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

Felipe Brugués†
(Job Market Paper 1)

This version: October 16, 2020
Link to the latest version

Abstract

Firms in developing countries often face concentrated input markets and contracting frictions. This paper studies the efficiency of self-enforced long-term relationships between buyers and sellers, a common solution to contracting frictions, when sellers have significant market power and trade-credit contracts cannot be enforced through courts. Using new transaction-level data from the Ecuadorian manufacturing supply chain, I document trade patterns consistent with these frictions. As a relationship ages, quantities rise, and prices fall more than can be explained by quantity discounts. Based on these facts, I develop and estimate a dynamic non-linear contracting model with limited enforcement in which buyers can default on their trade-credit debt without legal penalties. In the estimated model, sellers withhold trade in early periods of a relationship, and encourage trade in later periods, in order to give buyers an incentive to pay debts. My key finding is that bilateral trade is estimated to be inefficiently low in early periods of the relationship, but converges toward efficiency as relationships age, despite sellers' market power. Counterfactual simulations imply that both seller market power and limited enforcement contribute to inefficiencies in trade, as addressing either friction alone leads to welfare losses, and that relaxing both frictions can lead to significant efficiency gains.

*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, Jack Fanning, Andrew Foster, John Friedman, Samuele Giambra, Stefan Hut, Amanda Loyola, Teddy Mekonnen, Bobby Pakzad-Hurson, Elena Pastorino, Diego Ramos-Toro, Pau Roldan-Blanco, Bryce Steinberg, Marcel Peruffo, Julia Tanndal, and Matt Turner for helpful conversations and comments. I thank Cristian Chicaiza 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.

†Department of Economics, Brown University. E-mail: felipe_brugues@brown.edu

1 Introduction

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 between-firm organizational structures. Understanding the efficiency of these informal agreements is essential for policymakers 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. Notably, however, the very same economies where enforcement constraints are likely to matter may also experience additional frictions, such as high market concentration, making them second-best environments (Rodrik, 2008). Under the presence of seller market power, weak enforcement may improve the buyer's relative bargaining power, limiting downstream distortions while increasing the efficiency of a relationship relative to a perfect enforcement world (Genicot and Ray, 2006). Therefore, 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. I apply this modeling framework to study self-enforced relationships in the manufacturing supply chain in Ecuador, a middle-income country with slow commercial courts and concentrated manufacturing sectors.

The paper has two novel empirical contributions. Through the use of a structural econometric model, I provide the first evidence on the efficiency of long-term relationships. The results show that relationships tend to be inefficient early on. However, distortions vanish to zero over time—highlighting the role of repeated informal agreements for value creation. Next, I counterfactually consider applying best-practice institutions (e.g., eliminating contracting frictions) and find that, surprisingly, they generate welfare losses relative to the second-best equilibrium. On the contrary, by addressing all the modeled frictions at once, efficiency increases.

I start by documenting six patterns that motivate the key ingredients in the model. First, the majority of trade is channeled through repeated relationships. Second, most transactions are financed by the vendor using trade-credit, even in new relationships. Third, relationships grow, both in terms of quantity and value, as they age. Fourth, sellers offer significant quantity discounts—a 10% increase in quantity is associated with a 2% unit price decrease. Fifth, conditional on the quantity purchased, clients receive additional unit discounts as their relationship evolves—older buyers receiving up to 3% discounts relative to new ones. Given profit margins of 50%, these discounts on quantities and age of relationships are economically significant. Finally, the survival probability of relationships increases in quantity and as relationships mature.

Standard models in the literature 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 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 must be large enough to secure the payment. To prevent a *take the goods and run* scenario, the seller must share a larger portion of the surplus than otherwise. Thus, enforcement constraints could act against the seller's incentive 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 share of surplus captured by the buyer increase over time.

I employ a recursive Lagrangian approach (Pavoni et al., 2018; Marcet and Marimon, 2019) that allows me to characterize the optimal dynamic contract in terms of *past* and *present* limited enforcement Lagrange multipliers (LE multipliers). Current limited enforcement constraints are captured through present LE multipliers. Moreover, promises made about future levels of consumption to prevent default in the past are captured through past LE multipliers, which serve as promise-keeping constraints. In equilibrium, the optimal quantity allocations are then determined by a *modified virtual surplus*, which accounts for standard informational rents due to incentive compatibility as well as the shadow costs of binding enforcement constraints.

The paper specifies an econometric model directly from the theoretical model. It shows that the model's parameters can be identified using cross-sectional information on the distribution of prices, quantities, age of the relationships, and a measure of the marginal cost of one seller. The dynamic identification results rely 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). In the model, the seller offers prices to induce the revelation of types and discriminate across different buyers, implying the observed price schedules at different quantities reveal information about the buyers' heterogeneous types. Allocated quantities are determined through the seller's first-order condition and are constructed so the gap between marginal prices and marginal costs respond to the seller's current and past promises needed to satisfy enforcement constraints. Hence, conditional on past constraints and a measure for marginal costs, the variation in quantity and prices is informative about the shadow value of relaxing current enforcement constraints. Therefore, with observations on the cross-sectional paths of prices and quantities, it is possible to learn the distribution of unobserved buyer heterogeneity and the extent by which current quantities are distorted due to enforcement constraints.

I estimate the model using three administrative databases collected by the Ecuadorian government for tax purposes that match the objects in the theoretical model. I obtain pair-specific unit prices and quantities using a new electronic invoice database that contains all sales for 107 (all the available) manufacturing firms in the textile, pharmaceutical, and cement-product sectors for 2016-2017, each with a large number of buyers each year (mean of 418). 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 model fits the data well, and the estimation reveals that enforcement concerns are relevant throughout the life-cycle of a relationship. Specifically, an average of 80% of new relationships has binding enforcement constraints. As relationships age, these constraints are relaxed—by year 4 of a relationship, only 20% of pairs have binding constraints—reflecting an increase in quantities coming from past promises made by the seller. Given the large number of trading partners, I explore the heterogeneity of enforcement constraints and find they differ significantly by buyers' and sellers' characteristics. For example, they are more likely to bind when the buyer is local rather than multinational or when the seller and buyer's headquarters are far away.

I use the estimated parameters to assess the transactions' efficiency at any point in time and learn about surplus division. I find that new relationships are, on average, at 68% of their first-best level. Efficiency increases over time, and those relationships lasting four years or longer are able to reach levels close to full efficiency. Furthermore, while I find that sellers capture the majority of generated surplus (around 70%), some of the surplus is directed towards the buyers over time through lower unit prices.

The paper then discusses counterfactual scenarios with counterintuitive implications. First, addressing seller market power or enforcement constraints alone, without addressing the other, leads to a lower total surplus. These results are direct manifestations of the *theory of second-best* (Lipsey and Lancaster, 1956). In the presence of one friction, the effect on welfare from eliminating one friction on its own is a priory ambiguous. In my context, each friction serves to counteract the other one. Second, by addressing both frictions at once, the results show that most relationships achieve higher total surplus and lower surplus for the seller.

This paper contributes to several strands of the theoretical and empirical literatures. First, I contribute to the theoretical and empirical literature on imperfect lending and contracting.[1] 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, with frictions that are common in other 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. Relative to the empirical literature, this is, to my knowledge, the first empirical paper to quantify the dynamic efficiency of self-enforced relationships.

This work also follows the theoretical and empirical literature related to price discrimination (Maskin and Riley, 1984; Jullien, 2000; Villas-Boas, 2004; d'Haultfoeuille and Fevrier, 2011; Grennan, 2013; Luo et al., 2018; Attanasio and Pastorino, 2020; Marshall, 2020).[2] 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 dynamic setting.

More generally, this paper relates to works in finance and development studying manifestations of the *theory of second-best* (e.g., Petersen and Rajan, 1995; Genicot and Ray, 2006; Macchiavello and Morjaria, 2020; Liu and Roth, 2020).[3] My work contributes to this strand of literature by suggesting that, empirically, fixing only one market friction may lead to welfare losses and showing that fixing both enforcement and seller market power could increase welfare. My counterfactual results also relate to Genicot and Ray (2006), which shows that, theoretically, improving enforcement reduces the buyer's expected payoff whenever the seller has the bargaining power.

Lastly, some of the empirical facts documented in Section 3 have been documented, individually, by previous works. The fact of relationship dynamics in quantities and prices has been previously documented for international trade by Heise (2019) and, partially, by

[1] Theoretical model includes 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) and empirical applications include 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 (2018); Heise (2019); Ghani and Reed (2020); Casaburi and Reed (2020).

[2] This paper is related to the literature studying the durable/storable-goods monopolist (e.g. Coase, 1972; Bulow, 1982; Dudine et al., 2006; Hendel and Nevo, 2013; Hendel et al., 2014). However, it differs from it, as this paper treats inputs as non-durable and non-storable by assuming the buyer's production opportunity is time-specific.

[3] Macchiavello and Morjaria (2020) study the effects of increased competition in the coffee supply chain in Rwanda on welfare when trading partners engage in self-enforced agreements and find *adverse* effects of competition as it reduces parties ability to sustain the agreements. Liu and Roth (2020) offers a theory in which the creation of a micro-loan sector with profit-maximizing lenders that have market power over the terms of investment creates debt traps. Furthermore, increasing levels of borrowing patience amplify the adverse effects. Similarly, Petersen and Rajan (1995) shows that increasing competition in bank lending when buyers have limited commitment to paying their debts actually hurts the buyers by decreasing overall volumes of lending.

Monarch and Schmidt-Eisenlohr (2017). The persistence of intra-national links is documented by Huneeus (2018) for Chile. Grennan (2013) and Marshall (Forthcoming) have documented price discrimination in the context of medical devices and wholesale food, respectively. Antras and Foley (2015), Garcia-Marin et al. (2019), and Amberg et al. (2020) have documented similar patterns of trade-credit issuance. To my knowledge, this paper is the first documenting relationship dynamics regarding prices and quantities intra-nationally, as well as the first documenting all of these facts in the same setting.

The remainder of the paper is organized as follows. Section 2 describes the context and presents summary statistics of the data. Section 3 offers the motivating facts that the model needs to match. Section 4 presents the model and its implications for dynamics. Section 5 discusses identification, while Section 6 describes the estimation procedure. Section 7 offers the estimated results, model fit, and discusses the performance of alternative models. Section 8 discusses welfare and three counterfactual exercises. Section 9 concludes.

2 Context, Interviews, and Data

2.1 Context: Contract Enforcement and Market Power in Ecuador

Figure C.6 provides some context related to the enforcement of contracts using rankings from the World Bank Doing Business survey. The X-axis shows the ranking under the Contract Enforcement measure, capturing the courts' efficiency in solving a quality dispute. The Y-Axis shows the ranking under an Insolvency measure, which captures the courts' efficiency in solving a default in debts due to bankruptcy. Both rankings are defined so the 1st country is the most efficient one. As shown in the figure, Ecuador is a median country under the Contract Enforcement ranking—close to average in Latin America, the Middle East and North Africa, and East Asia and Pacific. At the same time, Ecuador is one of the worst performers in the Insolvency ranking. Taken together, Ecuador appears to a below-median performer in terms of enforcement.

Figure C.7 shows the distribution of Herfindahl-Hirschman Indices (HHI) for 6-digit manufacturing sectors in 2017. The figure shows that the manufacturing sectors in Ecuador are highly concentrated. The dashed line shows that the average for all sectors is close to 0.6, while the solid line shows that the average for the sectors in my sample is close to 0.4. The US Justice Department generally considers a market to be highly concentrated if the HHI is above 0.25.[4]

2.2 Interviews

In order to gain insider knowledge of how manufacturing firms in Ecuador manage their relationships, I conducted hour-long free-form interviews with high-rank managers in 10 man-

[4]See https://www.justice.gov/atr/herfindahl-hirschman-index.

ufacturing firms in the Spring of 2019. The following points summarize the main takeaways:[5]

- Relationships do not rely primarily on written contracts but rather on informal agreements. Although transactions are formally recorded when they occur, they tend to be managed without third-party enforcement; formal enforcement is costly and inefficient.[6]
- Quality issues from suppliers were not highly relevant, as most inputs used tend to be very standardized.
- Enforcing payment of trade-credit transactions do require some investments, in terms of time and personnel, to pressure buyers to pay their debts.
- Most firms are aware that cash transactions obtain discounts (relative to trade-credit) and would like to take advantage of them, but often rely on trade-credit due to the lack of liquidity in the short-term.

This paper will not attempt to explain why these features exist but rather rely on them to understand the way they shape how on-going relationships are managed.

2.3 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.

2.3.1 *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 B2B VAT system. I use the universe of business-to-business (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, age of relationship is censored at +9.

[5] I am planning on conducting a large scale survey to obtain more systematic evidence regarding how relationships are created and managed.

[6] The Judicial Magazine of the Ecuadorian Government, available here, also provides evidence about the efficiency of the court system. I found two recent cases related to buyer default. (Case 1) Company attempts to collect 10K in debt from an invoice from March 2005. Company brings the case to court in October 2006. Final date of the case: October 2012. (Case 2) Company tries to collect 210K USD in debt from an invoice from January 2009. Company brings case to court in October 2011. Final date of the case: June 2015. In 2016, a new reform to the *Código Orgánico General de Procesos* was set in place to speed up debt collection. In theory, firms could bring cases to collect debts of up to 18K USD (in 2017) for a speedy audience. In practice, from interviews with the managers, this route was used as a last resort. The route is not exceptionally fast either. Personal estimates from 7K cases in the Civil Court in Quito, the capital, in 2017 show that it takes around 2 years to enforcement payment through the new expedited court system.

2.3.2 *Electronic Invoicing*

The primary data source for the analysis is the electronic invoicing (EI) system for 2016–2017 for 107 manufacturing firms in textiles, pharmaceutical, and cement products.

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 5000 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.[7] For each sale done by a firm in the system, the EI collects product-level information, including a bar-code identifier, product description, unit price, quantities, discounts, as well as transaction-level information, such as buyer unique national identifier and method of payment. Method of payment can be: cash, check, credit card, trade credit offered by seller with trade credit payment terms, amongst others.

I have access to 107 firms in the textiles, pharmaceutical, and cement product manufacturing sectors, which represent the largest firms in their sectors and were all the available firms in the system for their sectors. The average firm in my sample has 12% of the market share in their 6-digit sector at the national level and 29% of the market share in their sector at their province level. The database's coverage is good, with the average selling firm in my sample having more than 90% of the reported sales captured by the EI system. Interviews with managers in my sample indicate that most firms are using the invoices sent and received for internal accounting.

I classify a *product* as a bar-code identifier and description combination. I allocate any discount given in a transaction equally to all products purchased in the transaction by adjusting the product unit price by the discount. For instance, if discount offered amount to 5% of the transaction, I adjust reported unit prices of each product by 5%. Let p_{ijgry} be the discount adjusted unit price and q_{ijgty} be the reported quantity for buyer i from seller j for good g in transaction r during year y

I defined standardized unit prices at the transaction-product level $\tilde{p}_{ijgry}$ as

$$\tilde{p}_{ijgry} = ln(p_{ijgry}) - \overline{ln(p_{jgy})}, \tag{1}$$

where $\overline{ln(p_{jgy})}$ is the average log discount adjusted price for the good g of seller j in year y. I define standardized quantity at the transaction-product level $\tilde{q}_{ijgry}$ in an analogous manner.

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

[7] See Figure A.1 in the Appendix for an example of an electronic invoice observed by the government.

Then, define pair-year level equivalents for the standardized prices and quantities as:

$$\tilde{p}_{ijy} = \sum_{r \in R_{ijy}} \sum_{g \in G_{ijry}} s_{ijgry} * \tilde{p}_{ijgry}, \tag{2}$$

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.

These measures of standardized prices and quantities will be used in presenting motivating evidence of dynamics and patterns in prices and quantities. Their use indicate that product-specific differences across buyers do not drive empirical facts.

Instead, for estimation, I will use the following definitions of prices and quantities, as they better match the structure of the model. For total quantity q_{ijy}, I sum over all reported quantities over all goods and all transactions:

$$q_{ijy} = \sum_{r \in R_{ijry}} \sum_{g \in G_{ijry}} q_{ijgry}. \tag{3}$$

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

$$p_{ijy} = V_{ijy} / q_{ijy}. \tag{4}$$

This definition of prices lines up well with the weighted average of product-level discount inclusive prices, using expenditure weights, as shown in Appendix Figure B.2.

The total 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.

Appendix Section B.2 presents summary statistics about quantities, values, and the number of buyers per seller obtained through this dataset. It also reports product-level variation of standardized prices and quantities.

2.3.3 *Financial Statements*

I complement this information with yearly data on expenditures and wage bill from financial statements for all sellers for 2016-2017.[8]

I use average variable cost avc_{jy} for a firm j in year y, defined as the sum of total expenditures and wages divided by total quantity Q_{jy}, as a proxy for marginal cost.

[8]In robustness exercises, I also use sales, exports, imports, total assets, total debt, total receivables, and total uncollectibles for all buyers and sellers in the data for 2008-2017. This data is obtained from the financial statements. I also add information on 6-digit sector code, GPS location of headquarters' neighborhood, year founded, type of ownership (multinational, local, part of a business group), and whether the buyer and seller are vertically integrated. Appendix B provides detailed description of the information captured in the databases. In summary, I find that sellers are larger, older, and have more direct contact with international trade than buyers. Moreover, as shown in Appendix Figure B.5, the average accounting markups for the sellers in the sample are close to 1.5 in 2017, which are defined as total sales over total variable costs.

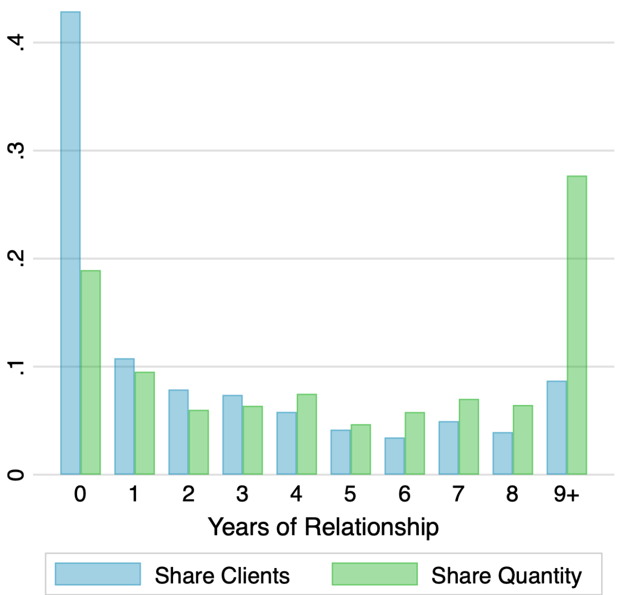
3 Motivating Evidence

In this section, I present evidence of how buyer-seller relationships work in my setup. The data highlights the three-main ideas: i) trade depends heavily on past relationships and trade credit, ii) as relationships age, quantities increase and prices decrease, and iii) at any given point in time, larger purchases are met with lower prices. In Section 4, I propose a model that captures these dynamics by using a long-term contract, in which the seller can price discriminate across buyers and time, and where buyers can default on trade credit debts with no legal repercussions.

Fact 1: Large amount of trade occurs via repeated relationships

Figure 1 shows that repeated relationships are important for the sellers in my sample. The blue bars show the average share of clients by length of relationship, whereas the green bars show the average share of total quantity sold. While around 40 percent of all buyer-seller pairs are with new buyers, only around 20 percent of trade is channeled through these new relationships. Instead, relationships that have been sustained for at least nine years represent less than 10 percent of all pairs but account for almost 30 percent of all trade.

Figure 1: **Share of Clients and Trade by Relationship Age**

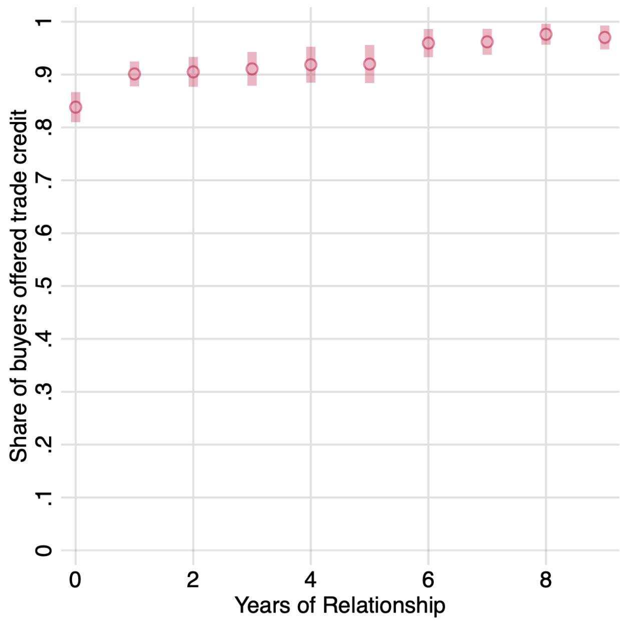


Notes: This figure presents the distribution of the average, across sellers, of the within-seller share of clients and quantity sold by age of relationship in 2016.

Fact 2: Most transactions occur via trade credit

The EI database contains information related to payment method, which specifies whether the transaction was financed by the seller and the terms of the credit in days. Here, I only consider whether any trade credit was offered to the buyer, regardless of the terms of the agreement.[9] Figure 2 plots the point estimate of the average, across sellers, of the share of relationships of a given age involving trade credit at some point during a given year. Trade credit usage is widespread, with around 85 percent of relationships receiving trade credit in their first year of contact. By age 8, almost all relationships receive trade credit during the year.

Figure 2: **Share of Relationships and Trade by Relationship Age**



Notes: The figure plots the point estimate and 90% confidence interval of the average, across sellers, of the share of relationships of a given age involving trade credit at some point during the year.

This fact has two important implications. First, the vendor is assuming a large share of the risks embedded in the transaction. In a weak legal enforcement framework, the buyer's opportunistic action would likely imply all the direct costs of such action have to be directly absorbed by the seller. Second, the seller's opportunistic actions, such as cheating in quality or quantity, are likely to be constrained (Smith, 1987). Post-delivery, the buyer can keep the value of the transaction as a guarantee of quality. For that reason, the terms of trade tilt in favor of the buyer when the seller finances transactions.

[9]Conditional on trade credit being issued, the average maturity of the agreement is 29 days.

Fact 3: Quantities increase as relationships age

I now turn to provide evidence regarding the life-cycle of quantities in relationships in Figure 3. First, in panel (a) of Figure 3, I plot a binscatter regression of standardized log quantities $\tilde{q}_{ijgry}$ on dummies for the different ages of relationships in the cross-section. The figure shows that older relationships purchase more of a given product within a given year than younger relationships.

In panel (b) of Figure 3, 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 age of relationship, controlling for pair fixed effects.[10] The figure plots the coefficients for the relationship age dummies and shows that the volume of total quantity purchased grows as relationships age.

Note, however, that I only observe at most two years per pair. For that reason, within-pair growth uses partial information to reconstruct the whole path of quantities. To verify that the partial panel of quantities is capturing correctly the growth of relationships, in Appendix Figure C.9, I plot the path of total value transacted in relationships using both the partial panel captured in the EI database as well as a longer panel using VAT data for years 2007-2015. To correctly measure the age of a relationship in the VAT data, I drop relationships that start during the first year that a seller appeared in the data. Moreover, to correct for partial-year effects in exit (Bernard et al., 2017), I drop the last observation available for each pair. The figure shows that under both databases, the value transacted within pairs increases as they age. Moreover, the EI database's partial panel accurately captures the full growth path observed in the VAT data.

It is worth highlighting that the quantity differences over time reflect a significant shift in the distribution of quantities by relationship age, rather than only an increase in dispersion. Appendix Figure C.10 shows the cumulative distribution function for standardized log quantities (product-level) and residualized log quantities (seller-level). Under both measures, the mass of quantities shifts to the right over time.

Fact 4: Quantity discounts for a given age of relationship

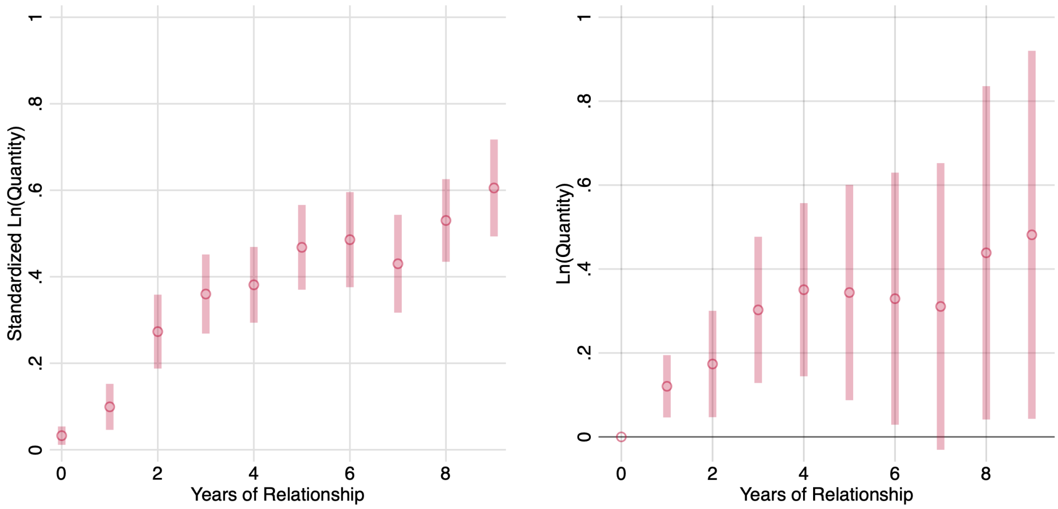
Figure 4 shows the relationship between prices and quantity purchased. Given the differences in quantities sold by different buyers, I present quantities as quantiles, calculated within each seller and the following history types: i) new relationships, ii) relationships age 1-3, iii) relationships age 4+.

Panel (a) presents a binscatter plot of the standardized unit price by quantiles of quantity. The standardization allows comparing quality-adjusted prices, as the variation is at the product-level. The figure shows that larger quantities receive lower quality-adjusted prices, regardless of the relationship's age.

[10] In Appendix Figure C.8, I report the regression of standardized log quantities within-pairs, which also shows a general increasing trend over time. However, the exact interpretation of these coefficients is more complicated, as they require across-year comparisons of within-year standardized units.

Figure 3: **Quantity and Age of Relationship**

(a) Standardized Log Quantity - Cross Section (b) Log Quantity - Within Pair



Notes: These figures plot the cross-sectional evolution of standardized log quantities (panel a) and the within-pair evolution of log total quantity (panel b), with their corresponding 90% confidence intervals. Panel (a) plots a binscatter regression of standardized log quantity on relationship age dummies. Standardized log quantity is obtained by netting out the log average quantity in a given year for each seller-product. Standard errors are clustered at the seller-year. Panel (b) 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) plots a binscatter regression of log average unit price on quantiles of quantity, controlling for seller-year fixed effects. The figure again documents the presence of quantity discounts, within relationship age.

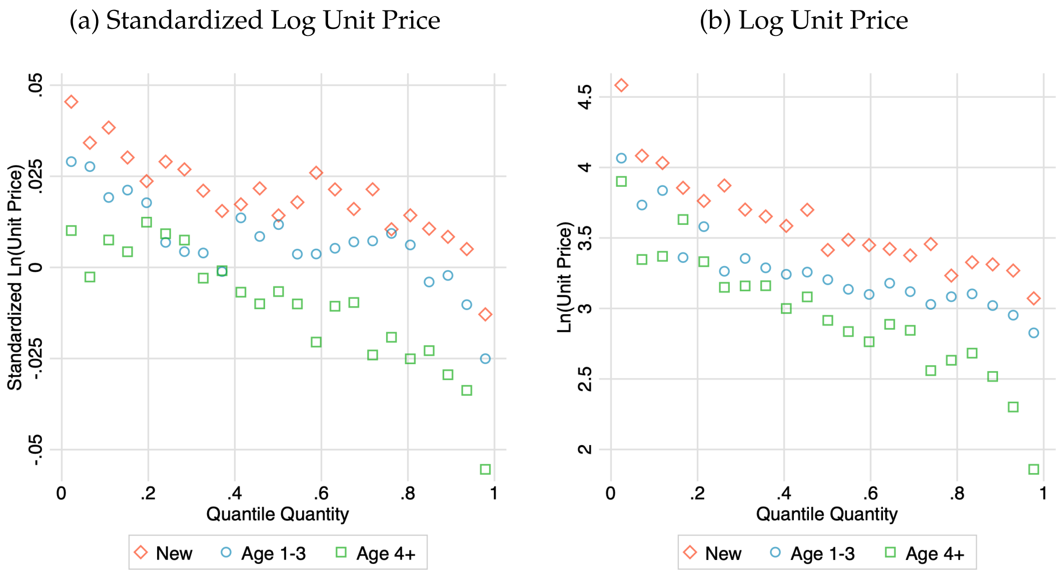
To benchmark the size of quantity discounts, a 10% increase in total quantity purchased is associated with a 2% average price decrease, as shown in Appendix Table C.4.

Fact 5: For a given quantity, older relationships pay lower unit prices

Figure 5 shows the relationship between unit prices and relationship age, for standardized log prices in panel (a) and log average prices in panel (b). Panel (a) presents a binscatter regression of standardized log prices on age of relationship dummies, controlling for a flexible spline of standardized log quantities. The figure reports that older relationships receive up to 3% quality-adjusted additional discounts as new relationships. Panel (b) shows a binscatter plot of log average prices on age of relationship, controlling for pair fixed effects. The figure shows that as relationships age, they receive around 1.5% additional discounts. That is, under both formulations, there are price discounts conceded to older clients. Given accounting profit margins of 50%, these discounts are economically significant as well.

