TI 2024-048/VII
Tinbergen Institute Discussion Paper

Bargaining Power and Quantity Discounts to Retailers: Evidence from India's Pharmaceutical Industry

Gianluca Antonecchia[1]
Ajay Bhaskarabhatla[2]
Enrico Pennings[3]

1 KU Leuven and Tinbergen Institute

2 Erasmus School of Economics and Tinbergen Institute

3 Erasmus School of Economics and Tinbergen Institute

Bargaining Power and Quantity Discounts to Retailers: Evidence from India's Pharmaceutical Industry*

Gianluca Antonecchia
KU Leuven†

Ajay Bhaskarabhatla
Erasmus School of Economics

Enrico Pennings
Erasmus School of Economics

July 12, 2024

Abstract

This paper develops a novel theory linking quantity discounts to bargaining power in scenarios where retailers, organized as a trade association, negotiate uniform wholesale prices with suppliers. Our theory predicts that suppliers offer greater quantity discounts in regional markets where they possess relatively less bargaining power, as a counterbalance to the higher national wholesale prices negotiated by the retailer trade association. We test these predictions using detailed product-level data from the Indian pharmaceutical industry, where significant geographic variations in quantity discounts are observed. Our findings provide empirical support for the proposed theory.

JEL Codes: L11, L42, D22
Keywords: quantity discounts, bargaining power, pharmaceuticals, India

*We thank Rubaiyat Alam, Tommaso Alba, Remi Avignon, Filippo Biondi, Jan De Loecker, Etienne Guigue, Liran Einav, Pranav Jindal, David Ronayne, Jo Van Biesebroeck, Frank Verboven and Weijie Yan for thoughtful comments; and the seminar and conference participants at Erasmus School of Economics, KU Leuven, University of Padua, EARIE (Rome) and RSA (L'Aquila), MaCCI (Mannheim), CEPR-Health (Toulouse), CIIP (Hong Kong), IIOC (Boston). We thank AIOCD for sharing the data and the Thakur Family Foundation for financial support.

†Contact: gianluca.antonecchia@kuleuven.be

1 Introduction

Suppliers rely on retailers to reach consumers across various industries. While suppliers hold market power, retailers possess countervailing buyer power, defined as the ability of buyers or retailers with bargaining strength to extract wholesale price concessions from suppliers (Galbraith, 1952). Retailers with greater buyer power can act as gatekeepers, securing larger price discounts by promising higher sales to suppliers and threatening to exclude those that do not provide concessions (e.g., Dobson and Waterson, 2007; Inderst and Mazzarotto, 2008; Wu, 2009). The growth of such confidential price concessions in the US pharmaceutical industry is well-documented by Dafny et al. (2017) and Ellison and Snyder (2010).[1] Beyond price concessions, suppliers often offer quantity discounts, which also remain undisclosed. The standard explanation for quantity discounts is second-degree price discrimination, where customers receive discounts for bulk purchases (e.g., Crawford and Yurukoglu, 2012). We depart from this literature by linking quantity discounts to bargaining power.

Our study focuses on the Indian pharmaceutical industry, characterized by a retail trade association of pharmacies with significant buyer power. We develop a theory to explain how the relative bargaining strength of the retailers in a geographic market influences the quantity discount they receive when suppliers negotiate a uniform wholesale price with the retailer association. The theoretical model employs a Nash bargaining framework in which suppliers and regional retailers negotiate over the wholesale price of a product variety. Given the wholesale price, retailers then set the retail price to maximize their profits. We distinguish between coordinated bargaining, where an association coordinates negotiations between a supplier and regional retailers over a uniform wholesale price, and decentralized bargaining, where regional retailers negotiate independently over their regional-specific wholesale price. In coordinated bargaining, suppliers use quantity discounts to compensate regional retailers with greater bargaining power who would have negotiated lower prices in decentralized bargaining. Consequently, quantity discounts are used as an instrument to guarantee resale price maintenance and are more likely in regions with stronger retailer bargaining power. Discounts do not reduce consumer prices but instead increase retailer profits and stabilize the trade association.

[1] Herper (2012) estimates that US pharmaceutical firms provide nearly 40 billion dollars annually in rebates to institutional payers such as insurance agencies, health management organizations, and government health agencies. Herper (2012) reports rebates for ten medicines, including Lipitor (35%), Plavix (3%), and Nexium (61%). Mattingly et al. (2018) find that discounted prices for the Veterans Administration in the US range from 9 to 74% of the wholesale acquisition cost.

We test our model using data on quantity discounts compiled by the retail pharmacies' trade association in India from March 2007 to June 2013. Annually, these discounts account for INR 31.6 billion (USD 620 million in 2012). Our model shows that higher retailer bargaining power decreases supplier concentration in regional markets and increases quantity discounts. We test the predicted negative relationship between supplier concentration and quantity discounts and find that higher supplier concentration decreases quantity discounts. For the same product variety, a one standard deviation increase in supplier concentration is associated with a 1.12 percentage point decrease in the probability of receiving quantity discounts and a 0.92 percentage point decrease in the share of quantity discounts provided. This relationship is stronger for medicines with lower elasticity of substitution and varieties from multiproduct suppliers. Our results are robust to alternative model specifications and proxies for supplier competition. Retail prices do not decrease in response to higher quantity discounts, confirming that benefits are not passed to consumers.

Our study makes several contributions. First, we contribute to the literature on buyer power by studying a case where suppliers compete, and retailers are effectively a monopsony. Buyer power usually results in price discounts for the retailers and its extent is related to the level of downstream concentration (e.g., Gaudin, 2018; Ho and Lee, 2017). In the case of monopsony, instead, the ability of retailers to extract price concession from the suppliers is related to the competition upstream. Using wholesale prices for antibiotics sold to US drugstores, Ellison and Snyder (2010) show that the extent of price discounts larger drugstores receive from pharmaceutical firms depends on the presence of competition among suppliers, indicating that buyer power does not always guarantee price discounts. Our paper complements the research on the drivers of price discounts to retailers by studying a context where only quantity discounts apply. We are the first to study quantity discounts for an entire industry and show that lower supplier bargaining power in a region leads to more quantity discounts but not lower retail prices. This original mechanism is not alternative to the standard second-degree price discrimination mechanism of quantity discounts, which might still be in place in our context (Allende et al., 2024).[2]

Second, we contribute to the literature of uniform pricing in heterogeneous geographic

[2]Although they might look equivalent, the quantity discounts that we observe are not volume-based price rebates paid for the retailer's purchase of one supplier's product, like, for example in Conlon and Mortimer (2021).

markets by studying the case of coordinated bargaining on prices in the Indian pharmaceutical industry. Uniform retail pricing is observed in many industries, across and within firms, and has relevant implications on consumer welfare (DellaVigna and Gentzkow, 2019; Antonecchia and Bhaskarabhatla, 2023; Daruich and Kozlowski, 2023). Uniform wholesale prices within product-across geographic markets is also frequently observed (e.g., Stroebel and Vavra, 2019). However, no study has focused on the vertical relationship between supplier and retailer in such a context with both a theoretical and empirical approach. We show that the bargaining power of the regional retailers that cannot be exerted on prices spills over into nonprice outcomes, such as quantity discounts. This mechanism contributes to offset the profit gap for the regional retailers created by the uniform pricing, without lowering the price for the consumers. Also, although coordinated/collective bargaining via trade unions and associations has a long tradition in the economic literature, empirical studies are mostly focused on wage setting (e.g., Bhuller et al., 2022), with no evidence on non-labor input or product wholesale price setting.[3] We model the uniform pricing in the Indian pharmaceutical industry as the outcome of a coordinated bargaining and show that quantity discounts help preserve resale price maintenance and the stability of the trade association.

Third, we contribute to the literature on vertical relations and market power in the pharmaceutical industry (Lakdawalla, 2018; Scott Morton and Kyle, 2011). Supplier-retailer contracting is central in healthcare and pharmaceutical markets because of the presence of large suppliers and large buyers (e.g., Grennan, 2013; Gowrisankaran et al., 2015; Ho and Lee, 2017; Dubois and Sæthre, 2020). We examine the case of India and the role of the retail trade association. The trade association is made up of 850,000 pharmacies whose entry is regulated and ownership is distributed among small mom-and-pop businesses. We show that its activity of coordinating the negotiation with suppliers might be beneficial for local retailers with small bargaining power. However, it limits competition downstream by sustaining uniform pricing and resale price maintenance. As individual pharmacies in India do not directly negotiate any quantity discounts with pharmaceutical firms, our study points to the role of the buyer power exercised by the trade association and the need for closer scrutiny of the association's vertical practices (Genesove and Mullin, 2001, 2007; Levenstein and Suslow, 2006; CCI, 2013).

[3] Allain et al. (2020) study the case of purchasing alliances among retailers from a theoretical perspective. Collective bargaining is also used as a counterfactual to evaluate the effect of price discrimination on welfare (e.g., Grennan, 2013).

The rest of the paper is organized as follows. Section 2 provides institutional background. Section 3 presents the model and its testable predictions. Section 4 details the data and descriptive statistics. Section 5 tests the model's predictions. Section 6 concludes.

2 Institutional Context

Trade associations wield significant buyer power by coordinating the decision-making of their members (Alé-Chilet and Atal, 2020; Genesove and Mullin, 2001; Levenstein and Suslow, 2006). In India, the All India Organization of Chemists and Druggists (AIOCD) orchestrates the activities of retail pharmacies. The AIOCD became the exclusive distribution channel for pharmaceutical firms in 1975 after two competing trade associations merged (Chaganti, 2006, p. 192). While this paper focuses on the AIOCD, similar trade associations with considerable countervailing power operate in various Indian sectors, often enforcing restrictive vertical trade practices such as price fixing, collective boycotts, exclusion from association membership, and refusal to deal (see, for a summary, Verma, 1981). Among the more than 2,000 trade associations in India, several have faced investigations for restrictive trade practices over the past five decades. The precursor to the current Competition Commission of India (CCI), the Monopolies and Restrictive Trade Practices Commission (MRTPC), investigated trade association activities in various sectors such as gas stations, film distribution, motor parts, jute products, yarn, timber, X-ray, alcohol, and tobacco (Verma, 1981, p. 65). Early cases include the Indian Jute Mills Association coordinating a 15 percent production reduction, the Indian Sugar Mills Association limiting sugar releases by 30 percent, radio and lamp sellers imposing uniform prices to eliminate competition, and cooking oil sellers fixing prices and boycotting those who did not comply (Verma, 1981). Rao and Sastry (1989) notes that between 1970 and 1983, 27,541 agreements on restrictive trade practices were registered with the commission, but only 0.76 percent were investigated due to limited resources.

