Importers, Exporters and the Division of the Gains from Trade*

Andrew B. Bernard†
Tuck School of Business at Dartmouth, CEPR & NBER

Swati Dhingra‡
LSE, CEP & CEPR

This Version: September, 2016

Abstract

This paper examines the microstructure of import markets and the division of the gains from trade among consumers, importers and exporters. When exporters and importers transact through anonymous markets, double marginalization and business stealing among competing importers lead to lower profits. Trading parties can overcome these inefficiencies by investing in richer arrangements such as bilateral contracts that eliminate double marginalization and joint contracts that also internalize business stealing. Introducing the microstructure of import markets into a trade model, we show that trade liberalization increases the incentive to engage in joint contracts, thus raising the profits of exporters and importers at the expense of consumer welfare. We examine the implications of the model for prices, quantities and exporter-importer matches in Colombian import markets before and after the US-Colombia free trade agreement. US exporters that started to enjoy duty-free access were more likely to increase their average price, decrease their quantity exported and reduce the number of import partners. These exporters increased their average tariff-inclusive price by as much as 25%, leading to a low average pass-through rate of less than 4% for all imports that received duty-free access into Colombia.

Keywords: Heterogeneous firms, exporters, importers, pass through, contracts, microstructure, consumer welfare.
JEL codes: F10, F12, F14.

*We are grateful to John Morrow, Emanuel Ornelas, Gianmarco Ottaviano, Chris Snyder, Catherine Thomas, Dan Trefler, Tony Venables, Alwyn Young and seminar participants at Aarhus, Berkeley-Stanford, Bologna, ERWIT, FGV Sao Paulo, IIFT, KU Milan, Oslo, Oxford, LSE, Paris, Princeton Summer Workshop, PSU, PUC, Salento, SED, Sussex, Toronto, Tuck (Dartmouth), World Bank and UEA for helpful comments. We thank Giovanni Maggi and Dan Trefler for detailed discussions and Angela Gu and Yan Liang for excellent research assistance.

†100 Tuck Hall, Hanover, NH 03755, USA, *tel:* +1 603 646 0302, *email:* andrew.b.bernard@tuck.dartmouth.edu

‡Houghton Street, London, WC2A 2AE UK; *tel:* +44(0)20-7955-7804, *email:* s.dhingra (at) lse.ac.uk

1 Introduction

Trade liberalization can generate substantial improvements in welfare through increased product variety, lower prices and the reallocation of resources from less to more efficient producers. Domestic consumers gain from easier access to imports, but they rarely have direct access to foreign products. Trade is typically between firms, who exercise market power in buying and selling foreign products. This paper examines how the behavior of importers and their relationships with exporters affect the division of the gains from trade. Trading partners make endogenous choices among different types of contracts, which gives them different degrees of market power in importing. Trade liberalization affects the incentives for different contracts. This leads to a change in the microstructure of import markets, which provides a qualitatively new channel for the division of the gains from trade. Foreign exporters and domestic importers potentially gain market power and benefit at the expense of domestic consumers.

In line with the concerns of the European Commission and the Federal Trade Commission, market integration can induce firms to replace trade barriers with microstructures that limit the gains from market access to consumers (Raff and Schmitt 2005). Typically, foreign competition is associated with lower price-cost margins (Tybout 2003; Levinsohn 1993; Harrison 1994) but a growing literature documents that the pass-through of border prices into home market prices is low in macro studies and in detailed micro work.[1] In our framework, the low pass-through of trade cost reductions into import prices is a result of the changes in the contracts between exporters and importers.

Exporters and importers operate in thin markets, that do not reflect anonymous transactions. Anonymous transactions lead to two inefficiencies in profit maximization. First, when exporters and importers engage through anonymous market-clearing prices, market power of importers leads to double marginalization and lower profits. Exporters can overcome this inefficiency by moving away from unit prices and offering bilateral contracts that specify total payments and quantities. Second, importers in an anonymous market are unable to internalize the business stealing effect of their sales on competing importers that sell varieties of the same foreign product. The exporter can mitigate this externality by investing in joint contracts with these competing importers. This leads to higher profits by restricting sales and driving up prices charged to consumers. We embed the microstructure of richer contracts into a standard trade model. The microstructure channel reinforces the gains

[1] See for example Engel and Rogers (1996); Burstein et al. (2003); Campa and Goldberg (2005) in the macro literature. Detailed micro studies include Ganslandt and Maskus (2004) for Swedish pharmaceuticals, De Loecker et al. (2012) and Mallick and Marques (2008) for Indian manufacturing, Badinger (2007) for manufacturing and services in the European Single Market, and Konings et al. (2005) for Bulgarian and Romanian manufacturing industries.

from trade when firms enter into bilateral contracts that mitigate double marginalization. But as firms start to invest in internalizing business stealing, trade liberalization increases their market power and affects the division of the gains from trade. They engage in complex relationships to overcome the inefficiencies of anonymous markets.

In our model, market expansion increases the returns to joint contracts, relative to bilateral contracts. This is because market expansion enables firms to better amortize the higher fixed costs of joint contracts and because import competition does less damage to firms with greater market power. As firms switch to joint contracts, they internalize the business stealing effect of their sales. They raise consumer prices, and get more of the surplus from consumers. We show that the microstructure effect is a new channel for the gains from trade, that goes beyond the standard gains from variety expansion and pro-competitive effects, and that increases profits at the expense of consumer welfare.

To test for the microstructure effect, we would ideally like to examine the impact of trade liberalization on the actual relationships between exporters and importers. Such data is unavailable so we examine the distinct observable implications for prices, quantities and exporter-importer matches during a period of trade liberalization. If the microstructure matters, trade liberalization will induce exporters to consolidate their import market. They will benefit from higher prices by scaling back on the quantity they sell and the number of importing partners that they sell to in the liberalizing country. We test these implications for Colombian imports before and after the US-Colombia Free Trade Agreement (FTA) using transaction-level matched importer-exporter data. In keeping with the theory, we find that US exporters whose products started to enjoy duty-free access through the FTA were more likely to simultaneously increase their import price, lower their quantity and reduce the number of their importer partners in Colombia. This confirms the Metzler paradox of domestic prices rising in response to tariff reductions and, we find that it persists after adjusting for quality in prices. Exporters in the first and second terciles of initial size greater than \$10,000 increased their free-on-board import prices by 8 to 30% on average. This contributed to a less than 4% average reduction in tariff-inclusive prices for all Colombian imports that received duty-free access.

The main theoretical contribution of the paper is to provide a model that embeds two-sided market power in a general equilibrium setting. We build on the vertical relations literature in industrial organization to model the relationship between exporters and importers (Hart and Tirole 1990). While a large literature examines these relationships, its focus is on firm behavior, typically in a stylized setting of two buyers and sellers. We embed the firm behavior in a general equilibrium setting with multiple firms, which is important for understanding the gains from trade because they are also affected by labor market clearing and firm entry and exit. This is also useful in taking the theoretical predictions to trade data that is usually available through firm-level customs transactions.

Our focus on vertical relations is related to work in international trade on intermediation (Akerman 2010; Bernard et al. 2010; Blum et al. 2010; Ahn et al. 2011; Atkin and Donaldson 2012), retailing (Eckel 2009; Raff and Schmitt 2012; Blanchard et al. 2013) and vertical integration (Feenstra et al. 2003; Antràs and Helpman 2004; Conconi et al. 2012). As is well-known, embedding two-sided market power in general equilibrium gives intractable models so many of these papers focus on firm characteristics and abstract from the market power of importers. Raff and Schmitt (2005, 2009) examine importer market power in an oligopoly model of trade and retailing to show trade liberalization can reduce welfare due to vertical restraints. We embed vertical restraints in a general setting with many heterogeneous exporters and importers to obtain predictions that can be taken to the data. We obtain tractability by using the tools developed in the variable markups literature under monopolistic competition (Parenti et al. (2014); Neary and Mrazova (2013); Dhingra and Morrow (2012); Mayer et al. (2014)). This enables us to generalize the results for firm behavior to a wide class of demand functions, which is important in revealing new gains from trade that do not arise under the knife-edge cases of standard demand systems.

The focus on consumption side gains is similar to a large literature on the impact of trade liberalization on markups and prices. Tybout (2003) surveys the research using industry-level data and plant-level panel data and concludes that most studies find higher industry-level exposure to foreign competition is associated with lower price-cost margins or markups, e.g. Levinsohn (1993); Harrison et al. (2005). The pass-through of reductions in trade costs to domestic prices is typically low, and recent studies find the behavior of domestic firms determines the extent to which trade policy affects prices at home. Mallick and Marques (2008) find low tariff rate pass-through into import prices in Indian manufacturing during the liberalization of 1991. De Loecker et al. (2012) estimate that on average, factory-gate prices fell by 18 percent despite average import tariff declines of 62 percentage points, as domestic Indian firms did not pass on the reductions in trade costs to consumers. Badinger (2007) finds the European Single Market led to an overall reduction in markups for manufacturing products, but markups rose in several manufacturing and services industries that also experienced an increase in industry concentration and average firm size. In early work on Japanese imports, Lawrence and Saxonhouse (1991) document that the presence of large conglomerates at home is associated with lower import penetration, suggesting the import-inhibiting effects of firms with high market power. Lawrence (1991) argues market power of intermediaries can explain why Japanese consumer prices were higher than German import prices for the same export from the US. Yeats () finds iron and steel prices are higher in more concentrated import markets. In early theoretical work, Venables (1985) shows unilateral reductions in trade barriers can increase consumer prices in the liberalizing country when entry and exit induce profit shifting across countries.

Our pricing predictions are consistent with these findings, and provide systematic evidence for the Metzler paradox (Venables 1985; Bagwell and Staiger 2012; Bagwell and Lee 2015). A reduction in tariffs induces more consolidation in the import market of medium-sized US exporters in Colombia. Previous work has suggested that trade policy is a substitute for competition policy. Small open economies can increase competition in the domestic market by integrating with world markets and benefiting from import competition. Our results show that when exporters and importers operate in thin markets, trade policy and competition policy are complements. Trade liberalization would be most beneficial when competition policy is used in conjunction with trade policy to encourage firms to pass on the cost savings to consumers.

We introduce contacting choices to see how market power in import markets changes the gains from trade. This is related to several different strains of work in international trade. Bernard et al. (2009), Castellani et al. (2010) and Muuls and Pisu (2009) document substantial heterogeneity across importing firms for the US, Italy and Belgium respectively and also show that importers differ from non-importers. Papers by Rauch (1999), Rauch and Watson (2004), Antràs and Costinot (2011), Chaney (2014) and Petropoulou (2011) model the formation of matches between exporters and importers. These papers adopt a search and matching approach to match formation. Our paper is also closely related to the recent set of papers using matched exporter-importer data. Blum et al. (2010, 2012) examine characteristics of trade transactions for the exporter-importer pairs of Chile-Colombia and Argentina-Chile while Eaton et al. (2014) consider exports of Colombian firms to specific importing firms in the United States. Carballo et al. (2013) and Bernard et al. (2014) use matched data to study the role of buyers in firm-level trade flows. We abstract from these mechanisms and focus on the microstructure of import markets to understand the resulting effects on prices, quantities and the division of the gains from trade. A growing literature also shows how imported inputs increase the productivity of importers, see Amiti and Konings (2007), Halpern et al. (2015) and Boler et al. (2015). We focus on the consumption side role of importers.

The rest of the paper is organized as follows: Section 2 introduces a contracting choice model where exporters and importers transact through anonymous markets, bilateral contracts or joint contracts. Trade liberalization affects the division of the gains from trade among domestic consumers, home importers, and foreign exporters. Section 3 develops the key predictions of the model on prices, quantities, and matches. Section 4 takes the observable implications to Colombian import data and Section 5 concludes.

2 Model of Exporters and Importers

This section describes the economy which consists of consumers, exporters and importers. We work with a standard trade model akin to Krugman (1979), and introduce importers with market power. As is well-known, two-sided market power in an industry equilibrium often leads to intractable models. This is why several papers in intermediation abstract from market power at least on one side of the market. As our focus is on the division of the gains from trade, we model market power of exporters and importers in an industry equilibrium. We achieve this by using the tools developed in Dhingra and Morrow (2012) for monopolistic competition trade models. For ease of exposition, we start with homogeneous firms in this Section. The next Section examines the testable variety-level predictions of the theory under firm heterogeneity and fixed entry.

In a standard setting, exporters sell to importers in anonymous markets which means they set a unit price at which any importer can buy from them. We first model the firm problem in an anonymous market with unit prices, and show that market power results in double marginalization and business stealing among importers of a product. When exporters are not constrained to set unit prices, exporters and importers choose payments and quantities that maximize their joint profits by overcoming the externalities induced by an anonymous market. We specify the simplest setting that departs from anonymous markets to provide a richer microstructure for the import market.

There are two countries, Home and Foreign (with x indexing exports from the foreign country to the home country). Differentiated products are produced at home and abroad, and are further differentiated by domestic firms (importers) who sell to consumers. The home country has L workers, each of whom is endowed with a unit of labor and has preferences over consumption goods. We specify the demand and technology in the next sub-sections, and then discuss firm decisions under anonymous markets.

2.1 Preferences

Each worker has identical preferences over varieties of a differentiated good. Preferences for differentiated goods take the form of nested variable elasticity of substitution (VES) utility. A differentiated product in the upper nest is indexed by i, and is a composite of further differentiated variants in the lower nest which are indexed by ij. The utility function is as follows:

$$U \equiv \int u(q_i)di, \qquad q_i = \int v(q_{ij})dj, \qquad u', v' > 0,\ u'', v'' < 0.$$

The VES inverse demand builds on a growing literature of variable markup frameworks like Parenti et al. (2014); Neary and Mrazova (2013); Dhingra and Morrow (2012); Mayer et al. (2014). The sub-utility functions u and v are thrice continuously differentiable, strictly increasing, strictly concave on $(0, \infty)$, normalized to zero at zero quantities and satisfy inada conditions. Concavity ensures that consumers purchase all available varieties and the inverse demand for variety ij is

$$p_{ij} = u'(q_j)v'(q_{ij})/\delta \tag{1}$$

where δ is the consumer's budget multiplier. For a worker with income I, the budget multiplier is

$$\delta = \int\int u'(q_j)v'(q_{ij})q_{ij}didj/I.$$

A well-defined equilibrium will require conditions on the demand elasticities across the two nests of preferences. Following Dhingra and Morrow (2012), the elasticity of utility is $\varepsilon_x(q) \equiv x'(q)q/x(q)$ for $x \in \{u, v\}$ and the elasticity of marginal utility is $\mu_x \equiv -x''(q)q/x'(q)$. These elasticities are bounded below by $b > 0$ and above by $1 - b < 1$, and are increasing in q. To fix ideas, CES demand is $x(q) = q^\rho$ which implies $\varepsilon_x(q) = \rho$ and $\mu_x(q) = 1 - \rho$ are between 0 and 1, and correspond to the special case with $\varepsilon_x^{'}(q), \mu_x^{'}(q) = 0$. While CES demand provides simple pricing decisions, we depart from CES demand because it gives knife-edge results that trade liberalization has no impact on the microstructure, similar to the lack of pro-competitive effects in Krugman (1979) and Melitz (2003) under CES demand.

2.2 Technology

Following Krugman (1980), there are M_c identical producers at home. Each producer supplies a unique differentiated product with a linear technology. Each producer faces a unit cost c. Producers are monopolistically competitive and pay fixed operation costs f_c. Producers cannot directly access final consumers. They must engage distributors to deliver their products to consumers.

There are M_d identical distributors at home. Producers at home and abroad must sell through these distributors to the consumers. Distributors are monopolistically competitive and transform the producer's product into a differentiated variety for final consumption. A distributor with unit cost d transforms producer c's product from $x(c)$ units of production into $y(c, d) = x(c)/d$ units of the final differentiated variety. If $p^x(c)$ is the unit price of variety c charged by the producer, then the unit cost of a distributor is $p^x(c)d$. Under this formulation, distributors perform the function of lowering the costs of delivery to consumers.

Before proceeding to the equilibrium, we need to specify where the producers and distributors fit into the nesting structure of demand. To capture rich substitutability patterns, we specify the upper nest quantity as $q_j = \theta q_c + (1 - \theta)q_d$ for $\theta \in \{0, 1\}$. When $\theta = 0$, $q_j = q_d$ and distributors are in the upper nest. For $\theta = 1$, $q_j = q_c$ and producers are in the upper nest. We will show that $\theta = 1$ leads to interesting results because the lower nest distributors do not internalize the impact of their sales on other distributors. This provides a role for richer contracting arrangements and gives new results for the impact of trade liberalization on the gains from trade. As we will initially focus on a free trade equilibrium, all firms will engage in international trade and we will use the term exporters to denote c firms and importers to denote d firms.

2.3 Anonymous Market Equilibrium

A natural way of introducing importers into a standard trade model is through anonymous market transactions. An exporter chooses the market price for her product and then takes her product to an import market. The importers choose how much to buy. They further differentiate the product and supply final varieties to consumers. Then cd indexes a final variety exported by c and imported by d. We start with this benchmark case of anonymous markets to illustrate the inefficiencies that lead to richer relationships between exporters and importers. To formalize the setting, the timing is as follows.

