Take the Goods and Run: Contracting Frictions and Market Power in Supply Chains*

Felipe Brugués†

This Version: March 11, 2025
First version: October 16, 2020

Abstract

This paper studies the efficiency of self-enforced relational agreements, a common solution to contracting frictions, when sellers have market power and contracts cannot be externally enforced. To this end, I develop a dynamic contracting model with limited enforcement in which buyers can default on their trade-credit debt and estimate it using a novel dataset from the Ecuadorian manufacturing supply-chain. The key empirical finding is that bilateral trade is inefficiently low in early periods of the relationship, but converges toward efficiency over time, despite sellers' market power. Counterfactual simulations imply that both market power and enforcement contribute to inefficiencies in trade.

*Thanks to Jesse Shapiro, Rafael La Porta, and Neil Thakral for their continued guidance and support, as well as Lorenzo Aldeco, Dan Bjorkegren, Joaquin Blaum, Javier Brugués, Pedro Dal Bó, Mert Demirer, Rebecca De Simone, Andrew Foster, John Friedman, Samuele Giambra, Stefan Hut, Diego Jimenez, Ken Kikkawa, Amanda Loyola, Teddy Mekonnen, Bobby Pakzad-Hurson, Elena Pastorino, Diego Ramos-Toro, Pau Roldan-Blanco, Bryce Steinberg, Marcel Peruffo, Arthur Taburet, Julia Tanndal, Marta Troya Martinez (discussant), Matt Turner, and members of audiences from multiple seminars for helpful conversations and comments. I also thank the editor Arnoud Costinot and the referees for their detailed comments. I thank Cristian Chicaiza, Alexiss Mejia, Sebastián Carvajal, and Maria Eugenia Andrade at the Servicios of Rentas Internas of Ecuador for help with the data. Carolina Alvarez provided excellent research assistance. I gratefully acknowledge financial support from the Bank of Spain and the Nelson Center for Entrepreneurship at Brown University.

†The Business School, Instituto Tecnológico Autónomo de México. E-mail: felipe.brugues@itam.mx

When courts cannot enforce contracts, trading partners often resort to long-term relational contracts, sustained through repeated interactions, to ease frictions and constrain opportunistic behavior (Johnson et al., 2002). As weak contract enforcement is a common feature of developing economies, relational agreements are highly relevant inter-firm organizational structures. Understanding the efficiency of these informal agreements is essential for policy-makers in developing countries, as they frequently have to make trade-offs regarding where to focus their reform efforts.

The traditional view sees contracting frictions as a hindrance that distorts productive decisions (La Porta et al., 1997; Nunn, 2007), implying that, as a standard solution, relational contracts may be inefficient. However, it is noteworthy that the same economies where enforcement constraints are likely to be a significant factor may also encounter additional frictions, such as high market concentration, making them second-best environments (Rodrik, 2008). In the presence of seller market power, weak enforcement may increase the buyer's relative bargaining power, thereby limiting downstream distortions while improving the efficiency of a relationship as opposed to a perfect enforcement world (Genicot and Ray, 2006).[1] Thus, the efficiency of relational agreements remains unclear.

This paper uses theory and data to quantify the static (period-by-period) efficiency of self-enforced long-term relationships in the presence of seller market power and limited external enforcement of contracts. I develop a novel long-term contracting model where 1) the seller can price discriminate across buyers and time, and 2) the buyer can act opportunistically and simply *take the goods and run* whenever the delivery of the goods occurs before payment. Without access to external enforcement, the seller uses the value of the relationship itself to discipline the buyer's behavior. The modeling framework is applied to examine self-enforced relationships in the manufacturing supply chain in Ecuador, a middle-income country with slow commercial courts and concentrated sectors.

The paper has two novel empirical contributions. First, by utilizing a structural econometric model, it provides the first empirical evidence regarding the efficiency evolution of long-term trade relationships. The findings demonstrate that relationships tend to be highly inefficient at the early stages, but over time, such inefficiencies diminish, indicating the crucial role of repeated informal agreements in creating surplus. Second, the study examines the counterfactual scenario of implementing best-practice institutions (e.g., eliminating contracting frictions) and

[1] Throughout the paper, the working definition of **seller market power** is the *seller's ability to price discriminate with prices above marginal costs*. This definition encapsulates the common one referring to the ability of sellers to price above marginal costs often used in the economics literature (e.g., De Loecker et al., 2020) and in economic law (e.g., Kaplow, 2016). Moreover, the common definition of market power is seen as a necessary condition for price discrimination (Varian, 1989; Stole, 2007). I do note, however, that in general price discrimination, relative to profit-maximizing uniform pricing, can be welfare-enhancing or welfare-decreasing (Varian, 1989). In the specific case of third-degree price discrimination (non-linear pricing or wholesale quantity discounts), price discrimination can be also welfare-increasing or welfare-decreasing relative to profit-maximizing uniform pricing (Katz, 1984; Varian, 1985). Furthermore, except the case of perfect price discrimination, market power (both in uniform prices or with price discrimination) generates quantity distortions relative to a competitive benchmark.

finds an intertemporal trade-off. In the short term, the implementation of best-practice institutions leads to an increase in welfare. However, in the medium and long term, such institutional changes are found to result in welfare losses when compared to the observed second-best equilibrium. In contrast, efficiency improves when all modeled frictions are addressed simultaneously.

I start by documenting six fundamental patterns that provide the basis for the key elements of the model. First, it is observed that most trade takes place through repeated relationships. Second, the vendor finances a substantial share of transactions using trade-credit, even in new relationships, indicating that the seller bears the risk of the transaction. Third, as relationships age, they exhibit growth in both quantity and value. Fourth, sellers offer considerable quantity discounts, with a 10% increase in quantity corresponding to a 2% decrease in unit price. Fifth, accounting for quantity discounts, older buyers receive up to a 3% discount compared to new buyers as the relationship matures. These discounts are observed only in cases where buyers use trade-credit as opposed to paying the full order amount upfront. Finally, the survival probability of relationships is observed to increase in quantity and as relationships mature. These patterns provide valuable insights into the nature of long-term relationships in the manufacturing supply chain in Ecuador, which are used to build the theoretical model and to inform the empirical analysis.

Standard models in the literature (such as efficiency gains, learning, demand assurance, or supply-side enforcement issues) are not able to capture all of these patterns under realistic assumptions. For that reason, to account for these patterns and assess the efficiency of relationships over time, I develop a novel dynamic contracting model by embedding a non-linear pricing model with heterogeneous participation constraints (Jullien, 2000; Attanasio and Pastorino, 2020) into an infinitely repeated game with limited enforcement (Martimort et al., 2017; Pavoni et al., 2018; Marcet and Marimon, 2019). In the model, sellers and buyers with private heterogeneous demand meet randomly and have the opportunity to engage in repeated trade. The seller has all the bargaining power and proposes a dynamic contract of prices and quantities, for which they have commitment. Consistent with the data, the seller in the model finances all the transactions using trade-credit. Buyer heterogeneity provides incentives to price discriminate, so the seller offers menus of quantities and prices that satisfy *incentive compatibility* and induce revelation of the buyer asymmetric information.

Crucially, the buyer cannot commit to paying their debts and is subject to forward-looking *limited enforcement* constraints. The future stream of benefits created by the relationship for the buyer must be large enough to secure the payment. To prevent a *take the goods and run* scenario, the seller must share a greater amount of surplus, through greater levels of future net returns, than otherwise. Thus, enforcement constraints could dynamically act against the seller's profit-maximizing incentives to distort trade downward through inefficiently low quantities. Matching the empirical picture described above, the optimal dynamic menu of quantities and prices in a setting with limited enforcement features *backloading*: both the total surplus generated by the relationship and the net return enjoyed by the buyer increase over time.

To determine the optimal quantity allocations in this setting, I use a recursive Lagrangian approach (Pavoni et al., 2018; Marcet and Marimon, 2019), which characterizes the optimal dynamic contract in terms of *past* and *present* limited enforcement Lagrange multipliers (LE multipliers). The present LE multipliers capture the current limited enforcement constraints, while past LE multipliers account for promises made in the past to prevent default and serve as promise-keeping constraints. In equilibrium, the optimal quantity allocations are then determined by a *modified virtual surplus*, which takes into account the standard informational rents due to incentive compatibility, as well as the shadow costs of binding enforcement constraints.

The paper proposes an econometric model that is directly derived from the theoretical model and shows that the parameters of the model can be identified using cross-sectional data on prices, quantities, age of relationships, and marginal costs for one seller. The model relies on the seller's optimality conditions and the buyer's dynamic first-order conditions for incentive compatibility (as in the static results of Luo et al., 2018 and Attanasio and Pastorino, 2020) to identify the dynamic effects of limited enforcement on trade. The identification intuition is twofold. First, the seller offers prices and quantities that induce the revelaxtion of information about buyers' types and discriminate across them. This implies that price and quantity variation across buyers is a signal of their underlying types. Second, the degree of trade distortion in quantities relative to the efficient outcome provides information on whether current or past enforcement concerns are constraining the trade relationship. By examining the difference between marginal prices and marginal costs, which indicates the presence of downward and upward distortions, we can identify the extent of additional distortions due to limited enforcement.

I estimate the model using three administrative databases collected by the Ecuadorian government for tax purposes that provide empirical analogs to the objects in the theoretical model. I obtain pair-specific unit prices and quantities using a new electronic invoice database that contains all domestic sales for 49 manufacturing firms in the textile, pharmaceutical, and cement-product sectors for 2016-2017, each with a large number of buyers each year (median of 600). The age of relationships is inferred through the universe of firm-to-firm VAT database, which tracks the total volume of bilateral trade from 2008-2015. Lastly, a measure of seller's costs comes from information on total variable costs (i.e., intermediate inputs expenditure and labor wages) contained in usual financial statements reported to the tax authority.

The estimated model fits the data well, and the estimation results reveals that enforcement concerns are relevant throughout the life-cycle of a relationship. Specifically, almost all new relationships have binding enforcement constraints, meaning that if the seller where to increase current prices without a corresponding future decrease in prices or increase in quantities, the buyer would default and exit the relationship. As relationships age, these constraints are relaxed, reflecting the increase in quantities coming from past promises made by the seller.

Using the estimated model parameters, I evaluate the efficiency of transactions at any given point and examine the division of surplus. My findings indicate that new relationships operate at approximately 30% of the optimal (i.e., frictionless) level, but efficiency increases as relationships age. Relationships lasting five years or more can achieve efficiencies upwards of 80%. In the aggregate, my analysis reveals that sellers heavily distort quantities early on. Specifically, only 5% of suppliers achieve levels of aggregate output that are indistinguishable from efficient output when dealing with new buyers, whereas 84% of sellers achieve long-term aggregate output levels that cannot be distinguished from efficient levels. Remarkably, these patterns hold for each industry studied, talking to the generality of the result. As for the division of surplus, I find that sellers capture the majority (around 80%) of the generated surplus, although some buyers may capture up to 30% of the total surplus.

The paper proceeds to investigate counterfactual scenarios that have surprising implications. First, the analysis shows that addressing enforcement constraints alone, without addressing market power, can lead to higher surplus in the short term, but result in a lower total surplus in the medium and long term. Similarly, only addressing market power leads to substantial welfare losses across different types and time periods. These findings are consistent with the *theory of second-best* (Lipsey and Lancaster, 1956), which suggests that in the presence of one friction, the effect on welfare of removing one friction alone is uncertain. In this particular case, each friction serves to counterbalance the other. Second, the paper explores the effects of addressing both frictions simultaneously. The results indicate that most relationships achieve a higher total surplus and lower surplus for the seller when both frictions are addressed together. Overall, these counterfactual analyses underscore the significance of recognizing the interplay between various frictions in markets. Simply addressing one friction in isolation may not produce the desired outcome and could result in unintended consequences.

This paper contributes to several strands of the theoretical and empirical literatures. First, I contribute to a vast and diverse theoretical literature on imperfect lending and contracting (Bull, 1987; MacLeod and Malcomson, 1989; Thomas and Worrall, 1994; Watson, 2002; Ray, 2002; Levin, 2003; Albuquerque and Hopenhayn, 2004; Board, 2011; Halac, 2012; Andrews and Barron, 2016; Martimort et al., 2017; Troya-Martinez, 2017). The closest theoretical paper to mine is Martimort et al. (2017), which provides a theory of a two-sided limited enforcement problem in which buyers can default on debts and sellers can cheat on quality. In their setting, the buyer is the principal and increasingly shares a greater amount of surplus with the seller, implying dynamics where quantities *and* prices both increase. These dynamics do not match those observed in the setting I study, which has frictions that are common in many parts of the developing world. In contrast, I consider a model where, besides the incentives to default, the buyer has private information about the value of the relationship and the seller has the bargaining power.

Second, I contribute to the empirical literature on imperfect lending and contracting (McMillan and Woodruff, 1999; Banerjee and Duflo, 2000; Karaivanov and Townsend, 2014; Antras and Foley, 2015; Macchiavello and Morjaria, 2015; Boehm and Oberfield, 2020; Startz, 2024; Blouin and Macchiavello, 2019; Heise, 2024; Ghani and Reed, 2020; Ryan, 2020; Harris and Nguyen, 2022). Several papers, including Blouin and Macchiavello (2019), Ryan (2020), Startz (2024), and Harris and Nguyen (2022) have previously estimated the efficiency losses arising from imperfect contracting. In particular, Blouin and Macchiavello (2019) analyze strategic default on forward-contracts by sellers in the international coffee market, Ryan (2020) focuses on contract renegotiation in public procurement, Startz (2024) studies weak contract enforcement concerning seller opportunism and the presence of search frictions, and Harris and Nguyen (2022) studies the interaction of relational contracts with the thickness of a spot market. To my knowledge, my paper is the first empirical study to quantify the evolution of efficiency in relationships over time and find that dynamics matter significantly. Moreover, relative to these papers, my contribution is to quantify the inefficiencies from buyer opportunism in conjunction with seller market power. As the use of trade-credit is highly common in developing and high-income countries (Murfin and Njoroge, 2015; Giannetti et al., 2021; Burstein et al., 2024), and trade-credit reliance appears to increase with seller market power (Giannetti et al., 2011; Garcia-Marin et al., 2023) my findings and methodology have a wide-scope applicability.

Within the same body of work, this study relates to an extensive literature on formal and informal contracts in agricultural supply chains (Jacoby et al., 2004; Barrett et al., 2012; Michelson, 2013; Bubb et al., 2016; Macchiavello and Miquel-Florensa, 2017, 2019; Michler and Wu, 2020).[2] The literature supports the notion that formal contracting positively impacts welfare levels through real effects on income (Barrett et al., 2012; Michelson, 2013; Macchiavello and Miquel-Florensa, 2019), and that relational contracts can generate efficiency gains in the presence of contracting frictions (Jacoby et al., 2004; Macchiavello and Miquel-Florensa, 2017; Banerji et al., 2012). While Banerji et al. (2012) finds that relational contracts achieve constrained-efficiency under external output distortions, these gains from relational contracting may be limited in the presence of monopoly power (Jacoby et al., 2004), perform worse than vertical integration (Macchiavello and Miquel-Florensa, 2017), or may even be non-existent in certain contexts (Bubb et al., 2016). My contribution lies in providing a further analysis of the interaction between seller market power and relational contracts, empirically demonstrating that in the Ecuadorian context, the influence of relational contracts drives contracts towards unconstrained efficiency in the medium and long term.

Third, this paper relates to the literature examining the effects of market power in developing settings. Some studies have found that low market competition negatively impacts welfare,

[2]The paper is also linked to the literature testing communal risk-sharing in villages, which constitute a form of relational agreement (Townsend, 1994; Udry, 1994; De Weerdt and Dercon, 2006; Mazzocco and Saini, 2012; Chiappori et al., 2014). This literature indicates that while full village insurance is often rejected, certain networks among households (e.g., caste) do share risk efficiently, aligning with my finding that informal agreements can be near-optimal in some settings.

as firms distort total output and do not pass on cost savings to consumers (Fisman and Raturi, 2004; Atkin and Donaldson, 2015; De Loecker et al., 2016; Bergquist and Dinerstein, 2020; Casaburi and Reed, 2022; Grant and Startz, 2022; Reed et al., 2022; Chatterjee, 2023; Brugués and De Simone, 2024). However, some of the literature has demonstrated that monopoly power can enhance welfare in the presence of additional frictions. Such manifestations of the *theory of second-best* suggest that market power enables suppliers to offer credit (McMillan and Woodruff, 1999; Emran et al., 2021) and generate sufficient surplus for sustaining repeated relationships (Macchiavello and Morjaria, 2021; Boudreau et al., 2023).[3] In a similar vein, my paper finds that market power, manifested in the seller's ability to price discriminate flexibly, allows them to offer contracts that overcome each buyer's specific contracting frictions and achieve trade levels that would otherwise be unattainable.

Fourth, this work also follows the theoretical and empirical literature related to price discrimination (Maskin and Riley, 1984; Jullien, 2000; Villas-Boas, 2004; Grennan, 2013; Luo et al., 2018; Attanasio and Pastorino, 2020; Marshall, 2020). The works by Luo et al. (2018) and Attanasio and Pastorino (2020) provide estimation methodology and identification results for static non-linear pricing problems, with and without binding participation constraints, respectively. This paper generalizes their models and estimation methods to a multi-period setting by the relying on the recursive Lagrangian approach, a tool typically used in sovereign-debt macroeconomic models (Aguiar and Amador, 2014). Furthermore, while Attanasio and Pastorino (2020) provide identification results for non-linear pricing models with participation constraints under constant participation multipliers, I extend their findings by showing that, for non-constant multipliers, these models are identified under a parametric assumption.

The paper is organized as follows. Section 1 provides a description of the context and data. Section 2 offers the motivating facts that the model needs to match. Section 3 presents the model. Section 4 discusses identification and Section 5 the estimation procedure. Section 6 presents the estimated results and model fit. Section 7 discusses welfare and three counterfactual exercises. Finally, Section 8 concludes the paper.

[3]Theoretical studies in the theory of second best include Petersen and Rajan (1995), who demonstrate that increasing competition in bank lending can harm buyers by reducing the overall volume of lending when buyers have limited commitment to repaying their debts. This paper contributes to this literature by empirically showing that addressing only one market friction can result in welfare losses, and that addressing both enforcement and seller market power simultaneously could increase welfare. Additionally, my counterfactual results align with the theoretical findings of Genicot and Ray (2006), who show that improving enforcement reduces the buyer's expected payoff when the seller has bargaining power, and of Troya-Martinez (2017), who find that total welfare decreases as enforcement quality increases beyond a certain level.

1 Context, Interviews, and Data

Ecuador is an upper-middle-income country with weak enforcement of contracts and concentrated manufacturing markets. According to the World Bank Doing Business survey, Ecuador ranks as a median country in terms of Contract Enforcement, measuring the efficiency of courts in resolving commercial disputes, and one of the worst in terms of Insolvency measures, reflecting the inefficiency of courts in dealing with debt defaults due to bankruptcy (Online Appendix Figure OA-2). Additionally, the country's manufacturing sectors exhibit high levels of concentration, with average Herfindahl-Hirschman Indices of 0.6 for 6-digit economic codes (Online Appendix Figure OA-3), which are significantly higher than the concentration threshold of 0.25 used by the US Justice Department to identify highly concentrated markets.

1.1 Interviews

To gain a deeper understanding of the relationship management practices of manufacturing firms in Ecuador, I conducted hour-long interviews with high-ranking managers from 10 *manufacturing* firms in my studied industries in the spring of 2019. The following are the key findings from these interviews, from the perspective of the seller:

- Relationships among firms are not primarily based on written contracts but rather on informal agreements. Although transactions are documented, they are usually managed without the involvement of third-party enforcement, as formal enforcement is seen as costly and inefficient.[4]
- Quality issues from upstream suppliers are not a major concern, as the inputs used are highly standardized.[5]
- Enforcing payment for trade-credit transactions requires some investment in terms of time and personnel to pressure buyers to pay their debts.
- Most firms are aware that cash transactions offer discounts compared to trade-credit, but they often resort to trade-credit due to a lack of short-term liquidity.

This paper will not attempt to explain the underlying causes of these features but instead will focus on how they shape ongoing relationships.

[4]The Judicial Magazine of the Ecuadorian Government, available here, provides further evidence of the inefficiency of the court system. Two recent cases of buyer default were found, one taking 6 years to resolve and the other 4 years. A 2016 reform was made to the *Código Orgánico General de Procesos* to speed up debt collection, but in practice, this route is used as a last resort and takes around 2 years to enforce payment, according to personal estimates from 7,000 cases in the Civil Court in Quito in 2017.

[5]For textiles, their main supplies include raw textiles, which in the case of the manufacturing firms in my sample, are often imported (Online Appendix Table OA-8). For pharmaceuticals, variable inputs include active components, again often imported (Online Appendix Table OA-8). For cement-products, the main components include gravel and cement.

1.2 Administrative Data

The data used in this paper come from various administrative databases collected by Ecuador's Servicio of Rentas Internas (IRS) for tax purposes.

VAT database. By law, since 2008, firms are required to report all of their firm-to-firm inputs and purchases with information on the identity of the buyer and seller through the business-to-business (B2B) VAT system. I use the universe of B2b VAT database for 2008-2015 to measure the lengths of relationships. In particular, I define *age of relationship* as the total number of years that the seller has sold some positive value to the buyer in the past. Given the first year of observation is 2008, the age of the relationship is censored at +9.

Electronic Invoicing. The primary data source for the analysis is the electronic invoicing (EI) system. In 2014, Ecuador started rolling out a new EI system to collect VAT information more consistently, requiring large firms to implement this new technology. By 2015, the largest 5,000 firms were required to use the EI system for all sales. This system would send a copy of the transaction information to the buyer and government immediately after the transaction occurs. For each sale done by a firm in the system, the EI collects product-level information, including a bar-code identifier, product description, listed unit price, quantities, and discounts relative to listed prices, as well as transaction-level information, such as the buyer's unique national identifier and method of payment.[6] Method of payment can be cash, check, credit card, trade-credit offered by the seller with trade-credit payment terms, among others.

The data collected for this study is drawn from the EI system and pertains to 49 manufacturing firms operating in the textiles, pharmaceuticals, and cement sectors for the years 2016-2017. These firms are large, with an average (median) of 8,000 buyers (600) and a market share of 24% in their 6-digit sector at the national level and 50% in their sector at the provincial level. The database coverage is considered to be good, with the average selling firm in the sample having more than 90% of its reported sales captured by the EI system. Managerial interviews also revealed that most of these firms use the invoices received and sent for internal accounting purposes.

Because the manufacturing firms in this study produce multiple products, I use two approaches to measure quantities and prices. First, when presenting the stylized facts in Section 2, I focus on quality-adjusted prices and quantities. Specifically, I standardize prices and quantities by netting out product-seller-year fixed effects in transaction-level regressions of log unit prices or log quantities and I aggregate them to average standardized units at the buyer-seller-year level. Second, for the structural model estimation in Section 5, I rely on total quantities aggregated across all products for each buyer-seller-year and on average unit prices computed by dividing the total value of transactions by total quantities. Further details on the construction

[6] Listed prices may differ across buyers within a particular week, so listed discounts are not the only source of price variation.

of these variables are provided in the Online Appendix Section OA-1.1.

Financial Statements. I complement this information with yearly data on expenditures and wage bill from financial statements for all sellers for 2016-2017, which will be used to obtain firm-level variable costs.

1.3 Overview of the data

Online Appendix Section OA-1.2 provides detailed descriptive statistics for the datasets used. These statistics show that the sellers in the sample are large, well-established firms that make extensive use of imported inputs and channel most of their sales domestically, whereas buyers tend to be smaller, younger, less capital-intensive, and less exposed to international trade. Sellers outside the sample but in the same industries are typically even smaller—often micro-entrepreneurs with minimal reliance on imported inputs. Industry-specific breakdowns across textiles, pharmaceuticals, and cement-products indicate that, in each selling sector, a sizable share of buyers operate in wholesale and retail trade, suggesting relatively linear input needs. The electronic invoice data reveal that sellers transact with numerous buyers, with a median (average) bill of around \$USD 9,000 (\$USD 44,000). Illustrative product-level information highlights substantial variation in prices and costs within industries—potentially reflecting local market power or product differentiation—and indicates that product units within a firm's portfolio are comparable.

2 Motivating Evidence

This section presents evidence on how buyer-seller relationships work in the Ecuadorian supply chain.[7] Based on the data analyzed, there are three key findings: i) Trade heavily relies on past relationships and trade-credit arrangements. ii) As relationships mature, the quantity of goods exchanged increases, while prices decrease. iii) At any given time, larger purchases are associated with lower prices. In Section 3, a long-term contract model is proposed to capture these dynamics. The model allows the seller to use price discrimination across buyers and time, and enables buyers to default on trade-credit debts without facing legal consequences.

Fact 1: Large amount of trade occurs via repeated relationships. Figure 1a demonstrates the significance of repeated relationships for the sellers included in this study. The blue bars represent the average proportion of clients by length of relationship, while the green bars indicate the average proportion of the total quantity sold. The results reveal that although roughly

[7] Some of these relationship patterns have been previously documented in the literature. Heise (2024) and, partially, Monarch and Schmidt-Eisenlohr (2023) have previously documented the fact of relationship dynamics in quantities and prices for international trade, and Burstein et al. (2024) for intra-national trade in Chile. The persistence of intra-national links has been documented by Huneeus (2018) for Chile. Price discrimination in the context of medical devices and wholesale food has been documented by Grennan (2013) and Marshall (2020), respectively. Similarly, Antras and Foley (2015), Garcia-Marin et al. (2023), Amberg et al. (2020), and Burstein et al. (2024) have documented similar patterns of trade-credit issuance.

35% of all buyer-seller pairs consist of new buyers, only about 10% of the total trade is conducted through these fresh relationships. In contrast, relationships that have endured for at least nine years constitute less than 10% of all pairs but contribute to over 30% of the total trade.

Fact 2: Large share of transactions occur via trade-credit. The EI database includes payment method information, specifying whether the seller financed the transaction and the credit terms in days. For this analysis, I only consider whether the buyer was offered trade-credit, irrespective of the terms of the agreement.[8] Figure 1b displays the average share of purchases by buyer, across sellers, of relationships of a certain age that involved trade-credit. The data shows that the use of trade-credit is widespread, with approximately 65% of all purchases conducted via trade-credit in the first year of contact. For older relationships, around 70 to 75% of the volume of purchases are conducted via trade-credit.[9,10]

This fact has two important implications. Firstly, the seller bears a substantial portion of the risks associated with the transaction. In the absence of a strong legal enforcement framework, any opportunistic action taken by the buyer would result in the direct costs being absorbed by the seller. Secondly, the seller's opportunistic actions, such as cheating in quality or quantity, are likely to be limited (Smith, 1987; Klapper et al., 2012; Antras and Foley, 2015). Post-delivery, the buyer may retain the value of the transaction as a guarantee of quality. Therefore, when the seller finances transactions, the risk in trade tends to favor the buyer.

Fact 3: Quantities increase as relationships age. Figure 1c plots empirical evidence on the life cycle of quantities in buyer-seller relationships. The figure shows a binscatter regression of standardized log quantities on dummies for different ages of relationships in the cross-section. I find that older relationships tend to purchase more of a given product within a given year than younger relationships. These patterns also hold within a relationship, using total quantity purchased while controlling for pair fixed effects (Online Appendix Figure OA-4a).

Fact 4: Quantity discounts for a given age of relationship. Next, I examine the link between prices and quantities, focusing on *quantity discounts*, a common term in the literature for non-linear quantity-dependent decreasing price schedules (Maskin and Riley, 1984; Katz, 1984).[11]

Given the differences in the quantities sold by different manufacturers, I present quantities as quantiles, calculated within each seller and across the following relationship categories: i)

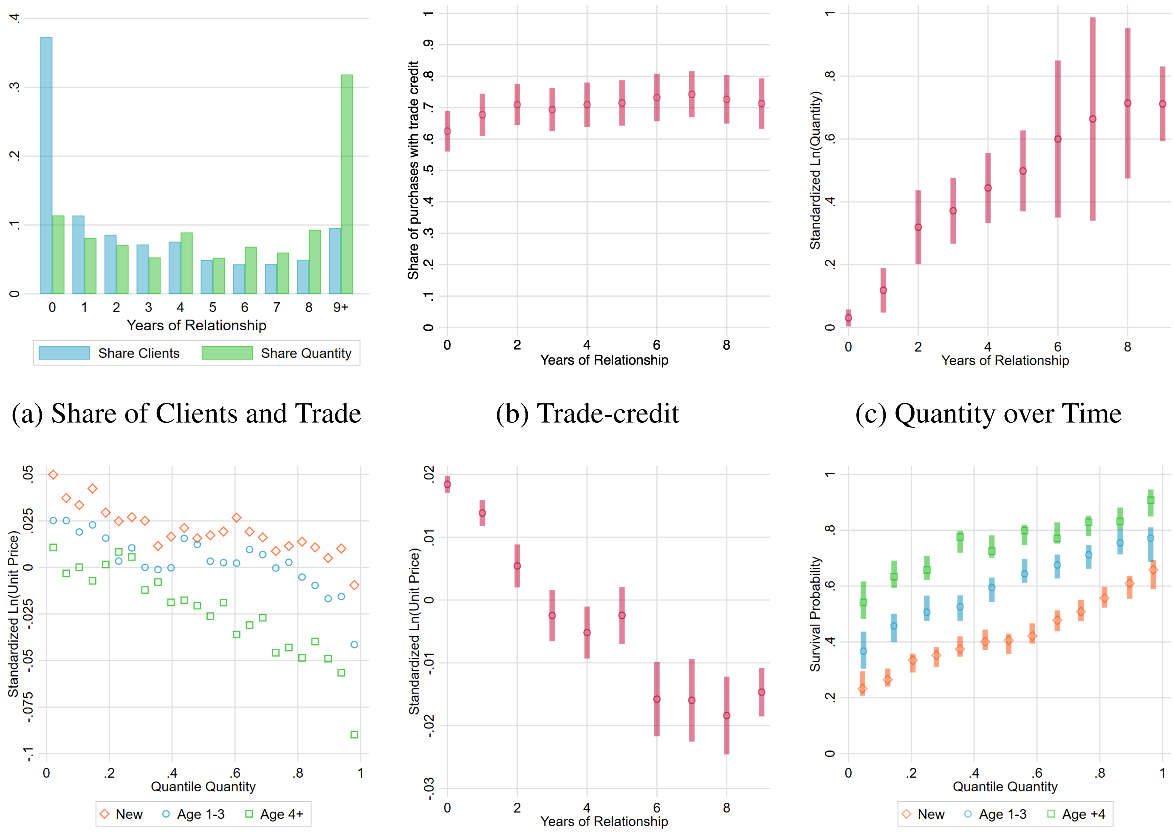
[8] On average, trade-credit agreements have a maturity of 40 days in textiles, 55 days in pharmaceuticals, and 40 days in cement products (Online Appendix Figure OA-9).

[9] These estimates are close in magnitude to inter- and intra-national figures from Chile, as reported by Garcia-Marin et al. (2023) and Burstein et al. (2024), respectively.

[10] It is possible that this empirical pattern for financing is valid for the sample of large manufacturing firms in my sector, but may not hold for smaller or informal firms. Reassuringly, using data from the World Bank, World Enterprise 2017 Survey for Ecuador, I find that 63% of retail firms and 77% of manufacturing firms use supplier or customer credit to finance working capital.

[11] The literature does not differentiate whether discounts come from a posted schedule or negotiated discounts. In this paper, I consider both sources by focusing on the effective price, which includes product-specific discounts as well as potential differences in posted prices across buyers. Also note, the term *quantity discounts* can also be seen in the literature as wholesale discounts.

Figure 1: **Motivating Facts**



(a) Share of Clients and Trade

(b) Trade-credit

(c) Quantity over Time

(d) Quantity Discounts

(e) Prices over Time

(f) Survival Rates

Notes: Subfigure (a) displays the distribution of the average of the share of clients and quantity sold by relationship age, calculated across all sellers in 2016. Sub-figure (b) displays the average of the share of purchases channeled through trade-credit, along with a 90% confidence interval, calculated across all sellers. Subfigure (c) displays the evolution of standardized log quantities, with their corresponding 90% confidence intervals, calculated across all sellers. The standardized log quantity is obtained by taking the average quantity sold in a given year for each seller-product and subtracting the log average quantity for that year. The standard errors are calculated at the seller-year level. Subfigure (d) shows the relationship between quantity purchased and standardized log unit price through a binscatter plot that displays the measure of unit price against the quantity sold, based on relationship age. The standardized log unit price is obtained netting out average log unit price for that year for each seller-product. The quantiles of quantity are calculated for each seller-relationship age combination. Subfigure (e) presents a binscatter plot of standardized log unit prices against years of relationship, controlling for a flexible spline of standardized log quantities. The standard errors are calculated at the seller-year level. Subfigure (f) displays a binscatter plot of the average survival rate of pairs at different ages and quantiles of quantity. The quantiles of quantities are calculated for each seller-age combination, and the error bars represent a 90% level of variation across all sellers.

new relationships, ii) relationships aged 1-3 years, iii) relationships aged 4 or more years. To compare quality-adjusted prices, the standardized unit price by quantiles of quantity is displayed as a binscatter plot in Figure 1d. The results demonstrate that, regardless of the relationship's age, larger quantities obtain lower quality-adjusted prices. This finding also holds true when considering average unit prices (Online Appendix Figure OA-4b). In terms of magnitude, a 10% increase in total quantity purchased is associated with an average price decrease of 2% (Online Appendix Table OA-5).

Fact 5: For a given quantity, older relationships pay lower unit prices. Figure 1e presents the relationship between unit prices and the age of the relationship. Using a binscatter regression of standardized log prices on age-of-relationship dummies, while controlling for a flexible spline of standardized quantities to account for potential quantity discounts, the figure reveals that older relationships receive up to 3% more quality-adjusted discounts compared to new relationships. These effects in standardized prices are comparable to those of moving from the median to the top percentile in quantity.

These dynamic discounts over time remain robust even after controlling for pair fixed effects in a regression of log average prices on relationship age (Online Appendix Figure OA-4c), indicating that the results are not driven by composition nor short-term fixed characteristics of the firm. Moreover, the results are robust and stable after including additional buyer and relationship-level controls, e.g., buyer's size or relationship demand and supply shares (Online Appendix Table OA-6). Interestingly, the discounts are only observed in trade-credit transactions and not in pay-in-advance ones (Online Appendix Table OA-9), supporting the interpretation of limited contract enforcement as the underlying mechanism for the observed price and quantity dynamics (over alternatives such as efficiency gains or demand assurance).

Fact 6: Relationships that trade more are more likely to survive. Figure 1f plots the share of relationships that survive from 2016 until 2017 by quantile of quantity in 2016 and age of relationship. The figure shows the survival rates of new links in red, links aged 1-3 years in blue, and links aged 4 years or more in green. I find that approximately 40 percent of new relationships survive at least one more year, 60 percent of relationships aged 1-3 years survive, and more than 75 percent of relationships aged 4 years or more survive. Moreover, within each relationship age category, pairs that trade higher volumes are more likely to survive from year to year.

While this paper does not focus on institutional differences among the sectors studied, it is important to highlight that the observed stylized facts (Facts 1 through 6) are consistent across all three industries analyzed (Online Appendix Section OA-3). Consequently, although specific primitives may vary by industry and seller, the underlying forces remain universally operative.

3 An Empirical Dynamic Contracting Model

This section introduces an empirical model of dynamic contracting with limited enforcement and seller market power from the perspective of a single seller. Through the first-order, necessary conditions for optimality of the seller and the buyers, I derive the key empirical equation.

3.1 Preliminaries

Setting. Consider an infinitely repeated relationship between a seller (the principal) and a buyer (the agent). Time is indexed by $\tau \geq 0$, and both parties discount future payoffs at a common factor $\delta < 1$. The buyer's preferences depend on a private type θ, which is drawn once at the outset from a continuous distribution with support $[\underline{\theta}, \overline{\theta}]$, where $\underline{\theta} = 1$ and $\overline{\theta} < \infty$, with cumulative distribution function $F(\theta)$ and density $f(\theta)$.[12] Although the type is privately observed, the distribution $F(\cdot)$ is common knowledge.

