmetic Act. I measure quality using the most serious types of disclosures: VAI and OAI inspection grades and class I and II recalls, which apply when a drug poses a "reasonable probability" of serious injury or death (FDA, 2024b).[4] Unlike inspection failures, recalls also require the removal of defective drug batches, often causing temporary stockouts. This creates a second channel—information about supply reliability—through which recalls may influence market outcomes.

The impact of the FDA's disclosure policies depends on the extent to which they increase demand for quality. However, patient awareness of generic quality is very low, for multiple reasons. Disclosures often fail to reach patients or omit actionable information, such as the drug's name (GAO, 2011). Most importantly, U.S. generics are poorly differentiated: they are typically unbranded and sold without identifying the exact product (i.e., the manufacturer). As a result, even if a patient should observe a disclosure, she would have almost no way of applying that information to her drug purchase.

In contrast, intermediaries such as wholesalers and pharmacies have strong incentives to monitor quality disclosures. Intermediaries purchase drugs from manufacturers to sell to patients, with the five largest firms collectively capturing over 95% of U.S. generic volume (Fein, 2023). These firms purchase generics through auctions, awarding supply contracts to a single manufacturer per market (for 1.5 years on average) (Cuddy, 2020, State of Connecticut, 2019). When choosing manufacturers, intermediaries weigh both price and quality, as quality failures impose significant costs on them. Stockouts reduce sales, and recalls require inventory tracking, with estimated labor costs of up to $230,000

[4] VAI and OAI inspection grades indicate "objectionable" or "unacceptable" manufacturing violations, respectively.

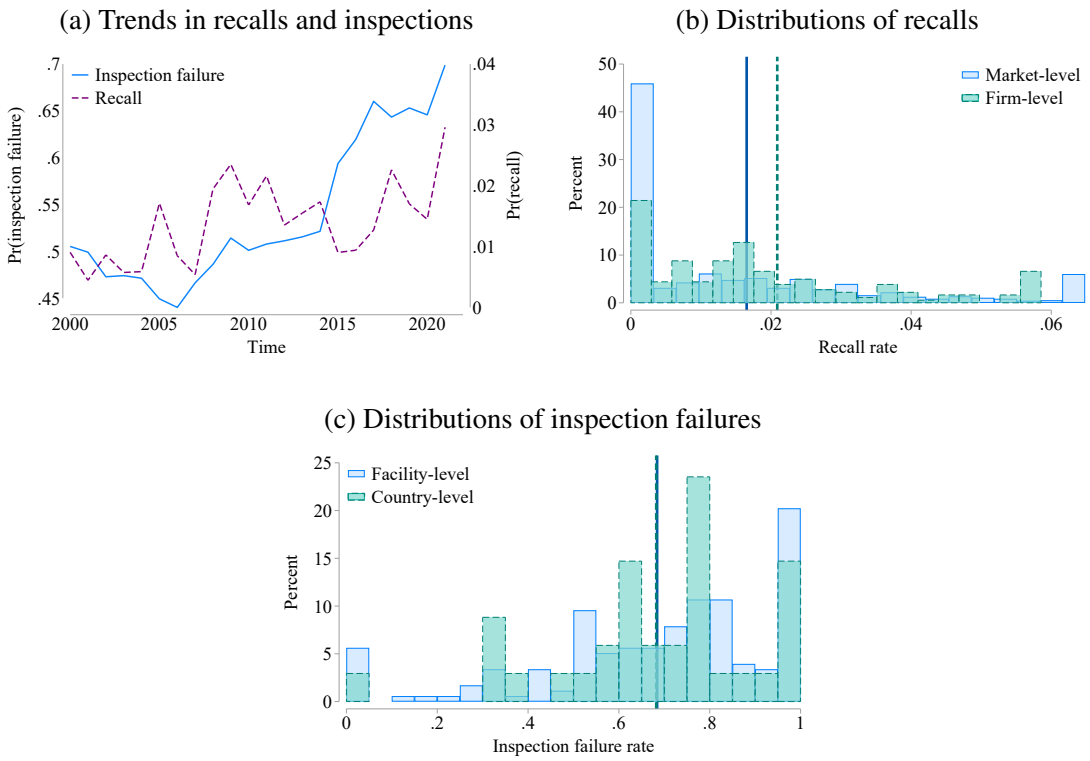


FIGURE 1.—Trends and variation in quality failures

Note: Panel (a) presents annual drug recall rates and facility inspection failure rates. Panels (b) and (c) present distributions of recall and inspection failure rates averaged over drug-years at the manufacturer, facility, market, and country levels. Distributions are winsorized at the 95th percentile, and vertical lines display mean rates.

per store annually (Tracelink, 2023). Consistent with this, intermediaries emphasize that "lapses in supply ruin relationship[s]," and their reference manuals specify the quality criteria used to "score" manufacturers (shown in Figure 5) (FTC, 2011, AmerisourceBergen, 2019, CardinalHealth, 2022).[5]

Since patients do not observe quality, intermediary choices determine the equilibrium quality and wholesale prices of nearly all U.S. generics. In theory, this gives intermediaries substantial scope to improve quality and mitigate welfare losses from asymmetric information. Intermediaries, however, face a classic agency problem and may choose socially suboptimal levels of quality, maximizing profits rather than total welfare (Gavazza and Lizzeri,

[5]Intermediaries do not release information about losing bids, and auction winners change rapidly—in certain markets, with every batch—implying limited roles for common cost information and inertia.

2021). This problem is particularly acute in the generics market, where intermediaries—unlike physicians or financial advisors—bear no legal responsibility to act in patients' interests. Moreover, since patients cannot observe or reward intermediaries' quality choices, it is not clear whether firms would respond to disclosure at all. In this context, whether intermediaries improve quality, and how close welfare is to first-best, are empirical questions.

2.2. *Data*

This paper combines three main data sources; the first is data on drug quality, including the universe of FDA quality disclosures and enforcement actions from 2000-2022. Inspection and enforcement records were obtained through FOIA requests, while recall data are from archived FDA Enforcement Reports. Linking drugs to quality disclosures typically requires drugs' manufacturing locations, which are trade secrets not releasable under FOIA (Conti et al., 2020). To overcome this, I use drug labels, trade manifests, and OECD administrative records to recover manufacturing locations.[6] For recalls, I observe announcement dates, classes, reasons, and National Drug Code (NDC) identifiers. For inspections, I observe dates, grades, citations, and whether the inspection was routine or "for cause."

The second source is data on drug prices, which measure weekly average acquisition costs for retail pharmacies over 2010-2022 (the "AAC" or "NADAC"). The data are from surveys conducted by fourteen state and federal Medicaid agencies and were collected through Open Records Requests and the Medicaid website. I exploit variation in pharmacy chain presence across states to recover intermediary-specific winning bids, which allows me to estimate my model non-parametrically.

The third source is administrative claims data from UnitedHealth, the largest private U.S. health insurer. The data cover a near census of U.S. pharmacies, as well as over 30 million individuals with pharmacy benefits per year from 2000-2022, including both employer-sponsored and Medicare Advantage enrollees. I use these data to identify intermediary auctions and to measure patient medication adherence. I identify auctions following Cuddy (2020): I infer winners from the drugs stocked by each pharmacy chain, and identify auction timing using manufacturer entry and changes in the composition of drugs across phar-

[6]This method was validated by CIDRAP; I scale it by two orders of magnitude to cover all generic drugs in a national claims dataset (Center for Infectious Disease Research & Policy, 2018).

macies. Manufacturers usually participate in each auction within a round, allowing me to identify the number of bidders from the manufacturers observed in each market.[7] I estimate the relative volumes of supply contracts using annual prescription counts per pharmacy. Finally, patient adherence is the Medication Possession Ratio (MPR), defined as the ratio of days supplied by any manufacturer to days prescribed. In addition, I use supplemental datasets on drug, pharmacy, and shortage characteristics from the FDA NDC Directory and Orange Book, the National Council for Prescription Drug Programs, and the University of Utah Drug Information Service (UUDIS).

My analysis sample consists of generic prescription drugs purchased by retail pharmacies over 2000-2022. I focus on unbranded, oligopoly ("multisource") drugs, which account for 84% of generic volume, and first recall and inspection announcements over 2001-2021 (1,159 and 1,483 events, respectively).[8] In total, I study the quality characteristics of 10,199 drugs across 1,204 molecule-form markets, and the purchasing responses of approximately 64,000 pharmacies and 63 million patients.

Table I summarizes the main characteristics of my sample. Column (1) reports values for all drugs, while columns (2) and (3) summarize drugs that were recalled and never-recalled during the study period, respectively. Each value represents the average over the full sample period (cols. (1) and (3)) or prior to the first recall (2). The typical drug is old and inexpensive: its first generics were approved more than 18 years ago, and it is priced at $0.70 per unit (e.g., pill). Columns (2) and (3) indicate that recalled drugs differ systematically from never-recalled ones, with greater volumes (56,651 vs. 14,765 prescriptions), larger markets (with 12.4 vs. 10.7 participants), and lower prices ($0.41 vs. $0.76). These differences motivate the use of not-yet-recalled drugs—rather than never-recalled ones—as a more appropriate control group for estimating the effects of recall disclosure.[9]

[7] I exclude collusive auctions, which I observe in unsealed court records (State of Connecticut, 2019).

[8] I focus on announcements from surveillance inspections, which are conducted quasi-randomly conditional on facility risk scores, to avoid contamination by other disclosure actions.

[9] Table VI presents the corresponding statistics for inspected drugs. This sample includes drugs that are matched to manufacturing locations and inspection data and represents 54% of drug volumes. It is observably similar to the main sample in Table I, with somewhat larger prescription volumes (28,668 versus 17,211).

TABLE I

DESCRIPTIVE STATISTICS

	(1) All drugs	(2) Recalled drugs	(3) Never recalled drugs
Drug characteristics			
Years since first generic entry	18.26	18.16	18.90
Fraction unit type:			
Oral pill	0.64	0.59	0.65
Injection	0.07	0.13	0.07
Topical & other	0.27	0.28	0.27
Extended release	0.10	0.10	0.10
Market shares & prices			
Share of pharmacies	0.10	0.12	0.10
	(0.18)	(0.15)	(0.18)
Mean annual prescriptions	17,211	56,651	14,765
	(66,340)	(168,526)	(54,001)
Mean price per unit	4.55	3.82	4.68
Median price per unit	0.70	0.41	0.76
	(20.67)	(15.72)	(21.43)
Mean drugs per market	10.73	12.44	10.66
N drugs	10,199	1,159	9,040

Note: The table presents descriptive statistics for multisource generic drugs in Optum claims data. Column (1) shows statistics for all drugs in the sample, column (2) shows recalled drugs, and column (3) shows drugs that were not recalled during 2000-2022. Statistics are drug-level means over all years (cols. (1) and (3)) or prior to the first recall (col. (2)). Corresponding standard deviations are reported in parentheses for key market share and price outcomes. Drug characteristics are from the FDA NDC Directory and Orange Book, and years since generic entry is as of 2022. Shares of pharmacies, annual volumes, and counts of drugs per market are observed in claims data. Prices are pharmacy acquisition costs in 2022 dollars from Medicaid surveys.

3. EVIDENCE ON THE EFFECTS OF QUALITY DISCLOSURE

The theoretical benefits of quality disclosure described in Section 2.1 rest on two assumptions: intermediaries must observe quality through disclosures and respond to incentives to select high-quality drugs. In this section, I provide evidence on the causal effects of recall and inspection announcements consistent with these assumptions. To estimate the effects of disclosure, I use a difference-in-differences framework ("DD") that compares outcomes for drugs that experienced disclosure to those that did not until later in the study period. For exposition, I refer to these drugs as "recalled" and "not-yet-recalled," respec-

tively, though I also apply the framework to study inspection disclosures. My primary event study specification is:

$$Y_{dt} = \alpha_d + \alpha_t + \sum_{s \neq -1} \beta_s \, D_{d,t+s} + \varepsilon_{dt}, \tag{1}$$

where Y_{dt} denotes the outcome for drug d in year t. A drug is a molecule-formulation made by a specific manufacturer, and a market (molecule-formulation) includes all generic bioequivalent drugs. I model the outcome as a function of drug α_d and year fixed effects α_t and unobserved time-varying factors ε_{dt}. $D_{d,t+s}$ is an indicator equal to one s years after the drug's quality is first disclosed; its coefficient β_s varies flexibly by year and is normalized to zero in the year before treatment. I estimate the model using drug-year-level data and cluster standard errors at the drug-level (the level at which treatment occurs). I also report the average effects on treated drugs by estimating the DD specification below:

$$Y_{dt} = \alpha_d + \alpha_t + \beta \, D_{dt} + \varepsilon_{dt}, \tag{2}$$

where the variable D_{dt} is equal to one after disclosure.

This design relies on the identifying assumption that the outcomes of recalled and not-yet-recalled drugs would have evolved in parallel in the absence of disclosure, and it exploits variation in disclosure timing to identify causal effects. While this assumption cannot be tested, the absence of differential trends prior to treatment provides evidence consistent with it. I plot the dynamic coefficients β_s and find no evidence of pre-existing trends that are correlated with disclosure timing. In formal tests of joint significance, I reject the null hypothesis of non-zero pre-trends for all key outcomes.

Disclosure is not random: for example, the FDA's risk-based inspection model provides one source of selection into treatment. The model prioritizes facilities based on drug characteristics such as "inherent risk," inspection history, and patient exposure. To control for the model's criteria, I include drug fixed effects to remove "inherent," unobserved, and persistent differences across drugs. In robustness checks, I add inspection grade covariates to control for time-varying differences in inspection histories, and I explore heterogeneity in the estimated effects by drug volumes as a proxy for patient exposure. Disclosure timing could also be correlated with outcomes through other channels. Intermediaries, for instance, might observe private quality signals (e.g., consumer complaints), allowing them to antic-

ipate disclosure, a possibility that I examine using more granular data. Alternatively, the FDA might relax quality enforcement for drugs at high risk of shortage (Brennan, 2016). In these cases, the estimates are likely to understate the true effects of disclosure, but other non-randomness might have opposite signed effects.

3.1. *Effects of Recall Announcements on Intermediary Purchases*

I begin by showing that recall announcements have large, negative, and persistent effects on intermediary purchases. Figure 3 panel (a) presents the event study coefficients from Equation (1) with each drug's share of pharmacies as the dependent variable.[10] Prior to disclosure, the estimated coefficients are insignificant and close to zero, suggesting no differential pre-trends ($p = 0.93$). Starting in the disclosure year, the trend for recalled drugs diverges sharply, with their share steadily declining over the next five years. Figure 8 provides additional detail on these dynamics using finer and coarser cuts of the data. Panel (a) zooms in on the *months* surrounding recalls. It shows that recalled and control drugs follow similar trends up until the last month before disclosure, suggesting anticipatory effects are minimal. Panel (b) zooms out to consider the *long-run* effects of recalls, showing that the recalled drug share continues to erode even a decade later (falling 68% below its pre-recall mean).

Table II presents the average effects, estimated using Equation (2). Recalled drugs experience a 43% decline in their pharmacy share (4.4pp, panel A, col. (1)), relative to not-yet-recalled drugs. While large, this estimate may actually understate the true effect, if drugs with larger responses are more likely to exit and are not observed in the sample.[11] In another sense, the magnitude of this effect is not surprising. Generic drugs are wholly undifferentiated apart from quality, which may cause potentially severe health consequences. Both features imply more aggressive substitution patterns than is typically observed for consumer products.[12]

---

[10] This measure captures intermediaries' "extensive margin" supplier choices, and not choices over volume.

[11] Consistent with this, excluding recalled drugs that eventually exit the sample yields significantly smaller effects (-2.5pp, Table IX, row 11).

[12] Volume should not be expected to fall by 100%, as recalled and inspected drugs need not exit the market.

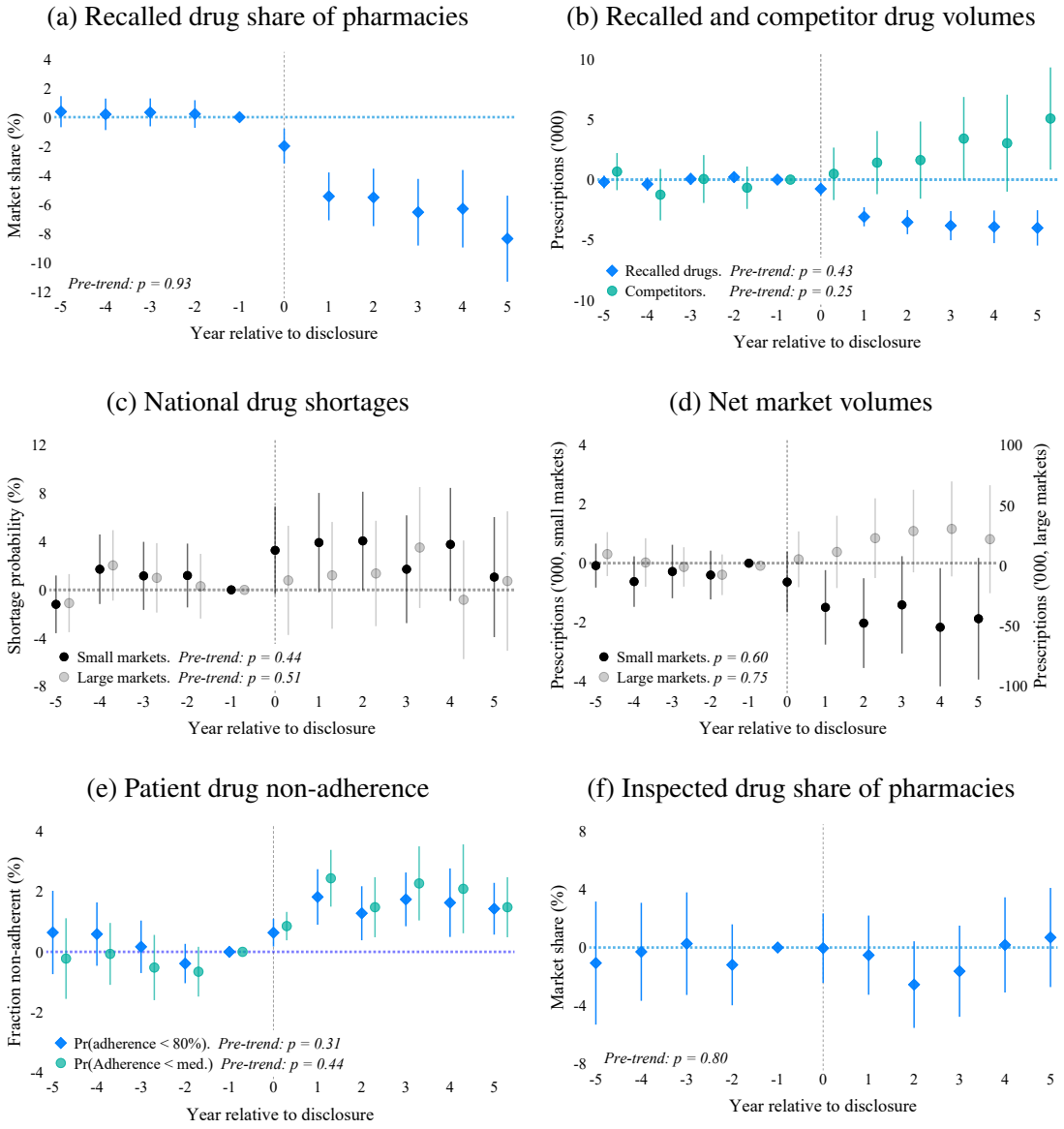


FIGURE 2.—Effects of quality disclosure on drug purchases

Note: The figure displays the estimated dynamic effects of quality disclosures, obtained using Equation (1). The sample includes multisource generic drugs in Optum claims, and the unit of observation is a drug- or market-year. The year prior to the first disclosure event is the omitted reference. The figures present the effects of recall disclosures on recalled drugs (panels (a)–(b)), bioequivalent competitors of the recalled drug (b), and net market effects ((c)–(e)). Panel (f) shows the effect of inspection failure disclosures on inspected drugs, obtained using a DDD specification (Equation (1) with an additional term, $\gamma\, D_{dt} \times 1\{\text{failed}_d\}$.) The dependent variables are each drug's share of pharmacies ((a), (f)) and prescriptions ((b), (d)), as well as the probabilities of new national drug shortages (c) and patient drug adherence (e). The figures present 95% confidence intervals with standard errors clustered by drug or market, and p-values from F-tests of the joint significance of pre-treatment coefficients.