- Firms pay their fixed costs of operation (f_c,f_d).
- c chooses her market price $p^x(c)$.
- d buys x_{cd} units at price $p^x(c)$.
- Quantities q_{cd} are supplied to final consumers.

Markets are segmented and we solve for an equilibrium by first determining the final quantities sold to consumers. Then we derive the demand for the producer's product and determine the optimal price chosen by the producer. We abstract from variety-specific search costs f_{cd} until the Section with heterogeneous firms where they lead to interesting selection effects.

2.3.1 Prices in Anonymous Markets

Importer d faces the inverse demand function $p_{cd} = v'(q_{cd})u'\left(\theta q_c + (1-\theta)q_d\right)/\delta$. He cannot influence the aggregate market conditions δ and the producer's total sales in the home country q_c. At unit price p^x_c of c's product, the importer chooses final quantities q_{cd} and his total quantity q_d to maximize profits. His variable profit is $\pi_d \equiv M_c \left(p_{cd}\left(q_{cd}, q_d, \hat{q}_c, \hat{\delta}\right) - p^x_c d\right) q_{cd} L_d$ where the hat denotes that the distributor takes these components of the inverse demand as given. Summing the demand of all importers in the home country $x_c \equiv M_d x_{cd} L_d = M_d d q_{cd}\left(p^x_c\right) L_d$, exporter c chooses p^x_c to maximize variable profits $\pi_c = M_d\left(p^x_c - c\right) d q_{cd}\left(p^x_c\right) L_d$.

To ensure a well-defined firm problem, we assume that marginal revenues are decreasing and that $\mu_v^{'} q < (1-\mu_v)(1-\mu_v-\mu_u\varepsilon_v)$. We will also assume that $\mu_u\varepsilon_v \leq \min \mu_v \cdot \min(1-\mu_v)$ which ensures that the direct impact of own quantity on price is greater than the indirect impact of other varieties through the upper nest quantities. This condition is equivalent to the substitutability restriction of a nested CES demand system and later we will show that it ensures prices are higher under anonymous markets compared to bilateral and multilateral profit maximization among firms.

Under these conditions, d's optimal price sets the marginal cost of a variety equal to the marginal revenue from that variety and from other varieties sold by d. The optimal price chosen by d is $p_{cd} = p^x_c d/\left(1 - \mu_v(q_{cd}) - \frac{(1-\theta)u'(q_d)}{\theta u'(q_c)+(1-\theta)u'(q_d)}\mu_u(q_d)\varepsilon_v(q_{cd})\right)$. Using subscripts for brevity, the exporter's optimal price is $p^x_c = c/\left(1-\gamma_c\right)$ where $\gamma_c \equiv \mu_{v_{cd}} + \mu^{'}_{v_{cd}} q_{cd}/\left(1 - \mu_{v_{cd}} - \frac{(1-\theta)u^{'}(q_d)}{\theta u^{'}(q_c)+(1-\theta)u^{'}(q_d)}\mu_u(q_d)\varepsilon_{v_{cd}}\right)$.

Putting the two optimal price functions together, the final price of variety cd under anonymous markets is

$$p_{cd} = cd/\left[\underbrace{1-\gamma_c}_{\text{Double marginalization}} \cdot \left(1-\mu_{v_{cd}} - \frac{(1-\theta)u'(q_d)}{\underbrace{\theta u'(q_c)}_{\text{Business Stealing}} + (1-\theta)u'(q_d)}\mu_u(q_d)\varepsilon_{v_{cd}}\right)\right]. \tag{2}$$

The first term in square brackets in Equation 2 is the markup charged by the exporter to the importer which reflects the classic double marginalization problem in anonymous markets. Exporters take into account the derived demand for their product and charge $p_c^x = c/(1-\gamma_c) > c$. Importers further mark up the price (with the term in parenthesis) and consumers end up with having to pay double markups. This double marginalization leads to lower bilateral profits for c and d due to reduced sales. If producers and distributors can engage outside of anonymous transactions, then they need not set market-clearing unit prices. They can specify payments and quantities that get rid of double marginalization and increase the bilateral profits from their relationship. We will consider these "bilateral contracts" in the next Section and show that trade liberalization has the usual effect of increasing sales and reducing prices when exporters and importers transact under bilateral contracts.

The term in parenthesis shows the markup charged by the importer. Importers account for the cannibalization of their own varieties on each other. This cannibalization translates into higher final markups through $\mu_u(q_d)$. But they do not account for the business stealing impact of their sales on competing distributors of an exporter's product, which can be seen from the denominator in parenthesis. There is no $\mu_u(q_c)$ in the optimal markup and competition among distributors in the final goods market implies that the total profit from an exporter's product is not maximized. We will therefore consider "joint contracts" that maximize the total profits of an exporter and her importers. While an importer can internalize cannibalization of her own varieties by virtue of being second in the chain of sales, exporters and importers are unable to overcome the business stealing externality through anonymous markets.

Anonymous markets therefore provide two reasons for switching to richer contracts. Unit pricing in anonymous markets leads to double marginalization which lowers the bilateral profit of an exporter and an importer. Bilaterally profit maximizing contracts that specify payments and quantities can overcome the double marginalization problem and reduce consumer prices. Importers internalize the cannibalization of their own sales. When importers are in the lower nest ($\theta = 0$), there is no other reason to enter into richer contracts with the exporters. Bilateral contracts solve the double marginalization problem and the cannibalization effect is already accounted for. When exporters are in the upper nest ($\theta = 1$), there is an incentive for an exporter and her importers to enter into joint contracts to get rid of business stealing. Consumer prices are higher compared to bilateral contracts because competition among importers falls.

Bilateral contracts reinforce the usual gains from trade because they reduce consumer prices further. Joint contracts have the opposite effects. Competition in the import market is reduced to increase the prices faced by consumers. We examine how these richer arrangements between exporters and importers determine the division of the gains from trade liberalization between exporters, importers and consumers. To avoid a taxonomical analysis, we focus on the case where exporters are in the upper nest and importers are in the lower nest, so that joint contracts are desirable. We explain the contracting arrangements that mitigate the externalities of anonymous markets and proceed to the industry equilibrium under this richer microstructure.

2.4 Microstructure of Import Markets

Exporters and importers often have long-standing and complex relationships. It is therefore likely that they engage in contracts that overcome the inefficiencies from anonymous transactions. Following the seminal work of Hart and Tirole (1990), we consider two distinct contracts that overcome these inefficiencies. The first type of contract overcomes double marginalization by specifying quantities and payments that maximize bilateral profits of the exporter-importer pair. The second type of contract overcomes double marginalization and business stealing by committing to all quantities and payments that maximize the joint profit of an exporter and her importers.

The main advantage of the Hart-Tirole approach is that we do not need to specify the methods through which firms maximize bilateral profits or joint profits, and instead can focus on the observable outcomes of quantities and payments that result from implementing profit-maximizing contracts. For instance, an exporter can maximize bilateral profits by setting a two-part tariff that charges the importer a price equal to her marginal cost ($p^x(c) = c$), and a fixed fee that recoups part or all of the bilateral profits without changing sales incentives. The exporter could also have maximized bilateral profits through resale price maintenance. The importer is then obliged to sell at a price chosen by the exporter. As we will show later in this Section, when the exporter chooses the final price, the resulting quantity allocation is the same as the two-part tariff case and bilateral profits of the exporter-importer pair are maximized.

The industrial organization literature provides different methods through which firms avoid double marginalization and business stealing, such as fixed fees, quantity discounts and resale price maintenance.[2] These are rarely observable, and these methods can often take the form of informal practices or implicit arrangements that are sustained through repeated interaction. Following Hart and Tirole (1990), we therefore abstract from the methods used to implement bilateral or joint profit maximization, and focus instead on the implications for observable outcomes, such as quantity allocations and payments, that are the same across different methods.

[2] This is an area of ongoing research in industrial organization and Miklos-Thal et al. (2010) provide an overview of the findings.

We start with bilateral profit maximization and then discuss joint profit maximization. The market does not maximize bilateral profits of an exporter-importer pair due to double marginalization. This problem is overcome when an exporter engages in a *bilateral private contract* with an importer. The bilateral contract specifies a final quantity and a payment to be made to the exporter (q_{cd} and T_{cd}). As the exporter no longer relies on unit prices that are marked up, the final quantity is chosen to maximize the bilateral profit of the exporter-importer pair. Following Horn and Wolinsky (1988), the exporter and the importer split the bilateral profits through Nash bargaining, and this determines the payments from the importer to the exporter. The bilateral contract overcomes double marginalization and results in higher quantities and lower prices for the final consumers.

But bilateral contracts do not mitigate the business stealing externality imposed by importers on each other. The competition between different importers implies that prices are lower than the monopolistic price that the exporter would have chosen. The key insight of Hart and Tirole (1990) is that a seller cannot commit to selling lower quantities of her product because she would prefer to bypass her existing buyers and sell more to earn higher profits. Even though the exporter could offer contracts that account for the business stealing externality, this opportunism of the exporter prevents importers from entering into such arrangments because they know the exporter would find it more profitable to deviate from these contracts. The exporter's opportunism prevents maximization of multilateral joint profits.

To overcome business stealing, the exporter must commit to restricting the total sales of her product to induce higher consumer prices by forming a *joint contract* with her importers.[3] Just like bilateral contracts, joint contracts specify bilateral quantities and payments, but now the total quantity of the exporter q_c is also specified in the contract and is not simply an outcome of the different bilateral contracts. Joint contracts can be implemented through various methods such as vertical integration that would eliminate the exporter's opportunism by aligning the interests of all parties. The integrated parties would internalize business stealing to maximize joint profits, and consumers would end up with lower quantities and higher prices than under bilateral contracts. The exporter could also implement joint profit maximization through other methods that do not involve ownership. For instance, importers would internalize business stealing if the exporter assigns them exclusive territories. The outcomes of the joint contract can also be replicated through "implicit exclusive dealing" when exporters have reputational concerns due to repeated interaction with their importers (Rey and Tirole 2007). We abstract from the methods through which joint profit maximization is implemented, and focus instead on observable outcomes such as prices and

[3] Joint contracts often refer to cross-ownership between the parties. We use the term joint contract to refer to a contracting choice between the exporter and importer which might, but does not necessarily, include cross-ownership. Martin et al. (2001) and Mollers et al. (2014) use experimental data to show vertical restraints of various forms (that do not entail vertical integration of firms) are sufficient to maximize joint profits.

quantities.

To formally model the different microstructures resulting from these contracts, we specify the timing as follows:

- Exporters and importers pay their sunk costs of operation (f_c, f_d)
- Exporters and importers meet each other.
- Exporters and importers decide whether to engage in bilateral or joint contracts and pay their contracting costs.
- Importers order quantities x_{cd} from the exporters and pay T_{cd}.
- Quantities q_{cd} are supplied to final consumers.

Exporters and importers need to pay fixed costs for contracting. Both will decide whether it is worthwhile to engage in richer contracts, but only one of their conditions will be binding. We assume that the exporter's contracting condition is binding, because this will give new results for the gains from trade. In the opposite case, when importer's contracting costs are binding, the change in the gains from trade goes in the same direction as the standard gains from trade. For simplicity, we set the contracting costs incurred by importers to zero because this does not change the qualitative results from the model.[4]

2.5 Bilateral Private Contracts

To overcome double marginalization, an exporter can make fixed investments in bilateral contracts that depart from unit pricing. After paying the fixed costs, an exporter engages in private contracts with each importer bilaterally. Importers hold passive beliefs which means that they expect the contracts offered to the other importers to be fixed at their equilibrium values. We first discuss the surplus division, and then proceed to determining optimal prices and entry into bilateral contracts.

2.5.1 Importer Payments

Under bilateral contracts, the importer chooses quantities q_{cd} to maximize his profit:

$$\pi_{cd}^{B} = \left(p(q_{cd}, q_d, \hat{q}_c, \hat{\delta}) - cd\right) q_{cd} - T_{cd}$$

where T_{cd} is the payment to exporter c for supplying $x_{cd} = dq_{cd}$ units of her product sent to d. The importer holds passive beliefs so he takes the contracts offered to his competitors as given $((\hat{T}_{cd'}, \hat{q}_{cd'})$ for any $d' \neq d)$.

[4] No new information is revealed after the contracting costs are paid. If the contracting costs of each party are sunk, then the opportunism problem remains and the qualitative results in the subsequent Sections are unaltered, but it would be more reasonable to think that the division of surplus is over the revenues, rather than the profits net of fixed costs.

For flexibility, we follow Horn and Wolinsky (1988) and assume that the payments are set through bilateral Nash bargaining, with β denoting the bargaining weight of the exporter and with zero disagreement payoffs.

The optimal payment is the solution to $\max_{T_{cd}} \left[\left(p(q_{cd}, q_d, \hat{q}_c, \hat{\delta}) - cd\right) q_{cd} - T_{cd}\right]^{1-\beta} T_{cd}^{\beta}$. As d takes the contracts offered to other importers as given, the optimal payment is $\hat{T}_{cd} = \left[\beta p(q_{cd}, q_d, \hat{q}_c, \hat{\delta}) + (1-\beta)cd\right] q$ which ensures the division of gross surplus is proportional to the bargaining weights.[5]

Substituting for the optimal payments, d chooses to sell quantities that maximize

$$\max_{q_{cd}, q_d} (1-\beta) \int_{c_{\min}(d)}^{c_{\max}(d)} \left[p(q_{cd}, q_d, \hat{q}_c, \hat{\delta}) - cd\right] q_{cd} dG_c.$$

The profit function shows that an importer ignores how his quantity affects the profit of other importers of exporter c's product (through q_c) but internalizes the cannibalization effect of his own quantities on each other (through q_d). The optimal final price of variety cd under bilateral contracts is

$$p_{cd}^{B} = cd / \left(1 - \mu_{v_{cd}} - \frac{(1-\theta)u'(q_d)}{\underbrace{\theta u'(q_c)}_{\text{Business Stealing}} + (1-\theta)u'(q_d)} \mu_u(q_d) \varepsilon_{v_{cd}} \right) \equiv cd / \left(1 - \mu_{v_{cd}} - (1-\theta)\tilde{\mu}_d\right). \tag{3}$$

This gets rid of the double marginalization problem because exporters do not charge unit prices that exceed their unit costs. In fact, optimal final quantities correspond to a two-part tariff where the exporters do not mark up their costs ($p_c^x = c$) and extract part of the surplus through payments T_{cd} from importers. We summarize this result in Proposition 1. The exporter therefore cannot force the "monopoly" price that maximizes multilateral profits.

Proposition 1. *Bilateral private contracts ensure lower prices than anonymous market transactions because they eliminate double marginalization.*

Bilateral contracts overcome double marginalization, but not business stealing as shown in Equation 3. The next sub-section discusses joint contracts that maximize multilateral profits from an exporter's product.

2.6 Joint Contracts

By partnering with her importers, an exporter can ensure that profits from her product are maximized by internalizing the business stealing effect imposed by importers on each other. Joint

[5] Using experimental data, Martin et al. (2001) show that the ability to reject an upstream firm's offer enables the downstream firm to get a positive share of the surplus from the relationship. In an alternative version of the model, we show that our results hold when the exporters and importers can replicate the other party's role in the relationship to some degree. This follows from the drop in delivered costs of exports after a trade liberalization which raise the ex-post bargaining share for exporters.

contracts provide higher prices, but involve fixed investments in building a relationship. This is because exporters need to demonstrate that they are committed to restricting quantities to their importers. We start with a discussion of the surplus division within the joint contract and then determine the optimal prices and the contracting decision.

2.6.1 Importer Payments

When the exporter negotiates jointly with her importers, the importer's profit function is the same as earlier but now the exporter's total quantity is no longer taken as given. The importers observes the exporter's quantity, and chooses q_{cd} to maximize $\pi_d^J = \int_{c_{\min}(d)}^{c_{\max}(d)} \left[\left(p(q_{cd}, q_d, q_c, \hat{\delta}) - cd\right) q_{cd} - T_{cd}\right] dG_c$. The split of profits and hence the payments T_{cd} are again determined by Nash bargaining where the exporter's bargaining weight is β and the disagreement payoffs are again zero. The optimal payment is once more $\hat{T}_{cd} = \left[\beta p_{cd}(q_{cd}, q_d, q_c, \hat{\delta}) + (1-\beta)cd\right] q_{cd}$.

Substituting for the optimal payments and summing over all importers, the exporter chooses quantities to maximize:

$$\max_{q_{cd}, q_c} \beta \int_{d_{\min}(c)}^{d_{\max}(c)} \left[p(q_{cd}, q_d, q_c, \hat{\delta}) - cd\right] q_{cd} dG_d.$$

Joint contracts enable importers to internalize the business stealing effect which implies the optimal price is

$$p_{cd}^J = cd/\left(1 - \mu_{v_{cd}} - (1-\theta)\tilde{\mu}_d - \theta\tilde{\mu}_c\right) \qquad \tilde{\mu}_c \equiv \frac{u'(q_c)}{\theta u'(q_c) + (1-\theta)u'(q_d)} \mu_u(q_c)\varepsilon_{v_{cd}}. \tag{4}$$

Prices are set at the profit-maximizing "monopoly" levels. The exporter ensures prices are higher than under bilateral contracts because importer competition is reduced. But under the earlier condition on elasticities $\mu_u\varepsilon_v < \min \mu_v \cdot \min(1-\mu_v)$, the joint contract price is lower than the price from anonymous markets because double marginalization is avoided. The optimal quantity sold to consumers corresponds to the monopoly quantity in the final goods market and is therefore lower than the quantity supplied under bilateral contracts where importers compete with each other. We summarize this result in Proposition 2.