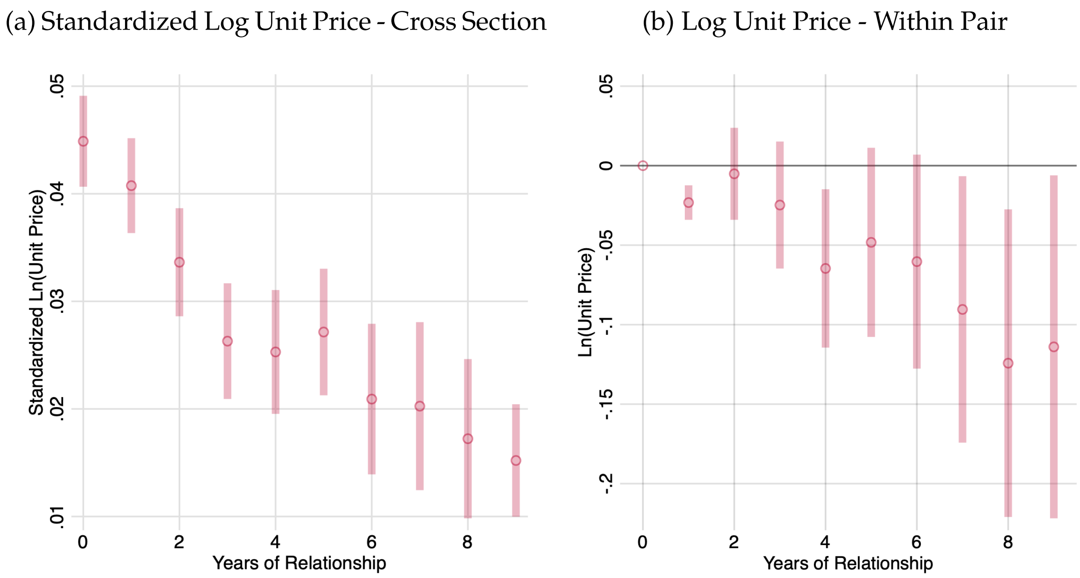
In Appendix Table C.6 I replicate Figure 5, panel (a), and find that the relationship is robust to additional controls to test for omitted variable bias. In particular, I control the buyer's age, distance in kilometers between headquarters, size of the buyer (in sales, number of em-

Figure 4: **Unit Prices by Quantile of Quantities and Age of Relationship**



Notes: These figures show the relationship between quantity purchased and standardized log unit price (left-panel) and average log unit price (right-panel) through binscatters of the measure of unit price against quantile of quantity by age of relationship. Standardized log unit prices are obtained by netting out the log average unit price in a given year for each seller-product. Quantiles of quantity are calculated for each seller-relationship age combination.

Figure 5: **Price by Relationship Age**



Notes: These figures show the relationship between age of relationship and standardized log unit price in the cross-sectional (left-panel) and average log unit price within-pair (right-panel). Panel (a) presents a binscatter of standardized log unit prices against years of relationship, controlling for a flexible spline of standardized log quantities. Standardized log unit prices are obtained by netting out the log average unit price in a given year for each seller-product. Standard errors are clustered at the seller-year level. Panel (b) 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.

ployees, assets), whether the buyer is a multinational, exporter, importer, or part of a business group. I also control for the importance of the relationship for the buyer (in terms of the supply share) or for the seller (in terms of demand share), in the spirit of Kikkawa et al. (2019), to capture possible heterogeneous markups stemming from bilateral market power. Regardless of the controls, the relationship between the age of relationship and unit prices is virtually unchanged. In Appendix Table C.3, I show that the effect is similar across the three types of industries considered.

In interpret this fact, together with the backloading of quantities, as evidence in favor of a model with limited enforcement of contracts. Such a model can accurately capture the price and quantity dynamics if the seller has a profit-maximizing incentive. By postponing the buyer's share of the surplus, the seller can discipline the buyer's behavior and maximize expected profits.

To facilitate the interpretation that the price dynamics are driven by enforcement. I study the price dynamics of multinational buyers in Appendix section D. I find that when the buyer is multinational, the dynamics of price discounts are muted. This effect is driven by multinationals with a *common law* origin, which tend to have better enforcement. I find similar patterns of prices when looking at exports from Ecuador, Peru, and Uruguay, and find that price discounts are less-steep or inexistent when the destination country has common law origin.

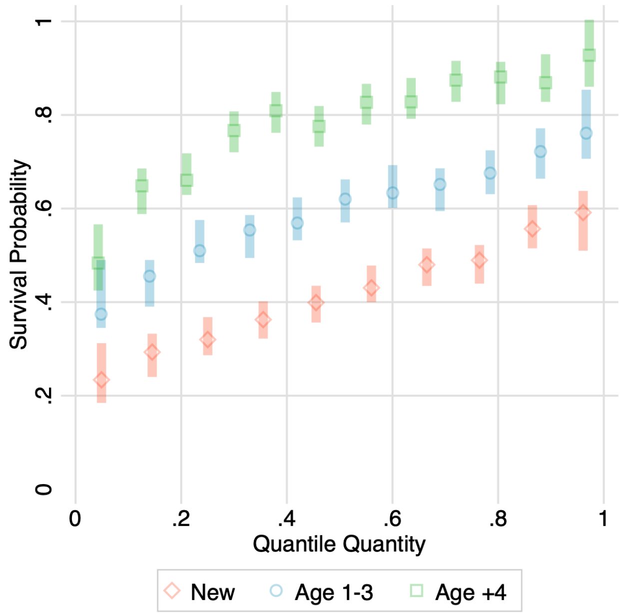
Fact 6: Relationships that trade more are more likely to survive

Lastly, relationships are persistent. Figure 6 plots the share of relationships that survive from 2016 until 2017 by quantile of quantity in 2016 and age of relationship. The figure reports the share of new links that survive in red, in blue for links age 1-3, and in green for links age 4 or older. I find that around 40 percent of new relationships survive at least one more year, 60 percent of relationships age 1-3 survive, and more than 75 percent of relationships of 4 years or more survive. Moreover, within a relationship age, pairs that trade more volume are also more likely to survive from year to year.

This last fact is important when using cross-sectional information in order to learn about panel dynamics. Given the positive selection pattern, cross-sectional variation in prices and quantities may not accurately reflect pair-specific variation, which would be accurately captured in a counterfactual world where all relationships persist. As an illustration to this issue, in Table C.5, I compare price discounts over time in the cross-section relative to pair-specific discounts. I find that the cross-sectional over-estates pair-specific price discounts by a factor of 3. However, by flexibly controlling of hazard rates by quantile of quantity (or prices) using a spline, the upward bias is eliminated.[11]

[11] I confirm in Monte-Carlo analysis not reported here that the upward bias in cross-sectional estimates is consistent with a positive selection process, and that controlling flexibly for hazard rates corrects the bias.

Figure 6: **Survival Probability by Quantile of Trade and Age**



Notes: This figure reports binscatters for the average survival rate of pairs at different ages and quantiles of quantity. Quantiles of quantities are created for each seller-age combination. Error bars are at the 90% level and reflect variation across sellers.

4 Model of a Dynamic Contract

This section introduces the dynamic model. The model has three primary purposes: 1) allow for dispersion in quantity, 2) capture quantity discounts at any point in time, and 3) obtain the backloading of prices and quantities. I accomplish the first two goals by using heterogeneous private information on the buyer's side. The model captures the backloading of prices and quantities by incorporating a limited enforcement constraint that prevents the buyer from defaulting on their trade-credit debts.

In this section, I also offer benchmark results with perfect enforcement. Moreover, I provide theoretical results to show the model can capture the desired dynamics. Finally, I also provide a theoretical discussion of the efficiency implications of limited enforcement and efficiency of contracts over time.

Preliminaries

Consider an infinitely repeated relationship between a seller (the principal) and a buyer (the agent). Time is indexed by $\tau \geq 0$ and we denote by $\delta < 1$ the common discount factor. Buyers' preferences depend on a private information match attribute (or type) θ, continuously distributed with support $[\underline{\theta}, \bar{\theta}]$, $\underline{\theta} > 0$, cumulative distribution function $F(\theta)$ and probability density function $f(\theta)$. This match attribute is drawn at the beginning of the relationship and is kept constant over time. Although the parameter is private information, the distribution $F(\cdot)$ is common knowledge.

Relationships end due to exogenous shocks that happen at every period τ with probability $X(\theta)$.[12] The exit probability $X(\cdot)$ is also common knowledge. Due to this, the type's distribution evolves over time. Define $f_\tau(\theta) = f(\theta)(1-X(\theta))^\tau / \int (f(m)(1-X(m))^\tau) dm$ as the probability density function for time τ and $F_\tau(\theta)$ as its corresponding density function.

A trade profile stipulates an infinite array of transfers t_τ and quantities q_τ for each time period τ, $\{t_\tau, q_\tau\}_{\tau=0}^{\infty}$.[13] The trade profile gives the following discounted payoff to the principal

$$\sum_{\tau=0}^{\infty} \delta^\tau (t_\tau - cq_\tau) \tag{5}$$

and to the buyer

$$\sum_{\tau=0}^{\infty} \delta(\theta)^\tau (\theta v(q_\tau) - t_\tau), \tag{6}$$

where $v(\cdot)$ is the base return function and $\delta(\theta) \equiv \delta(1 - X(\theta))$. I consider $v(\cdot)$ strictly increasing and strictly concave.[14]

4.1 Full Enforcement

As a benchmark, consider the case of full enforcement, both with symmetric and asymmetric information.

4.1.1 *Complete Information*

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

$$\theta v'(q^{1d}(\theta)) = c \quad \text{and} \quad t^{1d}(\theta) = \theta v(q^{1d}(\theta)).$$

The seller offers first-best quantities but extracts all the rents from the buyer. This allocation is infinitely repeated over time.

4.1.2 *Asymmetric Information*

The principal has commitment and wants to design a dynamic tariff scheme $t_\tau(\cdot)$ that maximizes their lifetime expected profit. The revelation principle applies to single-agent dynamic

[12] The model can accommodate for dynamic hazard rates $X_\tau(\theta)$.

[13] Throughout the next sections, I will use the terms transfers and tariffs interchangeably.

[14] This property 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).

setups (Baron and Besanko, 1984; Sugaya and Wolitzky, Forthcoming), so there is no loss of generality in restricting the study to an infinite sequence menu $\{t(\theta), q(\theta)\}_{\underline{\theta}, \bar{\theta}}$ that induces the agent to report their true type.

The theoretical insights from Baron and Besanko (1984) apply in this setup.[15] The optimal dynamic contract with full enforcement is equal to repeated Baron-Myerson static contracts with quantities determined by:

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

and tariffs such that

$$t_\tau^{fe}(\theta) = \theta v(q_\tau^{fe}(\theta)) - \int_{\underline{\theta}}^{\theta} v(q_\tau^{fe}(x))dx.$$

It is possible to show that 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 can be shown to increase.[16] Instead, without selection patterns (i.e., $X'(\theta) = 0, \forall \theta$), the optimal full enforcement contract with asymmetric information is stationary.

4.2 Limited Enforcement

While the seller can commit fully to the long-term contract, the buyer can act opportunistically. I assume that, in each period, the seller first delivers the goods and has to wait for the buyer to transfer the promised amount before the end of the period, effectively offering trade-credit to the buyer in every transaction. While this assumption is strong, it reduces the complexity of the problem and data shown in Section 3 shows trade credit is extremely common.

The direct mechanism $C(\theta) = \{q_\tau(\theta), t_\tau(\theta)\}_{\tau=0}^{\infty}$ stipulates quantities and post-delivery transfers in each period for agent reporting type θ. The seller offers the menu of $\{\theta, C(\theta)\}_{\underline{\theta}, \bar{\theta}}$, with combinations of available reporting types and corresponding allocations.

4.2.1 *Timing*

The contracting game takes places in the following order:

1. Prior to trade, at $\tau = 0$, the buyer observes their private type θ. The seller offers the mechanisms menu $\{C(\theta)\}$. The buyer either accepts or rejects the offer. If they accept, they report type $\widehat{\theta}$. If they reject, both the seller and buyer receive their outside options, normalized to 0.

2. In each trading period $\tau \geq 0$:

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

[16] 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.

- The seller produces and delivers $q_\tau(\widehat{\theta})$.
- The post-delivery payment $t_\tau(\widehat{\theta})$ is paid by the buyer, or they breach the contract.
- Following a breach on the buyer's side, the contract is terminated.

As it will made clear below, the contract considered is default-free, through the use of enforcement constraints, and features no renegotiation. Since default never occurs in equilibrium, there is no loss in assuming that the seller terminates trade following a breach (Abreu, 1988; Levin, 2003).

4.2.2 *Constraints*

Let us now characterize the set of constraints in the main problem. The set of constraints contain the usual individual rationality and incentive compatibility constraints of adverse selection problems. This setting's novelty is to include an additional enforcement constraint in each trading period, which acts as an endogenously determined participation constraint. Each of the enforcement constraints will ensure the buyer will not default in the specific time period.

Buyer's Incentive Compatibility

Under the assumption of perfectly persistent types, as in Martimort et al. (2017), incentive compatibility requires that the agent evaluates their lifetime return:

$$\sum_{\tau=0}^{\infty} \delta(\theta)^\tau u_\tau(\theta) \geq \sum_{\tau=0}^{\infty} \delta(\theta)^\tau [\theta v(q_\tau(\widehat{\theta})) - t_\tau(\widehat{\theta})] \ \ \forall \theta, \widehat{\theta}, \tag{IC-B}$$

where $u_\tau(\theta) = \theta v(q_\tau(\theta)) - t_\tau(\theta)$.

Buyer's Limited Enforcement Constraint

The key 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, a default-free menu satisfies the limited enforcement constraint of the buyer:

$$t_\tau(\theta) \leq \sum_{s=1}^{\infty} \delta(\theta)^s u_s(\theta) \ \ \forall \theta, \tau. \tag{LE-B}$$

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] For that reason, only LE-B and IC-B are consider. 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.

[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 $u(\theta) + t_\tau(\theta) = \theta v(q_\tau(\theta)) > 0$.

4.2.3 *Optimal Contract with Limited Enforcement*

Denote total surplus as $s(\theta, q) = \theta v(q) - cq$. The principal's problem becomes

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

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

The solution in a static setting (e.g., in Jullien (2000)) follows the first-order approach of Mirrlees (1971), which substitutes the global incentive compatibility constraint with a local one. Recent results in dynamic mechanism design (Pavan et al., 2014; Battaglini and Lamba, 2019) show that a dynamic envelope theorem for the relaxed problem can be used to characterize under certain conditions the global solution to the full contract. In particular, Battaglini and Lamba (2019) argue that if types are fully persistent, strictly monotonic contracts (i.e., those with $q'_\tau(\theta) > 0$ for all θ and τ) will be globally incentive compatible. Throughout this section, I will assume that allocated quantities satisfy this monotonicity property and will verify in estimation section 7 that observed allocations are consistent with this assumption.

Following Pavan et al. (2014), an implementable menu satisfies dynamic incentive-compatibility if it satisfies the dynamic envelope formula:

$$\sum_{\tau=0}^{\infty} \delta(\theta)^\tau u'_\tau(\theta) = \sum_{\tau=0}^{\infty} \delta(\theta)^\tau v(q_\tau(\theta)), \tag{7}$$

for any arbitrary $0 < \delta(\theta) < 1$ function and $u'_\tau(\theta) \equiv du_\tau(\theta)/d\theta$. Substituting the envelope condition 7 with $\delta(\theta) = \delta$ into the seller's problem SP yields

$$\sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\overline{\theta}} [s(\theta, q_\tau(\theta)) - \int_{\underline{\theta}}^{\theta} v(q_\tau(x))dx] f_\tau(\theta) d\theta - \sum_{\tau=0}^{\infty} \delta^\tau u_\tau(\underline{\theta}). \tag{8}$$

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

I follow Jullien (2000) and write the problem in Lagrangian-type form. For this formulation, the dynamic LE-B constraint at time τ is given by:

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

where $\Gamma_\tau(\theta) = \int_{\underline{\theta}}^{\theta} \gamma_\tau(x)dx$ is the cumulative LE multiplier with derivative $\gamma_\tau(\theta)$. The derivative $\gamma_\tau(\theta) > 0$ whenever the limited enforcement constraint binds and it captures the shadow value of the enforcement constraint at θ. The cumulative 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. As extending θ increases the set on which the enforcement constrained is relaxed, Γ_τ is nonnegative and nondecreasing. Moreover, by relaxing uniformly the constraints, the seller can reduce net returns by keeping quantities unchanged, so $_\tau(\bar{\theta}) = 1$.[18] Thus, the cumulative multiplier has the properties of a cumulative distribution function.

After manipulating the limited enforcement constraints,[19] one can obtain the full Lagrangian maximand:

$$\sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\bar{\theta}} [s_\tau(\theta, q_\tau(\theta)) - v(q_\tau(\theta)) \frac{\Gamma_\tau(\theta) - F_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s^\tau(\theta))\tilde{\Gamma}_s^\tau(\bar{\theta}) + \theta\gamma_\tau(\theta)}{f_\tau(\theta)}] f_\tau(\theta) d\theta, \tag{10}$$

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

$$\Gamma_s^\tau(\theta) = \frac{\int_{\underline{\theta}}^{\theta}(1 - X(x))^{\tau-s}\gamma_s(x)dx}{\tilde{\Gamma}_s^\tau(\bar{\theta})}, \tag{11}$$

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

The corresponding seller's first order condition determining the allocation rule at any relationship tenure τ is:

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

The allocation equation responds to intuitive forces. For expositional purposes, assume that the breakup probability is constant over types $X(\theta) = 0$ for all θ. Then, $\Gamma_s^\tau(\theta) = \Gamma_s(\theta)$, $\tilde{\Gamma}_s^\tau(\bar{\theta}) = 1$, $F_\tau(\theta) = F(\theta)$, $f_\tau(\theta) = f(\theta)$. Assume as well that $v(q) = kq^\beta$. The equation can be written as:

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

which resembles the usual solution to an adverse selection problem in which the allocation is determined by an inverse markup (μ) rule adjusted by the *modified virtual surplus*, which

[18] In Section F, I show formally that $\Gamma_\tau(\bar{\theta}) = 1$.

[19] 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:

$$\begin{aligned} &\sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\theta} v(q_\tau(x))dx \sum_{s=0}^{\tau-1}(1 - X(\theta))^{\tau-s} d\Gamma_s(\theta) \\ &- \sum_{\tau=0}^{\infty} \delta^\tau \int_{\underline{\theta}}^{\bar{\theta}} [\theta v(q_\tau(\theta)) - \int_{\underline{\theta}}^{\theta} v(q_\tau(x))dx] d\Gamma_\tau(\theta). \end{aligned} \tag{9}$$

Then integrate by parts.

accounts for necessary rents due to incentive compatibility and due to the limited enforcement constraint.

In particular, the incentive compatibility constraint forces the seller to give higher quantities to higher types through $F(\theta)$ as informational rents. Moreover, when the current limited enforcement constraint is binding ($\gamma_\tau(\theta) > 0$), it limits the volume of trade. To preserve incentive compatibility and prevent low-types from pretending to be higher-types, quantities are shifted upwards by $1 - \Gamma_\tau(\theta)$.

Importantly, relative to Jullien (2000), the critical distinction here is the addition of past cumulative multipliers, which generate backloading of quantities. A similar result is offered in Martimort et al. (2017) for a discrete number of types. This multiplier serves a promise-keeping constraint, where types for which their limited enforcement constraint was binding in the past, receive higher quantities in the present. With exogenous exit ($X > 0$), promises made in the distant past weigh less now. However, if relationships never end, promises made in the past shift trade levels forever, as in Marcet and Marimon (1992).

The equilibrium combination of $\Gamma_\tau(\theta)$, $\Gamma_s(\theta)$, and $\theta\gamma_\tau(\theta)$ determine whether quantity allocated is greater or lower than under full enforcement. Furthermore, as usual, allocated quantities decrease in the markup that a seller would charge under linear monopolist pricing.

The results of Pavan et al. (2014) allow us to construct the transfers $t_\tau(\theta)$ satisfying the necessary first-order conditions with the corresponding allocation rule specified in SFOC. Specifically, if the contract allocation satisfies a strict monotonicity assumption,[20] the following transfer rule satisfies the buyer's dynamic envelope formula:

$$t_\tau(\theta) = \theta v(q_\tau(\theta)) - \int_{\underline{\theta}}^{\theta} v(q(x))dx - u_\tau(\underline{\theta}), \tag{t-RULE}$$

and so the derivative of the transfer rule with respect to type is

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

4.2.4 *Non-Stationary Equilibrium*

This paper does not attempt to characterize the full optimal dynamic contract but instead conjecture about its existence and use data to learn about the primitives of the model. I prove the optimal contract cannot be stationary in Appendix E in two steps. First, I prove the existence of a unique stationary equilibrium. Second, I show that a non-stationary deviation exists that dominates the stationary equilibrium. For that reason, if an optimal contract exists, it must be non-stationary.

[20] Pavan et al. (2014) use a weaker assumption, integral monotonicity, which is implied by the strict monotonicity assumption.

4.3 Dynamics in the Limited Enforcement Model

The section offers details on how the model can rationalize the dynamics observed in Section 3. Proofs are available in Appendix Section G.

4.3.1 *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 1. *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.*

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,[21] the seller finds it optimal to offer quantity discounts at any relationship age. This is likely to be satisfied if the limited enforcement constraint is slack for some buyers already at their first interaction. Moreover, it also requires the enforcement constraint is slack for all buyers in the long run. This last requirement is in line with the model of Martimort et al. (2017), where buyers reach a *mature* phase in which the constraints no longer bind.

In terms of generality, the usual monopolist screening problem requires (or uses) log-concavity of $f(\theta)$.[22] 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 conditions on the dynamic distribution of types, in order to guarantee that the seller has the desire of price discriminating 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 $q'_\tau(\theta) > v'(q_\tau(\theta))/[-v''(q_\tau(\theta)\theta]$. As long as quantities increase by types fast enough, then 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.

4.3.2 *Evolution of Quantities*

Next, I discuss how quantities evolve in Proposition 2.

Proposition 2. *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 θ and $q_{\tau^*}(\theta) \geq q_\tau(\theta)$ for all $\tau < \tau^*$ and all 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

[21] In the data, this last condition holds for all but one seller.

[22] 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.

model predicts that quantities will reach a mature phase in which constraints no longer bind. At this mature phase, quantities will be at their highest level in the relationship.

4.3.3 *Discounts over time*

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

Proposition 3. *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 $p_{\tau+1}(q) \equiv T_{\tau+1}(q)/q < T_\tau(q)/q \equiv p_\tau(q)$.*

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.

4.4 Static Efficiency of Limited Enforcement

We now turn to analyzing the efficiency of contracts with limited enforcement. Relationship-specfic total surplus (and thus efficiency) is determined by the total quantity transacted at a point in time. I concentrate on static (period-by-period) efficiency, as it is common in the relational contracting literature (e.g., as in Fong and Li, 2017; Barron and Powell, 2019; Kostadinov and Kuvalekar, Forthcoming), rather than total lifetime efficiency.

For simplicity, suppose that $\theta\gamma_\tau(\theta)$ is small enough so the quantities allocated in the limited enforcement contract with no exit and the assumed parametrization of $v(\cdot)$ are given by:

$$q_\tau^{LEm}(\theta)^{1-\beta} = \frac{k\beta}{c}\Big[\theta - \frac{\Gamma_\tau(\theta) - F_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))}{f(\theta)}\Big].$$

Define the modified value of the cumulative multiplier at time τ as $\tilde{\Gamma}_\tau(\theta) = \Gamma_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))$, so the allocation is given by:

$$q_\tau^{LEm}(\theta)^{1-\beta} = \frac{k\beta}{c}\Big[\theta - \frac{\tilde{\Gamma}_\tau(\theta) - F(\theta)}{f(\theta)}\Big].$$

Moreover, recall that the first-best outcome is given by:

$$q_\tau^{FB}(\theta)^{1-\beta} = \frac{k\beta}{c}\theta.$$

If $\tilde{\Gamma}_\tau(\theta) < F(\theta)$, there is overconsumption relative to first best. If $\tilde{\Gamma}_\tau(\theta) > F(\theta)$, there is underconsumption. If $\tilde{\Gamma}_\tau(\theta) = F(\theta)$, trade is fully efficient. Therefore, this limited enforcement model allows for the possibility of efficient trade, as well as inefficient trade either through underconsumption or overconsumption.

For the case with underconsumption, i.e., $\tilde{\Gamma}_\tau(\theta) > F(\theta)$, efficiency increases over time if $\tilde{\Gamma}_\tau(\theta) < \tilde{\Gamma}_{\tau-1}(\theta)$. By reordering and eliminating repeated terms, the condition becomes

$\Gamma_\tau(\theta) < 1$. Thus, under the case with no exit and underconsumption, we expect efficiency to increase until pair-wise trade becomes unconstrained.

4.4.1 *Static Efficiency Relative to Perfect Enforcement*

Comparing equations SFOC and PE, in the case with no exit $X(\theta) = 0$ for all θ, the total quantity transacted is greater under full enforcement than under limited enforcement if:

$$(1 - \Gamma_\tau(\theta)) + \sum_{s=0}^{\tau-1}(1 - \Gamma_s(\theta)) - \theta\gamma_\tau(\theta) < 0. \tag{12}$$

For the types for which the limited enforcement constraint is not binding (so $\gamma_\tau(\theta) = 0$), except for the highest type, the inequality does not hold, and pair-wise welfare decreases under full enforcement. This will likely matter for middle-types early on. Moreover, it also applies for lower-types with binding constraints at the beginning for the contract, but that grow over time to become unconstrained in the long-run. Therefore, welfare can be greater under a long-term relational contract with limited enforcement than under perfect enforcement.

For types with $\gamma_\tau(\theta) > 0$, the inequality can be written as:

$$\theta - \frac{1 - \Gamma_\tau(\theta)}{\gamma_\tau(\theta)} > \frac{\sum_{s=0}^{\tau-1}(1 - \Gamma_s(\theta))}{\gamma_\tau(\theta)}.$$

The inequality above reminds us of a modified virtual surplus, where instead of the distribution of types we use the distribution of enforcement constraints. For perfect enforcement to be welfare increasing, the virtual surplus accounting for contemporaneous information rents of limited enforcement has to be greater than the information rents (promises to increase quantity) stemming from past enforcement constraints. Of course, while early on, perfect enforcement could be more efficient, as relationships age this might be more difficult to sustain.

In contrast with the arguments set forward in past literature, I have shown that in the interaction of market power and enforcement constraints could imply that weak legal enforcement is actually efficiency *increasing* at some points in time, and particularly so in the long-run. Intuitively, absent enforcement constraints, the seller is able to offer *the* profit-maximizing menu of quantities and prices. The buyer's ability to act opportunistically restricts how much the seller can extract and changes the surplus in favor of the buyer.

4.5 A Two-Type Illustrative Example

To illustrate the main forces behind the problem at hand, I discuss a two-type example. A reader may skip this section with little loss in continuity.

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 shares of surplus 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.

4.5.1 *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., $X(\theta) = 0$.

4.5.2 *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.[23]

To simplify the problem, assume that the IC-L and LE-H are slack while IC-H and LE-L are binding.[24] By using these assumptions on the constraints, one can obtain the optimal quantity

[23]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.

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

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 transfers:

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

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 such, the standard quantity allocation for θ_L (i.e., when $\delta = 1$), together with the optimal transfers 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 fixed (relative to the standard menu) and at the same time asks or lower transfers, 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.

4.5.3 *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 uses 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.

4.5.4 *A Visual Example*

To visualize the problem, I consider a numerical example with the following values for the parameters: $\beta = 0.25$, $c = 1$, $f_L = 0.95$, $\theta_L = 10$, $\theta_H = 20$, $\delta = 0.9$.

Figure 7 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, stationary with limited enforcement, and dynamic with limited enforcement.

With the solid lines, the figure shows the stationary solution both under weak enforcement and perfect 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.

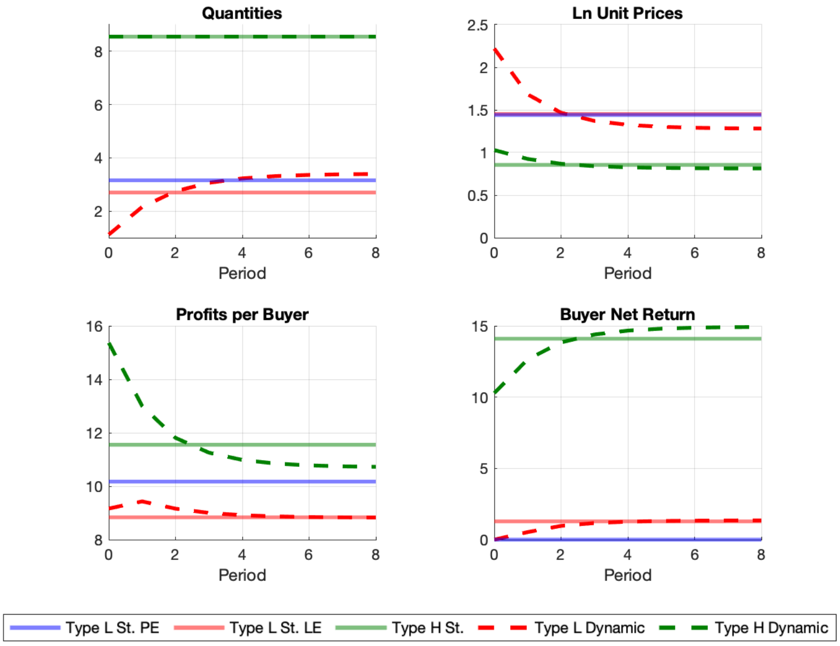
In solid red, the figure shows the allocation for type-L under limited enforcement. Relative to perfect enforcement, type-L sees a reduction in quantities and an increase in net return, in line with the logic explained above. Importantly, as the buyer's return function features diminishing returns in 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. 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 expected present-value lifetime profits from each buyer relative to the optimal stationary contract. The seller can effectively prevent default now and increase present-value lifetime profits by offering higher surplus levels to the buyers in the future.

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 and without limited enforcement. That is, through dynamic contracts, long-term allocations could potentially be more efficient than contracts under perfect enforcement.