Taken together, these results suggest that recall announcements sharply reduce pharmacy purchases of low-quality drugs, with effects that last far longer than the recall itself or most disclosure interventions (Ben-Shahar and Schneider, 2017). The persistence of these effects cannot be easily explained by supply disruptions—which are typically resolved within months—or by purchasing inertia, given the rapid supplier turnover observed in the data (GAO, 2011). (The average supply contract in my data lasts for 1.5 years, and the average supplier is chosen for 1.9 years in a row.) Similarly, I do not find significant effects for subsequent recalls. Persistent effects could, however, be explained by long-run changes in intermediary preferences or manufacturer costs, explanations I explore in Section 4.

Next, I consider intermediary substitution patterns and how competing manufacturers respond to disclosure. Figure 3 panel (b) presents event studies for recalled and competitor drugs, with drug volumes as the dependent variables.[13] After disclosure, the coefficients for recalled and competitor drugs break in similar but opposite-signed ways, consistent with intermediaries substituting from low- to high-quality drugs. Table II shows that competitor volumes increase by 12% on average, while recalled drug volumes decline by 60% (cols. (4), (2)). As a result, the average patient's probability of receiving a high-quality drug increases by 9% for 10 years—suggesting that recall announcements are strikingly effective, even when patients do not observe them.

In my robustness checks, I report coefficients from the estimators proposed by Callaway and Sant'Anna (2021) and Borusyak et al. (2024), which address concerns about DD estimates in staggered treatment designs. I also test the stability and significance of my estimates to using different outcome measures, market definitions, control groups, covariates, samples, weighting and clustering. The results, shown in Table IX, remain stable and typically fall within two standard errors of the baseline estimates.[14]

---

[13] Since prescription volumes are highly skewed, I estimate effects using logged values. Competitor and market level effects are estimated using data aggregated to the market-year level. My data capture a subset of U.S. drug volumes, implying that my estimates understate the effects of recalls on volume levels, while overstating their percentage effects.

[14] The two exceptions are from drugs that never exit the sample (see footnote 10) and when using a never-treated control group. The latter should be interpreted cautiously, as this group differs observably from treated drugs and exhibits significant pre-trends.

TABLE II: Effects of quality disclosures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Recalled drug		Competitor spillovers		Net market effect				Inspected drug	
Dependent var.	Share of pharms.	Log scripts	Share of pharms.	Log scripts	Log scripts	Pr(short -age)	Pr(non-adherence)	Log price	Share of pharms.	Log scripts
A - Average effects										
Disclosed * post	-0.044	-0.918	0.068	0.115	-0.037	0.021	0.021	-0.036	-0.005	-0.097
	(0.007)	(0.082)	(0.014)	(0.083)	(0.060)	(0.010)	(0.006)	(0.036)	(0.012)	(0.133)
Disclosed * post (%)	-43%	-60%	9%	12%	-4%	33%	5%	-4%	-8%	-9%
Disclosed * post (level)	–	-2,683	–	1,927	-1,071	–	–	–	–	-165
B - Heterogeneity by market structure										
Low participation: Disclosed * post	-0.069	-1.178	0.089	0.051	-0.141	0.027	0.031	-0.051	-0.033	-0.031
	(0.014)	(0.110)	(0.018)	(0.104)	(0.071)	(0.013)	(0.008)	(0.047)	(0.020)	(0.170)
High participation: Disclosed * post	-0.018	-0.602	0.015	0.153	0.075	0.015	0.012	-0.010	0.019	-0.118
	(0.005)	(0.117)	(0.014)	(0.097)	(0.069)	(0.018)	(0.011)	(0.039)	(0.013)	(0.203)
Observations	12,380	12,380	6,420	6,420	6,531	6,531	20,005	3,984	13,456	13,456
Dependent var. median	0.103	4,466	0.738	15,810	29,333	0.063	0.408	0.290	0.064	1,790

Note: The table reports the estimated average effects of disclosures obtained using Equation (2). It presents the effects of recall disclosures on recalled drugs (cols. (1)–(2)), bioequivalent competitors of the recalled drug ((3)–(4)), and net market effects ((5)–(8)). Columns (9)–(10) present effects of inspection failure disclosures on inspected drugs, obtained using a DDD specification (Equation (2) with $\gamma D_{dt} \times 1\{\text{failed}_d\}$.) The dependent variables are each drug's share of pharmacies and logged prescriptions, the probabilities of new national drug shortages and patient drug adherence, and logged volume-weighted average wholesale prices. To facilitate comparison, effects are presented in levels and as percentages of the dependent variable medians ("%"). Panel A presents average effects for all drugs in the sample, while Panel B presents results by baseline market structure, for low (N $\leq$10 drugs) and high participation (N > 10) markets. Standard errors are clustered at the drug- or market-level and reported in parentheses.

3.2. *Effects of Recall Announcements on Supply Reliability and Patients*

I now explore how disclosure affects the supply of drugs, using national shortages as a proxy for disruptions. Shortages are relatively rare—I observe fewer than 1,200 in my data—causing the effects to be imprecisely estimated. Nevertheless, I find that recalls increase the probability of a national drug shortage by 33% (significant at $p < 0.05$, Table II, col. (6)). Conceptually, these effects might be expected to be concentrated in small markets, where substituting between manufacturers requires proportionally greater capacity adjustments and fixed costs. Consistent with this, Figure 3 panel (c) shows that recall disruptions stem from small markets (with fewer than 10 manufacturers), which experience a 43% increase in shortages (significant at $p < 0.05$), while large markets experience no significant changes. The claims data confirm these patterns: panel (d) shows that small markets experience significantly larger declines in volume than large markets do (-13% versus 8% on average, Table II panel B, col. (5)). Overall, these results suggest that intermediary transitions between suppliers are not seamless and may generate short-run supply disruptions, particularly in concentrated markets.

To consider whether supply disruptions affect patient access to drugs, I look at changes in patient medication adherence. Following the literature, I consider patients adherent if their MPR exceeds 80% (Andrade et al., 2006). Baseline adherence is low—approximately half of patients regularly take their prescribed medications, in line with prior research (Watanabe et al., 2018). Following a recall, adherence declines rapidly, with the largest effects observed in the first post-recall year (Figure 3 panel (e)). On average, recalls reduce adherence by 2.1pp (5%), consistent with evidence that even modest hassle costs, such as a stockout at the nearest pharmacy, can meaningfully reduce adherence (McHorney and Spain, 2011). Non-adherence is associated with large morbidity and mortality costs of up to $500bn annually in the U.S. This suggests that supply disruptions caused by recalls may impose substantial patient health costs, which must be weighed against the benefits of improved drug quality. In contrast, I find little evidence that disclosure affects patients through drug prices. Recalls do not significantly affect prices (which fall by 4% on average; Table II, col. (8)), consistent with prior evidence from supply disruptions (Dave et al., 2018).[15]

[15]To facilitate comparison with the structural results, I present price results weighted by pre-recall drug volumes. Results using unweighted prices are similar (-5%).

I also examine whether the disclosure affects manufacturer participation. Table VIII col. (3) shows that markets gain 0.6 (10%) net participants on average after recalls. This increase, however, cannot explain the estimated declines in recalled drug purchases, as effects in markets with no net entry are statistically indistinguishable from the overall estimate (-0.029 versus -0.044; Table IX row 12 versus row 1).

3.3. *Mechanisms and Effects of Inspection Reports*

Finally, I consider potential mechanisms underlying the disclosure effects. To do this, I exploit differences in the information content between recall and inspection failure announcements. Both announcements reveal that quality falls below an acceptable threshold, but recalls also signal a heightened risk of supply disruption. Comparing the two effects therefore informs whether intermediary purchases respond to signals of drug *supply reliability* versus *quality* alone. To isolate the effects of negative quality information from inspections, I estimate a triple difference ("DDD") specification that compares outcomes for failing and passing drugs, before and after inspections. The specification is identical to Equation (2) but includes an additional term $(\gamma\, D_{dt} \times 1\{\text{failed}_d\})$, allowing me to difference out potential bias from selection into inspections. This strategy assumes that passing and failing drugs would have evolved similarly in the absence of inspections.[16] The coefficients from the event study (Figure VII panel (b)) are uniformly close to zero and insignificant (but the average effects are precise enough to rule out changes greater than 65% of the recall effects, Table II, cols. (5)–(6)). These results are robust to restricting failures to only the lowest grade, OAI, which signifies the FDA's most serious quality violations (Table VII).

In sum, disclosing even the most severe inspection failures has no detectable effect on intermediary purchases or drug quality. It is unlikely that intermediaries are unaware of inspection disclosures, or fail to interpret them as information about quality. Intermediaries state that they closely monitor inspection results, which are often widely publicized in industry press and can jeopardize government contracts (FDA, 2024a).[17] While recalls

[16] DD event studies for passing and failing drugs show very similar patterns before disclosure, supporting the use of the DDD design (Figure 9).

[17] See, for example, https://www.fiercepharma.com/warning-letters.

and inspection failures may signal somewhat different dimensions of quality, both indicate violations that may be "injurious" to patient health, implying that quality-seeking buyers should respond. Instead, one interpretation of the differential responses to recall and inspection announcements is that the market reacts to information about supply reliability rather than quality per se. Consistent with this, intermediary reference manuals describe penalties for recalls but otherwise make no mention of quality (CardinalHealth, 2022, AmerisourceBergen, 2019). From a policy design perspective, this suggests that intermediaries may face inefficiently low incentives to act in patients' best interests when selecting quality. Furthermore, disclosures about supply reliability information—rather than quality alone—may produce greater market responses and quality improvements.

3.4. *Summary of Disclosure Effects*

In summary, disclosure significantly improves the quality of generics available to patients, even when they cannot observe quality themselves. These improvements last for at least a decade—well beyond the duration of a typical recall—suggesting they may be driven by persistent changes in intermediary preferences or manufacturer costs. Finally, manufacturers are penalized more forcefully for signals of low supply reliability than low quality alone. Overall, this evidence is consistent with a story in which intermediaries observe quality disclosures, update their expectations about supply reliability, and respond by punishing low-quality manufacturers. Alternatively, manufacturers could adjust their costs and prices after recalls as they reoptimize capacity and positioning, responses that could also rationalize the effects. Since the DD framework identifies changes in equilibrium outcomes, it cannot disentangle these channels or predict how quality would evolve without disclosure. Doing so requires understanding how prices would evolve without intermediary and manufacturer responses; this, in turn, presupposes information about their preferences and costs. In the next section, I develop a structural model to recover these primitives and assess the welfare effects of intermediaries.

4. AN EMPIRICAL MODEL OF GENERIC PROCUREMENT

In this section, I develop and estimate a scoring auction model of generic procurement. The model characterizes how intermediary preferences and manufacturer costs jointly de-

termine equilibrium prices and quality. I reference the framework and notation developed by Laffont et al. (2020), which I adapt to an empirical setting with incomplete price data.

Suppose an intermediary wishes to procure a quantity of a generic drug. I assume it auctions off a single supply contract to $N \geq 2$ manufacturers in a first score, sealed bid auction (with no reservation score). Manufacturer bids are differentiated by two characteristics: the price b_i and the drug's historical quality q_i, which the intermediary weighs using a scoring rule. The intermediary observes manufacturer quality signals—and manufacturers observe intermediary scoring—through repeated interactions over years and across thousands of markets. By the start of the auction, quality and the scoring rule are thus assumed to be predetermined and common knowledge. The econometrician, however, does not observe the scoring rule and must estimate it.

Manufacturers are characterized by their quality type q_i and a random productivity shock θ_i, which jointly determine their cost $C(\theta_i, q_i)$ of producing a drug. Qualities may be freely correlated—for example, across drugs and over time—but manufacturer productivity shocks are assumed to be independent of their competitors'. That is, θ_i is independently distributed conditional on quality as $F(\cdot|q_i)$; furthermore, this distribution (but not its realization) is commonly known. Together, these assumptions situate the model within the conditional independent private values framework (Milgrom and Weber, 1982).

The intermediary receives a payoff of $U(q_i, b_i) = S(q_i) - b_i$ from awarding a drug contract with quality q_i. Given manufacturers' production technologies, Laffont et al. (2020) derive the intermediary's *optimal scoring* mechanism.[18] For each bid, the intermediary assigns the score

$$R_O(\theta_i, q_i) = S(q_i) - \frac{F(\theta_i|q_i)}{f(\theta_i|q_i)} C_\theta(\theta_i, q_i), \tag{3}$$

[18]The preference for quality $S(q_i)$ may potentially reflect expected supply disruption costs or altruism. The direct mechanism satisfies the manufacturers' incentive and rationality constraints within the space of first-score auctions. The empirical auction literature has focused on linear scoring rules, which are simple to implement and common in practice (e.g., Krasnokutskaya et al. (2020)). Since I do not observe the scoring rule, I check whether my results are robust to using both optimal and linear scores.

and selects the manufacturer with the highest R_O. In contrast to a *linear-in-price* score, $R_L(\theta_i, q_i) = S(q_i) - b_i$, the optimal score exploits the buyer's monopoly power to induce lower prices from low-cost manufacturers.

4.1. *Equilibrium*

Under the optimal score, Laffont et al. (2020) show that each manufacturer's probability of winning is determined by

$$P(\theta_i, q_i, q_{-i}) = \prod_{j \neq i} \left[1 - F\Big(\Gamma^{-1}(S(q_j) - S(q_i) + \Gamma(\theta_i, q_i), q_j)|q_j\Big)\right], \tag{4}$$

where $\Gamma(\theta_i, q_i) = C(\theta_i, q_i) + \frac{F(\theta_i|q_i)}{f(\theta_i|q_i)} C_\theta(\theta_i, q_i)$. Manufacturers choose prices to maximize their expected profits $[b_i(\theta_i, q) - C(\theta_i, q_i)]P(\theta_i, q_i, q_{-i})$. The first order condition yields the following type-symmetric Bayesian Nash equilibrium bidding strategy:

$$b_i = \sigma(\theta_i, q_i, q_{-i}) = C(\theta_i, q_i) + \frac{\int_{\theta_i}^{\overline{\theta}} C_\theta(u, q_i) P(u, q_i, q_{-i}) du}{P(\theta_i, q_i, q_{-i})}.^{19} \tag{5}$$

This equilibrium expression is similar to one from price-only auctions (e.g., Proposition 2 in Riley and Samuelson (1981)). It departs in two ways, which cause manufacturers of different quality types to bid different strategies in the same auctions. First, each manufacturer type is characterized by a potentially distinct cost technology $C(\theta_i, q_i)$ and productivity distribution $F(\cdot|q_i)$. Since the optimal bid is strictly increasing in costs and productivity, differences in these values generate asymmetric bidding behavior.[20] Second, each type faces different choice probabilities $P(\theta_i, q_i, q_{-i})$ reflecting intermediary preferences. For example, stronger preferences for quality increase the probability that high-quality types win, which enhances their differentiation and tends to increase their optimal prices. For low-quality types, quality preferences work in the opposite direction and tend to lower prices.

[19] Equilibrium relies on the following technical assumptions: manufacturers are risk neutral and know the number of bidders; costs increase monotonically in θ; and high- and low-quality firms do not lose with certainty.

[20] For exposition, I refer to the cost residual or inefficiency θ as "productivity."

4.2. *Equilibrium with quality disclosure*

Equilibrium prices and qualities are jointly determined by the buyer's scoring rule (Equation 3) and manufacturers' bidding strategies (Equation 5). The basic structure of the model changes in three ways when new information about quality is disclosed through recalls. First, recalls may alter intermediary preferences by allowing them to update their expectations about manufacturer reliability. Stronger preferences for high-quality (and in expectation, more reliable) manufacturers may reduce effective competition and increase prices. If quality is costly, disclosure may also lower the auction's productive efficiency by reallocating intermediary purchases toward higher-cost manufacturers. Second, disclosure may affect manufacturer costs. Recalled manufacturers may increase their effort or reduce their capacity to improve quality, increasing variable costs. By Equation (5), these changes are partly passed through to the intermediary—and ultimately to patients—in higher prices. Third, manufacturers may strategically adjust their prices in response to new information about quality and costs. For example, low-quality manufacturers may reduce their prices to rehabilitate their chances of winning. Whether disclosure is socially optimal depends on whether the observed quality gains outweigh the potential costs of reduced competition and productive efficiency.

I note that this analysis is only partial, since I do not model patient demand responses, or disclosure's long-run effects on manufacturer entry and investment. The quasi-experimental evidence, however, suggests that these forces are unlikely to have first order effects in the short- to medium-run. For example, the average market shows no detectable change in patient demand after disclosure, and changes in manufacturer participation are not correlated with intermediary purchases and drug quality. I leave the study of these long-run forces to future work.

4.3. *Estimation*

The key primitives of the scoring auction model are buyer preferences for quality $S(q)$, the cost function $C(\theta_i, q_i)$, and the distributions of manufacturer productivities $F(\theta_i|q_i)$. With complete data on prices, qualities, and the number of bidders, these objects are nonparametrically identified following the arguments of Guerre et al. (2000) and Laffont et al. (2020). My contribution is to show how the model can be identified using incomplete price data, which is common in empirical applications.

Identification with incomplete prices is non-standard because the model incorporates two forms of asymmetry: variation in *ex ante* bidder characteristics (costs and productivities) and in buyer preferences. Models with just the first form of asymmetry are identified from data on winning bids and winner's identities (Athey and Haile, 2002). I show that a model with both forms of asymmetry is also identified from these data (and complete data on qualities), for both optimal and linear scores. My approach is general and can be easily adapted to other scoring formats, such as polynomial scores (e.g., $S(q_i) - b_i^2$) and quality-to-price ratios.[21]

4.3.1. *Unobserved Prices*

First, I describe what information about the unobserved prices distributions can be recovered from winning prices. Winning prices are only observed when the bidder has the highest score, which in turn depends on all bidder types (Equation (3)). This suggests that the information that can be recovered will depend on the particular scoring rule and the composition of bidder types.

In symmetric auctions—where all bidders share the same type—the price distributions can be fully recovered from the winning prices. Here, the winning price is necessarily the minimum from each auction, and prices are independently and identically distributed (as bidder types are also identical). This allows me to recover the price distribution using the properties of the first order statistic (Athey and Haile, 2002). However, this method cannot be used to identify the price distribution in asymmetric auctions, where prices are neither independently nor identically distributed. To see this, note that equilibrium prices depend on q_i (Equation 5), implying different and potentially correlated distributions for each manufacturer type. Moreover, if buyers value quality, winning prices need not be minimum prices, unless they were submitted by the lowest-quality type.

For most scoring rules, identifying buyer preferences requires the price distributions from at least one asymmetric auction (i.e., one with variation in winner type). I now show that these price distributions can be non-parametrically recovered, up to a guess of the buyer preferences. For exposition, I focus on a *linear score* with two bidder types, $t \in \{H, L\}$. In the Supplementary Materials, I show the identification argument can be extended to other

[21] Derivations, proofs, and additional details on the estimation algorithm are in the Supplementary Materials.

scoring rules and an arbitrary, finite number of types. Let N_t refer to the number of type t bidders in an auction, and let prices be distributed as $b_t \sim G_t(\cdot|q_i = t, q_{-i})$. I first rewrite the observed winning price distributions using Bayes' Rule:

$$Pr(b_{t,i}|i \text{ wins}) = \frac{Pr(i \text{ wins}|b_{t,i})Pr(b_{t,i})}{Pr(i \text{ wins}),}$$

where I suppress conditioning on $q_i = t, q_{-i}$ throughout. After substituting in the linear scoring rule and rearranging terms, the following holds for high-quality types:

$$Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) = \big[1 - G_L(b_{H,i} - \Delta S)\big]^{N_L}\big[1 - G_H(b_{H,i})\big]^{N_H - 1} g_H(b_{H,i}).$$

Note that the buyer preferences for quality have been normalized relative to the lowest quality type, as $\Delta S = S(q_H) - S(q_L)$. The preferences are thus identified up to a location, as in discrete choice models. Similarly, for low-quality types, I can write:

$$Pr(b_{L,i}|i \text{ wins})Pr(i \text{ wins}) = \big[1 - G_H(b_{L,i} + \Delta S)\big]^{N_H}\big[1 - G_L(b_{L,i})\big]^{N_L - 1} g_L(b_{L,i}).$$

In these equations, the left-hand side (LHS) is the ratio of two probabilities: the likelihood that the winning price b_t is observed, and that the focal bidder i wins. Both objects can be estimated from the data using standard kernel methods. This implies that these two equations define a system of two nonlinear functional equations in the two unknowns I wish to recover, $g_L(b_L)$ and $g_H(b_H)$. Intuitively, this identification strategy exploits two pieces of information from each auction. First, knowing the type of the winning bidder tells me about the relative order of the lowest b^L and b^H, up to the guess of ΔS. Second, the observed winning price provides a lower bound for the losing one, allowing me to carve out the marginal probability of each price.