In addition to potential endogenous terminations of a relationship, relationships end exogenously every period from shocks that occur with common knowledge probability $X(\theta)$. As a result, the distribution of types evolves over time. Specifically, define the time-τ density as $f_\tau(\theta) = f(\theta)(1-X(\theta))^\tau / \int (f(m)(1-X(m))^\tau) dm$, with the associated cumulative distribution function $F_\tau(\theta)$.

A trade profile is defined by an infinite sequence of tariffs $\{t_\tau\}_{\tau=0}^{\infty}$ and quantities $\{q_\tau\}_{\tau=0}^{\infty}$. This profile delivers a discounted payoff to the seller of

$$\sum_{\tau=0}^{\infty} \delta^\tau (t_\tau - c_\tau q_\tau) \tag{1}$$

and to a buyer of type θ of

$$\sum_{\tau=0}^{\infty} \delta(\theta)^\tau (\theta v(q_\tau) - t_\tau), \tag{2}$$

where $v(\cdot)$ is a strictly increasing and strictly concave return function, $\delta(\theta) \equiv \delta(1-X(\theta))$, and c_τ denotes the constant marginal cost in period τ.[13]

Empirical evidence in Section 2 suggests that trade-credit is prevalent. Hence, I assume that the seller delivers goods before receiving payment, effectively extending trade-credit in every transaction. This assumption, while strong, streamlines the analysis by eliminating an additional choice variable, i.e., the choice to offer and accept trade-credit.

In line with the dynamic mechanism design literature (Pavan et al., 2014; Garrett et al., 2018), I assume that the seller can fully commit to a long-term contract. In particular, the seller does not alter the terms of the trade profile over time. This assumption is made for technical convenience, allowing me to concentrate on direct mechanisms thanks to the revelation principle, where the direct mechanism $C(\theta) = \{q_\tau(\theta), t_\tau(\theta)\}_{\tau=0}^{\infty}$ stipulates quantities and post-delivery tariffs in each period for agent reporting type θ.

Notably, while the seller has long-term commitment over the mechanism, the buyer can act opportunistically in the short-term, within each period. Namely, they can neglect payment and simply *take the goods and run*.

[12] The normalization $\underline{\theta} = 1$ is made without loss of generality.

[13] The concavity of the buyer's return function can be micro-founded by using diminishing returns in production for one input, keeping at least one other input fixed. This assumption is common in the literature. For instance, standard production function estimation generally assumes that capital is set one year in advance (e.g., Levinsohn and Petrin, 2003).

Timing. The contracting game unfolds as follows:

1. **Pre-trade (at $\tau = 0$):** The buyer observes their persistent private type θ. The seller offers the mechanisms menu $\{C(\theta)\}_{\underline{\theta}}^{\overline{\theta}}$, for which they have commitment. The buyer then either accepts or rejects the offer. Upon acceptance, the buyer reports a type $\widehat{\theta}$. If the buyer rejects the offer, both parties obtain their outside options, each normalized to zero.[14]
2. **Within each trading period $\tau \geq 0$:**
   - The seller first produces and delivers $q_\tau(\widehat{\theta})$.
   - The post-delivery payment $t_\tau(\widehat{\theta})$ is paid by the buyer, or the contract is breached.
   - When payment is made, the stage payoffs are $u_\tau(\theta, \widehat{\theta}) = \theta\, v(q_\tau(\widehat{\theta})) - t_\tau(\widehat{\theta})$ for the buyer and $\pi_\tau(\widehat{\theta}) = t_\tau(\widehat{\theta}) - c_\tau q_\tau(\widehat{\theta})$ for the seller.
   - In case of a breach, stage payoffs are $\theta v(q_\tau(\widehat{\theta}))$ for the buyer and $-c_\tau q_\tau(\widehat{\theta})$ for the seller.
3. **Between trading periods:**
   - If payment is made, the relationship may still be terminated exogenously with probability $X(\theta)$, in which case both parties revert to their outside options; otherwise, the relationship continues to the next period with probability $1 - X(\theta)$.
   - If a breach occurs, the seller terminates the contract and both parties receive their outside options in all subsequent periods.[15]

Equilibrium. The solution concept for this principal–agent game is a Perfect Bayesian Equilibrium in pure strategies. Under this equilibrium, the seller's contract $\{C(\theta)\}_{\underline{\theta}}^{\overline{\theta}}$ at $\tau = 0$ is profit-maximizing—subject to the relevant constraints—given their beliefs about the buyer's privately known type and the buyer's anticipated default decisions. In turn, the buyer's initial announcement and subsequent default or payment decisions each trading period form a subgame-perfect best response to the specified trade profiles, as well as the threat of termination in the event of default. Because the seller fully commits to the contract at $\tau = 0$, no further beliefs or actions on their part are required once the contract is in place.

3.2 Constraints

As usual, the set of constraints of the seller's problem contains the traditional individual rationality and incentive compatibility constraints of adverse selection problems.[16] However, this setting's novelty is to include additional enforcement constraints in each trading period, which

[14] For the buyer, this normalization is nonrestrictive under standard production function assumptions (e.g., linearity in variable inputs) or in a monopolistic supplier setting. Similarly, for the seller, constant returns to scale justify this normalization.

[15] Because enforcement constraints ensure that breaches never occur in equilibrium, there is no loss of generality in assuming termination as punishment. This *worst outcome* approach was introduced by Abreu (1988) and is standard in the relational contracting literature (Levin, 2003; Halac, 2012; Martimort et al., 2017).

[16] To make non-trivial theoretical predictions about the dynamics in the relational contract, one should add *interim* individual rationality constraint, $u_\tau(\theta) \geq \underline{u}$, for some lower bound $\underline{u}$. For the empirical estimating framework presented here, this additional assumption is not needed, as it enters into the limited enforcement multipliers used to satisfy the enforcement constraints.

act as endogenously determined participation constraints. Each of the enforcement constraints will ensure the buyer will not endogenously default in the specific time period.

Buyer's Incentive Compatibility. Under the assumption of perfectly persistent types, incentive compatibility requires that the agent evaluates their lifetime return:

$$\underbrace{\sum_{\tau=0}^{\infty} \delta(\theta)^{\tau} u_{\tau}(\theta)}_{\text{Lifetime truthful returns}} \geq \underbrace{\sum_{\tau=0}^{\infty} \delta(\theta)^{\tau} u_{\tau}(\theta, \widehat{\theta})}_{\text{Lifetime deviation returns}} \quad \forall \theta, \widehat{\theta}, \tag{IC-B}$$

where their period's net return is $u_{\tau}(\theta) \equiv u_{\tau}(\theta, \theta) = \theta v(q_{\tau}(\theta)) - t_{\tau}(\theta)$.

Buyer's Limited Enforcement Constraint. The novel friction in the model is the limited enforcement of the trade-credit contracts, which allows for the possibility of buyer's default. Under the assumption of contracting termination following a breach and the normalization of the buyer's outside option to zero, a *default-free* menu satisfies the limited enforcement constraint of the buyer:

$$\underbrace{t_{\tau}(\theta)}_{\substack{\text{Post-delivery}\\ \text{payment}}} \leq \underbrace{\sum_{s=1}^{\infty} \delta(\theta)^{s} u_{\tau+s}(\theta)}_{\substack{\text{Discounted future}\\ \text{truthful returns}}} \quad \forall \theta, \tau. \tag{LE-B}$$

The condition requires that the costs of breaking the relationship, in terms of the forgone opportunities of trade, have to be greater than the benefits from breaching the contract.

The buyer's LE-B constraint at $\tau = 0$ implies the individual rationality constraint required for buyer participation in trade.[17] From this, it follows that ex-ante trade under limited enforcement should leave participating buyers weakly better than under perfect enforcement whenever the seller has the bargaining power.

Buyer's Double-Deviation Constraint. The buyer could do a *double-deviation*, in which they announce type $\widehat{\theta}$ and default at some period τ. To prevent that, the truthful revelation menu must be appealing enough and satisfy

$$\underbrace{\sum_{\tau=0}^{\infty} \delta(\theta)^{\tau} u_{\tau}(\theta)}_{\text{Lifetime truthful returns}} \geq \underbrace{\delta(\theta)^{\tau} \theta v(q_{\tau}(\widehat{\theta}))}_{\substack{\text{Deviation + breach}\\ \text{stage return at } \tau}} + \underbrace{\sum_{s=0}^{\tau-1} \delta(\theta)^{s} u_{s}(\theta, \widehat{\theta})}_{\text{Deviation returns up to } \tau-1} \quad \forall \theta, \widehat{\theta}, \tau \tag{DD-B}$$

As the constraints IC-B and LE-B are necessary conditions for constraint DD-B, I concentrate on the relaxed problem and omit DD-B.[18]

[17] A mechanism C is individually rational if the participation constraint at $\tau = 0$ holds: $\sum_{\tau=0}^{\infty} \delta(\theta)^{\tau}(u_{\tau}(\theta)) \geq 0 \ \forall \theta$. To see how LE-B implies this, add $u_0(\theta)$ on both sides and note that $u_{\tau}(\theta) + t_{\tau}(\theta) = \theta v(q_{\tau}(\theta)) \geq 0$.

[18] For IC-B, simply consider $\tau \to \infty$ in DD-B. For LE-B, simply set $\widehat{\theta} = \theta$ in DD-B. Moreover, note that, for any $\widehat{\theta}$ such that $\delta(\theta)^{\tau}\theta v(q_{\tau}(\widehat{\theta})) + \sum_{s=0}^{\tau-1} \delta(\theta)^{s}[\theta v(q_s(\widehat{\theta})) - t_s(\widehat{\theta})] < \delta(\theta)^{\tau}\theta v(q_{\tau}(\theta)) + \sum_{s=0}^{\tau-1} \delta(\theta)^{s} u_s(\theta)$, condition LE-B implies DD-B, so for such $\widehat{\theta}$ the condition DD-B is irrelevant. For all other $\widehat{\theta}$, the condition is LE-B is a necessary condition for DD-B to hold. In particular, if DD-B holds, then $\delta(\theta)^{\tau} t_{\tau}(\theta) \leq \sum_{s=\tau+1}^{\infty} \delta(\theta)^{s} u_s(\theta) - \Big(\sum_{s=0}^{\tau-1} \delta(\widehat{\theta})^{s}[\theta v(q_s(\widehat{\theta})) - t_s(\widehat{\theta})] - \sum_{s=0}^{\tau-1} \delta(\theta)^{s}[\theta v(q_s(\theta)) - t_s(\theta)]\Big) \ \forall \theta, \widehat{\theta}, \tau$. As the term in the brackets is positive by assumption, LE-B holds.

3.3 The Firm's Problem

Denote total surplus as $s(\theta,q,c) = \theta v(q) - cq$. The principal's problem is to maximize their lifetime profits. As the buyer's type θ is unknown, their problem is set in expectation. The seller therefore chooses a direct mechanism that maximizes their expected lifetime profits:

$$\max_{\{u_\tau(\theta),q_\tau(\theta)\}} \sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\overline{\theta}} [s(\theta,q_\tau(\theta),c_\tau) - u_\tau(\theta)] f_\tau(\theta) d\theta, \tag{SP}$$

such that IC-B, LE-B, and DD-B are satisfied. That is, the objective of the seller is to maximize total surplus while reducing the share of surplus given to the buyer as much as possible without breaching the constraints.

3.4 Necessary First-Order Conditions

The next proposition provides the necessary conditions for the profit-maximization problem of the firm.

Proposition 1. *Suppose that the contract $C^*(\theta) = \{q^*_\tau(\theta), t^*_\tau(\theta)\}_{\tau=0}^{\infty}$ maximizes the lifetime profits of the firm subject to IC-B, LE-B, and DD-B. Then, it must be that the contract satisfies the first-order conditions of the seller's problem SP:*

$$\theta v'(q^*_\tau(\theta)) - c_\tau = \frac{\Gamma_\tau(\theta) - F_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma^\tau_s(\theta))\tilde{\Gamma}^\tau_s(\overline{\theta}) + \theta\gamma_\tau(\theta)}{f_\tau(\theta)} v'(q^*_\tau(\theta)), \tag{SFOC}$$

for each τ and θ, such that $\gamma_\tau(\theta)$ is the corresponding Lagrange multiplier for type's θ LE-B constraint at time for type θ; $\Gamma_\tau = \int_{\underline{\theta}}^{\theta} \gamma_\tau(x)dx$ is the cumulative multiplier on the constraint from $\underline{\theta}$ to θ, such that $\Gamma_\tau(\overline{\theta}) = 1$; $\Gamma^\tau_s(\theta)$ is the conditional cumulative multiplier $\tau - s$ periods ago from $\underline{\theta}$ to θ; and $\tilde{\Gamma}^\tau_s(\overline{\theta})$ the discounted cumulative multiplier $\tau - s$ periods ago from $\underline{\theta}$ to $\overline{\theta}$. Moreover, the tariffs satisfy the following local incentive compatibility condition:

$$t^{*'}_\tau(\theta) = \theta v'(q^*_\tau(\theta)) q^{*'}_\tau(\theta). \tag{t-RULE}$$

For a full derivation, refer to Appendix A.

The allocation equation SFOC responds to intuitive forces. For clarity, assume momentarily that $v(q) = kq^\beta$ and the breakup probability is zero for all types, i.e., $X(\theta) = 0$ for all θ. Under these assumptions, $\Gamma^\tau_s(\theta) = \Gamma_s(\theta)$, $\tilde{\Gamma}^\tau_s(\overline{\theta}) = 1$, $F_\tau(\theta) = F(\theta)$, and $f_\tau(\theta) = f(\theta)$. The equation SFOC simplifies to:

$$q_\tau(\theta)^{1-\beta} = \overbrace{\frac{k\beta}{c_\tau}}^{\text{Inv. } \mu} \left[\overbrace{\theta - \frac{1-F(\theta)}{f(\theta)}}^{\text{Virtual Surplus}} - \overbrace{\frac{\theta\gamma_\tau(\theta)}{f(\theta)}}^{\text{LE}} + \overbrace{\frac{(1-\Gamma_\tau(\theta))}{f(\theta)}}^{\text{LE+IC}} + \overbrace{\frac{\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))}{f(\theta)}}^{\text{Past LE + IC}} \right] \tag{3}$$

which resembles the typical solution to an adverse selection problem. In this solution, the allocation is determined by an inverse markup (μ) rule adjusted by the *modified virtual surplus*, which accounts for necessary rents due to incentive compatibility and the limited enforcement constraint.

First, as is typical, the amount of allocated quantities decreases as the inverse markup that a seller would charge under linear monopolist pricing (μ) increases.

Second, through the virtual surplus, higher types (θ) receive greater quantities, while the incentive compatibility constraint forces the seller to distort trade downward for lower types ($1-F(\theta)$), thus granting higher types informational rents. These are the common forces at play in non-linear pricing contracts (Maskin and Riley, 1984).

Third, when the current limited enforcement constraint is binding ($\gamma_\tau(\theta) > 0$), it restricts the volume of trade. Keeping the future stream of quantities constant, if the buyer is on the verge of defaulting, the seller needs to reduce tariffs immediately. However, to maximize profits by reducing total costs per dollar of revenue, the seller must also decrease quantities. Therefore, enforcement concerns lead to a reduction in contemporaneous quantities.

Fourth, a countervailing force exists: to maintain incentive compatibility and prevent low types from mimicking higher types, quantities are uniformly shifted upwards by $1-\Gamma_\tau(\theta)$. This countervailing force is also present in the static allocation equations in Jullien (2000) and Attanasio and Pastorino (2020).

Fifth, dynamic promises aimed at increasing future trade to incentivize the payment of debts are captured by the inclusion of past cumulative multipliers ($\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))$). These multipliers generate the backloading of quantities, acting as a promise-keeping constraint where types whose limited enforcement constraint was binding in the past receive higher quantities in the present.

The equilibrium combination of $\Gamma_\tau(\theta)$, $\Gamma_s(\theta)$, and $\theta\gamma_\tau(\theta)$ determines whether the allocated quantity is greater or lower than it would be under full enforcement.

Returning to the general equation SFOC, it is worth highlighting the role of selection implied by the exit probability $X(\theta)$ in the allocation of quantities, as it generates two opposing forces.

On the one hand, the selection functions exert downward pressure on quantities through the virtual surplus and the cumulative multipliers. In the virtual surplus, positive selection ($X'(\theta) < 0$) implies lower quantities over time, all else being equal. This occurs as the selection pattern concentrates the distribution towards higher types over time, forcing the seller to decrease future quantities for middle types to maintain incentive compatibility. Moreover, selection also influences the promises captured through past cumulative multipliers. Ceteris paribus, if the selection function $X(\theta)$ implies first-order stochastic dominance over another selection function $\tilde{X}(\theta)$, i.e., $F_\tau(\theta) \leq \tilde{F}_\tau(\theta)$, the past cumulative multipliers move closer to one for each type. This shift reduces the upward push that past promises would normally provide.

On the other hand, the heterogeneity in exit rates implies relatively less discounting of past multipliers for middle and upper types. This means that as their past multipliers are discounted less, the impact of earlier promises is stronger for them. Consequently, compared to low types, the selection mechanism leads to more significant backloading of quantities for middle or high types.

3.5 Model Properties

Next, I discuss how the model rationalizes the stylized facts in Section 2. Formal proofs for the statements in this subsection and a two-type solved example appear in Online Appendix Section OA-4. That appendix also examines equilibrium contracts under individual relaxations of the model's constraints, showing that only the complete framework—incorporating both limited enforcement and asymmetric information—can fully explain the stylized facts.

Non-Stationarity. The optimal contract must be non-stationary, driven by the forces implied by the limited enforcement constraints (Proposition 3). In particular, these constraints create a dynamic asymmetry in incentives between buyer and seller. The buyer evaluates current tariffs relative to future net returns, which, all else being equal, incentivizes the seller to reduce current quantities while keeping current tariffs constant, thereby increasing current profits and still satisfying the enforcement constraint. Relative to the optimal stationary contract, it is possible to construct non-stationary deviations that increase initial profits, even in the presence of the incentive compatibility constraint stemming from asymmetric information.

Quantity Discounts. At each relationship age, the seller offers quantity discounts to maintain incentive compatibility (Proposition 4). The conditions needed to support such discounts are strengthened forms of the usual assumptions in non-linear pricing models (e.g., Maskin and Riley, 1984). Specifically, the evolution of distribution types $F_\tau(\theta)$ must preserve log-concavity and satisfy a modified monotone hazard condition, in addition to the standard requirement that quantities be strictly increasing in the buyer's type for each relationship age ($q'_\tau(\theta) > 0$).

Backloading of Quantities. The model rationalizes increases in quantity over time ($q_\tau(\theta) \leq q_{\tau+1}(\theta)$) if and only if enforcement constraints are relaxed ($\gamma_\tau(\theta) \leq \gamma_{\tau+1}(\theta)$) (Proposition 5, i.). Thus, the model permits quantity dynamics.

Moreover, absent selection patterns, the model explicitly predicts backloading of quantities. (Proposition 5, ii.) There exists a finite time period τ^* such that enforcement constraints are no longer binding for any type θ, causing the contract to converge to a long-term stationary equilibrium. In this equilibrium, quantities reach their highest levels for each type ($q_{\tau^*}(\theta) \geq q_\tau(\theta)$ for $\tau^* \geq \tau$).

Backloading of Prices. The model accommodates backloading of prices (Proposition 6). In particular, if the quantity schedule (weakly) increases over time for all buyers (and strictly for the lowest type), then the resulting increase in quantities forces a global decrease in prices to preserve local incentive compatibility.

3.6 Discussion of Modeling Assumptions

Although standard in the literature (Pavan et al., 2014; Garrett et al., 2018), the assumption that the seller can commit to the mechanism might be unrealistic in a setting where the buyer can defect and default. Despite this limitation, I adopt the commitment assumption as it allows me to focus on direct mechanisms through the standard revelation principle approach.

Another assumption is that the seller does not cheat on quality. This issue, combined with the lack of enforcement on the buyer's side, has been theoretically explored for two seller types by Martimort et al. (2017). While their framework could technically be applied, I do not follow it for several reasons. First, a model featuring seller opportunism would generate front-loaded prices, which is inconsistent with the data. Second, because a large share of trade is conducted via trade-credit, buyers can withhold payment if quality is subpar, reducing the scope for seller cheating (Smith, 1987; Klapper et al., 2012; Antras and Foley, 2015). Finally, the sellers in this study are larger, more capital-intensive, and more directly involved in input sourcing than the average manufacturing firm, reducing the likelihood of quality issues arising from production errors.

Buyer types are assumed fully persistent due to data limitations. While the model can accommodate Markov types, empirical implementation would require tracking new buyer transactions over time, which is infeasible with the two years of data available.

4 Identification

In this section, I discuss the identification of the model primitives θ and $v(\cdot)$ and the auxiliary functions $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$. Each of these primitives and auxiliary functions are seller-year-specific. The results presented here build on the identification work of Luo et al. (2018) and Attanasio and Pastorino (2020), but extend the analysis to a multi-period framework rather than a single-period problem. To derive a key identifying equation that maps data into primitives, I rely on the necessary conditions for the seller and the buyer as outlined in Proposition 1.

4.1 Observables and Known Objects

For each seller in a given year, the observables are unit prices $p_\tau(q)$ (or tariffs $T_\tau(q)$) and quantities q_τ for different buyers with relationship age τ, as well as marginal costs c.[19] Throughout this section, I abstract from the possibility of exogenous breakups; the possibility of breakups will be reintroduced in estimation.[20]

[19] The price schedule $T_\tau(\cdot)$ and its derivatives are nonparametrically identified from information on prices and quantities alone (Perrigne and Vuong, 2011), so in this section, I treat them as known. Moreover, I treat c as known, as I can backout average cost (across all product varieties) using information on total variable costs and total seller output.

[20] As exogenous breakups can be directly estimated from the data, they are treated as known during identification. Their inclusion would only complicate the notation without providing substantial insights regarding identification.

4.2 Identification Assumptions

I now begin by stating the identification assumptions (IA).

Identification Assumption 1. *Each seller offers a unique menu of dynamic contracts to all buyers, and such menu satisfies equations SFOC and t-RULE for all θ and τ.*

Identification Assumption 2. *Within each period, quantity increases strictly monotonically with type θ: $q'_\tau(\theta) > 0$.*

Identification Assumption 3. *The return function is of the form $v(q) = kq^\beta$, for $k > 0$ and $\beta \in (0,1)$.*

IA 1 guarantees the existence and uniqueness of the contract.[21] Moreover, instead of relying on forward iteration to solve the problem, IA 1 allows me to collapse all information about future unobserved quantities and tariffs into the limited enforcement multipliers.

IA 2 directly links observed quantities with underlying unobserved types, allowing us to infer that buyers purchasing higher quantities have higher types. IA 2 may fail under certain conditions, leading to quantities bunching over different types: (1) If the distribution of types $F_\tau(\theta)$ is non-continuous, presenting masses (jumps) at some type θ. (2) If the exit probability $X(\theta)$ is not smooth, implying jumps in future distribution of types. (3) If the exit probability $X(\theta)$ implies that the distribution $F_\tau(\theta)$ becomes log-convex.[22] (4) If the return function $v(\cdot)$ is too inelastic, making it difficult to implement incentive compatibility without significantly changing quantities. In such cases, bunching may be desirable to reduce losses from informational rents while preserving incentive compatibility.

Finally, through IA 3, I consider constant-elasticity parametrization for the return function $v(q)$, which will be essential for point identification of the primitives and auxiliary functions by restricting the number of parameters that need to be identified.

4.3 Deriving the Key Identification Equation

Exploiting the fact that the mapping from agent type θ to quantity q_τ is strictly monotone (IA 2), one can write the seller's first-order condition SFOC in terms of quantiles α:

$$k\beta\theta_\tau(\alpha)q_\tau(\alpha)^{\beta-1} - c = \left[\Gamma_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha)) + \frac{\theta_\tau(\alpha)}{\theta'_\tau(\alpha)}\gamma_\tau(\alpha)\right] k\beta\theta_\tau(\alpha)q_\tau(\alpha)^{\beta-1}\frac{\theta'_\tau(\alpha)}{\theta_\tau(\alpha)}, \qquad \text{(I-Q)}$$

[21] Although uniqueness assumptions are strong, they are often used in the identification of dynamic games, as these types of games may have multiple equilibria (Aguirregabiria and Nevo, 2013).

[22] In simulations, negative selection patterns ($X'(\theta) > 0$), which are not consistent with the data but a theoretical possibility nonetheless, lead to this.

as well as the derivative of the buyer's tariff rule t-RULE:

$$T_\tau'(q_\tau(\alpha)) = k\beta\theta_\tau(\alpha)q_\tau(\alpha)^{\beta-1}, \tag{I-T}$$

where $\alpha \in [0,1]$, $\theta_\tau(\alpha)$ and $q_\tau(\alpha)$ are the α-quantiles of the agent's type and quantity at tenure τ, respectively, and I used the fact that the observed tariff schedule can be mapped to the model tariff schedule by $T_\tau(q_\tau(\theta(\alpha))) = t_\tau(\theta(\alpha))$. Notice as well that I have linked past multipliers $\Gamma_s(\alpha)$ with the buyer's current quantile α, as types are fully persistent and quantiles are held fixed across tenures, and used the following relationships derived from IA 2: (i) $F_\tau(\theta(\alpha)) = \alpha$, (ii) $f_\tau(\theta(\alpha)) = 1/\theta_\tau'(\alpha)$, (iii) $\Gamma_\tau(\theta_\tau(\alpha)) = \Gamma_\tau(\alpha)$, and (iv) $\gamma_\tau(\theta(\alpha))\theta_\tau'(\alpha) = \gamma_\tau(\alpha)$.

By relying on I-T and the parametrization in IA 3, one can obtain the following expression for the ratio $\theta_\tau'(\alpha)/\theta_\tau(\alpha)$ in I-Q:

$$\frac{\theta'(\alpha)}{\theta(\alpha)} = q_\tau'(\alpha)\left[\frac{T''(q(\alpha))}{T'(q(\alpha))} + \frac{1-\beta}{q_\tau(\alpha)}\right], \tag{4}$$

which depends on functions of tariffs and quantities, and only one unknown elasticity parameter β.

Substituting I-T into I-Q, the key identification equation becomes:

$$\frac{T_\tau'(q_\tau(\alpha)) - c}{T_\tau'(q_\tau(\alpha))} = \frac{\theta_\tau'(\alpha)}{\theta_\tau(\alpha)}\left[\Gamma_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))\right] + \gamma_\tau(\alpha), \tag{I-EQ}$$

where $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$ are unknown, and the ratio $\theta_\tau'(\alpha)/\theta_\tau(\alpha)$ is given in 4 under IA 3. This equation will be the base for the identification of all unknown functions and parameters. Note that the full persistency of types, combined with IA 1, implies that the dynamic contract is identified using cross-sectional variation within cohorts for a given seller-year.

4.4 Identification Results

I now present point identification results.

Proposition 2. *Under IA 1, 2, and 3, the auxiliary functions $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$, and the elasticity parameter β are identified from cross-sectional data on prices, quantities and marginal costs from one seller. Moreover, the functions for types $\theta_\tau(\cdot)$ and $\theta_\tau'(\cdot)$ are identified over $\alpha \in [0,1]$ and the return function scale parameter k is identified.*

The proof is relegated to Appendix B. The argument involves the following four steps: (1) For the highest types $\alpha \approx 1$ at $\tau = 0$, the identification equation I-EQ is shown to depend solely on one unknown elasticity parameter β. This is because the cumulative multiplier can be unconditionally shown to be $\Gamma_0(1) = 1$, which in turn implies that the observed difference in marginal prices $T'(1)$ and marginal costs c directly reveals $\gamma_0(1)$. To a first-oder approximation, the equation I-EQ is known for each $\alpha \approx 1$ up to the unknown parameter β. Therefore, pooling the equations across the highest types, the cross-sectional variation in prices identifies β. (2) Once β is identified, the functions $\Gamma_0(\alpha)$ and $\gamma_0(\alpha)$ can be recovered for all types. They are determined as the unique solutions to an ordinary differential equation whose other components are known, which implies that the multipliers are point-identified. (3) With β and the multipliers for tenures $s < \tau$ already identified, the multipliers at tenure τ are recovered as the unique solutions to the corresponding ordinary differential equation. (4) Finally, having identified all multipliers, simply apply the identification argument in Luo et al. (2018) for static non-linear pricing problems to identify the distribution of types $\theta_\tau(\alpha)$. The scale parameter k is then recovered using known elements via equation I-T.

Although parametrizing $v(\cdot)$ via IA 3 yields point identification, in the estimation procedure I choose to parametrize $\Gamma_\tau(\cdot)$ as a flexible function of q_τ, rather than imposing a parametric form on $v(\cdot)$. This approach simplifies solving the differential equations for $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$, as it confines them to a known family of functions. As shown below, the return function $v(\cdot)$ is recovered semi-parametrically, in a manner consistent with the chosen parametrization of the multiplier functions.

4.5 Discussion of Limitations in Identification

In my methodology, I leverage the fact that the seller knows the optimal contract solution, which must satisfy the first-order conditions of both the seller and the buyer. Besides the benefits of allowing estimation without solving the full model through forward iteration, this assumption also proves useful if the model is misspecified. Specifically, it allows the buyer to have outside options that the econometrician does not observe, provided the seller is aware of these outside options. They are then incorporated into the enforcement constraints. Although the econometrician may not distinguish outside options from future promises, these factors do not create identification issues for the welfare analysis primitives.[23]

A key limitation of this approach is that it cannot handle counterfactuals involving dynamic quantities. Solving for those would require forward iteration solution methods. Nevertheless, this methodology yields valuable insights into the efficiency of *actual* trade, which is the paper's main focus.[24]

Finally, my identification results rely on observing (or using proxies for) marginal costs. The gap between prices and marginal costs indicates whether trade for the highest type is distorted by enforcement constraints. Previous work by Attanasio and Pastorino (2020) infers unobserved

[23] Mispecification of the model will affect the equilibrium tariff solution. For example, if buyers have a constant outside option, the equilibrium tariffs will be lower by the value of the outside option. However, this does not affect marginal prices or the primitives (such as the base marginal return or the type) identified from them.

[24] Model mispecification regarding outside options also affects counterfactuals under different enforcement or pricing regimes. If the outside options are constant, the counterfactual outcomes remain correct in terms of efficiency, though surplus division is biased in favor of the seller. When outside options are heterogeneous, the counterfactual efficiency may also be affected; the direction of this bias is uncertain *ex-ante* and depends on the distribution of types and the curvature of the return function.

costs via the parametrization of the multipliers, thereby jointly identifying costs consistent with those multipliers. In the same vein, Luo et al. (2018) identifies costs by assuming that, in the absence of enforcement constraints, trade for the highest type is efficient, with marginal prices equating marginal costs. By drawing on production-cost data for sellers, I can relax these assumptions.

5 Estimation

In this section, I first describe the estimation sample and define relationship tenure used in estimation. Then, I present the Intermediate Steps used to estimate the objects that are assumed to be known for identification purposes but need to be estimated from finite data. Lastly, I describe the Main Steps in the estimation process to recover the primivities and auxiliary functions, which rely on a cross-sectional approach using the main identification equation with the available data for each seller-year separately.

5.1 Definitions of Relationship Tenure and Estimation Sample

To facilitate estimation and reduce measurement error in relationship ages, I impose two restrictions. First, I require that buyers have at least one previous relationship with some seller (not necessarily those in my sample) prior to 2016.[25] Second, I pool relationship ages using the following classification method and define *relationship tenure* between seller i and buyer j at year t as:

$$tenure_{ijt} = \begin{cases} \text{pair-age}_{ijt} & \text{if pair-age}_{ijt} < 5, \\ 5 & \text{if pair-age}_{ijt} \geq 5. \end{cases}$$

I bunch all older relationships together to ensure a sample large enough for estimation.[26]

The final sample with the estimated structural model consists of 24 sellers with information for both 2016 and 2017, and 25 sellers with information for either 2016 or 2017. I consider these 73 seller-year observations on their own, but use sellers that appear in multiple years to validate the fit over time.

5.2 Estimation of Objects Assumed as Known in Identification

Before reaching the key estimating equation, there are three intermediate steps to recover the objects assumed as known in identification. Namely, I detail the steps to recover the 1) tariff function, 2) the heterogeneous exit/survival rates, and 3) the marginal costs.

[25] I verify that this restriction is not driving the results by estimating the model with *all* available buyers, despite the possible measurement error in the age of the relationship. Overall, results are very consistent with those presented here. Results of this robustness check are available upon request.

[26] The threshold at +5 is not driving the results, as results are robust to using higher threshold values.

Intermediate Step 1: Tariff Function. For identification, I treated the tariff function $T_\tau(\cdot)$ as given. However, I observe only pairs of payments and quantities $(t_{i\tau}, q_{i\tau})$ for $i = 1,2,...,N_\tau$ for each tenure. The pricing model discussed in Section 3 implies that observed tariffs lie on the curve $t_{i\tau} = T_\tau(q_\tau(\theta_i\tau))$, as they are both functions of the type $\theta_{i\tau}$ in a given tenure. However, observed prices and quantities may not lie on the curve, if there is measurement error or further unobserved heterogeneity beyond quantity and relationship age, introducing additional randomness beyond $\theta_{i\tau}$.

To deal with this additional randomness, I follow Perrigne and Vuong (2011), who show that the tariff function is nonparametrically identified under the assumption that observed tariffs differ from optimal tariffs due to random measurement error. In particular, observed tariffs are a function of optimal tariffs $t_{i\tau} = T_\tau(q_{i\tau})e^{\upsilon_{i\tau}}$, such that $\upsilon_{i\tau}$ is independent of $q_{i\tau}$.

I consider a parametric version of the model, in which $T_\tau(q) = e^{\rho_{0\tau}}q^{\rho_{1\tau}}$. This leads to the estimation model with measurement error:

$$ln(t_{i\tau}) = \rho_{0\tau} + \rho_{1\tau} ln(q_{i\tau}) + \upsilon_{i\tau}, \tag{5}$$

where $t_{i\tau}$ is the observed tariff and $q_{i\tau}$ is the observed quantities for buyer i with tenure τ. Under the given assumption of independence, the tariff schedule can be estimated via ordinary least squares. The estimated tariff schedule linking observed quantities is $\widehat{T}_\tau(q_{i\tau}) = e^{\hat{\rho}_{0\tau}}q_{i\tau}^{\hat{\rho}_{1\tau}}$, while the marginal tariff is $\widehat{T}'_\tau(q_{i\tau}) = \hat{\rho}_{1\tau}t_{i\tau}/q_{i\tau}$. Note that I allow for differences in tariff schedules across τ, responding to the dynamic treatment of the problem, i.e., the same level of quantity q may have different associated tariffs if the buyer-seller relationship is new or has been sustained for some years.

Intermediate Step 2: Heterogeneous Survival Rates. I estimate heterogeneous survival rates $S(\cdot)$, i.e., $(1 - X(\cdot))$, at the percentile-tenure level. In particular, I rank buyers in percentiles of quantity for each tenure in 2016. I then calculate the share of buyers in each percentile that survived until 2017. To reduce noise and preserve monotonicity and smoothness of the survival rate, I then approximate the estimated nonparametric survival rates as a logistic function of percentiles:

$$S_\tau(r) = \frac{\exp(a_\tau + b_\tau r)}{1 + \exp(a_\tau + b_\tau r)} + \varepsilon^s_\tau(r), \tag{6}$$

where $S_\tau(r)$ is the share of buyers surviving from 2016 until 2017 in percentile rank r for tenure τ and $\varepsilon^s_\tau(r)$ is Gaussian noise orthogonal to r.

Intermediate Step 3: Marginal Cost. Marginal cost is estimated directly from the data under the assumption of constant marginal cost, which implies marginal cost is equal to average variable cost. I present validating exercises for this assumption in my setting in Online Appendix Section OA-7.[27] Therefore, I recover average variable cost by dividing the sum of total expenditures and total wages by the total quantity sold for each seller-year.

5.3 Estimation of the Primitives and Auxiliary Functions

Now, I detail how to recover the auxiliary functions $\Gamma_\tau(\cdot)$ and the primitives θ and $v(\cdot)$, as well as their derivatives, by relying on the key identification equation I-EQ.