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 7: **Example - Nonlinear Pricing and Limited Enforcement**



Notes: This figure shows Quantities, Prices, Profits, and Buyer Net Return for different enforcement and contract regimes. In solid green, the optimal stationary contract for type-H under perfect enforcement and limited enforcement. In dashed green, the optimal dynamic contract for type-H under limited enforcement. In solid blue, the optimal stationary contract for type-L under perfect enforcement. In solid red, 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: $\{\beta = 0.25, c = 1, f_L = 0.95, \theta_L = 10, \theta_H = 20, \delta = 0.9\}$.

5 Identification of Dynamic Contracts

This section discusses identification of the model primitives θ, $v(\cdot)$ and $\Gamma_\tau(\cdot)$. 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. Throughout this section, I abstract away from the possibility of exogenous breakups. The possibility of breakups will be reintroduced in estimation.

As shown in Section 4, the dynamic contract is a complex object. Rather than deriving the full equilibrium contract by forward-iteration, I rely on the following assumption for 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.*

Under this assumption, I can collapse all information about future unobserved quantities and transfers into the limited enforcement multipliers. Although the assumption is strong, it is often used in the identification of dynamic games, as these types of games may have multiple equilibria (Aguirregabiria and Nevo, 2013).

For identification, I exploit the fact that the mapping from agent type θ to quantity q_τ is strictly monotone and write the first-order condition of the seller SFOC and the derivative of the transfer rule of the buyer t-RULE in terms of quantiles (Luo et al., 2018; Luo, 2018):

$$\theta_\tau(\alpha)v'(q_\tau(\alpha)) - c =$$
$$\Big[\Gamma_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha)) + \frac{\theta_\tau(\alpha)}{\theta'_\tau(\alpha)}\gamma_\tau(\alpha)\Big]\theta_\tau(\alpha)v'(q_\tau(\alpha))\frac{\theta'_\tau(\alpha)}{\theta_\tau(\alpha)},$$
$$T'_\tau(q_\tau(\alpha)) = \theta_\tau(\alpha)v'(q_\tau(\alpha)),$$

where $\alpha \in [0,1]$ and I used the fact that observed price schedule can be mapped to the model tariff schedule by $T_\tau(q_\tau(\theta(\alpha))) = t_\tau(\theta(\alpha))$. Moreover, $\theta_\tau(\alpha)$ and $q_\tau(\alpha)$ are the α-quantiles of the agent's type and quantity at tenure τ, respectively. Notice as well that I have used $f_\tau(\theta(\alpha)) = 1/\theta'_\tau(\alpha)$ and $\gamma_\tau(\theta_\tau(\alpha)) = \gamma_\tau(\alpha)/\theta'_\tau(\alpha)$.

Together, the key identification equation becomes:

$$\frac{T'_\tau(q_\tau(\alpha)) - c}{T'_\tau(q_\tau(\alpha))} = \frac{\theta'_\tau(\alpha)}{\theta_\tau(\alpha)}\Big[\Gamma_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))\Big] + \gamma_\tau(\alpha), \tag{13}$$

where $\theta_\tau(\cdot)$, $\theta'_\tau(\cdot)$, $\Gamma_\tau(\cdot)$, and $\gamma_\tau(\cdot)$ are unknown. The price schedule $T_\tau(\cdot)$ and its derivatives are nonparametrically identified from information on prices and quantities alone, so in this section, I treat them as known. Moreover, I treat c as known.

Identification of the Limited Enforcement Multiplier $\Gamma_\tau(\theta)$

The identification argument is recursive and takes the the primitives at time $s < \tau$, and in particular, $\Gamma_s(\alpha)$, as known.

Define $\Xi_\tau(\alpha) = \Gamma_\tau(\alpha) + \theta_\tau(\alpha)/\theta'_\tau(\alpha)\gamma_\tau(\alpha)$. Substituting in and reordering, equation 13 becomes:

$$\Xi_\tau(\alpha) = \alpha + \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha)) + \frac{T'_\tau(q_\tau(\alpha)) - c}{T'_\tau(q_\tau(\alpha))}\frac{\theta_\tau(\alpha)}{\theta'_\tau(\alpha)}. \tag{14}$$

As $\theta_\tau(\alpha) > 0$ and $\theta'_\tau(\alpha) > 0$, $\Xi_\tau(\alpha)$ is set identified. In particular,

$$\Xi_\tau(\alpha) \in \begin{cases} [0, \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))) \text{ if } T'_\tau(q_\tau(\alpha)) < c, \\ [\alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha)), 1] \text{ if } T'_\tau(q_\tau(\alpha)) \geq c, \end{cases}$$

where $\Gamma_s(\alpha)$ is assumed to be known.

For every value $\Xi_\tau(\alpha)$, there is a unique value for $\theta_\tau(\alpha)/\theta'_\tau(\alpha)$. Therefore, for each combination of $\{\Xi_\tau(\alpha)\}_{\alpha\in[0,1]}$, $\Gamma_\tau(\alpha)$ is identified from the solution of the differential equation:

$$\gamma_\tau(\alpha) + \Gamma_\tau(\alpha)\frac{\theta_\tau(\alpha)}{\theta'_\tau(\alpha)} = \Xi_\tau(\alpha)\frac{\theta_\tau(\alpha)}{\theta'_\tau(\alpha)}, \tag{15}$$

after defining a boundary condition for $\Gamma_\tau(\alpha)$, which is shown in the Appendix F to be $\Gamma_\tau(1) = 1$. For that reason, $\Gamma_\tau(\cdot)$ is set identified.[25]

However, by making the parametric assumption on the return function $v(q) = kq^\beta$ for $k > 0$ and $\beta \in (0,1)$, the multipliers $\Gamma_\tau(\cdot)$ are point identified.[26] Appendix I provides the details. Generally speaking, $\Gamma_\tau(\cdot)$ is point identified up to a function $A(q) = -v''(q)/v'(q)$. By parametrizing $v(q) = kq^\beta$, the function $A(q) = (1-\beta)/q$ depends only on one parameter, which is identified from observations of prices, quantities, and marginal cost for the lowest and highest consumption buyers. Hence, through the assumed parametrization, $\Gamma_\tau(\cdot)$ is point identified from observations of prices, quantities, and marginal cost.

For estimation, I follow the approach of Attanasio and Pastorino (2020), which consider a parametrization of $\Gamma_\tau(\cdot)$ as a flexible function of q_τ rather than parametrizing $v(\cdot)$. Throughout the remainder of the section, I consider $\Gamma_\tau(\cdot)$ as point identified.

Identification of Types θ

Using the allocation equation at τ and the fact that $\delta \ln(\theta_\tau(\alpha))/\delta\alpha = \theta'_\tau(\alpha)/\theta_\tau(\alpha)$, I obtain the following expression for $\theta_\tau(\alpha)$:

$$\ln(\theta_\tau(\alpha)) = \ln(\underline{\theta}_\tau) + \int_\tau^\alpha \frac{\delta \ln(\theta_\tau(x))}{\delta x} dx \tag{16}$$

$$= \ln(\underline{\theta}_\tau) + \int_0^\alpha \frac{1}{\Gamma_\tau(x) - x - \sum_{s=0}^{\tau-1}(1-\Gamma_s(x))} \Big[\frac{T'_\tau(q_\tau(x)) - c}{T'_\tau(q_\tau(x))} - \gamma_\tau(x)\Big] dx, \tag{17}$$

which identifies the quantile function of type $\theta_\tau(\cdot)$ up to $\underline{\theta}_\tau$ and $\Gamma_\tau(\alpha)$. Making the scale normalization on types $\underline{\theta}_\tau \equiv \underline{\theta} = 1$, the quantile function for types becomes:

$$\theta_\tau(\alpha) = \exp\Big(\int_0^\alpha \frac{1}{\Gamma_\tau(x) - x - \sum_{s=0}^{\tau-1}(1-\Gamma_s(x))} \Big[\frac{T'_\tau(q_\tau(x)) - c}{T'_\tau(q_\tau(x))} - \gamma_\tau(x)\Big] dx\Big). \tag{18}$$

As Luo et al. (2018) show, $\underline{\theta} = 1$ is a normalization for a general function $v(\cdot)$. Under a parametrization $v(q) = kq^\beta$, which provides point identification for $\Gamma_\tau(\cdot)$, $\underline{\theta} = 1$ is also a normalization as it suffices to multiply k by the normalization constant to obtain an observa-

[25] If $\Gamma_s(\alpha)$ is taken to be a set, then the identification set for $\Xi_\tau(\alpha)$ should be defined as:

$$\Xi_\tau(\alpha) \in \begin{cases} [0, \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s^{SUP}(\alpha))) \text{ if } T'_\tau(q_\tau(\alpha)) < c, \\ [\alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s^{INF}(\alpha)), 1] \text{ if } T'_\tau(q_\tau(\alpha)) \geq c, \end{cases}$$

where $\Gamma_s^{SUP}(\alpha)$ is the supremum and $\Gamma_s^{INF}(\alpha)$ is the infimum in identified set for $\Gamma_s(\alpha)$. Although the bounds for $\Xi_\tau(\alpha)$ are wider, the identification argument for $\Gamma_\tau(\cdot)$ remains unchanged. For every value $\Xi_\tau(\alpha)$ and $\sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))$, there is a unique value for $\theta_\tau(\alpha)/\theta'_\tau(\alpha)$. Therefore, for each combination of $\{\Xi_\tau(\alpha), \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))\}_{\alpha\in[0,1]}\}$, $\Gamma_\tau(\theta)$ is identified from the solution of the differential equation 15.

[26] Luo (2018) studies the nonparametric identification of this model and of $\Gamma_\tau(\alpha)$ in particular with observations on prices and quantities alone. They find that this model is nonparametrically identified if one can find an alternative efficient market, for which $\Gamma_\tau(\alpha) = 1$ for all α, in order to learn about $\theta'_\tau(\alpha)/\theta_\tau(\alpha)$. With information on $\theta'_\tau(\alpha)/\theta_\tau(\alpha)$ in hand, $\Gamma_\tau(\alpha)$ is nonparametrically identified from information on tariffs and prices alone. This approach is not feasible in my setting as each seller is considered a market, and it is impossible to find something that could be regarded as an alternative efficient market for each seller.

tionally equivalent structure.

The distribution $f_\tau(\theta)$ is identified from $\theta'_\tau(\alpha)$ since $f_\tau(\theta) = 1/\theta'_\tau(\alpha)$ and $\theta'_\tau(\alpha)$ is obtained from

$$\theta'_\tau(\alpha) = \frac{\theta_\tau(\alpha)}{\Gamma_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\alpha))}\Big[\frac{T'_\tau(q_\tau(\alpha)) - c}{T'_\tau(q_\tau(\alpha)} - \gamma_\tau(\alpha)\Big]. \tag{19}$$

Identification of the Base Return Function $v(\cdot)$

The base return function $v(q_\tau(\alpha))$ is identified in two steps under the assumed parametrization. First, the elasticity β is identified from observations of quantities, prices, and marginal costs, as detailed in Appendix I. The level shifter k is identified using the derivative of the transfer rule $T'_\tau(q_\tau(\alpha)) = \theta_\tau(\alpha)v'(q_\tau(\alpha)) = \theta_\tau(\alpha)k\beta q_\tau(\alpha)^{\beta-1}$, as $\theta_\tau(\alpha)$ and β are identified, while $q_\tau(\alpha)$ and $T'_\tau(q_\tau(\alpha))$ are known.

6 Estimation

This section discusses the details of the estimation procedure. The main estimation steps are based on identification equations 13 and 18. I estimate the equations for each seller j and tenure τ using $N_{j\tau}$ observations. As I perform estimation for each seller j separately, I drop subscript j from now onward.

The general structure of the estimation procedure is as follows. First, I perform three intermediate steps: 1) I estimate the tariff functions using observations on payments and quantities for each tenure separately, 2) I use all information on sales and variable costs to estimate constant marginal costs, and 3) using pair-wise information I estimate heterogeneous hazard rates at the percentile-tenure level and obtain percentile-to-percentile transition matrices over time. Second, I perform the following steps iteratively starting at $\tau = 0$: 1) using the empirical analog of equation 13, I estimate enforcement multipliers, 2) using the transition matrix, I link estimated multipliers for $s < \tau$ to quantiles of quantity at τ. Third, using equation 18 and estimated objects, I obtain estimated types $\widehat{\theta}$ for each τ.

6.1 Tariff Function

In 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 4 implies that observed transfers lie on the curve $t = T(q)$, as they are both functions of the type $\theta_{i\tau}$ in a given tenure. As noted by Luo et al. (2018), observed prices and quantities may not lie on the curve, if there is measurement error or unobserved heterogeneity, introducing additional randomness beyond $\theta_{i\tau}$.

To deal with this additional randomness, I follow Perrigne and Vuong (2011), which 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 $t_{i\tau} = T(q_{i\tau})e^{v_{i\tau}}$, such that $v_{i\tau}$ is independent of $q_{i\tau}$.

I consider a parametric version of the model, in which $T_\tau(q) = e^{\beta_0\tau} q^{\beta_{1\tau}}$. This leads to the estimation model with measurement error:

$$ln(t_{i\tau}) = \beta_{0\tau} + \beta_{1\tau} ln(q_{i\tau}) + v_{i\tau}, \tag{20}$$

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 $\widehat{T}_\tau(q_{i\tau}) = e^{\hat{\beta}_{0\tau}} q_{i\tau}^{\hat{\beta}_{1\tau}}$, while the marginal tariff is $\widehat{T}'_\tau(q_{i\tau}) = \hat{\beta}_{1\tau} t_{i\tau}/q_{i\tau}$. 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 have been sustained for some years.

6.2 Heterogenous Hazard Rates

I estimate heterogenous hazard rates 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 the noise and preserve a monotonicity of hazard rate, I then approximate the estimated nonparametric hazard rates as a logistic function of percentiles:

$$S\tau(r) = \frac{exp(a_\tau + b_\tau r)}{1 + exp(a_\tau + b_\tau r)} + \varepsilon^s_\tau(r), \tag{21}$$

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.

6.3 Marginal Cost

Marginal cost is estimated directly from the data under the assumption that marginal cost is equal to average variable cost. As defined in Section 2, average variable cost is defined as total expenditures and total wages divided by total quantity sold.

6.4 LE Multipliers

Recall that the LE multiplier $\Gamma_\tau(\alpha)$ has the properties of a cumulative distribution function. Following Attanasio and Pastorino (2020), I parametrize the multiplier as a logistic distribution:[27]

$$\Gamma_\tau(\alpha) = \frac{exp(\phi_\tau(q_\tau(\alpha)))}{1 + exp(\phi_\tau(q_\tau(\alpha)))}, \tag{22}$$

[27] The multiplier function is the solution to a differential equation. As shown in Appendix H, 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).

where $\phi_\tau(q_\tau(\alpha))$ is a polynomial up to the second degree. Under this parametrization, the derivative of the multiplier is $\gamma_\tau(\alpha) = \phi'_\tau(q_\tau(\alpha))\Gamma_\tau(\alpha)(1-\Gamma_\tau(\alpha))$.

Moreover, I parametrize $\theta'(\alpha)/\theta(\alpha)$ as a inverse quadratic function of quantity:

$$\frac{\theta'(\alpha)}{\theta(\alpha)} = \frac{1}{d_0 + d_1 q_\tau(\alpha) + d_2 q_\tau(\alpha)^2}. \tag{23}$$

The key identification equation 13 provides the following estimating equation:

$$\begin{aligned}
&\frac{\hat{\beta}_{1\tau} p_\tau(\alpha) - \widehat{c}}{\hat{\beta}_{1\tau} p_\tau(\alpha)} = \qquad\qquad\qquad\qquad \text{(Main Est. Eq.)}\\
&\frac{1}{d_0 + d_1 q_\tau(\alpha) + d_2 q_\tau(\alpha)^2}\Big[\frac{exp(\phi_\tau(q_\tau(\alpha)))}{1+exp(\phi_\tau(q_\tau(\alpha)))} - \alpha - \sum_{s=0}^{\tau-1}(1-\widehat{\Gamma}_s(\alpha))\Big]\\
&+ \phi'_\tau(q_\tau(\alpha))\frac{exp(\phi_\tau(q_\tau(\alpha)))}{1+exp(\phi_\tau(q_\tau(\alpha)))}\Big(1 - \frac{exp(\phi_\tau(q_\tau(\alpha)))}{1+exp(\phi_\tau(q_\tau(\alpha)))}\Big) + \varepsilon^g_\tau(\alpha),
\end{aligned}$$

where I have used $p_{i\tau} = t_{i\tau}/q_{i\tau}$ and where ε^g is measurement error coming from the misspecification of Γ, the tariff function, or the marginal cost. $\widehat{\Gamma}_s(\alpha)$ for $s < \tau$ is 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 $\{\phi_\tau, d_0, d_1, d_2, \sigma^{\varepsilon^g}\}$.

To match previously estimated LE multipliers $\Gamma_s(\theta)$ to $\theta(\alpha)$ at tenure τ, I use the estimated hazard rates to generate a percentile-percentile transition matrix. Then, I can match percentiles matching α_s for $s < \tau$ to percentiles matching α_τ. Moreover, I use the estimated hazard rates for τ corresponding to α to properly discount past promises captured in past multipliers.

6.5 Buyer Types and Type Distribution

Once Γ_τ and γ_τ are estimated, the consumer type $\theta_\tau(\alpha)$ is obtained from

$$ln(\widehat{\theta}_\tau(\alpha)) = \tag{24}$$

$$\frac{1}{N_\tau}\sum_{k=1}^{N_\tau}\frac{1\{\alpha \geq k/N_\tau\}}{\widehat{\Gamma}_\tau(k/N_\tau) - k/N_\tau - \sum_{s=0}^{\tau-1}(1-\widehat{\Gamma}_s(k/N_\tau))}\Big[1 - \frac{\widehat{c}}{\hat{\beta}_{1\tau} p_\tau(k/N_\tau)} - \widehat{\gamma}_\tau(k/N_\tau)\Big], \tag{25}$$

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

$$\widehat{\theta'}_\tau(\alpha) = \frac{\widehat{\theta}_\tau(\alpha)}{\widehat{\Gamma}_\tau(\alpha) - \alpha - \sum_{s=0}^{\tau-1}(1-\widehat{\Gamma}_s(\alpha))}\Big[1 - \frac{\widehat{c}}{\hat{\beta}_{1\tau} p_\tau(\alpha)} - \widehat{\gamma}_\tau(\alpha)\Big]. \tag{26}$$

Finally, the density function $\widehat{f}_\tau(\theta(\alpha))$ is $1/\widehat{\theta'}_\tau(\alpha)$.

6.6 Base Marginal Return and Return Function

The derivative of the transfer 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

$$v'(\widehat{q_\tau(\alpha)}) = \frac{\hat{\beta}_{1\tau} p_\tau(\alpha)}{\widehat{\theta}_\tau(\alpha)}. \tag{27}$$

Following the discussion in the identification section, $v(\cdot)$ is estimated by

$$v(q_\tau(\alpha)) = \widehat{T}_\tau(q_\tau(0)) + \frac{1}{N_\tau} \sum_{k=1}^{N_\tau} v'(q_\tau\widehat{(k/N_\tau)}) 1\{\alpha \geq k/N_\tau\}. \tag{28}$$

In Appendix Section M, I show in Monte Carlo simulation that the estimation method can accurately recover the primitives $\{\Gamma_\tau, v(\cdot), \theta\}$ for a two-period dynamic contract.

6.7 Parametrization of $v(\cdot)$ for Counterfactual Analysis

To calculate pair-specific efficient (first-best) quantities, I require estimated buyer types θ, base marginal returns $v'(\cdot)$ and seller marginal costs c. The range of optimal quantities may not be covered by the range of realized quantities, and thus, base marginal returns may be undefined for some quantities. For that reason, during counterfactual analysis, I parametrize the seller-specific marginal return functions $v(\cdot)$ as $v(q) = kq^\beta$, for $k > 0$ and $\beta \in (0,1)$ and estimate these functions for each seller using linear least squares and the values of estimated marginal returns $\widehat{v'(\cdot)}$.

7 Empirical Results

In this section, I first explain the definition of relationship tenure, then discuss the estimates of primitives of the model, and show the data fit. I present the results pooling all sellers together but conduct estimation at the seller level.

7.1 Definition of Relationship Tenure

For now, I do not have enough observations at each seller-relationship age to estimate the dynamic model.[28] For this reason, I pool the following relationship ages together and define a *relationship tenure* between seller i, buyer j and year t as:

$$tenure_{ijt} = \begin{cases} 0 & \text{if pair-age = 0} \\ 1 & \text{if } 1 \leq \text{pair-age} \leq 3 \\ 2 & \text{if pair-age} > 3. \end{cases} \tag{29}$$

[28] [During COVID, I only had access to the small sample offered by the government to develop my codes at home.]

I restrict estimation to sellers with at least 50 observations of tenure 0, 25 of tenure 1, and 25 of tenure 2. This leaves me with 11 sellers with information for 2016 and 2017 as well as 12 sellers with information for either 2016 or 2017.

7.2 Estimation Results

My model relies on the following seller-dependent ingredients: initial distribution of private types θ, the base return function $v(\cdot)$, and the limited enforcement multipliers $\Gamma_\tau(\cdot)$ for tenure $\tau \in \{0, 1, 2\}$.

Figure 8 shows the average estimated log type θ by quantile of quantity for tenure 0. Recall that for identification, I normalized the lowest type $\underline{\theta}$ to 1. The figure shows that types monotonically increase with quantity purchased on average, with a larger increase in the level of types for the top quantiles of quantities. This figure is consistent with the monotonicity assumption $q'_\tau(\theta) > 0$ required for dynamic incentive-compatibility and identification. Monotonicity is not by construction, as it possible for equation 26 to be negative in various ranges when either marginal price $T'_\tau(q_\tau(\alpha))$ is lower than the marginal cost c or when the buyer is overconsuming relative to first-best.

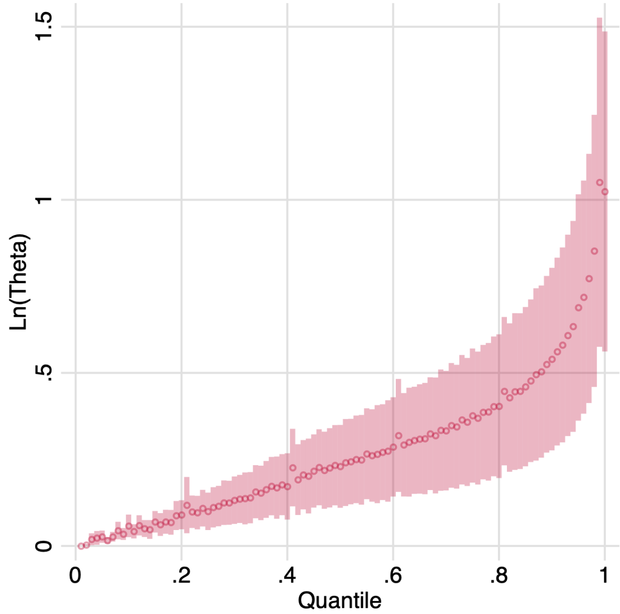
Figure 9 plots the average estimated base marginal return $v'(\cdot)$ by quantity quantile and relationship tenure. Consistent with the model, the base marginal return function $v'(\cdot)$ decreases as quantity increases for all. Moreover, the figure shows that older tenures experience a downward shift in their functions $v'(\cdot)$ for a large number of quantiles, reflecting the higher levels of quantity consumption as time increases.

The estimated results have an intuitive economic interpretation, as $v'(\cdot)$ captures, for a given type, the marginal revenue for the buyer of an additional unit of the good. For the median quantity, an additional unit of the good generates 1.67 dollars of revenue for the buyer for each dollar spent on manufacturing the good by the seller (see Appendix Figure K.20).

Lastly, Figure 10 plots the average estimated limited enforcement multiplier $\Gamma_\tau(\cdot)$ in panel (a) and the difference in multipliers within θ over time, $\Delta\Gamma_\tau(\theta) \equiv \Gamma_\tau(\theta) - \Gamma_{\tau-1}(\theta)$, in panel (b). Panel (a) shows that, on average, across-sellers, 80 percent of new pairs are constrained, as the average multiplier $\Gamma_0(\cdot)$ is only equal to 1 for the top 20 percent of pairs. As relationships age, the estimated results imply that the limited enforcement constraint is less restrictive. Over time, the average multiplier approaches 1 at lower quantiles of trade. In fact, by tenure 2, the estimates indicate that only 40 percent of transactions face distortions related to the limited enforcement constraint. Panel (b) indicates that most of the action occurs in the earlier years of a relationship.

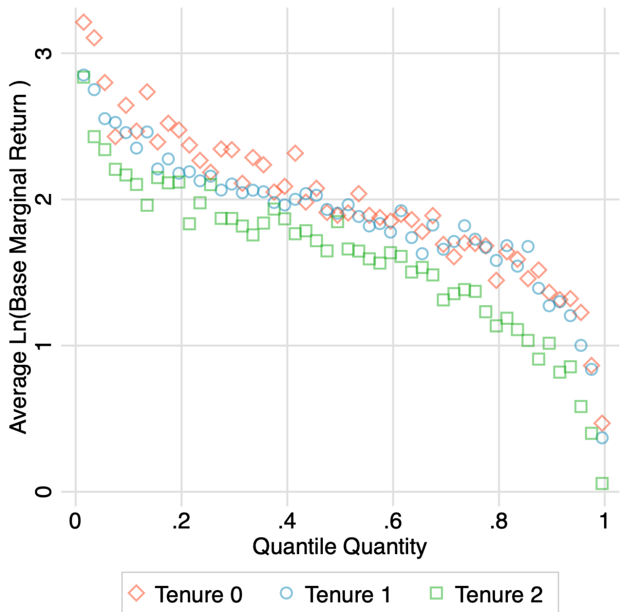
The multiplier $\Gamma_\tau(\theta)$ captures the shadow value of relaxing the enforcement constraint uniformly from $\underline{\theta}$ to θ. The evolution of the multipliers over tenures indicate that the benefits of relaxing the enforcement constraints move towards lower types. For instance, a uniform reduction of $1/\delta$ dollars in the tariffs for all buyers in tenure 1 generates on 80 cents per buyer in tenure 0, while the same reduction in tenure 3 generates only 40 cents per buyer in tenure 2. However, by reducing tariffs by $1/\delta$ dollars in tenure 1 for the lowest 20% of buyers, the

Figure 8: **Average Type by Quantile**



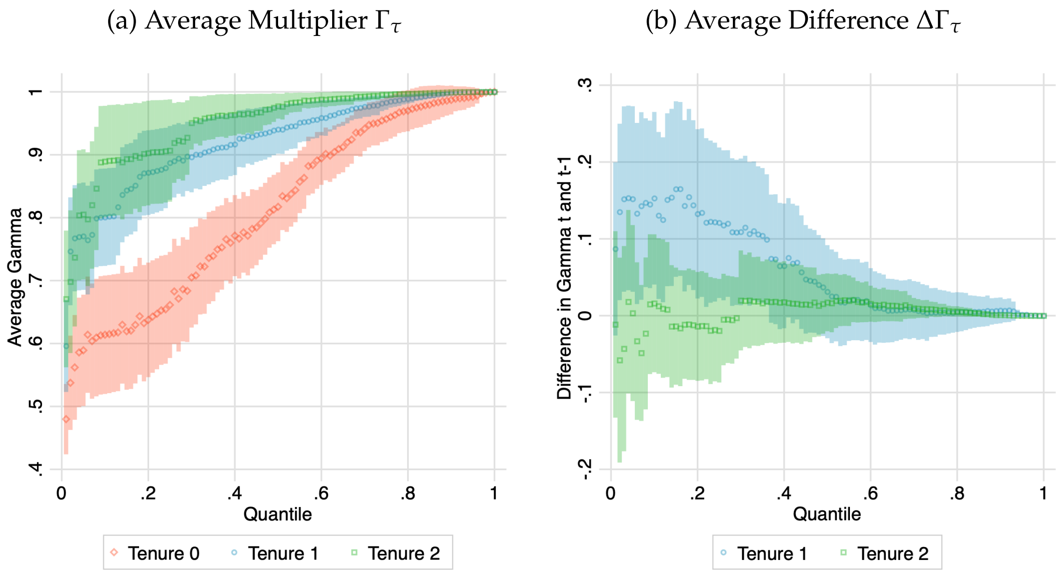
Notes: This figure plots the average log type $ln(\theta)$ by quantile of quantity, across-sellers, with error bars corresponding to $\pm$ 1.96 standard errors.

Figure 9: **Average Base Marginal Return by Quantile**



Notes: This figure plots the average base marginal returns, across-sellers, quantile of quantity for the different estimation tenure groups.

Figure 10: **Limited Enforcement Multiplier**



Notes: These figures show the average estimated limited enforcement multiplier by tenure and quantile of quantity across sellers (panel a) and the average difference in multipliers across type within a type θ by tenure (panel b). The error bars reflect $\pm$ 1.96 standard errors for each quantile across sellers.

seller captures an average of 65 cents per buyer in tenure 0, while doing the same in tenure 3 generates 90 cents per buyer in tenure 2.

Appendix K.1 shows the distribution of t-statistics for the LE multiplier at tenure 0 (Γ_0) for a test against a standard model null hypothesis. Based on the significance of the parameters of estimated $\Gamma_0(\theta)$, I reject the null that the standard nonlinear pricing model applies in my setup.

Appendix K.11 reports the estimated values for k and β of the parametrization of $v(\cdot)$, which will be used to obtain quantities in counterfactual simulations.

7.3 Model Fit

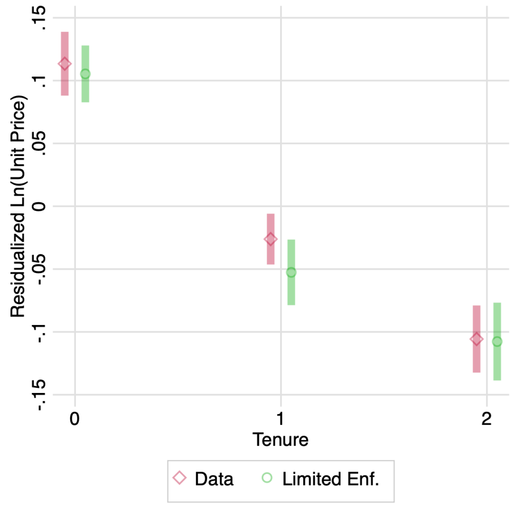
7.3.1 *Cross-Sectional Fit*

Next, I consider four different measures of cross-sectional model fit. First, Appendix Figure K.21 shows the model has good statistical fit across tenures.