Following economic liberalization in 1991, India introduced the Competition Act in 2002 and established the CCI in 2003, replacing the MRTPC. Unlike the MRTPC, which could only issue cease and desist orders, the CCI can impose deterrent penalties on trade associations for anticompetitive conduct. Efforts by pharmaceutical firms to bypass the AIOCD and sell directly through online pharmacies, physicians, or hospitals led to coor-

dinated nationwide boycotts by the AIOCD.[4] Consequently, pharmaceutical firms adhere to the AIOCD's restrictive trade practices (Bhaskarabhatla et al., 2016; Chaganti, 2006). The CCI investigated and penalized the AIOCD for restrictive practices such as uniform pricing, boycotts, and refusal to deal.[5]

The AIOCD plays a crucial role in coordinating its members' pricing decisions. Operating as an "open price" trade association, it imposes minimum trade margins and negotiates with suppliers to set uniform wholesale and retail prices, ensuring guaranteed retailer margins (CCI, 2013).[6] The AIOCD enforces these margins through product sales boycotts, cutting off supplies from non-compliant firms (Bhaskarabhatla et al., 2016). Reflecting its bargaining power, the pharmaceutical industry trade associations for domestic and foreign suppliers agreed in 1982 to guarantee minimum margins of 30 percent for unregulated medicines and 24 percent for price-controlled medicines (Bhaskarabhatla et al., 2016).

Our data reveal little regional variation in product prices, consistent with the uniform pricing imposed by the AIOCD. This raises important questions for policymakers and competition authorities about whether the concentration of buyer power in the retail trade association benefits consumers, particularly in countries like India (Dobson and Waterson, 1997). Pharmaceutical firms bear transportation costs, aligning with our theoretical model, and provide quantity discounts directly to retailers. These discounts are distributed to individual pharmacies alongside purchased quantities. Although the quantity discount might depend on regional bargaining strength and order size, we lack pharmacy-level data to assess the role of price discrimination.[7] These discounts are mea-

[4]Since at least 1980, no pharmaceutical firm in India could launch a new drug without a "no objection certificate" from the AIOCD (Bhaskarabhatla, 2020; Singh, 1981, 1984). Firms offering lower-than-acceptable retailer margins faced boycotts (Singh, 1984): "The Sandoz boycott, typical of the 'disciplinary' actions by Shah and Umedchand, began when Wander Pharmaceuticals, an associate company—both owned by the multinational Sandoz—sacked three wholesalers unwilling to accept reduced margins on some drugs." These practices extended beyond pharmaceuticals to other industries (Rao and Sastry, 1989).

[5]In two separate cases filed in 2011, the CCI found AIOCD vertical practices to be anticompetitive. In the first, the AIOCD instructed USV Pharmaceuticals, to refuse to supply to a distributor in Odisha and in another case Janssen Cilag Pharmaceuticals was instructed to boycott a distributor in Kerala.

[6]Open price trade associations, like the Sugar Institute in the US (1928-1936), required members to sell at open prices and publicly announced terms, eliminating price discrimination and prohibiting rebates so that large customers like Coca Cola and Kroger paid the same price as smaller ones (Genesove and Mullin, 2001, 2007). In India, the open price regime is facilitated by maximum resale price maintenance laws mandating the printing of the maximum resale price (MRP) on medicine packaging (Bhaskarabhatla, 2020). Maximum resale price maintenance specifies a retail price ceiling, preventing retailers from selling above a set amount (Shaffer, 1991).

[7]Pharmacies in India are typically small, family-owned businesses specializing in medicine sales and

sured by the AIOCD's subsidiary, which compiles data on the pharmaceutical industry, aggregated at the regional level with prices and quantities on a monthly basis. In the empirical section, we use data from this subsidiary, which competes with other providers like IMS Health, to test our theoretical model's predictions.

3 Theory

Our theory is broadly related to models of bargaining between suppliers and retailers taking into account the specific context of the pharmaceutical industry in India. Broadly consistent with our institutional setting and empirical framework, many suppliers produce one unique variety of a medicine that is sold to retailers in different regions. So, unlike Ho and Lee (2017) and Gaudin (2018) where there is strategic interaction between actors, we assume vertical bargaining with monopolistic competition. Both coordinated and decentralized bargaining models are examined where (i) decentralized bargaining involves bargaining between regional retailer pharmacies and suppliers over the wholesale price of a variety of a medicine, and (ii) coordinated bargaining refers to bargaining which takes place at the national level. With decentralized bargaining, the wholesale and retail prices are region-specific while coordinated bargaining results in a nationwide wholesale and retail price. As the bargaining powers of retail pharmacies and the pharmaceutical industry appear to be region-specific, optimal retail and wholesale prices differ across regions with decentralized bargaining. Coordinated bargaining subsequently yields unstable outcomes, as regions with high retail bargaining power will opt out and start bargaining on their own.[8] To avoid such a scenario, suppliers will try to compensate retail pharmacies in regions with high retail bargaining power. Here is where the specific focus and contribution of this paper come in: this compensation is done through quantity discounts of the pharmaceutical suppliers to retailers.

The theoretical part is structured as follows. In Section 3.1 we model how suppliers and retailers set the wholesale and the retail price of each product variety with decentralized

licensed by the government. Despite some size variation, most pharmacies serve local neighborhoods, hold inventory for a week, and do not place large orders, suggesting second-degree price discrimination is not a major factor in our data.

[8]Coordinated bargaining is different from centralized bargaining in the sense that the payoff with centralized bargaining is at the national level and regions would not matter. A union would maximize the national payoff with respect to a decision variable, e.g., wages. With coordinated bargaining, an association maximizes the sum of the regional Nash bargaining payoffs with respect to the same decision variable; see Holden (1988) for further discussion of central and local bargaining.

bargaining. In Section 3.2 we examine the same setup but with coordinated bargaining. Section 3.3 introduces quantity discounts as a means to compensate powerful retailers which are worse off with coordinated bargaining. In Section 3.4 we develop additional implications of the model and discuss an extension to multiproduct suppliers in Section 3.5.

3.1 Decentralized bargaining over wholesale price

Demand. A monopolistically competitive manufacturer produces one variety i of a medicine and sells it to a number of regional retailers. Retailers, located in the region j and selling the variety, bargain with the manufacturer on a regional wholesale price (w_{ij}) and set a regional retail price (p_{ij}). Regional consumers maximize their utility subject to a budget constraint and with a constant elasticity of substitution between the differentiated varieties of the medicine. Demand for a variety i of a medicine in region j can be written as:

$$q_{ij}(p_{ij}) = A_j p_{ij}^{-\sigma} \tag{1}$$

where σ stands for the country-wide elasticity of substitution across varieties within the medicine, as well as the price elasticity of demand. In this equation, the constant can be written as $A_j = Y_j P_j^{\sigma-1}$, where P_j equals the regional price index of the medicine and Y_j is the exogenously determined total amount spent in the medicine market in line with Dixit and Stiglitz (1977).[9]

Suppliers and regional retailers. The regional retailer bargains with each supplier over the wholesale price (w_{ij}) and subsequently maximizes profit with respect to the retail price, knowing w_{ij}. Therefore, the optimal outcome for the wholesale price and the retail price can be seen as the solution to a two-stage game where in the first stage the supplier and the regional retailer bargain over the regional wholesale price, and in the second stage the retailers set the regional retail prices.[10] For variety i of the medicine in region j the retailer bears a variable cost corresponding to the wholesale price (w_{ij}), while the supplier bears a region-variety-specific fixed cost (F_{ij}) - which includes the negotiation cost - and

[9]The regional price index can be written as $P_j = \left(\sum_{i=1}^{N_j} p_{ij}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$. The number of firms/varieties in a region (N_j) is assumed to be large enough so that the firms take the price index as given. The exogeneity assumption of Y_j is appropriate for the pharmaceutical industry where demand for a medicine is due to the therapeutic need of the medicine and not to the number of firms serving the market.

[10]The outcome of the model does not depend on the timing of the game. A simultaneous bargaining over w_{ij} and p_{ij} would give the same result.

a marginal cost of production (c) - which is the same across all varieties and regions.[11] Similar to the model of Montagna (1995), only those suppliers for whom the operating profit is higher than the fixed cost will be active in the market.[12]

As the game is sequential, where there is first bargaining over the wholesale price and then retailers set the retail price, it is solved by backward induction. The regional retailer sets the regional price conditional on the regional wholesale price bargained with the supplier. Retailer's regional profit can be written as

$$\pi_{ij}^{R} = p_{ij} q_{ij}\left(p_{ij}\right) - w_{ij} q_{ij}\left(p_{ij}\right) \tag{2}$$

Maximizing the regional profit of the retailer with respect to its price gives:

$$p_{ij}^{*} = \frac{\sigma}{\sigma - 1} w_{ij} \tag{3}$$

Demand at the optimal price is equal to $q_{ij}^{*} = A_j \left(\frac{\sigma w_{ij}}{\sigma - 1}\right)^{-\sigma}$ and the maximized retailer's profit is equal to $\pi_{ij}^{R*} = \frac{A_j}{\sigma - 1}\left(\frac{\sigma}{\sigma - 1}\right)^{-\sigma}\left(w_{ij}\right)^{1-\sigma}$.

Bargaining. In the first stage of the game, the regional retailer and the supplier negotiate over the regional wholesale price for each variety i simultaneously. In the Nash bargaining game, regional retailers make profit π_{ij}^{R*} and the supplier makes profit $\pi_{ij}^{S*} = \left(w_{ij} - c\right) A_j \left(\frac{\sigma w_{ij}}{\sigma - 1}\right)^{-\sigma} - F_{ij}$. The Nash bargaining outcome for w_{ij} would be the outcome of:

$$\max_{w_{ij}} \left(\pi_{ij}^{R*} - \overline{\pi}_0^{R}\right)^{\beta_{ij}} \left(\pi_{ij}^{S*} - \overline{\pi}_0^{S}\right)^{1-\beta_{ij}} \tag{4}$$

where β_{ij} denotes the bargaining power of the retailers in variety i in region j and $1 - \beta_{ij}$ equals the bargaining power of the supplier in variety i in region j. Note that the threatpoint, or the disagreement profit, for the supplier is equal to $\overline{\pi}_0^{S} = -F_{ij}$ as the fixed cost is already sunk when the firm starts the negotiations. As the retailer's profit

[11] The fixed cost (F_{ij}) is composed of a variety-specific cost - exemplified by the negotiation cost with retailers - and a regional-specific cost - exemplified by the cost of market analysis and network maintenance in a region. Fixed costs are sunk at the start of the negotiation and contribute to the profitability, entry and survival of the variety in the region. We can let the marginal cost of production include an iceberg transportation cost that varies across regions without altering the implications of our results.

[12] Our model introduces heterogeneity in the variable costs of the retailer, but, unlike in the model by Melitz (2003) who introduced random productivity differences across firms as random differences in variable costs, the model endogenizes differences in variable costs by accounting for differences in firm bargaining power. Our model is similar in nature to Eckel and Egger (2009) who examined firms producing differentiated goods and negotiating with unions over the wage cost.

from not selling a variety is zero and this has no effect on the negotiations over prices of other varieties when there is monopolistic competition and simultaneous bargaining for alternative varieties, the threatpoint for the retailer is zero ($\overline{\pi}_0^R = 0$).[13]

Equilibrium. Maximizing Equation (4) leads to

$$\max_{w_{ij}} (w_{ij})^{\beta_{ij}-\sigma} (w_{ij} - c)^{1-\beta_{ij}} \tag{5}$$

whose first order condition gives

$$w_{ij}^* = \frac{\sigma - \beta_{ij}}{\sigma - 1} c \tag{6}$$

showing that wholesale prices are a decreasing function of the elasticity of substitution. The regional retail price can now be written as $p_{ij}^* = \frac{\sigma}{\sigma-1}\frac{\sigma-\beta_{ij}}{\sigma-1}c$, which is also a decreasing function of the elasticity of substitution. If the regional retailers have all the bargaining power ($\beta_{ij} = 1$), then they will pay the supplier only the marginal cost. If the supplier has all the bargaining power in the region ($\beta_{ij} = 0$), then the supplier will charge the monopoly price, and there is double marginalization for the retail price.