Proposition 2. *Joint contracts eliminate double marginalization and business stealing in the final goods market, leading to $p^M > p^J > p^B$.*

We are interested in the impact of trade liberalization on firm's ability to engage in richer contracts and will examine the qualitatively new results of the importing microstructure on the division of the gains from trade. The business stealing externality arises when importers' varieties of an exporter's products are more substitutable than the exporters' products. Then as firms move from bilateral contracts to joint contracts, they increase the prices charged to consumers. The opposite happens - prices fall - when firms switch from anonymous transaction to bilateral contracts. We therefore focus on the case of exporters in the upper nest and the choice between bilateral contracts and joint contracts because the other cases reinforce the standard gains from trade by reducing prices further. In the remainder of this Section, we derive the market equilibrium and then discuss the implications of opening to free trade in the next Section.

2.7 Equilibrium

We start with specifying the equilibrium of a closed economy. As we are interested in the division of the gains from trade, the mass of firms is fixed so that they earn positive profits. Assuming a symmetric equilibrium, we can write $p_{cd} = p$ and $q_{cd} = q$ for brevity. Let $m \equiv \mu_v(q) + 1_J \mu_u(q_c)\varepsilon_v(q)$ denote the markup charged in the final goods market, where 1_J is 1 under joint contracts and 0 otherwise. From optimal pricing, $p = cd/\left(1-m\right)$ where the inverse demand is $p = u'\left(M_d v(q)\right) v'(q)/\delta$ and the budget constraint is $M_c M_d pq = 1$ (which gives the consumer's budget multiplier δ). Then the optimal contracting decision is to invest in joint contracts if $M_d \beta \left(\pi^J_{cd} - \pi^B_{cd}\right) = M_d \beta \left[\frac{m^J}{1-m^J} q^J - \frac{\mu_v}{1-\mu_v} q^B\right] cdL \geq f^J_c$, where the superscript J denotes the quantity implied under joint contracts and B denotes the quantity implied under bilateral contracts from the optimal pricing condition. The importer always finds the joint contracting decision viable when this condition holds as long as his relative contracting costs are low $(f^J_d/M_c(1-\beta) < f^J_c/M_d\beta)$. The four conditions - optimal pricing, the inverse demand, the budget constraint and the optimal contracting decision - summarize the market equilibrium under fixed entry.

2.8 Opening to International Trade

When the economy opens up to free trade, all producers and distributors engage in international trade. We show that looking at the microstructure of import markets reveals new sources of gains from trade, beyond the usual gains from variety and pro-competitive effects. Opening to trade makes joint contracts more profitable, and exporters and importers become more likely to enter into joint contracts. Consumers still experience the usual gains from access to foreign varieties, but they do not get the full benefit of openness on prices because firms reduce competition in the import market through joint contracts. Firms gain at the expense of consumers, and later we show that these forces are present even under free entry of firms.

Opening to free trade increases the mass of consumers available to producers from L to sL where s is the scaling factor for the increase in consumers from trade. The equilibrium conditions for optimal pricing and inverse demand are unchanged. The budget constraint now accounts for the fact that consumers have access to $sM_c > M_c$ producers and their budget constraint is $sM_c M_d pq = 1$. The optimal contracting decision of firms is to invest in joint contracts if $sM_d \beta \left(\pi^J_{cd} - \pi^B_{cd}\right) \geq f^J_c$.[6]

[6] The contracting cost can also be flexibly specified as $f^J_c(s)$ such that $d \ln f^J_c(s)/d \ln s < 0$ so that there is some economies of scale in contracting.

As long as profits are supermodular in markups and market size, international trade makes joint contracts more likely. . We specify the supermodularity condition in Assumption 2.8 below and provide the conditions on primitives in the Appendix.

Assumption 1. Profits $\pi_c(m,s)$ are supermodular in markups m and market size s.

CES demand ensures profits are modular and we get the knife-edge case where profits rise at the same rate for all profit levels. This implies that profits rise at the same rate under joint contracts and bilateral contracts, so that opening to trade has no impact on the likelihood of entering into joint contracts (that provide higher profits). To overcome this kinfe-edge result, we assume that the rate of change of markups is such that profits are supermodular in markups and market size. Then profits rise more after free trade trade for higher profit levels, as is standard in most trade models with variable markups. The reasoning behind Assumption 2.8 is as follows. As market size expands, the direct impact is to increase profits which are greater under joint contracts. The indirect impact is that profits earned from an individual consumer decline due to competition. This profit drop is smaller when markups are higher and firms have greater ability to absorb the negative profit effect of import competition. But the rate of change in markups differs across contracts. The drop in profits is bigger when the rate of change in markups is greater. The rate of change in markups under joint contracts includes the rate of change in own markups $\mu'(q)$ and the rate of change in markups from internalizing business stealing $(\mu_u\varepsilon)'$. The supermodularity condition therefore ensures that the rate of change in markups from business stealing is small enough to not completely mitigate the direct impact of market expansion and the indirect impact of higher markups. Having specified the supermodularity condition, we discuss the impact of trade on the division of the gains from trade under fixed entry, and then show that these results are robust to free entry where profits are wiped away.

2.8.1 The Division of the Gains from Trade

Opening to trade gives foreign exporters access to home consumers, so the size of the market expands. Consumers get access to imported varieties and the change in welfare is $\Delta \ln U = \Delta \ln M_c + \Delta \ln u(M_d v(q))$ where Δ denotes changes. Under fixed entry, the change in variety is $sM_cM_d - M_cM_d$ as there is no feedback effect into entry of firms. Quantities per variety q change due to the usual forces of import competition and the new source of changes in contracting choice. We decompose the change in consumer welfare into the standard gains from variety and pro-competitive effects, and the new gains from changes in the microstructure of import markets.

Let $\Delta_s x(s) \equiv x(s) - x(1)$ denote the change in outcome x when the market size rises from L to sL, under a fixed microstructure. Let $\Delta_V x^V(s) \equiv x^J(s) - x^B(s)$ for $V \in \{B.J\}$ denote the change in outcome x when firms move from bilateral contracts to joint contracts for a fixed level of market size s. Then the change in consumer welfare can be decomposed as:

$$\begin{aligned} \Delta \ln U = & \underbrace{\Delta_s \ln sM_c u\left(M_d v(q^B(s))\right)}_{\text{Gain from Variety \& Pro-Competitive Effects}>0} \\ + & \underbrace{\Delta_V \ln u\left(M_d v(q^V(s))\right)}_{\text{Gain from Change in Microstructure}<0} \end{aligned} \tag{5}$$

The first line of Equation 5 shows the standard gains from trade for consumers for a given contracting form. Entry of foreign varieties provides higher welfare through Gains from Imported Variety and Gains from Pro-Competitive Effects. Prices fall due to foreign competition because $d\ln p/d\ln s = -1 + 1/\left(1+\mu_v' q/(1-\mu_v)\right) < 0$ and consumers get access to foreign varieties so that welfare increases by $d\ln sM_c u\left(M_d v(q^B(s))\right)/d\ln s = 1 - \varepsilon_u \varepsilon_v/\left(1+\mu_v' q/(1-\mu_v)\right) > 0$.

The second line shows a new source of change in consumer welfare arising from changes in the microstructure of the import market. Increase in the market size available to producers makes joint contracts more likely.

This change in the microstructure of the import market increases the price of a variety. Combining the budget constraint, the demand function and the pricing condition, optimal quantity is determined by $\left(1-\mu_v(q) - 1_J \mu_u\left(M_d v(q)\right)\varepsilon_v(q)\right)/q = sM_c M_d cd$. The LHS is decreasing in quantity, and is smaller under joint contracts because firms account for the business stealing effect through $\mu_u \varepsilon_v$. Firms reduce their quantities under joint contracts and prices rise (because $p = u'\left(M_d v(q)\right) v'(q)/\delta = 1/sM_c M_d q$). The consolidation of the import market therefore lowers the gains from trade that are passed on to consumers.

Total profits for firms in a country are $\Pi = sM_c M_d \pi_{cd}^V - M_d f_d - M_c f_c - 1_J M_c f_c^J$, and the change in profits after opening to trade can be decomposed into the standard pro-competitive effects from foreign varieties and the new gains from the change in the microstructure of the import market. Firms gain access to more consumers but also face competition from imports which leads to lower profits, as shown in the first line below. The second line is the rise in profits when firms move from bilateral contracts to joint contracts. Firms gain at the expense of consumers, as the import market is consolidated.

$$\begin{aligned} \Delta \Pi = & \underbrace{\Delta_s \left(sM_c M_d \pi_{cd}^B(s) - M_d f_d - M_c f_c\right)}_{\text{Gains from Market Access and Pro-Competitive Effects}<0} \\ + & \underbrace{\Delta_V \left(sM_c M_d \pi_{cd}^V(s) - 1_J M_c f_c^J\right)}_{\text{Gains from Change in Microstructure}>0} \end{aligned}$$

The first line follows from the drop in markups after opening to trade. Substituting for the inverse demand and the budget constraint, the optimal quantity sold to each worker is $(1-\mu_v)/q = sM_c M_d cd$, and foreign competition lowers the quantity by $d\ln q/d\ln s = -1/\left(1+\mu_v^{'}q/\left(1-\mu_v\right)\right)$. Markups therefore fall by $d\ln\mu_v/d\ln s = \left(\mu_v^{'}q/\mu_v\right)\left(d\ln q/d\ln s\right)$ and the profit from home consumers decreases by $d\ln s\pi/d\ln s = -\left(\mu_v^{'}q/\mu_v\right)\left(1-\mu_v\right)/\left(1-\mu_v+\mu_v^{'}q\right) < 0$. The second line is the gain in profits from investing in joint contracts, which is positive for all switches because firms internalize business stealing.

We summarize the results for the division of the gains form trade in Proposition 3, and relegate details of the proof to the Appendix.

Proposition 3. *Under Assumption 2.8, opening to trade makes joint contracts more likely. This change in the microstructure of import markets increases firms profits at the expense of consumer welfare.*

2.8.2 The Division of the Gains from Trade under Free Entry

We will also show that the qualitative results for consumer welfare persist under free entry of firms when profits are wiped away. For simplicity, we specify a CES upper nest $u(q_c) = q_c^{\rho}$ and examine the free entry equilibrium. When firms can enter freely, there will be two additional conditions for equilibrium in the economy. Distributors can enter freely and this will drive down their profits net of entry costs to zero, $sM_c\left(1-\beta\right)\pi_{cd}^{V} = f_d$ for the optimal contract $V \in \{B,\, J\}$. Likewise, producers enter till their profit margin net of entry costs is driven down to zero, $sM_d\beta\pi_{cd}^{V} = f_c + 1_J f_c^{J}$.

Opening to trade increases the size of the market available to producers. As earlier, the change in consumer welfare can be decomposed into the standard gains from trade and the gains arising from the change in the microstructure of markets:

$$\begin{aligned}\Delta\ln U = \quad & \underbrace{\Delta_s \ln sM_c u\left(M_d v(q^B(s))\right)}_{\text{Gain from Variety \& Pro-Competitive Effects}>0} \\ + \quad & \underbrace{\Delta_V \ln u\left(M_d v(q^V(s))\right)}_{\text{Gain from Change in Microstructure}<0}\end{aligned} \tag{6}$$

The first line of Equation 6 shows the standard gains form trade for consumers for a given contracting form. Entry of foreign varieties provides higher welfare through Gains from Imported Variety and Gains from Pro-Competitive Effects. Although the mass of importers per market falls, consumers get access to new foreign varieties and firms charge lower markups. And the net effect is an increase in consumer welfare from the standard gains from trade. The second line, as earlier, is the impact of trade through the microstructure of the import market. Consumers lose out as firms switch from bilateral to joint contracts. Although more importers enter, there is a drop in the mass of producers because they need higher scale to justify the bigger fixed investments. The rise in markups from the switch to joint contracts and the fall in producer entry overwhelm the entry of importers, and consumer welfare falls because of the change in the microstructure. Market expansion makes joint contracts more likely, under the supermodularity condition of Assumption 2.8. Firm profits are unchanged because of free entry. Therefore, the aggregate impact of opening to trade on welfare is positive through the standard channels of variety and pro-competitive effects, but negative through the new channel of market microstructure. Details are in the Appendix.

Having discussed the division of the gains from trade from changes in the microstructure of markets, we proceed to testing the observable implications of the model. As we will work with firm-level data during an episode of tariff cuts, the next Section will introduce firm heterogeneity and trade costs.

3 Firm-Level Predictions

Section 2 highlights how the microstructure of importing affects the division of the gains from trade. Ideally we would test the model by examining how the microstructure changes between exporters and importers in a period of falling trade costs. Data on contracting choices are not available in standard datasets. Further, informal arrangements can replicate bilateral contracts and joint contracts, as discussed earlier. So, we focus on the unique predictions of the model for observable outcomes - prices, quantities and importer-exporter matches. In this section, we start by incorporating differences in importer cost cutoffs across exporters. We then provide "difference-in-difference" predictions for changes in prices, quantities and the number of importers per exporter under different contracts.

3.1 Microstructure of Import Markets

Continuing with the framework of Section 2, we examine a model with heterogeneous exporters and importers. For simplicity, we assume that there is an outside good q_0 that is freely traded and produced one for one with labor. As is well-known, the outside good mitigates income effects and we focus on the price effects arising from changes in the microstructure of the import market for differentiated varieties. The welfare function is $W = q_0 + Q^\eta$ for $0 < \eta < 1$, where the differentiated varieties are $Q = U = \int u(q_j)dj$ and $q_j = \int v(q_{ij})di$. For exporters in the upper nest and importers in the lower nest, the inverse demand is $p_{cd} \equiv Q^{\eta-1}u'(q_c)\nu'(q_{cd})$ which is similar to Section 2. We will look at reductions in tariffs and specify $\tau > 1$ as the ad valorem trade cost incurred on payments made to exporters. Specifically, importer d pays τT_{cd} to exporter c, who receives just T_{cd}.

Firms differ in the unit costs c and d which are drawn from cumulative densities $G_c(c)$ and $G_d(d)$. As earlier, contracts specify quantities and payments and the surplus from a variety is split through bilateral Nash bargaining. As earlier, firms chose quantities that maximize bilateral profits under bilateral contracts and that maximize multilateral profits under joitn contracts. Then the optimal final price for variety cd is $p_{cd} = cd/\left(\mu_v(q_{cd}) + 1_J\mu_u(q_c)\bar{\varepsilon}_{v_c}\right)$ where $\bar{\varepsilon}_{v_c} \equiv \int_0^{d_m} \varepsilon_v(q_{cd}) \frac{v(q_{cd})}{\int_0^{d_m} v(q_{cd})dG_d} dG_d$.

As firms differ in costs, we introduce matching costs because this will generate differences in the range of exporters and importers that transact with each other. Exporters and importers face a matching cost $f_{cd} > 0$.[7] So matches will need to be productive enough to justify transactions, and this leads to an endogenous cost cutoff for import partners of each exporter, which we can take to the data. Exporter c supplies to all importers with $d \leq d_m^V(c)$ for $V \in \{B, J\}$. The bilateral profit from a match is $\pi_{cd}\left(q_{cd}, \hat{q}_c, \hat{Q}\right) - f_{cd}$ so exporter c sells to all importers with variable profits greater than the fixed cost of matching. Importers in joint contracts account for the business stealing effect of their actions, and this is reflected in the optimal importer range decision as $\pi_{cd_m}\left(q_{cd_m}, q_c, \hat{Q}\right) + \left(\partial\pi_{cd_m}/\partial q_c\right)\left(\partial q_c/\partial d_m\right) = f_{cd}$. Substituting for the change with respect to upper nest quantities, the optimal importer cost cutoff is given by $\left(p_{cd_m} - cd_m\right) q_{cd_m} - 1_J p_{cd_m} q_{cd_m} \mu_u(q_c) \bar{\varepsilon}_{v_c}/\varepsilon_v(q_{cd_m}) = f_{cd}/L$.

As firms move from bilateral contracts to joint contracts, they extract a higher markup by selling fewer units. Consumer prices rise, total quantities of an exporter fall and the importer cost cutoff of exporters falls. Exporters switching from bilateral to joint contracts therefore consolidate their import market to extract more of the consumer surplus. As is well-known, lower trade costs reduce prices directly. The new finding is that the price response to trade liberalization differs across the contractual choice of exporters and importers. Exporters and importers moving from bilateral contracts to joint contracts increase the final markup charged to consumers by overcoming business stealing among importers. The import price paid by the importer to the exporter also rises. Let p_{cd}^x denote the unit "price" paid to exporter c by importer d. In anonymous markets, there is a straightforward market-clearing price per unit paid by the importer. Under bilateral contracts, the unit price is the variable component of the payments to exporters, which rises as firms extract more of the consumer surplus.[8] Having incorporated the extensive margin of importers, we can determine the change in the range of importers chosen across different contracts. Exporters that switch to joint contracts lower their importer range because this reduces business stealing and increases final prices. For the same reason, these exporters also scale back on their quantities. Therefore, moving to joint contracts results in a consolidation of the import market for the product. The import price is increased, quantities are reduced and fewer importers carry the variety under joint contracts.

