Product Bundling, Joint Markups and Trade Liberalization

Ji Hye Heo*

PRELIMINARY AND INCOMPLETE

PLEASE DO NOT CIRCULATE

August 15, 2022

Abstract

Product bundling is a frequent practice by multi-product firms to increase their profit. This paper examines how product bundling affects a firm's markups at various levels in international trade. Joint pricing decisions for multi-product firms with product bundling entail information about joint markups in their profit maximization problem. Utilizing this information, I propose a method to estimate joint markups for multi-product firms with product bundling at the transaction level. Focusing on Chinese exporter firms, the Chinese Customs data show that multi-product firms with product bundling enjoy higher markups than separate pricing firms on average. Exploiting China's WTO accession as a trade liberalization event, further analysis shows that multi-product firms with product bundling also benefit the most due to improved market access by the trade liberalization and have the lowest pass-through rate of costs to prices.

Keywords: Joint markups, multi-product firms, product bundling, pass-through, consumer valuation estimation, trade liberalization

JEL Codes: D22, D24, L11, F13, L60

*Vanderbilt University. e-mail : ji.hye.heo@vanderbilt.edu.

1 Introduction

Multi-product firms are generally larger and hold a significant share of transactions in international trade, as shown by Bernard et al. (2000). This implies that understanding the behavior of multi-product firms is central to characterizing the costs and benefits of trade policy change. However, while multi-product firms dominate trade, little is known about how they utilize their position as a multiple goods provider to retain their market power in a trade policy change. This paper examines product bundling as a possible channel for multi-product firms to retain their market power, captured by markups, in the event of trade liberalization.

One of the main advantages multi-product firms have over single-product firms is the ability to engage in product bundling, a practice whereby firms sell multiple goods in a single package. As a multiple goods provider, multi-product firms can sell their products independently with separate pricing or jointly with product bundling. If they allow an option for buyers to buy products separately and as a bundle, it is called *mixed bundling* practice. If only a bundle option is available, it is called *pure bundling* practice. Depending on the market structure and the firm's market power, a firm engaging in a bundling practice could price a bundle at a higher (*bundling premium*) or a lower (*bundling discount*) price.[1]

By studying how firm-level markups differ across a firm's pricing decisions, i.e., separate pricing firms such as single-product firms and multi-product firms without product bundling versus joint pricing firms with product bundling, this paper contributes to the literature studying firm-level international trade on three fronts. The first contribution is a novel methodology to flexibly identify joint markups at the transaction level for multi-product firms accounting for mixed bundling practice. The multi-product firm with mixed bundling practice makes joint pricing decisions for all of its single-product goods and bundles to maximize its profit at the firm level. Thus, the firm's first-order conditions from the profit maximization problem reflect a firm's joint pricing decision, which creates a markup linkage across goods. These expressions for markups are in terms of prices and the distribution of consumer valuations across goods. Transaction data can be used to estimate the distribution

[1] This paper focuses on mixed product bundling practice with price discount, the most prominent case in product bundling.

of consumer valuations, and then the joint markups can be recovered numerically using the information from the first-order conditions.

There have been broadly two approaches to estimate markups, which capture a firm's market power to price goods over marginal costs. The traditional approach is structural, where discrete choice models are used to construct the consumer's utility and estimate demand. Under the profit maximization assumption, product-level markups can be expressed in terms of price elasticity which can be recovered from the demand. The demand can be estimated using Berry et al. (1995) method, which requires data on sales, product characteristics, and market shares. The second approach, which utilizes information from the production side, is widely used due to less burden on the data requirement. This approach is pioneered by De Loecker and Warzynski (2012) and De Loecker et al. (2016), which use the firm's cost minimization problem to express firm-level and product-level markups in terms of the ratio between input cost to revenue shares and output elasticity. Then, only production data containing output and inputs are required to estimate the output elasticity and recover markups.[2]

The method this paper proposes to recover markups departs from the previous literature in that both approaches regard markups of each product for multi-product firms to be *independent* of other product markups. Therefore, characterizing the difference in markup strategy across firms with separate pricing and multi-product firms with joint pricing decisions has not been addressed by existing work. Product bundling is a prominent example where multi-product firms make joint pricing decisions. By recovering markups jointly for multi-product firms with mixed bundling practices, I can recover markups at the transaction level and examine systematic differences in markups between firms with different pricing decisions.