Assuming the elasticity of substitution between varieties of the same medicine is greater than one ($\sigma > 1$), these results show that varieties with higher retail bargaining power in the region have lower wholesale and retail prices, lower producer markup ($\frac{w_i}{c}$), unaltered retail markup ($\frac{p_i}{w_i}$), lower supplier profits and higher retail profits. These results are valid both within variety across regions—variety i has a lower wholesale price in the regions where it has higher retail bargaining power than in the regions where it has lower bargaining power—and within a region across varieties—variety i has lower a wholesale price than variety h if in the same region variety i has higher retail bargaining power than variety h.

Competition. One variety is supplied in a region only if it generates profits for the supplier. Variety i is supplied in region j only if the bargaining power β_{ij} is such that the wholesale price w_{ij} is high enough to cover the fixed costs of the supplier F_{ij}. This requires β_{ij} to be lower than a $\overline{\beta}_{ij}$ threshold above which variety i has negative supplier profits in region j and would not be supplied. The higher this threshold the higher the probability that $\beta_{ij} < \overline{\beta}_{ij}$. The number of varieties supplied in region j (N_j) is given by

[13]Even in case of sequential bargaining, any disagreement profit for the retailers that is proportional to the retailer profit would not affect the solution of the optimal wholesale price.

the number of varieties with $\beta_{ij} < \overline{\beta}_{ij}$.

$$N_j = \sum_i 1[\pi_{ij}^M(\beta_{ij}, F_{ij}) > 0] = \sum_i 1[\beta_{ij} < \overline{\beta}_{ij}] \tag{7}$$

Therefore, the regions with higher average $\overline{\beta}_{ij}$ and, consequently, higher average β_{ij}, are also those served by more varieties and with lower supplier concentration.[14] To formalize this result, consider β_{ij} as a function of three components: i) β_i = average retail bargaining power of the variety across all regions; ii) β_j = average retail bargaining power of the region across all varieties; iii) b_{ij} = variety-region specific additive component with expected value zero. In formula:

$$\beta_{ij} = f(\beta_i, \beta_j) + b_{ij} \tag{8}$$

where β_i can be interpreted as the appeal of the variety in the country. Since β_i does not vary across regions and b_{ij} is expected to be zero, the expected number of varieties supplied in a region is a function of the average bargaining power of the regional retailers:

$$E[N_j] = \phi(\beta_j) \tag{9}$$

which entails that supplier concentration is an inverse function of β_j.[15]

3.2 Coordinated bargaining over wholesale price

Demand. Whereas the previous subsection examined bargaining at the regional level, we now consider the case of coordinated bargaining where a nationwide retail trade association coordinates bargaining efforts for a nationwide wholesale price (w_i) and sets a uniform country-level retail price (p_i). Demand for variety i of a medicine in region j remains as in Equation (1) with the only difference that the retail price is the same in all

[14] As an example, suppose that the fixed cost is the same for all the varieties within the region, $F_{ij} = F_j$. In such a case $\overline{\beta}_{ij}$ would be the same for all the varieties within the region, $\overline{\beta}_{ij} = \overline{\beta}_j$. All the varieties with a β_{ij} lower than the threshold will be supplied in the region. The higher this threshold, the higher the average β_{ij} of the region, the more the varieties that will be available in the region, the lower the concentration.

[15] Our analysis considers the steady state. Montagna (1995) explains how in a setting like ours firms incumbents stay in the market as long as it is profitable and potential entrants all face the same ex-ante uncertainty about their level of bargaining power and fixed cost. Uncertainty disappears once firms pay an entry cost and values for their bargaining power and fixed cost are drawn from a random distribution.

regions.[16]

Suppliers and retailer association. Also supplier cost conditions (c and F_{ij}) are the same as in the decentralized case.[17] The main difference is that the retail trade association maximizes the aggregated profit with respect to the uniform retail price, after having bargained over a nationwide w_i. Therefore, the optimal outcome for the wholesale price and the retail price can be seen again as the solution to a two-stage game where in the first stage the supplier and the retail association bargain over the manufacturer price, and in the second stage, the retail association sets the price.

With a uniform wholesale and retail price, the retailer's regional profit can be written as $\pi_{ij}^R = p_i q_{ij}(p_i) - w_i q_{ij}(p_i)$. The trade association maximizes total national retail profit with respect to its price:

$$\max_{p_i} \sum_j \pi_{ij}^R \tag{10}$$

The optimal uniform retail price for variety i is:

$$p_i^* = \frac{\sigma}{\sigma - 1} w_i \tag{11}$$

Regional demand at the optimal uniform price is equal to $q_{ij}^* = A_j \left(\frac{\sigma w_i}{\sigma - 1}\right)^{-\sigma}$ and the maximized retailer's profit is equal to $\pi_{ij}^{R*} = \frac{A_j}{\sigma - 1} \left(\frac{\sigma}{\sigma - 1}\right)^{-\sigma} (w_i)^{1-\sigma}$.

Bargaining. In the first stage of the game, the retail trade association and the supplier negotiate the wholesale price. In the Nash bargaining game, both the suppliers and the retail association consider the sum of all their regional profits over variety i. The regional retailers have profit π_{ij}^{R*} and the supplier has profit ${\pi^M}_{ij}^* = (w_i - c) A_j \left(\frac{\sigma w_i}{\sigma - 1}\right)^{-\sigma} - F_{ij}$. The Nash bargaining outcome for w_i would be the outcome of

$$\max_{w_i} \sum_j \left(\pi_{ij}^{R*} - \overline{\pi}_0^R\right)^{\beta_{ij}} \left({\pi^M}_{ij}^* - \overline{\pi}_0^S\right)^{1-\beta_{ij}} \tag{12}$$

where we assume variation in β_{ij} across regions. Note again that the threat point, or disagreement profit, for the supplier is equal to $\overline{\pi}_0^S = -F_{ij}$ and remains zero for the retailer.

[16] Despite the country-wide uniform (wholesale and retail) price, the regional price index of the medicine (P_j) varies across regions because of different number of varieties sold in each region.

[17] Since the supplier negotiates the wholesale price with the retail trade association, the negotiation cost component of F_{ij} is paid only once and not for every region, as in the decentralized case. Splitting this negotiation cost across all regions implies that the fixed cost per region F_{ij} in the coordinated case is smaller than that in the decentralized case.

Equilibrium. Maximizing Equation (12) corresponds to:

$$\max_{w_i} \sum_j Aj \left(w_i\right)^{\beta_{ij}-\sigma} \left(w_i - c\right)^{1-\beta_{ij}} \tag{13}$$

which, comparing the equation with equation (5), means maximizing a weighted average of the regional Nash bargaining payoffs. The maximization function can be approximated by

$$\max_{w_i} \sum_j \left(w_i\right)^{\beta_i^A-\sigma} \left(w_i - c\right)^{1-\beta_i^A} \tag{14}$$

where β_i^A, with $\min_j \beta_{ij} < \beta_i^A < \max_j \beta_{ij}$, denotes the model-implied bargaining power of the retailer association in variety i across all regions and $1 - \beta_i^A$ equals the bargaining power of the supplier in variety i across all regions.[18] The first order condition for maximization gives

$$w_i^* = \frac{\sigma - \beta_i^A}{\sigma - 1} c \tag{15}$$

The equation shows again that wholesale prices are a decreasing function of the elasticity of substitution. Note that equation (15) corresponds to equation (6) if $\beta_{ij} = \beta_i^A$. The retail price can now be written as:

$$p_i^* = \frac{\sigma}{\sigma - 1} \frac{\sigma - \beta_i^A}{\sigma - 1} c \tag{16}$$

which is again a decreasing function of the elasticity of substitution. If regional retailers have heterogeneous bargaining power over variety i, whatever the functional form of β_i^A, the uniform wholesale price of the variety would be different from the optimal wholesale price of the variety if decentralized bargaining was in place. In fact, retailers a region with a retail bargaining power higher than that of the association ($\beta_{ij} > \beta_i^A$) would be forced to pay a higher price than their regional optimal wholesale price ($w_{ij} < w_i$) and set a higher than optimal retail price ($p_{ij} < p_i$). This would reduce regional retailers' profits and make coordinated bargaining unstable as the regions with the highest bargaining power would be better off with decentralized bargaining.

As in the decentralized bargaining case, assuming that the elasticity of substitution is greater than one ($\sigma > 1$), varieties in which the trade association has a higher bargaining

[18] β_i^A can be interpreted as the bargaining power of the national retail association in a centralized Nash bargaining game. The functional form of β_i^A can be general: $\beta_i^A = f(\beta_i) + b_i^A$, where β_i is the average retail bargaining power of the variety across all regions - or the variety appeal, like in Equation (8) - and b_i^A a variety-specific additive component.

power have lower wholesale and retail prices, a lower supplier markup ($\frac{w_i}{c}$), an unchanged retail markup ($\frac{p_i}{w_i}$), a lower supplier profits and higher retail profits. These results are valid only across varieties within a region as one variety has a uniform price across regions.

Competition. Likewise the decentralized bargaining case, one variety is supplied in a region only if it generates profits for the supplier. The difference is that here the bargaining power that determines the wholesale price is uniform across all regions (β_i^A) but the threshold above which variety i has negative supplier profits in region j varies across regions ($\overline{\beta}_{ij}$) as it depends on the regional-specific fixed costs (F_{ij}). The number of varieties supplied in region j (N_j) is given by the number of varieties with $\beta_i^A < \overline{\beta}_{ij}$.

$$N_j = \sum_i 1[\pi_{ij}^M(\beta_i^A, F_{ij}) > 0] = \sum_i 1[\beta_i^A < \overline{\beta}_{ij}] \tag{17}$$

This result does not alter the result in Equation (9) for which the higher the average bargaining power of the region, the more the varieties supplied. The only difference is that, in the coordinated bargaining scenario, the average bargaining power of the region is that of the trade association and no longer that of the regional retailers. The average bargaining power of the varieties changes across regions because every region has a different set of varieties supplied.

3.3 Quantity discounts

As coordinated bargaining yields an unstable outcome, suppliers may compensate the regional retailers with greater bargaining power. In this section, we will consider a quantity discount, which has the significant advantage of maintaining uniform pricing, for such an implicit monetary transfer. When the supplier gives a quantity discount of $100x_{ij}\%$ to the regional retailer j for variety i, the retailer can sell a fraction x_{ij} of the quantity sold without paying w_i.