To estimate $g_L(b_L)$ and $g_H(b_H)$, I first recover the LHS objects using standard non-parametric kernel density estimators (with triweight kernels and rule-of-thumb bandwidths, Li et al. (2000)). I then discretize the prices and solve for $g_L(b_L)$ and $g_H(b_L + \Delta S)$ at each point in the price grid, starting with the lowest price $\underline{b}$. A solution to this system exists at each price, up to defined cutoffs, as long as the grid is sufficiently fine.[22] Above these

[22] The cutoffs are $\overline{b} - \Delta S$ for b_L and $\overline{b} + \Delta S$ for b_H.

cutoffs, one bidder type wins with certainty, implying the losers' bids are not defined by the model. Monte Carlo simulations (presented in Supplemental Figures 2–3) indicate that this procedure recovers the true bids and preferences without bias.

4.3.2. *Costs and Productivities*

With complete price data, manufacturer costs and productivities can be identified by extending the arguments of GPV (Laffont et al. (2020)). Since the bidding function (5) is strictly monotonic in θ_i, a firm's probability of winning given θ_i is equivalent to its probability given $b_i(\theta_i)$; in equations, that is $P(\theta_i, q_i, q_{-i}) = H(b_i(\theta_i), q_i, q_{-i})$. The bidding strategy can therefore be rewritten, after differentiating with respect to θ_i, as:

$$b(\theta_i, q_i, q_{-i}) = C(\theta_i, q_i) - \frac{H(b_i, q_i, q_{-i})}{H_b(b_i, q_i, q_{-i})}. \tag{6}$$

To apply Equation (6) with partial price data, I first rewrite H using Bayes' Rule:

$$H(b_i, q_i, q_{-i}) = \frac{Pr(b_i | i \text{ wins}, q_i, q_{-i}) Pr(i \text{ wins} | q_i, q_{-i})}{Pr(b_i | q_i, q_{-i})}, \tag{7}$$

where each object on the right-hand side (RHS) was already recovered when estimating the price distributions. Using $\hat{H}(b_i, q_i, q_{-i})$ and prices simulated from $\hat{G}(b|q_i, q_{-i})$, I then apply Equation (6) to recover the distribution of costs for each bidder type. To recover productivities, I follow Laffont et al. (2020) in assuming that the cost function is multiplicatively separable in θ_i and q_i. Productivities are then estimated as the residual from projecting logged costs on bidder quality:

$$\log \hat{C}(\theta_i, q_i) = \log C_0(q_i) + \log \theta_i.$$

4.3.3. *Intermediary Preferences*

In principle, having recovered the cost functions and productivity distributions, buyer preferences could be estimated by simulating winning prices for each guess of ΔS, and then selecting the value that minimizes the distance between the observed and simulated prices (yielding a nested fixed point algorithm (Rust, 1987)). In practice, solving for equilibrium prices in asymmetric auctions (the "inner loop") is computationally demanding

and typically requires numerical methods that introduce approximation error (Hubbard and Paarsch, 2014).[23]

Instead, it is computationally convenient to reformulate the estimation problem from one in *price space* to another in *productivity space*. This approach is valid because each guess of ΔS maps to a unique equilibrium price solution (Lebrun, 1999). Conceptually, my approach parallels two-step methods for estimating dynamic games, which approximate a challenging inner loop by mapping it to probability space and forward simulating it (Hotz et al., 1994, Aguirregabiria and Mira, 2007). In my inner loop, I map the prices from asymmetric auctions to estimated productivities, which I then compare against the "true" productivities from symmetric auctions in my outer loop. While my approach is tractable and avoids approximation error, it introduces an additional source of estimation error that decreases in the sample size.

For the *optimal score*, I follow Laffont et al. (2020) and repose the estimation problem as a discrete choice one. I rewrite the optimal scoring rule in Equation (3) as follows: a buyer selects bid i if and only if

$$S(q_i) - \hat{\eta}_i \geq S(q_j) - \hat{\eta}_j \quad \forall j \in \{1, ..., N\}, \;\; j \neq i, \tag{8}$$

where the distribution of the "random term" $\hat{\eta}_i = \Gamma(\hat{\theta}_i, q_i)$ is a known function of the productivities θ_i and has thus already been recovered. In this case, Matzkin (1991) shows that the buyer preferences ΔS are non-parametrically identified and can be estimated by maximizing the following likelihood equation:

$$\hat{\Delta S} = \text{argmax } \mathcal{L}(\Delta S) = \sum_{a=1}^{A} \sum_{i=1}^{N} 1(i \text{ wins } a) \log \hat{P}(i \text{ wins} | q_i, q_{-i}, \Delta S),$$

[23] The prices from a linear score solve a non-linear system of ordinary differential equations, where the Lipschitz condition does not hold near $\overline{\theta}$, implying standard numerical methods do not apply. Hubbard and Paarsch (2014) compare numerical strategies for solving this type of system, which introduce approximation error and are not guaranteed to be stable given estimation error in $\hat{f}(\theta)$ (Kastl et al., 2025).

where A is the number of auctions, and the indicator $1(i \text{ wins } a)$ is observed in the data.[24] The final term, the *ex ante* choice probability $P(i \text{ wins}|q_i, q_{-i}, \Delta S)$, is equal to $\int_{\underline{\theta}}^{\overline{\theta}} \hat{P}(\theta_i, q_i, q_{-i}) \hat{f}(\theta_i|q_i) d\theta_i$, where $P(\theta_i, q_i, q_{-i})$ is defined in Equation (4).

This identification strategy rests on two assumptions. First, bidder productivities must be identically distributed across symmetric and asymmetric auctions, conditional on the bidder type and auction-level characteristics. This assumption is standard in the empirical auction literature, where it is invoked to allow researchers to pool data across auctions (GPV). I discuss potential threats and evidence for it later in the section. Second, the buyer's scoring rule must depend on only the focal bidder's characteristics. This assumption only pertains to the optimal score, as most scoring rules depend on all bidder characteristics through the price. Together, these assumptions imply that the distribution of optimal scores in symmetric auctions identifies those in asymmetric ones, allowing estimation via Equation (8).

Next, I propose a general method that identifies buyer preferences under alternative scoring formats, including the *linear score*. My approach allows scores to depend on all bidder characteristics.[25] To pin down preferences, I match the distributions of $\hat{\theta}_i$ from asymmetric auctions—which are conditional on the guess of ΔS—to the $\hat{\theta}_i$ recovered from symmetric auctions—which are not. As before, the identifying assumption is that the distribution of θ_i is conditionally identical across auctions. I then estimate ΔS using maximum likelihood:

$$\hat{\Delta S} = \text{argmax } \mathcal{L}(\Delta S) = \sum_{a=1}^{A} \sum_{i=1}^{N} \log \hat{f}_{\text{asym}}(\hat{\theta}_{\text{sym},i}|q_i, \Delta S).$$

I estimate the model using bid-level data on drug quality, auction characteristics, and winning prices. The sample is limited to oral generics, which are manufactured using similar production technologies and represent over 70% of volumes, and to auctions held by the five largest intermediaries. I assume that intermediaries evaluate quality using one collective, optimal score, and that historical quality is an indicator of whether a drug was recalled

[24] Matzkin's maximum likelihood procedure does *not* involve parametric assumptions on the error term, since its distribution has already recovered. In this sense, estimation of ΔS is fully non-parametric.

[25] In this case, the distribution of scores differs across symmetric and asymmetric auctions, so the distribution of $\hat{\eta}_i$ in Equation (8) is not known.

in the last five years.[26] In Section 4.4, I assess robustness to each of these assumptions by using alternative scoring rules and definitions of quality.

The auctioned supply contracts in my data are observably heterogeneous—differing, for instance, in their realized volumes and drug characteristics. I follow Haile et al. (2003) to semi-parametrically control for auction-level heterogeneity from realized volumes, intermediary identities, the number of bidders, and drug characteristics that affect costs, such as the therapeutic area and delivery mechanism.[27] Year fixed effects control for strong deflationary price trends. Table XI presents the estimated coefficients, which are consistent with significant and decreasing economies of scale, modest prices variation across intermediaries, and higher prices for extended-release drugs. Overall, the estimated patterns are very similar to those found by Cuddy (2020) in this market. After trimming the top 5% of homogenized prices and auctions by size ($N > 9$ bidders), my final sample includes 276,495 bids from 67,880 auctions over 2010-2022. From this, I set aside observations from 2022, which I use to evaluate the fit of my model.

Table XII presents the descriptive statistics for this sample. Low quality is relatively common: 27% auctions have at least one low-quality bidder, and 5% have two or more. Figure 11 plots the distributions of winning homogenized prices. The prices exhibit high variance (with a coefficient of deviation of 0.55), and high-quality prices (solid blue) nearly stochastically dominate low-quality ones (dashed green). These patterns could potentially reflect both sources of bidder asymmetry: low-quality manufacturers might have lower costs, and they might set lower markups in response to intermediary preferences for quality. A key objective of the estimation strategy is to disentangle these sources.

[26] This definition is consistent with the quasi-experimental evidence, which shows persistent effects from first recalls, but limited effects from subsequent ones. For the structural model, I define a market as a molecule-formulation-strength combination, reflecting the level at which generic auctions are held. This is narrower than the standard market definition (molecule-formulation), used to study substitution patterns in Section 3.

[27] The adjustment assumes auction-level heterogeneity enters prices in a separable fashion. It preserves equilibrium bidding for optimal and linear scores.

4.4. *Results*

Figure 3 presents the key objects recovered from the structural model: the distributions of manufacturer costs and productivities, as well as intermediaries' preferences for quality. Each object is shown in homogenized price terms for high- (solid blue) and low-quality (dashed green) manufacturers. Panel (a) shows the first object, the distribution of manufacturer unit costs $C(\theta_i, q_i)$. On average, high-quality drugs cost 31% more to produce, with costs that stochastically dominate those of low-quality drugs. The cost estimates are credible, in the sense that they imply price-cost margins close to those reported by generic manufacturers. This comparison is only approximate, since the reported margins include unrelated line items (e.g., specialty drug profits). Nevertheless, it is reassuring that the estimated margin (25%) falls within the manufacturers' reported range of 18-29%.[28] Panel (b) plots the second object, the estimated productivity distributions. After controlling for average cost differences, high- and low-quality manufacturers exhibit similar productivities. Above the 80th percentile, low-quality types appear slightly less efficient.

The third and most important object is the buyer's preference for quality, which governs the extent to which intermediaries discipline quality in my model. I estimate that intermediaries are willing to pay up to 13% more to purchase high-quality drugs (0.75 relative to the mean homogenized price of 5.88). Panel (c) illustrates how these preferences shape intermediary choices probabilities, showing that low-quality manufacturers' probability of winning (dashed line) declines farther and faster in prices, relative to high-quality types (solid line). One interpretation of the preference estimate is as intermediaries' expected value from avoiding future supply disruptions. Since recalled drugs are 3.4 times more likely to be recalled again within five years, this implies that intermediaries are willing to pay a 2% premium for each 10% reduction in their future recall risk.[29]

Intermediaries are estimated to place substantial weight on quality, breaking with standard characterizations of the generics market. The FDA, for example, has described the

[28] To calculate margins, I take the ratio of the estimated winning prices and costs and weight them by the realized volume of each contract (Starc and Wollmann, 2022). I compare the median estimated margin to the operating margins reported by the seven largest public manufacturers in my data, excluding a cartel period (2013-2015).

[29] Previously recalled drugs are recalled 15.5% of the time, versus 4.5% for not recalled drugs, implying an elasticity of $.13/((.045/.155)-1)*.1 = 2\%$.

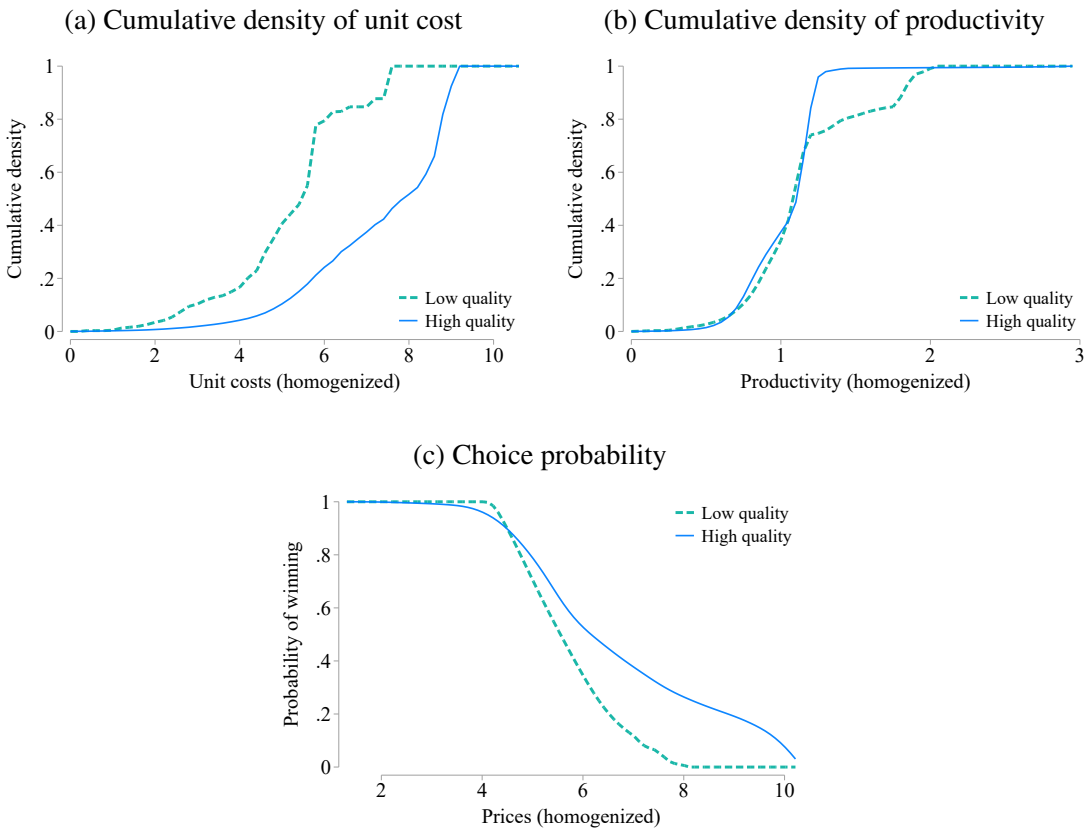


FIGURE 3.—Scoring auction model estimates.

Note: The figures show the key estimated objects from the optimal scoring auction model, measured in homogenized dollar terms. Panel (a) presents the kernel density of the cost per unit, while panel (b) presents the density of productivities θ_i. Panel (c) plots the choice probabilities for each bidder type as a function of the price.

market as "fundamentally" unable "to observe and reward quality" or to "penalize manufacturers that fail to invest in... a reliable supply" (FDA Drug Shortages Task Force, 2019). On the other hand, the estimates closely align with intermediaries' own statements about sourcing, which describe reliability as a first order concern (Section 2). They are also broadly consistent with quality preferences reported from other procurement auctions, which range from 19%-68% (Spagnolo, 2012, Krasnokutskaya et al., 2020, Jääskeläinen and Tukiainen, 2019, Albano et al., 2009). My estimates may be more conservative because generic intermediaries are relatively insulated from the costs of low quality: they face limited reputational exposure, are reimbursed for direct recall costs, and are not liable for patient injuries (Fink, 2012).

4.4.1. *Tests of the Identifying Assumption*

Identifying buyer preferences required the standard assumption that bidder productivities θ_i are conditionally identically distributed across symmetric and asymmetric auctions. I now discuss the empirical content of this assumption before providing evidence for it. For exposition, I consider an asymmetric auction with two quality types $A := \{N_H, N_L\}, N_H \in (0, N)$ and $N_H \in (0, N)$, which I compare to a symmetric, high-quality auction $S := \{N_H = N, N_L = 0\}$ that differs only in that it has not yet experienced a recall.

At least three sources of non-random variation threaten the identifying assumption. First, recalls may not be randomly assigned. For example, if low cost (and low θ_L) manufacturers are disproportionately recalled, then the productivities in auction A will systematically differ from those in S (in this case, $\theta_{H,A} > \theta_{H,S}$). To test for this source of *reclassification bias,* I compare the distributions of productivities in (not yet) symmetric and (not yet) asymmetric auctions, estimated using data from just before each market's first recall. Intuitively, this test exploits the same source of variation as the quasi-experimental design: if recall timing is exogenous, the closeness of the distributions in Figure 13 panel (a) (confirmed using a Wald test, $p = 0.42$) suggests that the identifying assumption holds.

Since I cannot directly compare the productivities in symmetric and asymmetric auctions after recalls, I use a complementary "balance test" approach during this period.[30] I compare firms' observable characteristics across auctions, focusing on features that are correlated with productivity but are not controlled for in the first stage homogenization. Panels (c)–(h) show the distributions of manufacturer sizes, which affect productivity through economies of scale. I find no systematic differences in manufacturer sizes across symmetric and asymmetric auctions, suggesting their productivities are also balanced after recalls ($p = 0.56$).

A second threat may arise if manufacturers *selectively participate* in auctions after recalls—for example, if recalls reduce competition and induce entry of high-cost types. To test this, I re-estimate the model using auctions that do not experience changes in bidder participation. The results are qualitatively similar to the baseline estimates (Table XIII row 5 versus row 1), suggesting that selective entry is not a major threat to identification. A final

[30] Estimating productivities in asymmetric auctions requires knowing buyer preferences, which in turn requires the identifying assumption.

concern is that high-quality manufacturers may *strategically adjust their costs* in asymmetric auctions. While I can estimate how low-quality manufacturers adjust their costs after being recalled (Table XIII row 4), I cannot separately identify adjustments for their high-quality competitors. I therefore assume that any changes in costs after recalls are attributed to low-quality manufacturers, and that these adjustments do not depend on the composition of competing bidder types.

4.4.2. *Model Fit and Robustness*

In addition to the cost and preference benchmarks discussed above, I assess model fit in two ways. First, I consider the model's ability to predict post-recall outcomes *in sample*. Using pre-recall data, I predict equilibrium prices and winning shares by type, which I then compare against the observed data. Figure 12 panel (a) presents the winning price distributions, and panel (c) the shares of high-quality drugs. The close correspondence between the model's predictions (dashed lines) and observed values (solid lines) indicates the good fit of the model in sample. Next, I repeat this exercise using post-recall data from the year 2022, which was not used in estimation. Panels (b) and (d) show that the predicted fit is also excellent *out-of-sample*.[31]

Finally, I assess the robustness of my results to perturbations in the sample construction and maintained assumptions. Table XIII presents intermediary preferences (col. (1)) and average differences in high- and low-quality costs (col. (2)) estimated with these perturbations. The baseline estimates assume optimal scoring and that "low-quality" drugs are recalled in the last five years (row 1). Alternative specifications apply broader definitions of quality (2), assess heterogeneity by auction size (3), consider auctions with only eventually recalled drugs (5) or fixed participation (5), or assume linear scoring (6). Across all specifications, I find positive preferences for quality and greater costs for high-quality drugs. The consistency of these patterns suggests that the model's key insights are robust.

[31] Confidence intervals are not typically reported in the empirical auction literature, since it is not clear that non-parametric estimators converge in a distributional sense. Ma et al. (2019) have recently proposed pointwise confidence intervals for price-only auctions, but not scoring ones.

5. COUNTERFACTUAL ANALYSES

In this section, I use the estimated model to decompose the equilibrium effects of quality disclosure and to study the welfare consequences of informed intermediaries. I then evaluate federal policy proposals that seek to improve generic quality and reliability. Throughout, my outcomes of interest are equilibrium quality, prices, and manufacturer and intermediary surplus.

