Online Appendix

OA1 Summary Statistics

Table OA1: **Summary Statistics - Sellers and Buyers in 2016**

	Sellers			***Buyers***		
	Mean	Median	SD	Mean	Median	SD
Total Sales (million USD)	14.95	8.26	24.33	2.35	0.20	24.33
Total Inputs (millon USD)	10.58	5.31	18.94	1.92	0.15	24.13
Accounting Markup	1.21	1.20	0.20	1.21	1.10	0.61
Age	30.47	29.00	19.16	15.18	14.00	9.75
Import Share (%)	24.47	21.38	22.96	3.82	0.00	13.49
Export Share (%)	5.81	0.00	19.11	1.06	0.00	8.87
Observations	49			28,138		

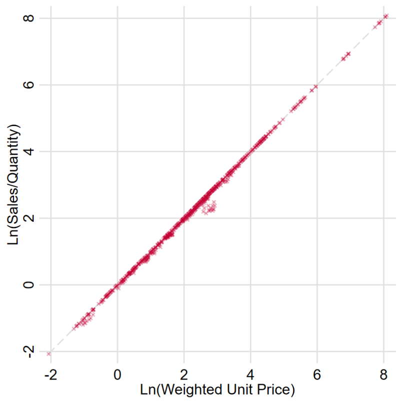
Notes: This table reports summary statistics about the size, age, and trade exposure of buyers and sellers in the sample for the year 2016. Monetary values are in U.S. dollars for 2016.

Table OA2: **Summary Statistics - Electronic Invoice Database**

	Mean	Median	SD
N. Buyers	8,028.41	613.50	25,078.11
N. Buyers (> USD250)	801.86	281.50	1,269.02
N. Buyers (> USD4,800)	160.49	73.00	265.19
Total Sales (million USD)	16.58	7.23	29.44
Total Q (million)	5.42	1.20	9.01
Q per Buyer	12,455.39	1,495.22	25,823.40
Bill per Buyer (USD)	43,490.37	9,067.65	105840.28
Bill per Buyer (USD) (> USD250)	61,756.13	14,898.73	143297.12
Bill per Buyer (USD) (> USD4,800)	176,632.43	46,899.44	444389.15
Observations	49		

Notes: This table reports summary statistics of the electronic invoice database. N. buyers refers to the number of unique buyers each seller in the sample has on average over 2016 and 2017. Quantity is the sum of all quantities across products. Bill per buyer is the total value of the transactions between buyer and seller. The thresholds > 250 and $> 4,800$ correspond to the data selection thresholds in Bernard et al. (2019) and Alfaro-Urena et al. (2022), respectively.

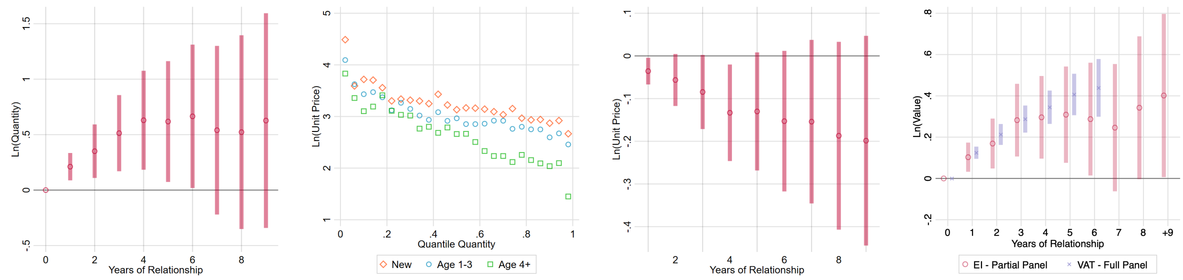
Figure OA1: **Average Price vs Weighted Price**



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.

OA2 Motivating Evidence - Robustness

Figure OA2: **Motivating Facts - Robustness**



(a) Quantity over Time (b) Quantity Discounts (c) Prices over Time (d) Prices over Time

Notes: Panel a) plots the coefficients of log total quantity on relationship age dummies controlling for pair fixed effects. Total quantity is obtained by aggregating all the reported units of sold goods. Standard errors are clustered at the pair-level. Panel b) shows the relationship between quantity purchased and average log unit price (right-panel) through binscatters of the measure of unit price against quantile of quantity by age of relationship. Quantiles of quantity are calculated for each seller-relationship age combination. Panel c) plots the regression coefficients of log unit prices on years of relationship, controlling for pair fixed effects. Standard errors are clustered at the pair-level. Panel d) plots regression coefficients for the 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 OA3: **Benchmark: Quantity Discounts**