The quantity discount can compensate for the missing profits of the regional retailers with higher bargaining power than that of the retail association. To fully restore the differences in regional bargaining power of those retailers, discounts should be the amount of quantity not paid by the regional retailers that yield the same profits as in the decentralized bargaining scenario:

$$p_{ij}(w_{ij})q_{ij}(p_{ij}) - w_{ij}q_{ij}(p_{ij}) = p_i(w_i)q_{ij}(p_i) - w_i q_{ij}(p_i)(1 - x_{ij})$$

Conditional upon $\beta_{ij} > \beta_i^A$, solving for x_{ij} gives:

$$x_{ij}^* = \frac{1}{\sigma - 1}\left[\left(\frac{\sigma - \beta_{ij}}{\sigma - \beta_i^A}\right)^{1-\sigma} - 1\right] \tag{18}$$

and $x_{ij}^* = 0$ for $\beta_{ij} <= \beta_i^A$. As the incidence of giving a quantity discount does not depend on the quantity sold, we provide an explanation for quantity discounts in addition to second degree price discrimination. Conditional upon giving a quantity discount, the discount as percentage of the quantity sold is independent of the quantity sold, but the total discount is increasing in the quantity sold.

In Figure 1, the y-axis reports the quantity discounts that should be given at different levels of β_{ij} (x-axis) when $\beta_i^A = 0.4$ (Panel A) and $\beta_i^A = 0.6$ (Panel B) for an elasticity of substitution of 4, 6 and 8.

The equation yields two hypotheses, which we will test in our empirical analysis.

Hypothesis 1: *A variety offers more quantity discounts in regions with a higher retailer bargaining power.*

Hypothesis 1 relates to the differences in bargaining power across regions and the model-implied bargaining power of the retail association for each variety. The equation shows that the smaller the bargaining power gap between the regional retailer and the trade association, the smaller the extent of quantity discounts given as compensation. Obviously, discounts are given only when $\beta_{ij} > \beta_i^A$. As an example, consider the case in which the bargaining power of the retail association is equal to 0.4 and the elasticity of substitution is equal to 4 (Figure 1, Panel A, solid line). If the bargaining power of the regional retailer is also 0.4 the discounts will be zero because the wholesale price has been set optimally for the regional retailers. If the bargaining power of the regional retailer is higher than 0.4, say 1 (that is, the bargaining power rests entirely with the retailers), then a 24 percent quantity discount should be given to the regional retailers to compensate for the higher wholesale price paid.

Hypothesis 2: *The difference in bargaining power between regional retailers and the trade association induces more quantity discounts in medicines with lower elasticity of substitution.*

Hypothesis 2 is based on the result that quantity discounts are a decreasing function of the elasticity of substitution. This originates from the result that wholesale and retail

prices are a decreasing function of this elasticity (Equation (15) and (16)). It is assumed that all varieties of the same medicine have the same elasticity of substitution and within and across regions. Therefore, regional retailers with significant bargaining power would obtain a relatively higher return when the elasticity of substitution is lower, and hence would need a higher compensation for giving up this higher regional bargaining power. It follows that medicines with a higher elasticity of substitution have a lower retail price at the national level and, given the bargaining power of the regional retailers and the trade association, have less discounts. This effect is exemplified in both panels of Figure 1. For any given bargaining power gap between regional retailers and the trade association, the higher the elasticity of substitution, the lower the discounts.

3.4 Other implications of the model

3.4.1 Quantity discounts versus price discounts

A natural question that arises is whether suppliers in our model can offer equivalent price discounts instead of quantity discounts. Quantity discounts offer unique and significant advantages over price discounts. To establish this, we quantify the excess profits that retailers in regions with strong bargaining power receive through quantity discounts. The unit value of the quantity discount is equal to the discount per unit quantity, as derived in Equation (18), multiplied by the wholesale price, as derived in Equation (15). This value is given by $\frac{1}{\sigma-1}\left[\left(\frac{\sigma-\beta_{ij}}{\sigma-\beta_i^A}\right)^{1-\sigma}-1\right]\frac{\sigma-\beta_i^A}{\sigma-1}c$. The total value of the discount is the unit value multiplied by the demand $q_{ij}^{\star}=A_j\left(\frac{\sigma}{\sigma-1}\frac{\sigma-\beta_i^A}{\sigma-1}c\right)^{-\sigma}$. This expression for the value of the quantity discount shows that it increases with relative regional bargaining power and regional demand, while decreasing with production cost and elasticity of substitution. Therefore, the monetary compensation accounts for regional bargaining power, regional demand level, production cost, and elasticity of substitution, which relates to the market power of the variety. Although increased elasticity of substitution reduces the price and increases demand, it also lowers the unit value of the quantity discount, resulting in a total value of the discount that decreases with elasticity of substitution.

Quantity discounts offer several advantages over price discounts. First, quantity discounts can incorporate cost, demand, and bargaining conditions without requiring an explicit financial transfer. Second, quantity discounts maintain uniform wholesale and retail prices. In contrast, price discounts would lower input prices for retailers, which would

then be passed on to consumers as retailers seek to increase market share and maximize profits. Furthermore, price discounts would be visible to all and could lead to undesirable renegotiations and inter-region trade to exploit price differences. Thus, quantity discounts offer a more advantageous mechanism than price discounts in this context.

3.4.2 Quantity discounts and consumer prices

The model yields additional insights that merit discussion. We begin with a summary of the winners and losers in the model. Coordinated bargaining with quantity discounts benefits all retailers, irrespective of their bargaining power, as the retailers across all regions are represented by a single association. Retailers with relatively weak bargaining power benefit from this coordinated approach, as they pay a lower wholesale price negotiated collectively. Retailers with relatively strong bargaining power maintain their profitability by receiving additional quantity discounts, even when the wholesale price remains at the coordinated level. Conversely, coordinated bargaining reduces suppliers' profits, as they transfer surplus to retailers with significant bargaining power. Surprisingly, consumers in our model are overall unaffected by these transfers, as they continue to pay a uniform price. This price is higher in regions with relatively strong retailer bargaining power and lower in regions with relatively weak bargaining power, compared to the case of decentralized bargaining. In other words, while the use of quantity discounts by suppliers enriches retailers with relatively strong bargaining power, the benefits of these discounts are *not* passed on to consumers through lower prices.

3.4.3 The incentives and costs of giving quantity discount

The previous discussion clarifies that retailers have an incentive to bargain in a coordinated manner over wholesale and retail prices. Moreover, suppliers also have a strong incentive to offer quantity discounts and prevent decentralized bargaining. Under decentralized bargaining, regional retailers would exploit their regional bargaining power, leading to different wholesale and retail prices. This would result in inter-region trade to exploit these price differences. Additionally, there would be high organizational costs associated with negotiating with numerous parties instead of a single coordinating organization. Different prices would also necessitate varied packaging and labeling, increasing operating costs for suppliers. Thus, consistent with DellaVigna and Gentzkow (2019), albeit through different reasoning, there is a strong incentive for suppliers to employ uni-

form pricing. Furthermore, suppliers prefer coordinated bargaining because a strong retail trade association, such as the AIOCD, ensures resale price maintenance through its retaliatory and exclusionary vertical practices. While we have derived the costs of providing quantity discounts, which are proportionate to the market, the benefits of offering these discounts to prevent a breakdown in negotiations, as discussed earlier, will be proportional to these costs. When there is a fixed cost associated with giving quantity discounts, discounts will not be provided if the demand value is below a certain threshold. In such cases, the market is not very profitable due to relatively low demand. Therefore, there is a positive relationship between regional sales and the strategy of providing quantity discounts.

3.4.4 Discounts help retail trade association stability

An important implication of our model concerns the role of quantity discounts in stabilizing the trade association. Regional retailers for whom $\beta_{ij} < \beta_i^A$ enjoy lower wholesale prices with coordinated bargaining. Regional retailers for whom $\beta_{ij} > \beta_i^A$ can obtain compensation for the higher wholesale price with centralized bargaining through quantity discounts. This allows for a trade association to form across regions with different strengths in terms of their individual bargaining power to organize as a single organization at the national level and operate under uniform pricing regime. The trade association would also proliferate more variety as firms would not sell in regions with strong bargaining power when there is decentralized bargaining. With coordinated bargaining, this region may sell the variety when it is compensated for the non-exerted bargaining power. If the region with strongest bargaining power does not sell the variety under coordinated bargaining, it may be the case that a variety does not offer any discounts, which is consistent with the evidence that not all varieties give discounts.

3.4.5 Model extension to multiproduct suppliers

In the pharmaceutical industry suppliers might produce more product varieties across different medicines. The model can be extended to study if discounts change when a supplier is a multiproduct firm. These suppliers bargain with the trade association on the wholesale price of multiple product varieties and are supposed to have a higher bargaining power compared to suppliers bargaining for one variety.

If we assume that the number of products of a supplier influence positively its bargaining power we can extend Hypothesis 1 as follows: *in regions with a higher retailer bargaining power a supplier serving relatively less varieties offers more quantity discounts than a supplier selling relatively more varieties.* This hypothesis would be valid for suppliers that offer more varieties both within the same medicine market and across different medicines.

4 Data and descriptive evidence

4.1 Data

We utilize data obtained from India's retail pharmacy trade association, the AIOCD, covering the period from March 2007 to June 2013. The AIOCD data represents a comprehensive census of pharmaceutical firms in India, encompassing the entire country and divided into 23 regional medicine markets. This dataset includes monthly wholesale and retail prices, quantities of drugs sold in each regional market, and their corresponding monthly quantity discounts over 76 months. These data are compiled by the AIOCD's subsidiary through sales audits of retailers. For each of the 85,384 varieties identified by a unique stock-keeping unit (SKU), we have detailed information on the medicine, including the active ingredient (e.g., paracetamol, atorvastatin), delivery form (e.g., injection, tablet), dosage strength (e.g., 10 mg, 100 ml), and package size (e.g., number of tablets, syringes). The dataset contains nearly 35.4 million variety-region-month observations, spanning 681 firms, 23 regions, and 18,079 medicine formulations, defined by the combination of active ingredients, dosage form, and drug strength (Table 1). The terms medicines, medicine formulations, and products are used interchangeably. Within each medicine market, varieties differ by brand or pack size.

4.2 Descriptive Statistics on Quantity Discounts

We define confidential quantity discounts as the quantity of units provided by the manufacturer to the retailer free of charge, as reported in the AIOCD data.[19] We measure quantity discounts in two ways: (i) at the extensive margin, as the share of variety-

[19]Our data are aggregated at the variety-region-month level, not at the pharmacy level, precluding analysis of how individual pharmacy characteristics such as buyer size influence quantity discounts.

region-month observations with quantity discounts (*Dummy discounted*); and (ii) at the intensive margin, as the ratio of units given as quantity discounts to those purchased at wholesale prices by retailers, considering only the observations that include quantity discounts (*Discounted-to-undiscounted ratio*). In our sample, 9.2 percent of the variety-region-month observations offer quantity discounts, and among these, the discounted-to-undiscounted ratio is 15.8 percent (meaning that for every six units sold, one is given for free). A variety does not uniformly offer discounts across all regions, displaying considerable cross-sectional heterogeneity in discounting practices within a variety. On average, 19.7 percent of the varieties offer quantity discounts monthly. Table 2 shows that quantity discounts are region-specific even when products are aggregated at the formulation or firm level. Discounts are given in 30.7 percent of the medicine formulations, but when considering formulations across regions, only 14.7 percent have varieties giving quantity discounts. At an aggregate level, almost two-thirds of firms provide quantity discounts at the country level (i.e., in at least one region), while 38.6 percent do so when considered at the firm-region level.