Main Step 1: Auxiliary Functions $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$. First, to recover the auxiliary functions $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$, I rely on an iterative approach, starting at $\tau = 0$, assuming that an estimate for $\Gamma_s(\cdot)$ for $\tau > s$ is already available from previous iterations.[28]

With the survival rates, marginal costs, and tariff functions in hand, the empirical analog of the key identifying equation I-EQ is given by:

$$\frac{\widehat{T}'_\tau(q_{i\tau}) - \widehat{c}}{\widehat{T}'_\tau(q_{i\tau})} = \frac{\theta'_\tau(\alpha)}{\theta_\tau(\alpha)}\left[\Gamma_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\widehat{\Gamma}_s^\tau(\alpha))\right] + \gamma_\tau(\alpha), \tag{7}$$

where the past conditional cumulative multiplier estimates $\widehat{\Gamma}_s^\tau(\alpha)$ for $s < \tau$ are obtained via numerical integration.[29]

Equation 7 contains multiple unknown functions. Above, I demonstrated that parametrizing $v(\cdot)$ (IA 3) is sufficient for nonparametrically identifying the auxiliary functions. However, instead of relying on the parametrization of the return function $v(\cdot)$, I leverage the fact that the LE multiplier $\Gamma_\tau(\alpha)$ possesses the properties of a cumulative distribution function.[30] Thus, I parametrize the multiplier as a logistic distribution:

$$\Gamma_\tau(\alpha) = \frac{exp(\phi_\tau(q_\tau(\alpha)))}{1 + exp(\phi_\tau(q_\tau(\alpha)))}, \tag{8}$$

where $\phi_\tau(q_\tau(\alpha))$ is a linear polynomial.[31] Under this parametrization, the derivative of the multiplier is $\gamma_\tau(\alpha) = \phi'_\tau(q_\tau(\alpha))\Gamma_\tau(\alpha)(1-\Gamma_\tau(\alpha))$. As mentioned above, the parametrization over $\Gamma_\tau(\cdot)$ simplifies the estimation approach, as the solution to the differential equations for $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$ are restricted to depend on the same parameters.

[27] In particular, I show that average variable costs are highly serially correlated within a given seller and that a test for constancy of marginal costs relying on demand-side instruments fails to reject constancy.

[28] Although the iterative approach may introduce propagation errors from earlier stages, joint estimation is infeasible because it requires numerically integrating past cumulative multipliers as an extra step during the parameter search. Joint estimation is viable if hazard rates are homogeneous or if relationships never terminate.

[29] Specifically, I numerically integrate the function $\left(\widehat{S}_\tau(\alpha)\right)^{\tau-s}\widehat{\gamma}_s^\tau(\alpha)$ using inverse transform sampling for each quantile α. Here, $\widehat{S}_\tau(\alpha)$ is the estimated survival rate from *Intermediate Step 2*, and $\widehat{\gamma}_s^\tau(\alpha)$ is the derivative of the cumulative multiplier at period s that corresponds to a buyer in quantile α during period τ. To match multipliers s periods ago, I rely on the estimated survival rates to generate a percentile-percentile transition matrix. This matrix allows me to align percentiles α_s for $s < \tau$ with percentiles α_τ.

[30] See Appendix A, which shows $\Gamma_\tau(\cdot)$ is non-negative, non-decreasing, and with boundary $\Gamma_\tau(1) = 1$.

[31] The multiplier function is the solution to a differential equation. As shown in Online Appendix Section OA-4.1.2, it is a function of the cumulative distribution of types θ, the marginal cost, and the expected base marginal return (i.e., depends on the curvature of the return function).

To compensate for the restrictions on the LE multipliers, I consider instead a flexible function for $\theta'(\alpha)/\theta(\alpha)$, specifically as an inverse quadratic function of quantity:

$$\frac{\theta'(\alpha)}{\theta(\alpha)} = \frac{1}{d_0 + d_1 q_\tau(\alpha) + d_2 q_\tau(\alpha)^2}. \tag{9}$$

All together, the key identification equation I-EQ is translated into the following estimating equation:

$$\frac{\hat{\rho}_{1\tau} p_\tau(\alpha) - \widehat{c}}{\hat{\rho}_{1\tau} p_\tau(\alpha)} = \tag{Main Est. Eq.}$$
$$\frac{1}{d_0 + d_1 q_\tau(\alpha) + d_2 q_\tau(\alpha)^2}\Big[\frac{exp(\phi_\tau(q_\tau(\alpha)))}{1 + exp(\phi_\tau(q_\tau(\alpha)))} - \alpha - \widehat{M}_\tau(\alpha)\Big]$$
$$+ \phi'_\tau(q_\tau(\alpha))\frac{exp(\phi_\tau(q_\tau(\alpha)))}{1 + exp(\phi_\tau(q_\tau(\alpha)))}\Big(1 - \frac{exp(\phi_\tau(q_\tau(\alpha)))}{1 + exp(\phi_\tau(q_\tau(\alpha)))}\Big) + \varepsilon^g_\tau(\alpha),$$

where I have used $p_{i\tau} = t_{i\tau}/q_{i\tau}$ and where ε^g is measurement error arising from the mispecification in the functional forms used in estimation. Moreover, past multipliers are captured by $\widehat{M}_\tau(\alpha) \equiv \sum_{s=0}^{\tau-1}(1 - \widehat{\Gamma}^\tau_s(\alpha))$ for $s < \tau$ estimated in earlier stages and taken in τ as given. The equation is estimated via maximum likelihood under the assumption that ε^g is drawn from a Gaussian with parameters $(0, \sigma^{\varepsilon^g})$. This step in the estimation process recovers the parameters $\{\phi_\tau, d_0, d_1, d_2, \sigma^{\varepsilon^g}\}$.

Main Step 2: Buyer Types θ. Once $\Gamma_\tau(\cdot)$ and $\gamma_\tau(\cdot)$ are estimated, the consumer type $\theta_\tau(\alpha)$ is obtained from

$$ln(\widehat{\theta}_\tau(\alpha)) = \frac{1}{N_\tau}\sum_{k=1}^{N_\tau}\frac{1\{\alpha \geq k/N_\tau\}}{\widehat{\Gamma}_\tau(k/N_\tau) - k/N_\tau - \widehat{M}_\tau(k/N_\tau)}\Big[1 - \frac{\widehat{c}}{\hat{\rho}_{1\tau} p_\tau(k/N_\tau)} - \widehat{\gamma}_\tau(k/N_\tau)\Big], \tag{10}$$

for $\alpha \in [0, (N_\tau - 1)/N_\tau]$ and where N_τ is the total count of buyers of tenure τ. The estimator for $\theta'_\tau(\alpha)$ is

$$\widehat{\theta'}_\tau(\alpha) = \frac{\widehat{\theta}_\tau(\alpha)}{\widehat{\Gamma}_\tau(\alpha) - \alpha - \widehat{M}_\tau(k/N_\tau)}\Big[1 - \frac{\widehat{c}}{\hat{\rho}_{1\tau} p_\tau(\alpha)} - \widehat{\gamma}_\tau(\alpha)\Big], \tag{11}$$

and corresponding density function $\widehat{f}_\tau(\theta(\alpha))$ is $1/\widehat{\theta'}_\tau(\alpha)$.

Main Step 3: Base Marginal Return $v'(\cdot)$ and Return Function $v(\cdot)$. The derivative of the tariff rule links the base marginal return with the marginal tariff and the consumer type: $v'(q_\tau(\alpha)) = T'_\tau(q_\tau(\alpha))/\theta_\tau(\alpha)$. Therefore, an estimator for the base marginal return is

$$v'\widehat{(q_\tau(\alpha))} = \frac{\hat{\rho}_{1\tau} p_\tau(\alpha)}{\widehat{\theta}_\tau(\alpha)} \tag{12}$$

and $v(\cdot)$ is estimated by

$$v(q_\tau(\alpha)) = \widehat{T}_\tau(q_\tau(0)) + \frac{1}{N_\tau}\sum_{k=1}^{N_\tau} v'(\widehat{q_\tau(k/N_\tau)})1\{\alpha \geq k/N_\tau\}. \tag{13}$$

These semi-parametric estimators for the return function are thus constructed to be consistent with the parametrization of the LE multipliers.

Parametrization of $v(\cdot)$ for Welfare and Counterfactual Analysis. To calculate pair-specific efficient (first-best) quantities, I require estimates of buyer types θ, baseline marginal returns $v'(\cdot)$, and seller marginal costs c. However, the range of optimal quantities may fall outside the range of realized quantities, potentially rendering baseline marginal returns undefined for certain values. To address this issue in the welfare and counterfactual analyses, I parametrize seller-specific marginal return functions, $v(\cdot)$, as $v(q) = kq^\beta$, where $k > 0$ and $\beta \in (0,1)$. These functions are then estimated for each seller-year using linear least squares with the estimated semi-parametric marginal returns, $\widehat{v'(\cdot)}$. Specifically, I estimate:

$$\ln(\widehat{v'}_{i\tau}) = \ln(k\beta) + (\beta - 1)\ln(q_{i\tau}) + \varepsilon_{i\tau},$$

using observations for buyer i and tenure τ and where $\varepsilon_{i\tau}$ is a Gaussian error term.

6 Estimation Results and Model Fit

In this section, present the estimates of primitives and auxiliary functions of the model, and show the data fit.[32] The results are shown pooling all sellers together but the estimation is conducted at the seller-year level. My model relies on the following seller-dependent ingredients: the initial distribution of private types θ, the base return function $v(\cdot)$, and the limited enforcement multipliers $\Gamma_\tau(\cdot)$ for tenure $\tau \in \{0, 1, \ldots, 4, 5\}$.

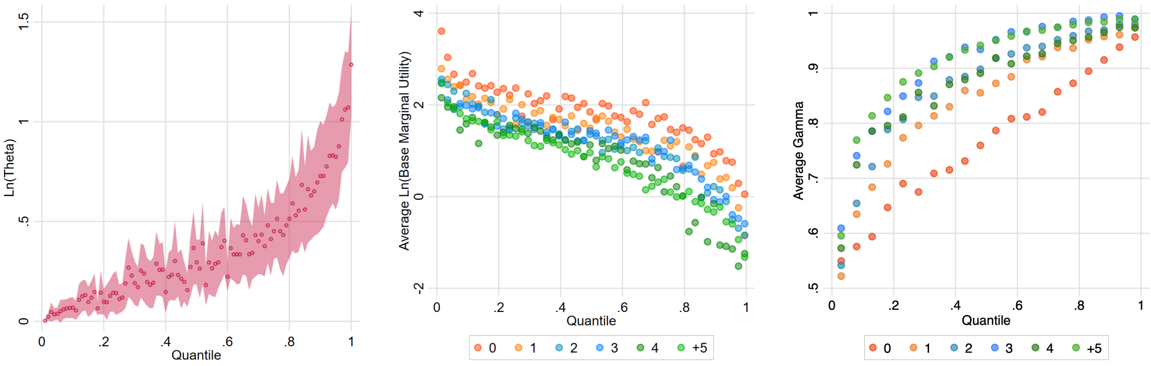
First, Figure 2a shows the average estimated log type θ by quantile of quantity for tenure 0, with error bars showing the dispersion across sellers for a given quantile.[33] The figure illustrates that, on average across sellers, types tend to increase with the quantity purchased, with a more significant increase in the top quantiles of quantities.

Next, Figure 2b plots the average estimated base marginal return $v'(\cdot)$ by quantity quantile and relationship tenure. Consistent with the model, the base marginal return function $v'(\cdot)$ decreases as quantity increases for all tenures. Additionally, the figure reveals that the functions $v'(\cdot)$ for older tenures shift downwards for many quantiles, reflecting the greater consumption levels as time goes by. These patterns suggest that the parametrization of the multiplier function

[32] Model fit of the tariff function is available in Online Appendix Section OA-8.0.1, while estimates of the survival functions by tenure in Online Appendix Section OA-8.1.

[33] For seller-year estimates of the distribution of types per seller-year, with confidence intervals constructed via bootstrap, refer to the Online Appendix Section OA-8.6.

Figure 2: **Estimated Primitives and Auxiliary Functions**



(a) Type (b) Base Marginal Return (c) Average Multiplier Γ_τ

Notes: Sub-figure (a) shows the average log type $ln(\theta)$ by quantile of quantity, across-sellers, with error bars representing the dispersion of ±1.96 standard errors for each quantile across sellers. Sub-figure (b) displays the average base marginal returns, across-sellers, for different estimation tenure groups, by quantile of quantity. Sub-figure (c) presents the average estimated limited enforcement multiplier by tenure and quantile of quantity, across-sellers.

rather than the base return function provides sensible results. Moreover, the estimated values have a clear economic interpretation, as $v'(\cdot)$ represents the marginal revenue for the buyer of an extra unit of the good for a given type. For the median new buyer (respectively, tenure 5), spending one dollar on manufacturing the good generates 2.5 (1.25) dollars of revenue for the buyer (Online Appendix Figure OA-17), which suggests that inefficiencies are more prevalent in new relationships than in older ones.

Since the buyer is purchasing inputs using trade-credit, it is possible to translate the figures into the marginal product of capital (MPK) per dollar price of credit (interest rate). The MPK measures the return the buyer would receive if given an extra unit of the input at their transaction price. I find a wedge of 40% between MPK and the transaction price for the median new relationship and 34% for the median tenure 5 relationship. Although these wedges are smaller than the gaps of 80% estimated for Indian firms by Banerjee and Duflo (2014), they are larger than the average gaps of 6% calculated by Blouin and Macchiavello (2019) in the international coffee market.[34]

Finally, Figure 2c presents the average estimated limited enforcement multiplier $\Gamma_\tau(\cdot)$. The figure indicates that almost all new pairs are constrained, as the average multiplier $\Gamma_0(\cdot)$ equals only 1 for the top 1% of pairs, on average across sellers. However, as time goes by, the average multiplier approaches 1 for lower quantiles of trade, suggesting that the limited enforcement constraint becomes less restrictive over time.[35]

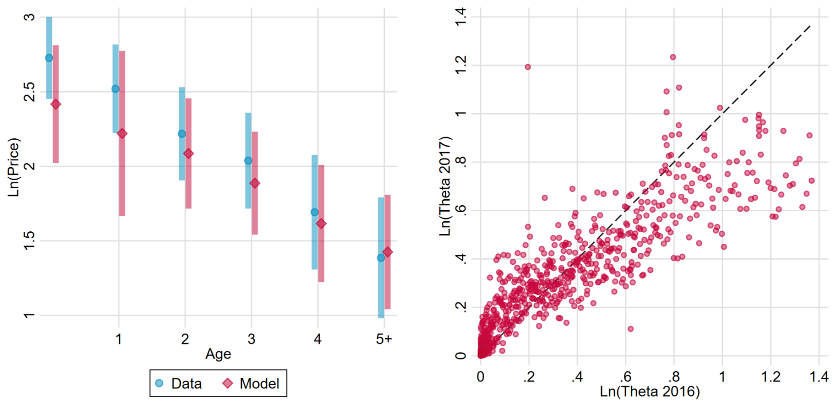
[34]It is important to note that the estimated gaps for micro-enterprises are even greater, ranging from 300% to 500% in Mexico (McKenzie and Woodruff, 2008). However, since the median buyer in my sample has total yearly sales of USD 200,000, they cannot be directly compared to micro-enterprises.

[35]The estimates for the multiplier allows us to test the model against the standard asymmetric information

6.1 Model Fit

I use five different measures to assess the fit of the model. First, the model has good statistical fit across tenures (Online Appendix Figure OA-18). While the fit does deteriorate over time and propagation bias is evident from a one-sided dispersion in the moment condition, it remains reasonable across tenures, with an average R-squared of 0.51 at tenure 0 and 0.42 at tenure 5. Second, I compare the observed quantities with model-predicted quantities. The predicted quantities, obtained using the closed-form solution of the seller's first-order condition under the parametrization of $v(\cdot)$, match well with the observed quantities across all tenures (Online Appendix Figure OA-19). Third, using the predicted quantities and the incentive-compatible tariff function (t-RULE), I generate predicted tariffs. The model-generated tariffs match the observed tariffs well across tenures (Online Appendix Figure OA-20).

Figure 3: **Non-targeted Moments**



(a) Price Discounts (b) Type in 2016 v. 2017

Notes: Sub-figure (a) presents a plot of unit prices by tenure over time using a binscatter plot, comparing prices in the data with model-generated prices. Model-generated unit prices are calculated by dividing model-generated tariffs by model-generated quantities. The error bars represent 95% confidence intervals, with standard errors clustered at the seller-year level. Sub-figure (b) shows the estimated types θ in 2017 plotted against those estimated in 2016, for buyer-seller pairs that appear in both years. These estimates were obtained through separate seller-specific estimations for each year using cross-sectional variation only. The dashed line represents the 45 degree line.

Fourth, I compare the non-targeted observed cross-sectional unit price discounts by tenure to those generated by the model in Figure 3a. The model replicates the observed discounts quite well.

To validate the model's within-pair dynamics, I consider a fifth validation exercise. I use the panel structure to verify that the primitives of the model are similar over time within pairs. Given that the model is estimated using cross-sectional information for each seller separately in 2016 and 2017, Figure 3b shows the value of estimated $\widehat{\theta}$ in 2017 against the value of estimated

model. Online Appendix Table OA-8.2 displays the distribution of t-statistics for the LE multiplier at tenure 0 (Γ_0) to test against the null hypothesis of a standard model. Based on the t-statistics, I reject the null that the standard non-linear pricing model applies in my setup for 86% of the markets (seller-years).

$\widehat{\theta}$ in 2016 for pairs that are active in both years. The figure illustrates a good correspondence between both estimated values, with the markers overlaying the diagonal in the graph. This result helps validate both the estimation procedure, as similar results are obtained via two independent estimation processes, and the persistency assumption for the types.

7 Welfare and Counterfactuals

In this section, I analyze the efficiency of relationships over time using the estimated model, while relying on the parametric estimates of $v(\cdot)$ (Online Appendix Table OA-13). Additionally, I evaluate the welfare performance of different pricing and enforcement schemes. I focus on three margins: (a) perfect enforcement with non-linear pricing, (b) limited enforcement with uniform pricing, and (c) perfect enforcement with uniform pricing.

7.1 Efficiency Relative to First-Best

Under the parametrization $v(q) = kq^{\beta}$, the first-best quantities for each pair are given by:

$$q^{fb}(\theta) = \left(\frac{k\beta\theta}{c}\right)^{1/(1-\beta)}. \tag{14}$$

Moreover, total surplus is a function of the buyer's type θ, quantity q, and seller's marginal cost c: $Surplus(\theta, q, c) = \theta kq^{\beta} - cq$. Hence, the static efficiency of allocation q for buyer type θ is defined as:

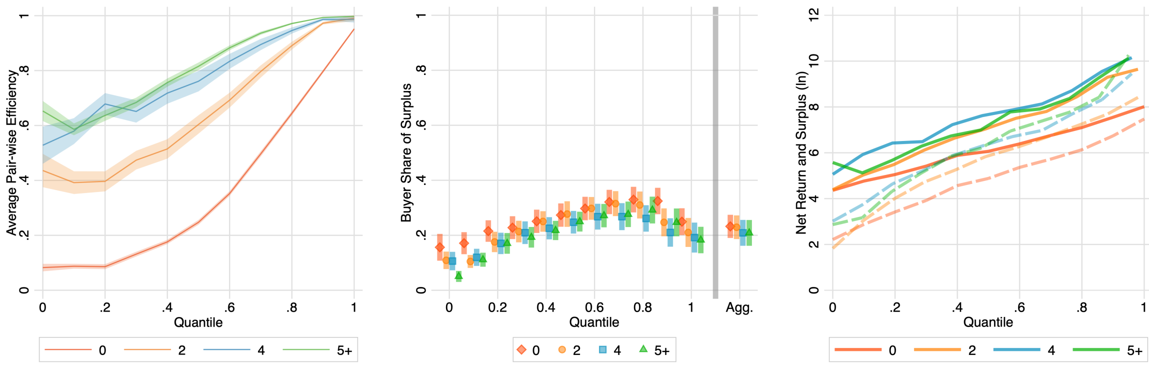
$$\text{Efficiency}(\theta, q, c) = \frac{Surplus(\theta, q, c)}{\max_q Surplus(\theta, q, c)}.$$

Figure 4a plots the average efficiency for each tenure across quantity deciles, averaging over all pairs, excluding tenure 1 and 3 for clearer visualization. The figure shows that new relationships are severely constrained, with the median buyer trading at only around 30% of their optimal level. However, as relationships age, efficiency increases. The median buyer trades at 60% of optimal levels at tenure 2, 75% at tenure 4, and over 80% at tenure 5. Additionally, the figure demonstrates significant heterogeneity in traded efficiency within relationship age: partners trading little experience greater distortions than partners trading more intensively.

While the general theoretical model does not yield precise estimates, the observed patterns of efficiency increasing with age and quantities can be explained by two key features of the model. First, to maintain incentive compatibility, higher types must receive higher quantities, resulting in lower distortions at the top compared to the bottom. Second, as trade is initially constrained and quantities are backloaded, increasing over time, efficiency is expected to increase.[36] However, the model does not mechanically imply an increase in efficiency over time.

[36]The backloading of prices and quantities is primarily concentrated among lower types, which helps explain the high efficiency of high types early on (Online Appendix Figure OA-5).

Figure 4: **Efficiency and Buyer Surplus**



(a) Efficiency (b) Buyer Surplus (c) Surplus & Net Return

Notes: Sub-figure (a) presents average efficiency by quantile of quantity and tenure across all buyer-seller pairs. Error bars show dispersion of ± 1.96 standard errors for each quantile across pairs. Sub-figure (b) shows average buyer share of surplus for quantile of quantity and tenure across all sellers. Error bars show ± 1.96 standard errors, clustered at the seller-year level. Sub-figure (c) plots average (log) surplus in solid lines and average (log) buyer net return in dashed lines by quantile of quantity and tenure, averaging across sellers.

If initial trade levels were close to or above efficient levels, higher quantities would lead to inefficient trade.

Of course, this characterization of efficiency might be too strict if the majority of trade is channeled through large buyers. To account for intensity-inclusive efficiency, I study the weighted average efficiency of all transactions per seller. This approach considers the potential efficiency losses and constructs weights using the share of total efficient quantities at a given tenure. Under this measure, the total output is inefficient early on but converges towards efficiency in the medium and long term. In Panel (a) of Table 1, I report the share of sellers trading at efficient levels, both in average total output and with the average buyer.[37] The results indicate that only 5% of sellers are trading efficiently with new buyers, but efficiency increases quickly, with 70% of sellers trading efficiently by tenure 2. In the long term, 84% of sellers transact with their buyers at efficient levels.

To better understand the long-term efficiency of relationships across different selling sectors, I present the share of sellers trading at aggregate efficient levels in Panel (b) of Table 1. While at the beginning of relationships almost no seller is trading efficiently, efficiency levels start to diverge at tenure 2. Starting at this point, Textiles shows slower growth in efficiency, while Pharmaceutical and Cement-Products continue to improve. By tenure 5, almost all Pharmaceutical and Cement-Products sellers are trading at aggregate efficient levels, while 70% of Textiles sellers do so. Despite this heterogeneity across sectors, the general takeaway is clear: even in different sectors, aggregate trade efficiency is high in the medium and long term.

[37] I test for seller-level efficiency via 30 bootstrap simulations and consider a seller's output efficient if the 95th percentile of weighted surplus is within 1% of efficiency.

Table 1: **% Share of Sellers with Efficient Trade**

	Tenure					
	0	1	2	3	4	5
Panel (a): All Sectors						
Weighted	5	41	70	79	75	84
Unweighted	5	23	32	37	38	30
Panel (b): Weighted, By Sector						
Textiles	6	45	59	64	64	68
Pharmaceutical	0	31	73	88	73	88
Cement-Products	13	50	75	87	87	95

Notes: This table reports the share of sellers that trade efficiently. Panel (a) presents results across all sectors. The first measure (Weighted) computes the share of sellers whose weighted average output cannot be rejected to be different from the efficient output at the 10% level. The weights are constructed over potential output for each seller-tenure. The second measures (Unweighted) computes the share of sellers for which the surplus created by the average buyer cannot be rejected to be different from efficient at the 10% level. Panel (b) presents results using the Weighted measure for each selling sector.

To provide a benchmark for the estimated inefficiencies due to imperfect contracting, it is helpful to compare these results to previous estimates in the literature. While the specific settings and frictions may vary, this comparison offers valuable insights. For instance, Blouin and Macchiavello (2019) find that strategic default reduces output by 16% for the mean relationship, with only 26% of relationships operating at first-best. Similarly, Ryan (2020) finds that weak contract enforcement reduces efficiency by 10% on average, while Startz (2024) finds that jointly contracting and search frictions reduce welfare by 9%. In contrast, the results presented in this paper offer a dire look at the relationship level, with average output at only 38% of first-best. However, when weighting for the size of relationships, the estimated inefficiencies are more moderate and in line with the literature, with a weighted average loss of 15%. It is worth highlighting that the previous studies only estimate efficiency for stationary relationships, whereas this paper offers efficiency estimates over the lifespan of a relationship. Additionally, these magnitudes of relationship-level inefficiencies may not be specific to developing countries, as contemporaneous work by Harris and Nguyen (2022) finds that the median relationship in the US trucking industry achieves only 44% of first-best output.

To analyze surplus division, I present Figure 4b. This figure displays the average share of surplus captured by buyers, across sellers, by bins over quantiles of quantity purchased at different tenures. The results show that sellers capture the majority of the surplus, with the median buyer in any tenure capturing around 25 percent of the generated surplus. The figure also reveals that, consistent with the non-linear pricing scheme, buyers who trade more intensively tend to capture a larger share of the surplus, up to 35 percent. However, the smallest buyers may capture less than 10 percent of the total surplus. In the aggregate, sellers capture an average of

80% of all surplus created, and this share is relatively constant over time. The combination of results showing that (1) sellers capture the majority of surplus and (2) sellers have the ability to extract different levels of surplus across different buyers can be seen as evidence that sellers indeed have market power in this setting.

The general flattening of the buyer share of surplus for the highest types does not reflect that middle types obtain greater net returns. Indeed, Figure 4c shows the net return of buyers in dashed lines as well as the total surplus in solid lines. Both total surplus and buyer's net return increase with quantile, with higher types obtaining higher net returns within tenure, in line with the requirements for incentive compatibility. Moreover, the total amount of net return captured by buyers grows over time. Instead, the non-linearity in the buyer share of surplus reflects the underlying distribution of types. In simulations not shown here, I find that extreme-valued distributions show the non-linear pattern in the buyer share of surplus, while for uniform distributions the buyer share of surplus increases monotonically with quantile. This is because the surplus at the highest types is growing faster than the amount of net return received.

A similar intuition helps explain why the aggregate buyer share of surplus is relatively constant over time. In particular, within a given type, if quantities increase relatively faster than prices decrease, the share of surplus can be kept constant or even decrease.

7.2 Counterfactuals

Next, I use the estimated model to explore the implications of improving the enforcement of trade-credit contracts and enforcing current Ecuadorian legislation that forbids price discrimination on identical transactions. I consider three counterfactual scenarios, explained below. Details on the computations of each counterfactual are provided in Online Appendix Section OA-9.1.

Counterfactual (a): Non-linear pricing with perfect enforcement. One natural question is to consider what the surplus would be in a world of perfect enforcement of contracts, mimicking a policy that improves court efficiency. This is implemented by allowing the seller to use non-linear pricing but forbidding the buyer to default.

Counterfactual (b): Uniform pricing with limited enforcement. Alternatively, one may address other frictions in the model. While asymmetric information is a friction generating distortions relative to the first-best (the seller distorts quantities for some buyers to incentivize the revelation of private information), another key friction is the ability of the seller to charge prices above marginal costs. Absent enforcement constraints, if the seller were not able to charge prices above marginal costs, trade would be efficient under incomplete information too. Thus, market power expressed as prices over marginal costs generates distortions. For a policymaker, policies addressing pricing power may be easier to design and enforce than policies addressing pair-specific information asymmetry.

Therefore, I consider a counterfactual policy aimed at addressing market power. Written law in Ecuador, the European Union, and the US forbids price discrimination that applies differential treatment to customers performing an otherwise equivalent transaction, including possibly preferential treatment due to tenure.[38] Under the model assumptions (constant marginal costs), any price discrimination would be unlawful and thus of interest to a policymaker as well. As such, this counterfactual studies the welfare effects of a policy that enforces uniform pricing but keeps the limited enforcement regime active.

Counterfactual (c): Uniform pricing with perfect enforcement. Lastly, I consider addressing both market power and enforcement.[39] The policy forbids price discrimination by enforcing uniform markups and forbids buyer default.

Discussion of Counterfactual Results. Table 2 shows the counterfactual results, displaying average surplus (as a percentage of the baseline) for each percentile group—formed by grouping quantiles of quantity—and tenure, and aggregate results weighted by observed quantities for each tenure.[40]

Counterfactual (a): Non-linear pricing with perfect enforcement. Panel (a) shows the results. The policy exercise generates an inter-temporal trade-off for middle and low types, as fixing enforcement generates massive gains for them in the early stages of the relationship. That is, weak enforcement forces the seller to create further downward distortions for low- and middle-types when buyers can default on trade. Fixing enforcement alone would increase surplus for 75% of the buyers in tenure 0 and 1. However, as relationships age, contract enforcement distortions become of second order. By tenure 3 and onward, limited enforcement contracts actually help discipline the downward distortions from non-linear pricing by the seller. Fixing enforcement would decrease the generated surplus in old relationships for essentially all buyers, as the seller increases quantities over time to incentivize debt repayment from the buyer side. In the long term, the threat of default is sufficient to overcome sellers' downward output distortions

[38]In Ecuador, Art. 9 of *Ley Orgánica de Regulación y Control del Poder de Mercado*. In the EU, Art. 102(c) of *Treaty on the Functioning of the European Union* (ex of Art. 82(c) of. *EC Treaty*). In the US, Section 2(a) of the *Robinson-Patman Act*. In practice, only the EU has enforced such a law in court. See, for instance, the cases *Hoffmann-La Roche v. Commission* and *Manufacture française des pneumatiques Michelin v Commission*. In the US, *some* variants of preferential pricing (such as loyalty discounts in multiproduct markets) have been upheld in court. See, for instance, cases *LePage's v 3M* and *SmithKline v Eli Lilly*. Moreover, in the US, discounts below cost are seen as anticompetitive (see *Eisai Inc. v. Sanofi-Aventis U.S., LLC*). In Ecuador, no cases have been brought to court regarding the specific Art 9.

[39]There are instances of real-world examples of policy reforms aimed at addressing payment enforcement and market power jointly or concurrently. For instance, the U.K. demonstrates concurrent independent reforms. The Late Payment of Commercial Debts (Interest) Act (1998, Amended 2013) addresses late payments by imposing penalties and establishing enforcement mechanisms. Separately, Section 18 of the Competition Act 1998 prohibits unequal treatment of equivalent transactions, targeting market power abuse. Conversely, the U.S. exemplifies joint reforms. The Packers and Stockyards Act of 1921 encompasses both payment enforcement and competition concerns. It mandates prompt payment for livestock sellers, prohibits unfair pricing practices that favor certain trading partners, and empowers the Department of Agriculture to enforce these provisions.

[40]Additional results for the three counterfactual exercises related to buyer net return, profits, and prices are presented in Online Appendix Section OA-9.2, where the table reports the percentage of observations in the baseline with a greater value in the specific category (e.g., prices) relative to the counterfactual.

Table 2: **Average Surplus as % of Baseline**

	10%	25%	50%	75%	100%	Agg.	10%	25%	50%	75%	100%	Agg.
	Panel (a): Non-linear + Perfect Enforcement						*Panel (c): Uniform + Perfect Enforcement*					
Tenure 0	1,508.4	1,419.0	628.0	150.3	56.5	67.9	46,633.7	42,233.1	8,487.5	1,083.0	64.0	192.6
Tenure 1	430.3	430.6	256.0	112.0	49.8	64.7	13,887.9	12,003.0	8,472.0	649.7	49.4	337.1
Tenure 2	164.8	139.9	102.6	59.7	44.2	46.7	5,399.0	4,161.9	1,531.8	97.7	35.9	75.3
Tenure 3	80.5	82.7	68.6	53.4	43.2	44.9	1,816.5	1,198.1	417.5	63.3	33.5	51.0
Tenure 4	72.4	72.7	67.9	54.0	45.2	47.9	745.0	624.2	294.0	60.8	35.1	53.5
Tenure 5	60.7	66.4	60.2	53.9	47.0	48.7	224.6	195.7	112.2	49.9	36.8	42.7
	Panel (b): Uniform + Limited Enforcement						*% Excluded*					
Tenure 0	1.0	1.3	1.5	2.1	3.2	3.1	97.3	96.4	95.8	94.1	90.5	90.9
Tenure 1	2.8	4.0	5.8	5.7	5.3	5.4	93.4	91.9	88.6	87.3	85.8	86.1
Tenure 2	12.2	14.1	18.4	16.7	15.4	15.5	81.5	77.8	70.1	65.7	61.3	61.9
Tenure 3	16.9	19.4	26.6	23.0	19.4	19.9	76.9	69.0	59.5	51.5	50.0	50.4
Tenure 4	17.7	25.3	33.4	28.9	24.6	23.6	66.8	58.1	47.5	44.7	43.5	50.0
Tenure 5	28.6	37.9	43.5	34.0	29.2	30.4	65.3	58.8	37.5	29.8	25.4	26.7

Notes: This table presents average efficiency measures as % of baseline (non-linear price with limited enforcement) of different pricing and enforcement regimes by percentile groups of quantity and tenure. Percentile groups are defined based on quantiles as follows: the 10% group includes all buyers within seller-year-tenure quantiles from 0 to 10% (non-inclusive), the 25% group includes buyers within quantiles from 10% to 25% (non-inclusive), and this pattern continues for all other percentile groups. Panel (a) reports results for non-linear pricing with perfect enforcement. Panel (b) reports optimal monopolistic uniform price with limited enforcement, with the subpanel reporting the share of excluded buyers in this counterfactual. Panel (c) reports results for optimal monopolistic uniform price with perfect enforcement. No buyer is excluded in Panels (a) and (c).

from prices above marginal costs.

For higher types, however, the policy is always welfare-reducing. To see why, consider equation 3 and the estimates for the multiplier $\Gamma_\tau(\cdot)$, which are less than 1 except for the highest type. Shutting down enforcement constraints sets $\Gamma_\tau(\cdot) = 1$ for all buyers. Thus, for any type such that $\Gamma_\tau(\theta) < 1$, total trade would tend to decrease. Furthermore, any past promise to increase trade, captured in past multipliers, would also disappear.

As a result, given that higher types trade efficiently across the board, the policy has an aggregate negative welfare effect, though the welfare losses are smaller for earlier periods, partially reflecting the inter-temporal trade-off of lower and middle types.

Counterfactual (b): Uniform pricing with limited enforcement. Panel (b) presents the results. The surplus ranges from 0 to 40 percent of the baseline surplus across time and types. The surprisingly low performance of this alternative regime is explained by the large share of buyers that would be excluded from trade, as some buyers cannot credibly commit to repaying their debts and the seller cannot use dynamic incentives to discipline their behavior. Thus, in the presence of limited enforcement, the seller's ability to price discriminate actually improves the situation for both buyers and sellers by increasing the share of buyers that can be credibly incentivized not to default. In the aggregate, results are similar: efficiency is extremely low but it increases over time, reflecting the positive selection of types.

These results on inefficiency hold even if prices are identical to marginal costs. The seller's ability to target each individual buyer's enforcement constraint through differentiated prices and quantities allows them to prevent default.

Counterfactual (c): Uniform pricing with perfect enforcement. Panel (c) reports the results.

The table shows that surplus increases relative to the baseline, except for the highest types. Welfare gains are concentrated among the lowest types (who see gains of up to 46,000%), although even median types also see large increases (from 12% up to 8,000%). Higher types, however, are negatively affected by the policy. Under a uniform markup, prices tend to be higher than in the baseline (Appendix Table OA-14), and consumption is now determined solely by prices, decreasing the quantity consumed by higher types and thus reducing the generated surplus relative to the baseline. This is reflected in aggregate surplus: as the policy does not improve efficiency for higher types, aggregate surplus under the policy is higher than under the baseline in the early stages of the relationships but lower in the medium and long term. The intuition for this result is simple: higher types face greater distortions under the constant monopolist markup than otherwise.

Given the large inter-temporal trade-off, the macro effect of this policy depends significantly on the weights assigned to each tenure. If the weights are based on the number of buyers, the policy improves welfare, with gains of approximately 40% relative to the baseline, as early tenures involve a larger number of buyers. Conversely, if the weights are derived from the quantities at baseline, the policy reduces welfare, resulting in a surplus of only 58% relative to the baseline, as older tenures carry greater importance.