Joint contracts become more likely after trade liberalization when the profit function $\left(\pi_c\left(m, -\tau\right)\right)$ is supermodular in markups and openness. This requires a condition similar to the one in Section 2 with modifications to reflect the cost distribution. We summarize this assumption and the comparison between bilateral contracts and joint contracts in Proposition 4 below and provide details in the Appendix.

[7] Bernard et al. (2014) introduce importer-specific fixed costs in a trade model with heterogeneity of both exporters and importers. They focus on the implications of variation in importer heterogeneity across destination markets and do not model contracting choice.

[8] In an earlier working paper with CES demand, we show that these predictions persist when disagreement payoffs are non-zero because the ex-post share of exporters under joint contracts rises after trade liberalization.

Proposition 4. *When exporter profits are supermodular in markups and openness, exporters switching from bilateral contracts to joint contracts increase their prices, reduce their sales, and sell to fewer importers.*

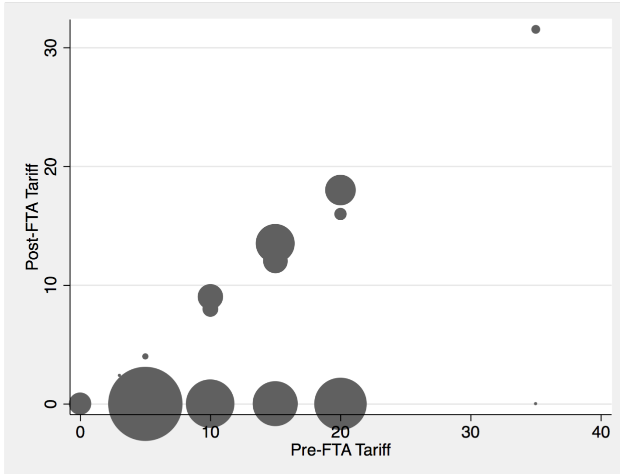
To understand the impact of trade liberalization through the microstructure, we examine these unique observable outcomes of higher import prices, lower imports and fewer import partners in response to a reduction in tariffs. Import prices and quantities are routinely observed in customs data, so we can use the above predictions to examine the main mechanism of the theory. The main prediction is that import prices rise after trade liberalization for varieties whose exporters reduce their importer cost cutoffs and import quantities. The prediction that trade liberalization induces exporters to simultaneously increase prices, reduce quantities and reduce importers does not arise in standard trade models. While the possibility of quality upgrading can lead to increased prices after trade liberalization, there is no clear theoretical prediction in the quality literature on how quantity varies with changes in quality (e.g. Eckel et al. 2015). Incorporating quality and quantity in different ways, Baller (2013) shows that standard trade models would predict a rise in quantities when there are economies of scale in quality upgrading. Going by this and ruling out other ways in which quality could be confounding our results, we will provide evidence for Proposition 4.

To directly test for the mechanism in the theory, the reduction in the number of import partners in Proposition 4 will be important for the empirical application. Recent models incorporating importer margins of trade also predict that trade liberalization induces exporters to increase their importer range as a result of higher export profitability. The range of import partners would rise rather than fall after trade liberalization in these models. Indeed, this is important because a fall in the number of importers is a direct measure of our theoretical mechanism that competition falls among importers of the product. We therefore test for the triple prediction of higher prices, lower quantities *and* fewer importers to isolate the quality upgrading effect from the contracting effect of trade liberalization. The subsequent section operationalizes the triple prediction by examining prices, quantities and importer matches following a major trade liberalization episode in Colombia.

4 Empirics

The microstructure effect of trade liberalization increases the probability that an exporter switches from using bilateral contracts to joint contracts. This implies an increase in the prevalence of the triple prediction of increased price, reduced quantities and fewer importers for exporters experiencing trade liberalization. We examine the empirical relevance of the triple prediction using data on Colombian imports after the implementation of the US-Colombia Free Trade Agreement (FTA). This Section contains a discussion of the data, the baseline empirical specification, and a series of robustness checks.

Figure 1: Colombian Tariffs on US Products in 2011 (Pre-FTA) and 2012 (Post-FTA)



Note: Size of the marker indicates the number of HS8 products.

4.1 Data

We examine the triple prediction of higher prices, lower quantities and fewer matches for imports into Colombia before and after the implementation of the US-Colombia FTA. Transaction-level import data for Colombia identifies the exporter and the importer for each import transaction as well as the total value, quantity, date and product. Aggregating the transaction-level import data to the exporter-product pair enables us to obtain the average import price of the product, the total quantity of the product shipped by the exporter to Colombia, X_{csht}, and the number of Colombian importers per exporter-country-product $d^{\max}_{csht}$.

While matched exporter-importer data is available for other countries, the Colombian import data is unusual because it covers a major trade liberalization with the largest trading partner. The US-Colombia FTA was implemented on May 15, 2012. The agreement immediately eliminated duties on 80 percent of US exports of consumer and industrial products to Colombia. Figure 1 shows the Colombian tariff rates charged to the US before and after the US-Colombia FTA. Most US exports obtain duty-free status after the FTA. Table 1 provides summary statistics for Colombian tariffs faced by US exporters and a comparable set of exporters from developed countries before and after the FTA. The pre-FTA tariff for 2011 fell from an average of 11 percentage points to an average of 4 percentage points in our sample. The tariff elimination was largely on the Colombian side because 90 percent of US imports from Colombia enjoyed duty free access before the FTA. The US is Colombia's leading trade partner and the FTA was a substantial step towards trade liberalization in Colombia. Between June 2012 and April 2013, US exports to Colombia increased by 14.2 percent over the same period a year earlier while total Colombian imports were up only 4.6

percent.

Table 1: Colombia Tariffs Before and After the FTA for Exporters in 2009

Variable	Obs.	Mean	S.D.	Min	Max
Tariffs Charged to US Exporters					
Pre-FTA Tariff in 2011	1,750	11.05	5.54	0	35
Post-FTA Tariff in 2012	1,750	4.27	6.37	0	31.5
Tariffs Charged to Non-US Developed Country Exporters					
Pre-FTA Tariff in 2011	1,747	6.57	3.52	0	30
Post-FTA Tariff in 2012	1,747	6.82	3.56	0	30

Note: Mean refers to an average across exporters in the sample, weighted by the initial value of Colombian imports from developed countries for each HS code in 2009.

The theoretical predictions of the model focus on the simultaneous rise in prices and drop in quantities and import partners. An exporter, c, increases the import price of its product ($\tilde{m}_c$), reduces its total quantity (q_c) and reduces the number of its importers ($d_c^{\max}$). We define the triple prediction as $\Delta\text{Triple}_{csht} = 1$ if exporter c from source country s selling product h at time t increases its average price across all its importers ($\triangle \ln \text{Import Price}_{csht} > 0$) *and* reduces the total quantity sold to all its importers ($\Delta X_{csht} < 0$) *and* reduces the number of importers of its product ($\Delta d_{csht}^{\max} < 0$). If any one of these events does not happen, then $\Delta\text{Triple}_{csht} = 0$.

To test the theoretical results, we need measures of $\Delta\text{Triple}_{csht}$ and the treatment variables. We use Colombian import transactions data recorded by its customs authority, which lists the name of the importer and the exporter for each import transaction.[9] We clean the names by harmonizing commonly occurring prefixes and suffixes and then aggregating imports to the exporter-country-product level in each time period ($csht$). A time period consists of the months from June to November for each year (2009-2012).[10] For each time period, import values and quantities are recorded at the 8-digit level under the Colombian implementation of Harmonized System (HS) classification. We compute the average import price (unit value in USD), the total quantity, and the total number of importers for each exporter-country-product-year, $csht$. We then compute $\Delta\text{Triple}_{csht} = 1$ for exporter-country-products that have a higher average price, lower total quantity and fewer importers compared to the previous period.

As customs data are known to be noisy, we typically separate the set of exporter-country-products with initial sales of less than US$10,000 and only include the set of country-product pairs with more than one exporter in 2009. The theory applies to incumbent exporters, so we focus on

[9] The raw data come from www.importgenius.com.

[10] The interval starts in June since the FTA came into force in the middle of May 2012. The period ends in November since data for December 2011 are missing. We also do not have data for July 2009 and quantities have been appropriately scaled to account for this. Results are robust to using 2010 as the initial period (available upon request).

exporters who sell from 2009 to 2012. Pierce and Schott (2012) show that HS codes change over time and this can lead to estimation bias. To account for this, we work with the set of products with HS codes that are unchanged between 2009-2012. This covers 84 percent of all products in our sample, alleviating concerns regarding generality of the results.

To construct the treatment variables, tariff data for the US-Colombia FTA is obtained from the Office of the US Trade Representative. Tariffs were reduced from their initial levels by one of four FTA multipliers: 0% of the initial tariff, 80% of the initial tariff, 87% of the initial tariff or 90% of the initial tariff. Table 2 summarizes the distribution of HS-codes by the FTA tariff multiplier. We classify $\text{Treat}_h = 1$ for product codes with a Tariff multiplier equal to zero. For all other product codes, $\text{Treat}_h = 0$.

Table 2: US-Colombia FTA Tariff Multipliers for Colombian Non-agricultural Tariffs

Tariff Multiplier	# of Products	% of Products
0	4,560	76.18
0.80	317	5.30
0.86	52	0.87
0.90	1,057	17.66

Note: Tariff Multiplier = Post-FTA Tariff/ Pre-FTA Tariff. Post-FTA Tariff is the tariff on a product from the US in the year after the implementation of the FTA

4.2 Baseline Empirical Specification

We examine whether the US-Colombia FTA induced US exporters to increase prices, reduce quantities and reduce the number of importers in Colombia. To control for underlying trends, we examine whether exporters from the US selling products that started to receive duty-free access into Colombia were more likely to increase their prices, reduce their quantities and reduce their number of importers relative to the previous period. We define $\text{Post}_t = 1$ for the period after the FTA and 0 for the period before the FTA. In addition, we compare the probabilities for US exporters to a control group of exporters that did not experience the FTA tariff reduction. We focus on exporters from developed countries to construct a suitable control group. Accordingly, $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from any other developed country.

Table 3 summarizes the triple dummy and exporter characteristics for exporters in the control and the treatment groups separately for 2012. The triple prediction is more prevalent for the treatment group, but this could be due to differences in product composition or due to other events specific to the post-FTA time period. In order to minimize these concerns, we proceed to a difference-in-difference estimation which accounts for product-specific and country-specific effects.

Table 3: Summary Statistics for Exporter-Products in 2012

Control Group: Exporter-Products with $\text{Treat}_h \cdot \text{USA}_s = 0$

Variable	Obs ($csht$)	Mean	S.D.	Min	Max
$\Delta\text{Triple}_{csht} = 1$	17,547	0.044	0.205	0	1
Initial trade $\text{value}_{csh,2009}$ ($mn)	17,547	0.054	1.298	<0.001	126
Initial number of $\text{importers}_{csh,2009}$	17,547	1.314	0.897	1	19
Size_{csh}	17,373	0.004	0.513	-1.593	6.522

Treatment Group: Exporter-Products with $\text{Treat}_h \cdot \text{USA}_s = 1$

Variable	Obs ($csht$)	Mean	S.D.	Min	Max
$\Delta\text{Triple}_{csht} = 1$	6,225	0.064	0.245	0	1
Initial trade $\text{value}_{csh,2009}$ ($mn)	6,225	0.107	1.101	<0.001	60.1
Initial number of $\text{importers}_{csh,2009}$	6,225	1.533	1.703	1	54
Size_{csh}	6,116	0.052	0.665	-1.485	5.911

Note: $\Delta\text{Triple}_{csht} = 1$ if $\triangle ImportPrice_{csht} > 0$ & $\triangle ImportQuantity_{csht} < 0$ & $\triangle \# Importers_{csht} < 0$ from period $t-1$ to t and 0 otherwise. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product for all developed countries in 2009.

We examine the prevalence of simultaneous increases in prices, reductions in quantities and reductions in the number of importers of an exporter. The estimating equation for the triple prediction for exporter c from source country s selling product h at time t is a linear probability model,

$$\Delta\text{Triple}_{csht} = \beta \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s + \gamma X_{csht} + \alpha_{st} + \alpha_{ht} + \varepsilon_{csht} \tag{7}$$

where ε_{csht} is a disturbance term while α_{st} and α_{ht} are source country-year and product-year fixed effects that account for changes such as exchange rate fluctuations and aggregate demand shocks. X_{csht} includes all interactions between Post_t, Treat_h, and USA_s, and other controls such as pre-trends and their interactions (where $\text{Pre}_t = 1$ for observations in 2011 and 0 otherwise). The coefficient of interest is β which we expect to be positive if the FTA led exporters to consolidate their import market resulting in higher import prices, lower import quantities and fewer importers.

From the theoretical results in Section 3, we expect the triple prediction to vary across exporters of different levels of productivity. High productivity exporters are expected to have already paid the costs of consolidating their import market, so bilateral trade liberalization would lead to consolidation in the import markets of less productive exporters from the US. To account for differences in responses across exporters, we allow the coefficient on $\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$ to vary with exporter size. For each exporter-country-product observation, exporter size is measured by the z-score of the initial value of sales of the exporter in 2009, relative to all exporters of the product from developed countries in 2009. Let $e_{csh,2009}$ denote the initial value of sales of exporter c from country s of product h in 2009. The z-score of initial size is $\text{Size}_{csh} \equiv (e_{csh,2009} - \mu_{h,2009}) / \sigma_{h,2009}$ which measures the initial sales of an exporter relative to exporters of that product from all developed countries.[11] Accounting for possible differential responses across firms, the estimating equation is

$$\Delta\text{Triple}_{csht} = \beta \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s + \beta_1 \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s \cdot \text{Size}_{csh} + \gamma X_{csht} + \alpha_{st} + \alpha_{ht} + \varepsilon_{csht} \quad (8)$$

where X_{csht} includes interactions of Size_{csh} with Post_t, Treat_h, USA_s and Pre_t as well. Coefficient β_1 allows the impact of the FTA liberalization to vary by initial size of the exporter. We cluster standard errors for each estimating equation by exporter-country pairs to account for correlation across time and products within an exporter-country pair.

4.3 Baseline Results

Table 4 summarizes the results from estimation of the linear probability models in Equations 7 and 8.

4.3.1 Baseline Specification

Column (1) shows that exporters who experienced duty-free access from the US-Colombia FTA are more likely to increase their average price, reduce their total quantities and sell through fewer importers, relative to a control group of developed country exporters who did not experience duty-free access for the product. The likelihood of a triple for the treated exporters rises 1.5 percentage points more after the FTA relative to the change for other exporters.

4.3.2 Triple Prediction by Exporter Type

Column (2) adds the interaction with the initial size of the exporter. From the theory, the sign of the interaction term is expected to be negative as less productive exporters are more likely to consolidate their import market after the FTA. The results in Column (2) show that exporters with smaller initial size are indeed more likely to consolidate their import market. Figure 2 plots the histogram of the estimated impact of the FTA across all exporters, $\hat{\beta}_0 \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s + \hat{\beta}_1 \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s \cdot \text{Size}_{csh}$. Most of the exporters from the US have an increase in the probability of higher prices, lower quantities and fewer partners as a result of the liberalization.

Another way of capturing the differences in the triple prediction is to estimate the treatment coefficient separately for exporters of different sizes. Instead of using a continuous measure of initial size, we estimate different coefficients for exporters by their size category. We divide all exporters of a product into three different terciles of Size_{csh} and estimate separate coefficients on $\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$ for each tercile. $\text{Tercile1}_{csh} = 1$ if exporter c's initial size measured by Size_{csh} is below the 33rd percentile of all exporters of that product from developed countries, and 0

[11] $\mu_{h,2009}$ is the mean of $e_{csh,2009}$ across all exporters of product h from developed countries and $\sigma_{h,2009}$ is the corresponding standard deviation.