The descriptive statistics in Table 2 indicate that quantity discounts are a significant phenomenon in the Indian pharmaceutical industry, though the practice varies greatly across regions. Figure 2 shows notable regional variation in both measures of quantity discounts. Larger firms and their best-selling medicines more commonly offer quantity discounts. Panels A and B of Figure 3 reveal that firms and medicine formulations offering discounts are, on average, 29 and 32 percent larger, respectively, than those not offering discounts.

We leverage the geographic disaggregation of the data to observe the same variety offering discounts in one geographic market but not in another. The model predicts this is influenced by the bargaining power of regional retailers. Following Galbraith (1952) and Ellison and Snyder (2010), we assume a negative relationship between retailer bargaining power and supplier concentration, expecting to observe a negative relationship between discounts and supplier concentration. Figure 4 plots the Herfindahl-Hirschman Index (HHI) at the regional level against the share of firms giving quantity discounts. Consistent with our expectations, the figure shows that higher levels of concentration correlate with lower shares of firms providing quantity discounts and a lower share of quantity discounts relative to undiscounted units sold. In Table A.1 in the Appendix, we demonstrate that this negative correlation holds at the formulation, active ingredient-dosage form, and active ingredient levels.

While a product variety may offer different discount shares across regions, its price remains consistent across all regions. Wholesale and retail prices vary significantly across varieties but are uniform within a variety across regions.[20]

5 Results

5.1 Testing the predictions of the model

We test the hypotheses derived from the model solution for discounts in Equation (18).

5.1.1 Hypothesis 1: How retail bargaining power affects discounts

Hypothesis 1 states that one variety offers more quantity discounts in regions with a higher retailer bargaining power. It follows from Equation (18), in which there is a positive relationship between the bargaining power gap between regional retailer and trade association and quantity discounts. Since the bargaining power of the trade association for one variety is invariant across regions (β_i^A), a variety would give more discounts in the regions whose retailers have more bargaining power (β_{ij}).

Regional bargaining power is not a variable that can be observed. To address this issue, we refer to the findings in Ellison and Snyder (2010) who show that supplier competition is a prerequisite for the retailers to have bargaining power. We proxy regional bargaining power with a supplier concentration index of the regional medicine market. High supplier concentration signals low average regional bargaining power of the retailers. An additional recommendation to use supplier concentration to proxy for regional bargaining power comes from the model. From Equation (8) we can assume that β_i^A is a component of β_{ij} and the difference between the two is a function of the average bargaining power of the varieties served in the region (β_j).[21] From Equation (9) and (17) we derive that the regional bargaining power of the medicine influences positively the number of varieties and, therefore, negatively the supplier concentration. In other words, the more concentrated is the formulation geographic market, the less retail bargaining power one variety has. Higher HHI (lower regional retail bargaining power), controlling for variety-time FE and

[20] Table A.2 in the Appendix reports the moments of the distributions of HHI and prices.

[21] It is not necessary that β_i^A is the average retailer bargaining power of the variety across across all the regions. If it is larger or smaller only the distribution of the error term would change.

region-time FE, is expected to correlate with lower discounts. To test Hypothesis 1, we estimate the following equation:

$$D_{ijt} = \alpha HHI_{gjt} + \gamma y_{gjt} + \theta_{it} + \delta_{jt} + \epsilon_{ijt}, \tag{19}$$

where D_{ijt} is either (i) a dummy taking value 1 when variety i in region j and month t gives quantity discounts; or (ii) the share of quantity discounts relative to the undiscounted units sold by variety i in region j and month t.[22] Our main explanatory variables are HHI_{gjt}, the HHI index at the formulation (g)-region(j)-month (t) level. The higher the regional HHI, the lower the bargaining power of the regional retailers, and the lower the quantity discounts they receive. We control for the regional sales (in logs) of the medicine formulation (y_{gjt}), as to control for the usual explanation for second degree price discrimination, and a set of fixed effects. Variety-month FE (θ_{it}) captures changes in the variety discount policy across regions or consumer preference for the variety or changes in the bargaining power of the association or changes in competition within a formulation at the national level (entry of national competitors). With the inclusion of variety-month FE we exploit the heterogeneity in discounts given by a variety across regions and relate it to the differences in regional supplier concentration of the formulation. We also control for region-month FE (δ_{jt}), capturing changes in regional policy (liberalization, taxes) and aggregate consumer tastes. According to Hypothesis 1 the parameter α is expected to be negative.

In Table 3, we report the estimated coefficients of Equation (19).[23] A variety is less likely to give quantity discounts in regions where the HHI index of the formulation is higher (Column 1-4). Similarly, on the intensive margin, a variety gives a lower share of discounts in regions where the HHI index of the formulation is higher (Column 5-8). The estimated coefficients remain negative and stable when introducing the explanatory variables one-by-one. Regional sales - exogenous variables in the model - are also positively correlated with the extent of quantity discounts given by a variety (Column 3 and 7). By including variety-region FE we control for variety-specific demand factors within the region and can capture the relationship between the changes in discounts and HHI over time for the same variety-region. Columns 4 and 8 confirm the negative relationship between HHI and

[22] The estimates in which the dependent variable is the discounted-to-undiscounted ratio are conducted on the sample of variety-region observations that have positive quantity discounts over the sample period.

[23] Despite both discount measures are limited dependent variables, linear models are recommended to avoid the incidental parameter problem, likely to occur when estimating many fixed effects using non-linear models on a large number of observations.

discounts. The coefficients of our preferred specification, as from Equation (19), should be interpreted as follows: a 1 percentage point increase in HHI in the formulation-region decreases the probability of a variety to be discounted in the region by 0.34 percentage points (Column 3) and the share of discounted units by 0.28 percentage points (Column 7). Alternatively, a variety supplied in a formulation-region with one standard deviation (across regions and formulations) higher HHI has 1.12 percentage point lower probability of giving discounts and, when it gives discounts, they are 0.92 percentage point lower.[24] These estimates are sizeable, considering that 9.2 percent of the observations offer quantity discounts and, among them, the discounted-to-undiscounted ratio is 15.8 percent.

5.1.2 Hypothesis 2: The effect of retail bargaining power depends on the elasticity of substitution

According to model prediction 2, the lower the elasticity of substitution in the formulation, the larger the effect of non-exerted bargaining power of regional retailers on discounts. Therefore, we test if higher concentration decreases discounts to a lesser extent when varieties in a formulation are more substitutable. In the model, the elasticity of substitution is the same for all varieties in the same formulation and is given by objective characteristics (exogenous in the model) that drive the consumer preferences. One can make a distinction using the state of matter of the drugs. The literature suggests that solid drugs (tablets) are more substitutable than liquid (syrups), which are in turn more substitutable than injections.[25]

To test Hypothesis 2 we estimate the following equation:

$$D_{ijt} = \alpha_1 HHI_{gjt} + \alpha_2 HHI_{gjt} \times Liq_g + \alpha_3 HHI_{gjt} \times Inj_g + \gamma y_{gjt} + \theta_{it} + \delta_{jt} + \epsilon_{ijt}, \qquad (20)$$

[24] The intention of this empirical study is not to identify the causal relationship between HHI and discounts, as in the model they both depend on the bargaining power of the variety. The OLS estimates in Table 3 might suffer from a simultaneity bias lead by the mechanism for which discounts are given to stimulate product demand. Since the products that receive more discounts are those with larger demand and prices (see Table A.3 in Appendix), higher discounts would increase HHI and this simultaneity bias is positive. The coefficients estimated in Table 3 can be considered upper bounds of the causal effect.

[25] Solid drugs (especially orally delivered ones) are the mostly used drugs due to characteristics such as dosage form variety, convenience, self-administration, high safety, and patient compliance (Raj et al., 2019). Liquid drugs have less varieties as they are only suitable for specific active principles. Injection drugs, instead, are often administered with the help of a healthcare professional. Antonecchia and Bhaskarabhatla (2022) estimate the elasticity of substitution separately for solid (9.6), liquid (6.2) and injectable (3.5) drugs. Solid drugs are almost 50 percent more substitutable than liquid drugs, that are in turn almost twice as substitutable as injectable drugs.

where Liq_g and Inj_g take value 1 when the drug formulation is liquid or injectable, respectively. Coefficient α_1 captures the effect of HHI when the drug formulation is solid. The coefficients α_2 and α_3 are also expected to be negative with $\alpha_3 < \alpha_2$.

Table 4 reports the results of the Hypothesis 2 test for both the extensive and intensive margin of discounts. HHI coefficient is negative in solid formulations (high substitutability) and larger in absolute terms in formulations with lower substitutability. This confirms the prediction stated in Hypothesis 2.

5.2 Robustness analysis

Robustness of H1 test. From Equation (17) we derive that the regional bargaining power of the medicine influences positively the number of varieties and, therefore, negatively the supplier concentration in the regional market. As an alternative to supplier HHI, the number of product varieties competing in the region-formulation would also proxy for the differences in regional bargaining power of retailers. In this case, however, a higher number of varieties implies stronger competition and we expect it to be positively correlated with the discounts. The results reported in Table A.4, in Appendix, confirm the model prediction.

As described in Section 3.4.5, the model can be extended to incorporate the presence of multiproduct suppliers, which we assume to have higher bargaining power compared to single-product suppliers. The extended Hypothesis 1 states that the bargaining power of regional retailers induces less quantity discounts for varieties belonging to suppliers that produce more varieties. We test this hypothesis considering two variables for multiproduct firms: (i) the number of varieties of the supplier in the formulation-region; (ii) the number of formulations of the supplier in the region. We distinguish the regions where the suppliers produce more varieties/formulations than the median supplier and indicate it with a dummy. Then we interact these variables with formulation-region HHI and report the results in Table A.5, in Appendix. The coefficients of the interacted variables are negative and significant, meaning that in regions where a supplier offers a higher number of varieties in the same formulation, a higher HHI is associated with lower discounts (Columns 1 and 3). Similarly, in regions where a supplier offers a higher number of formulations, a higher HHI is associated with lower discounts (Columns 2 and 4). This evidence supports the hypothesis of the model extension to multiproduct suppliers: producing more varieties increases supplier bargaining power which in turn decreases the effect of non-exerted

regional retail bargaining power on discounts. It is worth noting that the main effect of the number of varieties and formulations have both positive and significant coefficients. This confirms what we have observed above in the descriptive statistics: large firms (which produce more varieties and more formulations) give more discounts. Discounts might be given conditional on the number of varieties the retailers carry: "The more brands the retailer stocks, the higher its discount" (Shaffer, 1991). To support this evidence there is also the cost mechanism that we point out in Section 3.4.3, according to which in regions with larger sales (i.e., where suppliers offer more varieties and formulations) suppliers give more discounts.