Since this counterfactual allows the seller to set the monopolist's uniform price, a policy that addresses multiple frictions simultaneously would perform better with stronger measures to curb seller market power by further reducing markups.

8 Conclusion

This paper demonstrates that frictions in the manufacturing supply chain significantly affect long-term relationships. The novel theoretical model shows that limited enforcement constraints compel the seller to offer a larger net return to the buyer than under perfect enforcement, which in turn distorts trade inter-temporally by promising larger future quantities at lower prices to boost current profits. Using a unique intra-national trade database from Ecuador, I estimate a structural model of relational contracting with seller market power and quantify the efficiency of dynamic trade. The results reveal that although trade is initially inefficient, transacted quantities approach full efficiency in the long run despite the seller's market power, highlighting both the value and fragility of informal relational contracts. These findings suggest that unilateral reforms aimed at improving enforcement or modifying antitrust policies could inadvertently undermine long-term efficiency, whereas addressing multiple frictions simultaneously may yield significant welfare gains.

APPENDIX

A Proof of Proposition 1: *Model's First-Order Conditions*

Here, I walk through the characterization of the firm's problem subject to the constraints, deriving Proposition 1.

A.1 Relaxed Problem for Incentive-Compatibility

First, I focus on the relaxed problem, which replaces the global incentive compatibility constraints IC-B with a dynamic envelope formula. Specifically, any implementable dynamic incentive-compatible menu must satisfy (Theorem 1, Pavan et al., 2014):

$$\sum_{\tau=0}^{\infty} \delta(\theta)^{\tau} u_{\tau}'(\theta) = \sum_{\tau=0}^{\infty} \delta(\theta)^{\tau} v(q_{\tau}(\theta)), \tag{15}$$

for any arbitrary function $0 < \delta(\theta) < 1$ and $u_{\tau}'(\theta) \equiv du_{\tau}(\theta)/d\theta$. Substituting the envelope condition 15 with $\delta(\theta) = \delta$ into the seller's problem SP and integrating by parts yields:

$$\sum_{\tau=0}^{\infty} \delta^{\tau} \int_{\underline{\theta}}^{\overline{\theta}} \left[s(\theta, q_{\tau}(\theta), c_{\tau}) - \int_{\underline{\theta}}^{\theta} v(q_{\tau}(x))dx \right] f_{\tau}(\theta) d\theta - \sum_{\tau=0}^{\infty} \delta^{\tau} u_{\tau}(\underline{\theta}). \tag{16}$$

The return term of the buyer acknowledges the rents that must be given to higher types to preserve incentive compatibility.

It is well known that the solution to the full program might not match the solution to the relaxed program, as the dynamic envelope condition is only a necessary condition (Stantcheva, 2017). However, if the optimal contract is strictly monotonic (i.e., those with $q_{\tau}'(\theta) > 0$ for all θ and τ) for fully persistent types, then the contract is globally incentive compatible (Battaglini and Lamba, 2019).

A.2 Limited Enforcement Constraints in the Relaxed Problem

I write the problem in a Lagrangian-type form (in the spirit of the static problem in Jullien (2000)). For this formulation, the dynamic LE-B constraint for type θ at time τ is given by:

$$\left\{ \sum_{s=1}^{\infty} \delta^{s} (1 - X(\theta))^{s} u_{\tau+s}(\theta) - [\theta v(q_{\tau}(\theta)) - u_{\tau}(\theta)] \right\} \gamma_{\tau}(\theta) = 0, \tag{17}$$

where $\gamma_{\tau}(\theta)$ is the corresponding limited enforcement Lagrange (LE) multiplier for type θ's enforcement constraint at time τ. The LE multiplier is positive ($\gamma_{\tau}(\theta) > 0$) whenever the limited enforcement constraint binds, capturing the shadow value of the enforcement constraint for θ. To include the constraint across types, we integrate over all types to obtain:

$$\int_{\underline{\theta}}^{\overline{\theta}} \left\{ \sum_{s=1}^{\infty} \delta^{s} (1 - X(\theta))^{s} u_{\tau+s}(\theta) - [\theta v(q_{\tau}(\theta)) - u_{\tau}(\theta)] \right\} d\Gamma_{\tau}(\theta) = 0, \tag{Lagrangian-D-LE}$$

where $\Gamma_{\tau}(\theta) = \int_{\underline{\theta}}^{\theta} \gamma_{\tau}(x)dx$ is the *cumulative* LE multiplier with derivative $\gamma_{\tau}(\theta)$. The cumulative LE multiplier $\Gamma_{\tau}(\theta)$ captures the extent by which trade is distorted by limited enforcement. It represents the shadow value of relaxing the enforcement constraints uniformly from $\underline{\theta}$ to θ, capturing the amount of profits lost by the seller due to enforcement incentives.

The cumulative multiplier has the properties of a cumulative distribution function. Extending θ increases the set on which the enforcement constraint is relaxed, so Γ_τ is nonnegative and nondecreasing. By relaxing the constraints uniformly, the seller can reduce the buyers' net returns by keeping quantities unchanged, hence $\Gamma_\tau(\overline{\theta}) = 1$.[41]

After manipulating the limited enforcement constraints, one can obtain the full Lagrangian maximand:

$$\sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\overline{\theta}} \left[s(\theta, q_\tau(\theta), c_\tau) - v(q_\tau(\theta)) \frac{\Gamma_\tau(\theta) - F_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s^\tau(\theta))\tilde{\Gamma}_s^\tau(\overline{\theta}) + \theta\gamma_\tau(\theta)}{f_\tau(\theta)} \right] f_\tau(\theta) d\theta, \tag{18}$$

with the corresponding slackness condition Lagrangian-D-LE where $\Gamma_s^\tau(\theta)$ is the conditional cumulative LE multiplier constraint defined by:

$$\Gamma_s^\tau(\theta) = \frac{\int_{\underline{\theta}}^{\theta} (1 - X(x))^{\tau-s} \gamma_s(x) dx}{\tilde{\Gamma}_s^\tau(\overline{\theta})}, \tag{19}$$

for $\tilde{\Gamma}_s^\tau(\overline{\theta}) = \int (1-X(\theta))^{\tau-s}\gamma_s(\theta) d\theta$.[42] The conditional cumulative multiplier constraint adjusts for the likelihood that a given θ has survived $\tau - s$ periods, assigning lower weights to θs that are less likely to survive.

A.3 Relaxing the Double-Deviation Constraint

The problem is further relaxed by omitting the Double-Deviation Constraint DD-B. This is sensible as both IC-B and LE-B are necessary conditions for the constraint.

First, to see that DD-B implies IC-B, consider the limit as $\tau \to \infty$:

$$\begin{aligned} \sum_{\tau=0}^{\infty} \delta(\theta)^\tau u_\tau(\theta) &\geq \lim_{\tau\to\infty} \left\{ \delta(\theta)^\tau \theta v(q_\tau(\widehat{\theta})) + \sum_{s=0}^{\tau-1} \delta(\theta)^s \left[\theta v(q_s(\widehat{\theta})) - t_s(\widehat{\theta}) \right] \right\} \forall \theta, \widehat{\theta}, \\ &\geq \sum_{s=0}^{\infty} \delta(\theta)^s \left[\theta v(q_s(\widehat{\theta})) - t_s(\widehat{\theta}) \right] \forall \theta, \widehat{\theta}, \end{aligned} \tag{21}$$

thus IC-B is a necessary condition for DD-B.

---

[41] In Online Appendix Section OA-5, I show formally that $\Gamma_\tau(\overline{\theta}) = 1$.

[42] Pre-multiply each constraint by δ^τ and sum over τ. Reorder internal summations, substitute in the dynamic envelope condition, and eliminate constant terms to obtain:

$$\begin{aligned} &\sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} v(q_\tau(x)) dx \sum_{s=0}^{\tau-1} (1 - X(\theta))^{\tau-s} d\Gamma_s(\theta) \\ &- \sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\overline{\theta}} \left[\theta v(q_\tau(\theta)) - \int_{\underline{\theta}}^{\theta} v(q_\tau(x)) dx \right] d\Gamma_\tau(\theta). \end{aligned} \tag{20}$$

Then integrate by parts.

Second, to see that DD-B implies LE-B, simply set $\widehat{\theta} = \theta$:

$$\sum_{\tau=0}^{\infty} \delta(\theta)^{\tau} u_{\tau}(\theta) \geq \delta(\theta)^{\tau} \theta v(q_{\tau}(\theta)) + \sum_{s=0}^{\tau-1} \delta(\theta)^{s} \left[\theta v(q_s(\theta)) - t_s(\theta)\right] \forall \theta, \tau \Leftrightarrow \tag{22}$$

$$\sum_{s=\tau+1}^{\infty} \delta(\theta)^{s} u_s(\theta) + \delta(\theta)^{\tau} u_{\tau}(\theta) \geq \delta(\theta)^{\tau} \theta v(q_{\tau}(\theta)) \forall \theta, \tau \Leftrightarrow \tag{23}$$

$$\sum_{s=1}^{\infty} \delta(\theta)^{s} u_{\tau+s}(\theta) \geq t_{\tau}(\theta) \forall \theta, \tau. \tag{24}$$

Therefore, LE-B is a necessary condition for DD-B.

Furthermore, for any $\widehat{\theta}$ such that:

$$\delta(\theta)^{\tau} \theta v(q_{\tau}(\widehat{\theta})) + \sum_{s=0}^{\tau-1} \delta(\theta)^{s} \left[\theta v(q_s(\widehat{\theta})) - t_s(\widehat{\theta})\right] < \delta(\theta)^{\tau} \theta v(q_{\tau}(\theta)) + \sum_{s=0}^{\tau-1} \delta(\theta)^{s} u_s(\theta), \tag{25}$$

condition LE-B implies DD-B, so for such $\widehat{\theta}$ the condition DD-B is irrelevant.

For all other $\widehat{\theta}$, the condition LE-B is necessary for DD-B to hold. In particular, if DD-B holds, then:

$$\delta(\theta)^{\tau} t_{\tau}(\theta) \leq \sum_{s=\tau+1}^{\infty} \delta(\theta)^{s} u_s(\theta) - \left(\sum_{s=0}^{\tau-1} \delta(\widehat{\theta})^{s} \left[\theta v(q_s(\widehat{\theta})) - t_s(\widehat{\theta})\right] - \sum_{s=0}^{\tau-1} \delta(\theta)^{s} \left[\theta v(q_s(\theta)) - t_s(\theta)\right] \right) \forall \theta, \widehat{\theta}, \tau. \tag{26}$$

As the term in the brackets is positive by assumption, LE-B holds.

A.4 The Seller's First-Order Condition

All in all, the corresponding seller's first-order condition for the relaxed problem determining the allocation rule at any relationship tenure τ is:

$$\theta v'(q_{\tau}(\theta)) - c = \frac{\Gamma_{\tau}(\theta) - F_{\tau}(\theta) - \sum_{s=0}^{\tau-1}(1 - \Gamma_s^{\tau}(\theta)) \tilde{\Gamma}_s^{\tau}(\overline{\theta}) + \theta \gamma_{\tau}(\theta)}{f_{\tau}(\theta)} v'(q_{\tau}(\theta)). \tag{SFOC}$$

Therefore, if the quantity profile $\{q_{\tau}^{*}(\theta)\}$ maximizes lifetime profits for the firm subject to IC-B, LE-B, and DD-B, it must also satisfy SFOC.

A.5 Tariffs

Tariffs are then constructed to satisfy the dynamic envelope formula 15 for the optimal quantity profile $\{q_{\tau}^{*}(\theta)\}$ solving the seller's problem:

$$t_{\tau}'^{*}(\theta) = \theta v(q_{\tau}^{*}(\theta)) q_{\tau}'^{*}(\theta). \tag{t-RULE}$$

B Proof of Proposition 2: *Point Identification*

In this section, I detail how $\Gamma_{\tau}(\cdot)$ is point identified with observations of prices, quantities, and marginal cost for one seller under the parametrization of $v(q) = kq^{\beta}$ for $k > 0$ and $\beta \in (0, 1)$.

As a preliminary step, I state the following lemma.

Lemma 1. $\Gamma_\tau(\overline{\theta}) = 1, \forall \tau$.

The proof is relegated to Online Appendix Section OA-5. The intuition is that marginal uniform relaxation of the enforcement constraint does not optimally affect quantities across buyers but rather simply shifts the tariffs upward by the same amount. Thus, the shadow cost of a marginal uniform relaxation of the enforcement constraints is exactly the marginal uniform relaxation.

B.1 Step 1: Show β is identified

We first show that β is identified from observations on prices, quantities, and marginal cost for $\tau = 0$. In this step, we omit subscripts $\tau = 0$.

Consider $\rho(\alpha) = \partial \ln(\theta(\alpha))/\partial\alpha = \theta'(\alpha)/\theta(\alpha)$. Substituting in, the key identification equation I-EQ becomes

$$\frac{T'(q(\alpha)) - c}{T'(q(\alpha))} = \rho(\alpha)\Big[\Gamma(\alpha) - \alpha\Big] + \gamma(\alpha). \tag{27}$$

Evaluating at $\alpha = 1$ and using the fact that $\Gamma(1) = 1$ (Lemma 1), yields

$$\gamma(1) = \frac{T'(q(1)) - c}{T'(q(1))}. \tag{28}$$

Therefore, all parameters, except $\rho(\alpha)$, are known at the boundary $\alpha = 1$.

As an auxiliary result, note that:

$$\gamma'(1) = \frac{cT''(q(1))}{\Big((T'(q(1))\Big)^2}, \tag{29}$$

which is known.

Then consider the first-order condition at $\alpha = 1 - \varepsilon$ using Taylor approximations for the enforcement multipliers:

$$\frac{T'(q(1-\varepsilon)) - c}{T'(q(1-\varepsilon))} \approx \rho(1-\varepsilon)\Big[\Gamma(1) - \gamma(1)\varepsilon - 1 + \varepsilon\Big] + \gamma(1) - \gamma'(1)\varepsilon, \tag{30}$$

under the assumption that Γ is regular and second-order differentiable as it approaches $\alpha = 1$. From this equation, the value for $\rho(1-\varepsilon)$ is identified.

Using the derivative of the tariff rule I-T, obtain

$$\rho(\alpha) = \theta'(\alpha)/\theta(\alpha) = q'(\alpha)[T''(q(\alpha))/T'(q(\alpha)) + A(q(\alpha))], \tag{31}$$

where $A(q(\alpha)) = -v''(q(\alpha))/v'(q(\alpha))$. The assumed parametrization in IA 3 implies $A(q) = (1-\beta)/q$. As $T'(\cdot)$, $T''(\cdot)$, $q(\cdot)$, and $q'(\cdot)$ are known, $\rho(\cdot)$ depends on only one unknown parameter β, which is identified from the value of $\rho(1-\varepsilon)$ above.

B.2 Step 2: Show Γ_0 is identified from β

Consider equation 27 and use the parametrized version of $\rho_0(\alpha)$ for 31:

$$\Gamma_0(\alpha)+\gamma_0(\alpha)\left[q_0'(\alpha)\left(\frac{T_0''(q_0(\alpha))}{T_0'(q_0(\alpha))}+\frac{1-\beta}{q_0(\alpha)}\right)\right]^{-1} = \tag{32}$$
$$\alpha+\frac{T_0'(q_0(\alpha))-c}{T_0'(q_0(\alpha))}\left[q_0'(\alpha)\left(\frac{T_0''(q_0(\alpha))}{T_0'(q_0(\alpha))}+\frac{1-\beta}{q_0(\alpha)}\right)\right]^{-1}.$$

The LE multiplier $\Gamma_0(\alpha)$ is identified from the solution to the differential equation above using the boundary condition $\Gamma_0(1)=1$ (Lemma 1), and the fact that $T_0''(\cdot)$, $T_0'(\cdot)$, $q_0(\cdot)$, $q_0'(\cdot)$, and β are known or identified.

B.3 Step 3: Show Γ_τ is identified from β and Γ_s for $s<\tau$

To identify $\Gamma_\tau(\alpha)$, we start recursively from $\tau=1$. With knowledge of $\Gamma_s(\cdot)$ for $s<\tau$ and β, we note that from:

$$\Gamma_\tau(\alpha)+\gamma_\tau(\alpha)\left[q_\tau'(\alpha)\left(\frac{T_\tau''(q_\tau(\alpha))}{T_\tau'(q_\tau(\alpha))}+\frac{1-\beta}{q_\tau(\alpha)}\right)\right]^{-1} = \tag{33}$$
$$\alpha+\sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))+\frac{T_\tau'(q_\tau(\alpha))-c}{T_\tau'(q_\tau(\alpha))}\left[q_\tau'(\alpha)\left(\frac{T_\tau''(q_\tau(\alpha))}{T_\tau'(q_\tau(\alpha))}+\frac{1-\beta}{q_\tau(\alpha)}\right)\right]^{-1},$$

$\Gamma_\tau(\alpha)$ is identified from the solution to the differential equations above with the boundary condition $\Gamma_\tau(1)=1$ and the fact that $\Gamma_s(\cdot)$, c, $T_\tau'(\cdot)$, $T_\tau''(\cdot)$, $q_\tau(\cdot)$, $q_\tau'(\cdot)$, and β are known or identified.

B.4 Step 4: Show $\theta_\tau(\cdot)$, $\theta_\tau'(\cdot)$, and k are identified

With known multipliers and separately by tenure, I-EQ is equivalent to the non-linear pricing problem in Luo et al. (2018). Therefore, their results imply that the distribution of types is identified. The intuition for the identification result is that the incentive compatibility constraints for truthful revelation imply a monotonic relationship between quantities and types, which means the underlying distribution of unknown types can be recovered from the observed distribution of quantities.

Finally, the scale parameter k is identified I-T, as all elements are now identified or observed.

References

Abreu, Dilip, "On the theory of infinitely repeated games with discounting," *Econometrica*, 1988, pp. 383–396.

Acemoglu, Daron, Ufuk Akcigit, and William Kerr, "Networks and the macroeconomy: An empirical exploration," *Nber macroeconomics annual*, 2016, *30* (1), 273–335.

Aguiar, Mark and Manuel Amador, "Sovereign debt," in "Handbook of International Economics," Vol. 4, Elsevier, 2014, pp. 647–687.

Aguirregabiria, Victor and Aviv Nevo, "Recent developments in empirical IO: Dynamic demand and dynamic games," *Advances in Economics and Econometrics*, 2013, *3*, 53–122.

Albuquerque, Rui and Hugo A Hopenhayn, "Optimal lending contracts and firm dynamics," *Review of Economic Studies*, 2004, *71* (2), 285–315.

Alfaro-Urena, Alonso, Isabela Manelici, and Jose P Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," *Quarterly Journal of Economics*, 2022.

Alviarez, Vanessa I, Michele Fioretti, Ken Kikkawa, and Monica Morlacco, "Two-sided market power in firm-to-firm trade," Technical Report, National Bureau of Economic Research 2023.

Amberg, Niklas, Tor Jacobson, and Erik von Schedvin, "Trade Credit and Product Pricing: The Role of Implicit Interest Rates," *Journal of the European Economic Association*, 2020.

Andrews, Isaiah and Daniel Barron, "The allocation of future business: Dynamic relational contracts with multiple agents," *American Economic Review*, 2016, *106* (9), 2742–59.

Antras, Pol and C Fritz Foley, "Poultry in motion: a study of international trade finance practices," *Journal of Political Economy*, 2015, *123* (4), 853–901.

Armstrong, Mark, "Nonlinear pricing," *Annual Review of Economics*, 2016, *8*, 583–614.

Atkin, David and Dave Donaldson, "Who's getting globalized? The size and implications of intra-national trade costs," Technical Report, National Bureau of Economic Research 2015.

Attanasio, Orazio and Elena Pastorino, "Nonlinear pricing in village economies," *Econometrica*, 2020.

Banerjee, Abhijit V and Esther Duflo, "Reputation effects and the limits of contracting: A study of the Indian software industry," *Quarterly Journal of Economics*, 2000, *115* (3), 989–1017.

— and — , "Do firms want to borrow more? Testing credit constraints using a directed lending program," *Review of Economic Studies*, 2014, *81* (2), 572–607.

Banerji, Abhijit, JV Meenakshi, and Gauri Khanna, "Social contracts, markets and efficiency: Groundwater irrigation in North India," *Journal of Development Economics*, 2012, *98* (2), 228–237.

Baron, David P and David Besanko, "Regulation and information in a continuing relationship," *Information Economics and Policy*, 1984, *1* (3), 267–302.

Barrett, Christopher B, Maren E Bachke, Marc F Bellemare, Hope C Michelson, Sudha Narayanan, and Thomas F Walker, "Smallholder participation in contract farming: comparative evidence from five countries," *World Development*, 2012, *40* (4), 715–730.

Battaglini, Marco and Rohit Lamba, "Optimal dynamic contracting: The first-order approach and beyond," *Theoretical Economics*, 2019, *14* (4), 1435–1482.

Bergquist, Lauren Falcao and Michael Dinerstein, "Competition and entry in agricultural markets: Experimental evidence from Kenya," *American Economic Review*, 2020, *110* (12), 3705–3747.

Bernard, Andrew B, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes, "The origins of firm heterogeneity: A production network approach," *Journal of Political Economy*, 2022, *130* (7), 1765–1804.

Blouin, Arthur and Rocco Macchiavello, "Strategic default in the international coffee market," *Quarterly Journal of Economics*, 2019, *134* (2), 895–951.

Board, Simon, "Relational contracts and the value of loyalty," *American Economic Review*, 2011, *101* (7), 3349–67.

Boehm, Johannes and Ezra Oberfield, "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," *Quarterly Journal of Economics*, 2020, *135* (4), 2007–2058.

Boudreau, Laura, Julia Cajal-Grossi, and Rocco Macchiavello, "Global value chains in developing countries: a relational perspective from coffee and garments," *Journal of Economic Perspectives*, 2023, *37* (3), 59–86.

Brugués, Felipe and Rebecca De Simone, "Taxation when markets are not competitive: Evidence from a loan tax," 2024.

Bubb, Ryan, Supreet Kaur, and Sendhil Mullainathan, "Barriers to contracting in village economies: A test for enforcement constraints," 2016.

Bull, Clive, "The existence of self-enforcing implicit contracts," *Quarterly Journal of Economics*, 1987, *102* (1), 147–159.

Burstein, Ariel, Javier Cravino, and Marco Rojas, "Input price dispersion across buyers and misallocation," Technical Report, Central Bank of Chile 2024.

Casaburi, Lorenzo and Tristan Reed, "Using individual-level randomized treatment to learn about market structure," *American Economic Journal: Applied Economics*, 2022, *14* (4), 58–90.

Chatterjee, Shoumitro, "Market Power and Spatial Competition in Rural India," *Quarterly Journal of Economics*, 2023, *138* (3), 1649–1711.

Chiappori, Pierre-André, Krislert Samphantharak, Sam Schulhofer-Wohl, and Robert M Townsend, "Heterogeneity and risk sharing in village economies," *Quantitative Economics*, 2014, *5* (1), 1–27.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform pricing in us retail chains," *Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, and Glenn Magerman, "Imperfect Competition in Firm-to-Firm Trade," *Journal of the European Economic Association*, 03 2022. jvac013.

Emran, M Shahe, Dilip Mookherjee, Forhad Shilpi, and M Helal Uddin, "Credit rationing and pass-through in supply chains: theory and evidence from Bangladesh," *American Economic Journal: Applied Economics*, 2021, *13* (3), 202–236.

Fisman, Raymond and Mayank Raturi, "Does competition encourage credit provision? Evidence from African trade credit relationships," *Review of Economics and Statistics*, 2004, *86* (1), 345–352.

Garcia-Marin, Alvaro, Santiago Justel, and Tim Schmidt-Eisenlohr, "Trade credit and markups," 2023.

Garrett, Daniel, Alessandro Pavan, and Juuso Toikka, "Robust predictions in dynamic screening," *Manuscript, Northwestern Univ*, 2018.

Genicot, Garance and Debraj Ray, "Bargaining power and enforcement in credit markets," *Journal of Development Economics*, 2006, *79* (2), 398–412.

Ghani, Tarek and Tristan Reed, "Relationships on the rocks: Contract evolution in a market for ice," *American Economic Journal: Microeconomics*, 2020.

Giannetti, Mariassunta, Mike Burkart, and Tore Ellingsen, "What you sell is what you lend? Explaining trade credit contracts," *The Review of Financial Studies*, 2011, *24* (4), 1261–1298.

_, Nicolas Serrano-Velarde, and Emanuele Tarantino, "Cheap trade credit and competition in downstream markets," *Journal of Political Economy*, 2021, *129* (6), 1744–1796.

Grant, Matthew and Meredith Startz, "Cutting out the middleman: The structure of chains of intermediation," Technical Report, National Bureau of Economic Research 2022.

Grennan, Matthew, "Price discrimination and bargaining: Empirical evidence from medical devices," *American Economic Review*, 2013, *103* (1), 145–77.

Halac, Marina, "Relational contracts and the value of relationships," *American Economic Review*, 2012, *102* (2), 750–79.

Harris, Adam and Thi Mai Anh Nguyen, "Long-term Relationships and the Spot Market: Evidence from US Trucking," 2022.

Heise, Sebastian, "Firm-to-firm relationships and the pass-through of shocks: Theory and evidence," *Review of Economics and Statistics*, 2024.

Huneeus, Federico, "Production network dynamics and the propagation of shocks," *Princeton University*, 2018.

Jacoby, Hanan G, Rinku Murgai, and Saeed Ur Rehman, "Monopoly power and distribution in fragmented markets: The case of groundwater," *Review of Economic Studies*, 2004, *71* (3), 783–808.

Johnson, Simon, John McMillan, and Christopher Woodruff, "Courts and relational contracts," *The Journal of Law, Economics, and Organization*, 2002, *18* (1), 221–277.

Jullien, Bruno, "Participation constraints in adverse selection models," *Journal of Economic Theory*, 2000, *93* (1), 1–47.

Kaplow, Louis, "On the relevance of market power," *Harv. L. Rev.*, 2016, *130*, 1303.

Karaivanov, Alexander and Robert M Townsend, "Dynamic financial constraints: Distinguishing mechanism design from exogenously incomplete regimes," *Econometrica*, 2014, *82* (3), 887–959.

Katz, Michael L, "Price discrimination and monopolistic competition," *Econometrica*, 1984, pp. 1453–1471.

Klapper, Leora, Luc Laeven, and Raghuram Rajan, "Trade credit contracts," *Review of Financial Studies*, 2012, *25* (3), 838–867.

Levin, Jonathan, "Relational incentive contracts," *American Economic Review*, 2003, *93* (3), 835–857.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Lipsey, Richard G and Kelvin Lancaster, "The General Theory of Second Best," *Review of Economic Studies*, 1956, *24* (1), 11–32.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The rise of market power and the macroeconomic implications," *Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

—, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik, "Prices, markups, and trade reform," *Econometrica*, 2016, *84* (2), 445–510.

Luo, Yao, Isabelle Perrigne, and Quang Vuong, "Structural analysis of nonlinear pricing," *Journal of Political Economy*, 2018, *126* (6), 2523–2568.

Macchiavello, Rocco and Ameet Morjaria, "The value of relationships: evidence from a supply shock to Kenyan rose exports," *American Economic Review*, 2015, *105* (9), 2911–45.

— and —, "Competition and relational contracts in the Rwanda coffee chain," *Quarterly Journal of Economics*, 2021, *136* (2), 1089–1143.

— and Josepa Miquel-Florensa, "Vertical integration and relational contracts: Evidence from the Costa Rica coffee chain," 2017.

— and —, "Buyer-driven upgrading in gvcs: The sustainable quality program in colombia," 2019.

MacLeod, W Bentley and James M Malcomson, "Implicit contracts, incentive compatibility, and involuntary unemployment," *Econometrica*, 1989, pp. 447–480.

Marcet, Albert and Ramon Marimon, "Recursive contracts," *Econometrica*, 2019, *87* (5), 1589–1631.

Marshall, Guillermo, "Search and wholesale price discrimination," *RAND Journal of Economics*, 2020, *51* (2), 346–374.

Martimort, David, Aggey Semenov, and Lars Stole, "A theory of contracts with limited enforcement," *Review of Economic Studies*, 2017, *84* (2), 816–852.

Maskin, Eric and John Riley, "Monopoly with incomplete information," *RAND Journal of Economics*, 1984, *15* (2), 171–196.

Mazzocco, Maurizio and Shiv Saini, "Testing efficient risk sharing with heterogeneous risk preferences," *American Economic Review*, 2012, *102* (1), 428–468.

McKenzie, David and Christopher Woodruff, "Experimental evidence on returns to capital and access to finance in Mexico," *The World Bank Economic Review*, 2008, *22* (3), 457–482.

McMillan, John and Christopher Woodruff, "Interfirm relationships and informal credit in Vietnam," *Quarterly Journal of Economics*, 1999, *114* (4), 1285–1320.

Michelson, Hope C, "Small farmers, NGOs, and a Walmart world: Welfare effects of supermarkets operating in Nicaragua," *American Journal of Agricultural Economics*, 2013, *95* (3), 628–649.

Michler, Jeffrey D and Steven Y Wu, "Relational contracts in agriculture: Theory and evidence," *Annual Review of Resource Economics*, 2020, *12*, 111–127.

Monarch, Ryan and Tim Schmidt-Eisenlohr, "Longevity and the Value of Trade Relationships," *Journal of International Economics*, 2023.

Murfin, Justin and Ken Njoroge, "The implicit costs of trade credit borrowing by large firms," *The Review of Financial Studies*, 2015, *28* (1), 112–145.

Nunn, Nathan, "Relationship-specificity, incomplete contracts, and the pattern of trade," *Quarterly Journal of Economics*, 2007, *122* (2), 569–600.

Pavan, Alessandro, Ilya Segal, and Juuso Toikka, "Dynamic mechanism design: A myersonian approach," *Econometrica*, 2014, *82* (2), 601–653.

Pavoni, Nicola, Christopher Sleet, and Matthias Messner, "The dual approach to recursive optimization: theory and examples," *Econometrica*, 2018, *86* (1), 133–172.

Perrigne, Isabelle and Quang Vuong, "Nonparametric identification of a contract model with adverse selection and moral hazard," *Econometrica*, 2011, *79* (5), 1499–1539.

Petersen, Mitchell A and Raghuram G Rajan, "The effect of credit market competition on lending relationships," *Quarterly Journal of Economics*, 1995, *110* (2), 407–443.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert W Vishny, "Legal determinants of external finance," *Journal of Finance*, 1997, *52* (3), 1131–1150.

Ray, Debraj, "The time structure of self-enforcing agreements," *Econometrica*, 2002, *70* (2), 547–582.

Reed, Tristan, Mariana Pereira López, Ana Urrutia Arrieta, and Leonardo Iacovone, "Cartels, Antitrust Enforcement, and Industry Performance: Evidence from Mexico," 2022.

Rodrik, Dani, "Second-best institutions," *American Economic Review*, 2008, *98* (2), 100–104.

Ryan, Nicholas, "Contract enforcement and productive efficiency: Evidence from the bidding and renegotiation of power contracts in India," *Econometrica*, 2020, *88* (2), 383–424.

Seierstad, Atle and Knut Sydsaeter, *Optimal control theory with economic applications*, Elsevier North-Holland, Inc., 1986.

Smith, Janet Kiholm, "Trade credit and informational asymmetry," *Journal of Finance*, 1987, *42* (4), 863–872.

Stantcheva, Stefanie, "Optimal taxation and human capital policies over the life cycle," *Journal of Political Economy*, 2017, *125* (6), 1931–1990.

Startz, Meredith, "The value of face-to-face: Search and contracting problems in Nigerian trade," 2024.

Stole, Lars A, "Price discrimination and competition," *Handbook of Industrial Organization*, 2007, *3*, 2221–2299.

Sugaya, Takuo and Alexander Wolitzky, “The Revelation Principle in Multistage Games,” *Review of Economic Studies*, 2021, *88* (3), 1503–1540.

Thomas, Jonathan and Tim Worrall, “Foreign direct investment and the risk of expropriation,” *Review of Economic Studies*, 1994, *61* (1), 81–108.

Townsend, Robert M, “Risk and insurance in village India,” *Econometrica*, 1994, pp. 539–591.

Troya-Martinez, Marta, “Self-enforcing trade credit,” *International Journal of Industrial Organization*, 2017, *52*, 333–357.

Udry, Christopher, “Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria,” *Review of Economic Studies*, 1994, *61* (3), 495–526.

Varian, Hal R, “Price discrimination and social welfare,” *American Economic Review*, 1985, *75* (4), 870–875.

_ , “Price discrimination,” *Handbook of Industrial Organization*, 1989, *1*, 597–654.

Villas-Boas, J Miguel, “Price cycles in markets with customer recognition,” *RAND Journal of Economics*, 2004, pp. 486–501.

Watson, Joel, “Starting small and commitment,” *Games and Economic Behavior*, 2002, *38* (1), 176–199.

Weerdt, Joachim De and Stefan Dercon, “Risk-sharing networks and insurance against illness,” *Journal of Development Economics*, 2006, *81* (2), 337–356.

Online Appendix

OA-1 Data Construction and Summary Statistics

OA-1.1 Data Construction

The study defines a *product* as a bar-code identifier and description combination. While discounts are observed at the product level, I allocate the discounts offered in a transaction equally across all the products purchased in that transaction by adjusting the listed product unit prices. For example, if a 5% discount is offered on the total bill, the reported unit prices of all the products are adjusted by 5%. I do this, rather than performing analysis with observed discounts to average out managerial mistakes, such as assigning all discounts to a single product, while in principle the agreed discounts were on the total bill.[43]

Let p^l_{ijgry} be the listed unit price and q_{ijgry} be the reported quantity for buyer i from seller j for good g in transaction r during year y, and d_{ijry} be the total discount share in the transaction. Then, the effective unit price is defined as $p_{ijgry} = (1-d_{ijry}) \times p^l_{ijgry}$. Following DellaVigna and Gentzkow (2019), I define standardized unit prices at the transaction-product level $\tilde{p}_{ijgry}$ as:

$$\tilde{p}_{ijgry} = ln(p_{ijgry}) - \overline{ln(p_{jgy})}, \tag{34}$$

where $\overline{ln(p_{jgy})}$ is the average log effective unit price for the good g of seller j in year y. The standardized unit price captures the percentage price difference for a given product in a transaction relative to its average yearly price. I define standardized quantity at the transaction-product level $\tilde{q}_{ijgry}$ in an analogous manner:

$$\tilde{q}_{ijgry} = ln(q_{ijgry}) - \overline{ln(q_{jgy})}, \tag{35}$$

where $\overline{ln(q_{jgy})}$ is the average log quantity for the good g of seller j in year y. As with prices, standardized quantities measure the percentage quantity difference for a given product in a transaction relative to its average quantity sold in the year. Note that these definitions for standardized units are equivalent to netting out product-seller-year fixed effects in a regression of log effective unit prices or log quantities.

To obtain pair-year-level values of the standardized prices and quantities, I aggregate them by the respective share of total expenditures, which provides a common weight for prices and quantities. Define V_{ijy} as the total value of transactions between buyer i and seller j in year y. Let $s_{ijgry} = v_{ijgry}/V_{ijy}$ be the share of expenditure that good g in transaction r represents for the pair and $v_{ijgry} = p_{ijgry} * q_{ijgry}$ be the transaction value.[44] Then, define pair-year level

[43] An alternative method would be to use observed discount shares at the product level and adjust listed prices by the product-specific discount share. In practice, the reduced-form facts hold using either method. See, for instance, Online Appendix Table OA-7 for a robustness exercise using product-level vs bill-level allocation of discounts.

[44] Again, reduced-form results are robust to relying on quantities as weights, rather than values (Online Appendix Table OA-7).

equivalents for the standardized prices and quantities as:

$$\tilde{p}_{ijy} = \sum_{r \in R_{ijy}} \sum_{g \in G_{ijry}} s_{ijgry} * \tilde{p}_{ijgry}, \tag{36}$$
$$\tilde{q}_{ijy} = \sum_{r \in R_{ijy}} \sum_{g \in G_{ijry}} s_{ijgry} * \tilde{q}_{ijgry},$$

where R_{ijy} is the set of all the transactions between i and j in year y and G_{ijry} is the set of all goods in transaction r. The pair-level standardized price then captures the average relative price a buyer has in a given year. For instance, if $\tilde{p}_{ijy} = 0.1$, then the buyer pays on average 10% on their products than other buyers. The pair-level quantities capture the average relative quantity a buyer purchases in a given year. Thus, if $\tilde{q}_{ijy} = 0.1$, then the buyer purchases 10% more in quantity than other buyers.