Secondly, this paper is among the first to study strategic bundling practices by multi-

[2]While the production side approach is widely used due to its simplicity, there have been many challenges. For example, if the production data does not contain price information but only revenue, only revenue elasticity can be obtained instead of the output elasticity. With the revenue elasticity, the expression for markups collapses to one; hence does not entail any information about markups. See Klette and Griliches (1996) and Bond et al. (2021). Also, there are discussions on the nonidentification issue where the markup is not identified using the proxy model (See Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2015) for the proxy method.) to estimate a production function and hence the output elasticity. See Flynn et al. (2019), Doraszelski and Jaumandreu (2019), and Jaumandreu (2018) for relevant discussions.

product firms in an international trade context. After the basic framework of Stigler (1963), Adams and Yellen (1976), and McAfee et al. (1989), the bundling literature has focused on either theoretically extending the basic framework[3] or analyzing the bundling practice in retail, telecommunication, and software products markets[4]. However, Iyoha et al. (2022) document that product bundling is also prevalent in international markets. Specifically, they find that 37.76% of transactions are for bundles, accounting for 43.49% of import values in Columbia between 2015 and 2019. Also, the data sample that I use for the main analysis shows that out of 1663 firms, the majority of firms (99.22%) have been multi-product firms for a given market in a given year. Among those, 71 firms have engaged in bundling for CPU and GPU products. The empirical analysis for computer parts(CPU and GPU) for Chinese exporters from 2000 to 2006 shows that multi-product firms with product bundling enjoy roughly 21.8% higher markups at the firm-yearly level compared to their counterparts without bundling practice. These differences in markups across multi-product firms with and without bundling practices may plausibly reveal how multi-product firms use bundling practices to retain their market power. Nevertheless, to my knowledge, there has yet to be a study of product bundling in an international trade setting to date.

Lastly, this paper adds to the literature studying changes in markups and competition in response to trade reforms. Changes in the competitiveness of a market force firms to revisit pricing decisions, particularly when firms exert market power. Trade policy changes, such as trade liberalization, can bring substantial changes to firm-level market power. De Loecker et al. (2016) study the impact of India's trade liberalization on markups, prices, and costs and found that while incomplete pass-through of input cost declines to price occurs, conditional on marginal cost, trade liberalization brings pro-competitive effects on markups. However, in their setting, each product's markup was assumed to be *independent* of other products'

[3]See Zhou (2017), and Zhou (2021) for pure and mixed bundling practice in a competitive setting where there is an arbitrary number of firms and Chen and Riordan (2013) for general conditions for the profitability of product bundling where copula is used to model the stochastic dependence of consumer values.

[4]Regarding software, see *United States v. Microsoft Corporation, 253 F.3d 34* court case where the U.S government accused Microsoft of illegally maintaining its monopoly position primarily through bundling PC with Internet Explorer. Also, Crawford and Yurukoglu (2012) study the short-run welfare in the television channel market when á la carte policies which require distributors to offer individual channels for sale to consumers are introduced. Simulation results show increased input costs offset consumer benefits from purchasing individual channels.

markups.

By recovering markups *jointly* for multi-product firms with product bundling, I determine how joint pricing determination impacts the firm's profitability after trade liberalization. Empirical analysis shows that for computer parts, trade liberalization brought an increase in markups for all types of firms, but the magnitude was the highest for bundling firms. This directly shows how multi-product firms may utilize bundling practices to retain their markups and market share after a trade policy change. This is partly due to the fact that multi-product firms with bundling practices have the lowest pass-through rate of costs to prices, and they are able to turn a decrease in input costs from trade liberalization to increase markups rather than decrease prices.