Robustness of H2 test. In Equation (20) we use the state of matter of the drug to proxy for different elasticities of substitution across drugs. Another possibility is to distinguish between medications used to treat acute and chronic diseases using a classification developed by the AIOCD based on short-term versus long-term use and interact HHI with a dummy signalling drugs that treat an acute disease. Medicines for acute disease treatment are supposed to have a lower elasticity of substitution as they are used for immediate or emergency use. Medicines for chronic diseases are, instead, likely to be purchased in advance and the constant use of the drug generates incentives to search for cheaper alternative brands. This makes drug varieties for chronic diseases more substitutable with each other within the formulation. In Table A.6, in Appendix, columns 1 and 3, we report the results. A marginally higher concentration of suppliers reduces quantity discounts (both the probability and the amount) more for drugs treating acute diseases than for drugs treating chronic diseases. Another test of H2 can be conducted by distinguishing drugs that need a doctoral prescription to be purchased and those available over the counter. Medicines that require prescription are supposed to have lower elasticity of substitution as their consumption is mediated by the doctor that can recommend specific brands. Substitutability among over-the-counter medicines is, instead, more driven by consumer preferences. This makes over-the-counter drug varieties more substitutable with each other within the formulation. In Table A.6, in Appendix, columns 2 and 4, we report the results. A marginally higher concentration of suppliers reduces the amount of discounts more for prescription than for over-the-counter drugs. This is in line with the prediction of the model, where lower elasticity of substitution expands the effect of higher regional bargaining power of the retailers on discounts.[26]

[26]The elasticity of substitution σ in the model is assumed to be region invariant and not dependent on the number of competitors. It follows the other assumption that the number of competitors in all regions is large enough. Besides facilitating the aggregation of results across regions, these assumptions

6 Conclusion

In high-income countries, the buying power of third-party organizations that pool consumer risk and demand, such as health insurance providers, care providers, and large drugstores, is considered important for curbing the market power of pharmaceutical firms and controlling overall healthcare costs. This dynamic implies that price concessions from suppliers are, to some extent, passed on to consumers (e.g., Graf, 2014; Morgan et al., 2017). Nevertheless, the confidentiality of these concessions has led to policy challenges, including inflated "list prices" for uninsured and underinsured individuals in high-income countries (Morgan et al., 2013b). Since high-income countries use international drug prices to set domestic reference prices, these secretive concessions obscure the true costs, complicating the establishment of reference prices and exacerbating regional inequities (Morgan et al., 2013a). Furthermore, recent studies raise concerns about the anticompetitive use of such price concessions extracted by retailers with significant buyer power (Doyle and Inderst, 2007; Carlton and Israel, 2011; Grant, 2017).[27]

In low- and middle-income countries, where institutional payers are largely absent, insurance coverage is minimal, and consumers pay out of pocket for healthcare (Van Doorslaer et al., 2006; Leive and Xu, 2008; Shahrawat and Rao, 2012), retailer trade associations such as the AIOCD wield significant countervailing buyer power relative to the pharmaceutical firms. However, individual consumers have negligible bargaining power against pharmaceutical firms.

To our knowledge, no previous studies have examined the antecedents and consequences of pharmaceutical firms providing quantity discounts to retailers in low- and middle-income countries. Using novel data on wholesale and retail prices of medicines and quantity discounts given to retailers in the Indian pharmaceutical industry between March 2007 and June 2013, we document several key findings: pharmaceutical firms offer more quantity discounts to retailers in regions with higher bargaining power; these discounts lower supplier profits while increasing retailer profits; and they help maintain the stability of the retailer union's restrictive vertical trade practices. We discuss the advantages of quantity discounts over price discounts.

follow what we observe in the data, where the median number of suppliers in the region-formulation is 5. In Table A.7, in Appendix, we show that the relationship between HHI and discounts remains negative and significant also in region-formulations with a lower number of competitors than the median.

[27] Grant (2017) reports on the lawsuit filed by Pfizer claiming that Johnson & Johnson has used a system of rebates and discounts anticompetitively, limiting the sale of Pfizer's cheaper substitute.

Specifically, quantity discounts maintain uniform pricing within a variety while considering cost, demand, and bargaining conditions without an explicit financial transfer from the supplier to the retailer. We also extend the model to show that quantity discounts are lower for multiproduct firms, which can leverage their greater bargaining power. Although the use of quantity discounts to preserve resale price maintenance is generally successful, it is part of AIOCD's broader conduct, which is accused of being anti-competitive. Overall, our results indicate that confidential quantity discounts are likely detrimental to consumers, as these discounts are not passed on but instead increase retailer profitability.

As usual, our study is not without limitations. First, our theoretical model assumes simultaneous bargaining across varieties without renegotiation. Though this has enabled closed-form solution for prices, this may be an oversimplification of reality where outside options when negotiations break down depend on the number of alternative varieties. We leave a more rigorous theoretical model for further research. From a conceptual point of view, coordinated bargaining where regions matter and lobby at the retail association lead to a model-implied bargaining power of the association which is weaker than the bargaining power of the strongest region. This implication is fundamental for our research and is consistent with our findings. A more detailed examination in how the negotiations work out in practice is beyond the scope of our paper. Most importantly, however, we do not have data at the individual retailer pharmacy level, limiting our analysis of the individual retailer's buyer power. Specifically, we cannot estimate the impact of buyer size and local market power on the level of quantity discounts. While this limitation may be reasonable in our context due to the nature of entry and competition among drugstores in India, it could be a significant factor in other contexts.

References

Alé-Chilet, J. and Atal, J. P. (2020). Trade associations and collusion among many agents: evidence from physicians. *The RAND Journal of Economics*, 51(4):1197–1221.

Allain, M.-L., Avignon, R., and Chambolle, C. (2020). Purchasing alliances and product variety. *International Journal of Industrial Organization*, 73:102641.

Allende, C., Atal, J. P., Carril, R., Cuesta, J. I., and González-Lira, A. (2024). Drivers of public procurement prices: Evidence from pharmaceutical markets. *International Journal of Industrial Organization*, page 103086.

Antonecchia, G. and Bhaskarabhatla, A. (2022). How do wholesale prices and retail markups affect market shares in the pharmaceutical industry? In *Productivity and Strategies of Multiproduct Firms*, number 801, pages 53–88. Erasmus School of Economics dissertation series.

Antonecchia, G. and Bhaskarabhatla, A. (2023). How firms compete when they set identical prices: Nonprice strategies in the indian biscuit industry. *Journal of Economics & Management Strategy*, 32(4):733–756.

Bhaskarabhatla, A. (2020). Maximum resale price maintenance and retailer cartel profits: evidence from the indian pharmaceutical industry. *Antitrust Law Journal*, 83(1):41–74.

Bhaskarabhatla, A., Chatterjee, C., and Karreman, B. (2016). Hit where it hurts: Cartel policing using targeted sales and supply embargoes. *The Journal of Law and Economics*, 59(4):805–846.

Bhuller, M., Moene, K. O., Mogstad, M., and Vestad, O. L. (2022). Facts and fantasies about wage setting and collective bargaining. *Journal of Economic Perspectives*, 36(4):29–52.

Carlton, D. W. and Israel, M. (2011). Proper treatment of buyer power in merger review. *Review of Industrial Organization*, 39(1):127–136.

CCI (2013). Aiocd vs. m/s santuka associates pvt. ltd. *Competition Commission of India. https://www.cci.gov.in/sites/default/files/202011*$_0$*.pdf*.

Chaganti, S. R. (2006). *Pharmaceutical marketing in India*. Excel Books India.

Conlon, C. T. and Mortimer, J. H. (2021). Efficiency and foreclosure effects of vertical rebates: Empirical evidence. *Journal of Political Economy*, 129(12):3357–3404.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643–685.

Dafny, L., Ody, C., and Schmitt, M. (2017). When discounts raise costs: the effect of copay coupons on generic utilization. *American Economic Journal: Economic Policy*, 9(2):91–123.

Daruich, D. and Kozlowski, J. (2023). Macroeconomic implications of uniform pricing. *American Economic Journal: Macroeconomics*, 15(3):64–108.

DellaVigna, S. and Gentzkow, M. (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American economic review*, 67(3):297–308.

Dobson, P. W. and Waterson, M. (1997). Countervailing power and consumer prices. *The Economic Journal*, 107(441):418–430.

Dobson, P. W. and Waterson, M. (2007). The competition effects of industry-wide vertical price fixing in bilateral oligopoly. *International Journal of Industrial Organization*, 25(5):935–962.

Doyle, C. and Inderst, R. (2007). Some economics on the treatment of buyer power in antitrust. *European Competition Law Review*, 28(3):210.

Dubois, P. and Sæthre, M. (2020). On the effect of parallel trade on manufacturers' and retailers' profits in the pharmaceutical sector. *Econometrica*, 88(6):2503–2545.

Eckel, C. and Egger, H. (2009). Wage bargaining and multinational firms. *Journal of International Economics*, 77(2):206–214.

Ellison, S. F. and Snyder, C. M. (2010). Countervailing power in wholesale pharmaceuticals. *The Journal of Industrial Economics*, 58(1):32–53.

Galbraith, J. K. (1952). *American capitalism: The concept of countervailing power*. Boston: Houghton Miffin.

Gaudin, G. (2018). Vertical bargaining and retail competition: What drives countervailing power? *Economic Journal*, 128(September):2380–2413.

Genesove, D. and Mullin, W. P. (2001). Rules, communication, and collusion: Narrative evidence from the sugar institute case. *American Economic Review*, 91(3):379–398.

Genesove, D. and Mullin, W. P. (2007). *3. The Sugar Institute Learns to Organize Information Exchange*. University of Chicago Press.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105(1):172–203.

Graf, J. (2014). The effects of rebate contracts on the health care system. *The European Journal of Health Economics*, 15(5):477–487.

Grant, C. (2017). The murky world of drug rebates gets airing in lawsuit. *Wall Street Journal, https://www.wsj.com/articles/the-murky-world-of-drug-rebates-gets-airing-in-lawsuit-1505929432*.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–177.

Herper, M. (2012). Inside the secret world of drug company rebates. *Forbes, https://www.forbes.com/sites/matthewherper/2012/05/10/why-astrazeneca-gives-insurers-60-discounts-on-nexiums-list-price/?sh=7d64ca5d2b25.*

Ho, K. and Lee, R. S. (2017). Insurer competition in health care markets. *Econometrica*, 85(2):379–417.

Holden, S. (1988). Local and central wage bargaining. *Scandinavian Journal of Economics*, 90(1):93–99.

Inderst, R. and Mazzarotto, N. (2008). Buyer power in distribution. In *ABA Antitrust Section Handbook, Issues in Competition Law and Policy*, volume 3, pages 1953–1978. ABA.

Lakdawalla, D. N. (2018). Economics of the pharmaceutical industry. *Journal of Economic Literature*, 56(2):397–449.

Leive, A. and Xu, K. (2008). Coping with out-of-pocket health payments: empirical evidence from 15 african countries. *Bulletin of the World Health Organization*, 86:849–856C.

Levenstein, M. C. and Suslow, V. Y. (2006). What determines cartel success? *Journal of economic literature*, 44(1):43–95.

Mattingly, T. J., Levy, J. F., Slejko, J. F., Onwudiwe, N. C., and Perfetto, E. M. (2018). Estimating drug costs: how do manufacturer net prices compare with other common us price references? *PharmacoEconomics*, 36(9):1093–1099.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica*, 71(6):1695–1725.

Montagna, C. (1995). Monopolistic competition with firm-specific costs. *Oxford Economic Papers*, 47(2):318–328.

Morgan, S., Daw, J., and Thomson, P. (2013a). International best practices for negotiating 'reimbursement contracts' with price rebates from pharmaceutical companies. *Health Affairs*, 32(4):771–777.

Morgan, S. G., Thomson, P. A., Daw, J. R., and Friesen, M. K. (2013b). Canadian policy makers' views on pharmaceutical reimbursement contracts involving confidential discounts from drug manufacturers. *Health Policy*, 112(3):248–254.