Second, I compare observed quantities with model predicted quantities. Quantities are constructed using the closed form solution of the seller's first-order condition under the parametrization of $v(\cdot)$. Figure Appendix K.22 plots observed quantities on the X-axis and model predicted quantities on the Y-axis. Predicted quantities match well observed ones in all tenures.

Third, using predicted quantities and the incentive-compatible tariff function *t*-RULE, I generate predicted tariffs. Appendix Figure K.23 plots observed tariffs on the X-axis and model generated tariffs on the Y-axis for all different tenures. Again, the model has a good performance fitting the observed tariffs.

Figure 11: **Non-Targeted Cross-Sectional Moment: Price Discounts over Time**



Notes: This figure presents a binscatter of unit prices by tenure over time, both of prices in the data and model generated prices. Model-generated unit-prices are obtained by dividing model-generated tariffs by model-generated quantities. To correct for price level differences across sellers, as well as data relative to model, I remove seller-year fixed effects for data and model separately. Error bars capture $\pm$ 1.96 standard errors across pairs (buyer-seller).

Fourth, in Figure 11 I compare the non-targeted observed cross-sectional unit price discounts by tenure to those generated by the model, which the model replicates quite well.

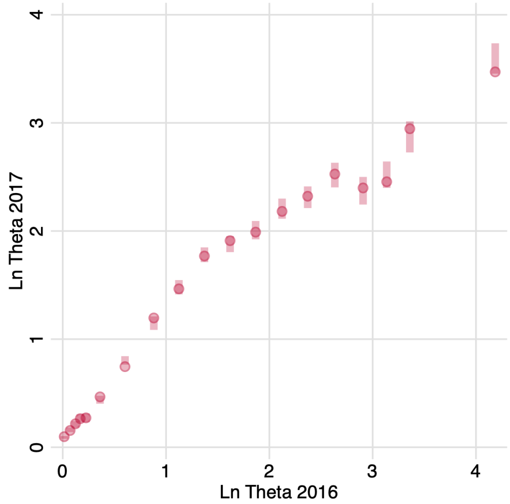
7.3.2 *Panel Fit*

One may worry that the model may fail to capture within-pair dynamics, despite performing well on cross-sectional measures. For that reason, I consider the following validation exercises.

First, given that my model is estimated using cross-sectional information for each seller separately in 2016 and 2017, I can use the panel structure to verify that the primitives of the model are similar over time within pairs. Figure 12 shows the value of estimated $\widehat{\theta}$ in 2017 against the value of estimated $\widehat{\theta}$ in 2016 for pairs that are active on both years. The figure shows a good correspondence between both estimated values, with the markers overlaying the diagonal in the graph. In fact, the estimated value of the slope using the corresponding regression framework is 0.99 and hypothesis testing fails to reject that it is different from 1 at any relevant value.

Furthermore, the model considers within-pair and within-type dynamics equivalent. For that reason, one may compare the within-pair and within-type evolution of prices, quantities, or elements that depend on both (such as profits) as well on additional estimated primitives (such as types or return functions). Table 1, panel a, shows the within-pair evolution using the panel structure of the data and measuring quantities and prices directly in the data. Over

Figure 12: **Estimated Type in 2017 against Estimated Type in 2016**



Notes: This figure shows estimated types θ in 2017 against those estimated in 2016 for buyer-seller pairs that appear on both years, which were obtained through separate seller-specific estimations for each year using cross-sectional variation alone. The error-bars represent 90% confidence intervals for average value in the bin at the buyer-level.

time, quantity increases by 6.2%, raising pair-specific surplus by 12.6%. Most of the gains are captured by buyers, who see an increase in net returns of 16.9%, while total profits for sellers increase only 4%. Panel b shows the evolution within estimated type θ for the equivalent model generated variables using the cross-sectional structure. Dynamics are remarkably similar. The cross-sectional model replicates well the panel dynamics of prices, quantities, and surplus. The model underestimates the extent to which surplus is shifted towards buyers over time relative to the data, as it underestimates the buyer's return growth rate and overestimates seller profits growth rate.

Table 1: **Panel and Cross-Sectional Fit**

	(1)	(2)	(3)	(4)	(5)
	Ln(Unit Price)	Ln(Quantity)	Ln(Surplus)	Ln(Return)	Ln(Profits)
Panel A: Quantities and prices observed in the data					
Tenure	-0.0221	0.0620**	0.126***	0.169***	0.0434
	(0.0162)	(0.0279)	(0.0197)	(0.0258)	(0.0339)
Panel A: Quantities and prices generated by model					
Tenure	-0.0295***	0.104***	0.125***	0.0786***	0.116***
	(0.00743)	(0.0187)	(0.0167)	(0.0193)	(0.0179)

Notes: This table reports the evolution of different contract measures as relationships age. Surplus and returns are generated using estimated values for types and base returns. Panel A uses observed quantities and prices together with estimated parameters. Clustered standard errors at the pair-level are reported in parenthesis. Panel B reports the equivalent regressions using instead model generated prices and quantities. Standard errors clustered at the type θ are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

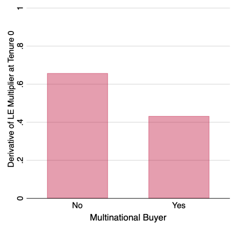
7.4 Qualitative Results

To further explore the estimated model's implications, I consider heterogeneity in limited enforcement multipliers by seller and buyer characteristics.

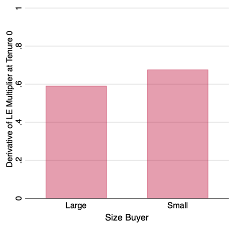
Recall that $\gamma_0(\cdot) > 0$ implies that the buyer's limited enforcement constraint is binding. Figure 13 shows the probability that the constraint is binding at tenure 0 by differ buyer characteristics, offering qualitative differences consistent with previous literature on enforcement constraints. Consistent with the literature finding that multinational buyers are more reliable and thus less likely to see their enforcement constraint bind ex-ante (e.g. Alfaro-Urena et al., 2019), multinational buyers are less likely to have a binding enforcement constraint. Similarly, larger firms, exporters, importers, or firms in business groups are less likely to have a binding constraint. Consistent with a property-rights approach (e.g. Grossman and Hart, 1986), vertically integrated firms are less likely to have a binding constraint. Moreover, buyers that might find it hard to locate an alternative supplier, for instance, those that depend heavily on the seller as measured by their supply share, also are less likely to have a binding constraint (e.g., as in McMillan and Woodruff, 1999). Lastly, distant buyers, which plausibly impose higher enforcement costs for the seller, are more likely to see their enforcement constraint bind (e.g., as in Antras and Foley, 2015).

Figure 14 plots the coefficients of regressions of the share of constrained buyers at tenure 0 on different sellers' characteristics. Larger sellers measured either by total sales, total assets or size of cash holdings correlate with a lower share of constrained buyers. Instead, sellers with higher levels of leverage (measured as debt over assets), higher maturities of trade credit (measured by the ratio of receivables over sales) or higher reported defaulted (measured by the ratio of uncollectibles over sales) correlate with a greater share of constrained buyers, though the coefficients are noisy and imprecisely estimated.

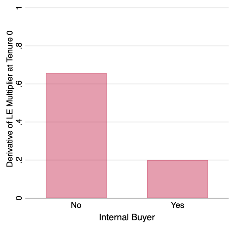
Figure 13: **Enforcement Constraints and Buyer Characteristics**



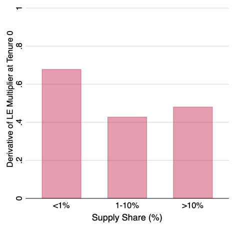
(a) Multinational



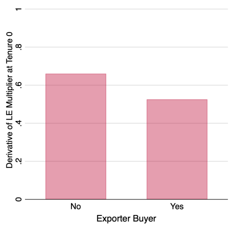
(b) Size



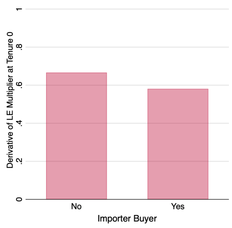
(c) Vertically Integrated



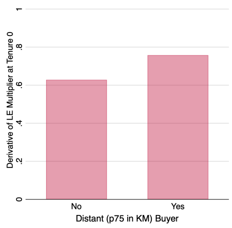
(d) Supply Share



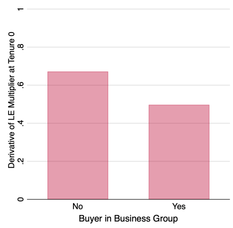
(e) Exporter



(f) Importer



(g) Distance



(h) Business Group

Notes: These figures present heterogeneity of estimated limited enforcement multipliers by buyer's characteristics. The figures shows the share of buyers in each group with positive enforcement constraint $\gamma_0(\cdot)$ in tenure 0. I take classification for multinational from the firm registry in the Servicio de Rentas Internas. I classify a firm as large if they are in the top 25 of sales from the set of buyers. I classify a pair as vertically integrated if they have any common owner with at least 1% of shares in each firm. Supply share is defined dividing total expenditure on seller by total expenditures in all intermediate inputs. Buyer is classified as exporter if they report at least $5,000 USD of exports and importer if they report at least $5,000 USD of imports. Distance between headquarters is calculated as kilometers between neighborhoods as the crow flies, for the neighborhoods appearing the firm registry in the Servicio de Rentas Internas. I classify a buyer as part of a business group if they have at least link with another firm in the economy through an shareholder that owns at least 1% shares in each firm.

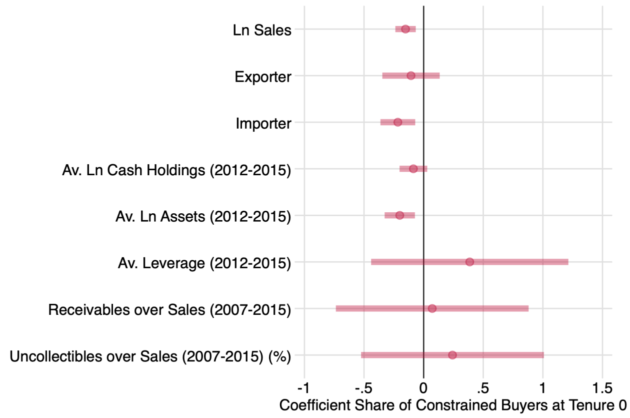
7.5 Performance of Alternative Models

While no theory is likely the only explanation behind an empirical phenomenon, I detail how relevant alternative models fail to match the observed dynamics in this section.

7.5.1 *Seller's Opportunistic Behavior*

I have abstracted away from the possibility of seller's opportunistic behavior. Martimort et al. (2017) developed a theory of contracts with limited enforcement, where the buyer can default on debts and the seller can act opportunistically by cheating on quality. Their model also features backloading of quantities, which increase over time, but predict increasing unit prices. The evidence regarding prices dynamics, therefore, rejects this extension of the model.[29]

Figure 14: **Correlation Seller Characteristics and Share of Constrained Buyers**



Notes: This figure plots the estimated coefficients of a regression of the share of constrained buyers for each seller-year on different seller's characteristics. Sales refer to total sales. I classify a seller as an exporter if they report exports of at least $5,000 USD and as an importer if they report imports of at least $5,000 USD. Cash holdings, receivables, uncollectibles, and total assets are obtained through the financial statements. Leverage is estimated as total debt over total assets.

7.5.2 *Customer Base*

Another strand of literature (Gourio and Rudanko, 2014; Roldan and Gilbukh, 2018; Fitzgerald et al., 2019; Piveteau, 2019) has emphasized the role of customer accumulation incentives on the dynamics of prices and quantities. These models are able to capture increases in quantities as relationships age but also predict increasing unit prices. The evidence, again, would reject this type of extension of the model.

[29] Also, interviews with managers in the field suggested that supplier misbehavior was not a main issue of concern, but rather opportunistic behavior on the buyer's side. Furthermore, the trade credit literature (e.g. Smith, 1987; Breza and Liberman, 2017) argue that trade credit itself serves a mechanism to guarantee product quality.

7.5.3 *Pair-specific Productivity Improvements*

Productivity improvements could drive the growth in quantities and decline in prices (Heise, 2019). If bilateral trade becomes more efficient as relationships age, pair-specific marginal costs decrease, leading to lower prices. As discussed in Appendix Section D, multinational buyers see lower discounts over time. Given that the empirical evidence highlights the *productivity* effects of becoming the supplier of a multinational (Alfaro-Urena et al., 2019), it seems unlikely that the price dynamics explained solely by productivity improvements would fail to find those improvements when trading with multinationals.

7.5.4 *Learning about Reliability*

Some works in international trade (Macchiavello and Morjaria, 2015; Antras and Foley, 2015; Araujo et al., 2016; Monarch and Schmidt-Eisenlohr, 2016) have highlighted the possible role of learning about an agent's reliability in explaining price and quantity dynamics. Suppose the seller learns about the buyer's reliability over repeated interactions, then the optimal price decreases because the risk of default decreases. In Appendix Section J, I formally consider the model used in the previous literature and show it is not able to reconcile the patterns in the data. The associated price discounts imply default rates at least 20 times larger than those observed in the data. Moreover, an estimated model is not able to recover the observed price discounts in relationships age, neither by allowing default rates to be free nor calibrating them to observed rates. Lastly, the limited enforcement model offers better statistical fit.

7.5.5 *Price Dynamics in Estimated Alternative Models*

In Appendix L, I consider the empirical fit of alternative models in terms of the price dynamics in the cross-section. The models I consider are: standard model, learning about reliability, and the estimated limited enforcement model where I shut down the promises captured through past LE multipliers. As mentioned above, the limited enforcement model fits the data well. However, the standard model cannot match the extent of price discounts, whereas the learning model predicts weakly increasing prices. Interestingly, shutting down the promises in the estimated limited enforcement model generates increasing prices.

8 Welfare and Counterfactuals

In Section 4, I provided theoretical results linking the increase in quantities with a decrease in the binding constraints. Similarly, I showed through an example, that if trade were inefficient early on, efficiency would continue to increase until trade became unconstrained. Now, in this section, I use the estimated model to perform dynamic efficiency analysis of the relationships and show that over time efficiency indeed increases.

Furthermore, I consider the efficiency of alternative pricing and enforcement schemes. In particular, I consider three margins: i) perfect enforcement + full price discrimination, ii) lim-

Table 2: **Model Efficiency: Surplus relative to pair-specific first-best**

	Tenure		
	0	1	2
As % of First-Best	68.32	87.70	98.38
(s.e.)	(1.47)	(3.04)	(4.15)

Notes: This table reports unweighted average efficiency measures (relative to pair-specific first-best) of model generated quantities for tenure 0, 1, and 2. Standard errors clustered at the type θ level are reported in parenthesis. Unit of observation is type-tenure-seller-year.

ited enforcement + uniform pricing, and iii) perfect enforcement + uniform pricing.

8.1 Efficiency Relative to First-Best

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

$$q^{fb}(\theta) = \left(\frac{k\beta\theta}{c}\right)^{1/(1-\beta)}. \tag{30}$$

Moreover, total surplus is a function of buyer's type θ, quantity q and seller's marginal cost c: $Surplus(\theta, q, c) = \theta v(q) - cq$.

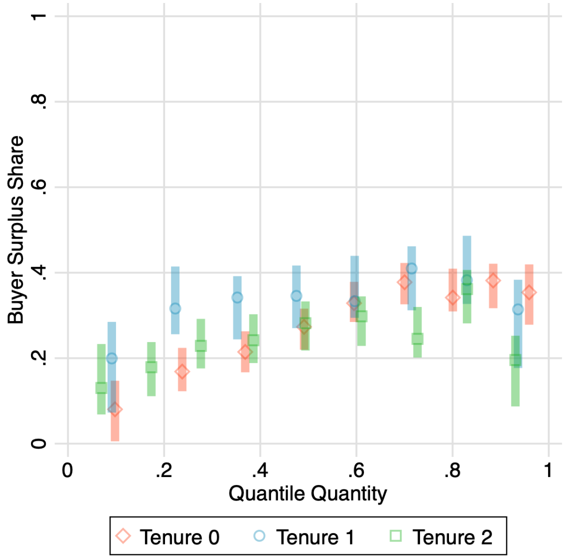
Table 2 shows the *average* evolution of pair-specific surplus relative to first-best over time. On average, new relationships start trading at 68 percent of their efficient value. As expected, efficiency increases over time, and relationships that survive until tenure 2 tend to reach efficiency, as the surplus generated by the relationship is 98 percent that of first-best and cannot be statistically rejected different from optimal surplus.[30]

I next consider surplus division in Figure 15. The figure shows the average share of surplus captured by the buyer, across sellers, by bins over quantiles of quantity purchased at different tenures. The figure shows that the median buyer tends to capture around 30% of the surplus generated in any point in time. Moreover, the buyer's share of surplus generally increases in the amount of quantity they purchase, from around 15% for the lowest quantiles to up to 40% for the highest. Lastly, price discounts over time imply surplus shares evolve in favor of the buyer. In Table 3, I report within-buyer and type-specific evolution of buyer surplus. Using quantities and prices observed in the data, I find as statistically significant shift in buyer surplus. Using simulated prices and quantities, I find a weakly positive, albeit statistically insignificant, growth in buyer surplus.

As a last exercise to understand the welfare implications of relationships, I compare allocations for each θ at tenure 1 and tenure 2 that would be generated by the model if the past promises are forgotten relative to their actual values at those tenures. To do so, I set $1 - \Gamma_s$ to zero for $s < \tau$. Table 4 shows the results. Over all buyers, eliminating past promises reduces

[30]These values are similar when restricting to types that survive all periods. Similarly, Table K.12 presents the equivalent exercise aggregating at the seller-year level and finds similar results.

Figure 15: **Buyer Share of Surplus by Quantile and Tenure**



Notes: This figure presents binscatters of the average buyer surplus share by quantiles of quantity and tenure, across types and sellers. Error bars reflect ±1.96 standard errors.

surplus generated by 15% in tenure 1 and 14% in tenure 2. Aggregating over buyers for each seller, I find that eliminating past memory decreases the total surplus generated by the seller by 6% in tenure 1 and 8% in tenure 2. The difference between buyer-specific and seller-specific losses reflects that low and medium-types are those benefiting over time from the repeated interactions. In dollar terms, the difference in total surplus change from shutting down memory about past promises represents around USD 200,000 per seller per year under surplus prices at tenure 0. Interestingly, the seller's discounts are cost-effective in terms of welfare for society, as, on average, the dollar gain in surplus over profit loss is 2.

Table 3: **Evolution of Buyer Surplus**

	Buyer Surplus Share
Panel A: P and Q from Data	
Tenure	0.011***
	(0.0036)
Panel B: P and Q from Model	
Tenure	0.004
	(0.2253)

Notes: This table reports regressions for the average change in buyer surplus share by tenure. Panel A presents results using prices and quantities as observed in the data while relying on estimated types and base return functions to estimate surplus. Regression controls for pair fixed effects. Standard errors are clustered at the pair-level. Panel B presents results using only model generated objects. Regression controls for type fixed effects. Standard errors clustered at the type-level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: **Value of Relationships - Shutting memory off**

	Pair-level	Seller-level
Tenure 1	85%	94%
Tenure 2	86%	92%

Notes: This table reports the surplus losses (in percentage) from eliminating past promises captured in past limited enforcement multipliers $(1-\Gamma_s)$ for $s<\tau$ at tenure 1 and 2. The first column reports the average surplus loss across all buyers, whereas the second column reports mean of seller-level measures of losses, which were calculated as weighted average (weighted using quantity) of the losses from all buyers in the respective tenure.

8.2 Counterfactuals

Next, I use the estimated model to assess the implications of i) improving enforcement of the trade credit contracts and ii) enforcing current legislation forbidding price discrimination on otherwise identical transactions (e.g., transactions purchasing the same quantity of the same product in a specific time period). I do so in three counterfactuals: 1) maintain price discrimination but eliminate limited enforcement, 2) maintain limited enforcement but eliminate price discrimination, and 3) eliminate both limited enforcement and price discrimination.

Counterfactual 1: *Perfect Enforcement + Price Discrimination*

A natural question that arises is what would the generated surplus and corresponding surplus shares be in a world of perfect enforcement of contracts. To obtain quantities under perfect enforcement, I use the distribution of types at tenures τ and equation Q-CES, setting $\Gamma_\tau(\cdot)$ to 1 and $\gamma_\tau(\cdot)$ to 0, as well as $\Gamma_s(\cdot)$ to 1.

Counterfactual 2: *Limited Enforcement + Uniform Pricing*

Written law in Ecuador, the European Union, and the US forbid price discrimination that applies differential treatment to customers performing an otherwise equivalent transaction, including possibly preferential treatment due to tenure.[31] This counterfactual studies the welfare effects of a policy that enforces uniform pricing but keeps the limited enforcement regime active.

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 $q^l(\theta) = (\alpha\beta\theta/p^l)^{1/(1-\beta)}$. This stationary menu will be insufficient for some enforcement constraints. Given exogenous hazard rates $X(\theta)$, the stationary enforcement constraint will be given by:

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

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.

Under a monotonicity assumption on $X(\theta)$,[32] the seller will set a minimum quantity $\underline{q}^l$ that the buyer needs to announce in order to be served. In particular, it will only serve $q(\theta) \geq \underline{q}^l$,

[31] 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.

[32] The monotonicity on the hazard rate $X'(\theta) < 0$ is observed in the data.

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 $q^l(\theta) < \underline{q}^l$.[33]

Counterfactual 3: *Perfect Enforcement + Uniform Pricing*

Lastly, I consider optimal uniform pricing under perfect enforcement. I use quantities and prices as in counterfactual 2 above. However, as buyers are precluded from the possibility of default, the seller serves all buyers. Thus, no quantity is set to zero.

8.2.1 *Discussion of Counterfactual Results*

Table 5 presents the results.[34] The table shows the average surplus in the counterfactual scenario as percentage of baseline for each percentile group in quantity and tenure.

Panel A shows the results for counterfactual 1 (nonlinear pricing with perfect enforcement). Across time and types, the surplus is between 40 to 60 percent of the baseline surplus. In the presence of market power, the limited enforcement of contracts actually helps discipline the seller, increases overall efficiency, and shifts the terms of trade in favor of the buyer.

Panel B presents the results for counterfactual 2 (uniform pricing with limited enforcement). Across time and types, the surplus is between 0 to 20 percent of the baseline surplus. The surprisingly low performance of this alternative regime is explained by the large share of buyers that would be excluded from trade. 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.

Panel C shows the results for counterfactual 3 (uniform pricing with perfect enforcement). The results show that surplus increases relative to baseline, except for the highest types. Welfare gains are concentrated in the lowest types (who see gains of up to 500%), although even median types also see large increases (of around 15 to 30%).

The counterintuitive results that solving only one friction at once may lead to welfare losses is a direct manifestation of the *theory of second best* (Lipsey and Lancaster, 1956). In the presence of multiple market frictions, eliminating one friction will not necessarily lead to higher welfare. In fact, in the presence of one market friction, an additional friction might be necessary to reach second-best.

[33] In this counterfactual exercise, I use an additional assumption: buyer's demand the optimal level of quantity that is consistent with prices and full enforcement.

[34] Section K.6 presents additional results for the three counterfactual exercises related to buyer net return, profits, and prices.

Table 5: **Average Surplus as % of Baseline (Nonlinear Price + Limited Enforcement)**

	10%	25%	50%	75%	100%
Panel A: Nonlinear Price + Perfect Enforcement					
Tenure 0	64.0	51.2	44.2	50.5	49.4
Tenure 1	42.2	64.7	62.7	60.3	49.9
Tenure 2	16.9	60.7	71.3	61.3	46.4
Panel B: Uniform Price + Limited Enforcement					
Tenure 0	0.7	14.0	18.5	21.6	18.4
Tenure 1	0.7	13.0	18.9	20.6	18.4
Tenure 2	2.1	13.2	19.7	20.4	20.4
	% Excluded				
Tenure 0	99.1	79.9	67.7	57.0	57.0
Tenure 1	98.2	75.5	63.5	57.0	56.3
Tenure 2	95.5	75.5	62.4	57.0	52.3
Panel C: Uniform Price + Perfect Enforcement					
Tenure 0	617.5	359.2	138.7	134.6	52.0
Tenure 1	259.1	206.6	99.0	133.0	52.1
Tenure 2	458.5	204.0	114.7	183.7	44.9

Notes: This table presents average efficiency measures as % of baseline (nonlinear price with limited enforcement) of different pricing and enforcement regimes by percentile groups of quantity and tenure. For instance, 10% collects all buyers between percentiles 0 and 10%. Panel A reports results for nonlinear pricing with perfect enforcement. Panel B reports optimal monopolistic uniform price with limited enforcement. Subpanel reports 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 Panel A and C.

9 Conclusion

This paper studies how frictions distort and shape long-term relationships in the manufacturing supply chain. Through the lens of a novel theoretical model, the paper shows that allocating the bargaining power to the seller but allowing the buyer to *take the goods and run* has exciting implications for surplus division as well as price and quantity dynamics. In particular, I show that by introducing a limited enforcement constraint to prevent the buyer from defaulting on their debts, the seller must share a larger portion of the surplus than required in a perfect enforcement world. This generates incentives for the seller to distort trade intertemporally by promising larger quantities and lower prices to the seller in the future to reap larger shares of profit now.

Using a unique intra-national database from Ecuador, I document new empirical patterns consistent with the model but hard to reconcile with existing bilateral trade models. To quantify the extent by which trade is affected by enforcement constraints, I estimate the model structurally using information on prices, quantities, and age of buyer-supplier relationships. The estimates reveal that trade is significantly constrained by enforcement concerns early on, but that distortions vanish over time.

The estimated model allows me to perform various relevant counterfactual exercises. First, the estimated coefficients are sufficient to characterize first-best trade levels. As such, they allow me to perform efficiency analysis for relationships at any point in time. On average, I find that relationships that are four years or older reach trade levels close to full efficiency.

Second, the model also allows me to explore different pricing and enforcement regimes. Interestingly, simulated results show that fixing the enforcement problem alone would generate *greater* distortions, as the seller behaves as a full monopolist. Moreover, limiting seller market power by forcing linear pricing alone would not necessarily lead to welfare gains. Under limited enforcement of contracts, linear pricing would force the seller to exclude a majority of buyers who cannot credibly commit to paying their debts. Instead, addressing seller market power and enforcement simultaneously could lead to beneficial gains in terms of surplus and lower prices.

References

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

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," *Available at SSRN 3376129*, 2019.

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.

Araujo, Luis, Giordano Mion, and Emanuel Ornelas, "Institutions and export dynamics," *Journal of International Economics*, 2016, *98*, 2–20.

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

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.

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

Barron, Daniel and Michael Powell, "Policies in relational contracts," *American Economic Journal: Microeconomics*, 2019, *11* (2), 228–49.

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

Bernard, Andrew B, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes, "The origins of firm heterogeneity: A production network approach," Technical Report, National Bureau of Economic Research 2019.

— , **Esther Ann Boler, Renzo Massari, Jose-Daniel Reyes, and Daria Taglioni**, "Exporter dynamics and partial-year effects," *American Economic Review*, 2017, *107* (10), 3211–28.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen, "Americans do IT better: US multinationals and the productivity miracle," *American Economic Review*, 2012, *102* (1), 167–201.

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," Technical Report 4 2020.

Breza, Emily and Andres Liberman, "Financial contracting and organizational form: Evidence from the regulation of trade credit," *Journal of Finance*, 2017, *72* (1), 291–324.

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

Bulow, Jeremy I, "Durable-goods monopolists," *Journal of Political Economy*, 1982, *90* (2), 314–332.

Casaburi, Lorenzo and Tristan Reed, "Interlinked Transactions and Competition: Experimental Evidence from Cocoa Markets," 2020.

Coase, Ronald H, "Durability and monopoly," *The Journal of Law and Economics*, 1972, *15* (1), 143–149.

d'Haultfoeuille, Xavier and Philippe Fevrier, "The provision of wage incentives: A structural estimation using contracts variation," Technical Report 2011.

Dudine, Paolo, Igal Hendel, and Alessandro Lizzeri, "Storable good monopoly: The role of commitment," *American Economic Review*, 2006, *96* (5), 1706–1719.

fai Fong, Yuk and Jin Li, "Relational contracts, limited liability, and employment dynamics," *Journal of Economic Theory*, 2017, *169*, 270–293.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi, "How exporters grow," Technical Report, National Bureau of Economic Research 2019.

Garcia-Marin, Alvaro, Santiago Justel, and Tim Schmidt-Eisenlohr, "Trade credit and markups," 2019.

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.

Gourio, Francois and Leena Rudanko, "Customer capital," *Review of Economic Studies*, 2014, *81* (3), 1102–1136.

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

Grossman, Sanford J and Oliver D Hart, "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of Political Economy*, 1986, *94* (4), 691–719.

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

Heise, Sebastian, "Firm-to-firm relationships and the pass-through of shocks: Theory and evidence," *FRB of New York Staff Report*, 2019, (896).

Hendel, Igal, Alessandro Lizzeri, and Nikita Roketskiy, "Nonlinear pricing of storable goods," *American Economic Journal: Microeconomics*, 2014, *6* (3), 1–34.

_ **and Aviv Nevo**, "Intertemporal price discrimination in storable goods markets," *American Economic Review*, 2013, *103* (7), 2722–51.

Hjort, Jonas, Xuan Li, and Heather Sarsons, "Across-Country Wage Compression in Multinationals," Technical Report, National Bureau of Economic Research 2020.

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

Johnson, Simon, John McMillan, and Christopher Woodruff, "Courts and relational contracts," *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.

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

Kikkawa, Ayumu Ken, Glenn Magerman, and Emmanuel Dhyne, "Imperfect competition in firm-to-firm trade," *Available at SSRN 3389836*, 2019.