The structure of this paper is as follows. Section 2 presents an empirical framework to recover (joint) markups using information from transactions and a firm's pricing decisions. Section 3 describes the data sets used in the empirical analysis and describes China's WTO accession features with respect to computer parts. In section 4, empirical results are presented, and section 5 concludes.

2 A Framework to Estimate Markups

To estimate joint markups for multi-product firms with product bundling, I introduce an empirical model from the bundling literature. The multi-product firm with joint pricing decision will choose prices for all of its single-product goods and bundles to maximize its profit across all products simultaneously. Therefore, the firm's first-order conditions from the profit-maximizing problem reflect a markup linkage across goods the multi-product firm sells. This linkage expression for markups is expressed in terms of consumers' valuation across goods and the optimal price levels that the firm chose. Using transaction data and assumptions on the parametric structure of the consumer's valuations, the consumer's valuation across products can be estimated by inference functions for margins(IFM) procedure as Letham et al. (2014), which allows for an arbitrary correlation on the valuation of goods. Once consumer valuations across goods are obtained, markups for separate pricing firms are explicitly calculated, and joint markups for bundling firms are solved numerically.

For the rest of the paper, denote the set of individual goods (bundles) for firm f at year t as $\mathcal{G}_{ft}(\mathcal{B}_{ft})$ and let the number of components in the set be $G_{ft}(B_{ft})$ respectively. Let J_{ft} be the total number of products the multi-product firm sells, either as an individual products or a product bundle. For example, when firm f produces two discrete products and sells three products - both individual products and one product bundle of both single-products - we have the following : $J_{ft} = 3$, $\{1, 2\} \in \mathcal{G}_{ft}$ and $\{12\} \in \mathcal{B}_{ft}$. Theoretically, for a total number of individual products G_{ft}, the number of possible bundles is at most, $\sum_{b=2}^{G_{ft}-1} \binom{G_{ft}}{b}$.

Let c_{fjt} be a firm f's constant marginal cost for single product $j \in \mathcal{G}_{ft}$ at year t.[5] The marginal cost of a bundle is the sum of the marginal costs of its single product components. The price of a bundle is potentially offered at a discount relative to the sum of its components.[6] For example, in the case of $G_{ft} = 2$, $c_{fbt} = c_{f1t} + c_{f2t}$ and $P_{fbt} = P_{f1t} + P_{f2t} - d_{fbt}$ with $d_{fbt} > 0$ where subscript b refers to a bundled product made of product 1 and product 2. The multi-product firms that do not engage in bundling practices could be interpreted as having $d_{fbt} = 0$, effectively selling both goods simultaneously. Thus, while the subsequent discussion assumes multi-product firms bundle individual products, it could easily be applied to multi-product firms without bundling by setting $d_{fbt} = 0$.

Assume consumers for each firm desire at most one unit of each good and demands each good independently of their consumption of the other goods.[7] These consumers are residual demand of firm f at year t. For these consumers, consider consumer valuations of G_{ft} goods $\mathbf{v}_{ft} = (v_{f1t}, ..., v_{fG_{ft}t})$ that are distributed according to the unknown distribution function $\Psi_{ft}(\mathbf{v}_{ft})$.[8] Let $\psi_{ft}(\mathbf{v}_{ft})$ and $\psi_{fkt}(v_{fkt})$ be the probability density function and marginal density functions for product k for $\Psi_{ft}(\mathbf{v}_{ft})$. To avoid trivial cases, a positive measure of

[5]The assumption that marginal costs are constant is needed to construct the marginal cost for the bundle. This assumption can be relaxed for separate pricing firms to incorporate any returns to scale.

[6]In this framework, bundling premium in which a bundle is offered at a higher price compared to the sum of its component goods is not considered. Intuitively, consumers will always have the option to buy single product goods together rather than a bundle when there is a mixed bundling practice.

[7]The unit demand assumption is relaxed in the estimation for consumer valuation section by utilizing quantity information from the transaction data. This is explained in section 2.3.