Morgan, S. G., Vogler, S., and Wagner, A. K. (2017). Payers' experiences with confidential pharmaceutical price discounts: a survey of public and statutory health systems in north america, europe, and australasia. *Health Policy*, 121(4):354–362.

Raj, G. M., Raveendran, R., et al. (2019). *Introduction to basics of pharmacology and toxicology: Volume 1: General and molecular pharmacology: Principles of drug action.* Springer.

Rao, P. K. and Sastry, K. (1989). Restrictive trade practices policy in india. *The Journal of Industrial Economics*, pages 427–435.

Scott Morton, F. and Kyle, M. (2011). Markets for pharmaceutical products. In *Handbook of Health Economics*, volume 2, pages 763–823. Elsevier.

Shaffer, G. (1991). Capturing strategic rent: Full-line forcing, brand discounts, aggregate rebates, and maximum resale price maintenance. *The Journal of Industrial Economics*, pages 557–575.

Shahrawat, R. and Rao, K. D. (2012). Insured yet vulnerable: out-of-pocket payments and india's poor. *Health Policy and Planning*, 27(3):213–221.

Singh, U. C. (1981). Aiocd demand increase trade margins to wholesalers and re-tailers of pharmaceutical product. *India Today*.

Singh, U. C. (1984). Bombay retailers demand raise in margins by manufacturers,threaten series of boycott. *India Today*.

Stroebel, J. and Vavra, J. (2019). House prices, local demand, and retail prices. *Journal of Political Economy*, 127(3):1391–1436.

Van Doorslaer, E., O'Donnell, O., Rannan-Eliya, R. P., Somanathan, A., Adhikari, S. R., Garg, C. C., Harbianto, D., Herrin, A. N., Huq, M. N., Ibragimova, S., et al. (2006). Effect of payments for health care on poverty estimates in 11 countries in asia: an analysis of household survey data. *The Lancet*, 368(9544):1357–1364.

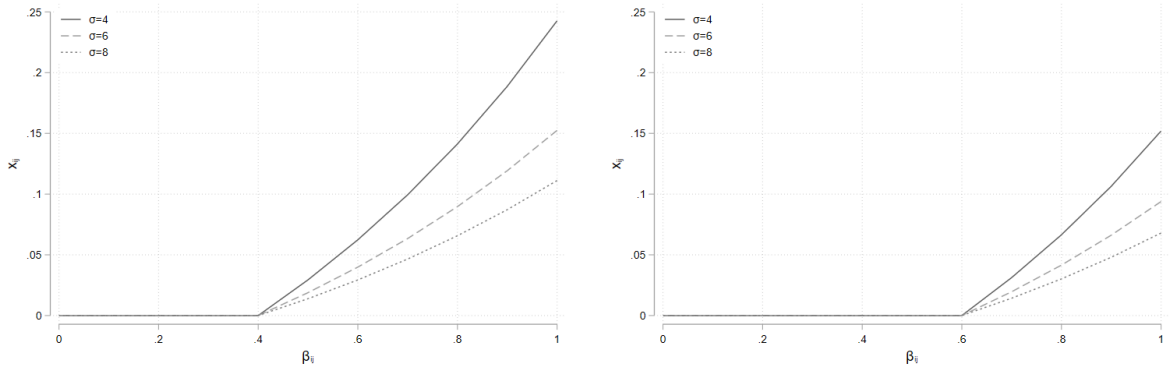
Verma, D. (1981). Regulation of trade association activities. *Economic and Political Weekly*, pages M59–M68.

Wu, V. Y. (2009). Managed care's price bargaining with hospitals. *Journal of Health Economics*, 28(2):350–360.

Figures

Figure 1: Discounts for different elasticities of substition (σ)

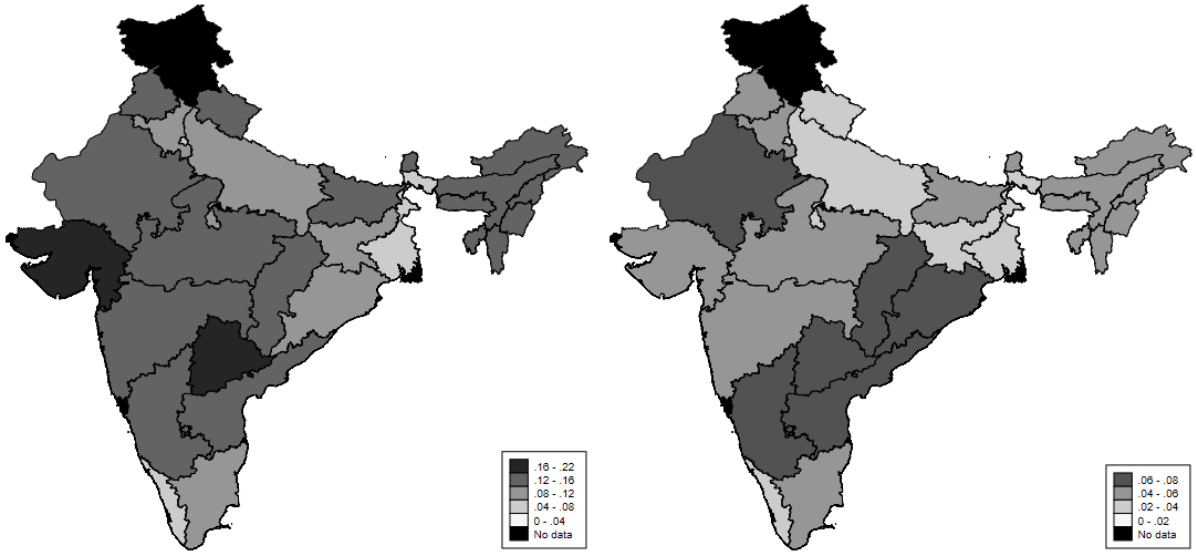
(A) if $\beta_i^A = 0.4$ (B) if $\beta_i^A = 0.6$



Notes: Quantity discounts offered by the suppliers for different values of β_{ij}, β_i and σ

Figure 2: Quantity discounts across regions

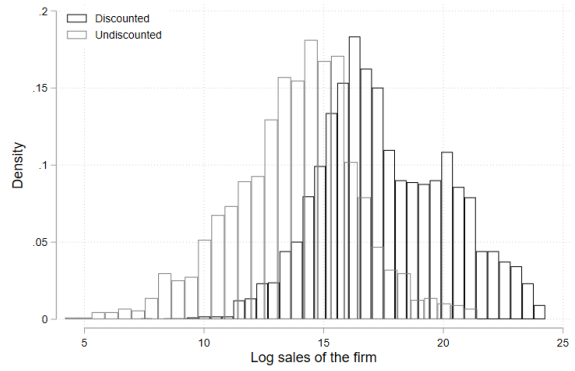
(A) Share of varieties discounted (B) Discounted-to-undiscounted ratio



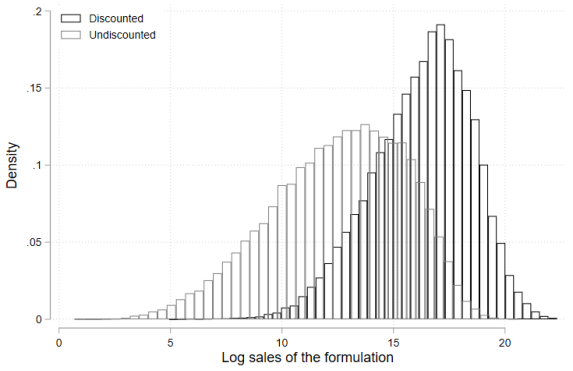
Notes: *Share of varieties discounted* is the share of varieties that give discounts. *Discounted-to-undiscounted ratio*, measured as the ratio between the units given as quantity discounts and those purchased at wholesale price by the retailers. Both variables are aggregated across formulation at the region-year level for year 2012. The area of Maharasthra is colored considering the simple average of the variable for the regions of Mumbai, Vidarbha and Marathwada. The area of West Bengal is colored considering the simple average of the variable for the regions of Kolkata and West Bengal Rest. This figure is based on AIOCD data of year 2012.

Figure 3: Firms and formulations that give quantity discounts by size

(A) Firm size



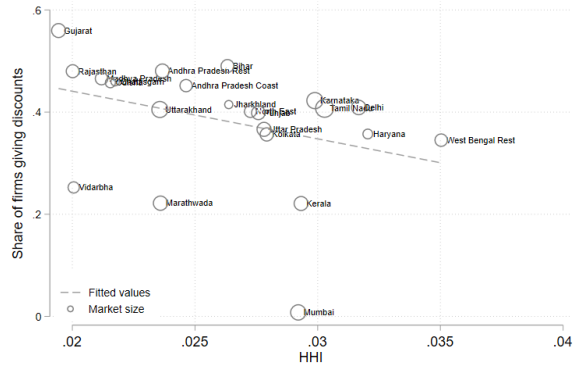
(B) Formulation size



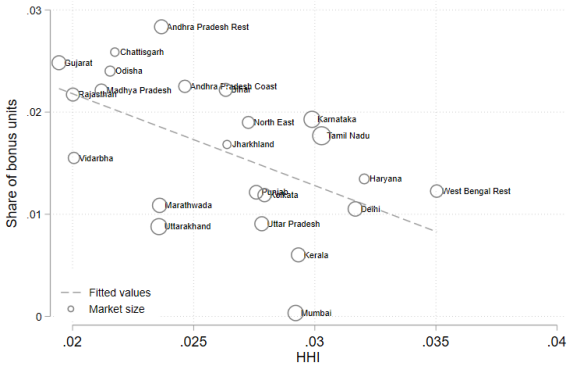
Notes: Panel (A): Size distribution of the firms giving quantity discounts (*Discounted*) and of the firms that do not (*Undiscounted*). Panel (B): Size distribution of the formulations giving quantity discounts (*Discounted*) and of the formulations that do not (*Undiscounted*). This figure is based on AIOCD data from April 2007 to June 2013.

Figure 4: Quantity discounts and HHI across regions

(A) Extensive margin



(B) Intensive margin



Notes: *Share of firms giving discounts* is the share of firms that give discounted varieties. *Discounted-to-undiscounted ratio* is measured as the ratio between the units given as quantity discounts and those purchased at wholesale price by the retailers. *HHI* is the Herfindahl-Hirschman Index. All variables are first aggregated at the region-month level and later averaged across the months. This figure is based on AIOCD data from April 2007 to June 2013.

Tables

Table 1: Variables and definitions

	Number	Definition
Regions	23	Andhra Pradesh Coastal, Andhra Pradesh Rest, Bihar, Chattisgarh, Delhi, Gujarat, Haryana, Jharkhand, Karnataka, Kerala, Kolkata, Madhya Pradesh, Marathwada, Mumbai, North East, Odisha, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh East, Uttarakhand & Uttar Pradesh West, Vidarbha, West Bengal Rest
Suppliers	681	Pharmaceutical manufacturers
Formulations	18,079	Active ingredient - Dosage form - Strength
Varieties	85,384	Stock Keeping Unit
Months	76	March 2007 - June 2013
Observations	35,347,564	Variety-region-month level

Notes: This table is based on AIOCD data from March 2007 to June 2013.