To assess patient welfare, I focus on quality and prices. Quality affects patient well-being through multiple channels: low quality can cause serious injury and increases the risk of supply disruptions, which indirectly reduces patient adherence and health outcomes. I approximate the value of quality using laboratory estimates of health effects from recent recalls. Drug prices affect patients in the form of higher retail prices and insurance premiums.

To measure producer surplus, I consider changes in manufacturer and intermediary variable profits, net of one-time reimbursement costs.[32] Intermediary surplus is computed from their payoffs $U(q_i, b_i) = S(q_i) - b_i$, which are a function of *ex post* revealed quality (as quality is stable over the study horizon). Since intermediary preferences are identified up to a location normalization, I solve for the location $S(q_i = 0)$ that maximizes surplus under the estimated parameters.[33] All outcomes in expressed dollars and weighted by realized contract volumes.

For each counterfactual environment, I simulate 50 draws of productivities per bidder from $F(\cdot|q_i)$. I calculate bidder costs and intermediary choice probabilities, and then integrate over these objects to find equilibrium prices. Winners are determined by the optimal scoring rule, which fits the data well and is computationally convenient. Finally, I compute manufacturer and intermediary surplus, holding patient demand, manufacturer participation, and drug quality fixed.

---

[32] Intermediary contracts typically stipulate "failure-to-supply" clauses that require manufacturers to reimburse direct recalled drug costs. I assume 10% of recalled drug volumes are returned, based on volumes in the recall data; results are not directionally sensitive to this assumption.

[33] This value (7.40 in homogenized dollars) is close to one obtained by calibrating surplus to match intermediaries' average reported profit margins (7.85) (Sood et al., 2017).

5.1. *Quality Disclosure Mechanisms*

I begin by studying the FDA's status quo disclosure policy and decomposing its supply- and demand-side mechanisms. Section 3 showed that disclosure improves quality without significantly changing prices. To quantify how it affects producer surplus, I simulate equilibria before and after recall announcements, updating q and redrawing θ to reflect the new quality information.[34]

Consistent with the quasi-experimental estimates, Table III shows that disclosure increases the share of high-quality drugs (by 6%), with limited effects on prices (-11% on average and 0% at the median, column (3)). On the supply-side, disclosure raises realized manufacturer costs (by 6%), implying a modest decline in productive efficiency. Manufacturer profits fall substantially (by 28%), as disclosure increases price competition and reimbursement costs. These losses help explain why manufacturers do not voluntarily disclose their quality types, as predicted by the "unraveling" result of Grossman (1981) and Milgrom (1981). Intermediary surplus approximately doubles, as they benefit from lower prices and better information. In total, producer surplus declines by 14%, suggesting that recall announcements benefit patients but have ambiguous welfare effects.

To decompose these effects, I next simulate the model while updating one mechanism—intermediary preferences, manufacturer costs, and prices—at a time. For each outcome y, I implement the following decomposition:

$$\begin{aligned}\Delta y = &\underbrace{y[R_O(\theta_i, q_i'), C(\theta_i, q_i), b(\theta_i, q)] - y[R_O(\theta_i, q_i), C(\theta_i, q_i), b(\theta_i, q)]}_{\Delta \text{ due to intermediary preferences}} + \\ &\underbrace{y[R_O(\theta_i, q_i'), C(\theta_i', q_i'), b(\theta_i, q)] - y[R_O(\theta_i, q_i'), C(\theta_i, q_i), b(\theta_i, q)]}_{\Delta \text{ due to manufacturer costs}} + \\ &\underbrace{y[R_O(\theta_i', q_i'), C(\theta_i', q_i'), b(\theta_i', q')] - y[R_O(\theta_i, q_i'), C(\theta_i', q_i'), b(\theta_i, q)]}_{\Delta \text{ due to prices}}.\end{aligned} \tag{9}$$

I first update intermediary preferences (the q_i in the first RHS term), holding the other mechanisms fixed. The results suggest that intermediary responses explain more than 100% of policy's quality benefits, raising the share of high-quality drugs by 18% (col. (4)). By

[34] To make the pre- and post-recall periods comparable, I use markets with fixed manufacturer participation.

TABLE III

DECOMPOSITION OF RECALL EFFECTS

		(1)	(2)	(3)	(4)	(5)	(6)
		Total equilibrium effects			Decomposition: %Δ due to		
	Segment	Pre-recall	Post-recall	%Δ	Buyer preferences	Bidder costs	Price responses
Bid price	All	2.01	1.68	-16%	–	–	-16%
	Recalled	1.17	1.01	-14%	–	–	-14%
	Competitors	2.31	1.87	-19%	–	–	-19%
Unit costs		0.17	0.18	6%	24%	-6%	-12%
Welfare-relevant effects							
High-quality share		0.57	0.60	6%	18%	–	-11%
Winning price		0.37	0.33	-11%	22%	–	-32%
Manufacturer profits	All	184	132	-28%	7%	6%	-42%
	Recalled	23	20*	-12%*	-20%	47%	8%
	Competitors	162	101	-37%	11%	0%	-49%
Intermediary surplus		23	46	100%	5%	0%	95%
Total producer surplus		208	178	-14%	7%	5%	-26%

Note: The table presents the simulated effects and mechanisms of recall disclosures. Columns (1) and (2) present pre- and post-recall outcomes, and col. (3) the percentage changes. Cols. (4)–(6) decompose the relative contributions of three mechanisms: changes in intermediary preferences (4), manufacturer variable costs (5), and manufacturers' strategic pricing responses (6). All monetary values are in 2022 dollars and weighted by realized volumes; profits and surplus are aggregated across markets and expressed in millions of dollars per year. *Post-recall manufacturer profits deduct one-time recall reimbursement costs, which are excluded from the last three columns.

selecting more expensive suppliers, intermediaries also modestly increase costs (by 24%) and prices (by 22%).

Manufacturer responses partially offset these effects, reducing equilibrium quality, costs, and prices. Updating the model's cost function (the second RHS term) while holding prices fixed, I find that average costs fall by 6% after disclosure (col. (5)), driven by recalled manufacturer responses.[35] Finally, I update manufacturer *prices* (the third RHS term). I find that recalled firms lower their prices by 14% (col. (6)), fully passing through cost savings and strategically compensating for their lower-quality type. This, in turn, intensifies price

[35] I estimate similar changes in costs using a sample of eventually recalled drugs, implying that changes occur after recalls and do not stem from pre-existing differences in recalled and never recalled drugs (Table XIII row 4).

competition, reducing high-quality prices by 19%. These responses transfer surplus from manufacturers (whose profits fall by 42%) to intermediaries (whose surplus increases by 95%), and lower total producer surplus (by 26%). Intermediaries respond to lower prices by shifting purchases toward low-quality drugs, lowering the equilibrium share of high-quality drugs by 11% and undoing much of the policy's benefit.

In summary, disclosure benefits patients and intermediaries but substantially penalizes manufacturers. The observed quality gains stem wholly from intermediary, rather than supply-side, responses, and I find no evidence that disclosure increases manufacturer effort to produce quality.[36] Taken together, these results highlight how intermediaries reshape the equilibrium effects of disclosure in a setting where patients cannot observe quality.

5.2. *The Role of Informed Intermediaries*

Next, I formally quantify the role of informed intermediaries, by asking how equilibrium outcomes would change if they did not care for (or equivalently, never observed) quality. In theory, informed intermediaries can raise equilibrium quality relative to the no-information benchmark. However, they may also increase manufacturer differentiation and prices, or decrease allocative efficiency by shifting demand to high-cost manufacturers. To quantify these effects, I simulate a counterfactual environment in which intermediary preferences for quality are reduced to zero ($\Delta S' = 0$). Table IV column (1) presents the results relative to the status quo in column (2). Intermediaries substantially increase the share of high-quality drugs (by 14pp or 27%), with minimal price effects. High-quality prices increase by \$0.01 (3%, or 0% at the median), consistent with increased differentiation, while low-quality prices remain unchanged. Intermediaries reward high-quality manufacturers (whose profits increase by 28%) and penalize low-quality ones (whose profits decrease by 37%, but fail to offset gains for high-types). With better information, intermediaries capture slightly more surplus, and total producer surplus increases by 1%.

Overall, these results suggest that informed intermediaries improve static welfare. By forcing manufacturers to compete on quality, and not just price, intermediaries may also encourage manufacturer investment in quality and the entry of high-quality drugs in the long-run. An alternative interpretation of this counterfactual is that it compares the per-

[36] My estimates capture variable costs and do not account for potential changes in manufacturer investment.

TABLE IV

COUNTERFACTUAL POLICIES: PRICE-ONLY AND PAY-FOR-PERFORMANCE SCORES

		(1)	(2)	(3)	
	Manufacturer type	Price-only scoring	Baseline scoring	Subsidized scoring	
Quality premium		0%	13%	26%	
Average bid price	All	1.66	1.66	2.19	
	Low	1.12	1.12	1.25	
	High	1.85	1.85	2.53	
Average winning price	All	0.34	0.35	0.45	
	Low	0.35	0.35	0.42	
	High	0.34	0.35	0.46	
Welfare-relevant effects					
High-quality share		0.51	0.64	0.80	
Average winning price		0.34	0.35	0.45	
Manufacturer profits	All	672	675	846	
	Low	227	143	75	
	High	315	402	642	
				20% penalties	*100% incentives*
Program cost				166	207
Intermediary surplus		125	126	239	274
Total producer surplus, less program cost		797	802	920	913

Note: The table presents equilibrium outcomes from counterfactual simulations that vary intermediary preferences for quality. Column (1) presents a "price-only" environment ($\Delta S = 0\%$ of average prices), column (2) presents the status quo ($\Delta S = 13\%$), and column (3) presents preferences under quality subsidies ($\Delta S = 26\%$). All monetary values are in 2022 dollars and weighted by realized volumes; profits and surplus are aggregated across markets and expressed in millions of dollars per year.

formance of price-only and scoring auctions. Viewed through this lens, it reproduces a key prediction from the theoretical scoring auction literature: buyer surplus is greater in scoring auctions than in price-only ones (Asker and Cantillon, 2008).[37] Intuitively, this is because price-only auctions limit the space over which buyers and bidders can maximize, thereby reducing the potential surplus.

[37] Surplus in scoring auctions also exceeds that from a closely related format, price-only auctions with minimum quality standards (discussed in subsection 5.3). This holds true even if bidders are allowed to costlessly "convert" to high-quality types, as the surplus gains from higher quality do not offset losses from higher prices.

5.3. *Policy Proposals*

In my final counterfactuals, I examine the welfare implications of recent federal proposals intended to improve drug quality and supply reliability in the U.S. These proposals fall into two broad classes. The first class seeks to increase demand for high-quality drugs and thus incentivize quality investment throughout the supply chain (HHS, 2024). For example, the Manufacturer Resiliency Assessment Program would develop new manufacturing measures to increase intermediary and consumer information about quality. Another proposal, the Hospital Resilient Supply Program, would subsidize purchases of high-quality drugs (or penalize low-quality purchases), providing an instrument to align intermediary preferences with social objectives. The second class of policies would instead alter the supply of drugs, by raising minimum quality standards or reshoring production of essential medicines to the U.S. or Europe (GAO, 2024, The White House, 2025).[38]

The quasi-experimental evidence suggests that the first policy, targeting transparency, may be redundant in this market, where intermediaries already closely monitor quality. Put another way, the "market information failures" that the policy targets are less severe when intermediaries are the primary decision-makers. The second policy, targeting intermediary preferences, could improve welfare if intermediaries undervalue quality relative to the social planner. To assess this possibility, I simulate counterfactual equilibria in which intermediary preferences are perturbed by various combinations of subsidies and penalties. I calculate the policy's social costs assuming a public cost of funds of 20% and consider two designs: 100% subsidies and a mix of 80% subsidies with 20% penalties. Table IV column (3) reports the results for one version, which subsidizes high-quality drug purchases at 13% of the average price, effectively doubling intermediary preferences for quality. Relative to the status quo, the policy increases the share of high-quality drugs by 16pp (24%), a magnitude comparable to intermediary effect documented in cols. (1)–(2). Under subsidies, however, equilibrium prices rise by substantially more (29% versus 3%), as increased

[38]The HHS proposal focuses on physician-administered drugs, drawing on existing Centers for Medicare and Medicaid Services authorities. It states that the policies may be extended to the broader generics market, where supply issues are also pervasive and my analysis is directed. While there may be many motivations to reshore manufacturing, I focus exclusively on the potential effects on short-run prices, costs, and quality.

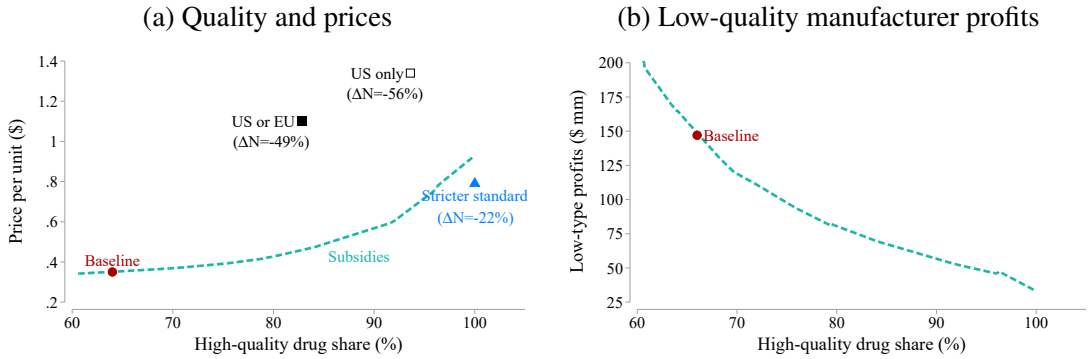


FIGURE 4.—Quality, prices, and profits under counterfactual policies

Note: The figures show equilibrium outcomes from counterfactual simulations of a range of quality policies. Panel (a) plots average equilibrium prices against the share of high-quality drugs, simulated under various levels of quality subsidies, increased minimum quality standareds, and reshored manufacturing. Panel (b) plots average profits for low-quality manufacturers against quality under subsidies. All monetary values are in 2022 dollars and weighted by realized volumes.

quality differentiation weakens price competition. On the supply-side, producer surplus increases by 14-15%, depending on the mix of subsidies and penalties used.

Figure 4 presents equilibrium outcomes under alternative subsidy levels and compares them against proposals to raise quality standards and reshore manufacturing. I begin by simulating intermediary preferences under various subsidies between 0% and 80% of average prices, plotting the equilibrium share of high-quality drugs (on the $x-$axis) against prices (on the $y-$axis). The figure shows that quality gains are increasingly costly. Prices increase gradually around the status quo, but reaching an 80% high-quality supply through subsidies would require price increases of over 30%. On the supply side, subsidies also impose substantial costs on low-quality manufacturers. Panel (b) plots their profits on the y-axis, which decrease steadily in the quality share (and reach Y% at an 80% high-quality supply). While this analysis is static, the magnitude of these losses suggests that low-quality manufacturers are likely to exit the market at high levels of subsidies. This has the potential to increase supplier churn and reduce aggregate manufacturing capacity, and thus increase the risk of shortage, a primary target of the policy.

The third policy, which would raise minimum quality standards, nearly doubles the share of high-quality drugs but sharply reduces competition, reducing the average number of bidders by 22% (Figure 4, blue triangle). As a result, prices more than double (increasing by 125%, or 32pp less than comparable quality improvements achieved through

subsidies). The last policy, reshoring manufacturing, produces the least efficient quality gains. Reshoring requires nearly 2.8 times higher prices to achieve quality improvements of only 30pp (94% high-quality, white square).[39] The structural model helps decompose why this intervention is relatively expensive: domestic production entails 9% higher estimated marginal costs and reduces bidder participation by about one-half. Furthermore, the policy is poorly targeted, since baseline U.S. quality only slightly exceeds that of foreign drugs, on average (by less than 1pp). Consequently, equilibrium quality gains are only possible with large price increases.

A final question concerns whether these changes in price and quality are likely to improve welfare. Estimating consumers' willingness-to-pay for quality is challenging, as they do not observe variation in quality (or even price, if insured), and no systematic evidence exists on the value of manufacturing quality. Instead, I draw on laboratory evidence from recent recalls to approximate the marginal value of higher quality. For example, the FDA estimates that 145 recently recalled drugs increased the incidence of cancer by roughly 1 in 8000, twelve times the agency's benchmark for acceptable risk (FDA, 2020). At standard values of QALYs, this implies that higher-quality drug are worth at least 20% more.[40] This value is likely conservative, as it reflects estimates from less severe (Class II) recalls, and patients may be willing to pay substantially more to avoid cancer (LaPenta et al., 2024). Nevertheless, it suggests that only policies with moderate price effects, such as small subsidies, are likely to improve consumer welfare at even generous values of quality.

6. DISCUSSION AND CONCLUSION

This paper studies how informed intermediaries and regulatory disclosure discipline quality in the U.S. generics market—a setting where quality has high stakes for health but is not observed by patients. Using comprehensive data on manufacturing quality over 2000-2022, I show that quality failures are widespread, and that disclosing these failures meaningfully improves the quality of drugs consumed by patients for at least a decade. By modeling generic procurement through scoring auctions, I show that intermediaries prefer

[39] A more conservative policy, reshoring to the U.S. *and* Europe, produces qualitatively similar results (plotted at the black square).

[40] This estimate reflects the expected increase in cancer incidence from four years of daily use, 10 QALYs lost per cancer diagnosis, and $150,000 per QALY.

high-quality, reliable drugs, which in turn encourages manufacturers to compete on quality. In equilibrium, intermediaries increase the share of high-quality drugs by over a quarter and amplify the effects of FDA disclosure, generating substantial welfare gains for patients.

These findings challenge the conventional wisdom that generics are perfectly substitutable and that the market is "unable to observe and reward quality" (FTC, 2011, FDA Drug Shortages Task Force, 2019). Instead, they reveal large and "potentially injurious" variation in generic quality, which is primarily disciplined by informed intermediaries. The results also suggest that quality policy design should explicitly account for intermediary responses. In particular, since intermediaries already closely monitor quality signals, additional disclosures may yield limited welfare gains. Modest quality subsidies could reduce intermediary distortions and improve quality, with greater welfare improvements than policies that raise quality standards or reshore manufacturing.

While this paper captures key mechanisms through which disclosure and intermediaries shape equilibrium outcomes, it abstracts away from manufacturer participation, capacity adjustment, and quality investment. Incorporating these long-run forces would inform the design of quality policy, efficient auctions, and the scope of intermediary antitrust enforcement (FTC, 2024). Another natural extension would be to link quality failures and disclosure directly to patient health outcomes, providing a fuller account of the welfare consequences of drug quality.