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

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

Table OA4: **Robustness - Standardized Log Price**

VARIABLES	(1) Stdz. ln(Price)	(13) Stdz. ln(Price)
Stdz. ln(Quantity)	-0.0463***	-0.0454***
	(0.00722)	(0.00631)
Age of Relationship	-0.00552***	-0.00395***
	(0.00146)	(0.00134)
ln(Assets Buyer)	0.00131***	0.000832***
	(0.000318)	(0.000230)
Supply Share	0.0262*	0.0143
	(0.0157)	(0.0145)
Demand Share	0.0119	0.00402
	(0.0486)	(0.0454)
Observations	73,633	73,626
R-squared	0.082	0.091
Controls	Yes	Yes
Year FE	Yes	Yes
Buyer Sector FE	No	Yes

Notes: This table presents regressions regressions of standardized unit prices on age of relationship, standardized quantity, and different buyer characteristics. Controls include Exporter, Importer, Business Group, Multinational Dummies, as well as Distance between HQs, Sales, Age, and Number of Employees of Buyer. Standard errors are clustered at the seller-year level. *** p<0.01, ** p<0.05, * p<0.1

Table OA5: **Robustness - Seller's Sector**

VARIABLES	(1) Stdz. ln(Price)	(2) ln(Price)
Textiles		
Age of Relationship	-0.00319*	-0.0973***
	(0.00180)	(0.0305)
Pharmaceuticals		
Age of Relationship	-0.00691***	-0.0308**
	(0.00185)	(0.0139)
Cements		
Age of Relationship	-0.00447***	-0.0387***
	(0.00141)	(0.00728)
Seller-Year FE	No	Yes
Controls	Yes	Yes
Observations	73,633	73,633
R-squared	0.083	0.608

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 as well as all variables in Table OA4. Standard errors are clustered at the seller-year level. *** p<0.01, ** p<0.05, * p<0.1

Table OA6: **Price Dynamics by Payment Method**

VARIABLES	(1) Stdz. ln(Price)	(2) Stdz. ln(Price)	(3) ln(Price)	(4) ln(Price)
Payment Method	Pay-in-advance	Trade-Credit	Pay-in-advance	Trade-Credit
Age of Relationship	0.000360	-0.00248***	0.0314	-0.0240***
	(0.00215)	(0.000258)	(0.0266)	(0.00516)
Quantity Control	Yes	Yes	Yes	Yes
Seller-Year FE	Yes	Yes	No	No
Pair FE	No	No	Yes	Yes
Observations	8,777	68,730	1,548	33,680
R-squared	0.148	0.134	0.947	0.937

Notes: This table presents a regression of unit prices on age of relationship, controlling for quantity, by payment modality. The sample excludes buyers that switch between modalities. Columns (1) and (2) uses standardized prices and controls for seller-year fixed effects, while Columns (3) and (4) relies instead on average unit prices while controlling for seller-buyer fixed effects. Standard errors are clustered at the seller-buyer level. *** p<0.01, ** p<0.05, * p<0.1

OA3 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. (2022) 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. (2022), 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. 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.

Online Appendix Table OA7 show the results of the tests, restricting the sample of sellers to firms that have at least one multinational buyer.[38] Columns (1) and (2) test whether multinationals from common and other legal origin, respectively, are subject to backloading. While other law multinational experience price discounts, common law multinationals do not. Column (3) confirms these results in a pooling regression. Lastly, Columns (4) and (5) include observations for all buyers and establishes again that common

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

[38] 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 sample, I observe 161 multinational buyer observations (there are only 673 multinationals active in the whole economy in 2016), 21 of which have HQ in common law countries and 140 in other countries.

Table OA7: **Price Dynamics and Legal Origin of Buyer**

VARIABLES	(1) Stdz. ln(Price)	(2) Stdz. ln(Price)	(3) Stdz. ln(Price)	(4) Stdz. ln(Price)	(5) Stdz. ln(Price)
Multinational Buyers Only	Yes	Yes	Yes	No	No
Legal Origin	Common	Other	All	All	All
Age of Relationship	0.0114*	-0.00627*	-0.00627*	-0.00210***	-0.00210***
	(0.00628)	(0.00328)	(0.00330)	(0.000235)	(0.000235)
1{Multinational Buyer}				0.0226*	
				(0.0133)	
1{Multinational Buyer} X Age of Relationship				-0.00137	
				(0.00315)	
1{Common Law}			-0.0500		-0.0221
			(0.0364)		(0.0330)
1{Common} X Age of Relationship			0.0177**		0.0135**
			(0.00689)		(0.00598)
1{Other Law}					0.0279*
					(0.0142)
1{Other Law} X Age of Relationship					-0.00417
					(0.00326)
Constant	-0.00962	0.0404***	0.0404***	0.0125***	0.0125***
	(0.0347)	(0.0143)	(0.0144)	(0.000175)	(0.000175)
Observations	21	140	161	741,897	741,897
R-squared	0.071	0.008	0.016	0.000	0.000

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

law multinationals do not experience price discounts over time.