[8]Consumer valuation distribution function Ψ can vary at various dimensions. The choice heavily depends on the observation number the data provides. In this paper, Ψ varies by firm, and yearly level to capture demand characteristics at the firm, and year level.

consumers exists such that $v_{fjt} \geq c_{fjt}$ for all j and resale by consumers is not possible.

This paper will focus on a benchmark case where $G_{ft} = 2$ to build on key intuition as transparently as possible. Then, I outline how to generalize the estimation procedure for the $G_{ft} > 2$ cases. Generalizing the estimation process for an arbitrary number of single products and bundles is a straightforward extension of the $G_{ft} = 2$ setting, albeit with substantially more derivations. In practice, bundled products do not typically contain many individual products, which eases the burden of derivation and any data restrictions.[9].

The subsequent section describes the framework under the assumption that transactions occur yearly, thus all subscripts are *ft*. However, suppose the data is detailed, and the transaction can be defined more precisely. In that case, the assumptions of units for marginal costs, consumer valuations, and prices could be relaxed. Hence markups can be recovered at a detailed level from firm-market-product-monthly to firm-yearly level depending on how precisely 'transaction' is defined.[10]

2.1 *Recovering Joint Markups with Product bundling*

To identify the joint markups, I introduce a framework from the bundling literature. The approach employs a model setting similar to McAfee et al. (1989), Chen and Riordan (2013) and builds upon Letham et al. (2014) for the estimation strategy. While bundling literature is introduced to help recover joint markups, this method can also recover markups for separate pricing firms.

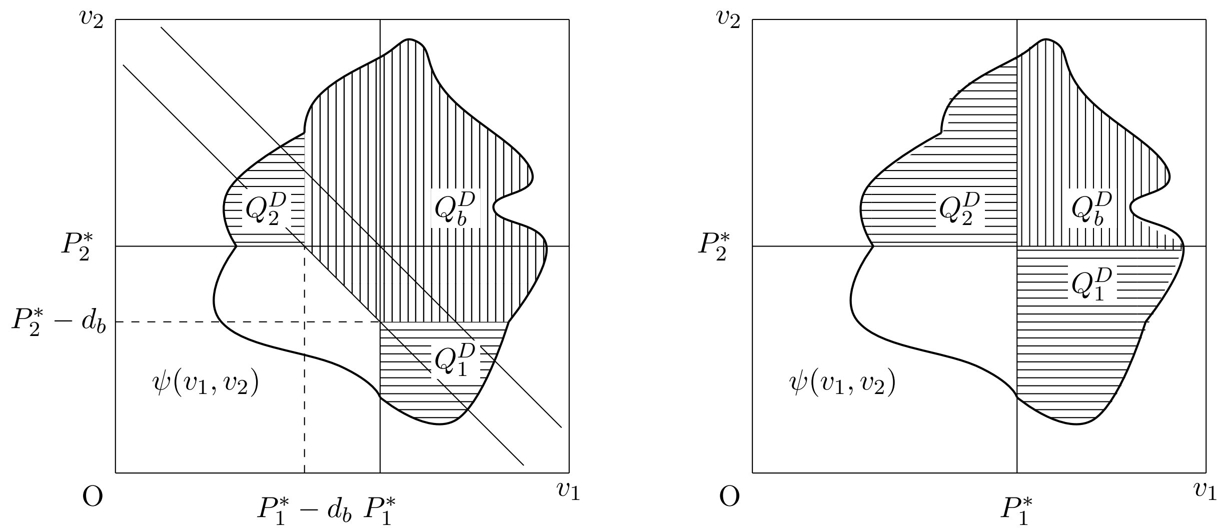
2.1.1 $G_{ft} = 2$ *Case*

Since consumers are rational, a given consumer will purchase product k from a firm f only if it gives her the highest utility among all other options. This enables me to write the

[9]Iyoha et al. (2022) find that most multi-product transactions have products less than four.