REFERENCES

AGUIRREGABIRIA, VICTOR AND PEDRO MIRA (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75 (1), 1–53. [27]

AKERLOF, GEORGE A (1970): "The market for lemons: quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, 84 (3), 488–500. [1, 6]

ALBANO, GIAN LUIGI, FEDERICO DINI, AND ROBERTO ZAMPINO (2009): "Bidding for complex projects: Evidence from Italian government's acquisitions of IT services," in *Electronic Government: 8th International Conference, EGOV 2009, Linz, Austria, August 31-September 3, 2009. Proceedings 8*, Springer, 353–363. [31]

ALLEN, JASON, ROBERT CLARK, BRENT HICKMAN, AND ERIC RICHERT (2024): "Resolving Failed Banks: Uncertainty, Multiple Bidding and Auction Design," *Review of Economic Studies*, 91 (3), 1201–1242. [6]

AMERISOURCEBERGEN (2019): "Manufacturer Packaging and Logistics Requirements Guide," Accessed October 29, 2024. [4, 9, 20]

ANDRADE, SUSAN E, KRISTIJAN H KAHLER, FERIDE FRECH, AND K ARNOLD CHAN (2006): "Methods for evaluation of medication adherence and persistence using automated databases," *Pharmacoepidemiology and Drug Safety*, 15 (8), 565–574. [18]

ASKER, JOHN AND ESTELLE CANTILLON (2008): "Properties of scoring auctions," *The RAND Journal of Economics*, 39 (1), 69–85. [38]

ATAL, JUAN PABLO, JOSE IGNACIO CUESTA, AND MORTEN SAETHRE (2022): Tech. rep., National Bureau of Economic Research. [6]

ATHEY, SUSAN AND PHILIP A HAILE (2002): "Identification of standard auction models," *Econometrica*, 70 (6), 2107–2140. [4, 24]

BARAHONA, NANO, CRISTÓBAL OTERO, AND SEBASTIÁN OTERO (2023): "Equilibrium effects of food labeling policies," *Econometrica*, 91 (3), 839–868. [6]

BEN-SHAHAR, OMRI AND CARL E SCHNEIDER (2017): "The failure of mandated disclosure," *Russian Journal of Economics and Law*, (4 (44)), 146–169. [1, 16]

BORUSYAK, KIRILL, XAVIER JARAVEL, AND JANN SPIESS (2024): "Revisiting event-study designs: robust and efficient estimation," *Review of Economic Studies*, 91 (6), 3253–3285. [16, 59]

BRENNAN, ZACHARY (2016): "FDA Allows Banned Chinese Firm to Ship Chemotherapy to US as Shortage Threat Looms," Accessed October 29, 2024. [14]

CALLAWAY, BRANTLY AND PEDRO HC SANT'ANNA (2021): "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 225 (2), 200–230. [16, 59]

CARDINALHEALTH (2022): "Manufacturer Reference Manual," Accessed October 29, 2024. [4, 9, 20]

COHEN, ROBIN A AND LARYSSA MYKYTA (2023): "QuickStats: Percentage of Adults Aged $\geq$ 18 Years Who Took Prescription Medication During the Past 12 Months, by Sex and Age Group - National Health Interview Survey, United States, 2021," *MMWR Morb Mortal Wkly Rep*. [2, 7]

CONTI, RENA M, ERNST R BERNDT, NERIMAN BESTE KAYGISIZ, AND YASHNA SHIVDASANI (2020): "We still don't know who makes this drug," *Health Affairs Forefront*. [3, 10]

CUDDY, EMILY (2020): "Competition and collusion in the generic drug market," *Mimeo*. [6, 8, 10, 29]

DAVE, CHINTAN V, AJINKYA PAWAR, ERIN R FOX, GREGORY BRILL, AND AARON S KESSELHEIM (2018): "Predictors of drug shortages and association with generic drug prices: a retrospective cohort study," *Value in Health*, 21 (11), 1286–1290. [18]

DRANOVE, DAVID AND GINGER ZHE JIN (2010): "Quality disclosure and certification: Theory and practice," *Journal of Economic Literature*, 48 (4), 935–963. [6]

DUBOIS, PIERRE, GOSIA MAJEWSKA, AND VALENTINA REIG (2023): "Drug shortages: empirical evidence from France," *TSE Working Paper*. [6]

FDA (2020): "Recalls of Angiotensin II Receptor Blockers (ARBs) including Valsartan, Losartan and Irbesartan," Accessed October 29, 2024. [41]

——— (2023): "Import Alert 66-40: Intas Pharmaceuticals Limited," Accessed October 29, 2024. [7]

——— (2024a): "Pharmaceutical Inspections and Compliance," Accessed October 29, 2024. [19]

——— (2024b): "Recalls Background and Definitions," Accessed October 29, 2024. [3, 8]

FEIN, ADAM J (2023): "The 2023-2024 Economic Report on Pharmaceutical Wholesalers and Specialty Distributors," Drug Channels Institute. [2, 8]

FINK, JOSEPH L. (2012): "Legal Duty to Notify Patient of Product Recall?" *Pharmacy Times*, 78 (3). [31]

FREEDMAN, SETH, MELISSA KEARNEY, AND MARA LEDERMAN (2012): "Product recalls, imperfect information, and spillover effects: Lessons from the consumer response to the 2007 toy recalls," *Review of Economics and Statistics*, 94 (2), 499–516. [6]

FTC (2011): "Authorized generic drugs: short-term effects and long-term impact," Retrieved October 29, 2024. [2, 9, 42]

——— (2024): "FTC, HHS Seek Public Comment on Generic Drug Shortages and Competition Amongst Powerful Middlemen," . [42]

GALDIN, ANAÏS SOLEDAD (2024): "Trade and Information Provision Across Geographical Space," Ph.D. thesis, Princeton University. [6]

GANAPATI, SHARAT AND REBECCA MCKIBBIN (2023): "Markups and fixed costs in generic and off-patent pharmaceutical markets," *Review of Economics and Statistics*, 105 (6), 1606–1614. [6]

GAO (2011): "Medical Devices: FDA Should Enhance Its Oversight of Recalls," Accessed October 29, 2024. [8, 16]

——— (2024): "FDA Has Faced Persistent Challenges Overseeing Foreign Drug Manufacturing," Accessed October 29, 2024. [5, 39]

GAVAZZA, ALESSANDRO AND ALESSANDRO LIZZERI (2021): "Frictions in product markets," in *Handbook of Industrial Organization*, Elsevier, vol. 4, 433–484. [6, 9]

GRENNAN, MATTHEW, KYLE MYERS, ASHLEY SWANSON, AND AARON CHATTERJI (2024): "No free lunch? welfare analysis of firms selling through expert intermediaries," *Review of Economic Studies*, rdae090. [6]

GROSSMAN, SANFORD J (1981): "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 24 (3), 461–483. [5, 35]

GUERRE, EMMANUEL, ISABELLE PERRIGNE, AND QUANG VUONG "GPV" (2000): "Optimal nonparametric estimation of first-price auctions," *Econometrica*, 68 (3), 525–574. [4, 7, 23]

HAILE, PHILIP, HAN HONG, AND MATTHEW SHUM (2003): "Nonparametric tests for common values at first-price sealed-bid auctions," *National Bureau of Economic Research Working Paper 10105*. [29]

HHS (2024): "White Paper: Policy Considerations to Prevent Drug Shortages and Mitigate Supply Chain Vulnerabilities in the United States," Accessed October 29, 2024. [5, 39]

HOTZ, V JOSEPH, ROBERT A MILLER, SETH SANDERS, AND JEFFREY SMITH (1994): "A simulation estimator for dynamic models of discrete choice," *The Review of Economic Studies*, 61 (2), 265–289. [27]

HUBBARD, TIMOTHY P AND HARRY J PAARSCH (2014): "On the numerical solution of equilibria in auction models with asymmetries within the private-values paradigm," in *Handbook of Computational Economics*, Elsevier, vol. 3, 37–115. [27]

JÄÄSKELÄINEN, JAN AND JANNE TUKIAINEN (2019): "Anatomy of public procurement," *VATT Institute for Economic Research Working Papers*, 118. [31]

JEWETT, CHRISTINA (2023): "Drug Shortages Near an All-Time High, Leading to Rationing," *The New York Times*, accessed October 29, 2024. [7]

KASTL, JAKUB, ERIC RICHERT, AND JESPER RUDIGER (2025): "Efficient Grid Search to Solve Static Games with Private Information," *Kilts Center at Chicago Booth Marketing Data Center Paper*, available at http://dx.doi.org/10.2139/ssrn.5187301. [27]

KRASNOKUTSKAYA, ELENA, KYUNGCHUL SONG, AND XUN TANG (2020): "The role of quality in internet service markets," *Journal of Political Economy*, 128 (1), 75–117. [6, 21, 31]

LAFFONT, JEAN-JACQUES, ISABELLE PERRIGNE, MICHEL SIMIONI, AND QUANG VUONG (2020): "Econometrics of scoring auctions," in *Essays in Honor of Cheng Hsiao*, Emerald Publishing Limited, vol. 41, 287–322. [4, 21, 22, 23, 26, 27]

LAPENTA, MATTHEW, LAUREN MASATSUGU, L ELIZABETH RIESS, AND EMMA GLIDDEN-LYON (2024): "Estimated Values of Avoiding Cancer Risks by Cancer Site and Population," *Journal of Benefit-Cost Analysis*, 15 (3), 395–439. [41]

LEBRUN, BERNARD (1999): "First price auctions in the asymmetric N bidder case," *International Economic Review*, 40 (1), 125–142. [27]

LI, TONG, ISABELLE PERRIGNE, AND QUANG VUONG (2000): "Conditionally independent private information in OCS wildcat auctions," *Journal of Econometrics*, 98 (1), 129–161. [25]

MA, JUN, VADIM MARMER, AND ARTYOM SHNEYEROV (2019): "Inference for first-price auctions with Guerre, Perrigne, and Vuong's estimator," *Journal of Econometrics*, 211 (2), 507–538. [33]

MARRA, MARLEEN R, NATHAN H MILLER, AND GRETCHEN SILEO (2024): "Phoning Home: The Procurement of Telecommunications for Incarcerated Individuals in the United States," . [6]

MATZKIN, ROSA L (1991): "Semiparametric estimation of monotone and concave utility functions for polychotomous choice models," *Econometrica: Journal of the Econometric Society*, 1315–1327. [27]

AMERICAN CANCER SOCIETY CANCER ACTION NETWORK (2023): "Survivor Views: Drug Shortages, Telehealth, & Biomarker Testing," Accessed October 29, 2024. [7]

CENTER FOR INFECTIOUS DISEASE RESEARCH & POLICY (2018): "Resilient Drug Supply: Critical Acute Drugs," Accessed October 29, 2024. [10]

FDA DRUG SHORTAGES TASK FORCE (2019): "Drug Shortages: Root Causes and Potential Solutions," Accessed October 29, 2024. [31, 42]

STATE OF CONNECTICUT (2019): "Connecticut et al. v. Teva Pharmaceuticals USA, Inc. et al. (Unredacted complaint)," *U.S. District Court for the District of Pennsylvania*, 2:19-cv-02407. [8, 11]

THE WHITE HOUSE (2024): "Biden Cancer Moonshot Announces New Pilot to Mitigate Pediatric Cancer Drug Shortages," Accessed October 29, 2024. [7]

——— (2025): "Executive Order on Regulatory Relief to Promote Domestic Production of Critical Medicines," Accessed July 1, 2025. [5, 39]

MCHORNEY, COLLEEN A AND CHARLES V SPAIN (2011): "Frequency of and reasons for medication non-fulfillment and non-persistence among American adults with chronic disease in 2008," *Health Expectations*, 14 (3), 307–320. [18]

MILGROM, PAUL R (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391. [5, 35]

MILGROM, PAUL R AND ROBERT J WEBER (1982): "A theory of auctions and competitive bidding," *Econometrica: Journal of the Econometric Society*, 1089–1122. [21]

PELTZMAN, SAM (1973): "An evaluation of consumer protection legislation: the 1962 drug amendments," *Journal of Political Economy*, 81 (5), 1049–1091. [6]

RILEY, JOHN G AND WILLIAM F SAMUELSON (1981): "Optimal auctions," *The American Economic Review*, 71 (3), 381–392. [22]

RUST, JOHN (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, 999–1033. [26]

SCOTT MORTON, FIONA (2000): "Barriers to entry, brand advertising, and generic entry in the US pharmaceutical industry," *International Journal of Industrial Organization*, 18 (7), 1085–1104. [6]

SOOD, NEERAJ, TIFFANY SHIH, KAREN VAN NUYS, AND DANA GOLDMAN (2017): "The flow of money through the pharmaceutical distribution system," *Health Affairs Blog*. [34]

SPAGNOLO, GIANCARLO (2012): "Reputation, competition, and entry in procurement," *International Journal of Industrial Organization*, 30 (3), 291–296. [31]

STARC, AMANDA AND THOMAS G WOLLMANN (2022): "Does entry remedy collusion? Evidence from the generic prescription drug cartel," *National Bureau of Economic Research Working Paper 29886*. [6, 30]

TRACELINK (2023): "Pharmacies and Recalls: Understanding the Impact-and Changes to Come," Accessed October 29, 2024. [9]

VATTER, BENJAMIN (2022): "Quality disclosure and regulation: Scoring design in medicare advantage," *Available at SSRN 4250361*. [6]

WATANABE, JONATHAN H, TERRY MCINNIS, AND JAN D HIRSCH (2018): "Cost of prescription drug–related morbidity and mortality," *Annals of Pharmacotherapy*, 52 (9), 829–837. [18]

WOODCOCK, JANET AND MARTA WOSINSKA (2013): "Economic and technological drivers of generic sterile injectable drug shortages," *Clinical Pharmacology & Therapeutics*, 93 (2), 170–176. [7, 8]

YURUKOGLU, ALI, ELI LIEBMAN, AND DAVID B RIDLEY (2017): "The role of government reimbursement in drug shortages," *American Economic Journal: Economic Policy*, 9 (2), 348–382. [6]

APPENDIX: ADDITIONAL FIGURES AND TABLES

Supply Chain Excellence Award scoring

SCE	Maximum points	Score structure					
Manufacturer fill rate	30 pts	97-100% = 30 95-96% = 15 <95% = 0					
Year-end fill rate	10 pts	97-100% = 10 95-96% = 5 <95% = 0					
Item level lead time	10 pts	**Consumer health:** 0-10 days = 10 11-14 days = 5 >15 days = 0	**Brand:** 0-4 days = 10 5-7 days = 5 >8 days = 0	**Private label:** 0-21 days = 10 22-28 days = 5 >29 days = 0	**Generic:** 0-6 days = 10 7-10 days = 5 >11 days = 0	**SPD (oncology):** 0-4 days = 10 5-7 days = 5 >8 days = 0	**SPD (non-oncology):** 0-15 days = 10 16-20 days = 5 >21 days = 0
Sourcing performance	15 pts	**Consumer health:** **Account management:** 15 • Sales = 5 • Inventory efficiency = 5 • DSA and market support = 5 Best Practices and Collaboration: +5 (bonus)		**Brand and generics:** **Account management:** 15 Best Practice and Collaboration: +5 (bonus)		**SPD:** **Account management:** 15 • Responsiveness = 3 • Dispute resolution = 3 • Pipeline engagement = 3 • Service fee value = 3 • Timely WAC change = 1 • Permit off cycle order planning = 1 • Provide compensation for enhances distribution services = 1	
Product Availability Report (PAR)	10 pts	9.5-10 average pts = 10 8.5-9.4 average pts = 9 7.5-8.4 average pts = 8 6.5-7.4 average pts = 7		5.5-6.4 average pts = 6 4.5-5.4 average pts = 5 3.5-4.4 average pts = 4 2.5-3.4 average pts = 3		1.5-2.4 average pts = 2 0.5-1.4 average pts = 1 <0.4 average pts = 0	
• • •							
Supply chain excellence quality	20 pts	**Deductions based on:** Class 3 recall/withdraw: -2 Class 2 recall: -4 Class 1 recall: -6		Incomplete recall notices: -2 Concealed shortage on control substances or list chemicals: -2 Complete, accurate and prompt recall notices: +1 (bonus)			

FIGURE 5.—Intermediary scoring criteria

Note: The figure presents an example of an intermediary's supplier scoring criteria (Cardinal Health, 2020). For brevity, it is shortened at the ellipsis. Approximately 60% of the possible points are related to the manufacturer's historical recalls and supply disruptions.

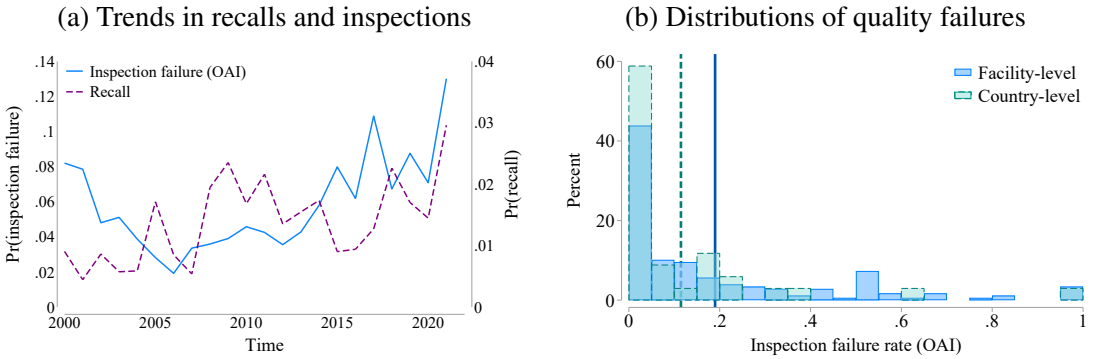


FIGURE 6.—Trends and variation in quality failures

Note: Panel (a) presents annual drug recall rates and facility inspection failure rates, where inspection failures are limited to "Official Action Indicated" (OAI) grades. Panel (b) presents distributions of inspection failure rates averaged over drug-years at the facility and country levels. Distributions are winsorized at the 95th percentile, and vertical lines display mean rates.

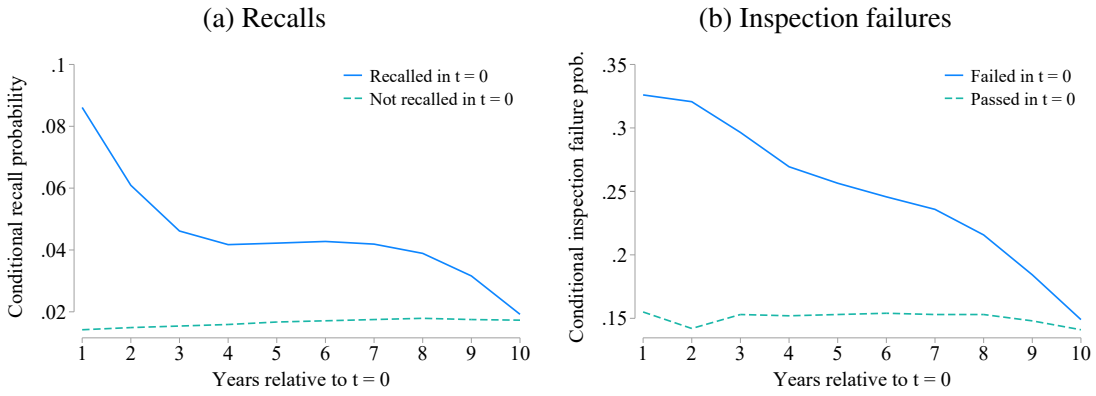


FIGURE 7.—Persistence in drug manufacturing quality

Note: The figure plots persistence in manufacturing quality failures within drugs over time, as observed in the FDA and Optum claims data. It shows the probability of a failing signal in each year for drugs with a previously disclosed failure in year 0 (solid lines), or without such a disclosure (dashed lines). Panel (a) presents the probabilities for recall announcements, and panel (b) the probabilities for inspection failures. Recalls are observed in FDA Enforcement Reports, and inspection grades are from internal FDA data obtained through FOIA requests. Failed inspections are those graded "official" or "voluntary action indicated."

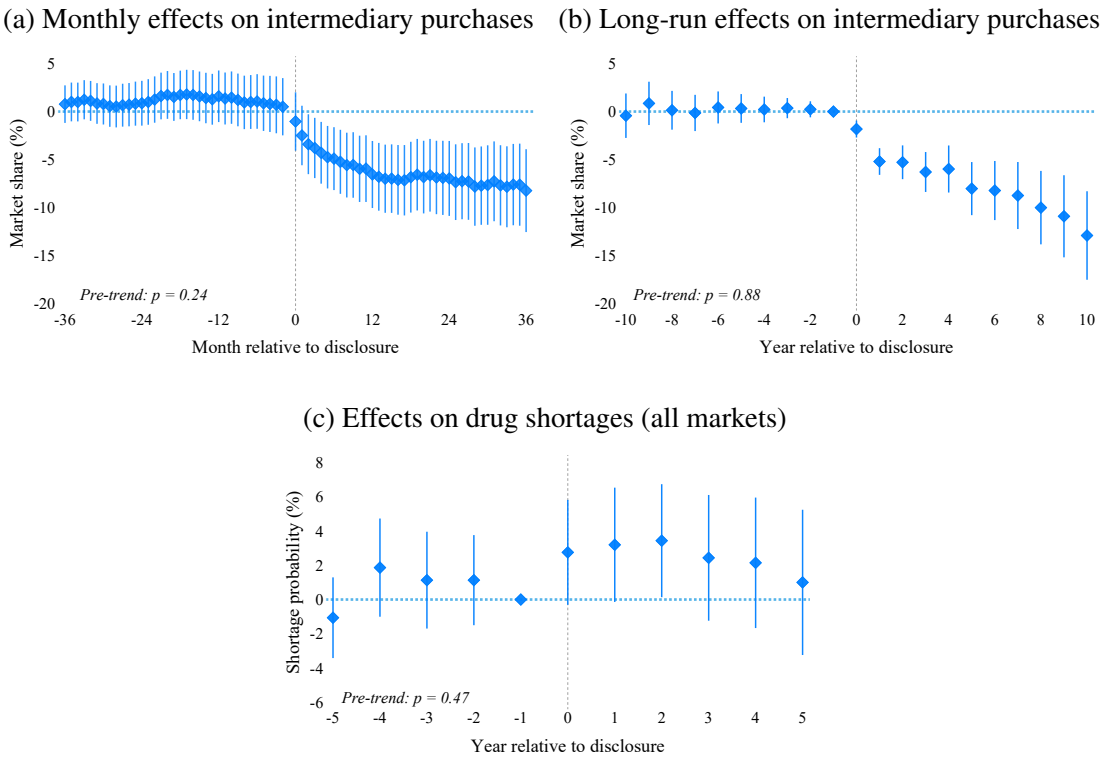


FIGURE 8.—Quasi-experimental effects of recall announcements: Timing and shortages

Note: The figure displays the estimated dynamic effects of recall disclosures on intermediary drug purchases and shortages. The coefficients were obtained by estimating Equation (1) on drug-level data constructed from Optum claims. The year or month prior to the first disclosure event is the omitted reference period. Panels (a)–(b) present the dynamic effects from Figure 3 panel (a) in additional detail. Both present the effects for recalled drugs with the share of pharmacies as the dependent variable. Panel (a) zooms in on the three years surrounding the recall using drug-month level data. Panel (b) presents the long-run effects of recalls, extending the event study to include ten years before and after the recall. Finally, panel (c) shows the estimated effects of recalls on drug shortages, measured using the UUDIS data. The figures present 95% confidence intervals with standard errors clustered by drug or market, and p-values from F-tests of the joint significance of pre-treatment coefficients.