Kostadinov, Rumen and Aditya Kuvalekar, "Learning in Relational Contracts," *American Economic Journal: Microeconomics*, Forthcoming.

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.

Liu, Ernest and Benjamin Roth, "Contractual Restrictions and Debt Traps," *Harvard Business School Entrepreneurial Management Working Paper*, 2020, (20-088).

Luo, Yao, "Identification of participation constraints in contracts," *Economics Letters*, 2018, *173*, 84–87.

_ , 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," 2020.

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

Marcet, Albert and Ramon Marimon, "Communication, commitment and growth," *Journal of Economic Theory*, 1992, *58* (2), 219–249.

_ and _ , "Recursive contracts," *Econometrica*, 2019, *87* (5), 1589–1631.

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

_ , "Search and wholesale price discrimination," Technical Report, RAND Journal of Economics Forthcoming.

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.

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

Mirrlees, James A, "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Monarch, Ryan and Tim Schmidt-Eisenlohr, "Learning and the value of relationships in international trade," *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-11*, 2016.

_ and _ , "Learning and the value of trade relationships," *FRB International Finance Discussion Paper*, 2017, (1218).

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.

_ and _ , "Trade credit: theories and evidence," *Review of Financial Studies*, 1997, *10* (3), 661–691.

Piveteau, Paul, "An empirical dynamic model of trade with consumer accumulation," Technical Report, Working Paper 2019.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny, "The quality of government," *Journal of Law, Economics, and Organization*, 1999, *15* (1), 222–279.

_ , _ , _ , 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.

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

Roldan, Pau and Sonia Gilbukh, "Firm dynamics and pricing under customer capital accumulation," 2018.

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.

Startz, Meredith, "The value of face-to-face: Search and contracting problems in Nigerian trade," 2018.

Sugaya, Takuo and Alexander Wolitzky, "The Revelation Principle in Multistage Games," *Review of Economic Studies*, Forthcoming.

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

Urena, Alonso Alfaro, Mariany Fuentes Fuentes, Isabela Manelici, and José P Vásquez, "Costa Rican Production Network: Stylized Facts," 2018.

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.

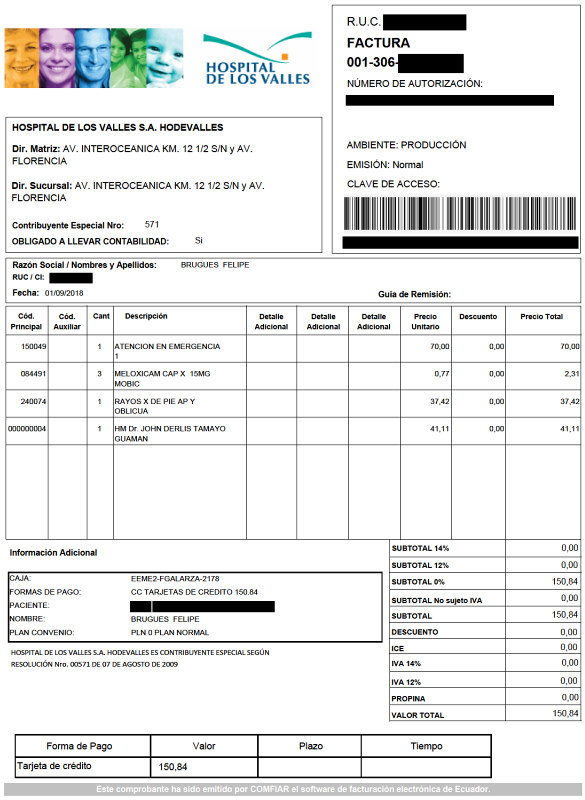
Appendix

A Example of Electronic Invoice

In Figure A.1, I present an example of an electronic invoice (EI) that is sent to buyers and to the government. The EI has a unique identifier (*Número de Autorización*). It lists the name of the seller (*Hospital de los Valles S.A. HODEVALLES*) as well as the unique firm ID (*R.U.C.*). It collects the name of the buyer (*Razón Social / Nombres y Apellidos*) as well as their unique firm/person ID (*RUC/CI*). It lists the date of transaction (*Fecha*).

In terms of the transaction itself, the EI collects the unique internal product barcode (*Cód. Principal* together with *Cód Auxiliar*, if any), as well as the written description (*Descripción*). It lists the quantity for that product (*Cant*), the listed unit price (*Precio Unitario*), any reported discounts (*Descuento*), and the total value of the product in the transaction (*Precio Total*). The EI lists the pre-tax value of the transaction (*SUBTOTAL*) as well as the total post-tax value (*VALOR TOTAL*). It also contains information on how the transaction is paid (*Forma de Pago*) and the terms of payment (*Plazo*), if accorded between the buyer and seller .

Figure A.1: **Example of an Electronic Invoice**



HOSPITAL DE LOS VALLES

R.U.C. [redacted]
FACTURA
001-306-[redacted]
NÚMERO DE AUTORIZACIÓN:
[redacted]

AMBIENTE: PRODUCCIÓN
EMISIÓN: Normal
CLAVE DE ACCESO:
[redacted]

HOSPITAL DE LOS VALLES S.A. HODEVALLES
Dir. Matriz: AV. INTEROCEANICA KM. 12 1/2 S/N y AV. FLORENCIA
Dir. Sucursal: AV. INTEROCEANICA KM. 12 1/2 S/N y AV. FLORENCIA
Contribuyente Especial Nro: 571
OBLIGADO A LLEVAR CONTABILIDAD: Si

Razón Social / Nombres y Apellidos: BRUGUES FELIPE
RUC / CI: [redacted]
Fecha: 01/09/2018 **Guía de Remisión:**

Cód. Principal	Cód. Auxiliar	Cant	Descripción	Detalle Adicional	Detalle Adicional	Detalle Adicional	Precio Unitario	Descuento	Precio Total
150049		1	ATENCION EN EMERGENCIA 1				70,00	0,00	70,00
084491		3	MELOXICAM CAP X 15MG MOBIC				0,77	0,00	2,31
240074		1	RAYOS X DE PIE AP Y OBLICUA				37,42	0,00	37,42
000000004		1	HM Dr. JOHN DERLIS TAMAYO GUAMAN				41,11	0,00	41,11

Información Adicional

CAJA:	EEME2-FGALARZA-2178
FORMAS DE PAGO:	CC TARJETAS DE CREDITO 150.84
PACIENTE:	[redacted]
NOMBRE:	BRUGUES FELIPE
PLAN CONVENIO:	PLN 0 PLAN NORMAL

HOSPITAL DE LOS VALLES S.A. HODEVALLES ES CONTRIBUYENTE ESPECIAL SEGÚN RESOLUCIÓN Nro. 00571 DE 07 DE AGOSTO DE 2009

SUBTOTAL 14%	0,00
SUBTOTAL 12%	0,00
SUBTOTAL 0%	150,84
SUBTOTAL No sujeto IVA	0,00
SUBTOTAL	150,84
DESCUENTO	0,00
ICE	0,00
IVA 14%	0,00
IVA 12%	0,00
PROPINA	0,00
VALOR TOTAL	150,84

Forma de Pago	Valor	Plazo	Tiempo
Tarjeta de crédito	150,84		

Este comprobante ha sido emitido por COMFIAR el software de facturación electrónica de Ecuador.

Notes: This figure presents an example of an electronic invoice that would be received by the buyer and the government after a transaction occurred.

B Summary Statistics

In this section, I report summary statistics of the characteristics of buyers and sellers in my sample, the electronic invoice database, dispersion of prices and quantities, as well as accounting markups.

B.1 Buyers and Sellers

Table B.1 shows basic descriptive statistics for buyers and sellers. Sellers are larger, older, and have more direct contact with international trade than buyers. The median seller in my sample has yearly sales of more than 3.5 million USD, while the median buyer has yearly sales of 0.17 million USD.[35] The median seller is 19 years old, while median buyer is 15 years old.

Table B.1: **Summary Statistics - Sellers and Buyers in 2017**

Sellers			
	Mean	Median	SD
Total Sales (million USD)	9.39	3.68	18.31
Total Inputs (millon USD)	6.83	2.37	14.94
Age	25.15	19.00	18.99
Import Share (%)	15.75	5.02	21.52
Export Share (%)	5.09	0.00	17.11
Observations	107		
Buyers			
	Mean	Median	SD
Total Sales (million USD)	2.35	0.17	46.46
Total Inputs (millon USD)	1.84	0.13	29.02
Age	16.02	15.00	9.80
Import Share (%)	4.37	0.00	14.60
Export Share (%)	1.13	0.00	9.37
Observations	40,005		

Notes: This table reports summary statistics about the size, age, and trade exposure of buyers and sellers in the sample for the year 2017. Monetary values are in U.S. dollars for 2017.

B.2 Electronic Invoice Data

Table B.2 presents basic summary statistics related to the electronic invoice database.[36]. The average (median) seller in my sample has 420 (70) buyers. This degree distribution is larger than that observed in Belgium and Costa Rica, where the average (median, p90) seller has 123 (26, 245) and 21 (6, 29) buyers, respectively (Bernard et al., 2019; Urena et al., 2018). These databases have a minimum cutoff of yearly transactions amounting to 250 euros in Belgium

[35] Given the roll-out method of the EI system, the firms in my sample are also large compared the median firm in manufacturing, which has yearly sales of 0.41 million USD in 2017. The median firm in the studied sectors not in my EI sample has yearly sales of 0.22 million USD in 2017

[36] Need to update this table with the full dataset in the government

and USD 4,800 in Costa Rica. Restricting to those minimums yield degree numbers more in line with other countries: in my database, average (median) sellers have 136 (48) buyers with total transactions greater than USD 250 and 49 (17) buyers with total transactions greater than USD 4,800. Notice as well that my EI database captures information of firms in the top 10 percentile of sales in their respective sectors, so their out-degree statistics are more likely to correspond to the top percentiles in other countries.

The average (median) seller reports 2.4 (0.14) million units of their products, and 23 (11) percent of the units go to new buyers. The average (median) buyer purchases more than 100K (1K) units and has a yearly bill of around 200K (19K) USD.

Table B.2: **Summary Statistics - Electronic Invoice Database**

	Mean	Median	SD
Over 2016-2017			
N. Buyers	418.58	70.00	1,402.99
N. Buyers (>USD 250)	136.33	48.00	209.31
N. Buyers (>USD 4,800)	49.31	17.00	89.02
Total Q (million)	2.40	0.14	4.83
Share Q New Buyers	0.23	0.11	0.28
Q per Buyer	104,758.68	1,880.28	347,033.74
Yearly Bill per Buyer (USD)	197,653.95	19,061.65	453,356.05
Yearly Bill per Buyer (USD) (>USD 250)	221,715.69	21,902.35	506,962.14
Yearly Bill per Buyer (USD) (>USD 4,800)	379,997.93	60,371.27	958,042.15
Over 2007-2017			
Uncollectibles/Sales	0.66%	0.36%	0.81%
Observations	107		

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 refers to the. Yearly bill is the total value of the transactions between buyer and seller. Uncollectibles are the receivables claimed by the seller to be impossible to collect after 5 years since the debt was emitted. Uncollectibles over sales is calculated by summing all uncollectibles claims from 2007 until 2017 and diving over all sales from the same period.

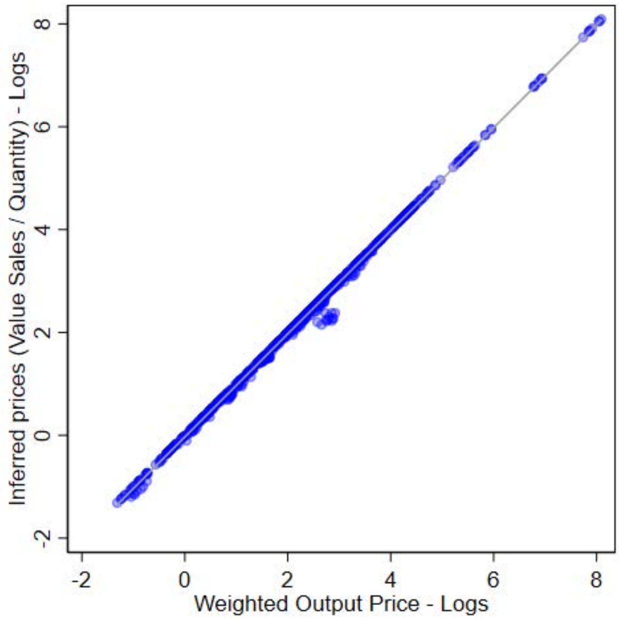
B.3 Average Price vs Weighted Price

Figure B.2 presents a scatter plot of average unit price obtained through equation 4 against the weighted average discount inclusive unit price. The weighted average discount inclusive unit price p_{ijy} is defined as:

$$p_{ijy} = \sum_{r \in R_{ijy}} \sum_{g \in G_{ijry}} s_{ijgry} * p_{ijgry},$$

for set of goods G_{ijry} in transaction r with share of expenditure s_{ijgry} summed over all transactions r between i and j in a given year y.

Figure B.2: **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 total value of yearly transactions by total quantity purchased (pooling different products). Weighted prices are calculated by summing transaction-product-level unit prices by total expenditure share.

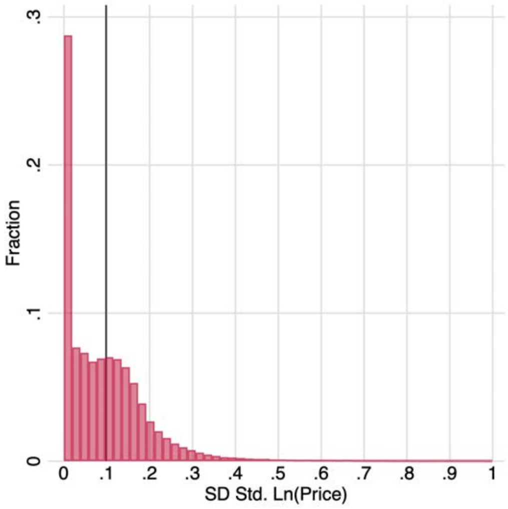
B.4 Price and Quantity Dispersion

Figures B.3 and B.4 show the dispersion of standardized log prices and quantities, respectively.

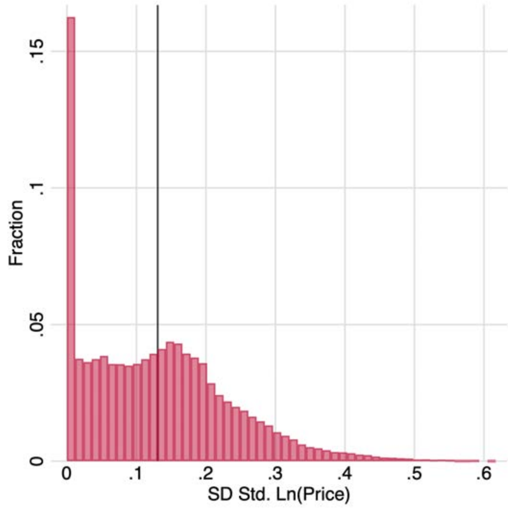
Figure B.3 shows that the average product has an average standard deviation of prices close to 0.10. This implies that in a given, the same product could have prices that are 10% higher or lower than the average price more than 30% of the time. Similarly, Figure B.4 shows that the average standard deviation of quantities for a given product in a month is close to 0.4.

Figure B.3: **Product-level Price Dispersion within Month and Year**

(a) Month



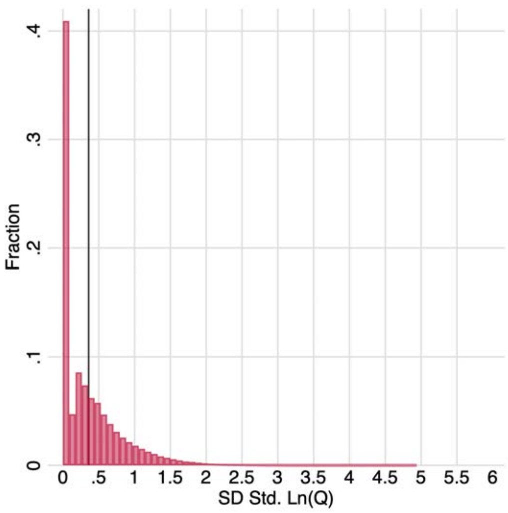
(b) Year



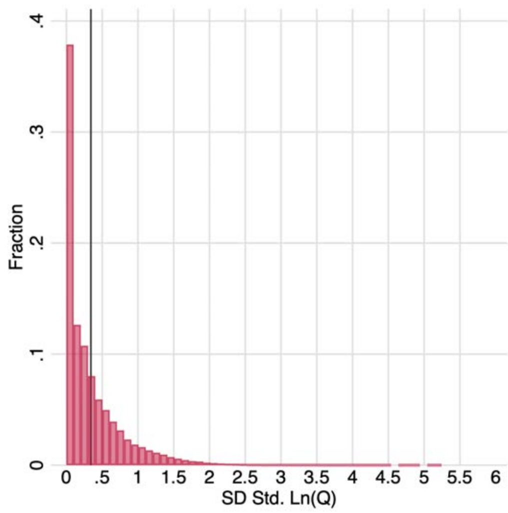
Notes: These figures plot histograms of the standard deviation of standardized log prices by month and year, for products that have at least 5 distinct buyers in time window.

Figure B.4: **Product-level Quantity Dispersion within Month and Year**

(a) Month



(b) Year

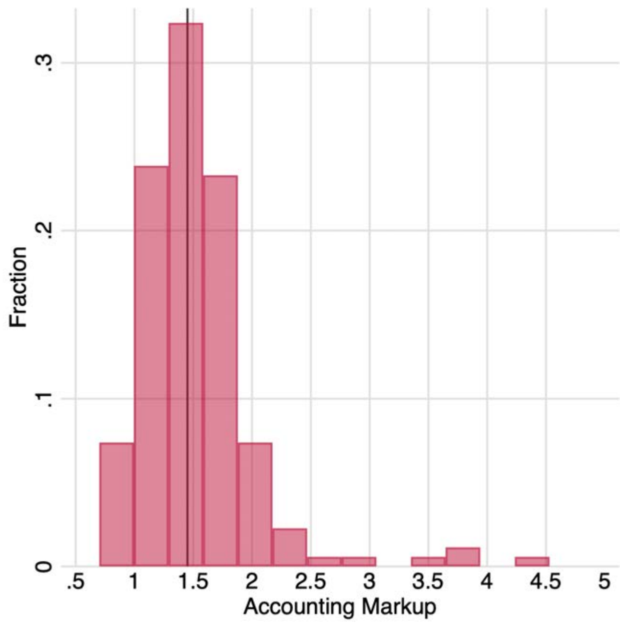


Notes: These figures plot histograms of the standard deviation of standardized log quantity by month and year, for products that have at least 5 distinct buyers in time window.

B.5 Accounting Markups

Figure B.5 shows the distribution of accounting markups, which are defined as total sales over total variable costs. Markups are relatively high, with the average markup being 50% the value of average variable costs.

Figure B.5: **Distribution of Accounting Markups**



Notes: This figure presents histogram of estimated accounting markups in 2016 and 2017. Accounting markups are calculated as total sales over total variable costs (total wage plus total expenditure on intermediate inputs)

C Additional Tables and Figures

C.1 Motivating Evidence

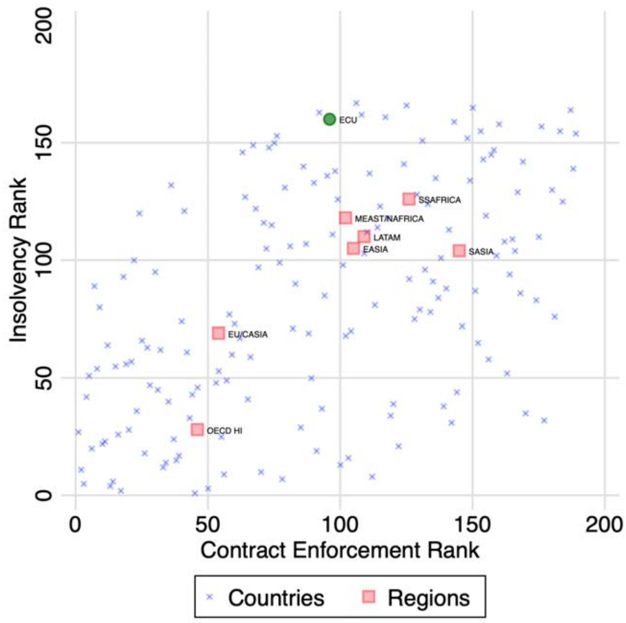
Figure C.6 shows the position of Ecuador in terms of contract enforcement and insolvency in the World Bank Doing Business report. Lower numbers represent better institutions to enforce contracts or solve insolvency cases.

Figure C.7 shows the distribution of Herfindahl-Hirschman Indices (HHI) for manufacturing 6-digit sectors in 2017. HHI_s for sector s is estimated using the following formula:

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

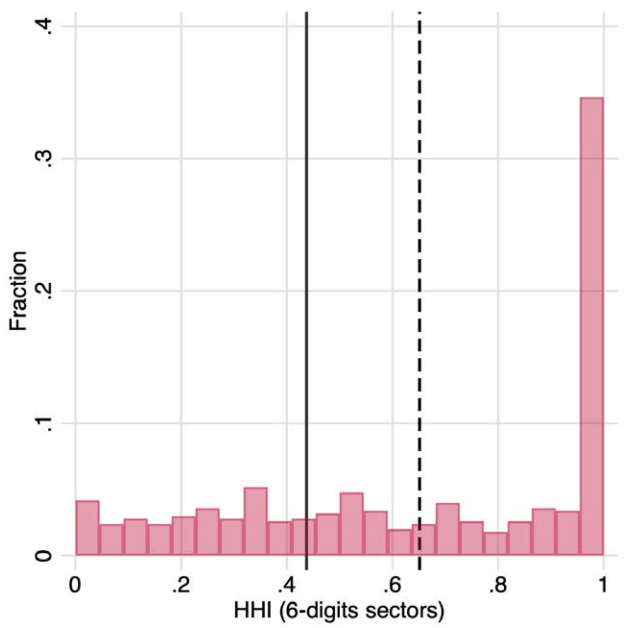
where m_j is the market share of firm j, J_s is the set of active firms in sector s. The market share of firm j is obtained by dividing total revenue of firm j by the sum total revenue of all firms in sector s.

Figure C.6: **Ranks Insolvency and Enforcement**



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

Figure C.7: **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.

C.2 Robustness - Stylized Facts

Table C.3 shows the results of a regression of standardized log prices or log average prices on age of the relationship interacted by an indicator for seller sector (4-digits). The regression controls for quantity.

Table C.3: **Robustness - Seller's Sector**

VARIABLES	(1) Std. ln(Price)	(2) ln(Price)
Textiles		
Age of Relationship	-0.00296*	-0.0867***
	(0.00172)	(0.0291)
Pharmaceuticals		
Age of Relationship	-0.00597***	-0.0158*
	(0.00183)	(0.00893)
Cements		
Age of Relationship	-0.00436***	-0.0275***
	(0.00126)	(0.00714)
Seller-Year FE	No	Yes
Controls	Yes	Yes
Observations	87,882	87,882
R-squared	0.043	0.543

Notes: This table presents regression of prices on age of relationship by sector of the seller. Column (1) presents results for the standardized log prices. Column (2) presents results for log average price, controlling for seller-year fixed effects. Both columns control for standardized quantities. Standard errors are clustered at the seller-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure C.8 presents the coefficients of a regression of standardized log quantity on age of the relationship controlling for pair fixed effects.

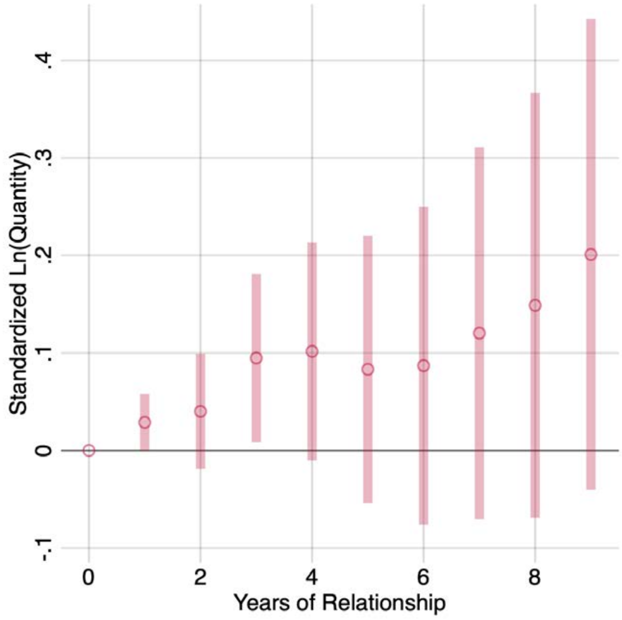
Figure C.9 shows the coefficients of a regression on log total value on age of relationship controlling for pair-fixed effects. The red results show the path of value traded within relationship using a partial panel with only two observations per pair from the EI database. In contrast, the purple results show the path using a full panel of up to 6 years from the VAT database.

Table C.4 shows the results of a regression on log average price on log quantity, controlling for seller-year fixed effects.

Table C.5 present regression results of log average price on age of relationship under different models. The results indicate that the cross-sectional model with seller-year fixed effects fails to capture the within-pair dynamics of prices. However, by using a spline on the hazard probability by quantile of quantity, the cross-sectional model is able to match the dynamics in the panel model.

Table C.6 presents robustness results for the relationship between prices and age of relationship.

Figure C.8: **Standardized Log Quantity and Age of Relationship - Within Pair**



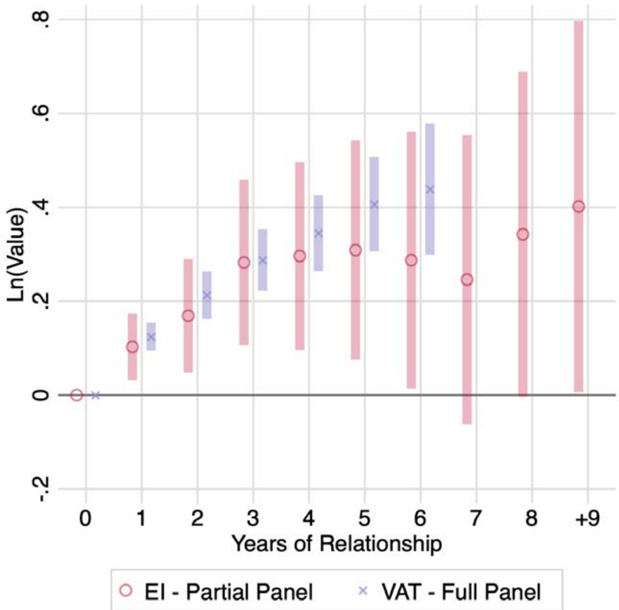
Notes: This figure plots coefficients for a regression of standardized log quantity on age of relationship, controlling for pair fixed effects. Error bars represents ± 1.96 standard errors, clustered at the pair-level.

Table C.4: **Benchmark: Quantity Discounts**

VARIABLES	(1) ln(Price)
ln(Quantity)	-0.209***
	(0.0243)
Constant	3.155***
	(0.0641)
Seller-Year FE	Yes
Observations	88,801
R-squared	0.611

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

Figure C.9: **Total Value and Age of Relationship - Within Pair**



Notes: This figure plots regression coefficients for the the value of total sales between buyer and supplier on age of 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 years 2016-2017. The purple marks are constructed using multiple observations of buyer-seller pairs from the VAT B2B database for years 2007-2015 for the sellers in the electronic invoice database.