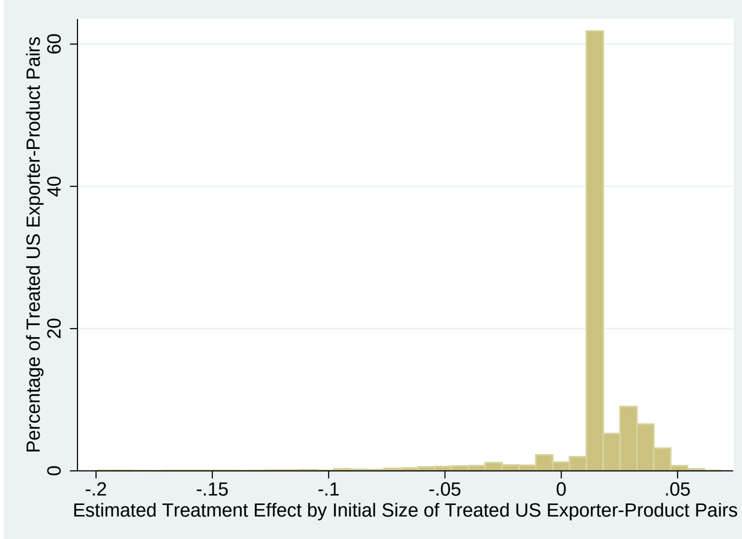
Table 4: Baseline results: Triple Prediction

Dependent Variable: $\Delta\text{Triple}_{csht}$

Variable	**(1)** **Coef.** **(S.E.)**	**(2)** **Coef.** **(S.E.)**	**(3)** **Coef.** **(S.E.)**	**(4)** **Coef.** **(S.E.)**
$\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.015**	0.015**	0.053*	0.015**
	(0.007)	(0.007)	(0.028)	(0.007)
$\text{Pre}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.004	0.004	0.047	0.004
	(0.008)	(0.007)	(0.032)	(0.008)
$\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s \cdot \text{Size}_{csh}$		-0.036**	-0.092***	-0.036**
		(0.017)	(0.032)	(0.016)
$\text{Pre}_t \cdot \text{Treat}_h \cdot \text{USA}_s \cdot \text{Size}_{csh}$		0.004	0.002	0.004
		(0.017)	(0.050)	(0.015)
$\text{Treat}_h \cdot \text{USA}_s$	-0.002	-0.002	-0.022	-0.002
	(0.005)	(0.004)	(0.019)	(0.004)
$\text{Treat}_h \cdot \text{USA}_s \cdot \text{Size}_{csh}$		0.005	0.027	0.005
		(0.010)	(0.024)	(0.009)
$\text{Post}_t \cdot \text{Treat}_h \cdot \text{Size}_{csh}$		0.029***	0.072***	0.030***
		(0.011)	(0.027)	(0.011)
$\text{Post}_t \cdot \text{USA}_s \cdot \text{Size}_{csh}$		0.021**	0.051**	0.021*
		(0.013)	(0.020)	(0.011)
$\text{Post}_t \cdot \text{Size}_{csh}$		-0.007	-0.019	-0.007
		(0.007)	(0.017)	(0.007)
$\text{Treat}_h \cdot \text{Size}_{csh}$		-0.009	-0.036*	-0.009
		(0.007)	(0.020)	(0.007)
$\text{USA}_s \cdot \text{Size}_{csh}$		-0.001	-0.019	-0.001
		(0.007)	(0.014)	(0.008)
$\text{Pre}_t \cdot \text{Treat}_h \cdot \text{Size}_{csh}$		-0.0003	-0.004	-0.003
		(0.001)	(0.025)	(0.011)
$\text{Pre}_t \cdot \text{USA}_s \cdot \text{Size}_{csh}$		-0.011	0.0003	-0.010
		(0.015)	(0.040)	(0.013)
$\text{Pre}_t \cdot \text{Size}_{csh}$		0.009	0.024	0.013
		(0.009)	(0.024)	(0.010)
Size_{csh}		0.005	0.008	0.005
		(0.006)	(0.018)	(0.007)
$\Delta\text{MFN}_{ht} \cdot \text{Size}_{csh}$				0.001
				(0.001)
Product-Year FE α_{ht}	yes	yes	yes	yes
Country-Year FE α_{st}	yes	yes	yes	yes
N	71,316	70,643	14,205	70,589
R^2	0.119	0.114	0.254	0.114

Note: Note: $\Delta\text{Triple}_{csht} = 1$ if $\triangle ImportPrice_{csht} > 0$ & $\triangle ImportQuantity_{csht} < 0$ & $\triangle \#Importers_{csht} < 0$ from period $t-1$ to t and 0 otherwise. $\text{Treat}_h = 1$ for HS-Codes that received duty-free access from the US-Colombia FTA and 0 otherwise. The sample covers changes in import outcomes in 2010, 2011 and 2012. $\text{Post}_t = 1$ for 2012 and 0 otherwise while $\text{Pre}_t = 1$ for 2011 and 0 otherwise. $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from other developed countries. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product from developed countries in 2009. ΔMFN_{ht} is the change in MFN tariffs from period $t-1$ to t. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

Figure 2: Distribution of Estimated Treatment Effects in 2012



Note: Horizontal axis shows the magnitude of the estimated treatment effect.

otherwise. Similarly, $\text{Tercile2}_{csh} = 1$ if c's initial size is between the 33rd and 66th percentile, and $\text{Tercile3}_{csh} = 1$ if c's initial size is greater than the 66th percentile for exporters of that product. We separate the Tiny exporter-country-products with initial sales less than US$10,000 into another size bin ($\text{Tiny}_{csh} = 1_{e_{csh,2009} \leq 10000}$). Column (1) of Table 5 shows that the triple prediction continues to hold and it is the medium-sized exporters in the first and second terciles that increase their prices, reduce quantities and reduce their number of importers. The coefficient on the second tercile is less precisely estimated, and the pre-treatment variables are always statistically insignificant. It is reassuring that tiny exporters show the opposite results which suggests that it is not purely selection effects at the bottom end of the exporter distribution that drives the results. Theoretically, the selection effects would work to increase exports of the tiny exporters. Column (2) of Table 5 shows that the triple prediction holds when tariff-inclusive prices are used instead of free-on-board (FOB) prices. So the rise in import prices is large enough to offset the direct impact of a reduction in tariffs after the FTA.

One shortcoming of measuring consolidation in the import market through the $\Delta\text{Triple}_{csht}$ indicator is that it will always be zero for exporters that sell to only one importer in the previous period. This underestimates the relative prevalence of the triple prediction among the set of exporters that sell to more than one importer initially. Column (3) restricts the estimation sample to exporter-country-products with more than one import partner initially. As expected, the coefficient on $\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$ increases to over 5 percentage points.[12] Exporters who gained duty-free

[12] Another concern might be that the higher prevalence of the triple prediction could be due to an increase in the variance of price changes, quantity changes and changes in the number of importers. While increases in the variances

Table 5: Tercile results: Triple Prediction by Size Bins

Dependent Variable: $\Delta \text{Triple}_{csht}$

Variable	**(1) fob** **Coef.** **(S.E.)**	**(2) w/tariff** **Coef.** **(S.E.)**
$\text{Tercile3}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.021	-0.031
	(0.021)	(0.022)
$\text{Tercile2}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.043	0.037
	(0.027)	(0.024)
$\text{Tercile1}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.077***	0.068**
	(0.024)	(0.027)
$\text{Tiny}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.067 **	-0.064**
	(0.027)	(0.028)
Controls (including Pre-treatment)	yes	yes
Product-Year FE α_{ht}	yes	yes
Country-Year FE α_{st}	yes	yes
N	71,316	71,316
R^2	0.121	0.117

Note: Note: $\Delta\text{Triple}_{csht} = 1$ if $\triangle ImportPrice_{csht} > 0$ & $\triangle ImportQuantity_{csht} < 0$ & $\triangle \# Importers_{csht} < 0$ from period $t-1$ to t and 0 otherwise. Column (1) uses FOB prices and Column (2) uses prices inclusive of effective duty rates. $\text{Treat}_h = 1$ for HS-Codes that received duty-free access from the US-Colombia FTA and 0 otherwise. The sample covers changes in import outcomes in 2010, 2011 and 2012. $\text{Post}_t = 1$ for 2012 and 0 otherwise while $\text{Pre}_t = 1$ for 2011 and 0 otherwise. $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from other developed countries. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product from developed countries in 2009. $\text{Tercile1}_{csh} = 1$ for the 33rd tercile of Size_{csh}, $\text{Tercile2}_{csh} = 1$ for the 66th tercile and $\text{Tercile3}_{csh} = 1$ for the top tercile. $\text{Tiny}_{csh} = 1$ for all exporters from developed countries with total sales of less than USD10,000 in 2009. Controls include all interactions between Post_t, Treat_h, USA_s, Pre_t and indicators for initial size bins. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

access from the FTA are more likely to increase prices, lower quantities and sell to fewer importers. This effect varies across exporters of different initial size with smaller exporters more likely to engage in consolidation of their import markets.

4.3.3 Quantitative Interpretation

An advantage of testing the model with the triple prediction is that it lets us examine a unique prediction which does not arise in standard models. A limitation of the triple prediction however is that it does not let us put a number on the price increases faced by the importers. To get a sense of the magnitude of the changes, we examine how unit values respond to duty-free access to the Colombian market.

Table 6: Unit value results: Changes in Exporter-level Prices
Dependent Variable: $\Delta \ln \text{Import Price}_{csht}$

Variable	**(1) fob Coef. (S.E.)**	**(2) w/tariff Coef. (S.E.)**
$\text{Tercile3}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.025	-0.069
	(0.104)	(0.104)
$\text{Tercile2}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.084*	0.040
	(0.133)	(0.133)
$\text{Tercile1}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.314**	0.257*
	(0.139)	(0.139)
$\text{Tiny}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.353**	-0.347**
	(0.149)	(0.149)
Controls (including Pre-treatment)	yes	yes
Product-Year FE α_{ht}	yes	yes
Country-Year FE α_{st}	yes	yes
N	71,316	71,285
R^2	0.089	0.089

Note: Note: $\Delta \ln \text{Import Price}_{csht}$ is the log change in the exporter-level average unit value. Column (1) uses FOB prices and Column (2) uses prices inclusive of effective duty rates. $\text{Treat}_h = 1$ for HS-Codes that received duty-free access from the US-Colombia FTA and 0 otherwise. The sample covers changes in import outcomes in 2010, 2011 and 2012. $\text{Post}_t = 1$ for 2012 and 0 otherwise while $\text{Pre}_t = 1$ for 2011 and 0 otherwise. $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from other developed countries. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product from developed countries in 2009. $\text{Tercile1}_{csh} = 1$ for the 33rd tercile of Size_{csh}, $\text{Tercile2}_{csh} = 1$ for the 66th tercile and $\text{Tercile3}_{csh} = 1$ for the top tercile. $\text{Tiny}_{csh} = 1$ for all exporters from developed countries with total sales of less than USD10,000 in 2009. Controls include all interactions between Post_t, Treat_h, USA_s, Pre_t and indicators for initial size bins. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

of these outcomes is consistent with the theory, we also re-estimate the regressions controlling for changes in the variances. The main results are unchanged but our sample is smaller because computing changes in variances for the US and other developed countries separately is more demanding on the data (available upon request).

Using a specification similar to the baseline, we regress changes in unit values to determine the coefficient on $\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$ for each tercile of US exporters. The dependent variable in Column (1) of Table 6 reports the change in the average FOB price charged by exporter c for product h and Column (2) reports the change in the tariff-inclusive price where the tariff is the MFN or preferential average applied duty rate. As earlier, tiny exporters and large exporters (in the third tercile) reduce their prices. From the theory, these are the firms that will not switch contracts after trade liberalization and their decisions will reinforce the standard gains from trade. Medium-size importers (in the first and second terciles of initial size greater than \$10,000) show the opposite effect. These are the firms that are expected to consolidate their import market by switching to joint contracts. And we find that they indeed increase their prices between 8 to 30%.

To interpret the price changes, we regress the log change in tariff-inclusive unit values of each exporter on the indicator for $\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$ for all exporters, without any size conditioning. The coefficient is -0.0425 which is estimated imprecisely. This implies that duty-free access to the Colombian import market induced exporters to reduce their tariff-inclusive prices by 4.25%. This low pass-through reflects the fact that the medium-size exporters increase their prices.

Having discussed the baseline results, we examine the robustness of these predictions with a series of checks on our results.

4.4 Robustness Checks

The robustness checks in this sub-section can be broadly divided into three types. First, we control for other explanations such as contemporaneous tariff cuts for trade partners to other developed countries and quality upgrading to minimize bias in the baseline estimates from omitted variables. Second, we look at a quality-based explanation for the triple prediction. Finally, we examine whether the triple prediction continues to hold when we look within exporter-importer pairs.

4.4.1 Multilateral tariff cuts

The time period that we study also includes a unilateral tariff reform in Colombia that reduced the average tariff from 12.2 percent to 8.2 percent on November 5 2010 (WTO 2012). The stated aim of this tariff reform was to reduce the dispersion in tariffs and eliminate the negative effective protection that existed in some industries. The reform reduced most-favored-nation (MFN) tariffs on a subset of products. For products in our sample, the average MFN duty was 10.3 percentage points before the reform and the duty fell by 1.3 percentage points as a result of the reform. To account for unilateral changes in Colombia's MFN tariffs, Column (4) of Table 5 includes an interaction between the change in MFN tariffs and initial size of the exporter. As the MFN tariff change applies to all trade partners with MFN status, the effect of ΔMFN_{ht} is subsumed in the product-time fixed effects α_{ht}. Across different specifications, we find that our main findings are unaffected, and the interaction between MFN tariff changes and initial size is small and statistically insignificant for the triple prediction.

4.4.2 Revenue Predictions

Increases in unit values following a trade liberalization are often interpreted as quality upgrading by exporters (Verhoogen 2008; Khandelwal 2010). Colombia is a small country relative to other destinations of US exporters, making it less likely that the unit values pick up investments in quality upgrading of US exporters following duty-free access to the Colombian market. In fact, lower trade costs would make it more likely for lower quality US exporters to be able to sell to the Colombian market, because the exporting cost cutoff typically rises after trade liberalization in heterogenous firm trade models. Empirically, higher quality products are more likely to be shipped out when trade barriers are higher (Hummels and Skiba). To ensure however that our results are not driven by quality increases, we examine another triple prediction which looks at a decline in total revenue rather than total quantity of the exporter. If the observed price rise is accompanied by a reduction in revenues and fewer importers, then it is more straightforward to explain it with a reduction in competition rather than an imrpovement in quality which would be expected to increase revenues.[13]

The variety-level model of Section 3 predicts that exporters who switch from bilateral contracts to joint contracts increase their profits but sell lower quantities to fewer importers. An extreme prediction of the model is that these exporters might also lower their sales values to earn higher profit margins per unit. The main intuition here is that under joint contracts, the firm need not sell many units of the product if the markup is large and the fixed costs of joint versus bilateral contracts are relatively low. So if we find that US exporters are more likely to reduce revenues when they increase prices and sell to fewer importers, it would be another piece of evidence that duty-free access induced exporters to increase markups rather than quality.

We examine the revenue-based triple prediction of higher prices, lower revenues and fewer importers in a way similar to the baseline analysis. Specifically, $\Delta\text{Triple}_{csht}^{Revenue} = 1$ if exporter c from source country s selling product h at time t increases its average price across all its importers ($\triangle\tilde{m}_{csht} > 0$) *and* reduces the total revenue received from all its importers *and* reduces the number of importers of its product ($d_{csht}^{\max} < 0$). If any one of these events does not happen, then $\Delta\text{Triple}_{csht}^{Revenue} = 0$. As in the baseline results, Table 7 shows US exporters who receive duty-free access are more likely to have higher prices, lower revenues and fewer importers, and these exporters tend to be the smaller exporters. As expected, the revenue-based triple prediction becomes more likely after the FTA and its increased likelihood is slightly smaller than that of the quantity-based triple prediction in the baseline analysis.[14]

[13] We are grateful to Penny Goldberg and Marc Melitz for pointing us in this direction.

[14] The main results are similar when we only use a double prediction of lower revenue and fewer importers on the HS (available upon request).

Table 7: Estimation results: Revenue-Based Triple Prediction

Dependent Variable: $\Delta \text{Triple}^{Revenue}_{csht}$

Variable	**(1) fob Coef. (S.E.)**	**(2) w/tariff Coef. (S.E.)**
$\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.016 **	0.012*
	(0.006)	(0.007)
$\text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s \cdot \text{Size}_{csh}$	-0.038***	-0.041***
	(0.014)	(0.014)
Controls (including Pre-treatment)	yes	yes
Product-Year FE α_{ht}	yes	yes
Country-Year FE α_{st}	yes	yes
N	71,316	70,612
R^2	0.125	0.113

Note: Note: $\Delta \text{Triple}^{Revenue}_{csht} = 1$ if $\triangle ImportPrice_{csht} > 0$ & $\triangle ImportRevenue_{csht} < 0$ & $\triangle \# Importers_{csht} < 0$ from period $t-1$ to t and 0 otherwise. Column (1) uses FOB prices and revenues, and Column (2) uses prices and revenues, inclusive of effective duty rates. $\text{Treat}_h = 1$ for HS-Codes that received duty-free access from the US-Colombia FTA and 0 otherwise. The sample covers changes in import outcomes in 2010, 2011 and 2012. $\text{Post}_t = 1$ for 2012 and 0 otherwise while $\text{Pre}_t = 1$ for 2011 and 0 otherwise. $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from other developed countries. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product from developed countries in 2009. Controls include all interactions between Post_t, Treat_h, USA_s, Size_{csh} and Pre_t. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

4.4.3 Quality Estimation

To look further into the quality explanation, we estimate quality based on standard methods and a flexible approach that is consistent with our theory. We show that the triple prediction using quality-adjusted prices (instead of simple unit values) continues to hold. Exporters who gained duty-free access became more likely to increase prices, reduce quantities and reduce the number of their importers, even after adjusting for estimated quality.