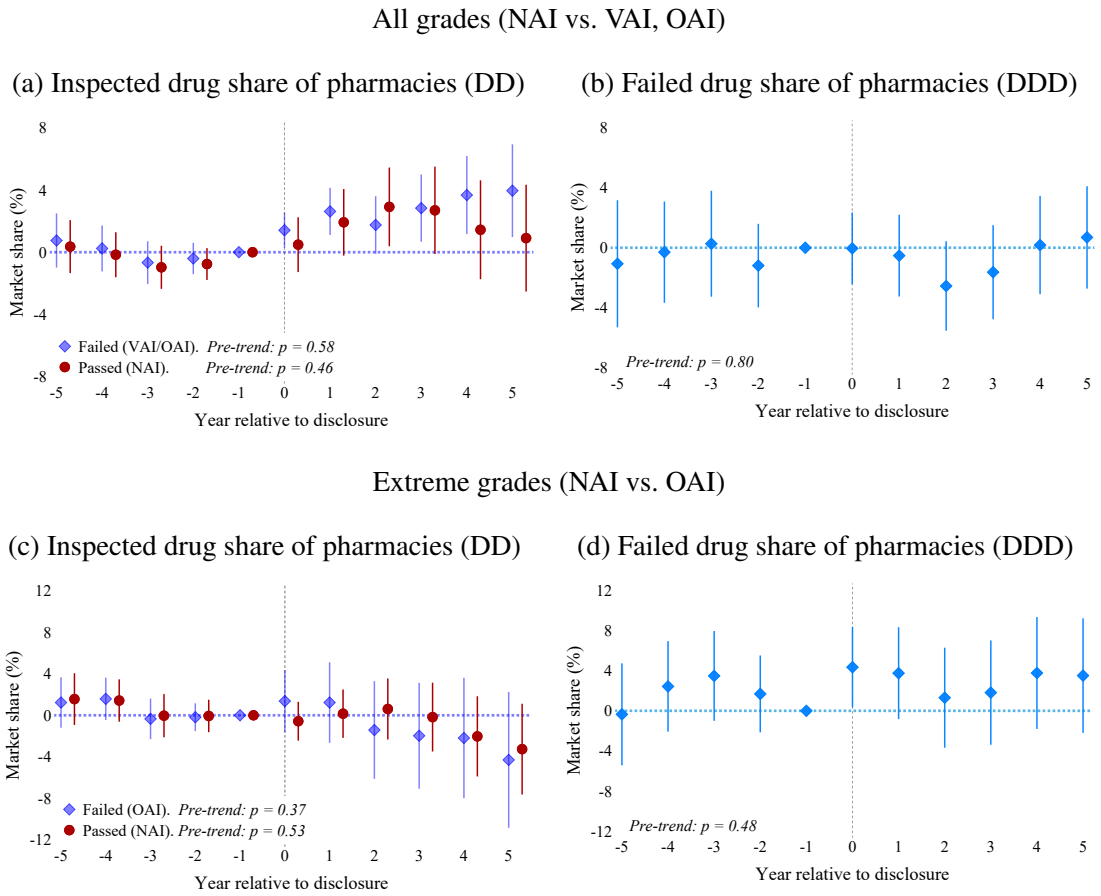


FIGURE 9.—Effects of inspection reports on intermediary purchases

Note: The figure displays the estimated dynamic effects of inspection disclosures on intermediary drug purchases. The coefficients were obtained using drug-year level data constructed from Optum claims. The year prior to the first disclosure event is the omitted reference year. Panels (a) and (c) present the effects of passing (red dots) and failing (purple diamonds) inspection reports, estimated using Equation (1) with the share of pharmacies as the dependent variable. The share of pharmacies is defined as the fraction of U.S. retail pharmacies supplied by a drug manufacturer, as observed in the Optum data. Panels (b) and (c) present the dynamic coefficients from a DDD specification estimated using Equation (1) with an additional term, $(\gamma\, D_{dt} \times 1\{\text{failed}_d\})$. Passing inspections are defined as those graded "No Action Indicated" (NAI) by the FDA. In panels (a)–(b), failing inspections are those graded "Voluntary" (VAI) or "Official Action Indicated" (OAI), while in panels (c)–(d), the definition is restricted to OAI grades only. In each plot, the control group is not-yet-inspected drugs. The figures present 95% confidence intervals with standard errors clustered by drug or market, and p-values from F-tests of the joint significance of pre-treatment coefficients. Table VII presents the corresponding DD coefficients and mean values.

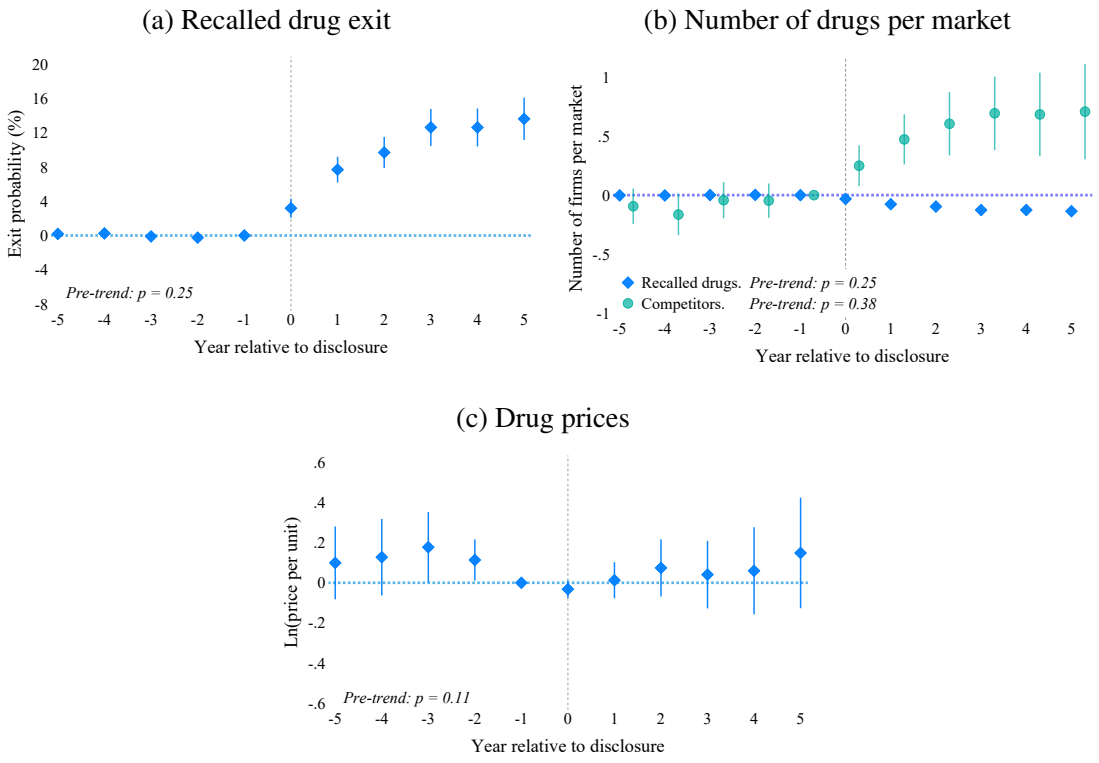


FIGURE 10.—Effects of recalls on market structure and patient welfare

Note: The figure displays the estimated dynamic effects of recalls on market structure and prices. The coefficients were obtained by estimating Equation (1) on drug or market-level data constructed from Optum claims. The year prior to the first disclosure event is the omitted reference year. Panel (a) displays the effects on the recalled drug's probability of exit. The dependent variable is an indicator equal to one in the last year before a drug disappears from the Optum data (for at least two consecutive years). Panel (b) shows the change in the total count of bioequivalent competitors of the recalled drug, estimated using market-level data. The competitor coefficients are plotted in green dots and overlaid on top of the recalled drug ones, which are in blue diamonds. Panel (c) presents the event study results for log average wholesale prices, weighted by baseline volume. The prices are market-level averages of pharmacy acquisition costs per unit from state and national surveys (described in Section 2.2) and inflated to 2022 dollar terms. The figures present 95% confidence intervals with standard errors clustered by drug or market, and p-values from F-tests of the joint significance of pre-treatment coefficients. Table VIII presents the corresponding DD coefficients and mean values.

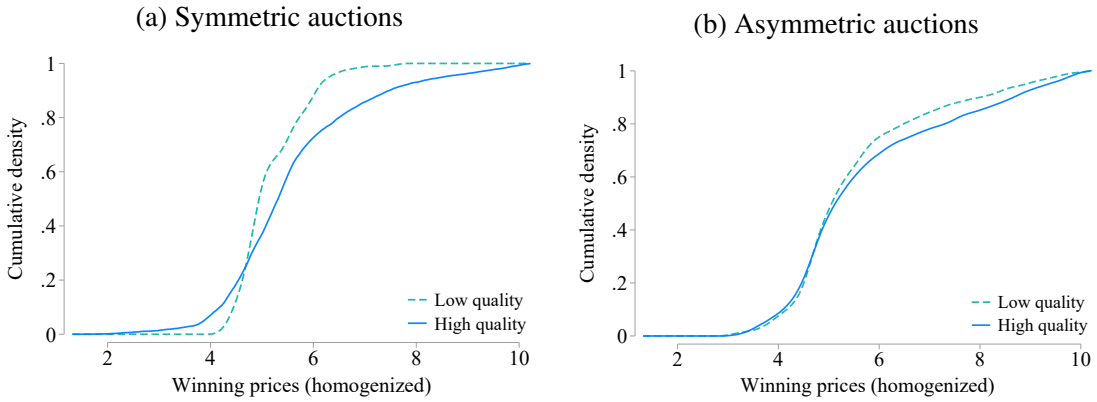


FIGURE 11.—Cumulative densities of winning prices by auction type

Note: The figures plot the kernel densities of homogenized winning prices from procurement auctions of unbranded, oral generic drugs in the Optum data with $N = 2$ bidders. Panel (a) plots the densities for symmetric auctions—with either high- (solid lines) or low-quality (dashed lines) bidders—while panel (b) plots the corresponding ones for asymmetric auctions. Prices represent intermediary acquisition costs per unit and are homogenized using the methodology described in Section 4.3. High-quality drugs have are those that have not experienced a recall within the previous five years, as observed in FDA Enforcement Reports.

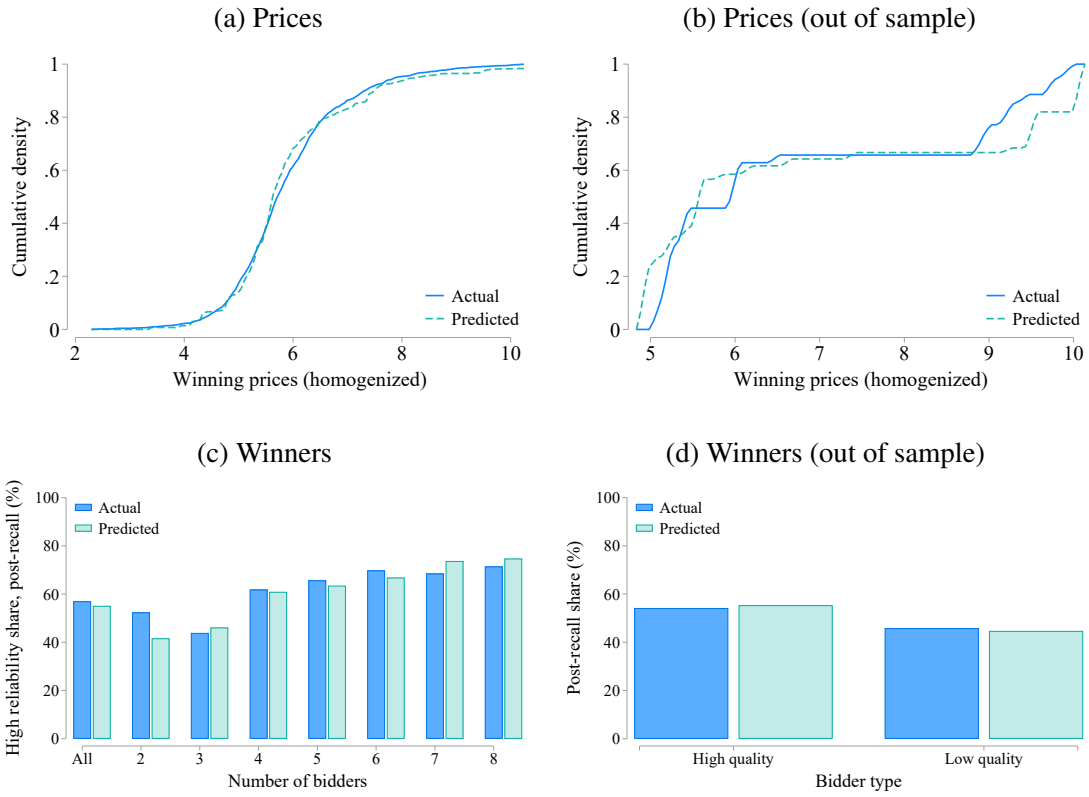


FIGURE 12.—Scoring auction model fit: predicted post-recall outcomes

Note: The figure presents the structural model's predictions of market outcomes after recalls, compared against the observed data. The model predictions use only *ex ante*, pre-recall data to predict post-recall outcomes. Throughout, model predictions are plotted in green, while observed outcomes are in blue. Panel (a) shows the kernel density of winning prices of recalled and competitor drugs after recalls. Panel (b) shows the high-quality drug share across auctions of different sizes. Panels (c) and (d) present the same outcomes but use only out-of-sample data from 2022 (that was not used in estimating the model). Out-of-sample fit was evaluated once. Since the test sample is smaller, panel (d) presents high-quality and recalled shares pooled across all auction sizes. The model fit exercises are described further in Section 4.4.

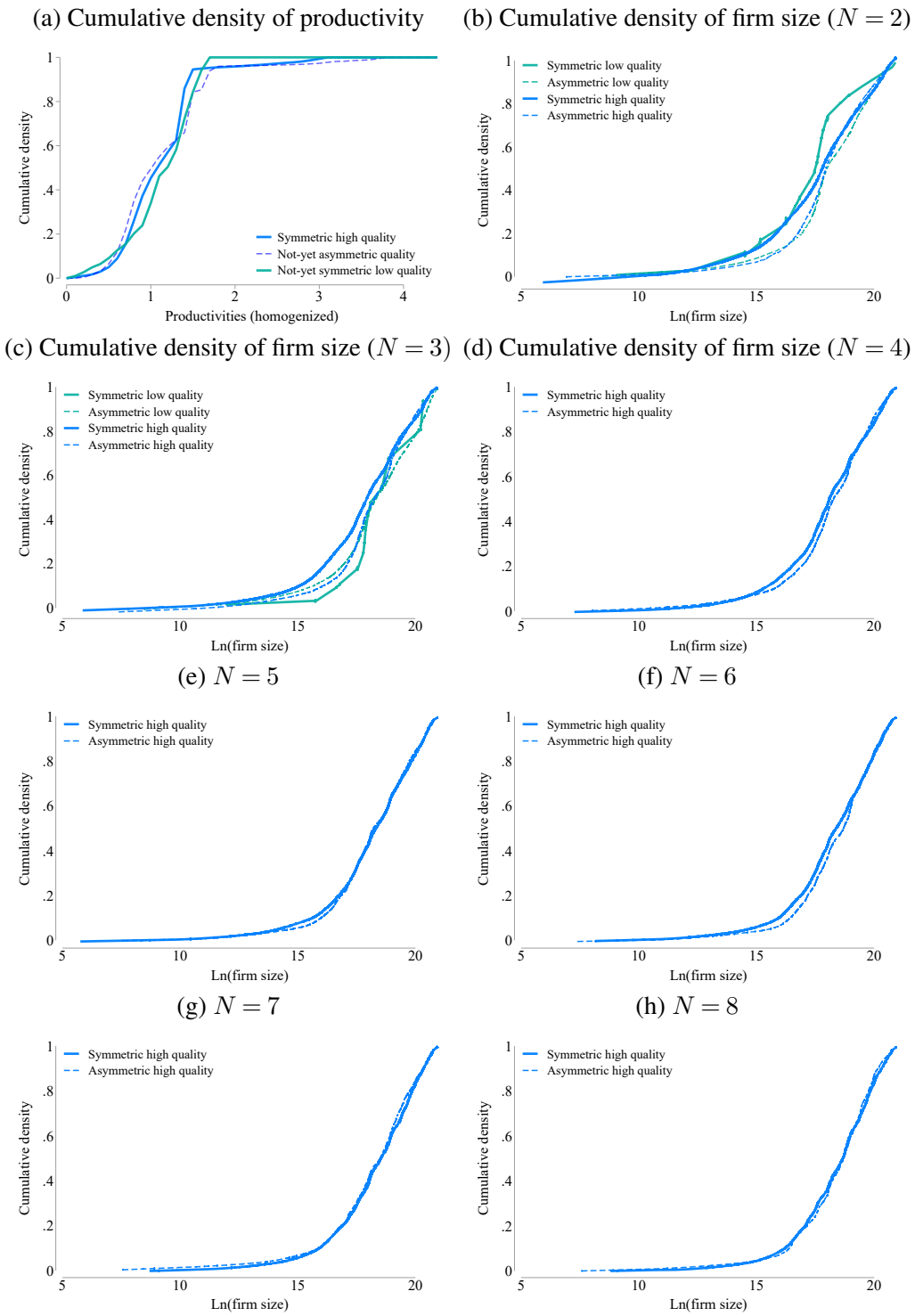


FIGURE 13.—Balance in productivity and firm size across symmetric & asymmetric auctions

Note: The figure plots the kernel densities of the drug-level estimated productivity shocks θ_i and firm size in auctions with symmetric (uniform bidder types, solid lines) and asymmetric (mixed types, dashed lines) auctions. Panel (a) presents the cumulative densities of productivities estimated from a subsample of not-yet-symmetric and not-yet-asymmetric auctions before the first recall, using the model and algorithm described in Section 4. Panels (b)–(h) present densities of firm size from the main estimation sample in increasing order of auction size (N bidders). Quality type is a binary indicator of whether a drug has been recalled in the last five years. Firm size is measured using annual log volumes observed in Optum claims data.