Table C.5: **Price Regressions: Controlling for Survival**

	(1)	(2)	(3)
	Cross-Sct.	Cross-Sct.	Panel
VARIABLES	ln(Price)	ln(Price)	ln(Price)
Age of Relationship	-0.052**	-0.014***	-0.017***
	(0.023)	(0.004)	(0.005)
Seller-Year FE	Yes	Yes	No
Flex. Hazard Control	No	Yes	No
Pair FE	No	No	Yes
Observations	88,801	44,339	39,812
R-squared	0.535	0.539	0.931

Notes: This table compares the regression coefficient of prices and age of relationship under different models. Column (1) presents regression model with log average unit price as dependent variable and controls for seller-year fixed effects. Column (2) adds to Column (1) a flexible spline for hazard probability by percentile of quantity. Column (3) presents regression model with log average unit price as dependent variable and controlling for pair fixed effects. Hazard probability is estimated for each percentile of quantity and seller as the probability that a buyer in such category in 2016 would not transact with the seller again in 2017. Standard errors are clustered at the seller-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

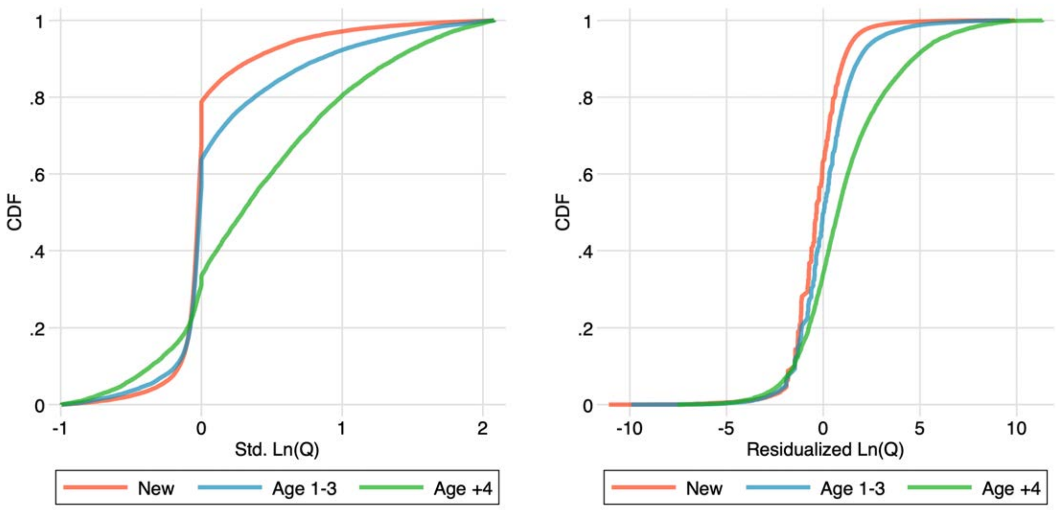
Table C.6: **Robustness - Standardized Log Price**

VARIABLES	(1) Std. ln(Price)	(2) Std. ln(Price)	(3) Std. ln(Price)	(4) Std. ln(Price)	(5) Std. ln(Price)	(6) Std. ln(Price)	(7) Std. ln(Price)	(8) Std. ln(Price)	(9) Std. ln(Price)	(10) Std. ln(Price)	(11) Std. ln(Price)	(12) Std. ln(Price)
Std. ln(Quantity)	-0.0425***	-0.0426***	-0.0427***	-0.0427***	-0.0429***	-0.0426***	-0.0425***	-0.0426***	-0.0425***	-0.0426***	-0.0425***	-0.0429***
	(0.00391)	(0.00392)	(0.00393)	(0.00394)	(0.00393)	(0.00391)	(0.00392)	(0.00393)	(0.00391)	(0.00391)	(0.00392)	(0.00397)
Age of Relationship	-0.00420***	-0.00419***	-0.00459***	-0.00444***	-0.00461***	-0.00419***	-0.00420***	-0.00426***	-0.00413***	-0.00424***	-0.00418***	-0.00464***
	(0.00148)	(0.00145)	(0.00146)	(0.00135)	(0.00139)	(0.00146)	(0.00144)	(0.00145)	(0.00140)	(0.00142)	(0.00147)	(0.00136)
ln(Age Buyer)	0.000984											-0.00149
	(0.000978)											(0.00102)
ln(Distance Km)		0.00103										0.00109
		(0.00217)										(0.00210)
ln(Sales Buyer)			0.00124***									0.000846***
			(0.000246)									(0.000254)
ln(N. Employees Buyer)				0.00121								-4.72e-05
				(0.00103)								(0.000742)
ln(Assets Buyer)					0.00105***							0.000900***
					(0.000316)							(0.000205)
1{Multinational Buyer}						0.0248**						0.0201*
						(0.0123)						(0.0111)
1{Exporter Buyer}							0.000612					-0.00285
							(0.00606)					(0.00483)
1{Importer Buyer}								0.00436*				0.00336
								(0.00239)				(0.00214)
1{BG Buyer}									-0.00225			-0.00540**
									(0.00397)			(0.00222)
Supply Share										0.00768		0.0130
										(0.0130)		(0.0134)
Demand Share											-0.00793	-0.0103
											(0.0246)	(0.0294)
Constant	0.0169**	0.0150***	0.00534	0.0178*	0.00864	0.0194**	0.0195**	0.0191**	0.0197**	0.0195**	0.0195**	3.47e-05
	(0.00850)	(0.00484)	(0.00889)	(0.00933)	(0.0106)	(0.00835)	(0.00838)	(0.00846)	(0.00865)	(0.00834)	(0.00834)	(0.00584)
Observations	88,767	87,914	88,801	88,801	88,801	88,801	88,801	88,801	88,801	88,801	88,801	87,882
R-squared	0.041	0.041	0.042	0.041	0.042	0.041	0.041	0.041	0.041	0.041	0.041	0.043
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	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

Figure C.10: **Cumulative Distribution Function of Quantities by Relationship Age**

(a) Product-level Standardized Ln(Q) (b) Seller-level Residualized Ln(Q)



Notes: These figures plot the cumulative distribution functions for standardized log quantities (left) and residualized log quantities (right) by different ages of relationship.

Figure C.10 presents the cumulative distribution of quantities, both in relative terms through standardized quantities and in absolute values through residualized log quantities, for different age of relationships groups.

D Evidence from Multinational Buyers

The model also allows for the limited enforcement constraint to be slack not only endogenously through the future surplus generated by the relationship but also to differ according to some visible category that exogenously shifts the difficulty of enforcing contracts.[37] In Antras and Foley (2015), the authors find that the use of trade credit is ex-ante higher when the buyer comes from a common-law country, arguing that common-law countries better protect property and contract rights. In the same line, although trade occurs within the Ecuadorian legal system, contracting with multinational buyers from common-law origin could mean the enforcement constraint is slack at the beginning of the relationship.

I offer two explanations for why this might be the case. First, contracting with multinationals is qualitatively different than contracting with domestic firms. Alfaro-Urena et al. (2019) find that firms improve estimated productivity and management practices reported by managers after the first year supplying to a multinational. Moreover, managers expect sales to multinationals to be markedly different from sales to domestic firms. In a survey conducted by Alfaro-Urena et al. (2019), they find that, only after the size of the purchase, managers expected to see the largest difference in the reliability of the payment relative to a domestic firm.

[37] In the model, this is accounted by allowing for a category-specific constant in the limited enforcement constraint.

Therefore, the limited enforcement constraint could possibly be slack at the beginning of the relationship for multinationals.

Second, multinationals differ from one another in their management practices and legal origin. Bloom et al. (2012) and Hjort et al. (2020) find that multinationals export their management practices to their foreign affiliates. Therefore, within Ecuador, we should expect differences across multinationals according to their HQ's origin.

Taken together, we should expect to see a weaker decrease (weaker backloading) of prices over time for multinationals relative to domestic buyers. Moreover, within the group of multinationals, we should see a faster decrease (more backloading) in prices in civil law or other legal origin multinationals relative to common law multinationals.

The Business Bureau in Ecuador (Superintendencia de Compañias) collects information on the ownership of all private firms in Ecuador, including country of origin for multinational companies. I obtain the legal tradition of the origin countries from La Porta et al. (1999). In my current sample, I observe 71 multinational buyers,[38] 11 of which have HQ in common law countries and 60 in other countries.

Table D.7 show the results of the tests, restricting the sample of sellers to firms that have at least one multinational buyer. Columns (1) and (2) test whether multinationals are subject to backloading. On both, the negative coefficient on age of relationship captures the general backloading either in the cross-section or within-pair. However, the coefficients on the interactions of age of relationship with the multinational dummy offset the coefficient on age of relationship, indicating that multinationals do not see backloading of prices.

Columns (2)-(8) studies the differences between multinationals from a common law country and other multinationals. Column (3) runs a regression of standardized log prices on age of the relationship for the set of common law multinationals and Column (4) for all other multinationals. I do not include any fixed effects as prices are standardized for each product-year. Column (3) shows that prices tend to increase as relationship ages for common law multinationals, while Column (4) indicates a slight backloading for other multinationals. Column (5) finds the same results when pooling all multinationals together and Column (6) shows a similar result when including all buyers. Column (7) shows that the result holds when using log unit prices with seller-year fixed effect instead. In this regression, however, there seem to be slightly weaker backloading for multinationals with other legal origin than domestic buyers, although the effect is not statistically significant. Although the results are noisy, Column (8) indicates that, within a pair, common law multinationals see no backloading while other multinationals see weaker backloading than domestic firms.

To explore the generality of my results, I conduct a similar exercise using export data at the firm-product-destination-year level for Peru (1993-2009), Uruguay (2001-2012), and Ecuador (2013-2018). I obtain the firm-level data for Peru and Uruguay from the Exporter Dynamics database of the World Bank and the data for Ecuador from the Servicio Nacional de Aduanas del Ecuador. I define the variable *tenure in sample* as the number of years since a firm entered

[38] There are only 673 multinationals active in the whole economy in 2016.

the destination country with a specific product. To account for inflation, I adjust prices using country of origin's consumer price index by the World Bank.

Table D.8 shows the results. In all three countries, prices decrease over time within firm-production-destination. However, the decrease in prices is slighter (Peru and Uruguay) or nonexistent (Ecuador) when the export destination is a common law country.

Table D.7: **Price Dynamics and Legal Origin of Buyer**

VARIABLES	(1) Std. ln(Price)	(2) ln(Price)	(3) ln(Price)	(4) Std. ln(Price)	(5) Std. ln(Price)	(6) Std. ln(Price)	(7) Std. ln(Price)	(8) ln(Price)	(9) ln(Price)
Multinational Buyers Only	No	No	No	Yes	Yes	Yes	No	No	No
Legal Origin	All	All	All	Common	Other	All	All	All	All
Age of Relationship	-0.00527***	-0.0795***	-0.0216***	0.0197*	-0.00527*	-0.00527*	-0.00527***	-0.0795***	-0.0216***
	(0.000284)	(0.00333)	(0.00633)	(0.0114)	(0.00308)	(0.00309)	(0.000284)	(0.00333)	(0.00633)
1{ Multinational Buyer }	0.00428	-0.0321							
	(0.0125)	(0.0875)							
1{ Multinational Buyer } X Age of Relationship	0.00508	0.0981**	0.0318						
	(0.00375)	(0.0383)	(0.0979)						
1{ Common Law }						-0.0471	-0.0373	-0.379	
						(0.0345)	(0.0313)	(0.240)	
1{ Common Law } X Age of Relationship						0.0250**	0.0250**	0.290**	0.106
						(0.0114)	(0.0109)	(0.139)	(0.0848)
1{ Other Law }							0.00979	0.0145	
							(0.0135)	(0.0947)	
1{ Other Law } X Age of Relationship							-2.26e-06	0.0479	0.0120
							(0.00307)	(0.0296)	(0.122)
Constant	0.0239***	2.846***	2.737***	-0.0135	0.0337**	0.0337**	0.0239***	2.846***	2.737***
	(0.000539)	(0.00349)	(0.00999)	(0.0327)	(0.0135)	(0.0136)	(0.000539)	(0.00349)	(0.00999)
Seller-Year FE	No	Yes	No	No	No	No	No	Yes	No
Pair FE	No	No	Yes	No	No	No	No	No	Yes
Observations	70,336	70,336	27,182	24	148	172	70,336	70,336	27,182
R-squared	0.005	0.411	0.899	0.147	0.007	0.025	0.005	0.411	0.899

Notes: Dependent variable noted as Std. ln(Price) is standardized log unit price and ln(Price) is log average unit price. Age of relationship is defined as the total number of years that the pair has transacted since the seller entered the VAT database. Indicator for multinational is obtained from registry in the Servicio de Rentas Internas. Common Law and Other Law dummies are obtained from La Porta et al. (1997). Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table D.8: **Price Dynamics and Limited Enforcement in Export Data**

	(1)	(2)	(3)
Country of Origin	**Peru**	**Uruguay**	**Ecuador**
VARIABLES	ln(Price)	ln(Price)	ln(Price)
Tenure in Sample	-0.0169***	-0.0168***	-0.0123***
	(0.00206)	(0.00173)	(0.00391)
$\mathbf{1}\{\text{Common Law}\}$ x Tenure in Sample	0.00859**	0.00292	0.0472***
	(0.00335)	(0.00392)	(0.00752)
Constant	2.323***	2.152***	0.487***
	(0.00510)	(0.00538)	(0.00681)
Firm-Product-Destination FE	Yes	Yes	Yes
Observations	467,765	97,784	88,683
R-squared	0.911	0.965	0.932

Notes: Dependent variable is log unit prices, calculated as total value exported over total quantity. Tenure is sample is calculated from the first observation observed for the specific product of a give firm in the destination country. Common Law dummy is obtained from La Porta et al. (1997). All regressions control for firm-product-destination fixed effects. Standard errors are clustered at the firm-destination level. *** p<0.01, ** p<0.05, * p<0.1

E Existence

To prove existence, I build on two results of the literature. First, I use the result of non-linear pricing of Jullien (2000) to prove the existence of a stationary optimal contract in the presence of heterogeneous participation constraints. I do so by showing the equivalence between the stationary contract with limited enforcement and a non-linear pricing problem with heterogeneous outside options. Then, similar to the argument in Martimort et al. (2017), I offer an simple non-stationary deviation that dominates the stationary optimal contract.

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 simply replace the discount factor δ for $\tilde{\delta} \equiv min\{\delta(\theta)\}$, where $\delta(\theta) = \delta(1 - X(\theta))$ is the discount factor that accounts for heterogeneous breakups. This change will only affect one of the assumptions discussed below and set an upper bound in the worse-case exit rate.

E.1 Existence of Stationary Contract

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

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

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

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

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

$$v_q(\theta, q(\theta)) - c = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta), \tag{31}$$

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

$$\int_{\underline{\theta}}^{\overline{\theta}} [u(\theta) - \bar{u}(\theta)] d\gamma(\theta) = 0. \tag{32}$$

Jullien (2000) shows that under three assumptions there exists a unique optimal solution in which all consumers participates, which is characterized by the first-order conditions 31 and complementary slackless condition 32 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

$$\frac{d}{d\theta}\Big(\frac{S_q(\theta, q)}{v_{\theta q}(\theta, q)}\Big) \geq 0$$
$$\frac{d}{d\theta}\Big(\frac{F(\theta)}{f(\theta)}\Big) \geq 0 \geq \frac{d}{d\theta}\Big(\frac{1 - F(\theta)}{f(\theta)}\Big).$$

The second and *key* assumption is homogeneity (H), requiring that there exists a quantity profile $\{\bar{q}(\theta)\}$ such that the allocation with full participation $\{\bar{u}(\theta), \bar{q}(\theta)\}$ is implementable in that $\bar{u}'(\theta) = v_\theta(\theta, \bar{q}(\theta))$ and $\bar{q}(\theta)$ 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) assumes 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. $s(\theta, \bar{q}(\theta)) \geq \bar{u}(\theta)$.

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:

$$\max_{\{t(\theta), q(\theta)\}} \frac{1}{1-\delta} \int_{\underline{\theta}}^{\overline{\theta}} [t(\theta) - cq(\theta)] f(\theta) d\theta \text{ s.t.} \tag{LE Problem}$$
$$v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\hat{\theta})) - t(\hat{\theta}) \;\; \forall \theta, \hat{\theta} \tag{IC}$$
$$\frac{\delta}{1-\delta}\Big(v(\theta, q(\theta)) - t(\theta)\Big) \geq t(\theta) = v(\theta, q(\theta)) - u(\theta), \;\; \forall \theta, \tag{LC}$$

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):

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

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

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

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

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

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

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

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 $s(\theta, \bar{q}(\theta)) \geq (1-\delta)\theta v(\bar{q}(\theta))$. The condition becomes:

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

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 31 and 32. Assuming that the $v(\cdot)$, $F(\theta)$, and δ are such that A1, A2, and A3 hold for $\{t^{st}(\theta), q^{st}(\theta)\}$, then $\{t^{st}(\theta), q^{st}(\theta)\}$ is uniquely optimal.

E.2 Optimality of Non-Stationary Contracts

Having established the existence of an optimal stationary contract, I now show that a non-stationary contract exists, which dominates the stationary contract. A similar argument was briefly discussed in the working paper version of Martimort et al. (2017).

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, $u_{st = \theta v(q^{st}(\theta)) - t^{st}(\theta)}$ and $t_0(\theta) = t^{st}(\theta)$. Define $q_0(\theta)$ to so it satisfies that the deviation defined above.. Under this deviation, the enforcement constraint at $\tau = 0$ is:

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

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 additional payoff ε while still satisfying both the incentive compatibility and limited enforcement constraints. Therefore, the optimal contract is non-stationary.

F Proof that Gamma Equals One for Highest Type

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 k, the outside option is uniformly shifted downward by > 0 for all θ, that is, $\overline{u}_k(\theta) = -\varepsilon$. The enforcement constraint at k is now given by:

$$\delta[\sum_{s=1}^{\infty} \delta^{s-1} u_{k+s}(\theta)] - \overline{u}_k(\theta) = \sum_{s=1}^{\infty} \delta^s u_{k+s}(\theta) + \varepsilon \geq t_k(\theta) = \theta v(q_k(\theta)) - u_k(\theta). \tag{33}$$

The seller's problem in the Lagrangian-form is

$$W(\varepsilon) = \max_{\{q_\tau(\theta), u_\tau(\theta)\}} \sum_{\tau=0}^{\infty} \delta^\tau \Big\{ \int_{\underline{\theta}}^{\overline{\theta}} [\theta v(q_\tau(\theta)) - cq_\tau - u_\tau(\theta)] f(\theta) d\theta + \tag{34}$$

$$\int_{\underline{\theta}}^{\overline{\theta}} [\sum_{s=1}^{\infty} \delta^s u_{\tau+s} + \varepsilon * 1\{\tau = k\} - t_\tau(\theta)] d\Gamma_\tau(\theta) \Big\} \tag{35}$$

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

$$\frac{dW(\varepsilon)}{d\varepsilon} = \delta^k \int_{\underline{\theta}}^{\overline{\theta}} d\Gamma_k(\theta), \tag{36}$$

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 transfers are all shifted upward by the constant . That is, if $q_\tau(\theta)$ is the solution for the problem with $\overline{u}_\tau(\theta) = 0$ for all θ and all τ with associated $t_\tau(\theta)$, $q_\tau(\theta)$ is also the solution for the problem with outside options $\overline{u}_\tau(\theta) = -\varepsilon 1\{\tau = k\}$ for all θ and all τ with associated transfers equal to $t_\tau(\theta) + \varepsilon 1\{\tau = k\}$. The value of the problem for the seller is:

$$W(\varepsilon) = \sum_{\tau=0}^{\infty} \delta^\tau \Big\{ \int_{\underline{\theta}}^{\overline{\theta}} [t_\tau(\theta) + \varepsilon 1\{\tau = k\} - cq_\tau] f(\theta) d\theta \Big\} \tag{37}$$

$$= \sum_{\tau=0}^{\infty} \delta^\tau \Big\{ \int_{\underline{\theta}}^{\overline{\theta}} [t_\tau(\theta - cq_\tau] f(\theta) d\theta \Big\} + \delta^k \varepsilon. \tag{38}$$

So

$$\frac{dW(\varepsilon)}{d\varepsilon} = \delta^k. \tag{39}$$

Therefore, the cumulative multiplier for any k will satisfy the following property:

$$\Gamma_k(\overline{\theta}) \equiv \int_{\underline{\theta}}^{\overline{\theta}} d\Gamma_k(\theta) = \frac{dW(\varepsilon)}{d\varepsilon}\frac{1}{\delta^k} = 1. \tag{40}$$

G Proofs - Model Dynamics

Proof of Proposition 1. 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:

$$T''_\tau(q) = \theta'_\tau(q)v'(q) + \theta_\tau(q)v''(q) = \theta(q)v'(q)\Big[\frac{\theta'_\tau(q)}{\theta_\tau(q)} + \frac{v''(q)}{v'(q)}\Big] \tag{41}$$

$$= T'(q)\Big[\frac{1}{\theta_\tau(q)q'_\tau(\theta)} - A(q)\Big], \tag{42}$$

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 Number we obtain an expression for $q'_\tau(\theta)$:

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

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

The denominator of the equation above is positive as $v'(q_\tau(\theta)) > 0$ and $c > 0$. As by assumption, strict monotonicity holds $q'_\tau(\theta) > 0$, then the numerator is also positive. Substituting in 41 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

$$\frac{\Big[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\Big]}{\theta\frac{d}{d\theta}\Big[\theta - \frac{\Gamma_\tau(\theta)-F_\tau(\theta)-\sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta))+\theta\gamma_\tau(\theta)}{f_\tau(\theta)}\Big]} \leq 1 \tag{43}$$

Define the $\Lambda_\tau(\theta) \equiv \Gamma_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta)) + \theta\gamma_\tau(\theta)$ and $\lambda_\tau(\theta) \equiv d\Lambda_\tau(\theta)/d\theta$. Inequality 43 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

$$[\Lambda_\tau(\theta) - F_\tau(\theta)][f_\tau(\theta) + f'_\tau(\theta)\theta] \geq \theta f(\theta)[\lambda_\tau(\theta) - f_\tau(\theta)]. \tag{44}$$

As noted above, $\theta f_\tau(\theta) \geq \Lambda_\tau(\theta) - F_\tau(\theta)$. Note that log-concavity of the density $F_\tau(\theta)$ is sufficient to satisfy the assumption of monotone hazard condition. For 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 43 as

$$(\theta - 1)f_\tau(\theta) + f_\tau(\theta) \geq [F_\tau(\theta) - \Lambda_\tau(\theta)]\Big(1 + \frac{f'_\tau(\theta)\theta}{f_\tau(\theta)}\Big) + \lambda_\tau(\theta). \tag{45}$$

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 $(\theta - 1)f(\theta) \geq F_\tau(\theta)(1 + \theta f'_\tau(\theta)/f_\tau(\theta))$, which can be expressed as:

$$\frac{d}{d\theta}\Big(\frac{F_\tau(\theta)}{f_\tau(\theta)}\Big) = \frac{f_\tau(\theta)^2 - F_\tau(\theta)f'_\tau(\theta)}{f_\tau(\theta)^2} \geq \frac{F_\tau(\theta)}{(\theta - 1)f_\tau(\theta)}. \tag{46}$$

□

Proof of Proposition 2. 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

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

which can be written as

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

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:

$$L(\theta, U, q, \lambda, \gamma) = (\theta v(q(\theta)) - cq(\theta) - U)f(\theta) + \lambda v(q(\theta)) + \gamma(U + \delta U^+ - \theta v(q(\theta))), \tag{47}$$

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 implies with a lower $\gamma(\theta)$.

For $\gamma_\tau(\theta) = 0$ for some finite $\tau > \tau^*$ for all θ. 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.

For $q_{\tau^*}(\theta) > q_\tau(\theta)$ 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, or when it binds in some period between τ and τ^* for some θ between θ and $\overline{\theta}$.

□

Proof of Proposition 3. 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. □

H Solution of Gamma Function

The seller's first order condition for each τ defines the following differential equation

$$\theta u'(q_\tau(\theta)) - c = \frac{\Gamma_\tau(\theta) - F_\tau(\theta) - \sum_{s=0}^{\tau-1}(1-\Gamma_s(\theta)) + \theta\gamma_\tau(\theta)}{f_\tau(\theta)} u'(q_\tau(\theta)). \tag{48}$$

For $\tau = 0$, the solution $\Gamma_0(\theta)$ to the equation above is given by:

$$\Gamma_0(\theta) = \frac{\int_{\underline{\theta}}^{\theta} [x f_0(x) - c(u'(q_0(x))^{-1} f_0(x) + F_0(x)] dx + K_0}{\theta}, \tag{49}$$

which by integration by parts reduces to:

$$\Gamma_0(\theta) = F_0(\theta)\Big(1 - \frac{c}{\theta E[(u'(q_0(\Theta)^{-1}|\Theta \leq \theta]^{-1}}\Big) + \frac{K_0}{\theta}. \tag{50}$$

The constant is obtained by using the boundary condition $\Gamma_\tau(\overline{\theta}) = 1$, which yields $K_0 = \frac{c}{E[(u'(q_0(\Theta)^{-1}]^{-1}}$.

More generally, the solution is given by:

$$\Gamma_\tau(\theta) = F_\tau(\theta)\Big(1 - \frac{c}{\theta E[(u'(q_\tau(\Theta)^{-1}|\Theta \leq \theta]^{-1}}\Big) + \frac{G_\tau(\theta)}{\theta} + \frac{K_\tau}{\theta}, \tag{51}$$

where $G_\tau(\theta) = \sum_{s=0}^{\tau-1} \int_{\underline{\theta}}^{\theta}(1-\Gamma_s(\theta))d\theta$, which can be written as $G_\tau(\theta) = \sum_{s=0}^{\tau-1} \Gamma_s(\theta) E_\Gamma[\Theta|\Theta \leq \theta]$. We again use the boundary condition $\Gamma_\tau(\overline{\theta}) = 1$ to obtain the integration constant $K_\tau = \frac{c}{E[(u'(q_\tau(\Theta)^{-1}]^{-1}} - G_\tau(\overline{\theta})$.

Given that $T'_\tau(q) = \theta u'(q_\tau(\theta)$, the multiplier of the limited enforcement constraint resizes the cumulative distribution of types by the average markup-up for lower types, which in itself, is related to the curvature of the return function. Intuitively, higher types, for which $F_\tau(\theta)$ is higher, have a larger share of other types for which the limited enforcement constraint must be binding. The extent by which the constraint binds for lower types is determined by the markups charged to those types. Under the constant marginal cost assumption, higher markups indicate higher transfers relative to the quantity obtained, which imply that all else equal, the limited enforcement constraint is more likely to bind.

I Point Identification of Gamma

In this section, I detail how $\Gamma_\tau(\cdot)$ is point identified with observations of prices, quantities, and marginal cost for one seller under two assumptions. The first assumption is the parametrization of $v(q) = kq^\beta$ for $k > 0$ and $\beta \in (0, 1)$. The second assumption requires to select one state of the world: $\{(\gamma(0) = 0, \Gamma(0) = 0); (\gamma(0) > 0, \Gamma(0) = 0); (\gamma(0) > 0, \Gamma(0) > 0)\}$. For my setting, I conjecture that the state of the world is $(\gamma(0) > 0, \Gamma(0) > 0)$ and derive sufficient conditions for such state of the world.

I.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 any possible state of the world. In this step, we omit subscripts $\tau = 0$.

Consider $\rho(\alpha) = d\ln(\theta(\alpha)) = \theta'(\alpha)/\theta(\alpha)$. Substituting in, the key identification equation 13 becomes

$$\frac{T'(q(\alpha)) - c}{T'(q(\alpha))} = \rho(\alpha)\Big[\Gamma(\alpha) - \alpha\Big] + \gamma(\alpha). \tag{52}$$

Reordering and differentiating by α yields

$$\frac{d\{(T'(q(\alpha)) - c)/T'(q(\alpha))\rho(\alpha)\}}{d\alpha}\frac{d\{(\gamma(\alpha)/\rho(\alpha)\}}{d\alpha} = \gamma(\alpha) - 1. \tag{53}$$

Integrating from 0 to 1 gives

$$\int^1 \frac{d\{(T'(q(x)) - c)/T'(q(x))\rho(x)\}}{dx}dx - \int^1 \frac{d\{(\gamma(x)/\rho(x)\}}{dx}dx = \int^1 \gamma(x)dx - 1 = 0, \tag{54}$$

where the last equality follows from $\int^1 \gamma(x)dx = 1$. Therefore,

$$\frac{T'(q(1)) - c}{T'(q(1))\rho(1)} - \frac{T'(q(0)) - c}{T'(q(0))\rho(0)} = \frac{\gamma(1)}{\rho(1)} - \frac{\gamma(0)}{\rho(0)}, \tag{55}$$

where by construction, $\gamma(1) = 0$. Reorder to obtain

$$\gamma(0) = \frac{T'(q(0)) - c}{T'(q(0))} - \frac{T'(q(1)) - c}{T'(q(1))}\frac{\rho(0)}{\rho(1)} \tag{56}$$

Use the derivative of the transfer rule to obtain $\rho(\alpha) = \theta'(\alpha)/\theta(\alpha) = T''(q(\alpha))/T'(q(\alpha)) + A(q(\alpha))$, where $A(q(\alpha)) = -v''(q(\alpha))/v'(q(\alpha))$. The assumed parametrization implies $A(q) = (1-\beta)/q$. Substituting $\rho(\cdot)$ above gives

$$\gamma(0) = \frac{T'(q(0)) - c}{T'(q(0))} - \frac{T'(q(1)) - c}{T'(q(1))}\frac{\frac{T''(q(0))}{T'(q(0))} + \frac{1-\beta}{q(0)}}{\frac{T''(q(1))}{T'(q(1))} + \frac{1-\beta}{q(1)}} \tag{57}$$

which shows a unique mapping between β and $\gamma(0)$, given knowledge of prices, quantities, and marginal cost.

I.1.1 *Case 1:* $\gamma(0) = 0$

If $\gamma(0) = 0$, equation 57 implies

$$1 - \beta = \left[\frac{T'(q(1)) - c}{T'(q(1))} \frac{T''(q(0))}{T'(q(0))} - \frac{T'(q(0)) - c}{T'(q(0))} \frac{T''(q(1))}{T'(q(1))} \right] \left[\frac{T''(q(0))}{q(0)T'(q(0))} - \frac{T''(q(1))}{q(1)T'(q(1))} \right]^{-1}. \tag{58}$$

Therefore, β is identified from observations in prices, quantities, and marginal cost when $\gamma(0) = 0$.

I.1.2 *Case 2:* $\gamma(0) > 0$ *and* $\Gamma(0) = 0$

For $\gamma(0) > 0$, substitute 57 in 13 evaluated at $\alpha = 0$, use $\rho(0)$ and rearrange to obtain:

$$\frac{T'(q(1)) - c}{T'(q(1))} \frac{1}{\frac{T''(q(1))}{T'(q(1))} + \frac{1-\beta}{q(1)}} = \Gamma(0) \tag{59}$$

If $\Gamma(\alpha)$ does not have a mass point at $\alpha = 0$, then $\Gamma(0) = 0$.[39] If $T'(q(1)) \neq c$, equation 59 implies:

$$1 - \beta = -\frac{T''(q(0))q(0)}{T'(q(0))}. \tag{60}$$

So β is identified when $\Gamma(0) = 0$ if $T'(q(1)) \neq c$. Given observations of prices and marginal costs, this last condition can be verified in the data to hold.

I.1.3 *Case 3:* $\gamma(0) > 0$ *and* $\Gamma(0) = \gamma(0)$

If $\Gamma(\alpha)$ has a mass point at $\alpha = 0$, then $\Gamma(0) = \gamma(0)$. Substitute 57 into 59 and rearrange to obtain:

$$1 - \beta = \left[\frac{T'(q(0)) - c}{T'(q(0))} \frac{T''(q(1))}{T'(q(1))} - \frac{T'(q(1)) - c}{T'(q(1))} \left(1 + \frac{T''(q(0))}{T'(q(0))} \right) \right] \left[\frac{T'(q(1)) - c}{q(0)T'(q(1))} - \frac{T'(q(0)) - c}{q(1)T'(q(0))} \right]^{-1}. \tag{61}$$

Therefore, β is identified with observations on prices, quantities, and marginal costs.

I.1.4 *A Conjecture*

By inspection of the solution to $\Gamma(\theta(0))$ in Section H, I conjecture that Case 3, i.e., $\Gamma_0(0) > 0$, is the relevant one for my setting.