To address the potential concern that cross-sectional differences in prices and quantities could be driven by variations in the bundles of goods purchased by buyers and over time, I report the main stylized facts on the patterns and dynamics of prices and quantities using standardized measures. The use of these measures indicates that differences in the products purchased by buyers do not influence the results.

For estimation purposes, however, I use the following definitions of prices and quantities, as they are better suited to the structure of the model. For total quantity q_{ijy}, I sum over all reported quantities over all goods and all transactions:

$$q_{ijy} = \sum_{r \in R_{ijry}} \sum_{g \in G_{ijry}} q_{ijgry}. \tag{37}$$

As discussed below in this Section, aggregation across products is not extremely problematic, as firms tend to produce either items or packages that can be summed over in a relatively consistent way.

For prices, I obtain the average unit price by dividing the total value of transactions by the total quantity:

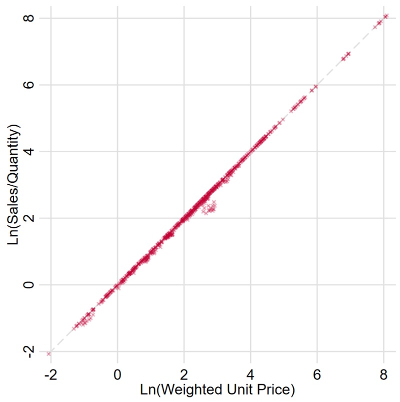
$$p_{ijy} = V_{ijy}/q_{ijy}. \tag{38}$$

This definition of prices is consistent with the weighted average of product-level effective prices, as demonstrated in Online Appendix Figure OA-1, which presents the fit between average unit prices and weighted effective unit prices.[45] The figure shows a strong fit between the two measures, with a correlation of 0.58 at the buyer-seller-year level.

The aggregate quantity produced by seller j in year y is given by $Q_{jy} = \sum_{i \in I_{jy}} q_{ijy}$, where I_{jy} is the set of all buyers that transacted with the seller in the year. While the measures of quantities differ between the model and the motivating evidence, all motivating facts hold when using total quantities, both in the cross-section and in the short panel structure (controlling for buyer-seller pair fixed effects). Robustness results using average unit prices and total quantity are also discussed below.

[45] Observed weighted prices are obtained by aggregating unit prices using the share of the total quantity of the goods sold as weights.

Figure OA-1: **Average Price vs Weighted Price**



Notes: This figure plots average unit prices (in logs) against weighted prices (in logs) across buyer-seller-year. Average unit prices are calculated by dividing the total value of yearly transactions by total quantity purchased (pooling different products). Weighted prices are calculated by summing transaction-product-level unit prices with total expenditure share as weights.

OA-1.2 Main Summary Statistics

Table OA-1 shows that the sellers in my sample are typically large and well-established, employ directly imported goods in their production, channel their sales through the local market rather than exporting. On the other hand, buyers are smaller, younger, and have limited direct contact with international trade. Moreover, buyers are less capital-intensive than sellers. At the same time, sellers in the same 6-digit industry but not in my sample are orders of magnitude smaller, younger, do not use imported inputs, and are much less capital-intensive than seller in sample.[46]

Table OA-1: **Summary Statistics - Sellers and Buyers in 2016**

	Sellers - Sample			*Buyers*			*Sellers - Not Sample*		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Total Sales (million USD)	14.95	8.26	24.33	2.35	0.20	24.33	0.10	0.00	3.04
Total Inputs (millon USD)	10.58	5.31	18.94	1.92	0.15	24.13	0.07	0.00	1.86
Age	30.47	29.00	19.16	15.18	14.00	9.75	9.24	7.00	8.88
Import Share (%)	24.47	21.38	22.96	3.82	0.00	13.49	0.30	0.00	3.91
Export Share (%)	5.81	0.00	19.11	1.06	0.00	8.87	0.10	0.00	2.81
Capital-Expenditures Ratio	0.27	0.30	0.18	0.16	0.05	0.23	0.02	0.00	0.10
Observations	49			28,138			28,424		

Notes: This table reports summary statistics about the size, age, capital intensity, and trade exposure of buyers and sellers in the sample for the year 2016. Monetary values are in U.S. dollars for 2016.

Table OA-2 presents the industrial composition of buyers categorized by selling sector. Buyers of Textile products are primarily from the Wholesale and Retail sector, followed by Manufacturing. Pharmaceutical products are mostly purchased by entities in the Wholesale and Retail sector, as well as the Human Health sector, which includes hospitals and doctors. Cement products, on the other hand, are mainly bought by businesses in Wholesale and Retail, Construction, and Professional Services, such as engineering and architectural firms. Across all

[46]The large number of sellers not in sample is driven primarily by thousands of micro-entrepreneurs in textiles. Online Table Appendix Table OA-8 presents the sample descriptive statistics by seller industry.

selling sectors, the predominance of buyers in Wholesale and Retail Trade suggests that buyers likely have linear input needs.

Table OA-2: **Industrial Composition of Buyers by Selling Sector**

Seller Industry	Ranking	Buyer Industry	Average % Share Pairs
Textiles	1	Wholesale & Retail	40
Textiles	2	Manufacturing	15
Textiles	3	Professional Activities	8
Textiles	4	Agriculture	5
Textiles	5	Other	31
Pharmaceutical	1	Wholesale & Retail	46
Pharmaceutical	2	Human Health	17
Pharmaceutical	3	Manufacturing	10
Pharmaceutical	4	Construction	4
Pharmaceutical	5	Other	23
Cement-Products	1	Wholesale & Retail	25
Cement-Products	2	Construction	20
Cement-Products	3	Professional Activities	16
Cement-Products	4	Manufacturing	8
Cement-Products	5	Other	31

Notes: This table presents a ranked breakdown of the industrial composition of buyers for each selling sector, organized by the highest share of buyers.

Table OA-3 presents summary statistics on quantities, values, and the number of buyers per seller obtained through the EI dataset. Notice that the reporting threshold is smaller than in previous work (Bernard et al., 2022; Alfaro-Urena et al., 2022), implying a larger number of buyers. Despite the large number of buyers, the yearly bills are not small for the country, with median (average) bill of 9K USD (44K USD).[47]

Table OA-3: **Summary Statistics - Electronic Invoice Database**

	Mean	Median	SD
N. Buyers	8,028.41	613.50	25,078.11
Total Sales (million USD)	16.58	7.23	29.44
Total Q (million)	5.42	1.20	9.01
Q per Buyer	12,455.39	1,495.22	25,823.40
Bill per Buyer (USD)	43,490.37	9,067.65	105,840.28
Observations	49		

Notes: This table reports summary statistics of the electronic invoice database. N. buyers refers to the number of unique buyers each seller in the sample has on average over 2016 and 2017. Quantity is the sum of all quantities across products. Bill per buyer is the total value of the transactions between buyer and seller.

The median (average) buyer purchases around 1.5K (12.5K) units of product. What are these products? Table OA-4 provides information on a random sample of products, including

[47] At the same time, due to the staggered rollout of the policy, data is sourced from the largest firms in the economy. Indeed, the size in number of buyers and total sales of the median manufacturing firm in my sample corresponds to size of manufacturing firms between the top 5 and 10 percent in Costa Rica (Alfaro-Urena et al., 2022) and between the top 25 and 10 percent in Belgium (Bernard et al., 2022).

their prices and average costs. The prices are obtained directly from invoices, while the average costs are imputed by dividing the total variable costs, including wages and intermediate inputs, by nthe aggregate output in units for each firm.

Table OA-4: **Example - Product Information, Prices, and Average Costs**

Industry	Firm-ID	Product Description	Observed Unit Price	Imputed Average Cost
Textiles	1	Teddy King, Size 55, Brim 7CM, Color-B02 [Panama Hat]	33.90	11.96
Textiles	2	Shirt, R:1931, Squares	19.34	9.85
Textiles	3	Tank Undershirt, Male, Size M, White	10.27	6.72
Textiles	4	Betty K246	19.44	16.94
Textiles	5	Bikini, Woman, 500306, Black, L	13.50	16.78
Textiles	6	Ribbon, Black, 30 mm X 700	26.62	1.86
Textiles	7	Skirt, Tropical Squares, Scottish	46.01	17.77
Textiles	8	Boots, LLN NG AM, Size 39	7.09	2.17
Textiles	9	Elastic Socks, Nylon and Cotton	16.56	8.48
Textiles	10	Jacket, Kids, Spiderman Print, Hoodie	18.30	7.11
Pharmaceutical	1	Nitazoxanida, 500mg X 6 tablets	5.27	4.83
Pharmaceutical	1	Clopidogrel Tarbis 75 mg film-coated tablets	12.90	6.57
Pharmaceutical	2	Losartan/Hydrochlorothiazide, 100mg X 28 tablets	5.04	0.78
Pharmaceutical	3	B Complex, Syrup 120 ml	2.32	0.81
Pharmaceutical	4	Sodium perborate, mint oil, saccharin	4.69	1.81
Pharmaceutical	5	Boldenone 50, Injectable, Bottle X 500 ml	123.12	3.01
Pharmaceutical	6	Pinaver, Film-coated, 100 mg X 20 tables	10.32	2.62
Pharmaceutical	7	Endobion X 60 tablets	14.83	5.49
Pharmaceutical	7	Prostageron X 60 capsules	14.75	7.04
Pharmaceutical	8	Oral rehydration solution, cherry, 500ml	2.67	1.80
Cement-Products	1	Gray French Pedestrian Paving Stone	11.28	18.11
Cement-Products	2	Corrugated Plate	23.73	9.56
Cement-Products	3	Polymer-modified adhesive mortar for ceramics, 25kg	6.31	2.99
Cement-Products	4	Polymer-modified adhesive mortar for ceramics, 25kg	6.94	12.36
Cement-Products	5	Polymer-modified adhesive mortar for ceramics, 25kg	6.65	3.45
Cement-Products	6	Straight Pole 21m x 1400kg, Reinforced Concrete	882.00	73.95
Cement-Products	6	Straight Pole 21m x 2400kg, Reinforced Concrete	1362.73	73.95
Cement-Products	7	Tile 50x50x2 cm (Color)	32.00	6.62
Cement-Products	8	MFC Concrete, 300, XXXXX XXXX-XXXX	94.00	50.34
Cement-Products	8	CFC Concrete, 240, XXXXX XXXX-XXXX	79.43	50.34

Notes: This table presents a sample of ten random products from each of the studied sectors (textiles, pharmaceutical, and cement-products), with product descriptions translated into English and sensitive information, such as brand names, removed to ensure confidentiality. The observed average unit prices reflect the listed prices reported by the firms, while the imputed average costs are estimated using the firms' total variable costs divided by total quantity.

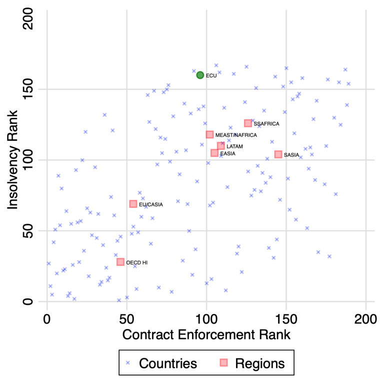
In the textiles industry, products may include shirts, skirts, hats, and others, with different patterns or sizes also considered separate products. Aggregation is thus over individual clothing *items*. Instead, in the pharmaceutical sector, products are typically packages of tablets or bottles, with aggregation across products being over *packages*. Comparing product-level prices with firm-level average costs yields reasonable estimates in both cases. For example, a shirt is sold for 19 USD and costs 9.85 USD to manufacture, and Vitamin B Syrup is sold for 2.3 USD, but it costs only 81 cents to manufacture.

In the cement-products industry, products may include stones, mortar, concrete, and the like. While aggregating over these types of products can be more challenging, it should be noted that firms producing products such as mortar do not typically produce tiles, poles, or stones. Two other notes are in order. First, there are three different firms selling mortar at similar prices, despite being headquartered in different and distant cities. This suggests that despite the products being substitutes, sellers may still have local market power due to transportation costs. Second, one firm produces two types of pole products, sold at different prices but with the same cost of production. Another firm produces two types of concrete products, sold at different prices but with the same cost of production. As costs will enter into the dependent variable in my main estimation process, possible mistakes in costs would enter as measurement error in the econometric model.

OA-1.3 Context Related Summary Statistics

Online Appendix Figure OA-2 shows Ecuador's position in terms of contract enforcement and insolvency in the World Bank Doing Business report. Lower numbers represent better institutions to enforce contracts or resolve insolvency cases.

Figure OA-2: **Ranks Insolvency and Enforcement**



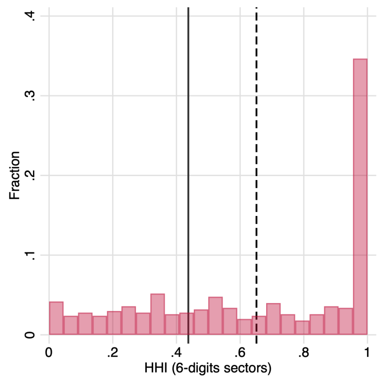
Notes: This figure presents Ecuador's rank in the World Bank Doing Business categories of Insolvency (Y-Axis) and Enforcement (X-Axis). The most efficient country in terms of enforcement ranks 1st.

Online Appendix Figure OA-3 shows the distribution of Herfindahl-Hirschman Indices (HHI) for 6-digit manufacturing sectors in 2017. HHI_s for sector s is estimated using the following formula:

$$HHI_s = \sum_{j \in J_s} m_j^2,$$

where m_j is the market share of firm j, and J_s is the set of active firms in sector s. The market share of firm j is obtained by dividing the total revenue of firm j by the total revenue of all firms in sector s.

Figure OA-3: **Distribution of Herfindahl-Hirschman Indices for Manufacturing in 2017**



Notes: This figure presents a histogram of estimated Herfindahl-Hirschman Indices (HHI) for 6-digit manufacturing sectors in 2017.

OA-2 Motivating Evidence - Robustness

Quantity Dynamics and Pair FE. In Online Appendix Figure OA-4a, I verify that the differences are not driven by selection, but rather reflect a real increase within pairs. To do so, I run a regression of total quantity q_{ijt} on dummies for the age of the relationship, controlling for pair fixed effects. The figure plots the coefficients for the relationship age dummies and shows that the volume of total quantity purchased grows as relationships age.

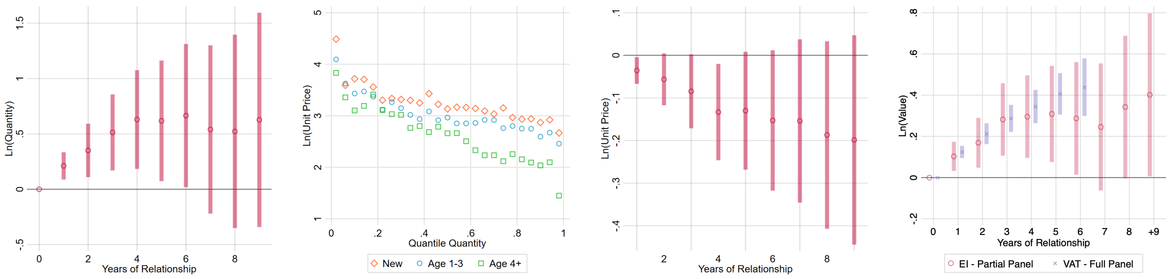
Quantity Discounts and Pair FE. Online Appendix Figure OA-4b plots a binscatter regression of log average unit price on quantiles of quantity, controlling for seller-year fixed effects. The figure documents the presence of quantity discounts within relationship age. *Relationship Discounts and Pair FE.* Online Appendix Figure OA-4c shows a binscatter plot of log average prices on the age of the relationship, controlling for pair fixed effects. The figure shows that as relationships age, they receive around 1.5% additional discounts per year. Under both formulations, there are price discounts conceded to older clients. These results indicate that the discounts are not driven by composition, nor by short-term fixed characteristics of the firm (such as location, managerial bargaining, size, etc.).

Relationship Value and Pair FE. Online Appendix Figure OA-4d plots regression coefficients for the value of total sales between buyer and supplier on the age of the relationship, controlling for pair fixed effects. The red figures use the electronic invoice database and are constructed using only a partial panel of two observations per pair for the years 2016-2017. The purple marks are constructed using multiple observations of buyer-seller pairs from the VAT B2B database for the years 2007-2015 for the sellers in the electronic invoice database. The figure confirms that relying on only two years of relationship data can properly capture full relationship dynamics observed in longer panel datasets.

Backloading in Prices and Quantities by Quantile and Relationship Age. Online Appendix Figure OA-5 presents pair-specific changes in prices and quantities between 2016 and 2017, by the age of the relationship in 2016, over quantiles of quantity purchased in 2016. The figures show that prices tend to decrease faster and quantities increase faster for lower quantiles. Over time, backloading in prices and quantities becomes weaker. By age 5, prices and quantities are relatively stable across quantiles.

Benchmarking Quantity Discounts. Online Appendix Table OA-5 shows the results of a regres-

Figure OA-4: **Motivating Facts - Robustness (Pair Fixed Effects)**



(a) Q. over Time (b) Q. Discounts (c) Prices over Time (d) Value over Time

Notes: Panel (a) plots the coefficients of log total quantity on relationship age dummies, controlling for pair fixed effects. Total quantity is obtained by aggregating all the reported units of sold goods. Standard errors are clustered at the pair level. Panel (b) shows the relationship between quantity purchased and average log unit price through binscatters of the measure of unit price against quantile of quantity by age of relationship. Quantiles of quantity are calculated for each seller-relationship age combination. Panel (c) plots the regression coefficients of log unit prices on years of relationship, controlling for pair fixed effects. Standard errors are clustered at the pair level. Panel d) plots regression coefficients for the value of total sales between buyer and supplier on the age of the relationship, controlling for pair fixed effects. The red figures use the electronic invoice database and are constructed using only a partial panel of two observations per pair for the years 2016-2017. The purple marks are constructed using multiple observations of buyer-seller pairs from the VAT B2B database for the years 2007-2015 for the sellers in the electronic invoice database.

sion on log average price on log quantity, controlling for seller-year fixed effects. The table presents a benchmark quantity discount measure of a 2% decrease in price for a 10% increase in total quantity purchased.

Table OA-5: **Benchmark: Quantity Discounts**

	(1)
VARIABLES	ln(Price)
ln(Quantity)	-0.220***
	(0.0238)
Constant	3.046***
	(0.0718)
Seller-Year FE	Yes
Observations	76,473
R-squared	0.666

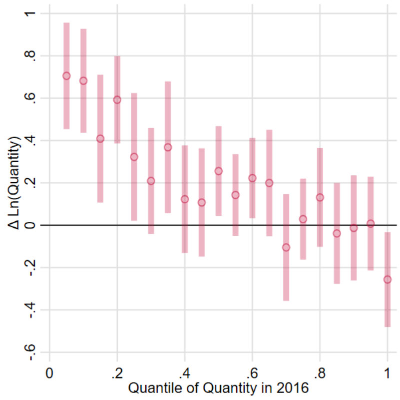
Notes: This table presents a regression of log average unit prices on log quantity, controlling for seller-year fixed effects. Standard errors are clustered at the seller-year level. *** p<0.01, ** p<0.05, * p<0.1

Relationship Discounts and Additional Controls. I replicate Figure 1e in Table OA-6 to assess the robustness of the results to additional buyer and relationship-level controls, obtained from the firms' financial statements. Relative to the base specification presented in Column (1), I find that the effects of relationship age and quantity discounts remain relatively unchanged after accounting for various buyer and pair characteristics.

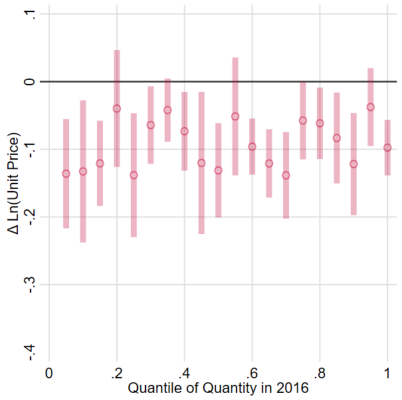
In Column (2), I control for buyer and pair characteristics such as age, distance between headquarters, size (measured by sales, number of employees, and assets), and whether the firm is a multinational, exporter, importer, or part of a business group. I also consider the impor-

Figure OA-5: **Backloading in Prices and Quantities by Quantile and Age**

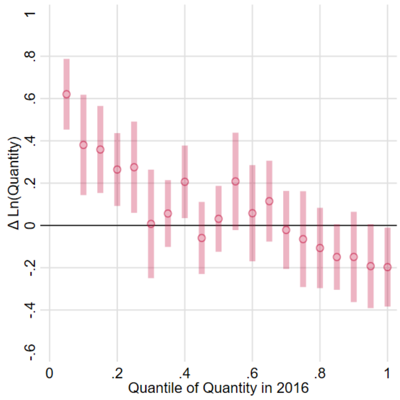
(a) Prices (New Buyers)



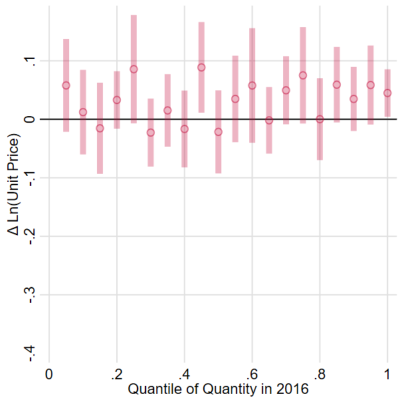
(b) Quantities (New Buyers)



(c) Prices (Age 1-4)



(d) Quantities (Age 1-4)



(e) Prices (Age 5+)

(f) Quantities (Age 5+)

Notes: This figure presents pair-specific year-to-year changes in unit prices and quantities from 2016 and 2017 for new buyers, ages 1 to 4, and age 5+, against the quantile of quantity purchased in 2016. The age of relationships is from 2016, and quantiles of quantity are measured in 2016 for each seller-relationship age. Error bars present variation at the 95% level, with standard errors clustered at the seller level.

tance of the relationship for both the buyer (in terms of supply share) and seller (in terms of demand share) to capture any potential asymmetries in bilateral bargaining power (Dhyne et al., 2022; Alviarez et al., 2023). In Column (3), I include further controls, such as buyer wages, expenditures, cash, fixed assets, debt, leverage, and export and import shares, as well as 6-digit sectoral fixed effects for buyers. The stability of the coefficients implies that buyer characteristics observed by the seller, but not accounted for in a model focusing solely on relationship and quantity variation, likely enter as measurement error rather than generating bias in the coefficients linking prices, quantities, and relationship age.

In Columns (4) and (5), I substitute the relationship age with its logarithmic form, rather than in levels, and again find robust results for both discounts over time and by quantity. Importantly, prices are most responsive to quantities and the age of the relationship. For instance, the coefficient for the (log) age of the relationship is 4 to 10 times larger than the coefficient for the (log) age of the buyer, and 15 to 20 times larger than the coefficient for the (log) sales of the buyer.

Relationship Discounts and Aggregation Weights. Online Appendix Table OA-7 presents the robustness of relationship discounts to the method of discount allocation (at the bill level vs. at the product level) as well as the weights used to aggregate prices at the seller-buyer-year level (quantities vs. values as weights).

Table OA-6: **Standardized Log Price - Robustness to Additional Controls**

VARIABLES	(1) Std. ln(Price)	(2) Stdz. ln(Price)	(3) Stdz. ln(Price)	(4) Stdz. ln(Price)	(5) Stdz. ln(Price)
Age of Relationship	-0.00554***	-0.00552***	-0.00480***		
	(0.00156)	(0.00146)	(0.00140)		
ln(Age of Relationship+1)				-0.0186***	-0.0161***
				(0.00500)	(0.00483)
Stdz. ln(Quantity)	-0.0472***	-0.0463***	-0.0420***	-0.0463***	-0.0420***
	(0.00780)	(0.00722)	(0.00414)	(0.00719)	(0.00414)
Supply Share		0.0262*	0.0183	0.0268	0.0184
		(0.0157)	(0.0137)	(0.0163)	(0.0148)
Demand Share		0.0119	0.0378	-0.00200	0.0250
		(0.0486)	(0.0400)	(0.0479)	(0.0388)
ln(Age Buyer)		-0.000836	-0.00368***	-0.00169	-0.00439***
		(0.00117)	(0.00117)	(0.00114)	(0.00124)
ln(Distance Km)		2.96e-05	0.000472	-2.77e-05	0.000424
		(0.00192)	(0.00194)	(0.00192)	(0.00194)
ln(Sales Buyer)		0.00108**	0.000774**	0.00110**	0.000804**
		(0.000469)	(0.000318)	(0.000477)	(0.000318)
ln(N. Employees Buyer)		0.000235	0.00161*	0.000256	0.00160*
		(0.000846)	(0.000932)	(0.000841)	(0.000936)
ln(Assets Buyer)		0.00131***	0.00228***	0.00132***	0.00230***
		(0.000318)	(0.000696)	(0.000319)	(0.000668)
ln(Wages Buyer)			-0.000472		-0.000471
			(0.000319)		(0.000322)
ln(Expenditures Buyer)			-0.000949**		-0.000926***
			(0.000377)		(0.000347)
ln(Cash Buyer)			0.000785**		0.000784**
			(0.000338)		(0.000335)
ln(Fixed Assets Buyer)			0.000507**		0.000501**
			(0.000253)		(0.000247)
ln(Debt Buyer)			-0.000456		-0.000453
			(0.000411)		(0.000410)
Leverage Buyer			0.000833		0.000835
			(0.00195)		(0.00194)
1{BG Buyer}		-0.00374	7.80e-05	-0.00383	-7.23e-06
		(0.00236)	(0.00207)	(0.00236)	(0.00207)
1{Multinational Buyer}		0.0194		0.0197*	
		(0.0120)		(0.0118)	
1{Exporter Buyer}		-0.00747	-0.0239**	-0.00721	-0.0239**
		(0.00566)	(0.00937)	(0.00553)	(0.00948)
Export Share Buyer			0.0321***		0.0314***
			(0.00793)		(0.00791)
1{Importer Buyer}		0.00369**	0.000511	0.00357*	-0.000107
		(0.00184)	(0.00440)	(0.00183)	(0.00442)
Import Share Buyer			0.0105*		0.0112*
			(0.00562)		(0.00579)
Observations	76,412	73,633	65,754	73,633	65,754
R-squared	0.075	0.082	0.048	0.083	0.048
Year FE	Yes	Yes	Yes	Yes	Yes
Buyer Sector FE	No	No	Yes	No	Yes

Notes: This table presents regressions regressions of standardized unit prices on age of relationship, standardized quantity, and different buyer characteristics. Standard errors are clustered at the seller-year level. *** p<0.01, ** p<0.05, * p<0.1

Table OA-7: **Robustness to Weights and Discount Allocation**

	(1)	(2)	(3)	(4)
Variable (Weighted Average)	Stdz. Price	Stdz. Price	Stdz. Price	Stdz. Price
Age of Relationship	-0.00690***	-0.00685***	-0.0106***	-0.00884**
	(0.00187)	(0.00177)	(0.00401)	(0.00433)
Weights	Values	Quantity	Values	Quantity
Discount Allocation	Bill	Bill	Product	Product
Observations	76,473	76,473	76,473	76,473
R-Squared	0.018	0.018	0.015	0.009
Pair FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Quantity Control	No	No	No	No

Notes: This table presents regressions of prices on the age of the relationship under different weights for aggregation and methods of allocating discounts. Column (1) is the benchmark and allocates discounts at the bill level, relying on the value share of total yearly transactions as aggregation weights. Column (2) allocates discounts at the bill level and uses total quantity as weights. Column (3) allocates discounts at the product level with values as weights. Column (4) allocates discounts at the product level with quantities as weights. Standard errors are clustered at the seller-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

OA-3 Motivating Facts by Seller Sector

In this section, I present the overall consistency of the motivating facts for each seller sector: namely, textile, cement-products, and pharmaceuticals.

Online Appendix Table OA-8 presents summary statistics by seller's sectors for *Sellers in Sample* (Panel (a)), *Sellers Not in Sample* (Panel (b)), which are sellers in the same industry but small enough that they were not covered by the EI seller's database, and *Buyers in Sample* (Panel (c)). The table demonstrates that the sellers in the sample are significantly larger than their non-sample competitors, with the mean sample seller being 272 times larger in the textile industry, 8 times larger in the pharmaceutical industry, and 32 times larger in the cement-products industry. Furthermore, the sample sellers exhibit a higher exposure to imported materials compared to their non-sample counterparts, with 113 times more reliance in textiles, 4 times more in pharmaceuticals, and 26 times more in cement-products. Additionally, the firms in the sample display a considerably higher capital intensity, with 18 times more capital per dollar in expenditure in textiles, 2 times more in pharmaceuticals, and 8 times more in cement-products. These patterns collectively suggest that (1) the manufacturing firms in the sample are preferred suppliers within their respective industries, (2) there is a degree of product differentiation, likely indicating higher quality given the increased reliance on imported inputs, and (3) the higher capital intensity relative to labor implies a reduced likelihood of production issues.

These statistics also help understand why sellers in the sample have market power in the first place. For textile-products: The average buyer in this industry is smaller than the sellers in the sample, yet significantly larger than the non-sample competitors (59 times larger). Furthermore, the average order in the industry, at 25,000 USD, is substantial relative to the size of the average (40,000 USD) and median ($<$ 9,000 USD) non-sample seller. Therefore, beyond the potential higher quality of goods offered by the sellers in the sample, the relatively large size requirements for the orders imply a scale advantage for in-sample sellers.

In the pharmaceutical-products industry: Products are generally horizontally differentiated, as active components are imperfect substitutes for the final consumer. The size, age, capital intensity, and reliance on imported inputs suggest that sellers in the sample are the preferred suppliers in this differentiated industry.

In the cement-products industry: Manufacturers likely benefit from local market power due to the high transportation costs associated with these types of goods. Additionally, the manufacturers in the sample are likely vertically differentiated due to their capital-intensive production. Similar to the textile industry, a scale argument is valid as well. The average buyer in the industry is 14 times larger than the average non-sample seller, and the orders are relatively large (45,000 USD) compared to the size of the non-sample seller (350,000 USD average; 10,000 USD median).

Decomposing Figure 1 by sector reveals that the vast majority of qualitative results hold individually in each sector, with Online Appendix Figure OA-6 for Textiles, Online Appendix Figure OA-7 for Pharmaceuticals, and Online Appendix Figure OA-8 for Cement-products.

First, a large share of trade is channeled through repeated relationships (Subfigure OA-6a; Subfigure OA-7a; Subfigure OA-8a), though pharmaceutical manufacturers have a lower share of new clients and quantity channeled through new buyers. Still, repeated transactions rather than spot transactions are thus likely important in each industry.

Second, at least 60% of all transactions are financed by trade-credit (Subfigure OA-6b; Subfigure OA-7b; Subfigure OA-8b). This implies that opportunism on the buyer side is relevant

Table OA-8: **Summary Statistics by Sector - Sellers, Buyers, and Other Competitors**

	Textiles			Pharmaceuticals			Cement-Products		
	Mean	Median	S.D.	Mean	Median	S.D.	Mean	Median	S.D.
Panel (a): Sellers in Sample									
Total Sales (million USD)	10.91	3.55	21.26	21.64	12.18	31.78	11.32	7.10	12.89
Expenditures (million USD)	6.48	2.38	12.90	16.13	6.29	27.10	8.73	5.58	8.68
Age	31.47	29.00	18.55	30.89	33.00	20.73	28.25	22.00	19.17
Import Share (%)	20.28	11.67	21.85	35.94	27.65	24.11	13.92	4.99	15.94
Export Share (%)	14.04	0.21	29.20	1.00	0.00	2.14	0.01	0.00	0.05
Capital Share of Expenditures	0.18	0.13	0.18	0.28	0.32	0.15	0.39	0.44	0.16
Observations	19			18			12		
Panel (b): Sellers Not in Sample									
Total Sales (million USD)	0.04	0.00	0.69	2.81	0.04	11.54	0.35	0.01	7.51
Expenditures (million USD)	0.03	0.00	0.42	2.44	0.03	12.59	0.22	0.00	3.79
Age	8.94	6.00	8.75	14.32	10.50	14.17	10.83	9.00	8.99
Import Share (%)	0.18	0.00	2.90	9.94	0.00	21.89	0.53	0.00	5.01
Export Share (%)	0.09	0.00	2.67	1.37	0.00	8.90	0.13	0.00	2.89
Capital Share of Expenditures	0.01	0.00	0.08	0.17	0.06	0.24	0.05	0.00	0.17
Observations	24,320			234			3,870		
Panel (c): Buyers in Sample									
Total Sales (million USD)	2.37	0.18	25.66	6.21	0.31	58.74	5.01	0.55	44.49
Expenditures (million USD)	1.93	0.13	25.64	5.27	0.28	55.61	4.00	0.50	35.29
Age	15.22	14.00	9.46	16.36	14.00	11.81	15.20	14.00	11.35
Import Share (%)	3.84	0.00	13.62	3.20	0.00	11.14	3.50	0.00	12.55
Export Share (%)	1.14	0.00	9.28	0.48	0.00	4.30	0.63	0.00	6.80
Capital Share of Expenditures	0.17	0.05	0.24	0.12	0.02	0.19	0.16	0.07	0.21
Observations	23,890			2,642			3,053		

Notes: This table reports summary statistics about the size, age, capital intensity, and trade exposure of buyers, sellers in the sample, and sellers not in the sample for the year 2016, separated by seller's sector. Monetary values are in U.S. dollars for 2016.

for all studied industries.

Third, quantities increase as relationships age, both measured as standardized quantities or total quantity (Subfigure OA-6c and OA-6i; Subfigure OA-7c and OA-7i; Subfigure OA-8c and OA-8i). Quantities grow faster in textiles than in other industries. Moreover, looking at both standardized quantity and total quantity demanded, buyers tend to buy more of the same product over time and in total. In pharmaceutical products, product-specific demand levels off after the first year, while total demand continues to increase; in cement-products, product-specific demand levels off after the second year, while total demand continues to increase. In any case, quantity backloading appears relevant across the board.

Fourth, quantity discounts are observed, both within product and in average prices (Subfigure OA-6d and OA-6g; Subfigure OA-7d and OA-7g; Subfigure OA-8d and OA-8g). Thus, a model with price discrimination in quantities is important.

Fifth, price discounts tend to be offered to older relationships (Subfigure OA-6d and OA-6h; Subfigure OA-7d and OA-7h; Subfigure OA-8d and OA-8h). However, in contrast to the main figure, product-specific discounts are not observed on average in pharmaceuticals, whereas they are present in textiles and cement-products. In terms of average prices, relational discounts are observed across all industries. The contrast between quality-adjusted prices and average prices for pharmaceuticals indicates that product bundles are likely switching over time, allowing buyers to purchase cheaper products either not available or desired at the beginning of the relationship. In any case, a model with dynamic discounts could rationalize observed dynamics for average prices for all industries, as well as for quality-adjusted prices for textiles and cement.

Sixth, relationships that trade more intensively are more likely to survive across all industries (Subfigure OA-6f; Subfigure OA-7f; Subfigure OA-8f), though the heterogeneity across ages is smaller in pharmaceutical products than in textiles and cement.

Online Appendix Table OA-9 shows price dynamics by payment modality, relying on the transaction-level data. Joining all industries together (Column 1), we observe that for transactions conducted via trade-credit, quality-adjusted prices decrease as relationships age, accounting for plausible quantity discounts by controlling for a flexible spline in quantity. Instead, when the transaction's modality is pay-in-advance (Column 2), standardized prices increase as relationships age. The same pattern holds for textiles (Columns 3 and 4), cement-products (Columns 7 and 8), and even so for pharmaceuticals (Columns 5 and 6), where quality-adjusted pair-specific prices do not decrease over time.

Online Appendix Figure OA-9 shows the distribution of trade-credit terms offered by industry. Textiles offer on average 40 days of trade-credit, with 7, 30, 45, and 60 days as common terms. Pharmaceutical products offer on average 55 days, with 30, 45, and 60 as common terms. Cement-products offer 40 days, with 30, 45, and 60 days as common terms.