TABLE V

TOP REASONS FOR RECALLS AND INSPECTION CITATIONS

A - Recall reason	*N*
1. Failure to comply with Good Manufacturing Practices	2,353
2. Cross contamination by chemical or drug	1,497
3. Lack of sterility	916
4. Active ingredient: sub or super potent	856
5. Lack of stability	764
6. Drug mixup or mislabeling	720
7. Drug tests out of specification	654
8. Foreign substance in drug	413
9. Improper storage	392
10. Defective packaging	347
11. Counterfeit or marketed without approval	291
12. FDA requested	201
13. Supplier manufacturing failure	190
14. Consumer complaints	186
15. Cross contamination by bacteria	148
B - Inspection citation reason	*N*
1. Failure to investigate out-of-specification results and batch discrepancies (211.192)	2,174
2. Inadequate validation of sterilization process (211.113(b))	1,586
3. Lack of scientifically sound laboratory controls (211.160(b))	1,128
4. No written procedures for process controls to assure drugs conform to standards of identity, strength, quality & purity (211.100(a))	925
5. Equipment is not cleaned at appropriate intervals to prevent contamination (211.67(a))	529
6. Records do not meet retention and audit trail requirements to ensure they are trustworthy (211.68(b))	503
7. Quality control unit lacks authority to review production records to investigate errors (211.22(a))	498
8. Equipment is not of appropriate design for its intended use, cleaning and maintenance (211.63)	461
9. Laboratory records do not include data from all tests necessary to assure compliance with standards (211.194(a))	368
10. Test methods have not established accuracy, sensitivity, specificity and reproducibility (211.165(e))	356
11. Sterile processing areas have deficient monitoring systems (211.42(c)(10)(iv))	332
12. Drug is released without appropriate laboratory testing to determine conformance to standards (211.165(a))	308
13. Deficient control systems to prevent contamination and mix-ups (211.42(c))	308
14. Lack of established time limits for production to assure drug quality (211.111)	258
15. Sterile processing areas have deficient cleaning systems (211.42(c)(10)(v))	232

Note: The table presents the 15 most common reasons for recalls (panel A) and inspection citations (panel B) in the Optum data and the corresponding number of drug-year observations. Recalls and inspections may have multiple reasons and citations, respectively. Recall reasons are manufacturer-reported and observed in FDA Enforcement Reports. Citations and descriptions are inspector-reported and from internal FDA data obtained through FOIA requests. The datasets are described in Section 2.2. To capture citations associated with failing inspections, the citation sample is limited to those with above-median shares of failing versus non-failing inspections. Citation numbers from the Code of Federal Regulations, Title 21 are presented in parentheses.

TABLE VI

DESCRIPTIVE STATISTICS: INSPECTIONS

	(1) All drugs	(2) Failing drugs	(3) Passing drugs
Drug characteristics			
Years since first generic entry	19.41	19.44	19.33
Fraction unit type:			
Tablet	0.53	0.48	0.54
Capsule	0.14	0.14	0.11
Oral liquid (ml)	0.10	0.10	0.12
Topical (gm)	0.15	0.20	0.14
Injection (ml)	0.07	0.07	0.09
Extended release	0.11	0.08	0.09
Volume			
Mean annual prescriptions	28,668	21,336	27,609
	(95,069)	(61,101)	(69,703)
Share of pharmacies	0.160	0.117	0.118
	(0.216)	(0.175)	(0.162)
Number of drugs	1,483	1,016	467

Note: The table presents descriptive statistics for generic drugs in the Optum claims data that are matched to FDA inspections. It presents all inspected drugs in column (1), drugs that receive grades of Voluntary or Official Action Indicated ("failing") in column (2) and drugs classified as No Action Indicated ("passing") in column (3). Each value represents the drug-level mean over all years prior to the first inspection. Corresponding standard deviations are reported in parentheses for the primary volume outcome variables. Years since generic entry is calculated as of 2022. Annual volumes and shares of pharmacies are observed in the claims data. The corresponding descriptive statistics for recalled drugs are presented in Table I.

TABLE VII

EFFECTS OF INSPECTION REPORTS ON INTERMEDIARY PURCHASES

	(1)	(2)	(3)	(4)	(5)	(6)
	Difference-in-difference estimates				Triple-difference estimates	
	Share of pharmacies		Ln(scripts)		Share of	Ln(scripts)
Dependent var.	Passed	Failed	Passed	Failed	pharmacies	
A - VAI & OAI effects						
Inspection * post	0.021**	0.020	0.256**	0.288	0.027	0.502
	(0.010)	(0.007)	(0.119)	(0.083)	(0.011)	(0.122)
Inspection * post * failed					-0.005	-0.097
					(0.012)	(0.133)
Observations	8,051	11,131	8,051	11,131	13,456	13,456
Dependent var. mean	0.168	0.168	1,295	1,295	0.168	1,295
B - OAI effects						
Inspection * post	-0.001	0.005	0.279	-0.273	0.003	0.307
	(0.011)	(0.018)	(0.126)	(0.158)	(0.012)	(0.129)
Inspection * post * failed					0.024	0.001
					(0.020)	(0.187)
Observations	4,900	3,604	4,900	3,604	5,898	5,898
Dependent var. mean	0.177	0.177	7.186	7.186	0.177	7.186

Note: The table reports the estimated average effects of inspection disclosures on intermediary drug purchases. Cols. (1)–(4) present the DD coefficients obtained by estimating Equation (2) on drug-year level data. Cols. (1) and (3) present the estimates for passing inspections, while cols. (2) and (4) present those for failing ones. Cols. (5)–(6) present the DDD coefficients that quantify the differences between the passing and failing estimates, obtained by estimating Equation (2) (with an additional term $\gamma\, D_{dt} \times 1\{\text{failed}_d\}$). The dependent variables are each drug's share of pharmacies (cols. (1), (2) and (5)) and logged prescriptions (cols. (4), (5) and (6)). The share of pharmacies is defined as the fraction of U.S. retail pharmacies supplied by a drug manufacturer, as observed in the Optum data. Throughout, passing inspections are defined as those graded "No Action Indicated" (NAI) by the FDA. In panel A, failed inspections are those graded "Voluntary" (VAI) or "Official Action Indicated" (OAI), while in panel B, the definition is restricted to OAI grades. The control group is not-yet-inspected drugs. Standard errors are clustered at the drug-level and reported in parentheses. The dependent variable mean is computed for treated drugs in the year prior to inspection; logged values are exponentiated for clarity.

TABLE VIII

EFFECTS OF RECALLS ON MARKET STRUCTURE AND PATIENT WELFARE

	(1)	(2)	(3)	(4)	(5)	(6)
	Market structure			Patient welfare		
Dependent var.	Recalled drug Pr(exit)	Competitors log N drugs	Net market log N drugs	Log price	Pr(adherence < 80%)	Pr(adherence < drug med.)
Recalled * post	0.067	0.125	0.094	-0.037	0.021	0.033
	(0.005)	(0.029)	(0.025)	(0.060)	(0.006)	(0.006)
Observations	11,673	6,419	6,530	3,984	20,005	20,005
Dependent var. mean	0.043	4.027	5.743	0.290	0.408	0.492
Recalled * post (%)	155%	13%	10%	-4%	5%	7%

Note: The table reports the estimated average effects of recalls on market structure, prices and patient drug adherence. It presents the coefficients obtained by estimating Equation (2) on drug (col. (1)) or market-level data (cols. (2)–(6)). The first three columns show the estimated effects on market structure, where the dependent variables are an indicator for recalled drug exit (col. (1)), and log counts of the number of bioequivalent competitors (col. (2)) and total market participants (col. (3)). The last three columns present outcomes relevant to patient welfare, where the dependent variables are log prices (col. (4)), weighted using baseline volumes), and the fraction of patients with medication adherence (MPR) less than 80% (col. (5)) or the drug median (col. (6)). Exit and the number of firms are observed in the Optum claims data. Log prices are pharmacy average acquisition costs per unit from state and national surveys (described in Section 2.2) and are inflated to 2022 dollar terms. Standard errors are clustered at the drug (cols. 1–2) or market (cols. 3–5) levels and reported in parentheses. The dependent variable mean is computed for treated drugs in the year prior to the recall; logged means are exponentiated for clarity. The corresponding event studies are presented in Figure 10.

TABLE IX

QUASI-EXPERIMENTAL ROBUSTNESS

	Recalled drug share of pharmacies
1. Baseline estimate	-0.044
	(0.007)
Alternative dependent variable	
2. Share of scripts	-0.065
	(0.008)
3. Share of first scripts	-0.055
	(0.007)
Alternative specifications	
4. Volume weighted	-0.051
	(0.012)
5. Inspection covariates	-0.045
	(0.012)
6. Never-recalled control group	-0.020
	(0.005)
7. Borusyak Jaravel Spiess	-0.052
	(0.009)
8. Callaway Sant'Anna	-0.031
	(0.005)
9. Cluster at firm level	-0.044
	(0.009)
10. Cluster at market-level	-0.044
	(0.007)
Alternative samples	
11. Exclude exiting drugs	-0.025
	(0.008)
12. Exclude growing markets	-0.029
	(0.010)
13. Exclude COVID-19	-0.047
	(0.008)
Observations	12,380
Dependent var. median	0.103

Note: The table reports the estimated average effects of recall disclosures on intermediary drug purchases. It presents the coefficients obtained by estimating Equation (2) with different dependent variables, specifications and samples. The first row reproduces the baseline estimate from Table II, using each drug's share of pharmacies as the dependent variable. Rows 2 and 3 vary the dependent variable, using the total share of prescriptions and first prescriptions, respectively. Rows 4 and 5 include pre-recall volume weights and inspection grade covariates, respectively. The sixth row changes the control group to never (as opposed to not-yet) treated drugs. Rows 7–8 include corrections for staggered treatment timing from Borusyak et al. (2024) and Callaway and Sant'Anna (2021), while rows 9–10 cluster standard errors at more aggregated levels than the treated unit (the drug). Rows 11—13 refine the estimation sample by excluding treated drugs that exit the market within 5 years; treated drugs in markets that experience net entry; and all observations after the COVID-19 pandemic (2019+), respectively. Unless otherwise noted, standard errors are clustered at the drug-level and reported in parentheses. The dependent variable median is computed for treated drugs in the year prior to the recall.

TABLE X

QUASI-EXPERIMENTAL ROBUSTNESS: ALTERNATIVE MARKET DEFINITIONS

	(1)	(2)	(3)	(4)	(5)	(6)
	Recalled drug			Competitor spillovers		
	NADAC market	Molecule formulation	ATC4	NADAC market	Molecule formulation	ATC4
Recalled * post	-0.043	-0.044	-0.005	0.060	0.068	0.052
	(0.006)	(0.007)	(0.003)	(0.011)	(0.014)	(0.013)
Observations	27,536	12,388	23,412	13,514	6,419	3,477
Dependent var. median	0.233	0.103	0.047	0.596	0.738	0.756
Recalled * post (%)	-18%	-43%	-12%	10%	9%	7%

Note: The table reports the estimated average effects of recall disclosures on intermediary drug purchases using different market definitions. It presents the coefficients obtained by estimating Equation (2) with each drug's share of pharmacies as the dependent variable. Cols. (1) and (3) report effects when the market is defined as a molecule-formulation-strength combination, as used by Medicaid in its NADAC survey. Cols. (2) and (4) present effects using the baseline market definition, the molecule-formulation. Cols. (3) and (6) use the larget market definition, the ATC-4 group. Cols. (1)–(2) present the effects for recalled drugs estimated using drug-year level data. Cols. (3)–(6) present the total effects for drugs in the same market that were not recalled ("competitors") using market-level data. Standard errors are clustered at the drug (cols. 1–2) or market (cols. 3–5) level and reported in parentheses. The dependent variable mean is computed for treated drugs in the year prior to the recall.

TABLE XI

SCORING AUCTION MODEL PRICE HOMOGENIZATION

Dependent variable	Price per unit
Volume	-0.166
Volume^2	0.000
Extended release	0.783
Intermediary 2	-0.147
Intermediary 3	-0.154
Intermediary 4	-0.173
Intermediary 5	-0.168
Therapeutic area (ATC-1)	
Blood and blood forming organs	2.471
Sensory organs	1.128
Antiparasitic products, insecticides and repellents	0.798
Genito urinary system and sex hormones	0.245
Nervous system	0.242
Antineoplastic and immunomodulating agents	0.139
Systemic hormonal preparations	0.087
Antiinfectives for systemic use	-0.031
Cardiovascular system	-0.444
Musculo-skeletal system	-0.509
Alimentary tract and metabolism	-0.599
Dermatologicals	-0.886
Respiratory system	-1.683
Other therapeutic areas	-0.389
N. bidders fixed effects	x
Year fixed effects	x
Observations	67,880

Note: The table reports the estimated coefficients from a linear regression of prices on auction characteristics using auction-level data. Annual drug volumes units are presented in millions. Drug volumes and characteristics are from the NDC Directory. Intermediaries are observed using Optum claims and NCPDP data. Apart from volume and volume^2, all regressors are indicator variables.

TABLE XII

DESCRIPTIVE STATISTICS: AUCTIONS

Variable	Mean
Winning price per unit	5.88
	(1.18)
Mean annual prescriptions ('000,000)	9.80
Mean bidders per auction	4.07
Fraction with $\geq$1 recalled participants	0.273
1 participant	0.211
2 participants	0.053
3 participants	0.006
4 participants	0.002
Number of auctions	67,880
Number of observations	276,495

Note: The table presents descriptive statistics for generic drug procurement auctions over 2010-2022, the sample used for estimating the structural model. Winning prices per unit represent intermediary acquisition costs and are homogenized using the methodology described in Section 4.3. Annual prescription volumes for each drug-intermediary combination are observed in the Optum claims data. The number of bidders are observed in the Optum claims and FDA Orange Book. Recalls are from FDA Enforcement Reports.

TABLE XIII

SCORING AUCTION MODEL ROBUSTNESS

	(1) Buyer preferences (ΔS / mean price)	(2) Bidder costs (high vs. low quality)
1. Optimal score (baseline)	0.13	0.31
2. Long-run quality	0.07	0.39
3. Auctions with $N = 2$ bidders	0.15	0.03
4. Recalled & not-yet-recalled drugs	0.37	0.47
5. Balanced panel	0.05	0.24
6. Linear score	0.18	0.77

Note: The table summarizes the estimates of buyer preferences and bidder costs obtained by changing the estimation sample or various assumptions of the structural scoring auction model. Column (1) presents estimates of buyer willingness to pay for quality drugs, expressed as a fraction of the mean price. Column (2) reports the average difference between high- and low-quality drug unit costs, expressed as a fraction of low-quality costs. In the first row, I present the baseline model estimates, assuming that buyers use an optimal scoring rule, and that "high-quality" is defined as not being recalled in the past five years. Row 2 reports estimates using an alternate definition of quality, which includes all drugs recalled within the prior decade, but excludes those that were most recently recalled (within the prior year). Row 3 tests for heterogeneity in buyer preferences by auction size and presents estimates from auctions with only $N = 2$ bidders. Row 4 limits the auction sample to recalled and not-yet-recalled drugs in order to isolate how recalled drug costs change after disclosure. Row 5 reports estimates using a "balanced" panel of auctions that do not experience entry or exit. Row 6 shows the estimates when buyers are assumed to use a linear, rather than an optimal, score.

SUPPLEMENT TO "IN DEFENSE OF THE MIDDLEMAN: QUALITY FAILURES IN THE GENERIC PHARMACEUTICAL MARKET"

CATHERINE ISHITANI
Yale University

1. DATA CONSTRUCTION

1.1. *Quality*

I use four secondary data sources to construct the manufacturing facility dataset. First, I obtain the names and locations of manufacturing facilities that produce drugs for the U.S. market from the FDA by FOIA. Second, I use drug labels scraped from the Drugs@FDA and National Library of Medicine's DailyMed websites. Manufacturers are required to list the manufacturer name and country of origin on imported drug labels under The Tariff Act of 1930 and The Food, Drug, and Cosmetic Act of 1938. In practice, they often include additional information regarding drug manufacturers, packagers and distributors and their locations. The FDA began publishing electronic copies of drug labels on these websites in 2006. Third, I obtain detailed customs data from ImportGenius for all countries that export drugs to the U.S. market from 2006-2022. The data contain the names of shipping and consigning firms, their locations and a detailed product description (e.g., "nifedipine extended-release tablets(procxl), usp 60mg (300ct) metoprolol succinate extended-release tablets, usp, 25mg (500ct) nifedipine extended-release, tablets(procxl), usp 90mg (100ct) metoprolol succinate extended-release tablets, usp, 50mg (1000ct)..."). Finally, I use information reported by Medsafe, the New Zealand Health Authority, which includes the drug names, sponsoring and manufacturing firms and their locations, for all drugs currently marketed in New Zealand. I confirm that drugs produced for the U.S. and New Zealand markets are manufactured in the same facilities using the label and import data.

These datasets were first used to match drugs to facilities by researchers at the University of Minnesota Center for Infectious Disease Research and Policy, who used them to identify the locations for 156 drugs (Center for Infectious Disease Research & Policy, 2018). Similar datasets have been used by Joon Noh (using label data) and Anais Galdin (using trade data for injectable drugs) (Noh et al., 2025, Galdin, 2024). The main advantages of using multiple data sources to identify locations are that they increase coverage and allow me to cross-check my matches.

To scale up the CIDRAP methodology, I rely on several machine learning packages. To convert label images to text, I use Optimal Character Recognition from Tesseract. To fuzzy match drug and firm names, I use string cosine distance measures based on NLP Term Frequency-Inverse Document Frequency and the Universal Sentence Encoder from TensorFlow. Lastly, to match locations, I use the Geocoding API from Google. Using this method, I am able to recover at least one facility location for 72% of generic NDCs in the Optum sample, representing 92% of generic volumes.

1.2. *Prices*

The national (NADAC) and state pharmacy average acquisition cost data are collected by a single firm, Myers and Stauffer, on behalf of CMS and state Medicaid agencies. Since the state price data have not been used in previous research, I describe them briefly here. Myers and

Stauffer conducts monthly, mandatory surveys of random samples of 250-350 independent and chain pharmacies per state, excluding 340B pharmacies. In comparison to the national data, the state data are updated more frequently (every week versus twice a week). They are also based on larger proportional sampling sizes—NADAC surveys 500-600 pharmacies *nationally*, versus 250-350 pharmacies per *state*—and mandatory rather than voluntary surveys.

The surveys capture invoice data, which represents acquisition costs net of rebates paid by manufacturers to plans and off-invoice volume discounts. The former are typically very small (< 5%) for generic drugs, and I am not aware of a research dataset that contains estimates of the latter. Since AAC data do not contain NDCs, I fuzzy match drug names, formulations, and strengths to those in the FDA NDC data.

2. QUALITY REGULATION

In a typical inspection, officials observe production for 2-5 days, interview employees and review the facility's quality management systems and historical records. At the start of a recall, the FDA convenes an expert committee to classify the defective drug's risks. While recalls, in principle, are voluntary, the FDA can file a court order if the firm should refuse a recall request. In practice, this sort of escalation is extremely rare, and fewer than 70 drug court orders have been granted since 2000.

FDA enforcement of recalls is described in two reports from the GAO and OIG, respectively: "Medical Devices: FDA Should Enhance Its Oversight of Recalls" (2011) and "The FDA's Food-Recall Process did not Always Ensure the Safety of the Nation's Food Supply" (2017). These reports identify gaps in the FDA's process to ensure timely and effective recalls of devices and food; no such report is available for drug recalls. With respect to food, the investigators found that the FDA could not always ensure that manufacturers initiated recalls promptly. It did not collect status updates or distribution data from the manufacturers, and it did not always perform audit checks consistent with its procedures. With respect to devices, the authors found that manufacturers often could not locate and remove all the defective products, even in high-risk recalls (manufacturers were successful in removing all products in just 52% of completed recalls). The FDA's procedures for monitoring recalls were "unclear," and officials sometimes reached opposite conclusions regarding whether recalls were effective. For example, it had no criteria for assessing whether firms removed a sufficient number of recalled devices. Inconsistent with its procedures, the FDA frequently did not conduct post-recall inspections to verify that the recall was complete and that sufficient actions had been taken to prevent a recurrence of the underlying quality problems (53% of completed recalls.)

While pharmacies are not legally obligated to inform patients about recalls, certain pharmacies have procedures to do so. For example, CVS Caremark and Walgreens state that they do so, depending on the recall classification. Other pharmacies' written procedures do not include steps to inform patients, e.g., UC San Diego Health System. Even pharmacies with formal outreach procedures may struggle to find the resources to do so: 76% of pharmacies reported being understaffed in 2022 (National Community Pharmacists Association, 2022). Anecdotes from lawsuits and the media suggest that information about recalls often fails to reach consumers (Newkirk and Berfield, 2019).