[10]The assumptions on the unit of the marginal costs, consumer valuations, and prices and hence the markups could be chosen appropriately depending on how detailed the 'transaction' is. For example, this paper defines the transaction as an exporter firm to the destination market at a monthly level. Thus, in this paper, the marginal costs can differ by destination market, i.e., c_{fdjt} to reflect shipping or market-specific marketing fee. However, it is not reasonable to assume that the marginal costs will differ at a monthly level; hence the time unit remains at a yearly level. The consumer valuations for each firm's residual demand also differ by destination market and year, i.e., Ψ_{fdt}, to reflect market-specific demand characteristics. The unit of price for each product follows the unit of transactions, i.e., P_{fdjm} where m is for a month.

Figure 1: Graphical illustrations : Joint Density Functions



(a) $\mathcal{G}_{ft} = 2$ with bundling (b) $\mathcal{G}_{ft} = 2$ without bundling

Note: This figure graphically depicts the quantity demanded for each j goods that depends on the joint density ψ and price variables. The left panel depicts case for multi-product firm with product bundling and the right panel shows the quantity demanded for a multi-product firm without bundling. For a multi-product firm without bundling, Q_{f1t}^D and Q_{f2t}^D refer to the quantities demanded for only products 1 and 2, respectively, and Q_{fbt}^D refers to the quantity demanded for both goods without discount. For both figures, subscripts f and t are dropped for space issues.

quantity demanded for each j goods(Q_{fjt}^D) in terms of prices and distribution of consumer valuations. For example, when $G_{ft} = 2$,

$$\begin{aligned}
&[Q_{f1t}^D \text{ when}] && v_{f1t} - P_{f1t} \geq max\{0, v_{f1t} + v_{f2t} - P_{fbt}\} \\
&[Q_{f2t}^D \text{ when}] && v_{f2t} - P_{f2t} \geq max\{0, v_{1t} + v_{f2t} - P_{fbt}\} \qquad (1) \\
&[Q_{fbt}^D \text{ when}] && v_{f1t} + v_{f2t} - P_{fbt} \geq max\{0, v_{f1t} - P_{f1t}, v_{f2t} - P_{f2t}\}
\end{aligned}$$

Denote a vector of prices for each j goods as $\mathbf{P}_{ft} = (P_{f1t}, P_{f2t}, P_{fbt})$ and combining each inequality and applying the definition of P_{fbt} gives the following expressions for the quantity demanded in equation (2), which are graphically illustrated in the left panel of Figure 1. The consumer whose valuation falls in the area marked as Q_1^D, Q_2^D, and Q_b^D is going to buy good 1, good 2, and the bundled good, respectively. Note that the quantity demanded for single goods $j = 1, 2$ are functions of not only its price but also the price of a bundled good b, and hence the price of the other good, explicitly showing the linkage across goods for firms with joint pricing.

$$
\begin{aligned}
Q^D_{f1t}(\mathbf{P}_{ft}) &= \int_{P_{f1t}}^{\infty}\int_0^{P_{fbt}-P_{f1t}} \psi_{ft}(x,y)dydx \\
Q^D_{f2t}(\mathbf{P}_{ft}) &= \int_0^{P_{fbt}-P_{f2t}}\int_{P_{f2t}}^{\infty} \psi_{ft}(x,y)dydx \\
Q^D_{fbt}(\mathbf{P}_{ft}) &= \int_{P_{f1t}}^{\infty}\int_{P_{fbt}-P_{f1t}}^{\infty} \psi_{ft}(x,y)dydx + \int_{P_{fbt}-P_{f2t}}^{P_{f1t}}\int_{P_{fbt}-x}^{\infty} \psi_{ft}(x,y)dydx
\end{aligned}
\tag{2}
$$

These expressions for the quantities demanded can be plugged into the firm's profit maximization problem. The profit-maximizing firm will simultaneously choose all prices $\mathbf{P}_{ft}$ to maximize its profit.