Let λ_c denote the quality of exporter c's variety. Then welfare from the differentiated goods is $Q = \int \lambda_c u(q_c) dG_c$. We are interested in quality-adjusted prices which net out changes in exporter-level quality λ_c before and after the trade liberalization. The quality adjusted price is $p_{cd}/\lambda_c = v'(q_{cd})u'(q_c)Q^{\eta-1}$ and we will infer it in two ways. First, we follow Khandelwal et al. (2013) and use the elasticities of substitution from the literature to recover quality-adjusted prices, based on the markup assumptions in the theory of Section 2. Then we will discuss a more flexible approach that enables us to directly approximate the lower nest marginal utility.

For an exporter c that sells to more than one importer d, the inverse demand for the variety is $\ln p_{cd} = \ln v'(q_{cd}) + (\ln \lambda_c + \ln u'(q_c) - (1-\eta)\ln Q) \equiv \ln v'(q_{cd}) + \alpha_c$ where α_c is an exporter-specific term. We typically have estimates for the elasticity in the upper nest of $u(q_c)$ but do not have readily available measures of the elasticity in the lower nest $v(q_{cd})$. To make progress, we augment the approach of Khandelwal et al. (2013) to recover quality using demand elasticities from Broda and Weinstein (2006) and Soderbery (2015). Assuming $v(q_{cd}) \equiv q_{cd}^{1-\mu_v}$, $\ln v'(q_{cd}) = \ln(1-\mu_v) - \mu_v \ln q_{cd}$. A key assumption in our theory is that the elasticity of demand in the upper nest is smaller than the elasticity in the lower nest ($\mu_u \varepsilon_v < \min \mu_v \min(1-\mu_v)$), so we get a lower bound for the elasticity in the lower nest. Let the upper nest utility function be CES with a differentiation parameter denoted by $1-\mu_u$. Then we know that the elasticity for the lower nest is bounded below by the upper nest elasticity so that $\mu_v = \kappa\mu_u$ for $0 < \kappa < 1$, and we will use the upper nest elasticities from the literature to estimate quality. We estimate the unit value of the variety of exporter-importer cd for each HS code h from source country s in each time period t as

$$\ln p_{cdsht} = -\kappa \cdot \mu_{uh} \ln q_{cdsht} + \alpha_{csht} + \epsilon_{cdsht}$$

where μ_{uh} is obtained from Broda-Weinstein-Soderbery elasticities ($\sigma_h \equiv 1/\mu_{uh}$). The key identifying assumptions are that κ does not vary across products and that quality varies at the exporter-product-time level. Having estimated $\hat{\kappa}\hat{\alpha}_{csht}$, we recover quality $\ln \hat{\lambda}_{csht}$ as a residual by regressing $\hat{\alpha}_{csht} + \mu_{uh} \ln q_{chst} - \ln(1-\mu_{uh})$ on product-time fixed effects and using μ_{uh} from the Broda-Weinstein-Soderbery estimates of elasticities of substitution across product groups. Then the quality-adjusted prices are $\ln p_{cdsht} - \ln \lambda_{csht}$ which we use to determine the changes in quality-adjusted price in the triple prediction.

Under the flexible approach, we move away from the lower bound method and approximate the lower nest marginal utility as a function of quantities. The first step is to estimate $v'(q_{cd})$ which we take to be a cubic function of $\ln q_{cd}$. The unit value of exporter-importer cd for each HS code h from source country s in each time period t is

$$\ln p_{cdsht} = \beta_1 \ln q_{cdsht} + \beta_2 \left(\ln q_{cdsht}\right)^2 + \beta_3 \left(\ln q_{cdsht}\right)^3 + \alpha_{csht} + \epsilon_{cdsht}.$$

Having approximated $\hat{v}'(q_{cd}) = \hat{\beta}_1 \ln q_{cd} + \hat{\beta}_2 \left(\ln q_{cd}\right)^2 + \hat{\beta}_3 \left(\ln q_{cd}\right)^3$, we proceed to recover λ_c. Assuming $u'(q_c) = q_c^{1-\mu_u}$, our estimated exporter-specific quality is $\ln \lambda_{csht} + \ln \alpha_{ht} = \ln p_{cdsht} - \ln \hat{v}'(q_{cdsht}) + \mu_{uh} \ln q_{csht} - \ln\left(1 - \mu_{uh}\right)$ which is identified up to a constant and where μ_{uh} is based on Broda-Weinstein-Soderbery estimates of elasticities of substitution for the product group h. Details of the quality estimation are provided in the Appendix and here we discuss the triple prediction using quality-adjusted prices.

The triple prediction continues to show similar results when quality-adjusted prices are used instead of unit prices. Column (1) of Table 8 is based on the quality-adjusted prices from the first method and Column (2) is based on the flexible method. We find once again that exporters in terciles 1 and 2 are more likely to consolidate their import market.

4.4.4 Within-Exporter-Importer Pairs

While the baseline results examine the triple prediction at the level of the exporter, the model predicts that the results should also hold for some exporter-importer pairs so we estimate the triple prediction at the exporter-importer pair level. Let d denote a Colombian importer that imports from exporter c. The triple variable is now defined as $\Delta\text{Triple}_{cdsht} = 1$ for exporter-importer pairs cd that experience an increase in their bilateral price, a fall in their bilateral quantity ($\Delta q_{cdsht} < 0$) and a decrease in the importers that buy from c ($\Delta d^{\max}_{csht} < 0$). If any one of these events does not happen, then $\Delta\text{Triple}_{cdsht} = 0$. Columns (1) and (2) of Table 9 show that the likelihood of increasing import prices, reducing quantities and reducing the number of importers goes up after the trade liberalization. As earlier, exporter-importer pairs in the middle of the exporter size distribution are more likely to see the triple prediction hold. The exporter-importer pairs of tiny and large exporters who are not expected to switch contracts show small effects or a reduction in the likelihood of the triple prediction.

5 Conclusion

This paper examines how the behavior of importers and their interaction with exporters affect the division of the gains from trade among consumers, importers and exporters. When an exporter sells a product to importers through anonymous markets, double marginalization and business stealing among competing importers lead to lower profits. Exporters and importers can invest in relationships that overcome these sources of lower profits.

Table 8: Tercile results: Quality-adjusted Triple Prediction by Size Bins

Dependent Variable: $\Delta\text{Triple}_{csht}$

Variable	**(1) Coef. (S.E.)**	**(2) Coef. (S.E.)**
$\text{Tercile3}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.001	-0.006
	(0.023)	(0.023)
$\text{Tercile2}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.026	0.019
	(0.021)	(0.023)
$\text{Tercile1}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.071**	0.061**
	(0.032)	(0.030)
$\text{Tiny}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.063**	-0.056*
	(0.031)	(0.030)
Controls (including Pre-treatment)	yes	yes
Product-Year FE α_{ht}	yes	yes
Country-Year FE α_{st}	yes	yes
N	57,480	57,480
R^2	0.112	0.124

Note: Note: $\Delta\text{Triple}_{csht} = 1$ if $\triangle QualityAdjImportPrice_{csht} > 0$ & $\triangle ImportQuantity_{csht} < 0$ & $\triangle \#Importers_{csht} < 0$ from period $t-1$ to t and 0 otherwise. $\text{Treat}_h = 1$ for HS-Codes that received duty-free access from the US-Colombia FTA and 0 otherwise. The sample covers changes in import outcomes in 2010, 2011 and 2012. $\text{Post}_t = 1$ for 2012 and 0 otherwise while $\text{Pre}_t = 1$ for 2011 and 0 otherwise. $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from other developed countries. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product from developed countries in 2009. $\text{Tercile1}_{csh} = 1$ for the 33rd tercile of Size_{csh}, $\text{Tercile2}_{csh} = 1$ for the 66th tercile and $\text{Tercile3}_{csh} = 1$ for the top tercile. $\text{Tiny}_{csh} = 1$ for all exporters from developed countries with total sales of less than USD10,000 in 2009. Controls include all interactions between Post_t, Treat_h, USA_s, Pre_t and indicators for initial size bins. Standard errors are bootstrapped and clustered by exporter. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

Table 9: Exporter-Importer results: Changes in the Triple Prediction Within Exporter-Importer Pairs

Dependent Variable: $\Delta \text{Triple}_{cdsht}$

Variable	**(1) fob Coef. (S.E.)**	**(2) w/tariff Coef. (S.E.)**
$\text{Tercile3}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.005	-0.017
	(0.025)	(0.025)
$\text{Tercile2}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.046*	0.028
	(0.028)	(0.027)
$\text{Tercile1}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	0.053**	0.051**
	(0.026)	(0.026)
$\text{Tiny}_{csh} \cdot \text{Post}_t \cdot \text{Treat}_h \cdot \text{USA}_s$	-0.0414	-0.049*
	(0.028)	(0.027)
Controls (inc. Pre-treatment)	yes	yes
Product-Year FE α_{ht}	yes	yes
Country-Year FE α_{st}	yes	yes
N	82,130	82,130
R^2	0.110	0.106

Note: Note: $\Delta \text{Triple}_{cdsht} = 1$ if $\triangle ImportPrice_{cdsht} > 0$ & $\triangle ImportQuantity_{cdsht} < 0$ & $\triangle \# Importers_{csht} < 0$ from period $t-1$ to t and 0 otherwise. Column (1) uses FOB prices and Column (2) uses prices inclusive of effective duty rates. $\text{Treat}_h = 1$ for HS-Codes that received duty-free access from the US-Colombia FTA and 0 otherwise. The sample covers changes in import outcomes in 2010, 2011 and 2012. $\text{Post}_t = 1$ for 2012 and 0 otherwise while $\text{Pre}_t = 1$ for 2011 and 0 otherwise. $\text{USA}_s = 1$ for exporters from the United States and 0 for exporters from other developed countries. Size_{csh} is the z-score of initial export value $e_{csh,2009}$ in 2009 relative to all exporters of the product from developed countries in 2009. $\text{Tercile1}_{csh} = 1$ for the 33rd tercile of Size_{csh}, $\text{Tercile2}_{csh} = 1$ for the 66th tercile and $\text{Tercile3}_{csh} = 1$ for the top tercile. $\text{Tiny}_{csh} = 1$ for all exporters from developed countries with total sales of less than USD10,000 in 2009. Controls include all interactions between Post_t, Treat_h, USA_s, Pre_t and indicators for initial size bins. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

We embed the choice of contracts between exporters and importers into a trade model with heterogeneous exporters and importers. Exporters can offer bilateral contracts that eliminate double marginalization, leading to higher profits and lower prices. Exporters can internalize business stealing across importers by investing in joint contracts. This enables an exporter to commit to mitigating competition among its importers allowing total profit from the product to rise at the expense of consumer welfare. When investments in joint contracts embody economies of scale and markups that rise with quantities, trade liberalization changes the relative incentives to engage in bilateral contracts and joint contracts. Lower trade costs increase the surplus from joint contracts, leading to an increase in firm profits at the expense of consumer welfare.

The model enables us to derive unique predictions for changes in import prices, quantities and the number of importers per exporter. Testing these implications empirically, we show that bilateral trade liberalization between Colombia and the US induced US exporters to consolidate their import market, increasing the probability of higher prices, reduced quantity and fewer Colombian importers per exporter. These observable predictions suggest that the actions of exporters and importers affect the ability of consumers to gain from trade. The estimated elasticity of import prices with respect to trade costs shows substantial increases in prices for certain products from trade liberalization. Future work can shed more light on how this translates into consumer price changes and its contribution to the aggregate pass-through of reduction in trade barriers into consumer prices.

References

Ahn, J., A. K. Khandelwal, and S. Wei (2011): "The role of intermediaries in facilitating trade," *Journal of International Economics*, 84(1), 73–85. 1

Akerman, A. (2010): "A theory on the role of wholesalers in International trade based on economies of scope," Research papers in economics, Stockholm University, Department of Economics. 1

Amiti, M. and J. Konings (2007): "Trade Liberalization, Intermediate Inputs and Productivity: Evidence from Indonesia," *American Economic Review*, 97, 1611–1638. 1

Antràs, P. and A. Costinot (2011): "Intermediated Trade," *The Quarterly Journal of Economics*, 126, 1319–1374. 1

Antràs, P. and E. Helpman (2004): "Global Sourcing," *Journal of Political Economy*, 112, 552–580. 1

Atkin, D. and D. Donaldson (2012): "Who's Getting Globalized? The Size and Nature of Intranational Trade Costs," *Working Paper*. 1

Badinger, H. (2007): "Has the EU's Single Market Programme Fostered Competition? Testing for a Decrease in Mark-up Ratios in EU Industries," *Oxford Bulletin of Economics and Statistics*, 69, 497–519. 1, 1

Bagwell, K. and S. H. Lee (2015): "Trade Policy under Monopolistic Competition with Firm Selection," *Wokirng Paper, Stanford University*. 1

Bagwell, K. and R. W. Staiger (2012): "The economics of trade agreements in the linear Cournot delocation model," *Journal of International Economics*, 88, 32–46. 1

Baller, S. (2013): "Product quality, market size and welfare: theory and evidence from French exporters," *Working Paper University of Oxford*. 3.1

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2010): "Wholesalers and Retailers in US Trade," *The American Economic Review*, 100, 408–413. 1

Bernard, A. B., J. B. Jensen, and P. K. Schott (2009): "Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods," in *Producer dynamics: New Evidence from Micro Data*, ed. by T. Dunne, J. Jensen, and M. Roberts, University of Chicago Press. 1

Bernard, A. B., A. Moxnes, and K. H. Ulltveit-Moe (2014): "Two-sided Heterogeneity and Trade," Working Paper 20136, NBER. 1, 7

BLANCHARD, E., T. CHESNOKOVA, AND G. WILLMANN (2013): "Private labels and international trade: Trading variety for volume," Tech. rep., Kiel Working Paper. 1

BLUM, B. S., S. CLARO, AND I. HORSTMANN (2010): "Facts and figures on intermediated trade," *The American Economic Review*, 100, 419–423. 1

BLUM, B. S., S. CLARO, AND I. J. HORSTMANN (2012): "Import Intermediaries and Trade Costs: Theory and Evidence," mimeo, University of Toronto. 1

BOLER, E. A., A. MOXNES, AND K. H. ULLTVEIT-MOE (2015): "Technological Change, International Sourcing and the Joint Impact on Productivity," *American Economic Review*, 105, 3704–39. 1

BRODA, C. AND D. E. WEINSTEIN (2006): "Globalization and the Gains From Variety," *The Quarterly Journal of Economics*, 121, 541–585. 4.4.3

BURSTEIN, A. T., J. C. NEVES, AND S. REBELO (2003): "Distribution costs and real exchange rate dynamics during exchange-rate-based stabilizations," *Journal of monetary Economics*, 50, 1189–1214. 1

CAMPA, J. M. AND L. S. GOLDBERG (2005): "Exchange Rate Pass-Through into Import Prices," *The Review of Economics and Statistics*, 87, 679–690. 1

CARBALLO, J., G. I. OTTAVIANIO, AND C. V. MARTINCUS (2013): "The Buyer Margins of Firms' Exports," Tech. rep., LSE. 1

CASTELLANI, D., F. SERTI, AND C. TOMASI (2010): "Firms in international trade: Importers' and exporters' heterogeneity in Italian manufacturing industry," *The World Economy*, 33, 424–457. 1

CHANEY, T. (2014): "The Network Structure of International Trade," *American Economic Review*, 104 (11), 3600–3634. 1

CONCONI, P., P. LEGROS, AND A. F. NEWMAN (2012): "Trade liberalization and organizational change," *Journal of international economics*, 86, 197–208. 1

DE LOECKER, J., P. K. GOLDBERG, A. K. KHANDELWAL, AND N. PAVCNIK (2012): "Prices, markups and trade reform," Tech. rep., National Bureau of Economic Research. 1, 1

DHINGRA, S. AND J. M. MORROW (2012): "Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity," *Working Paper*. 1, 2, 2.1, 2.1

EATON, J., M. ESLAVA, D. JINKINS, C. J. KRIZAN, AND J. TYBOUT (2014): "A Search and Learning Model of Export Dynamics," working paper. 1

Eckel, C. (2009): "International trade and retailing," Working Paper 2597, CESIfo. 1

Eckel, C., L. Iacovone, B. Javorcik, and J. P. Neary (2015): "Multi-product firms at home and away: Cost-versus quality-based competence," *Journal of International Economics*, 95, 216–232. 3.1

Engel, C. and J. H. Rogers (1996): "How Wide Is the Border?" *American Economic Review*, 86, 1112–25. 1

Feenstra, R. C., D.-S. Huang, and G. G. Hamilton (2003): "A market-power based model of business groups," *Journal of Economic Behavior & Organization*, 51, 459–485. 1

Ganslandt, M. and K. E. Maskus (2004): "Parallel imports and the pricing of pharmaceutical products: evidence from the European Union," *Journal of health economics*, 23, 1035–1058. 1

Halpern, L., M. Koren, and A. Szeidl (2015): "Imported Inputs and Productivity," *American Economic Review*, 12, 3660–3703. 1

Harrison, A. E. (1994): "Productivity, Imperfect Competition, and Trade Reform: Theory and Evidence," *Journal of International Economics*, 36 (1-2), 53–73. 1