3. SCORING AUCTION ESTIMATION ALGORITHM

3.1. *Unobserved prices.*

In this subsection, I first derive equations (4.3.1)–(4.3.1), which form a system of non-linear equations in the unobserved price distributions. Next, I describe the estimation algorithm and the conditions under which a solution to this system is guaranteed to exist.

Derivation. I rewrite the distribution of the observed winning prices using Bayes' Rule:

$$Pr(b_{t,i}|i \text{ wins}) = \frac{Pr(i \text{ wins}|b_{t,i})Pr(b_{t,i})}{Pr(i \text{ wins}),} \tag{1}$$

where I suppress conditioning on $q_i = t, q_{-i}$ throughout. Under the linear scoring rule, bidder i wins i.f.f.

$$S(q_i) - b_i \geq S(q_j) - b_j \;\; \forall j \in \{1, ..., N\}, \;\; j \neq i.$$

This is the linear score analogue of Equation (8). Without loss of generality, I focus on the case of $N = 2$ types, $t \in \{H, L\}$. After substituting the definition of the linear score into Equation (1) and rearranging terms, we have

$$\begin{aligned}
&Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) = \\
&Pr(b_{H,i})\big[Pr(S(q_i) - b_i \geq S(q_j) - b_j \;\; \forall j \in \{1, ..., N\}, \;\; j \neq i \mid b_{H,i})\big] = \\
&Pr(b_{H,i})\big[Pr(b_L \geq b_{H,i} - \Delta S \mid b_{H,i})\big]^{N_L}\big[Pr(b_H \geq b_{H,i} \mid b_{H,i})\big]^{N_H - 1} = \\
&g_H(b_{H,i})\big[1 - G_L(b_{H,i} - \Delta S))\big]^{N_L}\big[1 - G_H(b_{H,i})\big]^{N_H - 1}.
\end{aligned}$$

The second line holds because b_i is independently and identically distributed (conditional on q_i, q_{-i}). Note that b_L is a random variable corresponding to the price of another bidder of type L; b_H is similarly defined. The last line substitutes in the distributions of $b_t \sim g_t(\cdot)$. Equation (4.3.1) can be derived in the same way (after switching $t = H, L$.) For convenience, I reproduce Equation (4.3.1) below:

$$Pr(b_{L,i}|i \text{ wins})Pr(i \text{ wins}) = g_L(b_{L,i})\big[1 - G_H(b_{L,i} + \Delta S)\big]^{N_H}\big[1 - G_L(b_{L,i})\big]^{N_L - 1}.$$

Note that similar equations can be easily derived for quality-to-price scores, another common scoring format ($V = S(q_i)/b_i$) and non-linear scores (e.g., $V = S(q_i) - b_i^2$). For example, using a quality-to-price score, the analogues to Equation (4.3.1) and (4.3.1) are

$$\begin{aligned}
Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) &= g_H(b_{H,i})\big[1 - G_L(b_{H,i}r))\big]^{N_L}\big[1 - G_H(b_{H,i})\big]^{N_H - 1} \\
Pr(b_{L,i}|i \text{ wins})Pr(i \text{ wins}) &= g_L(b_{L,i})\big[1 - G_H(b_{L,i}/r))\big]^{N_H}\big[1 - G_L(b_{L,i})\big]^{N_L - 1},
\end{aligned}$$

where the normalized quality preference is $r = \frac{S(q_L)}{S(q_H)}$.

3.2. *Estimation algorithm.*

Together, Equations (4.3.1) and (4.3.1) define a system of two nonlinear functional equations in two unknowns, $g_L(b_L)$ and $g_H(b_H)$, at each point price in the support. To solve this system, I first estimate the LHS objects using the observed data and standard non-parametric kernel density estimators (Li et al., 2000). I then discretize the prices b and solve for $g_L(b_L)$ and $g_H(b_H)$ at each point in the price grid, starting with the lowest price $\underline{b}$. Let ε refer to the difference between two consecutive points in the discretized price grid.

For all $b_H < \underline{b} + \Delta S$, type H wins with certainty, implying that the system reduces to a single equation:

$$Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) = g_H(b_{H,i})\big[1 - G_H(b_{H,i})\big]^{N_H - 1}. \tag{2}$$

At the lowest price $b_H = \underline{b}$, we further have

$$Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) = g_H(\underline{b})\big[1 - g_H(\underline{b})\big]^{N_H - 1},$$

where I have used the fact that the empirical cdf in the leftmost bin is equal to the pdf in that bin. We can solve this equation for $g_H(\underline{b})$, as the root of an N_H^{th} degree polynomial in $g_H(\underline{b})$. We can then solve for subsequent $g_H(b_H)$ using Equation (2) and approximating $G_H(b_H)$ with the empirical cdf, namely the Riemann sum of the prior solutions (for $g_H(b_H)$ up to $b_H < \underline{b} + \Delta S$). Note that the boundary condition for $g_L(\underline{b} - \Delta S)$ is not yet defined. I now explain how to iteratively construct $g_H(b_H)$ and $g_L(b_H - \Delta S)$.

More generally, for all b, we can rewrite Equations (4.3.1) and (4.3.1) as

$$l = g_H(b_{L,i} + \Delta S)\big[k - g_L(b_{L,i})\varepsilon\big]^{N_L}\big[j - g_H(b_{L,i} + \Delta S)\varepsilon\big]^{N_H - 1} \tag{3}$$

$$m = g_L(b_{L,i})\big[j - g_H(b_{L,i} + \Delta S)\varepsilon\big]^{N_H}\big[k - g_L(b_{L,i})\varepsilon\big]^{N_L - 1}, \tag{4}$$

where we observe N_H, N_L and ε, have guessed ΔS, and have already estimated j, k, l and m (as defined below).

$$l = Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) \quad k = 1 - \sum_{x=\underline{b}}^{b_L - \varepsilon} g_L(x)\varepsilon$$

$$m = Pr(b_{L,i}|i \text{ wins})Pr(i \text{ wins}) \quad j = 1 - \sum_{x=\underline{b}}^{b_L + \Delta S - \varepsilon} g_H(x)\varepsilon.$$

This leaves two equations in two unknowns ($g_H(b_H = b_L + \Delta S)$ and $g_L(b_L)$), which can be solved for each price where bidding behavior is defined.[1] Note that objects j, k, l and m are all positive, since j and k are cdfs, and l and m are products of pdfs.

3.3. *Proofs of existence.*

I now show that a solution to the non-linear system defined by Equations (3) and (4) exists at each price, for up to $N = 3$ bidders. For larger auctions, the curse of dimensionality (described in Section 4.3) makes estimating the LHS objects increasingly (exponentially) impractical. To avoid this problem, Laffont et al. (2020) recommend using the estimates from only the smallest auction size, which is typically be $N = 2$. For $N = 2$ auctions, I also describe the conditions under which the solution is unique and bounded in $[0, 1]$.

Equations (3) and (4) can be rewritten, after some algebra, as (1) a polynomial of order N in the first unknown, and (2) an expression that defines the second unknown in terms of the first. From here, proving existence relies on well-known properties of polynomial functions. For example, for $N = 2$ ($N_H = 1, N_L = 1$), we can write

$$0 = k\varepsilon x^2 - (jk - m\varepsilon + l\varepsilon)x + jl \tag{5}$$

$$y = \frac{m}{j - \varepsilon x}. \tag{6}$$

[1] Bidding behavior is not defined for prices $b_L > \overline{b} - \Delta S$ and $b_H > \overline{b} + \Delta S$, where one bidder wins or loses with certainty.

For brevity, I have replaced $x = g_H(b_L + \Delta S)$ and $y = g_L(b_L)$. A real solution exists as long as $x \neq j/\varepsilon$ and the discriminant of the quadratic equation (5) is non-negative:

$$(jk - m\varepsilon + l\varepsilon)^2 - 4k\varepsilon jl \geq 0.$$

This condition is satisfied as the grid of prices becomes increasingly fine (as $\varepsilon \to 0$). Taking the limit of the Quadratic Equation as $\varepsilon \to 0$ implies that the two roots of x approach $\{\frac{l}{k}, \infty\}$, and $y = \frac{m}{j}$. Each root is positive since all of the quantities involved (j, k, l, m) are positive. Furthermore, x has a unique solution in $[0, 1]$ if $l \leq k$, or

$$Pr(b_{H,i}|i \text{ wins})Pr(i \text{ wins}) \leq 1 - G_L(b_L - \varepsilon).$$

Likewise, y is in $[0, 1]$ if $m \leq j$, or

$$Pr(b_{L,i}|i \text{ wins})Pr(i \text{ wins}) \leq 1 - G_H(b_L + \Delta S - \varepsilon).$$

In sum, a solution for $N = 2$ exists as long as the grid of prices is sufficiently fine, and a unique positive solution is bounded in $[0, 1]$ if $l \leq k$ and $m \leq j$.

For $N = 3$ ($N_H = 2, N_L = 1$), we can write

$$0 = k\varepsilon^2 x^3 - 2jk\varepsilon x^2 + (-m\varepsilon + j^2 k + l\varepsilon)x - jl \tag{7}$$

$$y = \frac{m}{(j - \varepsilon x)^2} \tag{8}$$

Similarly, for ($N_H = 1, N_L = 2$),

$$0 = j\varepsilon^2 y^3 - 2jk\varepsilon y^2 + (m\varepsilon + jk^2 - l\varepsilon)y - km \tag{9}$$

$$x = \frac{l}{(k - \varepsilon y)^2} \tag{10}$$

Since Equations (7) and (9) are odd degree polynomials, they each have at least one real solution. Furthermore, for small values of ε, they each have three sign changes, implying that they must have at least one positive root each (by Descartes's Rule of Signs). Equations (20) and (22) have positive solutions as long as $x \neq j/\varepsilon$ and $y \neq k/\varepsilon$.

3.4. *Monte Carlo studies.*

To illustrate my estimation algorithm, described in Section 4.3, I conduct a range of Monte Carlo studies. I first show that my algorithm for linear scores recovers the unobserved price distributions. Next, I compare the performance of my full estimation procedure (which uses only winning prices) to Laffont et al.'s (using complete data).

Each Monte Carlo experiment consists of 5,000 replications of N = 5,000 auctions, which corresponds to a realistic auction sample size.[2] To show that my estimation routine recovers the unobserved price densities for linear scores, I first draw prices from log-normal distributions conditional on bidder type. I then identify the winning bids using the linear scoring rule, remove all non-winning prices, and apply my estimation algorithm to recover the price densities. I

[2]Note, however, that this is an order of magnitude smaller than my data.

repeat this exercise using a range of parameters; for exposition, I show the results from one such experiment in Figure 3.4. This uses the following parameter values:

$$N = 2,\ \ (q_i, q_j) = (1, 0),\ \ b|q = 0 \sim \text{LogN}(\text{shape} = 1.0, \text{loc} = 0.0, \text{scale} = 0.9),$$
$$b|q = 1 \sim \text{LogN}(\text{shape} = 1.0, \text{loc} = 0.5, \text{scale} = 1.0),\ \ \Delta S = 0.7.$$

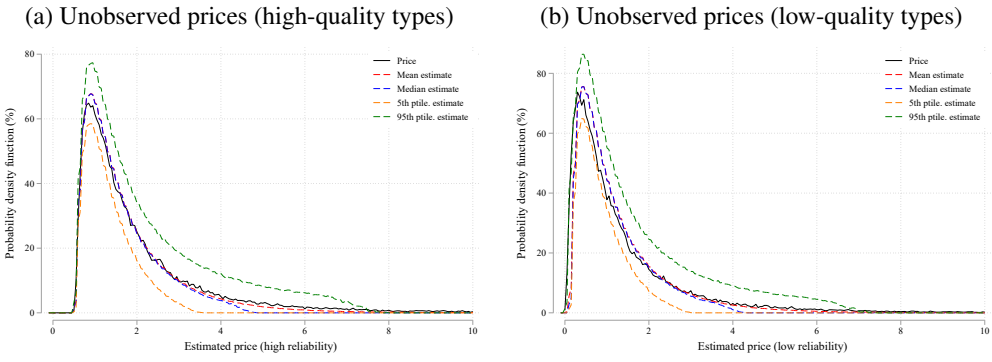


FIGURE 1.—Monte Carlo simulations of prices: Linear score

Figure 3.4 plots the true and estimated price densities for high- and low-quality types (in panels (a) and (b), respectively). The true densities are plotted using solid black lines, while the means, medians and pointwise 90% confidence intervals of the estimates are plotted using dashed lines. The true values are strikingly similar to the mean and median estimates (conditional on a guess of the true value of ΔS), and they appear to fall between the confidence intervals over the whole support.

To show that my full estimation routine recovers the model primitives, and to compare its performance against Laffont et al.'s, I conduct an additional set of experiments. For each replication, I draw productivities from log-normal distributions conditional on bidder type. I then numerically compute the corresponding bids using Equation (5) and apply Laffont et al.'s estimation algorithm. Finally, I leave out all non-winning prices—reducing the sample size by at least half—and apply my algorithm. Again, I repeat the experiment for a range of parameter values and auction sizes, which produces qualitatively similar results. I present the results from one such experiment, with the following parameter values:

$$N = 2,\ \ q \sim \text{Bernoulli}(0.7),\ \ \theta|q = 0 \sim \text{LogN}(0.2, 0.3),\ \ \theta|q = 1 \sim \text{LogN}(0.7, 0.3),\ \ \Delta S = 0.7.$$

The results are summarized in Figure 3.4, with estimates from Laffont et al.'s procedure on the left, and from mine on the right. Panels (a)—(d) report the true and estimated conditional densities of the productivities (in solid black and dashed lines, respectively). The figures show that the estimation routines recover the shape, scale and location of the underlying distributions: the true values are very close to the mean and median estimates, and lie in between the confidence intervals for all but small values of θ. These patterns are similar to those from GPV's Monte Carlo exercise, where even after truncation, extreme values are more difficult to capture due to the boundary bias associated with non-parametric estimators. Comparing estimates from Laffont et al.'s algorithm (on the left) and mine (right) suggests that my estimation procedure recovers the productivity distributions with comparable fidelity as the full-data routine.

This is consistent with the idea that the distribution of the minimum price is a sufficient statistic for the underlying price distribution, so leaving out non-winning prices (in this case, half of the sample) does not result in significant information loss about the underlying productivity distributions. Comparing the densities of high (panels (c)—(d)) and low (panels (e)—(f)) quality types provides additional color on this statement; the low-quality densities are somewhat more noisily estimated, but they represent just 4–8% of the original sample. My algorithm, however, does a poorer job than Laffont et al.'s of fitting the left tail of the distributions. This may be because, in limiting my sample to winning bids, I differentially discard the right tail of the distribution (high prices). Since each pdf must each sum to one, this may cause my results to overestimate sample in the left tails.

Panels (e)—(f) plot the distributions of the estimated ΔS parameters less the true value, with a vertical line at the mean estimate. Both distributions are centered around zero and approximately unbiased; the estimates from my algorithm are less precise, consistent with using a smaller sample size. The corresponding means (and standard deviations) are -0.005 (0.091) for Laffont et al. and 0.040 (0.189) for my algorithm.

REFERENCES

GALDIN, ANAÏS SOLEDAD (2024): "Trade and Information Provision Across Geographical Space," Ph.D. thesis, Princeton University. [1]

LAFFONT, JEAN-JACQUES, ISABELLE PERRIGNE, MICHEL SIMIONI, AND QUANG VUONG (2020): "Econometrics of scoring auctions," in *Essays in Honor of Cheng Hsiao*, Emerald Publishing Limited, vol. 41, 287–322. [4]

LI, TONG, ISABELLE PERRIGNE, AND QUANG VUONG (2000): "Conditionally independent private information in OCS wildcat auctions," *Journal of Econometrics*, 98 (1), 129–161. [3]

CENTER FOR INFECTIOUS DISEASE RESEARCH & POLICY (2018): "Resilient Drug Supply: Critical Acute Drugs," Accessed October 29, 2024. [1]

NATIONAL COMMUNITY PHARMACISTS ASSOCIATION (2022): "Survey of community pharmacy economic pressures 2022 report," Accessed October 29, 2024. [2]

NEWKIRK, MARGARET AND SUSAN BERFIELD (2019): "The FDA Drug Recall System is Voluntary, Haphazard, and Broken," *Bloomberg Businessweek*. [2]

NOH, IN JOON, JOHN GRAY, GEORGE BALL, ZACHARY WRIGHT, AND HYUNWOO PARK (2025): "Are all generic drugs created equal? An empirical analysis of generic drug manufacturing location and serious drug adverse events," *Production and Operations Management*, 10591478251319691. [1]

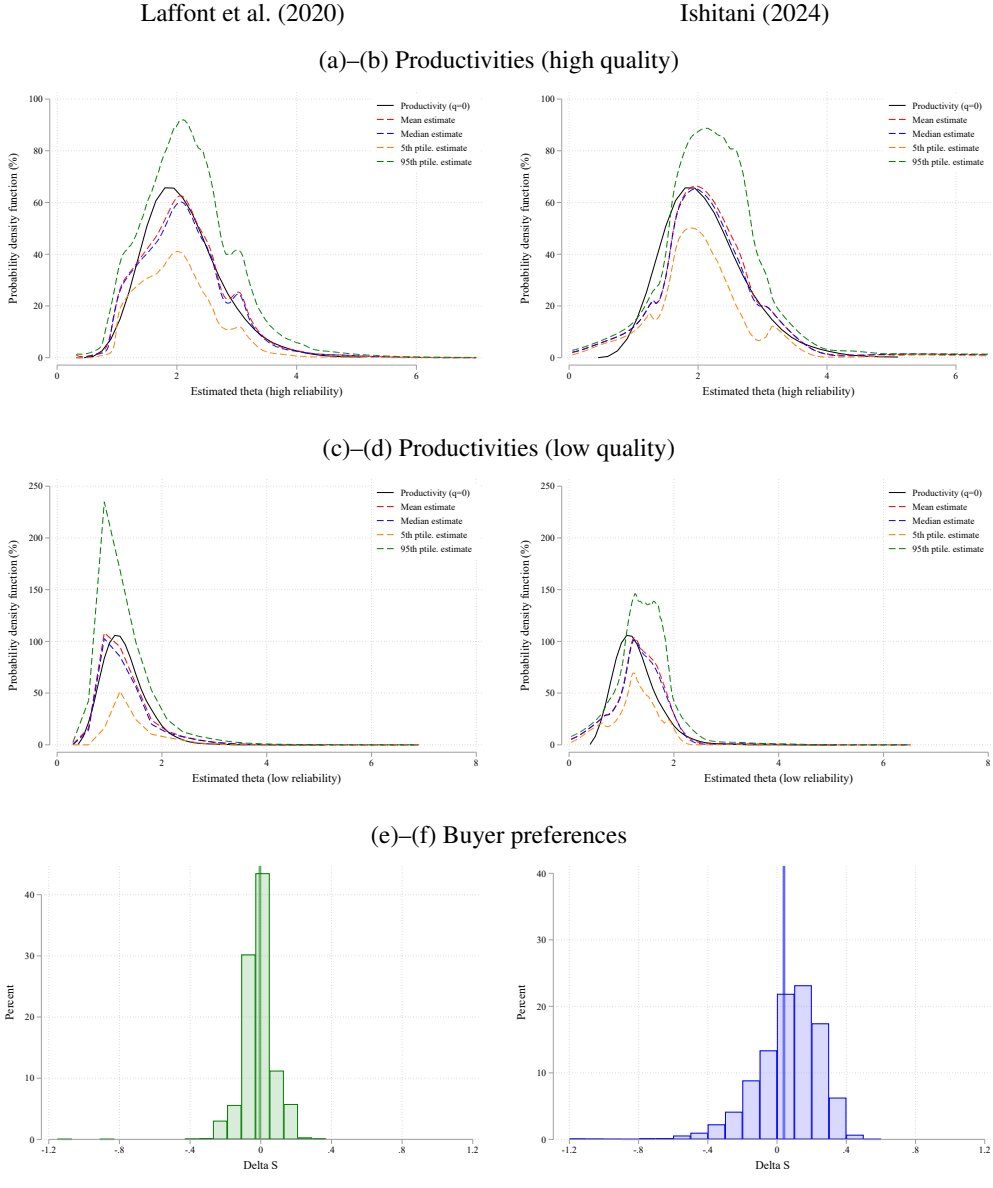


FIGURE 2.—Monte Carlo simulations of preferences and productivities