$$
\operatorname*{argmax}_{\mathbf{P}_{ft}} \Pi_{ft}(\mathbf{P}_{ft}) = \operatorname*{argmax}_{\mathbf{P}_{ft}} \Pi_{f1t} + \Pi_{f2t} + \Pi_{fbt} \tag{3}
$$

where $\Pi_{fkt} = (P_{fkt} - c_{fkt})Q^D_{fkt}(\mathbf{P}_{ft})$, for all $k \in \{1, 2, b\}$ and the analytical expression for $Q^D_{fkt}(\mathbf{P}_{ft})$ in terms of prices is derived from the rational consumer assumption as above. Thus, the profit function is as follows.

$$
\begin{aligned}
\Pi_{ft}(\mathbf{P}_{ft}) =& \sum_{j\in\{1,2,b\}} (P_{fjt} - c_{fjt})Q^D_{fjt}(\mathbf{P}_{ft}) \\
=&(P_{f1t} - c_{f1t})\int_{P_{f1t}}^{\infty}\int_0^{P_{fbt}-P_{f1t}} \psi_{ft}(x,y)dydx + (P_{f2t} - c_{f2t})\int_0^{P_{fbt}-P_{f2t}}\int_{P_{f2t}}^{\infty} \psi_{ft}(x,y)dydx \\
&+(P_{fbt} - c_{f1t} - c_{f2t})\left[\int_{P_{f1t}}^{\infty}\int_{P_{fbt}-P_{f1t}}^{\infty} \psi_{ft}(x,y)dydx + \int_{P_{fbt}-P_{f2t}}^{P_{f1t}}\int_{P_{fbt}-s}^{\infty} \psi_{ft}(x,y)dydx\right]
\end{aligned}
$$

Then the first-order conditions for price variables give the following three equations that express the relationship between markups across products in terms of consumer valuation $\psi_{ft}(v_{f1t}, v_{f2t})$ and price variables.[11]

$$
\begin{aligned}
&Q^D_{f1t}(\mathbf{P}_{ft}) - (1 - \mu^{-1}_{f1t})P_{f1t}\mathcal{A}_{ft} + ((1 - \mu^{-1}_{f2t})P_{f2t} - d_{fbt})\mathcal{B}_{ft} = 0 \\
&Q^D_{f2t}(\mathbf{P}_{ft}) - (1 - \mu^{-1}_{f2t})P_{f2t}\mathcal{C}_{ft} + ((1 - \mu^{-1}_{f1t})P_{f1t} - d_{fbt})\mathcal{D}_{ft} = 0 \\
&Q^D_{fbt}(\mathbf{P}_{ft}) - (1 - \mu^{-1}_{f1t})P_{f1t}(\mathcal{D}_{ft} + \mathcal{E}_{ft}) \\
&\qquad - (1 - \mu^{-1}_{f2t})P_{f2t}(\mathcal{B}_{ft} + \mathcal{E}_{ft}) + d_{fbt}(\mathcal{B}_{ft} + \mathcal{D}_{ft} + \mathcal{E}_{ft}) = 0
\end{aligned}
\tag{4}
$$

where for brevity, I denote $\mathcal{A}_{ft} = \int_0^{P_{fbt}-P_{f1t}} \psi_{ft}(P_{f1t}, y)dy$, $\mathcal{B}_{ft} = \int_{P_{f1t}}^{\infty} \psi_{ft}(x, P_{fbt} - P_{f1t})dx$,

[11] The derivation of these equations and equations used in the main analysis with a more detailed transaction level ($fdjm$) is included in the Appendix.

$\mathcal{C}_{ft} = \int_0^{P_{fbt}-P_{f2t}} \psi_{ft}(x, P_{f2t})dx$, $\mathcal{D}_{ft} = \int_{P_{f2t}}^{\infty} \psi_{ft}(P_{fbt} - P_{f2t}, y)dy$, and lastly $\mathcal{E}_{ft} = \int_{P_{fbt}-P_{f2t}}^{P_{f1t}} \psi_{ft}(x, P_{fbt} - x)dx$.[12] Once again, note that after estimating the consumer's valuation distribution $\psi_f(x, y)$, the first-order conditions provide the expression needed to identify joint markups.