Finally, Online Appendix Table OA-10 presents coefficients of variation of sales and expenditures (month-to-month) for sellers and buyers. We can see that sellers have lower variability both in sales and expenditures than buyers, though variability is still present for sellers, as the standard deviation is 25% of the mean sales for pharmaceuticals, 29% for cement-products, and 42% for textiles. Production expenses are also volatile, with the standard deviation representing 30% of mean expenditures, though the difference across industries is much more muted than in sales.

Table OA-9: **Price Dynamics and Payment Method**

	All		Textiles		Pharmaceuticals		Cement-Products	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Payment Method	TC	O	TC	O	TC	O	TC	O
Total Years	-0.00786***	0.00431***	-0.00358***	0.00380***	-0.00641***	0.00770***	-0.0291***	-0.00218
	(0.00214)	(0.00108)	(0.00125)	(0.00110)	(0.00215)	(0.00232)	(0.0106)	(0.00174)
Observations	3,383,399	608,318	2,249,157	305,517	742,940	240,995	391,302	61,806
R-squared	0.954	0.982	0.988	0.977	0.981	0.975	0.758	0.973
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quantity Control	Spline	Spline	Spline	Spline	Spline	Spline	Spline	Spline

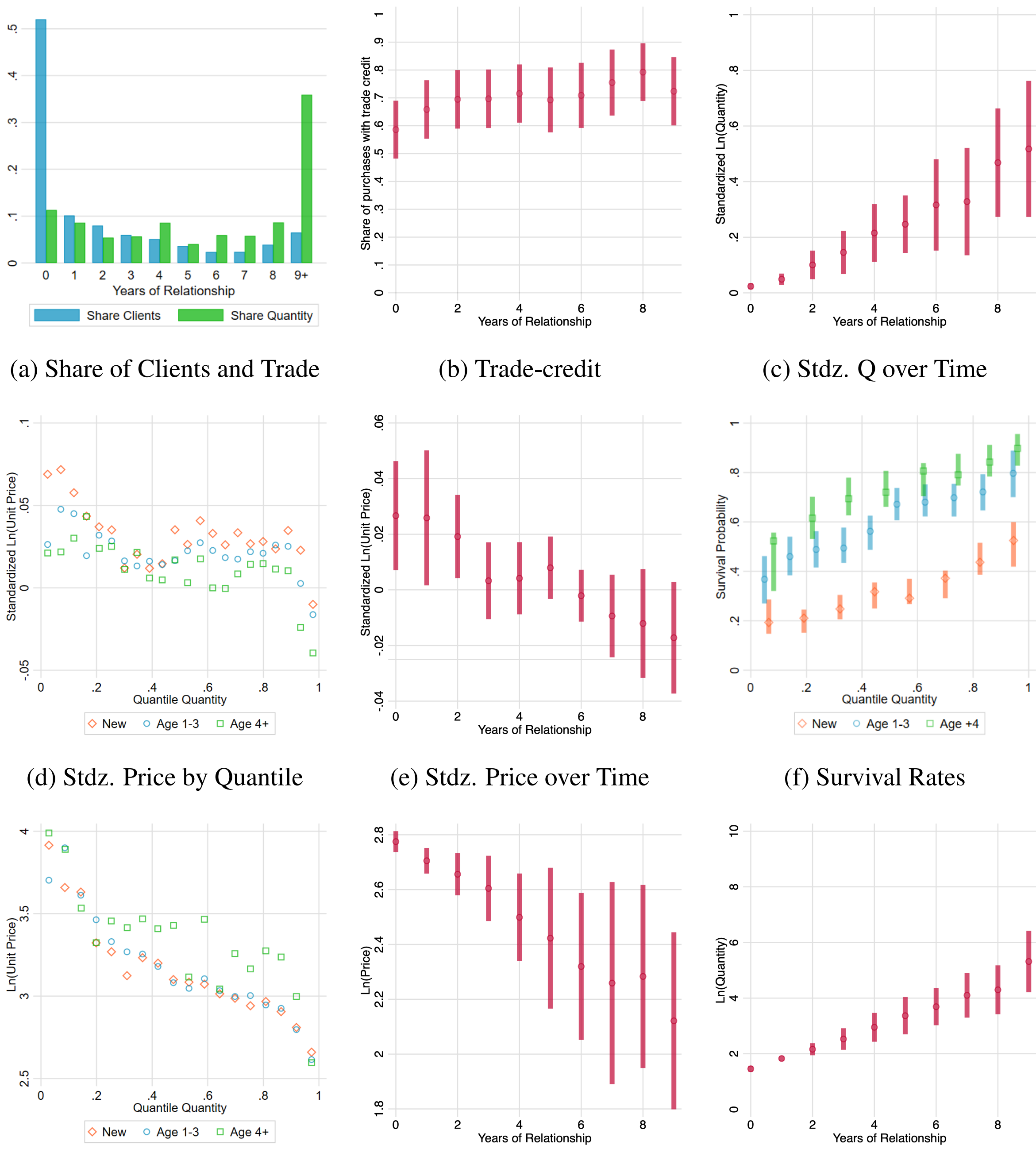
Notes: This table presents transactions-level regression of log unit prices on the age of relationship, controlling for a flexible spline of quantity and product-year fixed effects, by payment modality and sector. Columns (1) and (2) present results for all sectors, for trade-credit transactions and all others, respectively. Columns (3) and (4) report results for textiles, Columns (5) and (6) for pharmaceuticals, and Columns (7) and (8) for cement-products. Standard errors are clustered at the pair-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table OA-10: **Coefficient of Variation of Sales and Expenditures**

	CV Sales Seller	CV Expenditures Seller	CV Sales Buyer	CV Expenditures Buyer
Textiles	0.42	0.28	0.65	0.65
Pharmaceuticals	0.25	0.34	0.77	0.55
Cement-Products	0.29	0.31	0.86	0.68

Notes: This table presents coefficients of variation (CV) in monthly sales and expenditures for sellers and buyers between 2016 and 2017.

Figure OA-6: **Motivating Facts: Textile-Products**



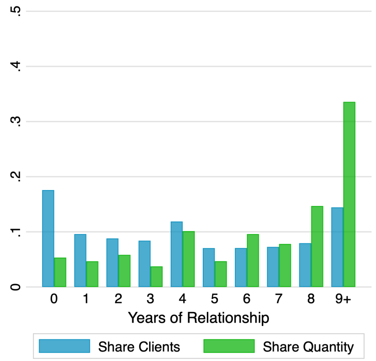
(a) Share of Clients and Trade (b) Trade-credit (c) Stdz. Q over Time

(d) Stdz. Price by Quantile (e) Stdz. Price over Time (f) Survival Rates

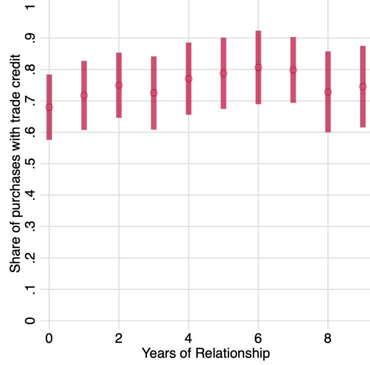
(g) ln(Price) by Quantile (h) ln(Price) over Time (i) ln(Q) over Time

Notes: This figure replicates Figure 1 for Textile-Products only. Subfigure (a) displays the distribution of the average of the share of clients and quantity sold by relationship age, calculated across all sellers in 2016. Subfigure (b) displays the average of the share of purchases channeled through trade-credit, along with a 90% confidence interval, calculated across all sellers. Subfigure (c) displays the evolution of standardized log quantities, with their corresponding 90% confidence intervals, calculated across all sellers. The standardized log quantity is obtained by taking the average quantity sold in a given year for each seller-product and subtracting the log average quantity for that year. The standard errors are calculated at the seller-year level. Subfigure d) shows the relationship between quantity purchased and standardized log unit price through a binscatter plot that displays the measure of unit price against the quantity sold, based on relationship age. The standardized log unit price is obtained by netting out the average log unit price for that year for each seller-product. The quantiles of quantity are calculated for each seller-relationship age combination. Subfigure e) presents a binscatter plot of standardized log unit prices against years of relationship, controlling for a flexible spline of standardized log quantities. The standard errors are calculated at the seller-year level. Subfigure f) displays a binscatter plot of the average survival rate of pairs at different ages and quantiles of quantity. Subfigure g) presents a binscatter of (log) average price on the quantile of quantity by relationship age, controlling for seller-year fixed effects. Subfigure h) presents a binscatter of (log) average price on years of relationship controlling for a flexible spline of quantity and seller-year fixed effects. Subfigure i) presents a binscatter of (log) total quantity on years of relationship controlling for seller-year fixed effects. The quantiles of quantities are calculated for each seller-age combination, and the error bars represent a 90% level of variation across all sellers.

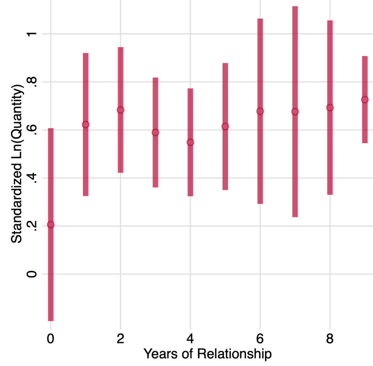
Figure OA-7: **Motivating Facts: Pharmaceutical-Products**



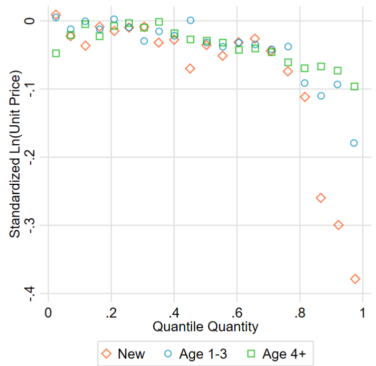
(a) Share of Clients and Trade



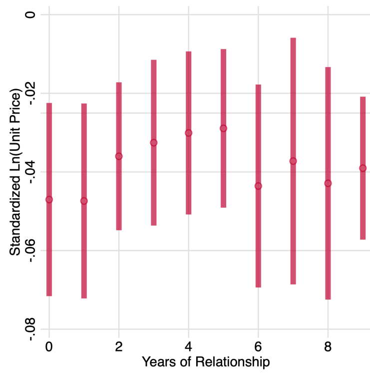
(b) Trade-credit



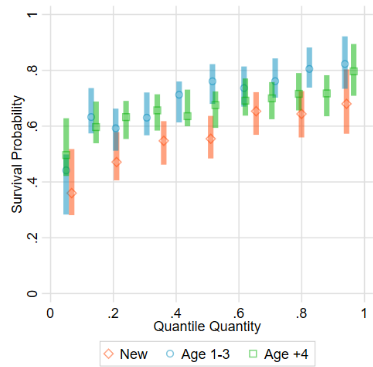
(c) Stdz. Q over Time



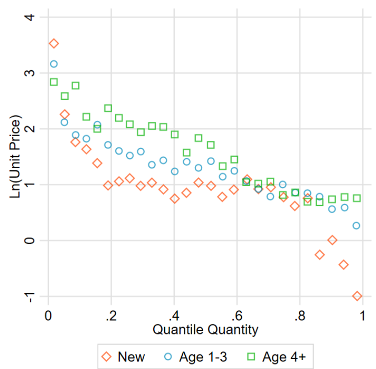
(d) Stdz. Price by Quantile



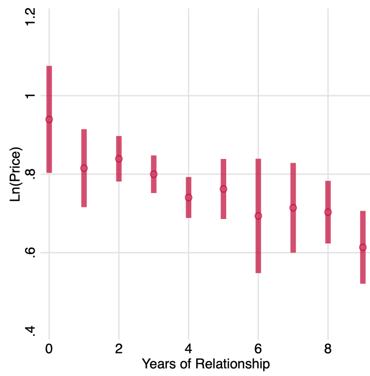
(e) Stdz. Price over Time



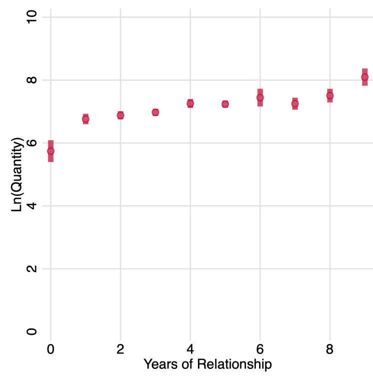
(f) Survival Rates



(g) ln(Price) by Quantile



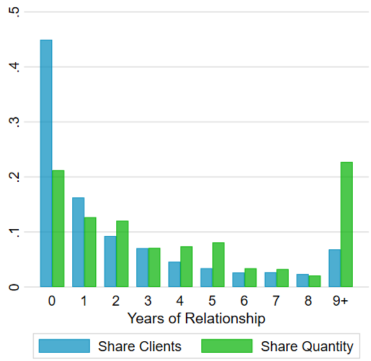
(h) ln(Price) over Time



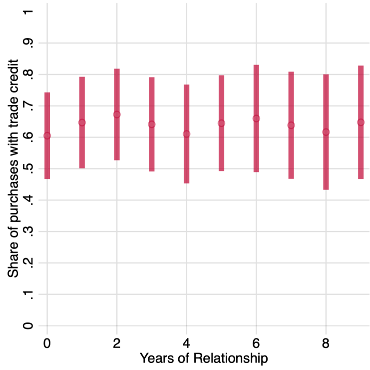
(i) ln(Q) over Time

Notes: This figure replicates Figure 1 for Pharmaceutical-Products only. Subfigure (a) displays the distribution of the average of the share of clients and quantity sold by relationship age, calculated across all sellers in 2016. Subfigure (b) displays the average of the share of purchases channeled through trade-credit, along with a 90% confidence interval, calculated across all sellers. Subfigure (c) displays the evolution of standardized log quantities, with their corresponding 90% confidence intervals, calculated across all sellers. The standardized log quantity is obtained by taking the average quantity sold in a given year for each seller-product and subtracting the log average quantity for that year. The standard errors are calculated at the seller-year level. Subfigure d) shows the relationship between quantity purchased and standardized log unit price through a binscatter plot that displays the measure of unit price against the quantity sold, based on relationship age. The standardized log unit price is obtained by netting out the average log unit price for that year for each seller-product. The quantiles of quantity are calculated for each seller-relationship age combination. Subfigure e) presents a binscatter plot of standardized log unit prices against years of relationship, controlling for a flexible spline of standardized log quantities. The standard errors are calculated at the seller-year level. Subfigure f) displays a binscatter plot of the average survival rate of pairs at different ages and quantiles of quantity. Subfigure g) presents a binscatter of (log) average price on the quantile of quantity by relationship age, controlling for seller-year fixed effects. Subfigure h) presents a binscatter of (log) average price on years of relationship controlling for a flexible spline of quantity and seller-year fixed effects. Subfigure i) presents a binscatter of (log) total quantity on years of relationship controlling for seller-year fixed effects. The quantiles of quantities are calculated for each seller-age combination, and the error bars represent a 90% level of variation across all sellers.

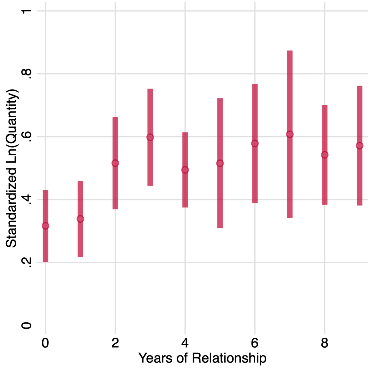
Figure OA-8: **Motivating Facts: Cement-Products**



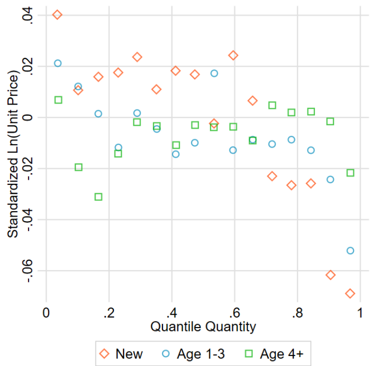
(a) Share of Clients and Trade



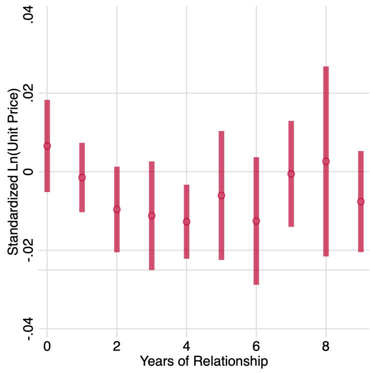
(b) Trade-credit



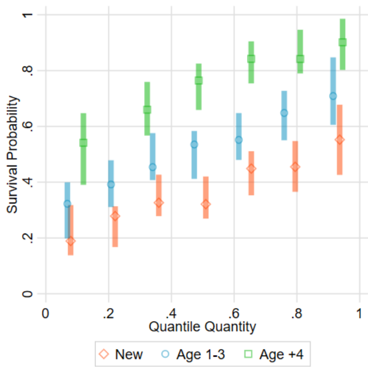
(c) Stdz. Q over Time



(d) Stdz. Price by Quantile

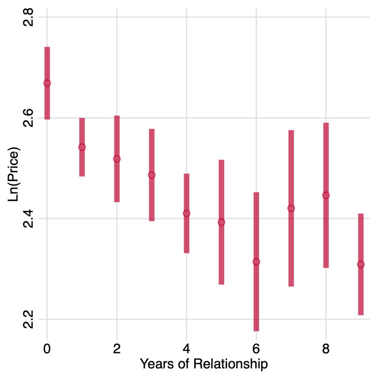


(e) Stdz. Price over Time

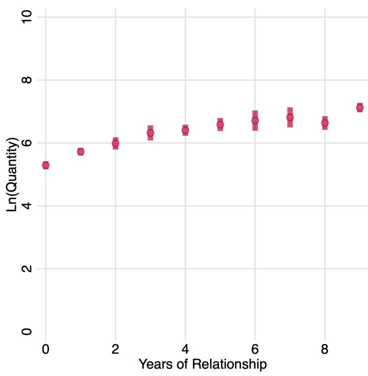


(f) Survival Rates

(g) ln(Price) by Quantile



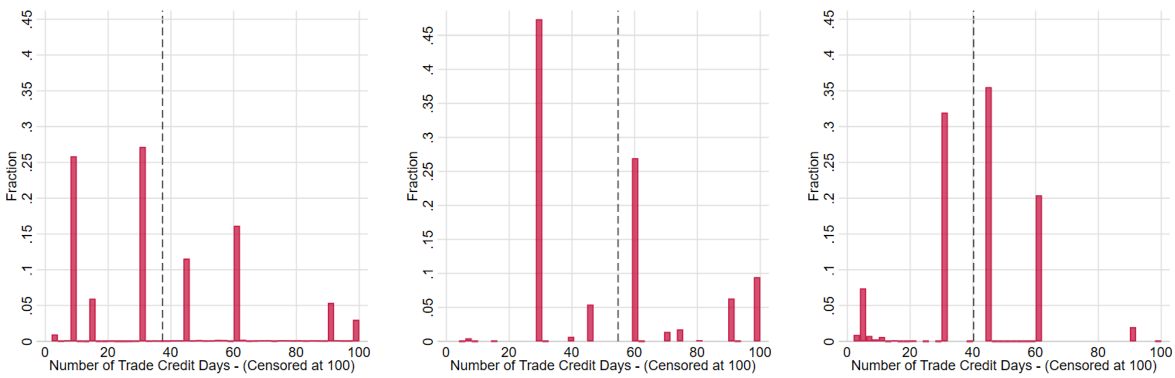
(h) ln(Price) over Time



(i) ln(Q) over Time

Notes: This figure replicates Figure 1 for Cement-Products only. Subfigure (a) displays the distribution of the average of the share of clients and quantity sold by relationship age, calculated across all sellers in 2016. Subfigure (b) displays the average of the share of purchases channeled through trade-credit, along with a 90% confidence interval, calculated across all sellers. Subfigure (c) displays the evolution of standardized log quantities, with their corresponding 90% confidence intervals, calculated across all sellers. The standardized log quantity is obtained by taking the average quantity sold in a given year for each seller-product and subtracting the log average quantity for that year. The standard errors are calculated at the seller-year level. Subfigure d) shows the relationship between quantity purchased and standardized log unit price through a binscatter plot that displays the measure of unit price against the quantity sold, based on relationship age. The standardized log unit price is obtained by netting out the average log unit price for that year for each seller-product. The quantiles of quantity are calculated for each seller-relationship age combination. Subfigure e) presents a binscatter plot of standardized log unit prices against years of relationship, controlling for a flexible spline of standardized log quantities. The standard errors are calculated at the seller-year level. Subfigure f) displays a binscatter plot of the average survival rate of pairs at different ages and quantiles of quantity. Subfigure g) presents a binscatter of (log) average price on the quantile of quantity by relationship age, controlling for seller-year fixed effects. Subfigure h) presents a binscatter of (log) average price on years of relationship controlling for a flexible spline of quantity and seller-year fixed effects. Subfigure i) presents a binscatter of (log) total quantity on years of relationship controlling for seller-year fixed effects. The quantiles of quantities are calculated for each seller-age combination, and the error bars represent a 90% level of variation across all sellers.

Figure OA-9: **Trade-credit Terms by Sector**



(a) Textiles (b) Pharmaceutical (c) Cement-Products

Notes: This figure plots the distribution of trade-credit days offered by the seller's sector.

OA-4 Model Propierties and a Solved Example

OA-4.1 Existence and Non-Stationarity

To prove existence, I build on two key results from the literature. First, I utilize the result of non-linear pricing from Jullien (2000) to demonstrate the existence of a stationary optimal contract in the presence of heterogeneous participation constraints. This is achieved by showing the equivalence between the stationary contract with limited enforcement and a non-linear pricing problem with heterogeneous outside options. Subsequently, similar to the argument in Martimort et al. (2017), I present a simple non-stationary deviation that outperforms the stationary optimal contract.

It is important to note that I will show existence results under the assumption of no exit, i.e., $X(\theta) = 0$ for all θ. To prove existence with exit, one must replace the discount factor δ with $\tilde{\delta} \equiv \min\{\delta(\theta)\}$, where $\delta(\theta) = \delta(1 - X(\theta))$ accounts for heterogeneous breakups. This adjustment only affects one of the assumptions discussed below and sets an upper bound on the worst-case exit rate.

OA-4.1.1 *Existence of Stationary Contract*

The model in Jullien (2000) solves the following problem:

$$\max_{\{t(\theta),q(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} [t(\theta) - cq(\theta)] f(\theta) d\theta \text{ s.t.} \tag{IR Problem}$$

$$v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\hat{\theta})) - t(\hat{\theta}) \ \ \forall \theta, \hat{\theta} \tag{IC}$$

$$v(\theta, q(\theta)) - t(\theta) \geq \overline{u}(\theta) \ \ \forall \theta. \tag{IR}$$

Under a modified first-order approach, the seller's first-order condition is given by:

$$v_q(\theta, q(\theta)) - c = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)), \tag{39}$$

for each type θ, and the complementary slackless condition on the IR constraints:

$$\int_{\underline{\theta}}^{\overline{\theta}} [u(\theta) - \overline{u}(\theta)] d\gamma(\theta) = 0. \tag{40}$$

Jullien (2000) shows that under three assumptions there exists a unique optimal solution in which all consumers participate. This solution is characterized by the first-order conditions 39 and complementary slackless condition 40 with $q(\theta)$ increasing.

The first assumption is potential separation (PS), which requires that the optimal solution is non-decreasing in θ, and satisfied under weak assumptions on the distribution of θ and the curvature of the surplus relative to the return of the buyer. In particular, it requires that

$$\frac{d}{d\theta}\left(\frac{S_q(\theta, q)}{v_{\theta q}(\theta, q)}\right) \geq 0$$

$$\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0 \geq \frac{d}{d\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right).$$

The second and *key* assumption is homogeneity (H), requiring that there exists a quantity profile $\{\overline{q}(\theta)\}$ such that the allocation with full participation $\{\overline{u}(\theta), \overline{q}(\theta)\}$ is implementable in that $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$ and $\overline{q}(\theta)$ is weakly increasing. This assumption implies that the reservation return can be implemented as a contract without excluding any type, ensuring that incentive compatibility is not an issue when the individual rationality constraint is binding.

Lastly, the assumption of full participation (FP) posits all types participate, and is satisfied when (H) holds and the surplus generated in the reservation return framework is greater than the private return to the buyer, i.e. $s(\theta,\overline{q}(\theta)) \geq \overline{u}(\theta)$.

I show that my setting can be rewritten in terms of Jullien (2000), implying that an optimal separating stationary contract exists. The seller chooses the optimal stationary contract $\{t(\theta),q(\theta)\}$ that satisfy incentive-compatibility and the limited enforcement constraint. Formally, the seller solves the problem:

$$\max_{\{t(\theta),q(\theta)\}} \frac{1}{1-\delta}\int_{\underline{\theta}}^{\overline{\theta}} [t(\theta)-cq(\theta)]f(\theta)d\theta \quad \text{s.t.} \tag{LE Problem}$$

$$v(\theta,q(\theta))-t(\theta) \geq v(\theta,q(\hat{\theta}))-t(\hat{\theta}) \;\; \forall\theta,\hat{\theta} \tag{IC}$$

$$\frac{\delta}{1-\delta}\Big(v(\theta,q(\theta))-t(\theta)\Big) \geq t(\theta) = v(\theta,q(\theta))-u(\theta), \;\; \forall\theta, \tag{LC}$$

where $u(\theta)$ is the return obtained by type θ. The limited enforcement constraint can be easily written as the IR constraint in Jullien (2000):

$$u(\theta) \geq (1-\delta)v(\theta,q(\theta)) \equiv \overline{u}(\theta) \;\; \forall\theta. \tag{LE'}$$

In my model, with $v(\theta,q)=\theta v(q)$, the first condition of assumption PS is always satisfied as

$$\frac{d}{d\theta}\left(\frac{S_q(\theta,q)}{v_{\theta q}(\theta,q)}\right) = \frac{d}{d\theta}\left(\theta-\frac{c}{v'(q)}\right) \geq 0 \Longleftrightarrow 1 \geq 0 \tag{A1}$$

As stated earlier, the second condition of assumption PS is satisfied for a wide-range of distributions for θ. Therefore, assumption PS is satisfied for any of those distributions.

Then, consider Assumption H. It requires that an allocation $\{\overline{q}(\theta)\}$ exists such that $\overline{u}'(\theta) = v_\theta(\theta,\overline{q}(\theta))$ and $\overline{q}(\theta)$ is weakly increasing. Notice that under LE', we can define $\overline{q}(\theta)$ as $\overline{u}'(\theta) = (1-\delta)[\theta v'(q(\theta))q'(\theta)+v(q(\theta)] = v(\overline{q}(\theta))$. Define $G(\overline{q},\theta) = v(\overline{q})-(1-\delta)[\theta v'(q(\theta))q'(\theta)+v(q(\theta))] = 0$. By the implicit function theorem, $\overline{q}(\theta)$ is weakly increasing if

$$\begin{aligned}
\overline{q}'(\theta) &= -\frac{dG/d\theta}{dG/d\overline{q}} \\
&= \frac{(1-\delta)[v'(q(\theta))q'(\theta)+\theta v''(q(\theta))(q'(\theta))^2+\theta v'(q(\theta))q''(\theta)+v'(q(\theta))]}{v'(\overline{q})} \geq 0 \\
&\Longleftrightarrow v'(q(\theta))[1+q'(\theta)+\theta q''(\theta))]+\theta v''(q(\theta))(q'(\theta))^2 \geq 0 \\
&\Longleftrightarrow \frac{q'(\theta)+\theta q''(\theta)+1}{\theta(q'(\theta))^2} \geq A(q) \\
&\Longleftrightarrow \left(\frac{T''(q)}{T'(q)}+A(q)\right)\Big(1+\theta(q)\theta'(q)r(q)+\theta'(q)\Big) \geq A(q) \\
&\Longleftrightarrow \frac{T''(q)}{T'(q)}\frac{M(q)}{M(q)-1} \geq A(q),
\end{aligned}$$

where $M(q) \equiv 1+\theta(q)\theta'(q)r(q)+\theta'(q)$ and $r(q) = g^{-1}(q)$ for $g(\theta) \equiv q''(\theta)$. As we expect

$T''(q) < 0$ and $T'(q) > 0$, it is necessary that $M(q)/(M(q)-1) < 0$. Such condition will be satisfied if $M(q) < 1$ and $M(q) > 0$, which imply that

$$r(q)\theta(q) < -1 \quad \text{and} \quad \theta'(q) < \frac{1}{\theta(q)|r(q)|-1}. \tag{A2}$$

The first condition sets restrictions on the rate of change of quantities, which requires $q''(\theta)$ to be negative, restricting how convex $u(\theta)$ can be. The second condition requires that quantities increase at a minimum rate. Moreover, the condition sets bounds on the price discounts offered relative to the buyers' return curvature at a given quantity.

Lastly, full participation requires H to hold as well as $s(\theta,\overline{q}(\theta)) \geq (1-\delta)\theta v(\overline{q}(\theta))$. The condition becomes:

$$\delta \geq \frac{c\overline{q}(\theta)}{\theta v(\overline{q}(\theta))}, \tag{A3}$$

which requires that agents value the future high enough, such that discount factor be greater than the ratio of average cost to average return.

Let $\{t^{st}(\theta), q^{st}(\theta)\}$ be the solution to the to the problem characterized by equations 39 and 40. Assuming that the primitives $v(\cdot)$, $F(\theta)$, and δ are such that conditions A1, A2, and A3 hold for $\{t^{st}(\theta), q^{st}(\theta)\}$, then $\{t^{st}(\theta), q^{st}(\theta)\}$ is uniquely optimal.

OA-4.1.2 *Solution to Stationary* $\Gamma^{st}(\theta)$

The seller's first-order condition defines the following differential equation in the stationary equilibrium

$$\theta u'(q^{st}(\theta)) - c = \frac{\Gamma^{st}(\theta) - F(\theta) + (1-\delta)\theta\gamma^{st}(\theta)}{f(\theta)} u'(q^{st}(\theta)). \tag{41}$$

The solution $\Gamma^{st}(\theta)$ to the equation above is given by:

$$\Gamma^{st}(\theta) = \frac{\int_{\underline{\theta}}^{\theta} x^{\delta/(1-\delta)}[xf(x) - c(u'(q^{st}(x))^{-1}f(x) + F(x)]dx + K}{\theta^{1/(1-\delta)}(1-\delta)}, \tag{42}$$

which by integration by parts reduces to:

$$\Gamma^{st}(\theta) = \frac{F(\theta)}{1-\delta} - \frac{\delta \int^{\theta} x^{\delta/(1-\delta)} F(x)dx}{(1-\delta)\theta^{1/(1-\delta)}} - \frac{cE[x^{\delta/(1-\delta)} u'(^{st}q(x))^{-1} | x \leq \theta]}{(1-\delta)\theta^{1/(1-\delta)}} + \frac{K}{(1-\delta)\theta^{1/(1-\delta)}} \tag{43}$$

The constant is obtained by using the boundary condition $\Gamma^{st}(\overline{\theta}) = 1$. Therefore,

$$K = cE[x^{\delta/(1-\delta)}u'(q^{st}(x))^{-1})] - \delta\overline{\theta}^{1/(1-\delta)} + \delta\int x^{\delta/(1-\delta)}F(x)dx. \tag{44}$$

OA-4.1.3 *Optimality of Non-Stationary Contracts*

Having established the existence of an optimal stationary contract, I show the optimality of non-stationary contracts.

Proposition 3. *If a non-stationary optimal contract exists, then it dominates the optimal stationary contract.*

Proof of Proposition 2. Consider the following deviation from the stationary contract, in which at tenure 0, the return obtained by the buyer is given by:

$$u_0(\theta) = u^{st}(\theta) - \varepsilon,$$

for some $\varepsilon > 0$ sufficiently small, where $u^{st}(\theta) = \theta v(q^{st}(\theta)) - t^{st}(\theta)$ and $t_0(\theta) = t^{st}(\theta)$. Define $q_0(\theta)$ to satisfy this deviation. Under this deviation, the enforcement constraint at $\tau = 0$ is:

$$t^{st}(\theta) \leq \frac{\delta}{1-\delta}\left[\theta v(q^{st}(\theta)) - t^{st}(\theta)\right],$$

which is identical to the one in the stationary contract, which we know $\{t^{st}(\theta), q^{st}(\theta)\}$ satisfy. Moreover, the incentive compatibility constraint is still satisfied as $\hat{\theta}$ maximizes

$$u_0(\theta,\hat{\theta}) + \frac{\delta}{1-\delta}u^{st}(\theta,\hat{\theta}) = \frac{\delta}{1-\delta}u^{st}(\theta,\hat{\theta}) - \varepsilon,$$

where $u_\tau(\theta,\hat{\theta}) \equiv \theta v(q_\tau(\hat{\theta})) - t_\tau(\hat{\theta})$.

Under this alternative scheme, the seller obtains an additional payoff ε per buyer while still satisfying both the incentive compatibility and limited enforcement constraints. Therefore, if it exists, the optimal non-stationary contract dominates the optimal stationary one. □

OA-4.2 Model Dynamics

Quantity Discounts

Define $T_\tau(q_\tau(\theta)) \equiv t_\tau(\theta_\tau(q))$, $\Lambda_\tau(\theta) \equiv \Gamma_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta)) + \theta\gamma_\tau(\theta)$, and $\lambda_\tau(\theta) \equiv d\Lambda_\tau/d\theta$. The price schedule is said to feature quantity discounts if $T''_\tau(q) < 0$.

Proposition 4. *Assume strict monotonicity of quantity $q'_\tau(\theta) > 0$ and that $\lambda_\tau(\theta) < f_\tau(\theta)$. If the densities $f_\tau(\theta)$ satisfy log-concavity and $d(F_\tau(\theta)/f_\tau(\theta))/d\theta \geq F_\tau(\theta)/[(\theta-1)f_\tau(\theta)]$, then the tariff schedule exhibits quantity discounts, $T''_\tau(q) \leq 0$ for each $q = q_\tau(\theta)$, $\theta \in (\underline{\theta},\overline{\theta})$ and τ.*

Proof of Proposition 3. Recall the quantity function $q_\tau(\theta)$ and its inverse function $\theta_\tau(q)$. Further differentiating the derivative of the incentive-compatible tariff schedule $T'_\tau(q_\tau(\theta)) = \theta v'(q_\tau(\theta))$ gives:

$$T''_\tau(q) = \theta'_\tau(q)v'(q) + \theta_\tau(q)v''(q) = \theta(q)v'(q)\left[\frac{\theta'_\tau(q)}{\theta_\tau(q)} + \frac{v''(q)}{v'(q)}\right] \tag{45}$$

$$= T'(q)\left[\frac{1}{\theta_\tau(q)q'_\tau(\theta)} - A(q)\right], \tag{46}$$

for $A(q) = -v''(q)/v'(q)$ and $\theta'_\tau(q) = 1/q'_\tau(\theta)$.

By implicit differentiation on the seller's first-order condition, we obtain an expression for $q'_\tau(\theta)$:

$$q'_\tau(\theta) = -\frac{\frac{d}{d\theta}\left[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\right]v'(q_\tau(\theta))}{\left[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\right]v''(q_\tau(\theta))}$$

$$= \frac{1}{A(q_\tau(\theta))}\frac{\frac{d}{d\theta}\left[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\right]}{\left[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\right]}$$

From SFOC, the denominator of the equation above is positive as $v'(q_\tau(\theta)) > 0$ and $c > 0$. By assumption, strict monotonicity holds ($q'_\tau(\theta) > 0$), which implies that the numerator is also positive. Substituting into (45) and using the fact that $T'_\tau(q) > 0$ and $A(q_\tau) > 0$, quantity discounts $T''_\tau(q) \leq 0$ hold if and only if

$$\frac{\left[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\right]}{\theta\frac{d}{d\theta}\left[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\right]} \leq 1 \tag{47}$$

Inequality 47 holds if

$$\theta - \frac{\Lambda_\tau(\theta) - F_\tau(\theta)}{f_\tau(\theta)} \leq \theta - \theta\frac{(\lambda_\tau(\theta) - f_\tau(\theta))f_\tau(\theta) - (\Lambda_\tau(\theta) - F_\tau(\theta))f'_\tau(\theta)}{f_\tau(\theta)^2}.$$

Rearranging, one obtains

$$[\Lambda_\tau(\theta) - F_\tau(\theta)][f_\tau(\theta) + f'_\tau(\theta)\theta] \geq \theta f(\theta)[\lambda_\tau(\theta) - f_\tau(\theta)]. \tag{48}$$

From the positive denominator above, one can obtain that $\theta f_\tau(\theta) \geq \Lambda_\tau(\theta) - F_\tau(\theta)$. Moreover, note that the log-concavity of the density $F_\tau(\theta)$ is sufficient to satisfy the standard assumption of the monotone hazard condition. So concentrating on log-concave densities, the following inequality holds: $f_\tau(\theta) \geq f'_\tau(\theta)\theta$. Therefore, if $\Lambda_\tau(\theta) > F_\tau(\theta)$, then a sufficient condition for quantity discounts is $\lambda_\tau(\theta) < f_\tau(\theta)$.