Table 2: Quantity discounts at various levels of analysis

Level	Observations	Extensive margin	Intensive margin
Variety-region-month	35,347,564	9.2	15.8
Variety-month	3,501,567	19.7	7.0
Formulation-region-month	12,128,758	14.7	9.5
Formulation-month	880,630	30.7	4.5
Firm-region-month	461,419	38.6	4.3
Firm-month	45,401	61.8	3.8

Notes: The *Extensive margin* is the share of varieties that give discounts. The *Intensive margin* is measured as the ratio between the units given as quantity discounts and those purchased at wholesale price by the retailers. This table is based on AIOCD data from March 2007 to June 2013.

Table 3: Testing Hypothesis 1: Supplier concentration and quantity discounts

	Dummy discounted				Discounted-to-undiscounted ratio			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Formulation-region HHI	-0.060***	-0.033***	-0.034***	-0.016***	-0.033***	-0.024***	-0.028***	-0.012***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.004)	(0.004)	(0.004)	(0.002)
Log sales			0.026***	0.010***			0.008***	-0.000
			(0.000)	(0.000)			(0.001)	(0.001)
Variety × Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region × Month FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Variety × Region FE	No	No	No	Yes	No	No	No	Yes
R-squared	0.511	0.531	0.538	0.732	0.532	0.540	0.541	0.677
Observations	34618025	34618025	34615256	34560592	8096304	8096304	8096095	8095803

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. The dependent variables are: *Dummy discounted* that takes value one when the variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio* is measured as the ratio between the units of variety given as quantity discounts and those purchased at wholesale price by the retailers. *Formulation-region HHI* is the formulation-region-month Herfindahl-Hirschman Index. *Log sales* is the logarithm of the formulation-region-month sales. The estimates reported in Column 1-4 are conducted on the full sample of variety-region-month observations; the estimates reported in Column 5-8 are conducted on the sample of variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.

Table 4: Testing Hypothesis 2: The role of elasticity of substitution

	Dummy discounted	Discounted-to-undiscounted ratio
	(1)	(2)
Formulation-region HHI	-0.021***	-0.021***
	(0.003)	(0.007)
Formulation-firm HHI × Liquid	-0.020***	0.011
	(0.006)	(0.007)
Formulation-firm HHI × Injection	-0.059***	-0.051***
	(0.008)	(0.014)
Variety × Month FE	Yes	Yes
Region × Month FE	Yes	Yes
R-squared	0.541	0.525
Observations	28586789	6683370

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. The dependent variables are: *Dummy discounted* that takes value one when the Variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio* is measured as the ratio between the units of Variety given as quantity discounts and those purchased at wholesale price by the retailers. *Formulation-region HHI* is the formulation-region-month Herfindahl-Hirschman Index. The estimates in both columns control for the formulation-region-month sales and the *Variety* × *month FE* control for the dummies *Liquid* and $Injection$. All estimates control for the logarithm of the formulation-region-month sales. The estimates reported in Column 1 are conducted on the full sample of variety-region-month observations; the estimates reported in Column 2 are conducted on the sample of Variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.

A Additional figures and tables

Table A.1: Supplier concentration and quantity discounts for various definitions of product market

	Share of products discounted			Discounted-to-undiscounted ratio		
	(1)	(2)	(3)	(4)	(5)	(6)
Formulation-region HHI	-0.019***			-0.018***		
	(0.005)			(0.003)		
Active ingredient-Dosage form HHI		-0.012**			-0.017***	
		(0.006)			(0.003)	
Active ingredient HHI			-0.023***			-0.019***
			(0.007)			(0.004)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.018	0.022	0.028	0.006	0.009	0.011
Observations	12128757	6507087	3189615	12127353	6505730	3189380

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation-region level. *Share of varieties discounted* is the share of varieties that give discounts within the market considered. *Discounted-to-undiscounted ratio* is the ratio between the units given as quantity discounts and those purchased at wholesale price by the retailers within the market considered. *HHI* is the market-region-month Herfindahl-Hirschman Index. The estimates are based on three samples with variety-region-month observation aggregated at the formulation-region-month level (Column 1 and 4), active ingredient-region-month level (Column 2 and 5), and active ingredient-dosage form-region-month level (Column 3 and 6). This table is based on AIOCD data from March 2007 to June 2013.

Table A.2: Distribution of prices and concentration

	Mean	Standard Deviation		
		Across regions Across formulations	Across regions Within formulation	
Formulation-region HHI	0.52	0.33	0.11	
		Across varieties	Across varieties	Within variety
Wholesale price	96.03	631.62	10.19	0.00
Retail price	121.85	791.73	12.79	0.00

Notes: Mean and standard deviation of the distributions of HHI, wholesale and retail prices. Price observations are at the variety-region-month level. HHI is calculated at the formulation-region-month level. *Mean* is the mean across all formulations, regions and months. *Standard deviation* is the average of all standard deviations of the distributions specified. This table is based on AIOCD data from March 2007 to June 2013.

Table A.3: Quantity discounts, prices and demand

	Log wholesale price		Log retail price		Log undiscounted units	
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy discounted	0.045***		0.027***		2.225***	
	(0.003)		(0.002)		(0.008)	
Discounted-to-undiscounted ratio		0.024***		0.017***		0.942***
		(0.006)		(0.006)		(0.096)
Region × Formulation × Month FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.816	0.859	0.821	0.860	0.358	0.386
Observations	28343873	6111974	28343992	6110667	28346765	6112152

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. *Log wholesale price* and *Log retail price* are measured as the logarithm of rupees per unit of variety. *Log undiscounted units* is measured as the logarithm of the units of variety purchased at wholesale price by the retailers. *Dummy discounted*, that takes value one when the variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio*, that is the ratio between the units of variety given as quantity discounts and those purchased at wholesale price by the retailers. In Column 1, 3 and 5 estimates are conducted on the full sample of variety-region-month observations. In In Column 2, 4 and 6 estimates are conducted on the subsample of variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.

Table A.4: Robustness analysis H1: Number of competitors and quantity discounts

	Dummy discounted		Discounted-to-undiscounted ratio	
	(1)	(2)	(3)	(4)
Log number of varieties	0.031***	0.003***	0.013***	0.003**
	(0.002)	(0.001)	(0.002)	(0.001)
Variety × Month FE	Yes	Yes	Yes	Yes
Region × Month FE	Yes	Yes	Yes	Yes
Variety × Region FE	No	Yes	No	Yes
R-squared	0.532	0.732	0.540	0.677
Observations	34618025	34563379	8096304	8096012

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. The dependent variables are: *Dummy discounted* that takes value one when the variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio* is measured as the ratio between the units of variety given as quantity discounts and those purchased at wholesale price by the retailers. *Log number of varieties* is the logarithm of the number of varieties in the formulation-region-month. The estimates in all columns control for the formulation-region-month sales. The estimates reported in Column 1-2 are conducted on the full sample of variety-region-month observations; the estimates reported in Column 3-4 are conducted on the sample of variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.

Table A.5: Robustness analysis H1: Multiproduct suppliers

	Dummy discounted		Discounted-to-undiscounted ratio	
	(1)	(2)	(3)	(4)
Formulation-region HHI	-0.033***	-0.024***	-0.028***	-0.024***
	(0.002)	(0.002)	(0.004)	(0.004)
HHI × High N varieties within firm-region-formulation	-0.022***		-0.015***	
	(0.005)		(0.006)	
HHI × High N formulations within firm-region		-0.020***		-0.008***
		(0.002)		(0.002)
High N varieties within firm-region-formulation	0.014***		0.008***	
	(0.003)		(0.002)	
High N formulations within firm-region		0.026***		0.011***
		(0.001)		(0.001)
Variety × Month FE	Yes	Yes	Yes	Yes
Region × Month FE	Yes	Yes	Yes	Yes
R-squared	0.538	0.538	0.541	0.541
Observations	34615256	34615256	8096095	8096095

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. The dependent variables are: *Dummy discounted* that takes value one when the variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio* is measured as the ratio between the units of variety given as quantity discounts and those purchased at wholesale price by the retailers. *Formulation-region HHI* is the formulation-region-month Herfindahl-Hirschman Index. *High N varieties within firm-region-formulation* is a dummy taking value 1 when in that region-formulation the supplier provides a number of varieties higher than its median across regions for the same formulation. *High N formulations within firm-region* is a dummy taking value 1 when in that region the supplier provides a number of formulations higher than its median across regions. The estimates in all columns control for the formulation-region-month sales. The estimates reported in Columns 1-2 are conducted on the full sample of variety-region-month observations; the estimates reported in Columns 3-4 are conducted on the sample of variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.

Table A.6: Robustness analysis H2: Acute Vs Chronic treatment medicines

	Dummy discounted		Discounted-to-undiscounted ratio	
	(1)	(2)	(3)	(4)
Formulation-region HHI	0.011***	-0.035***	-0.009***	-0.018***
	(0.002)	(0.003)	(0.003)	(0.002)
HHI × Acute treatment	-0.059***		-0.021***	
	(0.004)		(0.005)	
HHI × Prescription		0.003		-0.022***
		(0.004)		(0.008)
Variety × Month FE	Yes	Yes	Yes	Yes
Region × Month FE	Yes	Yes	Yes	Yes
R-squared	0.538	0.538	0.541	0.541
Observations	34615256	34615256	8096095	8096095

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. The dependent variables are: *Dummy discounted* that takes value one when the variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio* is measured as the ratio between the units of variety given as quantity discounts and those purchased at wholesale price by the retailers. *Formulation-region HHI* is the formulation-region-month Herfindahl-Hirschman Index. *Acute treatment* is a dummy taking value 1 when the drug is used for treating acute diseases (as opposed to chronic). *Prescription* is a dummy taking value 1 when the drug needs a doctoral prescription to be purchased. The estimates in both columns control for the formulation-region-month sales. The estimates reported in Column 1 are conducted on the full sample of variety-region-month observations; the estimates reported in Column 2 are conducted on the sample of variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.

Table A.7: Robustness analysis H2: elasticity of substitution and number of suppliers

	Dummy discounted	Discounted-to-undiscounted ratio
	(1)	(2)
Formulation-region HHI	-0.027***	-0.025***
	(0.002)	(0.004)
Formulation-firm HHI × High N suppliers	-0.020***	-0.010***
	(0.003)	(0.003)
High N suppliers	0.015***	0.005***
	(0.002)	(0.001)
Variety × Month FE	Yes	Yes
Region × Month FE	Yes	Yes
R-squared	0.538	0.541
Observations	34615256	8096095

Notes: OLS estimates, standard errors (in parentheses) clustered at the formulation level. The dependent variables are: *Dummy discounted* that takes value one when the variety gives quantity discounts and zero otherwise; *Discounted-to-undiscounted ratio* is measured as the ratio between the units of variety given as quantity discounts and those purchased at wholesale price by the retailers. *Formulation-region HHI* is the formulation-region-month Herfindahl-Hirschman Index. *High N suppliers* is a dummy taking value 1 when in that region the number of suppliers is higher than the median of the formulation across regions. The estimates in both columns control for the formulation-region-month sales. The estimates reported in Column 1 are conducted on the full sample of variety-region-month observations; the estimates reported in Column 2 are conducted on the sample of variety-region-month observations that give discounts at least once in the period considered. This table is based on AIOCD data from March 2007 to June 2013.