OA4 Existence and Non-Stationarity

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.

OA4.1 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:

$$v_q(\theta, q(\theta)) - c = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)), \tag{36}$$

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

$$\int_{\underline{\theta}}^{\overline{\theta}} [u(\theta) - \bar{u}(\theta)]d\gamma(\theta) = 0. \tag{37}$$

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 36 and complementary slackless condition 37 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:

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

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

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

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

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

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

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

where $M(q)\equiv 1+\theta(q)\theta'(q)r(q)+\theta'(q)$ and $r(q)=g^{-1}(q)$ for $g(\theta)\equiv q''(\theta)$. As we expect $T''(q)<0$ and $T'(q)>0$, it is necessary that $M(q)/(M(q)-1)<0$. Such condition will be satisfied if $M(q) < 1$ and $M(q) > 0$, which imply that

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

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

Lastly, full participation requires H to hold as well as $s(\theta, \bar{q}(\theta)) \geq (1-\delta)\theta v(\bar{q}(\theta))$. 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 36 and 37. 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.

OA4.2 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.

OA5 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:

$$\delta[\sum_{s=1}^{\infty} \delta^{s-1} u_{k+s}(\theta)] - \overline{u}_k(\theta) = \sum_{s=1}^{\infty} \delta^s u_{k+s}(\theta) + \varepsilon \geq t_k(\theta) = \theta v(q_k(\theta)) - u_k(\theta). \tag{38}$$

The seller's problem in the Lagrangian-form is

$$W(\varepsilon) = \max_{\{q_\tau(\theta), u_\tau(\theta)\}} \sum_{\tau=0}^{\infty} \delta^\tau \Big\{ \int_{\underline{\theta}}^{\overline{\theta}} [\theta v(q_\tau(\theta)) - cq_\tau - u_\tau(\theta)] f(\theta) d\theta + \tag{39}$$

$$\int_{\underline{\theta}}^{\overline{\theta}} [\sum_{s=1}^{\infty} \delta^s u_{\tau+s} + \varepsilon * 1\{\tau = k\} - t_\tau(\theta)] d\Gamma_\tau(\theta) \Big\} \tag{40}$$

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:

$$\frac{dW(\varepsilon)}{d\varepsilon} = \delta^k \int_{\underline{\theta}}^{\overline{\theta}} d\Gamma_k(\theta), \tag{41}$$

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:

$$W(\varepsilon) = \sum_{\tau=0}^{\infty} \delta^\tau \Big\{ \int_{\underline{\theta}}^{\overline{\theta}} [t_\tau(\theta) + \varepsilon 1\{\tau = k\} - cq_\tau] f(\theta) d\theta \Big\} \tag{42}$$

$$= \sum_{\tau=0}^{\infty} \delta^\tau \Big\{ \int_{\underline{\theta}}^{\overline{\theta}} [t_\tau(\theta - cq_\tau] f(\theta) d\theta \Big\} + \delta^k \varepsilon. \tag{43}$$

So

$$\frac{dW(\varepsilon)}{d\varepsilon} = \delta^k. \tag{44}$$

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

$$\Gamma_k(\overline{\theta}) \equiv \int_{\underline{\theta}}^{\overline{\theta}} d\Gamma_k(\theta) = \frac{dW(\varepsilon)}{d\varepsilon} \frac{1}{\delta^k} = 1. \tag{45}$$

OA6 Additional Theoretical Results

OA6.1 Model Dynamics

Proofs are available in Supplemental Material Section SM3.

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, 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, as well as Proposition 2 below, which also finds that trade reaches a mature phase.

In terms of generality, the usual monopolist screening problem requires (or uses) log-concavity of $f(\theta)$.[39] 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.

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 ($\gamma_\tau(\theta) \geq \gamma_{\tau+1}(\theta)$).*

[39] 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.

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 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.

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.

To further understand the dynamics in the model, I present a solved two-type example in Supplemental Material Section SM4. The example illustrates the backloading of prices and quantities together with quantity discounts as a way to maximize lifetime profits for the seller while preventing opportunistic behavior from the buyer.

OA6.2 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; 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 ($X(\theta) = 0$) and the assumed parametrization of $v(\cdot)$ are given by:

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

With some abuse of notation, 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. Note, however, that quantities may converge at inefficient levels.