Harrison, R., J. Jaumandreu, J. Mairesse, and B. Peters (2005): "Does Innovation Stimulate Employment? A Firm-level Analysis Using Comparable Micro Data on Four European Countries," *European research project Innovation and Employment in European Firms: Microeconometric Evidence (IEEF)*. 1

Hart, O. and J. Tirole (1990): "Vertical integration and market foreclosure," *Brookings papers on economic activity. Microeconomics*, 1990, 205–286. 1, 2.4

Horn, H. and A. Wolinsky (1988): "Bilateral monopolies and incentives for merger," *The RAND Journal of Economics*, 408–419. 2.4, 2.5.1

Khandelwal, A. K. (2010): "The Long and Short (of) Quality Ladders," *Review of Economic Studies*, 77, 1450–1476. 4.4.2

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013): "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters," *The American Economic Review*, 103, 2169–2195. 4.4.3, 3

Konings, J., P. V. Cayseele, and F. Warzynski (2005): "The Effects of Privatization and Competitive Pressure on Firms' Price-Cost Margins: Micro Evidence from Emerging Economies," *The Review of Economics and Statistics*, 87, 124–134. 1

Krugman, P. (1979): "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, 9, 469–479. 2, 2.1

Krugman, P. R. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70, 950–959. 2.2

Lawrence, R. Z. (1991): "How open is Japan?" in *Trade with Japan: has the door opened wider?*, University of Chicago Press, 1991, 9–50. 1

Lawrence, R. Z. and G. R. Saxonhouse (1991): "Efficient or exclusionist? The import behavior of Japanese corporate groups," *Brookings Papers on Economic Activity*, 311–341. 1

Levinsohn, J. (1993): "Testing the imports-as-market-discipline hypothesis," *Journal of International Economics*, 35, 1–22. 1

Mallick, S. and H. Marques (2008): "Passthrough of Exchange Rate and Tariffs into Import Prices of India: Currency Depreciation versus Import Liberalization*," *Review of International Economics*, 16, 765–782. 1, 1

Martin, S., H.-T. Normann, and C. M. Snyder (2001): "Vertical foreclosure in experimental markets," *RAND Journal of Economics*, 466–496. 3, 5

Mayer, T., M. J. Melitz, and G. I. Ottaviano (2014): "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 104(2). 1, 2.1

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725. 2.1

Miklos-Thal, J., P. Rey, and T. Verge (2010): "Vertical relations," *International Journal of Industrial Organization*, 28, 345–349. 2

Mollers, C., H.-T. Normann, and C. M. Snyder (2014): "Communication in Vertical Markets: Experimental Evidence," *Dartmouth College Working Paper*. 3

Muuls, M. and M. Pisu (2009): "Imports and Exports at the Level of the Firm: Evidence from Belgium," *The World Economy*, 32, 692–734. 1

Neary, J. P. and M. Mrazova (2013): "Not so demanding: Preference structure, firm behavior, and welfare," Tech. rep., University of Oxford, Department of Economics. 1, 2.1

Parenti, M., P. Ushchev, and J.-F. Thisse (2014): "Toward a theory of monopolistic competition," . 1, 2.1

Petropoulou, D. (2011): "Information costs, networks and intermediation in international trade," Globalization and Monetary Policy Institute Working Paper 76, Federal Reserve Bank of Dallas. 1

Pierce, J. R. and P. K. Schott (2012): "Concording US Harmonized System Categories over Time," *Journal of Official Statistics*, 28, 53–68. 4.1

Raff, H. and N. Schmitt (2005): "Endogenous vertical restraints in international trade," *European Economic Review*, 49, 1877–1889. 1

——— (2009): "Buyer Power in International Markets," *Journal of International Economics*, 79(2), 222–229. 1

——— (2012): "Imports and the Structure of Retail Markets," *Canadian Journal of Economics*, 45(4), 1431–1455. 1

Rauch, J. E. (1999): "Networks versus markets in international trade," *Journal of International Economics*, 48, 7–35. 1

Rauch, J. E. and J. Watson (2004): "Network Intermediaries in International Trade," *Journal of Economics & Management Strategy*, 13, 69–93. 1

Rey, P. and J. Tirole (2007): "A primer on foreclosure," *Handbook of industrial organization*, 3, 2145–2220. 2.4

Soderbery, A. (2015): "Estimating import supply and demand elasticities: Analysis and implications," *Journal of International Economics*, 96, 1–17. 4.4.3

Tybout, J. R. (2003): "Plant-and firm-level evidence on "new" trade theories," *Handbook of International Trade*, 1, 388–415. 1

Venables, A. J. (1985): "Trade and trade policy with imperfect competition: The case of identical products and free entry," *Journal of International Economics*, 19, 1–19. 1

Verhoogen, E. A. (2008): "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector*," *The Quarterly Journal of Economics*, 123, 489–530. 4.4.2

WTO (2012): "Trade Policy Review: Colombia," *WT/TPR/S/265*. 4.4.1

Appendix

Fixed Entry

We will show that the profits of the exporters are supermodular in market size and contracting choice under conditions on the rate at which markups change with quantities.

The exporter's optimal profit is $\pi_c = s\left(m(q)q/(1-m(q))\right) M_d\beta cdL$ where the markup is $m(q) \equiv \mu_v(q) + z\mu_u\left(M_d v(q)\right)\varepsilon_v(q)$ for $z \in \{0,1\}$ at the optimally chosen quantity q. This formulation lets us specify the different contracting choices through z. Profits under bilateral contracts correspond to $z = 0$ and profits under joint contracts correspond to $z = 1$. The optimal quantity is determined by $u'\left(M_d v(q)\right) v'(q)\left(1-m\right) = \delta cd$. The consumer budget multiplier is taken as fixed by the firm and is given by $\delta = sM_c M_d u'\left(M_d v(q)\right) v'(q)$ in equilibrium. As z rises, the profit function changes by $\frac{d\ln\pi}{d\ln z} = \left(1 + \frac{m_q q}{m(1-m)}\right)\frac{d\ln q}{d\ln z} + \frac{z\mu_u\varepsilon_v}{m(1-m)}$. From the optimal pricing condition, $\ln\delta + \ln u'\left(M_d v(q)\right) v'(q)\left(1-\mu_v - z\mu_u\varepsilon_v\right) = cd$ so that $-\left(\mu + \mu_u\varepsilon + \frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln z} = \frac{z\mu_u\varepsilon_v}{1-m}$ and quantities fall under joint contracts. Substituting in the profit derivative, $\pi_z = \beta L\tau cd\frac{(1-z)q}{(1-m)^2}\frac{sM_d(\mu_u\varepsilon_v)^2}{\mu_v + \mu_u\varepsilon_v + \frac{m_q q}{1-m}}$ so $s\pi_{zs}/\pi_z = 1 + \frac{d\ln q}{d\ln s}\left(1 + 2\frac{d\ln\mu_u\varepsilon_v}{d\ln q} + \frac{2m_q q}{1-m} - \frac{d\ln\left(\mu_v + \mu_u\varepsilon_v + \frac{m_q q}{1-m}\right)}{d\ln q}\right) \equiv 1 + \left(1 + A_q\right)\frac{d\ln q}{d\ln s}$.

From optimal contracting, $\pi_z = \beta L\tau cd\frac{(1-z)q}{(1-m)^2}\frac{xM_d(\mu_u\varepsilon_v)^2}{\mu_v + \mu_u\varepsilon_v + \frac{m_q q}{1-m}} = f_c^J$ so that differentiating with respect to market size gives $1 + \left(1 + A_q\right)\frac{d\ln q}{d\ln s} + \left(A_z - z/(1-z)\right)\frac{d\ln z}{ds} = 0$ for $A_z \equiv \frac{2z\mu_u\varepsilon}{1-m} - \frac{d\ln\left(\frac{m_q q}{1-m}\right)}{d\ln z}$. From optimal pricing and the budget multiplier, $1 - m = sqM_c M_d cd$ implying $-\left(1 + \frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln s} - \frac{z\mu_u\varepsilon_v}{1-m}\frac{d\ln z}{d\ln s} = 1$. Putting the two together, $-\left(1 + \frac{m_q q}{1-m} + \frac{\frac{z\mu_u\varepsilon_v}{1-m}(1+A_q)}{\frac{z}{1-z} - A_z}\right)\left(d\ln q/d\ln s\right) = 1 + \frac{\frac{z\mu_u\varepsilon_v}{1-m}}{z/(1-z) - A_z}$. Substituting for this, $d\ln\pi_z/d\ln s = \left(\frac{m_q q}{1-m} - A_q\right) / \left(1 + \frac{m_q q}{1-m} + \frac{\frac{z\mu_u\varepsilon_v}{1-m}(1+A_q)}{\frac{z}{1-z} - A_z}\right)$. So as long as $A_q, A_z \leq 0$, $d^2\pi_c/dzds > 0$ and market expansion makes higher values of z more desirable.

Free Entry

To examine the robustness of the qualitative predictions, we specify the upper nest as CES ($u(q_c) = q_c^\rho$). Under free entry, the additional conditions are $sM_c\left(1-\beta\right)\left[\frac{m}{1-m}\tau cdqL - f_{cd}\right] = f_d$ and $sM_d\beta\left[\frac{m}{1-m}\tau cdqL - f_{cd}\right] = f_c + zf_c^J$. Differentiating π_z with respect to market size, $d\ln\pi_z/d\ln s = 1 + \frac{d\ln M_d}{d\ln s} + \left(1 + A_q\right)\frac{d\ln q}{d\ln s}$. As z is chosen optimally, the free entry conditions imply $\frac{d\ln M_c}{d\ln s} = -1 - \left(1 + \frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln s} = \frac{d\ln M_d}{d\ln s}$. From optimal pricing and optimal contracting, $-\left(1 + \frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln s} - \frac{z\mu_u\varepsilon}{1-m}\frac{d\ln z}{d\ln s} - \frac{d\ln M_c}{d\ln s} - \frac{d\ln M_d}{d\ln s} = 1$ and $1 + \left(1 + A_q\right)\frac{d\ln q}{d\ln s} + \frac{d\ln M_d}{d\ln s} + \left(A_z - z/(1-z)\right)\frac{d\ln z}{d\ln s} = 0$. Solving for quantity changes using free entry, $-\left(1 + \frac{m_q q}{1-m}\left(1 + \frac{\frac{z\mu_u\varepsilon_v}{1-m}}{\frac{z}{1-z} - A_z}\right) - \frac{\frac{z\mu_u\varepsilon_v}{1-m}A_q}{\frac{z}{1-z} - A_z}\right)\frac{d\ln q}{d\ln s} = 1$. Then $d\ln\pi_z/d\ln s = 1 + \frac{d\ln M_d}{d\ln s} + \left(1 + A_q\right)\frac{d\ln q}{d\ln s} = -\left(\frac{m_q q}{1-m} - A_q\right)\frac{d\ln q}{d\ln s} > 0$.

Welfare

Welfare is $U = sM_c u\left(M_d v\left(q\right)\right)$ so that $\frac{d\ln U}{d\ln z} = \frac{d\ln M_c}{d\ln z} + \varepsilon_u \frac{d\ln M_d}{d\ln z} + \varepsilon_u \varepsilon_v \frac{d\ln q}{d\ln z}$. Under fixed entry, $d\ln U/d\ln z = \varepsilon_u \varepsilon\left(d\ln q/d\ln z\right)$ which is negative because firms scale back quantities under joint contracts as shown above. We have also shown above that $\pi_z > 0$ so profits rise under joint contracts. Under free entry with a CES upper nest, optimal pricing implies $-\left(1 + \frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln z} = \frac{z\mu_u \varepsilon}{1-m} + \frac{d\ln M_c}{d\ln z} + \frac{d\ln M_d}{d\ln z}$ and optimal contracting implies $0 = -\frac{z}{1-z} + \frac{2z\mu_u\varepsilon}{1-m} - \frac{\frac{m_q q}{1-m}}{\mu+\mu_u\varepsilon+\frac{m_q q}{1-m}}\frac{d\ln\frac{m_q q}{1-m}}{d\ln z} + \frac{d\ln M_d}{d\ln z} + \frac{d\ln q}{d\ln z}A_q$. From free entry,$-\frac{m_q q}{m}\frac{d\ln M_d}{d\ln z} = \left(1 + \frac{m_q q}{m(1-m)}\right)\frac{d\ln M_c}{d\ln z}$ and $\frac{\ln M_d}{d\ln z} = \frac{d\ln M_c}{d\ln z} + zf_c^J/\left(f_c + zf_c^J\right)$. Entry changes by $\frac{d\ln M_c}{d\ln z} = -\frac{m_q q/m}{1+\frac{2-m}{m}\frac{m_q q}{1-m}}\frac{zf_c^J}{f_c+zf_c^J}$ and $\frac{d\ln M_d}{d\ln z} = \frac{1+\frac{m_q q}{m(1-m)}}{1+\frac{2-m}{m}\frac{m_q q}{1-m}}\frac{zf_c^J}{f_c+zf_c^J}$. So the change in quantities is $-\left(1 + \frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln z} = \frac{z\mu_u\varepsilon}{1-m} + \frac{1+\frac{m_q q}{1-m}}{1+\frac{2-m}{m}\frac{m_q q}{1-m}}\frac{zf_c^J}{f_c+zf_c^J}$. Substituting in the welfare derivative,

$$
\begin{aligned}
-\frac{d\ln U}{d\ln z}/\frac{z\mu_u\varepsilon}{1-m} =& \frac{\varepsilon_u\varepsilon}{1+\frac{m_q q}{1-m}} + \frac{\varepsilon_u\varepsilon + \frac{m_q q}{m} - \varepsilon_u - \frac{\varepsilon_u}{1-m}\frac{m_q q}{m}}{1+\frac{2-m}{m}\frac{m_q q}{1-m}}\frac{(1-z)m\mu_u\varepsilon}{\mu+\mu_u\varepsilon+\frac{m_q q}{1-m}} \\
\geq& \frac{\varepsilon_u\varepsilon}{1+\frac{m_q q}{1-m}}\left(1 - \frac{m+m\frac{m_q q}{1-m}}{\mu+\mu_u\varepsilon+\frac{m_q q}{1-m}}\frac{1-\varepsilon+\frac{1}{m}\frac{m_q q}{1-m}}{1+\frac{2-m}{m}\frac{m_q q}{1-m}}(1-z)\mu_u\right) \\
\geq& \frac{\varepsilon_u\varepsilon}{1+\frac{m_q q}{1-m}}\left(1-(1-z)\mu_u\right) \geq 0
\end{aligned}
$$

A switch to joint contracts lowers consumer welfare.

Under free entry, the change in welfare for bilateral contracts is $d\ln U/d\ln s = 1 + d\ln M_c/d\ln s + \varepsilon_u d\ln M_d/d\ln s + \varepsilon_u\varepsilon_v d\ln q^B/d\ln s$. From optimal pricing, $-\left(1+\frac{m_q q}{1-m}\right)\frac{d\ln q}{d\ln s} = 1 + \frac{d\ln M_c}{d\ln s} + \frac{d\ln M_d}{d\ln s}$ and free entry implies $\frac{d\ln M_d}{d\ln s} = \frac{d\ln M_c}{d\ln s} = -1 - \left(1 + \frac{m_q q}{m(1-m)}\right)\frac{d\ln q}{d\ln s}$. Putting these conditions together, quantities and entry fall because $d\ln q/d\ln s = -1/\left(1 + \frac{m_q q}{1-m}\left(2/m - 1\right)\right)$ and $d\ln M_c/d\ln s = d\ln M_d/d\ln s = -1 + \frac{1+\frac{m_q q}{m(1-m)}}{1+\frac{m_q q}{1-m}\frac{2-m}{m}}$. Substituting in the welfare derivative, $\left(1 + \frac{m_q q}{1-m}\frac{2-m}{m}\right)\frac{d\ln U}{d\ln s} = 1 - \varepsilon_u\varepsilon + \frac{m_q q}{m(1-m)}\left(1 - \varepsilon_u(1-m)\right) > 0$.