[39] Recall that the measure $\gamma(\cdot)$ may have discrete jumps at some points. And specifically, I consider measures that may have discrete jumps at $\alpha = 0$.

For distributions with mass points at $\theta(0) = \underline{\theta}$, so $F(\underline{\theta}) > 0$, a sufficient condition for $\Gamma(0)$ to be positive is:

$$\underline{\theta}\beta k q(0)^{\beta-1} > c, \tag{62}$$

which is equivalent to say that serving the lowest type at observed quantity $q(0)$ is socially desirable. In practical terms, as $T'(q(0)) = \theta(0)\beta q(0)^{\beta-1}$, then the condition for identification is

$$T'(q(0)) > c, \tag{63}$$

when the distribution of θ is assumed to have a mass point at $\underline{\theta}$.

For cases with continuous distributions, so $F(\underline{\theta}) = 0$, $\Gamma(0) > 0$ will always be positive, as long as $\overline{\theta}$ is finite and c is positive.

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

Consider equation 14 and use the parametrized version of $\rho_0(\alpha)$:

$$\Xi_0(\alpha) = \alpha + \frac{T_0'(q_0(\alpha)) - c}{T_0'(q_0(\alpha))}\left[\frac{T_0''(q_0(\alpha))}{T_0'(q_0(\alpha))} + \frac{1-\beta}{q_0(\alpha)}\right]^{-1}. \tag{64}$$

As c, $T_0'(\cdot)$, $T_0''(\cdot)$, $q_0(\cdot)$ are known, $\Xi_0(\alpha)$ is identified up to β. As β is identified from observations of prices, quantities and marginal cost, then $\Xi_0(\alpha)$ is identified.

Then, $\Gamma_0(\alpha)$ is identified from the solution to the differential equation

$$\gamma_0(\alpha) + \Gamma_0(\alpha)\left[\frac{T_0''(q_0(\alpha))}{T_0'(q_0(\alpha))} + \frac{1-\beta}{q_0(\alpha)}\right]^{-1} = \Xi_0(\alpha)\left[\frac{T_0''(q_0(\alpha))}{T_0'(q_0(\alpha))} + \frac{1-\beta}{q_0(\alpha)}\right]^{-1}, \tag{65}$$

using the boundary condition $\Gamma_0(1) = 1$, and the fact that $T_0''(\cdot)$, $T_0'(\cdot)$, $q_0(\cdot)$, and β are known or identified.

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

Start recursively from $\tau = 1$. With knowledge of $\Gamma_s(\cdot)$ for $s < \tau$ and β, note that

$$\Xi_\tau(\alpha) = \alpha + \sum_{s=0}^{\tau-1}(1 - \Gamma_s(\alpha)) + \frac{T_\tau'(q_\tau(\alpha)) - c}{T_\tau'(q_\tau(\alpha))}\left[\frac{T_\tau''(q_\tau(\alpha))}{T_\tau'(q_\tau(\alpha))} + \frac{1-\beta}{q_\tau(\alpha)}\right]^{-1}. \tag{66}$$

is identified as $\Gamma_s(\cdot)$, c, $T_\tau'(\cdot)$, $T_\tau''(\cdot)$, $q_\tau(\cdot)$, and β are known or identified.

Then, $\Gamma_\tau(\alpha)$ is identified from the solution to the differential equation

$$\gamma_\tau(\alpha) + \Gamma_\tau(\alpha)\left[\frac{T''_\tau(q_\tau(\alpha))}{T'_\tau(q_\tau(\alpha))} + \frac{1-\beta}{q_\tau(\alpha)}\right]^{-1} = \Xi_\tau(\alpha)\left[\frac{T''_\tau(q_\tau(\alpha))}{T'_\tau(q_\tau(\alpha))} + \frac{1-\beta}{q_\tau(\alpha)}\right]^{-1}, \tag{67}$$

using the boundary condition $\Gamma_\tau(1) = 1$, and the fact that $T''_\tau(\cdot)$, $T'_\tau(\cdot)$, $q_\tau(\cdot)$, and β are known or identified.

J A Learning Model

In this section, I adapt the learning about reliability model common in the literature for a nonlinear price discrimination framework.

As in the limited enforcement models, buyers have heterogeneous tastes for the sellers product, captured in the parameter θ, distributed with cdf $F_\tau(\theta)$ and pdf $f_\tau(\theta)$. The base marginal return for the buyer of consuming q units is given by $v(q)$. At the same time, however, buyers can be of two additional types: reliable and unreliable. These types are orthogonal to the preference types θ. Reliable buyers pay their trade credit debts with probability 1 while unreliable buyers pay their debts with probability ϕ. The share of reliable buyers is χ. I assume the seller stops trading with the buyers whenever the buyer defaults on their debt. Through this screening mechanism, the belief at time τ that the seller has on the buyer being the reliable type is given by:

$$\hat{\chi}_\tau = \frac{\chi}{\chi + (1-\chi)\phi^\tau}. \tag{68}$$

Given this belief, at any time τ, the expected probability of payment is given by:

$$\psi_t = \hat{\chi}_\tau + (1-\hat{\chi}_\tau)\phi. \tag{69}$$

As in the limited enforcement model, the seller chooses transfers and quantities $\{T_\tau(q_\tau(\theta)), q_\tau(\theta)\}_{\tau=0}^{\infty}$ that maximize their expected profits subject to incentive-compatibility with respect to private types θ. I relax the problem by using the first-order approach and incorporate the incentive-compatibility constraints using their local equivalents.

The respective first-order conditions for the seller with marginal cost c and for the buyer are:

$$\theta v'(q_\tau(\theta)) - \frac{c}{\psi_t} = \frac{1-F_\tau(\theta)}{f_\tau(\theta)} v'(q_\tau(\theta)) \tag{70}$$

$$T'_\tau(q_\tau(\theta)) = \theta v'(q_\tau(\theta)). \tag{71}$$

Combining both equations yield the identification equation:

$$\frac{T'_\tau(q_\tau(\theta)) - c/\psi_\tau}{T'_\tau(q_\tau(\theta))} = \frac{1 - F_\tau(\theta)}{\theta f_\tau(\theta)}. \tag{72}$$

Nonparametric identification of the parameters follows from Luo et al. (2018).

As in the limited enforcement model, $T'(\cdot)$ is identified and estimated nonparametrically from data on prices and quantities alone. In my estimation exercise I approximate, for each seller and tenure, the transfer function using the following regression:

$$ln(T_{i\tau}) = \beta_0^\tau + \beta_1^\tau ln(q_{i\tau}) + \varepsilon_{i\tau}, \tag{73}$$

where T_{it} is the observed transfer for buyer i of tenure τ, $q_{i\tau}$ is the observed quantity. Using this equation, $\widehat{T}'_\tau(q) = \hat{\beta}_1^\tau T_\tau(q)/q = \hat{\beta}_1^\tau p(q)$. Reordering the identification equation and noting that $F_\tau(\theta) = G_\tau(q(\theta))$, where $G(\cdot)$ is the empirical distribution of quantities, we obtain the following estimation equation:

$$\frac{c}{\hat{\beta}_1^\tau p(q_{i\tau})} = \psi_\tau + \psi_\tau(1 - G_\tau(q_{i\tau}))(\kappa_0^\tau + \kappa_1^\tau q_{i\tau} + \kappa_2^\tau q_{i\tau}^2) + \tilde{\varepsilon}_{it}, \tag{74}$$

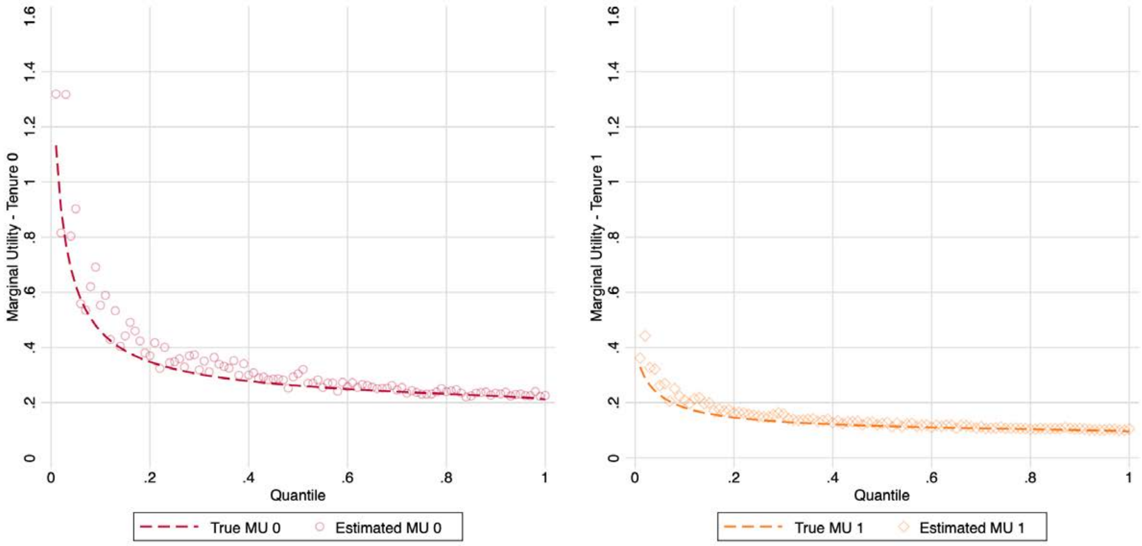
where $1/\theta f(\theta)$ is approximated using $\kappa_0 + \kappa_1 q_{i\tau} + \kappa_2 q_{i\tau}^2$. Thus, expected delivery probability ψ_τ can be estimated using information on quantities, transfers, and an estimate of marginal (or average) cost. To guarantee that ψ_τ is between zero and one, I parametrize it as a constant logit function.

J.1 Monte Carlo

To check that my estimator works well, I conduct a Monte Carlo experiment. I assume that at tenure 0, θ is drawn from a Weibull with minimum value 1.5, scale parameter 1 and shape parameter 2. To qualitatively match the fact that low types are less likely to survive, I assume that in tenure 1, θ is drawn from a distribution with minimum value 1.75, with the same parameters as in tenure 0. I assume marginal cost is 0.05 at both tenures but that delivery probability is 0.25 in tenure 0 and 0.55 in tenure 1. The return function is $v(q) = q^2$. I construct tariffs in each tenure using $t_\tau(\theta) = \theta q_\tau(\theta)^2 - \int^\theta q_\tau(x)^2 dx$. Moreover, I add Gaussian noise to the observed tariffs.

The Monte Carlo works well. Figure J.11 shows the estimated marginal return, i.e. the derivative of the tariff function, tracing the model marginal return well in both tenures. Figure J.12 shows that the estimated types at each quantile of the distribution matches well the model type for both tenures. Lastly, Figure J.13 shows the fit of estimating equation, with the left-hand side of equation 74 on the X-axis and the predicted values using the estimated parameters of the right-hand side of equation 74 on the Y-axis. Moreover, the estimated delivery probabilities

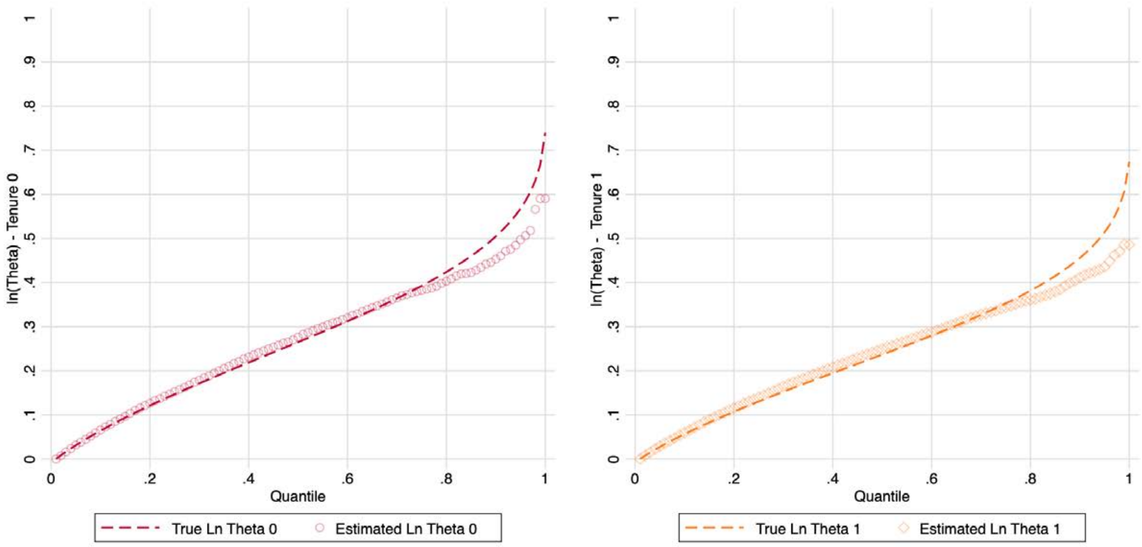
Figure J.11: **Estimates Marginal Return - Learning Model (Monte Carlo)**



Notes: This figure presents the results of the Monte Carlo for estimated base marginal return and true marginal return by quantile of quantity for tenure 0 (left) and tenure 1 (right) in the learning model.

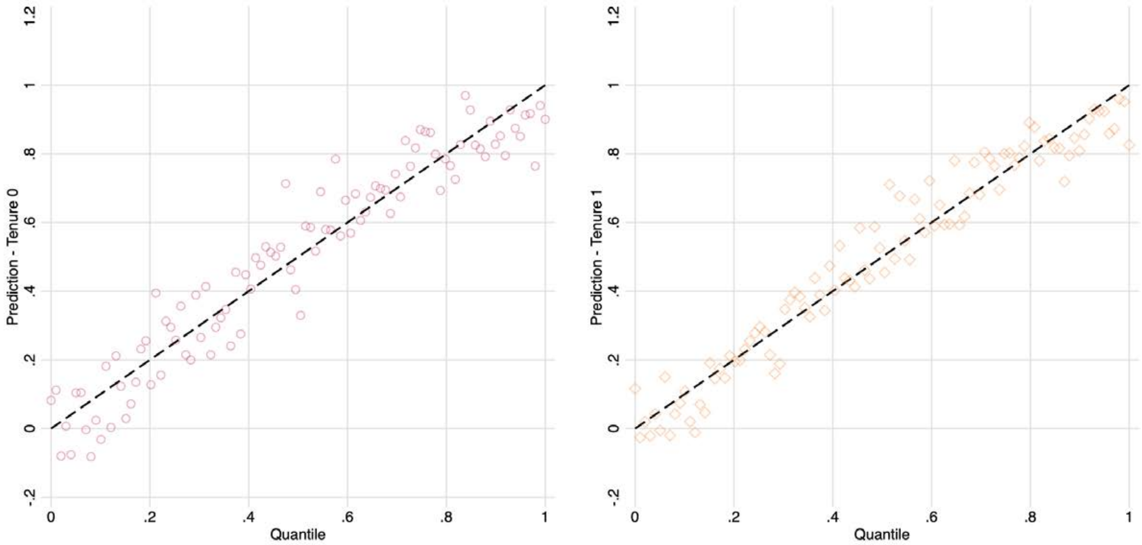
are 0.23 for tenure 0 and 0.51 for tenure 1.

Figure J.12: **Estimates Types - Learning Model (Monte Carlo)**



Notes: This figure presents the results of the Monte Carlo for estimated types and true types by quantile of quantity for tenure 0 (left) and tenure 1 (right) in the learning model.

Figure J.13: **Model Fit - Learning Model (Monte Carlo)**



Notes: This figure presents the statistical fit of the Monte Carlo by quantile of quantity for tenure 0 (left) and tenure 1 (right) in the learning model. The diagonal dashed line represens perfect fit.

J.2 Estimation Results

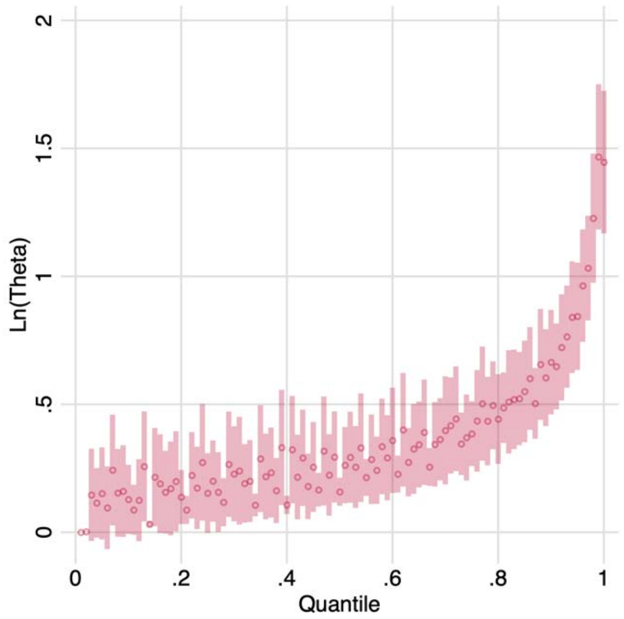
Next, I estimate the learning model for each seller. As in the limited enforcement model, I call new buyers tenure 0, buyers aged 1-3 tenure 1, and buyers aged 4+ tenure 2. I estimate equation 74 separately for each seller-tenure combination.

Figure J.14 shows the estimated values for types θ for each quantile. As required by incentive-compatibility, higher quantiles of quantities also have higher average θ. Figure J.15 plots the average log base marginal returns for each quantity quantile and tenure. As required by the model, log base marginal return decreases as quantity increases.

Figure J.16 plots the distribution of estimated delivery rates across sellers and tenures. The estimated distributions have a large support, starting at 20 percent up to 100 percent. The average for all tenures is around 80 percent. Although this is delivery rate is high, it requires that unreliable buyers default at least 20 percent of the time, and for the econometrician to observed reported default rates of 20 percent. As mentioned earlier, observed default rates are less than 1 percent.

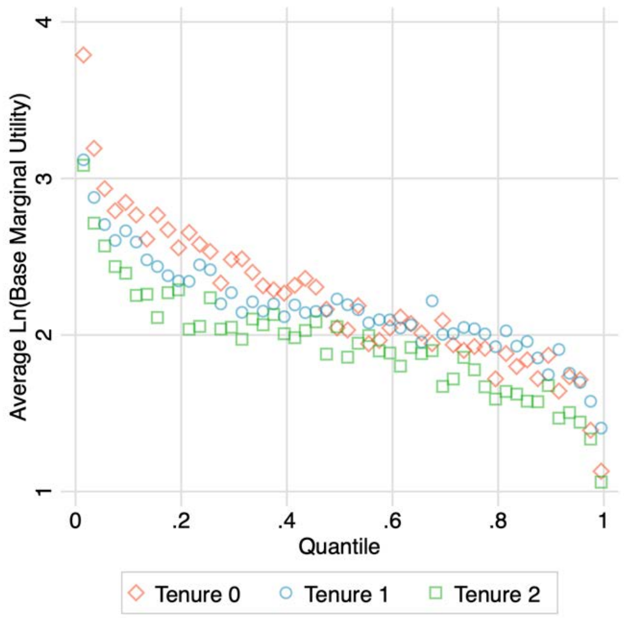
It is worth highlighting that expected delivery rates do not increase over tenures. In fact, as shown in Figure J.17, average change in estimated delivery probability within a seller is zero. This finding goes against the requirement of the learning model that expected delivery rates increase over time.

Figure J.14: **Average $Ln(\theta)$ - Learning Model**



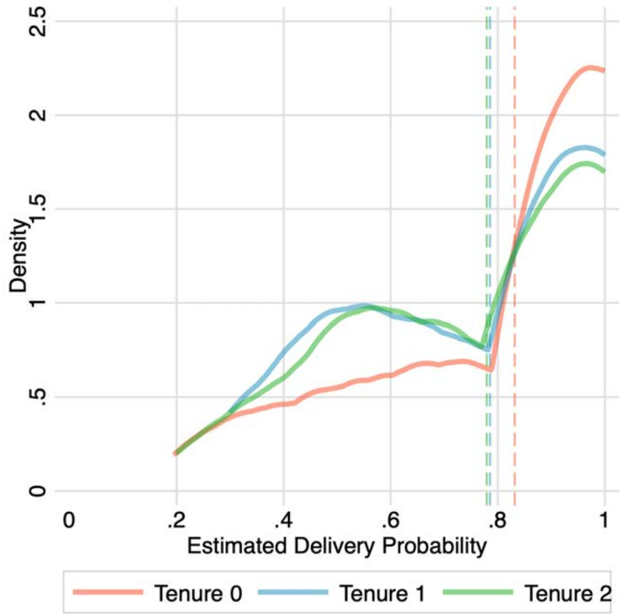
Notes: This figure presents the estimated average log types, across sellers, by quantile of quantity. Error bars represent $\pm$ 1.96 standard errors.

Figure J.15: **Average Base Marginal Return - Learning Model**



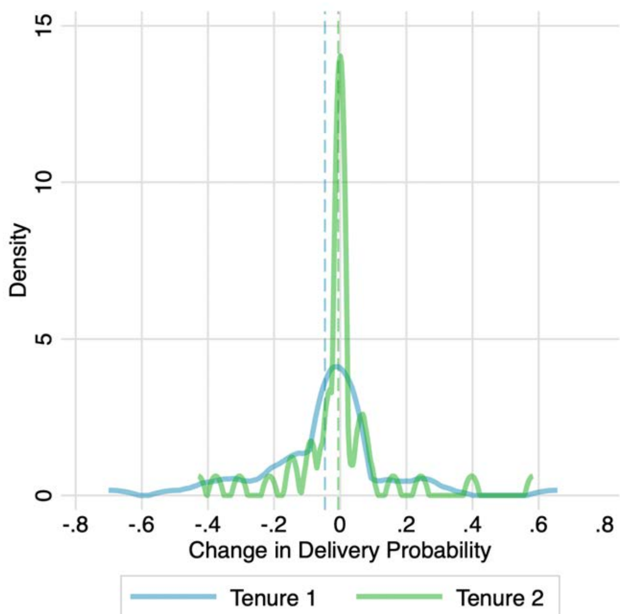
Notes: This figure presents average estimated (log) base marginal return, across sellers, by quantile of quantity.

Figure J.16: **Distribution of Delivery Rates ψ_τ - Learning Model**



Notes: This figure presents the distribution of estimated delivery rates, across-sellers, for different tenures.

Figure J.17: **Change Delivery Rates $\Delta\psi_\tau$ - Learning Model**



Notes: This figure presents the distribution of the changes in estimated delivery rates across-sellers.

J.3 Learning model vs. Limited enforcement model

I then verify the statistical fit of the learning model and the limited enforcement model. To compare both models, I rearrange the estimating equations in the following way. For the learning model:

$$G_\tau(q_{i\tau}) = 1 - \frac{1 - \frac{c}{\hat{\psi}_\tau \hat{\beta}_1^\tau p(q_{i\tau})}}{(\widehat{\kappa}_0^\tau + \widehat{\kappa}_1^\tau q_{i\tau} + \widehat{\kappa}_2^\tau q_{i\tau}^2)} \tag{75}$$

For the limited enforcement model:

$$G_\tau(q_{i\tau}) = \widehat{\Gamma}_{i\tau} - \sum_{s=0}^{\tau-1}(1 - \widehat{\Gamma}_{i\tau}) - \frac{1 - \frac{c}{\hat{\psi}_\tau \hat{\beta}_1^\tau p(q_{i\tau})} - \widehat{\gamma}_{i\tau}}{\widehat{\kappa}_0^{LE,\tau} + \widehat{\kappa}_1^{LE,\tau} q_{i\tau} + \widehat{\kappa}_2^{LE,\tau} q_{i\tau}^2}. \tag{76}$$

Figure J.18 shows the model fit for both models at different tenures. In all cases, the limited enforcement model fits better the data than the learning model. At all tenures, the prediction of the quantile trances well the diagonal line in the limited enforcement model. The learning model is noisier, which a greater share of their predictions falling far from the diagonal. As suggested by the graphical evidence, the Vuong test for model selection of non-nested models picks the limited enforcement model at all tenures at the 1% significance level.

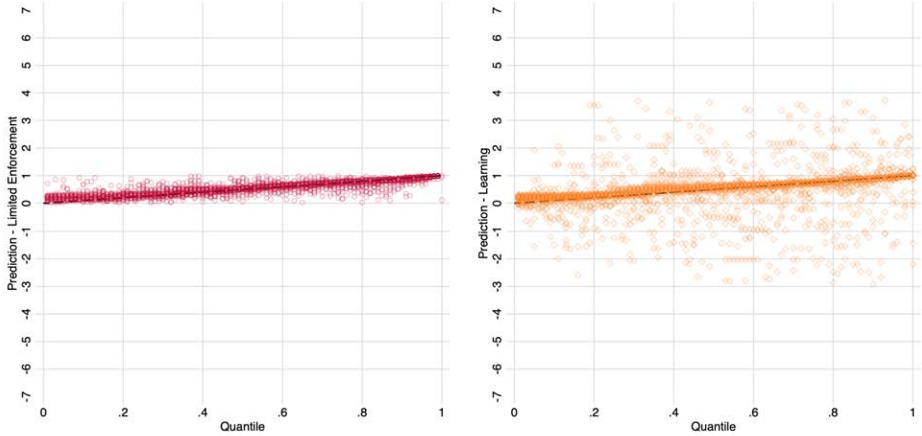
I conduct a last test of model selection by testing a non-targeted moment: the price decrease over time, which has not been used to estimate any parameter in the models. Using the estimated parameters, I obtain quantities, transfers, and unit prices for each estimated θ in both models at different tenures. As the types θ are normalized to obtain $\underline{\theta} = 1$ and transfers are defined by $\theta v(q(\theta) - \int^\theta v(x)dx$, the price levels will differ by model. To correct for that and make models comparable, I residualize the log unit prices and remove seller-model fixed effects. Figure L.24 plots the dynamics of discounts over time for the data, the limited enforcement model, the learning model, and a limited enforcement model with no memory. This last model uses the estimated parameters and multiplier constraints to obtain the path of quantities and prices but forces all past multipliers to be equal to 1. That is, I remove the memory of past constraints when defining quantities and prices.

Importantly, the limited enforcement model fits well the discounts over time observed in the data, while the learning model fails to replicate the dynamics. An explanation for this might be that the model estimates no change in the delivery probability in the data.[40] Interestingly, the limited enforcement model with no memory, where the past constraints are lifted, also fails to capture the dynamics observed in the data.

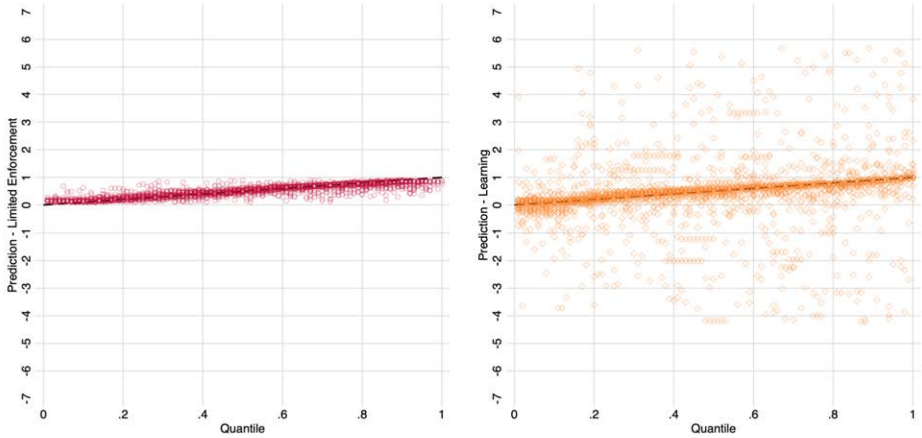
[40] Although not discussed here, a model without any friction other than adverse selection (the standard nonlinear pricing model) also fails to replicate these dynamics.

Figure J.18: **Model Fit: Limited Enforcement (Left) vs. Learning (Right)**

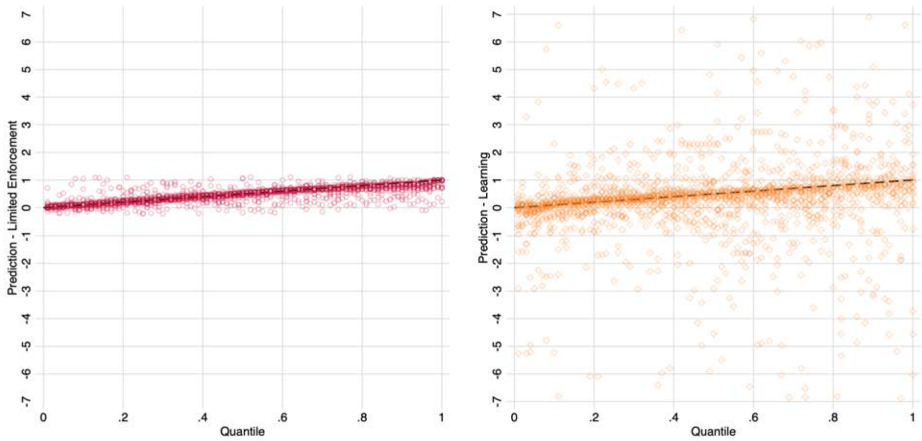
(a) Tenure 0



(b) Tenure 1



(c) Tenure 2



Notes: These figures compare the statistical model fit of the limited enforcement model (left) and the learning model (right) for the different tenures. Model fit comparisons use equations 76 and 75. Diagonal line represents perfect statistical fit.

J.4 Implied Default Rates

Lastly, I test if calibrated default rates that match price discounts also match aggregate reported default rates. In Ecuador, a firm reports the value of uncollectable receivables (bad-debt expenses) as part of their financial declarations, and these are defined as trade credit debt that are either: i) at least 3 years old, ii) from a officially bankrupt buyer, or iii) from a dissolved corporation. Although these reports are not a perfect measure of default, they proxy for defaults in a similar manner as account receivables proxy for trade credit in past literature (e.g. in Petersen and Rajan (1997)). I defined default rate as the rate of uncollectable receivables to sales, both yearly or over all available periods for each firm in the data. Table J.9 shows that under any definition reported default rates are low: between 0.41 percent and 0.86 percent for all firms not in my sample and 0.54 to 0.66 percent for firms in my sample.[41]

Under a learning model similar to those referenced above with CES demand functions, the unit price for a relationship that has lasted k years is given by:

$$p^l_k = \frac{\sigma c}{\sigma - 1}\frac{1}{\chi_k},$$

where χ_k is the expected probability of payment after k interactions, c the marginal cost, and σ is the elasticity of substitution. Commonly, χ_k is defined by

$$\chi_k = \frac{\lambda}{\lambda + (1-\lambda)P^k}(1-P) + P, \tag{77}$$

where λ is the share of reliable buyers in the population and P is the probability that the unreliable buyer will successfully pay their debt.