OA6.3 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:

$$(1 - \Gamma_\tau(\theta)) + \sum_{s=0}^{\tau-1}(1 - \Gamma_s(\theta)) - \theta\gamma_\tau(\theta) < 0. \tag{46}$$

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/high types early on. Moreover, it might apply too for lower types in the long-term that started with binding constraints at the beginning for the contract but that grew over time to become unconstrained. 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, early on, perfect enforcement could be more efficient, yet, 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.

OA7 Estimation

This section discusses the details of the estimation procedure.

OA7.1 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 $t_{i\tau} = T(q_{i\tau})e^{\upsilon_{i\tau}}$, such that $\upsilon_{i\tau}$ is independent of $q_{i\tau}$.

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

$$ln(t_{i\tau}) = \beta_{0\tau} + \beta_{1\tau} ln(q_{i\tau}) + \upsilon_{i\tau}, \tag{47}$$

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.

OA7.1.1 *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:

$$S_\tau(r) = \frac{exp(a_\tau + b_\tau r)}{1 + exp(a_\tau + b_\tau r)} + \varepsilon^s_\tau(r), \tag{48}$$

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.

OA7.2 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.

OA7.3 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:[40]

$$\Gamma_\tau(\alpha) = \frac{exp(\phi_\tau(q_\tau(\alpha)))}{1 + exp(\phi_\tau(q_\tau(\alpha)))}, \tag{49}$$

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:

$$\frac{\theta'(\alpha)}{\theta(\alpha)} = \frac{1}{d_0 + d_1 q_\tau(\alpha) + d_2 q_\tau(\alpha)^2}. \tag{50}$$

The key identification equation 12 provides the following estimating equation:

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

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

[40]The multiplier function is the solution to a differential equation. As shown in Supplemental Material Section SM2, it is a function of the cumulative distribution of types θ, the marginal cost, and the expected base marginal return (i.e., depends on the curvature of the return function).

To 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.

OA7.4 Buyer Types and Type Distribution

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

$$ln(\widehat{\theta}_\tau(\alpha)) = \tag{51}$$

$$\frac{1}{N_\tau}\sum_{k=1}^{N_\tau}\frac{1\{\alpha \geq k/N_\tau\}}{\widehat{\Gamma}_\tau(k/N_\tau) - k/N_\tau - \widehat{M}_\tau(k/N_\tau)}\Big[1 - \frac{\widehat{c}}{\hat{\beta}_{1\tau}p_\tau(k/N_\tau)} - \widehat{\gamma}_\tau(k/N_\tau)\Big], \tag{52}$$

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

$$\widehat{\theta'}_\tau(\alpha) = \frac{\widehat{\theta}_\tau(\alpha)}{\widehat{\Gamma}_\tau(\alpha) - \alpha - \widehat{M}_\tau(k/N_\tau)}\Big[1 - \frac{\widehat{c}}{\hat{\beta}_{1\tau}p_\tau(\alpha)} - \widehat{\gamma}_\tau(\alpha)\Big]. \tag{53}$$

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

OA7.4.1 *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

$$\widehat{v'(q_\tau(\alpha))} = \frac{\hat{\beta}_{1\tau}p_\tau(\alpha)}{\widehat{\theta}_\tau(\alpha)}. \tag{54}$$

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

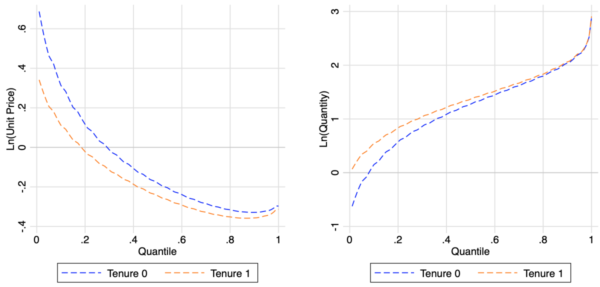
$$v(q_\tau(\alpha)) = \widehat{T}_\tau(q_\tau(0)) + \frac{1}{N_\tau}\sum_{k=1}^{N_\tau} v'(\widehat{q_\tau(k/N_\tau)})1\{\alpha \geq k/N_\tau\}. \tag{55}$$

OA7.5 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 $\widehat{v'(\cdot)}$.

Figure OA3: **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.