Instead if $\Lambda_\tau(\theta) < F_\tau(\theta)$, one can write 48 as

$$(\theta - 1)f_\tau(\theta) + f_\tau(\theta) \geq [F_\tau(\theta) - \Lambda_\tau(\theta)]\left(1 + \frac{f'_\tau(\theta)\theta}{f_\tau(\theta)}\right) + \lambda_\tau(\theta). \tag{49}$$

If $f'_\tau(\theta) < 0$, then a sufficient condition is $(\theta - 1)f_\tau(\theta) \geq F_\tau(\theta)$. If $f'_\tau(\theta) > 0$, then a sufficient condition is that $(\theta - 1)f(\theta) \geq F_\tau(\theta)(1 + \theta f'_\tau(\theta)/f_\tau(\theta))$. Both conditions can be expressed as:

$$\frac{d}{d\theta}\left(\frac{F_\tau(\theta)}{f_\tau(\theta)}\right) = \frac{f_\tau(\theta)^2 - F_\tau(\theta) f_\tau'(\theta)}{f_\tau(\theta)^2} \geq \frac{F_\tau(\theta)}{(\theta-1) f_\tau(\theta)}. \tag{50}$$

$\square$

Intuitively, the condition states that for a general class of distributions, as long as the incentive-compatibility marginal effects dominate those of the limited enforcement, the seller finds it optimal to offer quantity discounts at any relationship age. This condition is likely to be satisfied if the limited enforcement constraint is slack for some buyers even at their first interaction. Moreover, it also requires the enforcement constraint to be slack for all buyers in the long run. This last requirement aligns with the model of Martimort et al. (2017), where buyers reach a *mature* phase in which the constraints no longer bind. This is also consistent with Proposition 4 below, which finds that trade reaches a mature phase.

In terms of generality, the usual monopolist screening problem requires (or uses) log-concavity of $f(\theta)$.[48] I am strengthening the requirement that the evolution of the distribution also satisfies log-concavity, implicitly placing bounds on the distribution of exit rates over types.

The second condition strengthens the requirements on the dynamic distribution of types to ensure that the seller desires to price discriminate across types.

An alternative way to consider this property is to use (t-RULE) to obtain that the tariff schedule is concave if and only if $q_\tau'(\theta) > \frac{v'(q_\tau(\theta))}{-v''(q_\tau(\theta))\theta}$. As long as quantities increase by types fast enough, the seller will offer quantity discounts. The rate at which the quantities have to increase is determined by the level of the type and the curvature of the return function.

Evolution of Quantities

Next, I discuss how quantities evolve in Proposition 4.

Proposition 5. *For each θ, quantity increases monotonically in τ (i.e., $q_\tau(\theta) \leq q_{\tau+1}(\theta)$) if and only if the limited enforcement constraint is relaxed over time ($\gamma_\tau(\theta) \geq \gamma_{\tau+1}(\theta)$). Moreover, there is a time τ^* such that $\forall \tau \geq \tau^*$, $\gamma_{\tau^*}(\theta) = 0$ for all $\theta > \underline{\theta}$ and $q_{\tau^*}(\theta) \geq q_\tau(\theta)$ for all $\tau < \tau^*$ and all θ.*

Proof of Proposition 4. Notice that by the seller's first-order condition and $v'(\cdot) > 0$, $q_\tau(\theta) \leq q_{\tau+1}(\theta)$ holds if and only if

$$\begin{aligned} V_\tau(\theta) &\equiv \frac{\Gamma_\tau(\theta) - F_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta)) + \theta\gamma_\tau(\theta)}{f_\tau(\theta)} \\ &\geq \frac{f_\tau(\theta)}{f_{\tau+1}(\theta)} \frac{\Gamma_\tau(\theta) - F_{\tau+1}(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta)) + \theta\gamma_{\tau+1}(\theta)}{f_\tau(\theta)} + \frac{\Gamma_{\tau+1}(\theta) - 1}{f_{\tau+1}(\theta)} \equiv V_{\tau+1}(\theta) v, \end{aligned}$$

which can be written as

$$V_\tau(\theta) \geq \frac{f_\tau(\theta)}{f_{\tau+1}(\theta)} V_\tau(\theta) + \frac{\Gamma_{\tau+1}(\theta) - 1}{f_{\tau+1}(\theta)} + \frac{\theta[\gamma_{\tau+1}(\theta) - \gamma_\tau(\theta)]}{f_{\tau+1}(\theta)} - \frac{F_{\tau+1}(\theta) - F_\tau(\theta)}{f_{\tau+1}(\theta)}.$$

[48] Log-concavity of a density function $g(x)$ is equivalent to $g'(x)/g(x)$ being monotone decreasing. Families of density functions satisfying log-concavity include: uniform, normal, extreme value, exponential, amongst others.

With no selection pattern, i.e. $f_\tau(\theta) = f_{\tau+1}(\theta)$, the condition reduces to

$$\frac{1-\Gamma_{\tau+1}(\theta)}{f_\tau(\theta)} \geq \frac{\theta[\gamma_{\tau+1}(\theta)-\gamma_\tau(\theta)]}{f_\tau(\theta)}.$$

As $\gamma_\tau(\theta) > 0$ by assumption and the left-hand side is (weakly) positive due to $\Gamma_{\tau+1}(\theta) \leq 1$, a sufficient condition is that $\gamma_{\tau+1}(\theta) < \gamma_\tau(\theta)$. To obtain necessity, consider the Lagrangian keeping future return U^+ constant. The seller chooses $q(\theta)$ maximizing the following program:

$$L(\theta, U, q, \lambda, \gamma) = (\theta v(q(\theta)) - cq(\theta) - U)f(\theta) + \lambda v(q(\theta)) + \gamma(U + \delta U^+ - \theta v(q(\theta))), \quad (51)$$

where λ is the co-state variable for the incentive-compabilitity constraint and γ is the multiplier for the limited enforcement constraint. Noting that the necessary conditions are also sufficient (Seierstad and Sydsaeter, 1986) (pg. 276), the relevant optimality conditions are:

$$f(\theta)[\theta v'(q(\theta)) - c] + \lambda(\theta)v'(q(\theta)) = \gamma(\theta)\theta v'(q(\theta))$$

and

$$\dot{\lambda}(\theta) = f(\theta) - \gamma(\theta)$$

which imply

$$\gamma(\theta) = f(\theta) - \frac{cf(\theta)}{\theta v'(q(\theta))} + \frac{F(\theta)-\Gamma(\theta)}{\theta}.$$

Therefore, a higher level of quantity $q(\theta)$ is implied by a lower $\gamma(\theta)$.

Next, to obtain that $\gamma_\tau(\theta) = 0$ for some finite $\tau > \tau^*$ for all $\theta > \underline{\theta}$. Suppose otherwise, such that $\gamma_\tau(\tilde{\theta}) > 0$ for some $\tilde{\theta}$ and all τ. Then, $\Gamma_\tau(\theta) < 1$ for all $\theta \leq \tilde{\theta}$. Therefore, $1 - \Gamma_\tau(\theta) > 0$ for all $\theta \leq \tilde{\theta}$. Thus, as $\tau \to \infty$, $\sum_{s=0}^{\tau}(1-\Gamma_s(\theta)) \to \infty$ for all $\theta \leq \tilde{\theta}$. Thus, as long as $q_\tau(\theta) < \infty$ for all θ, τ, it must be the case that some finite τ^* exists such that $\gamma_\tau(\theta) = 0$ for all $\tau > \tau^*$ and for all θ. It is possible however for enforcement constraints to bind for $\underline{\theta}$, as in that case $\Gamma_\tau(\underline{\theta}) = 1$ and quantities would be finite.

Finally, to obtain $q_{\tau^*}(\theta) \geq q_\tau(\theta)$ for all $\tau < \tau^*$ and all θ. Notice that $q_{\tau^*}(\theta) \geq q_\tau(\theta)$ if and only if

$$\theta\gamma_\tau(\theta) + \sum_{s=\tau+1}^{\tau^*-1}(1-\Gamma_s(\theta)) \geq 0,$$

which always holds. It holds with strict inequality whenever the enforcement constraint binds at period τ, or when it binds in some period between τ and τ^* for some θ between $\underline{\theta}$ and θ.

□

In the model, quantities go hand-in-hand with enforcement constraints. Although the exact path depends on further assumptions on the return function and the distribution of types, the model predicts that quantities will reach a mature phase in which constraints no longer bind, except perhaps for the lowest type. At this mature phase, quantities will be at their highest level in the relationship.

Discounts over time

The model also offers conditions under which discounts over time are observed.

Proposition 6. *If $M_{\tau+1}(\theta) \equiv q_{\tau+1}(\theta) - q_\tau(\theta) \geq 0$ for all θ and with strict inequality for $\underline{\theta}$, then $p_{\tau+1}(q) \equiv T_{\tau+1}(q)/q < T_\tau(q)/q \equiv p_\tau(q)$.*

Proof of Proposition 5. Use the marginal price function $T'_\tau(q) = \theta_\tau(q)v'(q)$. Average unit prices $p_\tau(q)$ for $q > 0$ are given by:

$$p_\tau(q) = \frac{T_\tau(q)}{q} = \frac{\int_0^q \theta_\tau(x)v'(x)dx}{q},$$

where I have used the normalization $T_\tau(0) = 0$ and the inverse function $\theta_\tau(q)$. Average prices decrease over time if and only if

$$\int_0^q \theta_\tau(x)v'(x)dx > \int_0^q \theta_{\tau+1}(x)v'(x)dx$$

$$\iff$$

$$\int_0^q [\theta_\tau(x) - \theta_{\tau+1}]v'(x)dx > 0.$$

By assumption, $q_\tau(\theta) \geq q_{\tau+1}(\theta)$ (and strictly so for $\underline{\theta}$). Thus, $\theta_\tau(q) > \theta_{\tau+1}(q)$ for all q and the inequality holds.

$\square$

As long as quantities (weakly) increase from τ to $\tau+1$, unit prices at any given q decrease. The intuition behind this result is that marginal prices match marginal returns. A right-ward shift in quantities for (some) buyers further lowers marginal returns, requiring a decrease in marginal prices as well. As such, average prices will be lower at each q as well.

To further understand the dynamics in the model, I present a solved two-type example in Online Appendix Section OA-4.4. The example illustrates the backloading of prices and quantities together with quantity discounts as a way to maximize lifetime profits for the seller while preventing opportunistic behavior from the buyer.

OA-4.3 Equilibrium Contracts under Relaxation of the Constraints

OA-4.3.1 *Perfect Enforcement and Complete Information*

Under complete information and full enforcement, the seller acts as a monopolist practicing first-degree price discrimination with a stationary contract $(t^{1d}(\theta), q^{1d}(\theta))$, defined as

$$\theta v'(q^{1d}(\theta)) = c \quad \text{and} \quad t^{1d}(\theta) = \theta v(q^{1d}(\theta)). \tag{1D-Q \& 1D-T}$$

The seller offers first-best quantities but extracts all the rents from the buyer (subject to an interim individual rationality constraint, $u_\tau(\theta) \geq 0$). This allocation is infinitely repeated over time. In this model, quantities and prices are constant over time, hence there are no dynamics. Moreover, while quantities increase by type, prices may be constant under some parametrizations of $v(\cdot)$.

OA-4.3.2 *Perfect Enforcement and Incomplete Information*

This setting is similar to the canonical repeated adverse selection problem (Baron and Besanko, 1984; Sugaya and Wolitzky, 2021). As the seller has commitment, there is no loss of generality in restricting the study to an infinite sequence menu $\{t_\tau(\theta), q_\tau(\theta)\}_{\underline{\theta},\overline{\theta}}$ that induces the agent to report their true type. The problem of the seller is maximizing profits subject to IC-B and interim individual rationality constraints ($u_\tau(\theta) \geq 0$).

The theoretical insights from Baron and Besanko (1984) apply in this setup.[49] The optimal dynamic contract with full enforcement is equal to repeated Baron-Myerson static contracts with quantities determined by:

$$\theta v'(q_\tau^{pe}) = c_\tau - \frac{1 - F_\tau(\theta)}{f_\tau(\theta)} v(q_\tau^{pe}(\theta)), \tag{PE-Q}$$

and tariffs such that

$$t_\tau^{pe}(\theta) = \theta v(q_\tau^{pe}(\theta)) - \int_{\underline{\theta}}^{\theta} v(q_\tau^{pe}(x))dx. \tag{PE-T}$$

To preserve incentive compatibility of the buyer, the seller offers higher quantities to higher types within a given period. Moreover, the price schedule is shown to feature quantity discounts under common classes of assumptions on the curvature of demand and the distribution of types (Maskin and Riley, 1984) .

Under positive selection (i.e., $X'(\theta) < 0, \forall \theta$), average and type-specific quantities *decrease* over time. Similarly, average and type-specific unit prices *increase*.[50] Instead, without selection patterns (i.e., $X'(\theta) = 0, \forall \theta$), the optimal full enforcement contract with asymmetric information is stationary.

Therefore, while asymmetric information is able to rationalize the observed quantities discounts, on its own, it is not able to rationalize the dynamics of quantities and prices under observed selection patterns.

OA-4.3.3 *Limited Enforcement and Complete Information*

Next, consider a model without adverse selection, where the buyer can default on trade-credit at any time. In this context, the seller selects trade profiles $\{t_\tau(\theta), q_\tau(\theta)\}_{\underline{\theta},\overline{\theta}}$ that maximize lifetime profits, subject only to the limited enforcement constraint (equation LE-B). This model is reminiscent of the models in Thomas and Worrall (1994), Ray (2002), and Albuquerque and Hopenhayn (2004), which feature quantity and price backloading, as those described in the reduced form section.

In particular, the optimal contract quantities are determined by the following equation:

$$\theta v'(q_\tau^{le}) = \frac{c_\tau}{1 - \gamma_\tau(\theta)}, \tag{LE-Q}$$

where $\gamma_\tau(\theta)$ is the Lagrange multiplier on the limited enforcement constraint.

[49] Theorem 4' offers the results for fully persistent types in an infinite horizon model.

[50] With positive selection, informational rents given to middle-types decrease, as the distribution is shifting towards higher-types $F_\tau(\theta) > F_{\tau+1}(\theta)$. In order to incentivize the highest types still active, middle-types will be distorted downwards in the future. Marginal unit prices are given by $p(q(\theta)) = c + (1 - F_\tau(\theta)/f_\tau(\theta)$ (Armstrong, 2016), which will be generally larger for each θ, and as such, average price will be larger at each q.

Without the need of an interim individual rationality constraint, the limited enforcement constraint generates dynamics, features an *initial phase*, in which quantities are set to zero for all types, for which $\gamma_0(\theta) = 1$, and a stationary *mature phase*, in which $\gamma_\tau(\theta) = X(\theta)$. More patient buyers, those with smaller $X(\theta)$, are closer to their first-best. Additionally, all else equal, higher types receive higher quantities.

The enforcement constraints are always binding, and the optimal tariffs are set as follows:

$$t_\tau^{le}(\theta) = \delta(\theta)\theta v(q_{\tau+1}^{le}), \tag{LE-T}$$

so tariffs are constant over time, but prices decrease between the *initial* and *mature* phase. Prices may vary by type, but in a simple CES model, prices are constant across types.

Therefore, limited enforcement generates backloading. This backloading is not a result of unequal discount rates, as it also appears in cases without exit. The intuition is that limited enforcement constraints create an asymmetry: the buyer compares current tariffs to future returns, so *ceteris paribus*, there is an incentive to minimize current quantities to maximize current profits. However, trade converges to the mature phase almost immediately, by the second period.

By including the additional interim individual rationality constraint ($u_\tau(\theta) \geq 0$), the initial phase lengthens. The reason is that the additional limited liability constraint forces quantity changes between periods to be smaller. The length of the initial path is dependent on the parameters for the buyer's return function and the discount factor. The higher the common discount factor or the lower the exit rate (the more patient the buyer), the longer the path before the mature phase. Similarly, the more responsive the return function, the longer the path. Though the path to convergence is longer, under a CES model, prices are constant across types within a given period.

OA-4.4 A Two-Type Illustrative Example

The purpose of this example is four-fold. First, I illustrate how the introduction of the limited enforcement constraint may distort quantities relative to perfect enforcement. Second, I show that lower types unambiguously reap higher net returns due to the enforcement constraint. The introduction of the enforcement constraints effectively raises their reservation return to participate in trade, forcing the seller to offer larger net return values to lower types. Third, I demonstrate that the optimal contract must be non-stationary. Fourth, I show through a solved example that the optimal stationary contract features *backloading*: unit prices decrease while quantities increase as relationships age.

OA-4.4.1 *Buyer's Types*

A buyer type-θ gains a gross return θq^β from q units of the product sold by the seller. Assume there are positive, yet diminishing marginal returns, i.e., $\beta \in (0,1)$. The buyer types can take values $\{\theta_L, \theta_H\}$, such that $\theta_L < \theta_H$. Let f_L (resp. f_H) be the probability that buyer is type L (resp. type H) and assume no exit, i.e., $X(\theta) = 0$.

OA-4.4.2 *A Stationary Contract*

For now, consider the optimal *stationary* contract. The optimal choice gives the buyer the net return $R(\theta_i) = \theta_i q_i^\beta - T(q_i)$. The seller designs the scheme to maximize:

$$max_{\{T_i,q_i\}} f_L(T_L - cq_L) + (1 - f_L)(T_H - cq_H)$$

where $T_i \equiv T(q_i)$, subject to incentive-compatibility constraints:

$$R(\theta_H) \equiv \theta_H q_H^\beta - T_H \geq \theta_H q_L^\beta - T_L, \tag{IC-H}$$

$$R(\theta_L) \equiv \theta_L q_L^\beta - T_L \geq \theta_L q_H^\beta - T_H. \tag{IC-L}$$

as well as the limited enforcement constraint:

$$\frac{\delta}{1-\delta}(R(\theta_i)) \geq T_i \quad i = L, H. \tag{LE-i}$$

This last constraint effectively (weakly) raises the minimum net rent that each buyer needs to obtain to participate in trade. The usual nonlinear pricing problem only requires that $R(\theta_i) \geq 0$. Instead, the limited enforcement case requires that $R(\theta_i) \geq (1-\delta)/\delta T_i > 0$, where the minimum return is endogenously determined. Notice that as $\delta \to 1$, the limiting case becomes the standard nonlinear pricing problem.[51]

To simplify the problem, assume that the IC-L and LE-H are slack while IC-H and LE-L are binding.[52] By using these assumptions on the constraints, one can obtain the optimal quantity allocations:

$$q_H^* = \left(\frac{\beta}{c}\theta_H\right)^{\frac{1}{1-\beta}},$$
$$q_L^* = \left(\frac{\beta}{c}\left[\theta_L - \frac{(1-\delta)\theta_L}{f_L} - \frac{(1-f_L)(\theta_H-\theta_L)}{f_L}\right]\right)^{\frac{1}{1-\beta}},$$

and optimal tariffs:

$$T_H^* = \theta_H q_H^\beta + (\delta\theta_L - \theta_H) q_L^\beta,$$
$$T_L^* = \delta\theta_L q_L^\beta.$$

The tariffs are similar to that in the standard case, with the exception that the discount factor now enters the terms multiplying θ_L. Therefore, for a given quantity, tariffs are lower for both types.

The program's solution implies there is no distortion in quantities for type-H, as they purchase at the first-best level. However, type-L's purchases are shifted downwards. First, as is common in adverse selection problems, their purchases are distorted downwards to incentivize the revelation of type-H.

Second, contrary to the standard problem, extracting all rents from type-L is no longer feasible, as type-L would default. This generates a second downward pressure for quantities, as the standard quantity allocation for θ_L (i.e., when $\delta = 1$), together with the optimal tariffs for L under limited enforcement do not satisfy IC-H. To see this, notice that as IC-H was binding in the standard problem, type-H was on the margin between their standard bundle and the standard bundle for type-L. Thus, if the limited enforcement bundle for type-L keeps quantities

[51] The theoretical result that the buyer benefits from a deterioration of enforcement was previously discussed by Genicot and Ray (2006). In their model, they find that if better enforcement brings with it the deterioration of outside options and the seller has the bargaining power, the buyer will see their expected payoff increase. The opposite holds when the buyer has the bargaining power.

[52] All slack constraints are verified for the numerical example discussed below.

fixed (relative to the standard menu) and at the same time asks for lower tariffs, type-H buyers would now prefer the menu intended for type-L. As a result, the seller needs to reduce type-L's allocation, even further than would be required under the standard adverse selection problem.

OA-4.4.3 *Non-Stationarity*

Relative to the standard problem, the seller now needs to offer positive net returns to all buyers, in order to prevent default. Contrary to the results in Baron and Besanko (1984), the stationary contract is no longer the optimal contract. Instead, the seller could offer a dynamic contract with intertemporal incentives that use the promise of future returns to the buyer to discipline their behavior now. Through this approach, the seller can extract higher shares of surplus early on than would be feasible under a stationary contract, increasing their present-value lifetime profits.

The exact dynamic path depends on the return function and distribution of types of the buyer, as well as the marginal cost of the seller and the common discount factor. For that reason, I consider next a solved numerical example.

OA-4.4.4 *A Visual Example*

To visualize the problem, I consider a numerical example with the following values for the parameters: $\beta = 0.5$, $c = 1$, $f_L = 0.95$, $\theta_L = 1$, $\theta_H = 3$, $\delta = 0.9$.[53] Besides the incentive compatibility constraint and the limited enforcement constraint, I have also included the interim individual rationality constraint.

Online Appendix Figure OA-10 shows the levels of quantities, prices, profits per buyer, and buyer's net return for the example discussed above for different regimes: stationary with perfect enforcement (Baron-Myerson), stationary with limited enforcement, and dynamic with limited enforcement.

In solid green, the figure shows the allocation for type-H. As mentioned above, limited enforcement of contracts does not distort their consumption relative to perfect enforcement. In solid blue, the figure shows the allocation for type-L under perfect enforcement. Type-L receives lower quantities and higher prices than type-H and receives zero net return. In dashed-dot blue, the figure shows the stationary allocation for type-L under limited enforcement. Relative to perfect enforcement, type-L sees a reduction in quantities and an increase in net return, in line with the logic explained above. Importantly, as the buyer's return function features diminishing returns in q, lower levels of quantity for lower values of δ also imply the seller can charge *higher* unit prices to type-L.

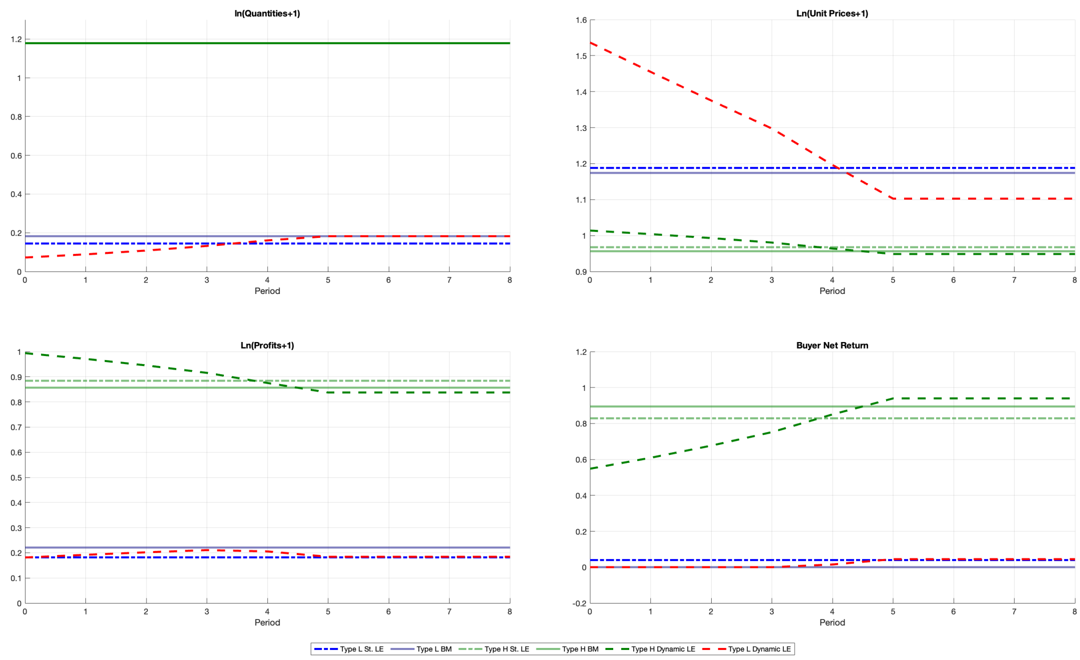
Lastly, the figure shows the optimal non-stationary path of prices and quantities in the dashed lines (red for type-L and green for type-H). The optimal path features *backloading* as quantities (weakly) increase and unit prices (weakly) decrease over time. As shown in the figure, this path of prices and quantities increases short-term expected profits from *each* buyer relative to the optimal stationary contract. Thus, the dynamics allow the seller to extract higher short-term profits for the high type as well. Indeed, in this example, the lifetime total profit in the dynamic case is 91% the level of the Baron-Myerson profit levels, whereas the stationary equilibrium reaches 88%. The seller can effectively prevent default now and increase present-value lifetime profits by offering higher surplus levels to the buyers in the future.

[53] The higher the difference between types, the higher the discount factor, the higher the elasticity β, or the bigger the share of high types, the longer the path to convergence.

Interestingly, the optimal path in the solved example features consumption for type-L in the long run that is greater than the stationary contracts with limited enforcement, as it converges to the Baron-Myerson allocation. Thus, in this case, the dynamics increase the long-term efficiency of the contracts.

In any case, the example shows that through the interaction market power on the seller side (which is reflected in the ability to offer incentive-compatible profit-maximizing menus) and the limited enforcement constraint, long-term contracts may display dynamics in which average quantities increase and unit prices decrease over time. Moreover, at any point in time, types consuming higher levels of quantities also enjoy lower unit prices. That is, this model of price discrimination with limited enforcement of contracts features i) *backloading* of prices and quantities, and ii) *quantity discounts* at any point in time.

Figure OA-10: **Example - Nonlinear Pricing and Limited Enforcement**



Notes: This figure shows Quantities, Prices, Profits, and Buyer Net Return for different enforcement and contract regimes. In dash-dot green, the optimal stationary contract for type-H under limited enforcement. In dashed green, the optimal dynamic contract for type-H under limited enforcement. In solid green, the optimal stationary contract for type-H under perfect enforcement. In solid blue, the optimal stationary contract for type-L under perfect enforcement. In dash-dot blue, the optimal stationary contract for type-L under limited enforcement. In dashed red, the optimal dynamic contract for type-H under limited enforcement. The parameters used in the example are: $\{\beta = 0.5,\ c = 1,\ f_L = 0.95,\ \theta_L = 1,\ \theta_H = 3,\ \delta = 0.9\}$.

OA-5 Proof of Lemma 1: $\Gamma_\tau(\overline{\theta}) = 1$

I prove that $\Gamma_\tau(\overline{\theta}) = 1$ for all τ. To begin, recall we assumed the outside option $\overline{u}_\tau(\theta)$ was equal to zero for all τ and all θ. Suppose instead that at some period k, the outside option is uniformly shifted downward by $\varepsilon > 0$ for all θ, that is, $\overline{u}_k(\theta) = -\varepsilon$. The enforcement constraint at k is now given by:

$$\delta[\sum_{s=1}^{\infty}\delta^{s-1}u_{k+s}(\theta)]-\bar{u}_k(\theta)=\sum_{s=1}^{\infty}\delta^{s}u_{k+s}(\theta)+\varepsilon\geq t_k(\theta)=\theta v(q_k(\theta))-u_k(\theta). \tag{52}$$

The seller's problem in the Lagrangian-form is

$$W(\varepsilon)=\max_{\{q_\tau(\theta),u_\tau(\theta)\}}\sum_{\tau=0}^{\infty}\delta^{\tau}\Big\{\int_{\underline{\theta}}^{\overline{\theta}}[\theta v(q_\tau(\theta))-cq_\tau(\theta)-u_\tau(\theta)]f(\theta)d\theta+ \tag{53}$$

$$\int_{\underline{\theta}}^{\overline{\theta}}[\sum_{s=1}^{\infty}\delta^{s}u_{\tau+s}+\varepsilon\times 1\{\tau=k\}-t_\tau(\theta)]d\Gamma_\tau(\theta)\Big\} \tag{54}$$

such that $u'_\tau(\theta)=\theta v'(q_\tau(\theta))$ for all τ,θ. The change in the value of the seller's problem given the uniform change in outside options is:

$$\frac{dW(\varepsilon)}{d\varepsilon}=\delta^{k}\int_{\underline{\theta}}^{\overline{\theta}}d\Gamma_k(\theta), \tag{55}$$

where the integral is the cumulative multiplier.

I argue that the quantities that solve the original problem still maximize the current one but that the tariffs are all shifted upward by the constant ε. That is, if $q_\tau(\theta)$ is the solution for the problem with $\bar{u}_\tau(\theta)=0$ for all θ and all τ with associated $t_\tau(\theta)$, $q_\tau(\theta)$ is also the solution for the problem with outside options $\bar{u}_\tau(\theta)=-\varepsilon\times 1\{\tau=k\}$ for all θ and all τ with associated tariffs equal to $t_\tau(\theta)+\varepsilon\times 1\{\tau=k\}$. The value of the problem for the seller is:

$$W(\varepsilon)=\sum_{\tau=0}^{\infty}\delta^{\tau}\Big\{\int_{\underline{\theta}}^{\overline{\theta}}[t_\tau(\theta)+\varepsilon\times 1\{\tau=k\}-cq_\tau(\theta)]f(\theta)d\theta\Big\} \tag{56}$$

$$=\sum_{\tau=0}^{\infty}\delta^{\tau}\Big\{\int_{\underline{\theta}}^{\overline{\theta}}[t_\tau(\theta)-cq_\tau(\theta)]f(\theta)d\theta\Big\}+\delta^{k}\varepsilon. \tag{57}$$

So

$$\frac{dW(\varepsilon)}{d\varepsilon}=\delta^{k}. \tag{58}$$

Therefore, from equations 55 and 58, the cumulative multiplier for any k will satisfy the following property:

$$\Gamma_k(\overline{\theta})\equiv\int_{\underline{\theta}}^{\overline{\theta}}d\Gamma_k(\theta)=\frac{dW(\varepsilon)}{d\varepsilon}\frac{1}{\delta^{k}}=1. \tag{59}$$

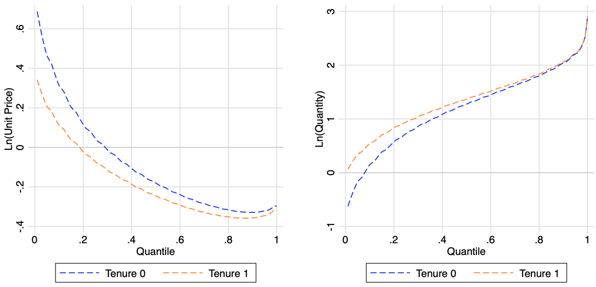
OA-6 Monte Carlo Study

The Monte Carlo studies the behavior of my estimators for two periods of a dynamic contract without breakups. I use the following design. The return function is $v(\theta,q)=\theta q^{1/2}$. The type distribution is Weibull with scale parameter equal to 1 and shape parameter equal to 2, $F(\theta)=1-exp(-(\theta-1)^{k})$, normalized so $\underline{\theta}=1$.[54] Marginal cost is 0.45. Although

[54]Recall that the model requires the type distribution to verify the monotone hazard rate condition, $\frac{d}{d\theta}\frac{F(\theta)}{f(\theta)}\geq 0\geq\frac{d}{d\theta}\frac{1-F(\theta)}{f(\theta)}$. Distributions that satisfy the monotone hazard rate condition include: Uniform, Normal, Logistic,

Figure OA-11: **Prices and Quantities by Quantile**

(a) **Unit Prices** (b) **Quantity**



Notes: These figures show the level of prices and quantities by quantile of quantity for tenure 0 and tenure 1 in the Monte Carlo simulation.

the multiplier function $\Gamma_\tau(\theta)$ is the solution to a differential equation linking the type distribution $F(\theta)$, the marginal cost, and the average base marginal return of types $\tilde{\theta} \leq \theta$, I parametrize it as a logistic distribution. In tenure 0, $\Gamma_0(\theta)$ has location parameter equal to 1 and scale parameter equal to 0.5. Instead, in tenure 1, $\Gamma_1(\theta)$ has location parameter 1 and scale 0.35. The lower scale parameter at tenure 1 reflects the idea that over time, the limited enforcement constraint is less binding. I construct the tariffs following Pavan et al. (2014): $t_\tau(\theta) = \theta q_\tau(\theta)^{1/2} - \int_{\underline{\theta}}^{\theta} q_\tau(x)^{1/2} dx$.

I randomly draw 1000 values of θ using $F(\theta)$ and obtain corresponding quantities $q_0(\theta)$ and $q_1(\theta)$ using the first-order condition of the seller and the assumed parametrizations of the return function, marginal cost, and multiplier at tenure 0 and 1. Then, I obtain the corresponding tariffs and I apply my estimator as defined in the previous sections to estimate $\{\theta, U(\cdot), \Gamma_\tau(\cdot)\}$. I repeat this 300 times to construct the dispersion for my estimates.

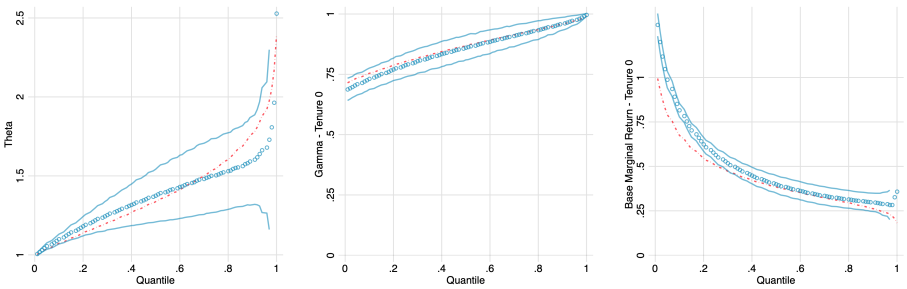
Online Appendix Figure OA-11 shows the (log) average prices and average quantities generated by the model for the two types of tenure. The model delivers quantity discounts (decreasing unit prices in θ), strict mononoticity of quantity (increasing quantities in θ), and backloading in the dynamic model, namely, further discounts and larger quantities in tenure 1 for each θ.

Online Appendix Figure OA-12 shows the results of the estimated Gamma distribution and the base marginal return, again in blue the estimated results and in red the true values. Both cases indicate good fit. Subfigure (a) shows the estimated $\hat{\theta}$ in blue and true θ in red by quantile. Dispersion at the 95 percent level are included for all except the top 2 quantiles, as they start to diverge. Overall, the figure shows a good fit, with most sections of including the true θ within their dispersion.

Next, I show the tenure 1's results estimates. Recall that the first-order condition of the seller now includes a backward-looking variable $1 - \Gamma_0(\theta)$ that keeps track of whether the limited commitment constraint was binding in the past. This variable is used by seller as a promise-keeping constraint that guarantees the seller delivers higher quantities and return in the future to prevent buyers from defaulting in the past. In my estimation, I use the tenure 0's predicted $\widehat{\Gamma}_0(\theta(\alpha))$ for each quantile α. Online Appendix Figure OA-13 shows the estimated Gamma distribution and the base marginal return. Although the fit is worse than in tenure 0, the dispersion of both gamma and the base marginal return include tend to include their true values.