Variety-Level Model

This section explains the variety-level predictions in a model with relationship-specific cost $f_{cd} > 0$, ad valorem trade costs $\tau > 1$ and fixed potential entry. The optimal price is $p_{cd}\left(1 - m_{cd}\right) = cd$ for $m = \mu + z$ and the optimal importer cost cutoff is $\left(m_{cd_m} - z/\varepsilon_m\right)cd_m q_{cd_m}/\left(1 - m_{cd_m}\right) = f_{cd}/L$. From optimal pricing, $-\left(\mu_{cd} + \frac{m_q(q_{cd})q_{cd}}{1-m_{cd}}\right)\frac{d\ln q_{cd}}{d\ln z} = \frac{z}{1-m_{cd}} + \mu_u(q_c)\frac{d\ln q_c}{d\ln z}$ and from the optimal cutoff, $\frac{1-z/\varepsilon_{cd_m}}{m_{cd_m}-z/\varepsilon_{cd_m}}\left(\mu_{cd_m} + \frac{m_q(q_{cd_m})q_{cd_m}}{1-m_{cd_m}}\right)\frac{d\ln q_{cd_m}}{d\ln z} + \frac{d\ln d_m}{d\ln z} + \frac{z-z/\varepsilon_{cd_m}}{m_{cd_m}-z/\varepsilon_{cd_m}} + \frac{z}{1-m_{cd_m}} = 0$. Substituting for the change in quantities and re-writing, the change in the importer cost cutoff is

$$
-\frac{d\ln d_m}{d\ln z} = \frac{z/\varepsilon_{cd_m}}{1-m_{cd_m}} + \frac{1-z/\varepsilon_{cd_m}}{1-m_{cd_m}}\mu_{u_c}\frac{d\ln q_c}{d\ln z}
$$

As $q_c = \int_0^{d_m} v(q)dG$, total quantity changes by $\left(\frac{q_c}{\mu_{u_c}} + \frac{v_{cd_m} d_m (1-z/\varepsilon_{cd_m})}{m_{cd_m} - z/\varepsilon_{cd_m}} + \int \frac{v'(q_{cd}) q_{cd} dG_d}{\mu_{cd} + \frac{m_q(q_{cd})q_{cd}}{1-m_{cd}}}\right) \mu_u \frac{d\ln q_c}{d\ln z} = -\left(v_{cd_m} d_m \frac{z/\varepsilon_{cd_m}}{1-m_{cd_m}} + \int \frac{v'(q_{cd}) q_{cd} dG_d}{\mu_{cd} + \frac{m_q(q_{cd})q_{cd}}{1-m_{cd}}} \frac{z}{1-m_{cd}}\right)$. So $d\ln q_c / d\ln z < 0$ and total quantities always fall after joint contracts. Substituting for the change in total quantity, the change in the importer cost cutoff is $(1-m_{cd_m}) \frac{d\ln d_m}{d\ln z} = -\varepsilon_{cd_m}^{-1} \frac{\int v'(q_{cd}) q_{cd} \left[1 + \frac{z}{\mu_{cd} + \frac{m_q(q_{cd})q_{cd}}{1-m_{cd}}} \frac{1-\mu_{cd}-\varepsilon_{cd_m}}{1-\mu_{cd}-z}\right] dG_d}{q_c/\mu_{u_c} + v_{cd_m} d_m \frac{1-z/\varepsilon_{cd_m}}{1-m_{cd_m}} + \int \frac{v'(q_{cd}) q_{cd} dG_d}{\mu_{cd} + \frac{m_q(q_{cd})q_{cd}}{1-m_{cd}}}}$. We will show that $(1-\mu_{cd} - \mu_{u_c}\bar{\varepsilon}_{v_c}) \mu_{cd} + (1-\mu_{cd} - \varepsilon_{cd_m}) \mu_{u_c}\bar{\varepsilon}_{v_c} + m_q(q_{cd}) q_{cd} > 0$ so that the importer cost cutoff falls under joint contracts.

First note that the LHS of the inequality is decreasing in $\mu_{u_c}\bar{\varepsilon}_{v_c}$ under $(1-\varepsilon)' > 0$. And we will need to show that $\mu_{u_c}\bar{\varepsilon}_{v_c} < \frac{(1-\mu_{cd})\mu_{cd}}{2\mu_{cd} + \varepsilon_{cd_m} - 1}$. The RHS of this inequality is decreasing in μ_{cd} because $\frac{\partial \ln RHS}{\partial \mu} = -\frac{\varepsilon_m + 1}{(1-\mu)(2\mu + \varepsilon_m - 1)} + \frac{1}{\mu} < -\frac{\varepsilon_m + 1}{2\mu + \varepsilon_m - 1} + \frac{1}{\mu} = \frac{-(1-\varepsilon_m)(1-\mu)}{\mu(2\mu + \varepsilon_m - 1)} < 0$. So the most binding inequality is at the highest μ_{cd} and ε_{cd_m}. At these values, the RHS is $\frac{b(1-b)}{2-3b}$. By assumption, $\mu_{u_c}\bar{\varepsilon}_v < \min \mu_v \min(1-\mu_v) \le b^2$ which is less than the RHS because $3b^2 - 3b + 1 > 0$ for $b \in (0,1)$. So $d\ln d_m / d\ln z < 0$ and the importer cutoff is lower for joint contracts. Finally, the price change is $d\ln p_{cd}/d\ln z = -d\ln(1-m_{cd})/d\ln z$ which gives $(1-m_{cd})\, d\ln p_{cd}/d\ln z = m_q q \frac{d\ln q}{d\ln z} + z = -\frac{m_q q}{\mu + \frac{m_q q}{1-m}} \left(\mu_u \frac{d\ln q_c}{d\ln z}\right) + \frac{\mu z}{\mu + \frac{m_q q}{1-m}} > 0$.

Having discussed the changes in observable outcomes across contracts, we will provide conditions for an increase in the cost cutoff for joint contracts after a reduction in import tariffs. We look at a reduction in ad valorem tariffs. If τT is the amount paid to exporters, then they receive just T. The exporter's problem can be written as $\Pi_c = \tau^{-1}\beta \int_0^{d_m} \left[(p_{cd} - cd)\, q_{cd} L_d - f_{cd}\right] dG_d + \tau^{-1}\beta\lambda_c L_d \left[z q_c - \int_0^{d_m} \nu(q_{cd}) dG_d\right] - z^2 f_c^J$ for a continuous choice of contracting denoted by z. Then the final markup is $m \equiv \mu_{v_{cd}} + \mu_{u_c}\bar{\varepsilon}_{v_c}/z \equiv \mu_v + y/z$. Firms will become more likely to enter into joint contracts if $\beta^{-1} \frac{d^2 \Pi_c}{d\tau^{-1} dz} = \frac{d \int_0^{d_m} [(p_{cd} - cd) q_{cd} L_d - f_{cd}]}{dz} - (1-\eta) \frac{d\ln Q}{d\ln \tau^{-1}} \frac{d \int_0^{d_m} pq}{dz} > 0$. We know that the first terms is positive for joint contracts that are viable. When revenues fall under joint contracts, as happens in our empirical work, we need to show that aggregate quantities rise with trade liberalization. Whenever aggregate trade volumes rise with trade liberalization, profits will be supermodular in contracting choice and the exporter cost cutoff will rise. We therefore provide conditions on primitives that ensure aggregate trade volumes rise after trade liberalization.

By definition, $\frac{d\ln y}{d\ln \tau} = \frac{\mu_u' q_c}{\mu_u} \frac{d\ln q_c}{d\ln \tau} + \int \frac{v' q (1-\mu-\bar{\varepsilon})}{\int v' q dG_d} \frac{d\ln q}{d\ln \tau} dG_d + \frac{(\varepsilon_m - \bar{\varepsilon}) v_m d_m}{\int v' q dG_d} \frac{d\ln d_m}{d\ln \tau}$. From optimal pricing, $-\left(\mu + \frac{\mu' q}{1-m}\right) \frac{d\ln q}{d\ln \tau} = (1-\eta) \frac{d\ln Q}{d\ln \tau} - \frac{y/z}{1-m} \frac{d\ln z}{d\ln \tau} + \frac{y/z}{1-m} \frac{d\ln y}{d\ln \tau} + \mu_u \frac{d\ln q_c}{d\ln \tau}$. From optimal contracting, $\ln y + (\eta - 1)\ln Q + \ln u' + \ln q_c - 2\ln z = \ln \tau + \ln 2 f_c^J / \beta L$ so that $2\frac{d\ln z}{d\ln \tau} = \frac{d\ln y}{d\ln \tau} - (1-\eta)\frac{d\ln Q}{d\ln \tau} + (1-\mu_u)\frac{d\ln q_c}{d\ln \tau} - 1$. From the optimal cutoff and substituting for optimal prices, $\frac{d\ln d_m}{d\ln \tau} = -\frac{1}{1-m_m}(1-\eta)\frac{d\ln Q}{d\ln \tau} - \frac{y/z\varepsilon_m}{1-m_m}\frac{d\ln y}{d\ln \tau} + \frac{y/z\varepsilon_m - 1}{1-m_m}\mu_u \frac{d\ln q_c}{d\ln \tau}$. From $\ln q_c = \ln \int_0^{d_m} v(q) dG_d$, the change in total quantities is $\frac{d\ln q_c}{d\ln \tau} = \frac{v_m d_m}{\int v dG_d} \frac{d\ln d_m}{d\ln \tau} + \int \frac{v' q}{\int v dG_d} \frac{d\ln q}{d\ln \tau}$. Substituting for the changes in optimal contracts, cutoffs and markups, we obtain the change in q_c in terms of changes in the aggregate quantity Q which can then be solved using the definition of $Q \equiv \int u(q_c)dG_c$. The change in the production cost cutoff is $\frac{d\ln c_m}{d\ln\tau}\int_0^{d_m} cdq_{cd}dG_d = -\int_0^{d_m}\left[(p_{cd}-cd)\,q_{cd}-f_{cd}/L\right]dG_d-(1-\eta)\,\frac{d\ln Q}{d\ln\tau}\int_0^{d_m} p_{cd}q_{cd}dG_d$. The change in exporter-level quantity is $\left[\ 1+\mu_u\frac{v_m d_m}{\int vdG_d}\frac{1-y/z\varepsilon_m}{1-m_m}+\mu_u\int\frac{v'q}{\int vdG_d}\frac{dG_d}{\mu+\frac{\mu' q}{1-m}}\left(1+\frac{1}{2}\left(1/\mu_u-1\right)\frac{y/z}{1-m}\right)+A_c\ \right]\frac{d\ln q_c}{d\ln\tau} = -A/2-A_Q\,(1-\eta)\,\frac{d\ln Q}{d\ln\tau}$ where

$A_y \equiv \left(\frac{v_m d_m}{\int vdG_d}\frac{y/z\varepsilon_m}{1-m_m}+\frac{1}{2}\int\frac{v'q}{\int vdG_d}\frac{1}{\mu+\frac{\mu' q}{1-m}}\frac{y/z}{1-m}dG_d\right)/\left(1+\int\frac{v'q(1-\mu-\bar{\varepsilon})}{\left(\mu+\frac{\mu' q}{1-m}\right)\int v'qdG_d}\frac{y/z}{1-m}dG_d+\frac{(\varepsilon_m-\bar{\varepsilon})v_m d_m}{\int v'qdG_d}\frac{y/z\varepsilon_m}{1-m_m}\right)$,

$A \equiv \int\frac{v'q}{\int vdG_d}\frac{1}{\mu+\frac{\mu' q}{1-m}}\frac{y/z}{1-m}dG_d + A_y\int\frac{v'q(1-\mu-\bar{\varepsilon})}{\left(\mu+\frac{\mu' q}{1-m}\right)\int v'q}\frac{y/z}{1-m}dG_d$,

$A_c \equiv A_y\left(\frac{\mu_u' q_c}{\mu_u}-\int\frac{v'q(1-\mu-\bar{\varepsilon})}{\left(\mu+\frac{\mu' q}{1-m}\right)\int v'qdG_d}\left(\mu_u-\frac{1}{2}\frac{y/z}{1-m}(1-\mu_u)\right)dG_d-\frac{(\varepsilon_m-\bar{\varepsilon})v_m d_m}{\int v'q}\frac{1-y/z\varepsilon_m}{1-m_m}\mu_u\right)$ and

$A_Q \equiv \frac{v_m d_m}{\int vdG_d}\frac{1}{1-m_m}+\int\frac{v'q}{\int vdG_d}\frac{1}{\mu+\frac{\mu' q}{1-m}}\left(1+\frac{1}{2}\frac{y/z}{1-m}\right)dG_d+A_y\left(\frac{v'q(1-\mu-\bar{\varepsilon})}{\left(\mu+\frac{\mu' q}{1-m}\right)\int v'qdG_d}\left(1+\frac{1}{2}\frac{y/z}{1-m}\right)dG_d+\frac{(\varepsilon_m-\bar{\varepsilon})v_m d_m}{(1-m_m)\int v'qdG_d}\right)$.

Then aggregate quantity rises after a reduction in tariffs as long as $A>0$ and $A_c>-1$ for $z\in[0,1]$.

Estimating quality

Table 10 shows the quality estimation procedure. Column (1) contains the Broda-Weinstein-Soderbery estimates and Column (2) contains the flexible method estimates. The elasticities of substitution are available at the HS 8-digit, and we take the mean for each HS 6-digit category in our data. We report estimates for the elasticities from the hybrid method of Soderbery. Results are similar with other elasticity estimates.

Table 10: Demand results: Quality Estimation

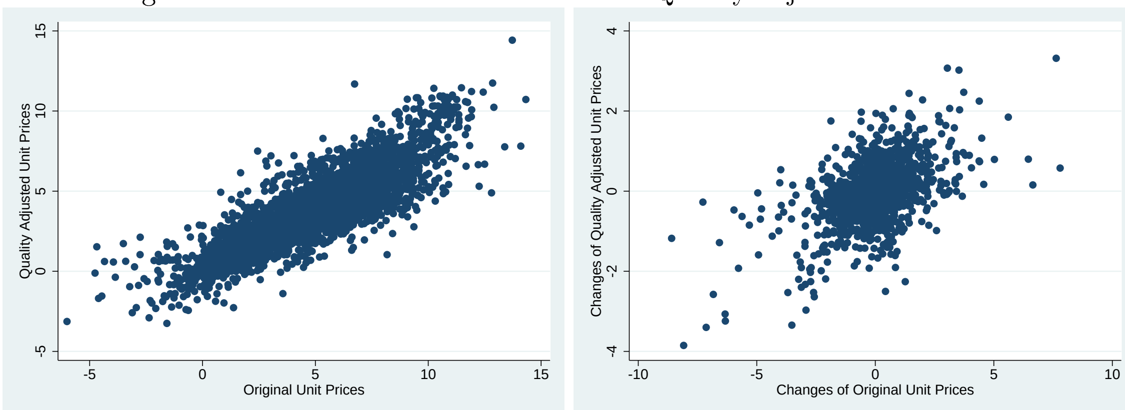
Dependent Variable: $\Delta \ln \text{Import Price}_{csht}$

Variable	**(1)** **Coef.** **(S.E.)**	**(2)** **Coef.** **(S.E.)**
$\ln \text{Import Quantity}_{cdsht}/\text{Elasticity}_h$	-0.377***	
	(0.002)	
$\ln \text{Import Quantity}_{cdsht}$		-0.297***
		(0.003)
$(\ln \text{Import Quantity}_{cdsht})^2$		0.009***
		(0.001)
$(\ln \text{Import Quantity}_{cdsht})^2$		-0.0003***
		(0.00003)
Exporter-Product-Year FE α_{csht}	yes	yes
Country-Year FE α_{st}	yes	yes
N	1,272,789	1,701,645
R^2	0.988	0.988

Note: Note: $\Delta \ln \text{Import Price}_{cdsht}$ is the change in import prices in logs from period $t-1$ to t and $\ln \text{Import Quantity}_{cdsht}$ is the change in import quantities. Elasticity_h is the Broda-Weinsten-Soderbery elasticity of substitution for product h. Significance levels are indicated by ***, **, and * for 1, 5 and 10 percent respectively.

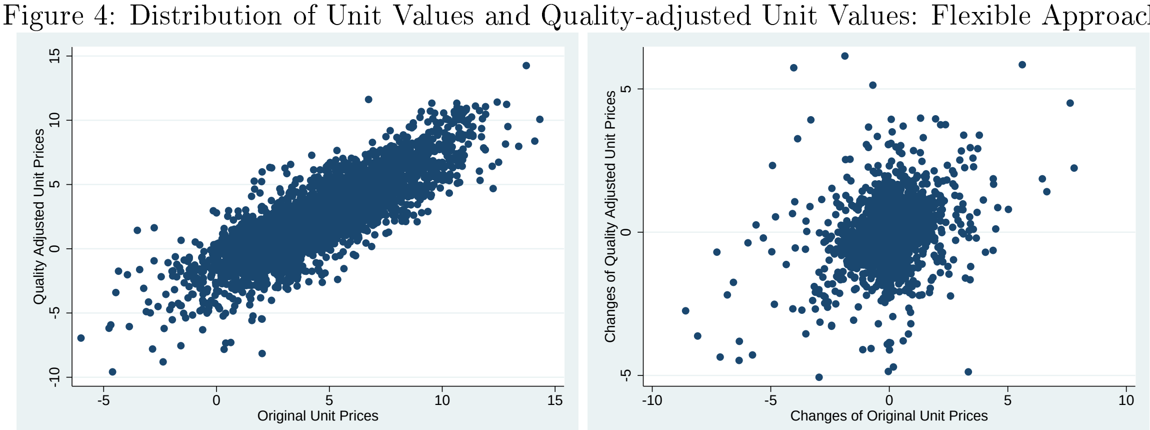
The correlation between the unit prices and the quality-adjusted units prices is about 0.8. This is shown in Figures 3 and 4 which show graph the scatter plots for the prices and the quality-adjusted prices for each method, and the changes in these variable which are used in the subsequent estimation.

Figure 3: Distribution of Unit Values and Quality-adjusted Unit Values



Note: Estimates based on the method of Khandelwal et al. (2013).

Figure 4: Distribution of Unit Values and Quality-adjusted Unit Values: Flexible Approach



Note: Estimates based on the flexible method of Section 4.