OA8 Monte Carlo Study

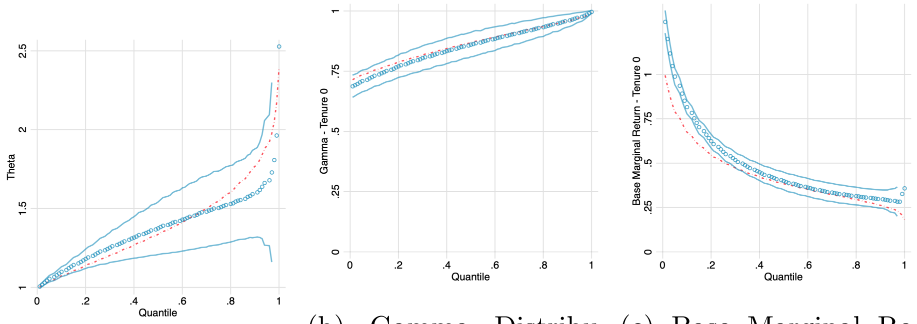
The Monte Carlo studies the behavior of my estimators for two periods of a dynamic contract without breakups. I use the following design. The return function is $v(\theta, q) = \theta q^{1/2}$. The type distribution is Weibull with scale parameter equal to 1 and shape parameter equal to 2, $F(\theta) = 1 - exp(-(\theta - 1)^k)$, normalized so $\underline{\theta} = 1$.[41] 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. In tenure 0, $\Gamma_0(\theta)$ has location parameter equal to 1 and scale parameter equal to 0.5. Instead, in tenure 1, $\Gamma_1(\theta)$ has location parameter 1 and scale 0.35. The lower scale parameter at tenure 1 reflects the idea that over time, the limited enforcement constraint is less binding. I construct the tariffs following Pavan et al. (2014): $t_\tau(\theta) = \theta q_\tau(\theta)^{1/2} - \int_{\underline{\theta}}^{\theta} q_\tau(x)^{1/2} dx$.

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

Online Appendix Figure OA3 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.

[41] 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.

Figure OA4: **Monte Carlo Results for Tenure 0**

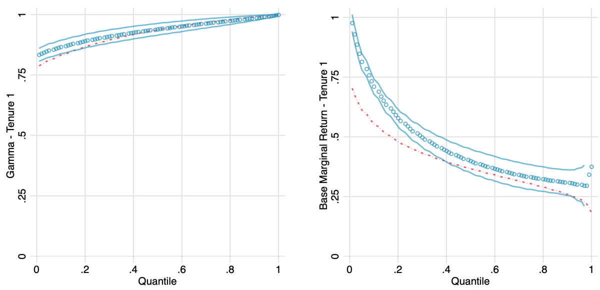


(a) Types θ (b) Gamma Distribution (c) Base Marginal Return

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

Figure OA5: **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.

Online Appendix Figure OA4 shows the results of the estimated Gamma distribution and the base marginal return, again in blue the estimated results and in red the true values. Both cases indicate good fit. Subfigure (a) shows the estimated $\hat{\theta}$ in blue and true θ in red by quantile. Dispersion at the 95 percent level are included for all except the top 2 quantiles, as they start to diverge. Overall, the figure shows a good fit, with most sections of including the true θ within their dispersion.

Next, I show the tenure 1's results estimates. Recall that the first-order condition of the seller now includes a backward-looking variable $1 - \Gamma_0(\theta)$ that keeps track of whether the limited commitment constraint was binding in the past. This variable is used by seller as a promise-keeping constraint that guarantees the seller delivers higher quantities and return in the future to prevent buyers from defaulting in the past. In my estimation, I use the tenure 0's predicted $\widehat{\Gamma}_0(\theta(\alpha))$ for each quantile α. Online Appendix Figure OA5 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, observations of quantity, and the nonparametric estimates of $v'(\cdot)$ as target. In particular, the estimated coefficient cannot be rejected to be different from 0.5 (the assumed value in simulation).

OA9 Additional Estimation Results and Model Fit

OA9.1 Distribution of t-Statistics against Standard Model Null

Online Appendix Table OA9.1 show the distribution of t-statistics for tests against a standard model null.

Table OA8: **Distribution of t-Statistics**

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

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

OA9.2 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. Online Appendix Table OA9 present the distribution of k and β.

Table OA9: **Parameters of Return Function**

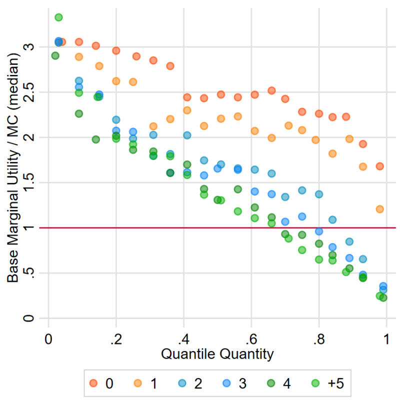
	mean	p10	p25	p50	p75	p90
β	0.56	0.30	0.48	0.61	0.76	0.82
k	171.23	9.00	17.24	39.64	86.61	282.40

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

OA9.3 Economic Magnitudes: Base Marginal Return

Online Appendix Figure OA6 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 OA6: **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 for each tenure.