This model is not able to match the observed price decrease of 1.5 percent per year at the observed default rates in the range of $[0.4\%, 0.9\%]$. I conducted a simple calibration exercise to obtain the price decrease by year for a combination of different parameters of default rates (1-P) between $[0.005, 0.25]$ and share of reliable buyers λ between $[0.01, 0.99]$. Figure J.19 below shows the price changes generated by the model for all the combinations of parameters. In order to match the 1.5 percent price change, the observed default rate should be of at least 20 percent. That is, at least 20 to 40 times larger than the default rates observed in the data.

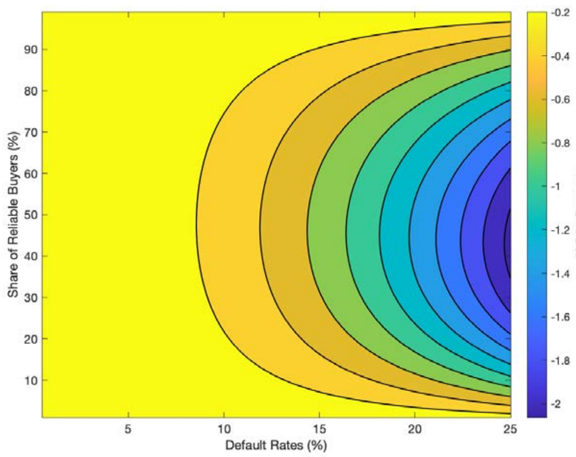
[41] Actual default rates on credit to firms reported by banks are around 1 percent in the years in my sample.

Table J.9: **Reported Default Rates in Financial Statements**

	Not in Sample	In Sample
Yearly	0.41%	0.54%
All 2007-2017	0.86%	0.66%
Num. of Firms	116,470	103

Notes: This table presents reported default rates as measured by *uncollectable receivables*. In Ecuador, a firm reports the value of uncollectable receivables (bad-debt expenses) as part of their financial declarations, and these are defined as trade credit debt that are either: i) at least 3 years old, ii) from a officially bankrupt buyer, or iii) from a dissolved corporation. *Not in Sample* refers to all formal firms with available data and positive value of receivables at any point in time. *In Sample* refers to sellers in this paper. Yearly reports the ratio of *uncollectable receivables* over total sales. All 2007-2017 reports the sum of all *uncollectable receivables* in that period over the sum of all sales in the same period.

Figure J.19: **Simulation of Price Changes in Learning Model**



Notes: This figure reports the level of price discounts over time consistent with different combinations of default rates and share of reliable buyers. Observed price discounts in data are around -1.5%.

K Additional Estimation Results

K.1 Distribution of t-Statistics against Standard Model Null

Table K.1 show the distribution of t-statistics for tests against a standard model null.

Table K.10: **Distribution of t-Statistics**

	p10	p25	p50	p75	p90
t-Statistics	4.34	8.37	15.31	29.94	131.77
Observations	33				

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

K.2 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 modelling assumptions $v'(\cdot) > 0$ and $v''(\cdot) < 0$. To estimate parameters, I consider tenure 0 transactions between buyer i and seller j at year t and perform the following uniform least squares regression:

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

using $v'(q) = k\beta q^{\beta-1}$, the estimated base marginal returns $\widehat{v'}_{ijt}$ and under the assumption that ε_{ijt} is Gaussian error. Table K.11 present the distribution of k and β.

Table K.11: **Parameters of Return Function**

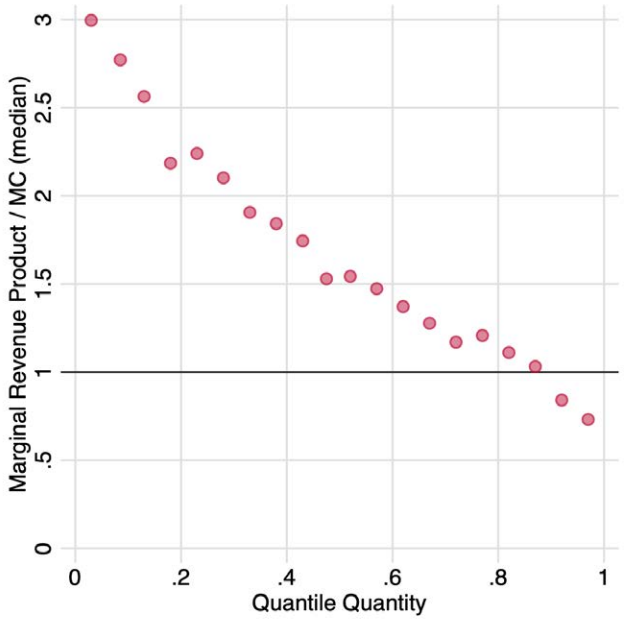
	mean	p10	p25	p50	p75	p90
β	0.67	0.41	0.53	0.71	0.81	0.89
k	89.39	14.40	19.77	42.73	60.57	169.30
Observations	33					

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

K.3 Economic Magnitudes: Base Marginal Return

Figure K.20 presents a binscatter of the ratio marginal revenue product (base marginal return) over marginal costs aginst 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 K.20: **Base Marginal Return over Marginal Costs**



Notes: This figure plots the median of the ratio of base marginal return by marginal costs across sellers by quantile of quantity of tenure 0.

K.4 Model Fit

Figure K.21 presents the statistical fit of the model across tenures. It plots a reordered equation 13'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.[42] Fit generally worsens for higher tenures; the results from Monte Carlo studies in Appendix M 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.

Figure K.22 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 Q-CES.

Figure K.23 shows the fit of tariffs. To generate tariffs in the model, I use the empirical equivalent of equation t-RULE.

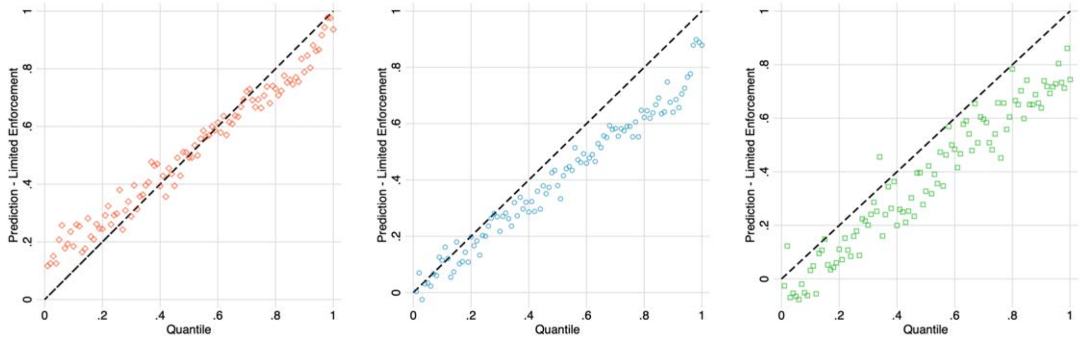
[42]Reorder equation 13 to obtain:

$$\alpha = \Gamma_\tau(\alpha) - \sum_{s=0}^{\tau-1}(1 - \Gamma_s(\alpha)) - \left[\frac{T'_\tau(q_\tau(\alpha)) - c_\tau}{T'_\tau(q_\tau(\alpha))} - \gamma_\tau(\alpha)\right]\frac{\theta_\tau(\alpha)}{\theta'_\tau(\alpha)},$$

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

Figure K.21: **Model Fit - Statistical**

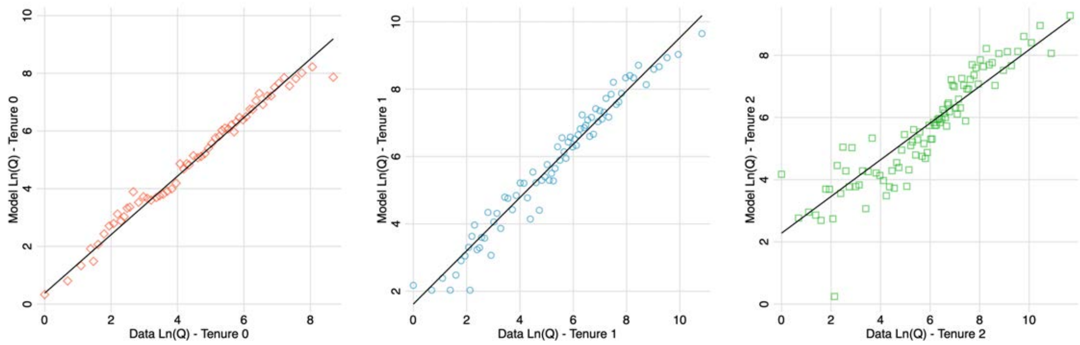
(a) Tenure 0 (b) Tenure 1 (c) Tenure 2



Notes: These figures show binscatters of statistical fit of the model across tenures. The X-axis plots the quantile of quantity and in the Y-axis it plots the predicted quantile using the estimated parameters of the limited enforcement model.

Figure K.22: **Model Fit - Quantities**

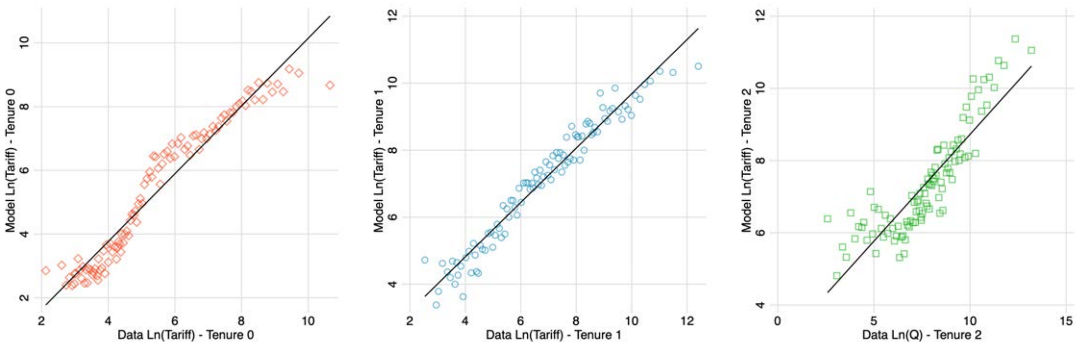
(a) Tenure 0 (b) Tenure 1 (c) Tenure 2



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 K.11.

Figure K.23: **Model Fit - Tariffs**

(a) Tenure 0 (b) Tenure 1 (c) Tenure 2



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

K.5 Sensitivity Analysis - Surplus

As distortions relative to first-best are type-dependent, overall efficiency at the seller level might differ. In panel b, I obtain seller-level weighted averages (weighted by quantity at first-best levels). Aggregate efficiency dynamics match closely those at the pair-level, with efficiency increasing over time. One difference is worth highlighting. When aggregating by quantity, the inefficiency in trade at tenure 0 decreases. Surplus relative to first-best moves from 68% at the pair-level to 79% at the seller-level. This difference indicates that most of the distortions are concentrated in buyers that purchase relatively little quantity. Importantly, however, total surplus *is* distorted downwards in new relationships.

Table K.12: **Model Efficiency: Surplus relative to pair-specific first-best**

	Tenure		
	0	1	2
	78.89	91.29	96.81
(s.d.)	(22.99)	(30.52)	(20.60)

Notes: This table shows average efficiency (relative to pair-wise first-best) at the seller-level, after calculating the seller-specific weighted mean efficiency, which uses share of total quantity as weight. Standard deviation of weighted average efficiency are reported in parenthesis.

K.6 Additional Counterfactual Results

This subsection presents comparisons of different counterfactual models relative to baseline nonlinear pricing regime with limited enforcement. The tables 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. Table K.13 shows the results for full nonlinear pricing and full enforcement. Table K.14 shows the results for optimal uniform pricing with limited enforcement. Table K.15 shows results for optimal uniform pricing and perfect enforcement.

Table K.13: **Nonlinear Price with Perfect Enforcement**

	10%	25%	50%	75%	100%
	Buyer's Net Return				
Tenure 0	50.8	66.9	79.8	59.2	32.8
Tenure 1	57.1	69.4	60.2	44.9	36.2
Tenure 2	57.1	71.1	61.3	48.3	52.6
	Seller's Profit				
Tenure 0	39.7	33.1	22.0	45.1	70.3
Tenure 1	50.8	36.4	42.1	58.0	64.0
Tenure 2	47.6	31.4	39.0	51.7	47.4
	Unit Price				
Tenure 0	50.8	33.1	18.8	40.3	70.5
Tenure 1	42.9	30.6	43.5	60.0	64.2
Tenure 2	46.0	29.8	38.0	51.5	47.4

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 optimal uniform monopoly pricing with limited enforcement of contracts. The figures are reported across different tenures and percentiles in the distribution of types.

Table K.14: **Uniform Price with Limited Enforcement**

	10%	25%	50%	75%	100%
	Buyer's Net Return				
Tenure 0	70.8	68.2	64.1	58.3	59.2
Tenure 1	68.9	66.1	58.7	57.8	58.5
Tenure 2	66.5	64.4	56.2	57.1	53.1
	Seller's Profit				
Tenure 0	95.0	97.9	97.8	90.3	88.0
Tenure 1	96.9	91.2	85.6	79.6	89.3
Tenure 2	90.7	78.7	81.7	89.3	97.6
	Unit Price				
Tenure 0	100.0	99.6	88.9	72.7	28.6
Tenure 1	95.7	88.3	78.2	56.4	28.8
Tenure 2	88.8	78.2	80.0	72.5	34.3
	% Excluded				
Tenure 0	99.1	79.9	67.7	57.0	57.0
Tenure 1	98.2	75.5	63.5	57.0	56.3
Tenure 2	95.5	75.5	62.4	57.0	52.3

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 optimal uniform monopoly pricing with limited enforcement of contracts. The figures are reported across different tenures and percentiles in the distribution of types. In the counterfactual example, buyers are excluded if they satisfying the enforcement constraint is unfeasible. For excluded buyers, I assign them $q = 0$ and missing price.

Table K.15: **Uniform Price with Perfect Enforcement**

	10%	25%	50%	75%	100%
			Buyer's Net Return		
Tenure 0	1.2	1.3	9.4	15.9	19.7
Tenure 1	4.3	5.0	8.4	19.4	17.5
Tenure 2	5.0	11.3	10.4	11.1	11.1
			Seller's Profit		
Tenure 0	49.1	59.0	56.7	58.8	72.1
Tenure 1	51.6	58.2	60.4	60.0	75.8
Tenure 2	59.0	54.8	61.6	69.0	88.9
			Unit Price		
Tenure 0	100.0	99.6	88.9	72.7	28.6
Tenure 1	95.7	88.3	78.2	56.4	28.8
Tenure 2	88.8	78.2	80.0	72.5	34.3

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 optimal uniform monopoly pricing with perfect enforcement of contracts. The figures are reported across different tenures and percentiles in the distribution of types. In the counterfactual example, buyers cannot default on their trade credit debts, and so, sellers do not exclude any buyer.

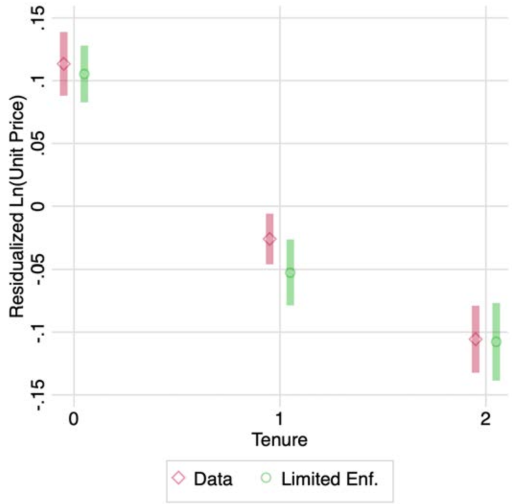
L Model Comparison: Non-targeted moment

This section compares model fit through the use of a non-targeted moment, namely, price discounts in tenure. I consider four models. First, I consider the limited enforcement model, as discussed in the main text. Second, I consider the standard nonlinear pricing model. I estimate the model using the same methodology as in the model with limited enforcement, but set $\Gamma_\tau(\cdot) = 1$ and $\gamma_\tau(\cdot) = 0$ for all τ. Third, I consider a learning about reliability model, where some share of buyers default with positive probability and the seller filters out unreliable buyers over time. The details for the model and its estimation procedure are offered in Section J. The estimated model attempts to estimate the rate of default. As reported default rates in financial statements of firms in my sample are remarkably low (less than 1%), the standard model would offer similar results as a learning model where the default rate is calibrated to match observed default rates. Lastly, I consider the primitives of the estimated limited enforcement model but erase all memory from past promises captured through $\Gamma_s(\theta)$ for $s < \tau$ by setting all of them equal to 1.

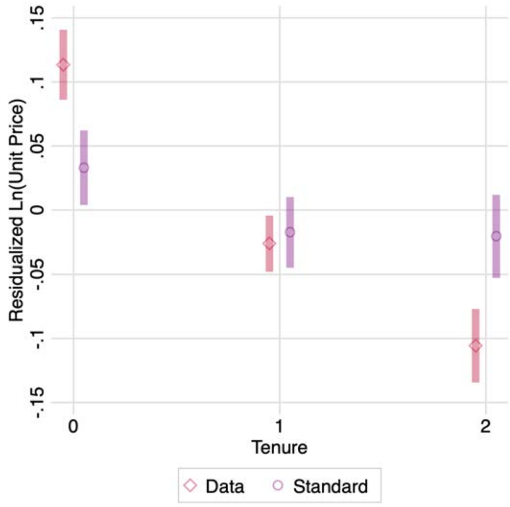
For each model, tariffs are generated within tenure using equation *t*-RULE, which relies on the estimated parameters for $v(\cdot)$, the distribution of θ, and the predicted values of $q_\tau(\theta)$. To correct for any differences in levels across models, I present log prices, residualized at the model-seller-year level.

Figure L.24 presents the results. Subfigure (a) shows that the limited enforcement captures well the backloading of prices. Subfigure (b) shows that the standard nonlinear pricing model does capture some of the discounts, but the fit is not as good as the limited enforcement model. Subfigure (c) shows that the estimated learning model fails to replicate any discounting. Lastly, subfigure (d) shows that by eliminating the memory in the estimated limited enforcement model, the model would predict increasing prices in tenure.

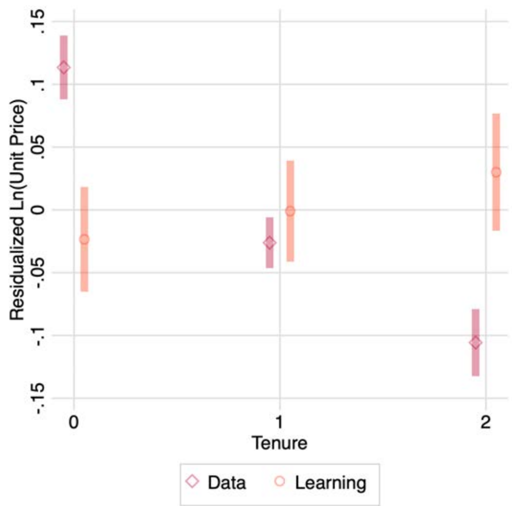
Figure L.24: **Model Comparison - Discounts**



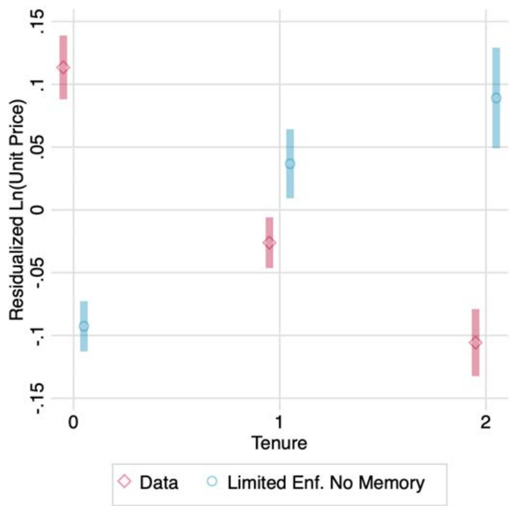
(a) Limited Enforcement



(b) Standard



(c) Learning



(d) Clearing Memory

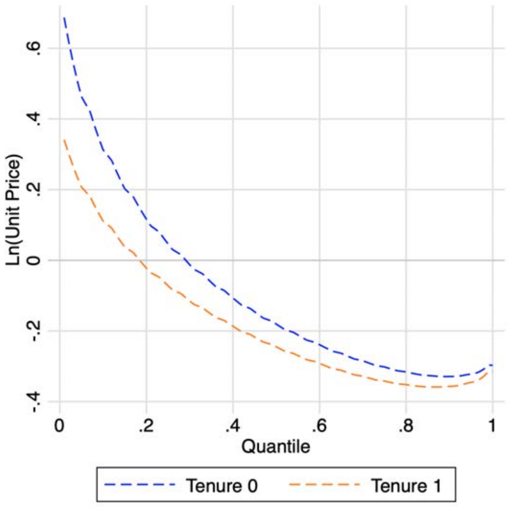
Notes: These figures compare the dynamics of prices across tenure as observed in the data against those generated in alternative models. Figure (a) shows results for the limited enforcement model, the main model in the paper. Figure (b) shows results for a standard nonlinear pricing model, which uses the same estimation methodology of the limited enforcement model but restricts $\Gamma_\tau(\theta) = 1$ for all τ and θ. Figure (c) shows an estimated learning about reliability model. Details for the model are presented in Appendix Section J. Figure (d) shows the performance of the estimated limited enforcement model but forces $\Gamma_s(\theta) = 1$ for all $s < \tau$. Error bars represent ± 1.96 standard errors. Unit of observation is seller-tenure-type.

M Monte Carlo Study

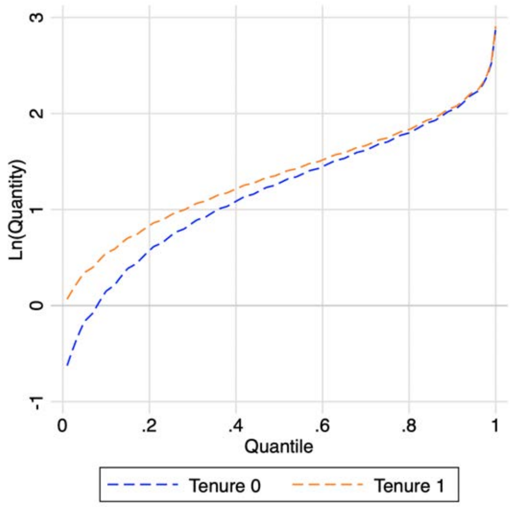
The Monte Carlo studies the behavior of my estimators for two periods of a dynamic contract without breakups. I use the following design. The return function is $v(\theta, q) = \theta q^{1/2}$. The type distribution is Weibull with scale parameter equal to 1 and shape parameter equal to 2, $F(\theta) =$

Figure M.25: **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.

$1 - exp(-(\theta - 1)^k)$, normalized so $\underline{\theta} = 1$.[43] Marginal cost is 0.45. Although 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.[44]. 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.

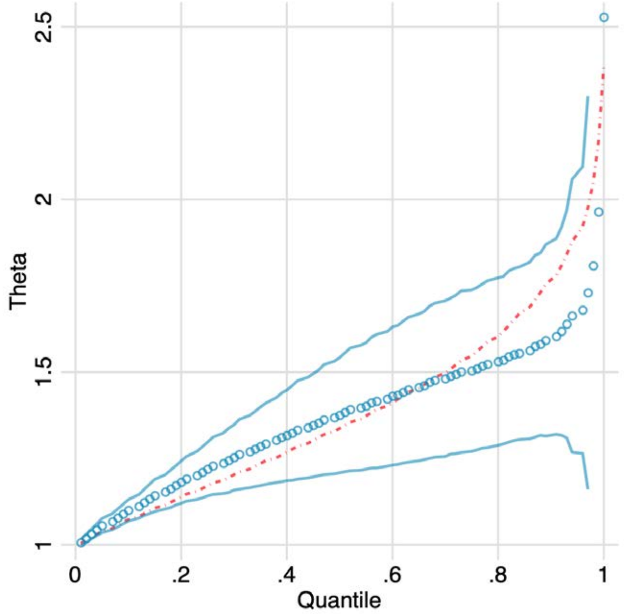
Figure M.25 shows the (log) average prices and average quantities generated by the model for the two types of tenure. The model delivers quantity discounts (decreasing unit prices in θ), strict mononoticity of quantity (increasing quantities in θ), and backloading in the dynamic model, namely, further discounts and larger quantities in tenure 1 for each θ.

Figure M.26 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,

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

[44]Section H provides more details about how to find the Gamma function as the solution of the differential equation

Figure M.26: **Quantile Types $\hat{\theta}$**



Notes: This figure plots the true (red) and estimated distribution (in blue) by quantile of quantity, with the error margins indicating ± 1.96 variation around estimated mean from 300 simulations.

the figure shows a good fit, with most sections of including the true θ within their dispersion.

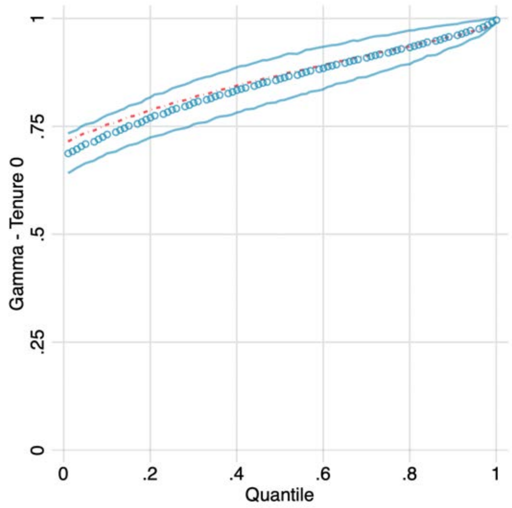
Figure M.27 shows the results of the estimated Gamma distribution and the base marginal return, again in blue the estimated results and in red the true values. Both cases indicate good fit.

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. Figure M.28 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.

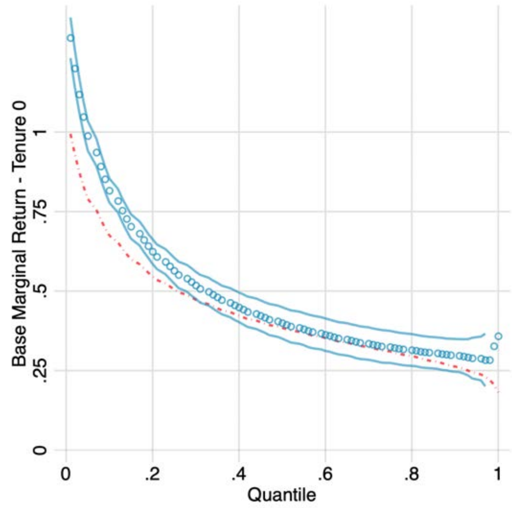
With respect to the differences between true and estimated functions, I find that the slight upward bias in the Gamma function for tenure 1 disappears if I use the true $\Gamma_0(\theta)$ function instead of the estimated $\widehat{\Gamma}_0$, suggesting that the bias is generated by sampling error in 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, observa-

Figure M.27: **Monte Carlo Results for Tenure 0**

(a) Gamma Distribution



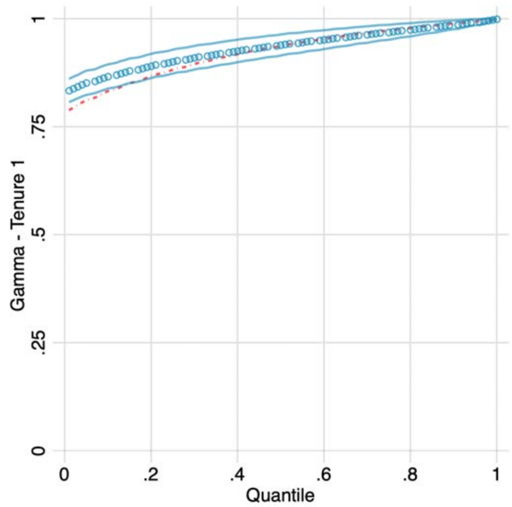
(b) Base Marginal Return



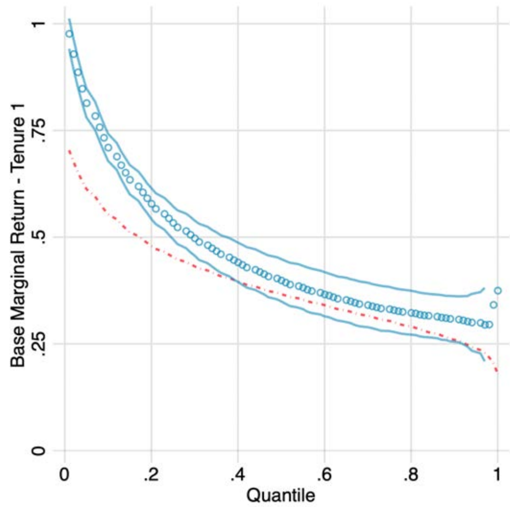
Notes: Panel (a) plots the true (red) and estimated value (blue) of the LE multiplier for tenure 0 by quantile of quantity, with error margins indicating ± 1.96 variation around the estimated mean. Panel (b) plots the true (red) and estimated value (blue) of the base marginal return for tenure 0 by quantile of quantity, with error margins indicating ± 1.96 variation around the estimated mean from 300 simulations.

Figure M.28: **Monte Carlo Results for Tenure 1**

(a) Gamma Distribution



(b) Base Marginal Return



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.

tions 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).