Extreme Value (including Frechet), Weibull (shape parameter ≥ 1), Exponential, and Power functions.

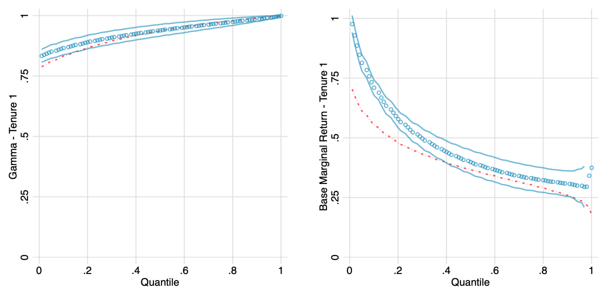
Figure OA-12: **Monte Carlo Results for Tenure 0**



(a) Types θ (b) Γ_0 Distribution (c) $v'(\cdot)$

Notes: Panel (a) plots the true (red) and estimated distribution of types (in blue) by quantile of quantity. Panel (b) plots the true (red) and estimated value (blue) of the LE multiplier for tenure 0 by quantile of quantity. Panel (c) plots the true (red) and estimated value (blue) of the base marginal return for tenure 0 by quantile of quantity. Error margins indicate ± 1.96 variation around estimated mean from 300 simulations.

Figure OA-13: **Monte Carlo Results for Tenure 1**



(a) Γ_1 Distribution (b) $v'(\cdot)$

Notes: Panel (a) plots the true (red) and estimated value (blue) of the LE multiplier for tenure 1 by quantile of quantity, with error margins indicating ± 1.96 variation around the estimated mean. Panel (b) plots the true (red) and estimated value (blue) of the base marginal return for tenure 1 by quantile of quantity, with error margins indicating ± 1.96 variation around the estimated mean from 300 simulations.

With respect to the differences between true and estimated functions, I find that the slight upward bias in the Gamma function for tenure 1 disappears if I use the true $\Gamma_0(\theta)$ function instead of the estimated $\widehat{\Gamma}_0$, suggesting that the bias is generated by sampling error in the tenure 0 estimates. Moreover, differences in the base marginal return for both tenure 0 and tenure 1 come from approximating the tariff function as log-linear. In the Monte-Carlo, the change in unit price is very steep for low-types, and this generates some approximation error for low-types in terms of the base marginal return function. Despite this error, the coefficient of the base return function is correctly estimated when using the assumed parametrization, observations of quantity, and the nonparametric estimates of $v'(\cdot)$ as target. In particular, the estimated coefficient cannot be rejected to be different from 0.5 (the assumed value in simulation).

OA-7 Evidence for Marginal Costs Constancy Assumption

I provide empirical support for the assumption of constant marginal cost in three ways.

First, I present evidence that average variable cost (AVC) is relatively constant over time.

For each seller i at time t, I construct *quarterly* measures of average cost by dividing total variable cost (intermediate inputs plus labor) in the quarter by total quantity sold in the quarter:

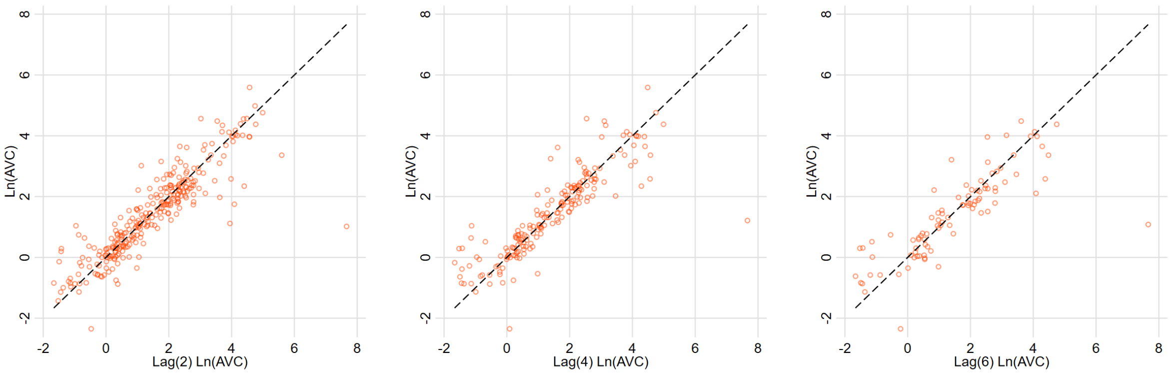
$$AVC_{it} = \frac{VC_i(Q_{it})}{Q_{it}},$$

where $VC_i(\cdot)$ is the variable cost function. Marginal cost is related to the previous equation via the derivative of the variable cost function: $MC_i(Q_{it}) = VC_i'(Q_{it})$. If marginal cost is constant, then $VC_i'(Q_{it}) = c_i Q_{it}$ and $AVC_{it} = c_i$. Therefore, strong serial correlation in AVC across periods indicates the following relationship:

$$AVC_{it} = c_i + \varepsilon_{it}.$$

Appendix Figure OA-14 presents a scatter plot of the (log) average variable cost on two, four, and six lags, with the dashed diagonal presenting a 1-to-1 fit. The figure shows that even after one and a half years apart, the average variable cost traces the diagonal fairly well.[55] This type of test is meaningful as sellers do experience variation in sales across months (Online Appendix Table OA-10), and therefore Q_{it} is non-constant.

Figure OA-14: **Serial Correlation**



(a) 2 Lags (b) 4 Lags (c) 6 Lags

Notes: These figures present the scatter plots of firm-level quarterly measures of average variable costs against 2 quarter lags (a), 4 quarters (b) and 6 quarters (c).

Second, I verify the constancy of average variable costs using a regression framework by regressing (log) average variable costs on seller fixed effects. I find that seller effects explain 87% of all variation using quarterly data and 84% using monthly data.

Third, under the assumption of constant marginal costs, we obtain the following accounting relationship for total variable costs: $VC_{it} = c_i Q_{it}$. Taking logs yields:

$$\ln(VC_{it}) = \ln(c_i) + \ln(Q_{it}).$$

This equation creates a testable framework for regression:

$$\ln(VC_{it}) = \beta_Q^c \ln(Q_{it}) + \ln(c_i) + \varepsilon_{it},$$

where $\beta_Q^c = 1$ under constant marginal costs, $\ln(c_i)$ is captured by a seller fixed effect, and ε_{it}

[55] A similar relationship exists if we focus only on monthly variation.

Table OA-11: Test for constancy of marginal cost

VARIABLES	(1) ln(VC)	(2) ln(VC)
ln(Q)	0.163**	0.757**
	(0.0723)	(0.302)
P-Value ($\beta_Q^c = 1$)	0.000	0.415
Observations	384	384
Seller FE	Yes	Yes
Time	Quarterly	Quarterly
Method	OLS	IV

Notes: This table presents the results of the test for constancy of marginal costs, of (log) total variables costs on (log) quantity. Column(1) reports OLS and Column (2) reports the instrumental variable results. Unit of observation is at the seller-quarter-level. Standard errors are clustered at the seller level. ***p<0.01, **p<0.05, *p<0.1

is noise, possibly stemming from model specification (i.e., true costs are non-constant and thus c_{it} is time-varying).

Notice that an OLS regression would not serve to test this equation if true marginal costs are time-varying, even if they are constant at the output level within the time period. An increase in true time-varying marginal cost is likely associated with a total decrease in quantity sold (as the seller increases prices to buyers). Thus, as quantity increases total variable costs, $\beta_Q^c > 0$, the negative relationship between costs and observed quantities implies downward bias in OLS due to omitted variable bias.

For that reason, I test this equation using an instrumental variable approach that exogenously shifts Q_{it} from changes in marginal costs captured by ε_{it}. Specifically, I use downstream demand shift-share style shocks in the spirit of Acemoglu et al. (2016) and Huneeus (2018). For a given selling firm i, I consider their 2015 demand share s_{ij}^{2015} over buyers j. Then, for each buyer, I regress their quarterly volume of log sales on buyer fixed effects and quarter-year fixed effects and collect the residuals as demand shocks $shock_{jt}^d$. For each seller, I obtain the weighted average of their exposure to potential demand shocks IV_{it}^d as follows:

$$IV_{it}^d = \sum_j s_{ij}^{2015} \times shock_{jt}^d.$$

I then run a regression for the testing equation using quarterly data at the seller level, using IV_{it}^d as an instrument for quarterly quantity Q_{it}.

Internet Appendix Table OA-11 shows the results. First, OLS (Column 1) shows a downward bias relative to the IV (Column 2), indicating some degree of model misspecification or measurement error in total quantity. Second, in the instrumental variable approach, we fail to reject that β_Q^c is equal to 1 (although, the point estimate is not precisely estimated at 1). Therefore, the test is again consistent with a constant marginal cost assumption.

Thus, all in all, the constant marginal cost assumption is not incredibly restrictive in this setting.

OA-8 Additional Estimation Results and Model Fit

OA-8.0.1 *Tariff Function*

Despite the simple approximation of the tariff function in equation 5, the within-tenure seller-specific tariff functions show a good fit. The average R-squared is close to 0.80, and the distribution of R-squared estimates for each seller-tenure (Figure OA-15) shows a good fit across the board.

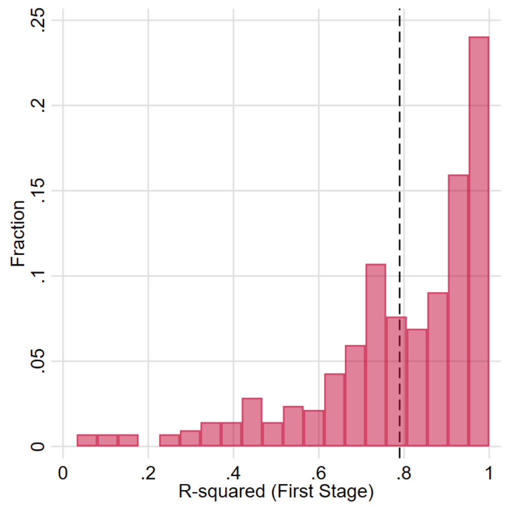


Figure OA-15: R-squared Distribution in the Estimation of the Tariff Function

Notes: This figure presents the distribution of R-squared values from seller-tenure-year regressions (equation 5).

Of course, as the fit is not perfect, it is worth highlighting some sources of measurement error in the tariff function. First, it is possible that the firm price schedule has higher-order terms, which would generate measurement error. However, this concern is small, as estimating a quadratic model only improves the R-squared on average by 0.008. Second, it is possible that, besides pricing on tenure and quantity, the firm is also pricing based on other unobservable characteristics (to the econometrician), which creates misspecification error, translating into measurement error. This would be particularly worrisome if the price schedule over quantities and tenure is not linearly separable from the other pricing characteristics. However, as shown in Table OA-6, the coefficients for prices on quantities and tenure are unaffected by the inclusion of a large set of buyer characteristics, supporting the assertion that pricing on other (plausibly unobserved) characteristics might enter as orthogonal measurement error.

OA-8.1 Survival Function Probability

Online Appendix Figure OA-16 presents estimated survival probabilities by age of relationship and quantile of quantity, with variation representing differences across seller-years.

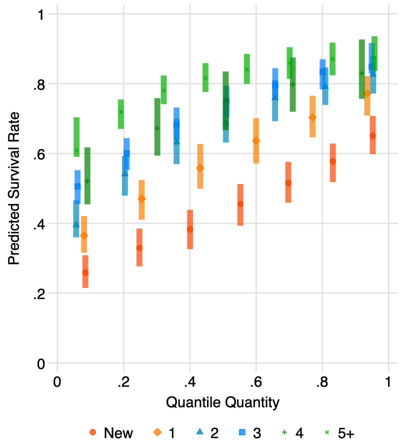


Figure OA-16: Survival Probability Function

Notes: This figure presents the estimated survival probability by quantile of quantity and age of relationship across seller-years. Confidence intervals represent the 90% level of variation across sellers, with standard errors clustered at the seller-year level.

OA-8.2 Distribution of t-Statistics against Standard Model Null

Online Appendix Table OA-12 shows the distribution of t-statistics for tests against a standard model null.

Table OA-12: **Distribution of t-Statistics**

	p10	p25	p50	p75	p90
Tenure 0	0.31	4.64	11.55	30.08	109.27

Notes: This table reports distribution of t-statistics for tests against a standard model null (e.g., $\Gamma_0(\cdot) = 1$).

OA-8.3 Parametrization of the Base Return Function

To conduct counterfactual experiments that consider quantities beyond those observed in the data, I parametrize the seller-specific buyer's return function $v(q) = kq^{\beta}$ for $k > 0$ and $\beta \in (0,1)$. This return function satisfies the modeling assumptions $v'(\cdot) > 0$ and $v''(\cdot) < 0$.

To estimate the parameters, I consider tenure 0 transactions between buyer i and the seller at a given year and perform the following linear least squares regression:

$$\ln(\widehat{v'}_i) = \ln(k\beta) + (\beta - 1)\ln(q_i) + \varepsilon_i,$$

using $v'(q) = k\beta q^{\beta-1}$, the estimated base marginal returns $\widehat{v'}_i$, and under the assumption that ε_i is Gaussian error.

Online Appendix Table OA-13 presents the distribution of k and β.

Table OA-13: **Parameters of Return Function**

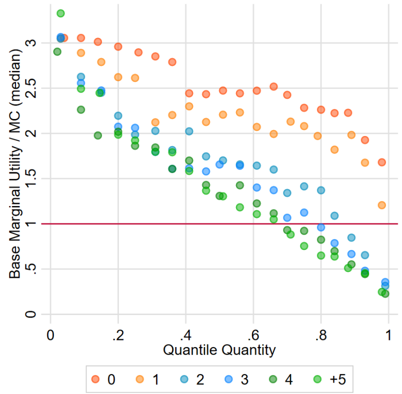
	mean	p10	p25	p50	p75	p90
β	0.56	0.30	0.48	0.61	0.76	0.82
k	171.23	9.00	17.24	39.64	86.61	282.40

Notes: This table reports distribution of estimated values for the ex-post parametrization of the return function.

OA-8.4 Economic Magnitudes: Base Marginal Return

Online Appendix Figure OA-17 presents a binscatter of the ratio of marginal revenue product (base marginal return) over marginal costs against the quantile of quantity, across sellers for tenure 0. It shows that the return of the input for the buyer is greater than the private marginal cost of providing it for the seller, for a majority of the buyers. For instance, the median buyer obtains 1.5 dollars of revenue for each dollar spent by the seller to produce the product.

Figure OA-17: **Base Marginal Return over Marginal Costs**



Notes: This figure plots the median of the ratio of base marginal return to marginal costs across sellers by quantile of quantity for each tenure.

OA-8.5 Model Fit

Online Appendix Figure OA-18 presents the statistical fit of the model across tenures. It plots a reordered equation I-EQ's left-hand side on the X-axis and the model's prediction using estimated coefficients of the right-hand side on the Y-axis.[56] Fit generally worsens for higher tenures; the results from Monte Carlo studies in Online Appendix OA-6 suggest that the decrease in statistical fit is driven by noise from using estimates for limited enforcement multipliers $\Gamma_s(\cdot)$ for earlier tenures s.

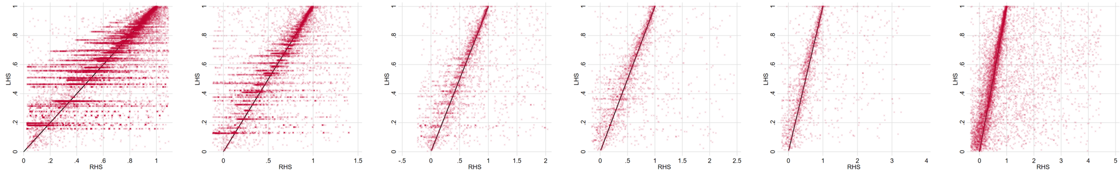
Online Appendix Figure OA-19 shows the fit in terms of quantities. To obtain quantities, I use the parametrization $v(q) = kq^{\beta}$, for $k > 0$ and $\beta \in (0,1)$ and the closed-form formula in 3.

[56]Reorder equation I-EQ to obtain:

$$\alpha = \Gamma_\tau(\alpha) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha)) - \left[\frac{T_\tau'(q_\tau(\alpha)) - c_\tau}{T_\tau'(q_\tau(\alpha))} - \gamma_\tau(\alpha)\right]\frac{\theta_\tau(\alpha)}{\theta_\tau'(\alpha)},$$

and use the estimated analogues of the right-hand side to make the predictions.

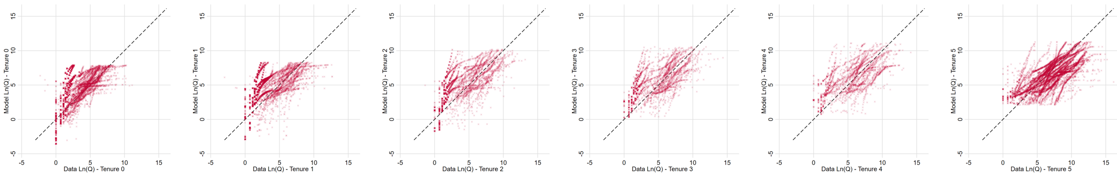
Figure OA-18: **Model Fit - Statistical**



(a) T0 (b) T1 (c) T2 (d) T3 (e) T4 (f) T5

Notes: These figures show binscatters of statistical fit of the model across tenures as implied by identification equation I-EQ. On the X-axis, it shows the predicted cumulative distribution function for the observation while on the Y-axis it plots the observed value.

Figure OA-19: **Model Fit - Quantities**

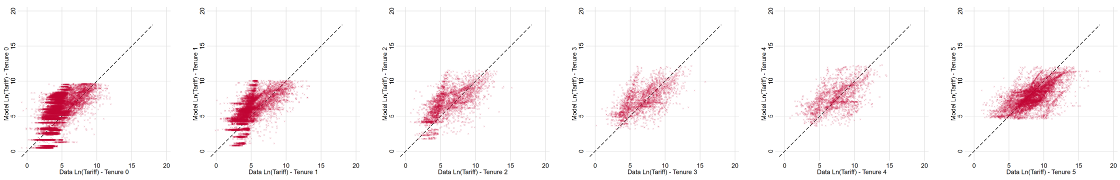


(a) T0 (b) T1 (c) T2 (d) T3 (e) T4 (f) T5

Notes: These figures display binscatters of model fit according to quantities. Estimated quantities use the close-form formula under the CES parametrization of the return function, as discussed in Appendix OA-13. The X-axis plots the observed (log) quantities and Y-axis model predicted (log) quantities.

Online Appendix Figure OA-20 shows the fit of tariffs. To generate tariffs in the model, I use the empirical equivalent of equation t-RULE.

Figure OA-20: **Model Fit - Tariffs**

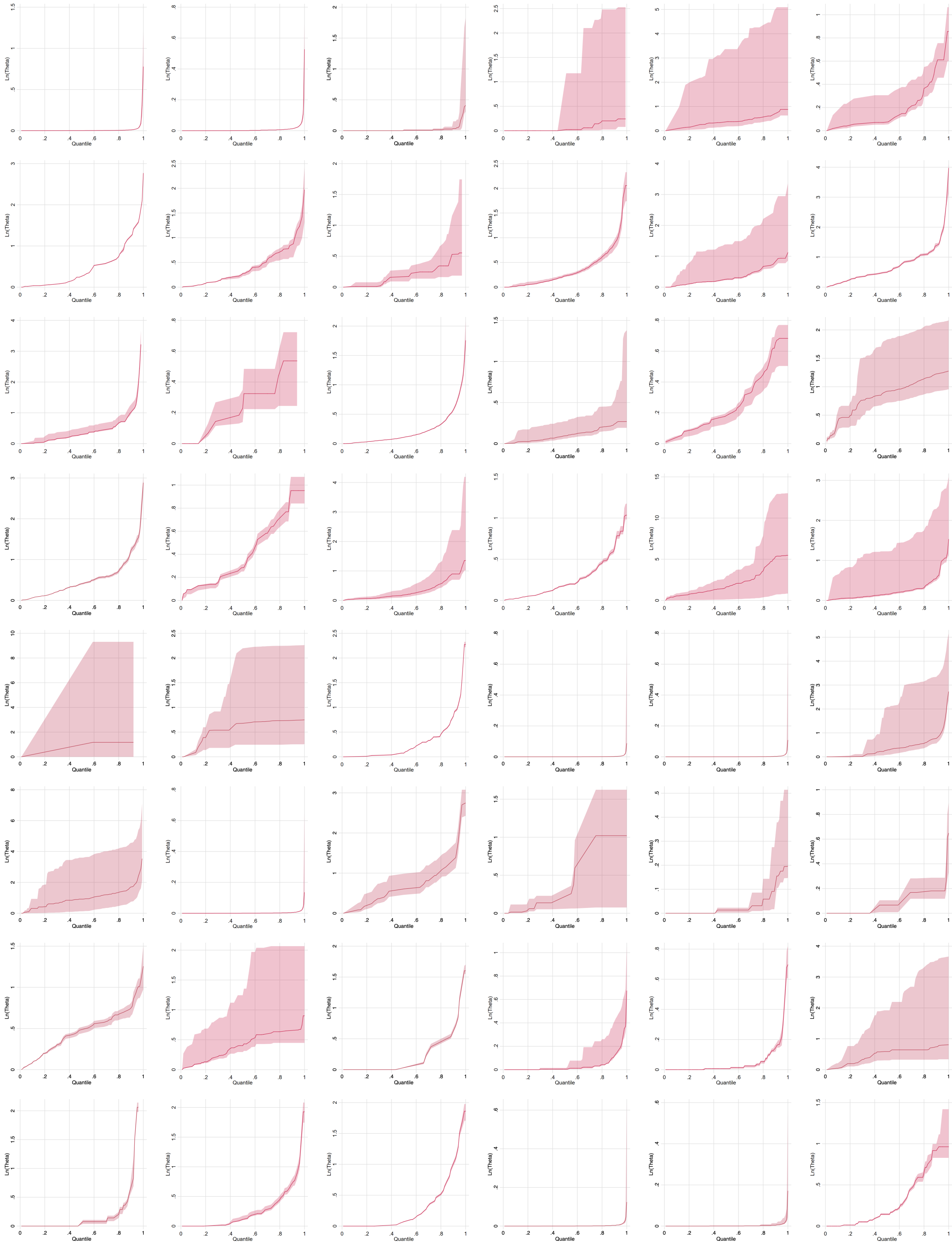


(a) T0 (b) T1 (c) T2 (d) T3 (e) T4 (f) T5

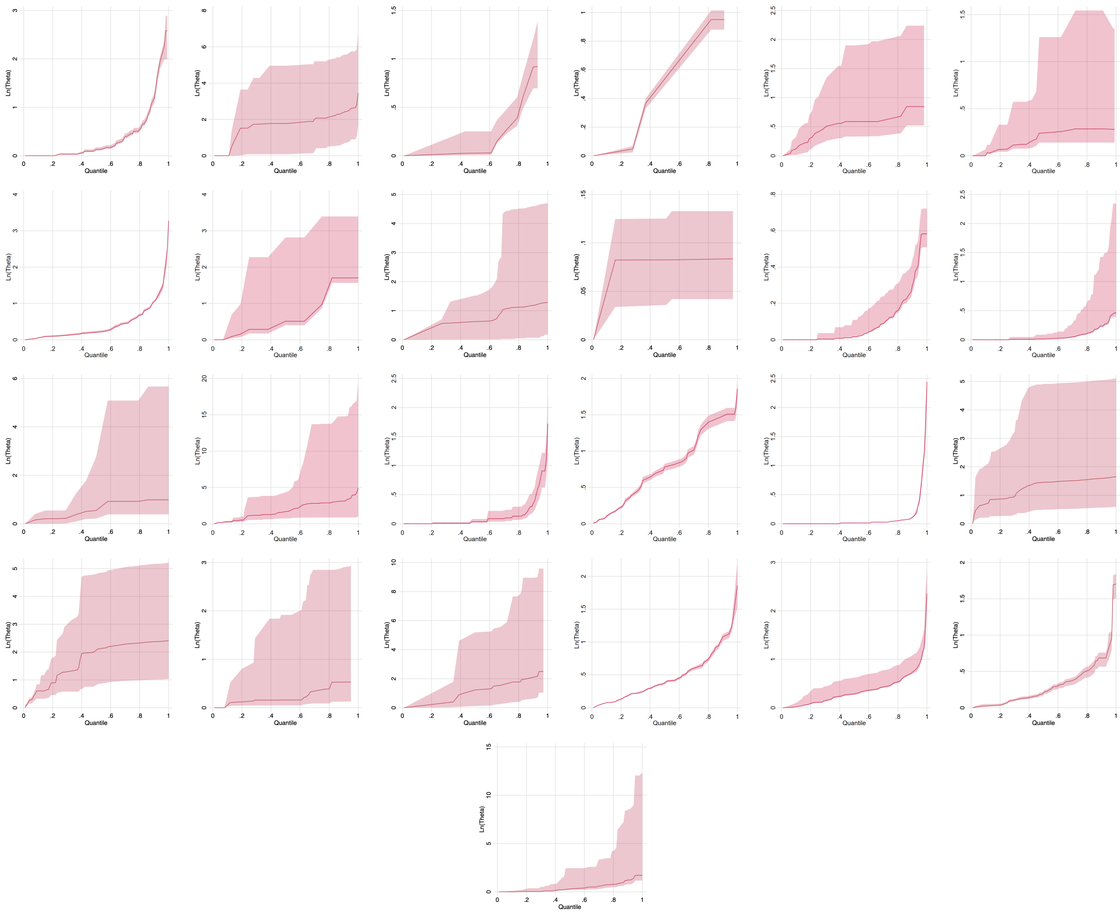
Notes: These figures display binscatters of model fit according to tariffs. Estimated tariffs are generated by using the empirical analogue of the tariffs rule t-RULE, taking as inputs estimated parameters θ, the parametrized return function $v(\cdot)$, and model generated quantities. The X-axis plots the observed (log) tariffs and Y-axis model predicted (log) tariffs.

OA-8.6 Bootstrapped Distribution of Types

Figure OA-21: Bootstrapped Distribution of Types



Bootstrapped Distribution of Types (Continued)



Notes: This figure plots distribution of types (log type $ln(\theta)$ by quantile of quantity) for each seller-year. The error bars show variation at the 90% confidence interval level, obtained from 30 bootstrapped simulations for each seller-year.

OA-9 Additional Counterfactual Results

OA-9.1 Computation of Counterfactuals

Counterfactual (a). I compute quantities based on the distribution of estimated types at different tenures and the quantity allocation equation 3 with $\Gamma_\tau(\cdot)$ set to 1 and $\gamma_\tau(\cdot)$ set to 0. I also set $\Gamma_s(\cdot)$ to 1 for $s < \tau$. With quantities in hand, the tariffs are set to satisfy incentive compatibility using equation t-RULE.

Counterfactual (b). Under the assumed base return function, the optimal uniform price is $p^l = c/\beta$ for any quantity. The corresponding type θ's demand is given by $q^l(\theta) = (k\beta\theta/p^l)^{1/(1-\beta)}$. This stationary menu will be insufficient for some enforcement constraints. Given exogenous hazard rates $X(\theta)$, the stationary enforcement constraint will be given by:

$$\delta(1 - X(\theta)) \geq \beta, \tag{U-LE}$$

which indicates that the rate of return captured by β has to be smaller than the buyer-specific discount rate. Notice that this limited enforcement constraint will hold for any other uniform price, so buyers who are willing to default at the optimal uniform price p^l will also be willing to default at any other alternative uniform price p^l_a, including $p^l_a = c$, which would generally imply an efficient allocation.

Under a monotonicity assumption on $X(\theta)$,the seller will set a minimum quantity $\underline{q}^l$ that the buyer needs to announce in order to be served.[57] In particular, it will only serve $q(\theta) \geq \underline{q}^l$, where $\underline{q}^l = min\{q^l(\theta) | \delta(1 - X(\theta)) \geq \beta\}$. In the counterfactual exercise, I set their quantities to zero to those θ with $q^l(\theta) < \underline{q}^l$.[58]

Counterfactual (c). Quantities and tariffs are those determined in Counterfactual 2. However, as buyers are precluded from the possibility of default, the seller serves all buyers. Thus, no quantity is set to zero.

OA-9.2 Results

This subsection presents comparisons of different counterfactual models relative to the baseline nonlinear pricing regime with limited enforcement. Online Appendix Table OA-14 shows all the results. The table present the *share of observations* in each percentile group for which each reported category (e.g., buyer's net return) is greater under the baseline than under the alternative. The main takeaways are the following.

Buyers. Small-quantity buyers tend to prefer limited enforcement of contracts over perfect enforcement. They can effectively use the threat of default to reap higher returns. In contrast, the median and top buyers prefer perfect enforcement in the short term but limited enforcement in the long term. Under weak enforcement of contracts, buyers prefer price discrimination over uniform pricing, as otherwise they would be excluded from trade (only median and top buyers prefer uniform pricing in the long term). However, if exclusion and default are restricted, most buyers prefer uniform pricing.

Sellers. Sellers prefer limited enforcement in the short term but perfect enforcement in the long term. Under weak enforcement of contracts, they enjoy the ability to price discriminate, as it allows them to sell to buyers that would otherwise be excluded from trade. In contrast, if

[57] The monotonicity on the hazard rate $X'(\theta) < 0$ is observed in the data.

[58] In this counterfactual exercise, I use an additional assumption: buyers demand truthfully the optimal level of quantity that is consistent with prices and full enforcement.

enforcement is strong, sellers prefer uniform pricing in the short term but price discrimination in the long term. This preference is driven by the rapid increase in quantities, despite the decrease in unit prices offered to most buyers as an incentive not to default.

Table OA-14: **Counterfactual Policies**

		Nonlinear + Perfect						**Uniform + Limited**						**Uniform + Perfect**					
		10%	25%	50%	75%	100%	Agg.	10%	25%	50%	75%	100%	Agg.	10%	25%	50%	75%	100%	Agg.
Buyer Return	*Tenure 0*	43.4	38.2	11.0	4.9	7.1	6.9	97.3	96.5	96.0	94.3	91.7	92.0	0.1	0.2	0.6	7.0	41.8	38.5
	Tenure 1	68.3	55.3	23.0	9.4	11.9	11.8	94.6	92.2	88.6	88.0	87.4	87.6	0.1	0.1	0.2	13.5	54.9	47.0
	Tenure 2	64.3	46.5	31.1	26.2	28.4	28.3	83.8	79.6	70.3	66.9	63.1	63.6	1.2	0.4	0.9	10.9	32.1	29.6
	Tenure 3	66.3	59.8	40.5	32.3	38.0	37.6	79.7	71.4	59.6	54.6	55.4	55.5	3.1	0.8	1.6	11.2	27.8	25.5
	Tenure 4	61.2	48.6	43.5	42.6	50.5	49.0	69.0	59.9	47.6	47.9	46.3	46.7	5.3	1.2	4.9	8.8	21.3	18.6
	Tenure 5	58.7	61.8	66.1	59.6	69.5	67.8	69.1	62.2	38.3	34.8	32.8	33.5	0.7	1.6	2.9	9.0	22.0	19.6
Seller Profit	*Tenure 0*	34.1	41.6	88.2	94.9	92.8	93.0	92.7	92.6	96.4	98.0	98.4	98.4	7.1	7.4	11.1	35.0	47.4	46.4
	Tenure 1	53.9	55.0	83.3	90.6	88.1	88.3	99.1	96.7	94.8	97.1	89.9	91.2	29.1	18.4	29.8	44.8	52.8	51.0
	Tenure 2	46.6	49.1	71.5	73.8	71.6	71.8	95.0	97.0	98.2	99.5	97.5	97.7	34.1	35.1	50.8	69.1	86.6	84.3
	Tenure 3	45.8	48.1	61.2	67.9	62.0	62.5	96.5	99.2	97.5	99.3	93.9	94.5	49.6	50.0	61.6	77.9	86.6	85.1
	Tenure 4	52.0	47.1	59.1	57.4	49.5	51.1	92.9	97.6	95.0	95.2	94.5	94.6	53.5	64.2	71.4	86.5	93.7	91.5
	Tenure 5	56.1	42.5	36.8	40.6	30.5	32.4	93.4	93.5	96.0	97.4	95.9	96.1	64.9	66.0	81.9	93.1	94.8	93.9
Surplus	*Tenure 0*	18.6	18.9	9.0	3.8	2.6	2.7	98.4	98.1	98.8	98.5	99.5	99.5	3.8	4.1	5.2	12.0	65.5	60.4
	Tenure 1	40.5	41.7	30.3	12.6	29.6	26.9	97.5	96.2	97.3	99.2	100.0	99.8	6.0	7.4	11.0	31.9	76.1	67.6
	Tenure 2	47.8	50.9	48.3	63.2	72.8	71.5	90.9	90.7	91.6	98.6	99.7	99.5	15.3	16.4	27.3	57.0	95.0	90.2
	Tenure 3	61.0	57.8	69.7	76.8	69.9	70.5	93.8	92.3	89.5	98.5	99.6	99.4	24.6	26.5	37.4	69.1	98.4	94.0
	Tenure 4	65.6	71.9	74.5	77.1	67.3	69.2	81.0	87.8	85.4	98.4	99.5	98.9	25.7	34.3	51.0	79.4	97.9	92.9
	Tenure 5	74.4	79.7	88.7	91.2	84.9	85.9	84.8	86.6	80.6	97.1	100.0	98.7	30.8	34.1	53.1	86.4	99.9	95.7
Unit Prices	*Tenure 0*	75.9	75.4	89.0	94.5	92.9	93.0	93.6	93.1	95.4	90.3	42.9	47.4	93.6	93.1	95.4	90.3	42.9	47.4
	Tenure 1	55.3	55.5	77.6	90.5	88.0	88.2	98.6	96.8	87.9	68.2	24.6	33.1	98.6	96.8	87.9	68.2	24.6	33.1
	Tenure 2	38.6	55.1	67.6	73.7	71.7	71.8	92.3	95.0	90.9	64.5	18.0	23.6	92.0	94.9	90.9	64.5	18.0	23.6
	Tenure 3	36.5	41.4	58.3	67.9	61.8	62.3	91.2	97.0	89.1	56.0	13.7	19.7	90.8	97.0	89.1	56.0	13.7	19.7
	Tenure 4	37.9	51.7	56.7	59.1	49.4	51.2	89.2	95.8	88.0	63.2	18.7	28.8	88.7	95.8	88.0	63.2	18.6	28.6
	Tenure 5	34.4	34.1	33.8	39.5	30.5	32.0	90.0	91.4	87.6	54.0	10.4	20.5	89.1	91.2	87.5	53.7	10.0	20.1
% Excluded	*Tenure 0*	-	-	-	-	-	-	97.3	96.4	95.8	94.1	90.5	90.9	-	-	-	-	-	-
	Tenure 1	-	-	-	-	-	-	93.4	91.9	88.6	87.3	85.8	86.1	-	-	-	-	-	-
	Tenure 2	-	-	-	-	-	-	81.5	77.8	70.1	65.7	61.3	61.9	-	-	-	-	-	-
	Tenure 3	-	-	-	-	-	-	76.9	69.0	59.5	51.5	50.0	50.4	-	-	-	-	-	-
	Tenure 4	-	-	-	-	-	-	66.8	58.1	47.5	44.7	43.5	50.0	-	-	-	-	-	-
	Tenure 5	-	-	-	-	-	-	65.3	58.8	37.5	29.8	25.4	26.7	-	-	-	-	-	-

Notes: This table reports the % share of observations for which the reported category (e.g., Buyer's Net Return) is greater under the observed nonlinear pricing regime than under the alternative policy. The values are reported across different tenures and percentile groups in the distribution of types. Percentile groups are defined based on quantiles as follows: the 10% group includes all buyers within seller-year-tenure quantiles from 0 to 10% (non-inclusive), the 25% group includes buyers within quantiles from 10% to 25% (non-inclusive), and this pattern continues for all other percentile groups. The policies considered are (a) Nonlinear pricing with perfect enforcement, (b) Uniform monopolist pricing with limited enforcement, and (c) Uniform monopolist pricing with perfect enforcement. The reported categories are Buyer's Net Return, Seller's Profits, Total Surplus, Unit Prices, and percentage of Excluded Buyers.