OA9.4 Model Fit

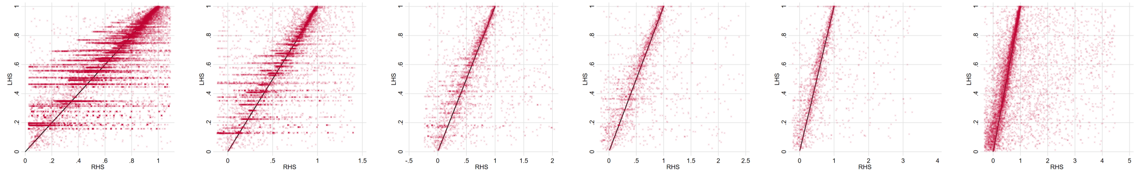
Online Appendix Figure OA7 presents the statistical fit of the model across tenures. It plots a reordered equation 12's left-hand side on the X-axis and the model's prediction using estimated coefficients of the right-hand side on the Y-axis.[42] Fit generally worsens for higher tenures; the results from Monte Carlo studies in Online Appendix OA8 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.

[42] Reorder equation 12 to obtain:

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

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

Figure OA7: **Model Fit - Statistical**

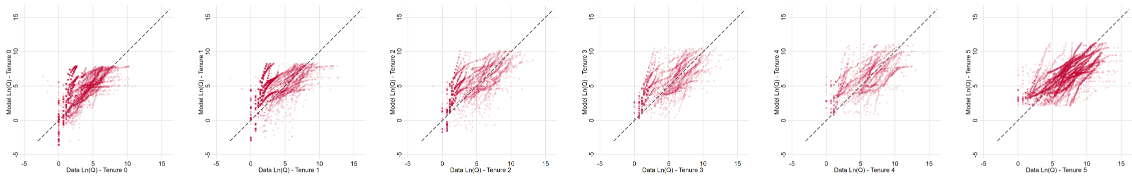


(a) T0 (b) T1 (c) T2 (d) T3 (e) T4 (f) T5

Notes: These figures show binscatters of statistical fit of the model across tenures as implied by identification equation 12. On the X-axis, it shows the predicted cumulative distribution function for the observation while on the Y-axis it plots the observed value.

Online Appendix Figure OA8 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 OA8: **Model Fit - Quantities**

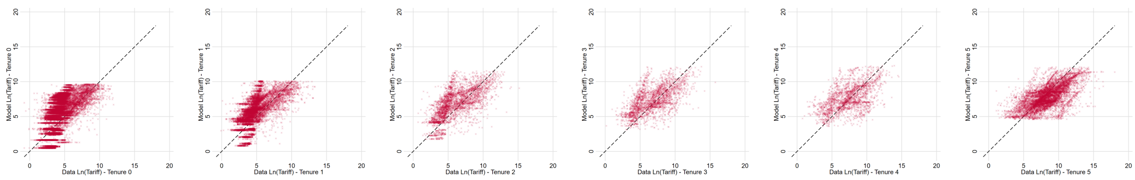


(a) T0 (b) T1 (c) T2 (d) T3 (e) T4 (f) T5

Notes: These figures display binscatters of model fit according to quantities. Estimated quantities use the close-form formula under the CES parametrization of the return function, as discussed in Appendix OA9. The X-axis plots the observed (log) quantities and Y-axis model predicted (log) quantities.

Online Appendix Figure OA9 shows the fit of tariffs. To generate tariffs in the model, I use the empirical equivalent of equation t-RULE.

Figure OA9: **Model Fit - Tariffs**



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

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

OA10 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.

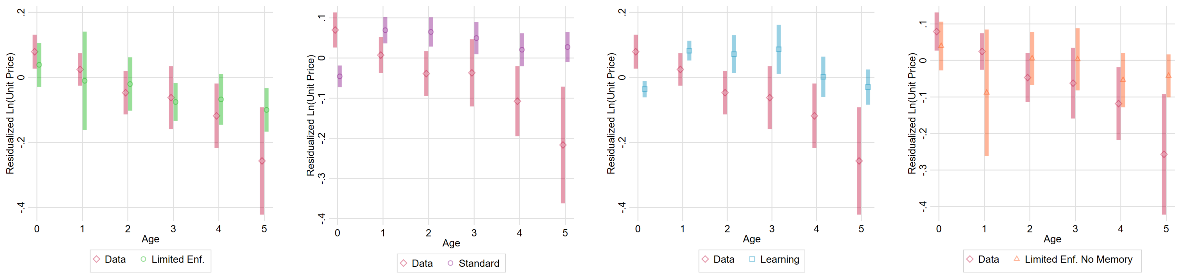
First, I consider the limited enforcement model, as discussed in the main text. Second, I consider the standard nonlinear pricing model. I estimate the model using the same methodology as in the model with limited enforcement, but set $\Gamma_\tau(\cdot) = 1$ and $\gamma_\tau(\cdot) = 0$ for all τ. Third, 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 Supplemental Material (available here). 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.

Online Appendix Figure OA10 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 is not able to capture the complete dynamics. Similarly, Subfigure (c) fails to replicate the correct discounting. Lastly, subfigure (d) shows that by eliminating the memory in the estimated limited enforcement model, model fit worsens, highlighting the importance of accounting for past promises in explaining the observed patterns.[43]

[43] These results are confirmed in regression form. I fail to reject different slopes between data and the limited enforcement model. However, the standard and learning models are rejected at the 1 significance level.

Figure OA10: **Model Comparison - Discounts**



(a) Limited Enf. (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 SM5. 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.

OA11 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. Online Appendix Table OA10 shows all the results.

The main takeaways are the following.

Buyers. Small-quantity buyers tend to prefer limited enforcement of contracts over perfect enforcement. They can effectively use the threat of default to reap higher returns. Instead, the median and top buyers prefer perfect enforcement in the short-term but limited enforcement in the long-term.

Under weak enforcement of contract, buyers prefer price discrimination than uniform pricing, as otherwise they would be excluded from trade (only median and top buyers prefer uniform pricing in the long-term). However, if exclusion and default are restricted, most buyers prefer uniform pricing.

Sellers. They prefer limited enforcement in the short-term but perfect enforcement in the long-term. Under weak enforcement of contract, they enjoy the ability to price discriminate, as it allows them to sell to buyers that would otherwise be excluded from trade. Instead, if enforcement is strong, they prefer uniform pricing in the short-term but price discrimination in the long-term. This is driven by the rapid increase in quantities, despite the decrease in unit prices offer to most buyers as incentive not to default.

Table OA10: **Counterfactual Policies**

		Nonlinear + Perfect					**Uniform + Limited**					**Uniform + Perfect**				
		10%	25%	50%	75%	100%	10%	25%	50%	75%	100%	10%	25%	50%	75%	100%
Buyer Return	*Tenure 0*	24.2	24.2	10.9	4.8	6.8	97.3	96.5	96.0	94.3	91.6	0.1	0.2	0.6	6.9	40.6
	Tenure 1	68.3	55.3	23.0	9.4	11.8	94.6	92.2	88.6	88.1	87.4	0.1	0.1	0.2	14.0	55.9
	Tenure 2	62.5	45.3	30.6	25.9	28.4	82.1	79.2	70.2	66.5	63.2	0.9	0.4	0.9	10.0	31.8
	Tenure 3	65.3	59.9	40.2	32.3	38.0	76.9	71.6	58.3	54.4	55.9	3.1	0.8	1.6	11.2	28.3
	Tenure 4	63.5	43.4	36.9	33.9	44.2	74.6	62.1	44.8	41.7	39.8	5.4	1.2	4.9	8.7	17.4
	Tenure 5	52.7	59.6	65.7	60.8	68.9	65.5	53.7	37.3	33.2	30.9	0.7	1.6	2.8	7.0	19.9
Seller Profit	*Tenure 0*	34.1	41.6	88.3	95.0	93.1	92.7	92.6	96.4	98.0	98.8	7.1	7.4	11.1	35.0	47.9
	Tenure 1	53.8	55.0	83.3	90.6	88.2	99.1	96.7	94.8	97.1	90.5	28.4	19.1	29.8	45.9	51.3
	Tenure 2	49.0	50.4	72.0	74.1	71.6	95.0	97.0	98.2	99.5	97.6	34.3	35.1	50.9	70.2	85.7
	Tenure 3	45.2	48.1	61.7	68.1	62.0	96.5	99.0	97.0	99.3	93.7	50.2	50.0	61.9	76.9	86.6
	Tenure 4	48.6	53.2	65.3	66.1	55.8	92.0	98.5	95.1	95.2	94.5	54.0	64.2	71.5	86.5	93.5
	Tenure 5	61.1	44.5	37.7	39.3	31.1	94.0	87.8	95.1	97.4	95.8	63.2	66.3	81.9	92.8	94.6
Surplus	*Tenure 0*	18.6	18.9	9.0	3.8	2.6	98.4	98.1	98.8	98.5	99.5	3.8	4.1	5.2	12.0	65.0
	Tenure 1	40.5	41.7	30.3	12.6	29.6	97.5	96.2	97.3	99.2	100.0	6.0	7.4	11.0	31.9	76.1
	Tenure 2	47.2	50.7	48.3	63.2	72.6	90.0	92.2	91.7	98.6	99.7	17.0	18.8	27.6	57.0	95.0
	Tenure 3	60.9	57.6	69.7	76.6	70.1	90.6	92.5	88.1	98.4	99.6	24.3	26.0	37.4	69.1	98.4
	Tenure 4	66.7	71.9	74.5	77.1	67.3	86.2	89.6	85.9	98.3	99.5	28.1	40.1	53.8	79.2	97.9
	Tenure 5	74.2	79.6	88.6	91.4	84.8	82.0	78.3	80.3	96.9	100.0	30.5	34.2	53.7	86.1	99.9
Unit Prices	*Tenure 0*	76.3	75.7	89.1	94.6	93.1	93.6	93.1	95.4	90.3	43.6	93.6	93.1	95.4	90.3	43.6
	Tenure 1	56.0	55.7	77.8	90.5	88.1	98.6	96.8	87.9	68.2	25.0	98.6	96.8	87.9	68.2	25.0
	Tenure 2	44.0	56.0	68.3	74.0	71.7	92.6	94.8	90.9	64.0	18.3	92.3	94.7	90.9	64.0	18.3
	Tenure 3	37.3	41.5	58.8	67.9	61.8	91.2	96.7	89.2	55.1	13.3	90.8	96.7	89.2	55.1	13.3
	Tenure 4	38.2	56.4	63.1	67.6	55.6	88.3	95.5	88.0	65.3	20.0	87.7	95.5	88.0	65.2	20.0
	Tenure 5	43.5	38.2	35.9	38.2	31.1	90.6	91.3	87.7	53.5	10.5	89.8	91.1	87.6	53.2	10.0
% Excluded	*Tenure 0*	-	-	-	-	-	97.3	96.4	95.8	94.1	90.4	-	-	-	-	-
	Tenure 1	-	-	-	-	-	93.4	91.9	88.6	87.3	85.8	-	-	-	-	-
	Tenure 2	-	-	-	-	-	79.8	77.4	70.0	65.6	61.3	-	-	-	-	-
	Tenure 3	-	-	-	-	-	74.4	69.2	58.0	51.4	50.0	-	-	-	-	-
	Tenure 4	-	-	-	-	-	71.3	60.1	44.7	38.5	37.4	-	-	-	-	-
	Tenure 5	-	-	-	-	-	62.3	50.7	36.4	30.3	25.8	-	-	-	-	-

Notes: This table reports the % share of observations for which the reported category (e.g., Buyer's Net Return) is greater under the observed nonlinear pricing regime than under the alternative policy. The values are reported across different tenures and percentiles in the distribution of types. The policies considered are 1) Nonlinear pricing with perfect enforcement, 2) Uniform monopolist pricing with limited enforcement, and 3) Uniform monopolist pricing with perfect enforcement. The reported categories are Buyer's Net Return, Seller's Profits, Total Surplus, Unit Prices, and percentage of Excluded Buyers.

Online Appendix References

Alfaro-Urena, A., Manelici, I., and Vasquez, J. P. (2022). The effects of joining multinational supply chains: New evidence from firm-to-firm linkages. *The Quarterly Journal of Economics*.

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

Attanasio, O. and Pastorino, E. (2020). Nonlinear pricing in village economies. *Econometrica*.

Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., and Moxnes, A. (2019). The origins of firm heterogeneity: A production network approach. Technical report, National Bureau of Economic Research.

Bloom, N., Sadun, R., and Van Reenen, J. (2012). Americans do it better: Us multinationals and the productivity miracle. *American Economic Review*, 102(1):167–201.

Fong, Y. F. and Li, J. (2017). Relational contracts, limited liability, and employment dynamics. *Journal of Economic Theory*, 169:270–293.

Hjort, J., Li, X., and Sarsons, H. (2020). Across-country wage compression in multinationals. Technical report, National Bureau of Economic Research.

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

Kostadinov, R. and Kuvalekar, A. (Forthcoming). Learning in relational contracts. *American Economic Journal: Microeconomics*.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and Organization*, 15(1):222–279.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Legal determinants of external finance. *Journal of Finance*, 52(3):1131–1150.

Luo, Y., Perrigne, I., and Vuong, Q. (2018). Structural analysis of nonlinear pricing. *Journal of Political Economy*, 126(6):2523–2568.

Martimort, D., Semenov, A., and Stole, L. (2017). A theory of contracts with limited enforcement. *Review of Economic Studies*, 84(2):816–852.

Pavan, A., Segal, I., and Toikka, J. (2014). Dynamic mechanism design: A myersonian approach. *Econometrica*, 82(2):601–653.

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