Denote equations from equation (4) as $\mathbf{\Gamma}(P_{fj\in\{1,2,b\}t}, Q^D_{fj\in\{1,2,b\}t}, \psi_{ft}(\mathbf{v}_{ft}); \mu_{fj\in\{1,2\}t}) = 0$. Note $\mathbf{\Gamma}(\mu_{fj\in\mathcal{G}_{ft}t}) = 0$, is a three($J_{ft}$) by one vector of equations. Because we have two(G_{ft}) unknown joint markup parameters and three(J_{ft}) number of equations, it is over-determined. I propose to recover joint markups by solving $\mathbf{\Gamma}(\mu_{fj\in\mathcal{G}_{ft}t}) = 0$ numerically and choosing the set of $\mu_{fj\in\mathcal{G}_{ft}t}$ that minimizes error below a given threshold level. The existence of a sufficiently small threshold level will filter out any cases where there is no solution for $\mu_{fj\in\mathcal{G}_{ft}t}$.

2.1.2 *General Case with* $G_{ft} > 2$

Here I provide a general approach for deriving markup expressions across goods in the $G_{ft} > 2$ cases. As noted before, once the number of single product goods exceeds two, the total number of possible combinations of single goods to make a bundled product becomes $\sum_{b=2}^{G_{ft}-1} \binom{G_{ft}}{b}$. This means even if firms have identical $\mathcal{G}_{ft}$, i.e., the same individual goods, they might have different bundled goods, i.e., different $\mathcal{B}_{ft}$. Thus, when G_{ft} exceeds two, treat it as if all firms offer all possible combinations of a bundle. That is $B_{ft} = \sum_{b=2}^{G_{ft}-1} \binom{G_{ft}}{b}$ for all firms. Then the discount value for combinations of goods that are not bundled can be set at zero like before in the case of multi-product firms without bundling practices.

Thus, for the $G_{ft} > 2$ cases, we can follow the same steps as $G_{ft} = 2$ case. First, construct

[12] For multi-product firms without bundling, taking first-order conditions with respect to only (P_{f1t}, P_{f2t}) or plugging in $d_{fbt} = 0$ to equation (4) results in an identical result to the following two first-order conditions. In this case, the joint pricing from product bundling is removed; hence, product markups are independent of one another as in previous literature.

$$
\begin{aligned}
Q^D_{f1t}(P_{f1t}, P_{f2t}) + Q^D_{fbt}(P_{f1t}, P_{f2t}) - (1 - \mu^{-1}_{f1t})P_{f1t}(\mathcal{A}_{ft} + \mathcal{D}_{ft}) = 0 \\
Q^D_{f2t}(P_{f1t}, P_{f2t}) + Q^D_{fbt}(P_{f1t}, P_{f2t}) - (1 - \mu^{-1}_{f2t})P_{f2t}(\mathcal{B}_{ft} + \mathcal{C}_{ft}) = 0
\end{aligned}
$$

the following profit-maximizing problem for a firm f.

$$\underset{\mathbf{P}_{ft}}{\operatorname{argmax}}\, \Pi(\mathbf{P}_{ft}) = \underset{\mathbf{P}_{ft}}{\operatorname{argmax}} \sum_{k \in \mathcal{G}_{ft} \cup \mathcal{B}_{ft}} \Pi_{fkt} \tag{5}$$

where $\Pi_{fkt} = (P_{fkt} - c_{fkt})Q^D_{fkt}$, for all $k \in \mathcal{G}_{ft} \cup \mathcal{B}_{ft}$. Second, using the rational consumer assumption, derive expressions for quantity demanded, i.e., $Q^D_{fkt}(\mathbf{P}_{ft})$ for $k \in \mathcal{G}_{ft} \cup \mathcal{B}_{ft}$. Note that for $k \in \mathcal{G}_{ft}$, $Q^D_k(\mathbf{P}_{ft})$ should be expressed in terms of its price, and the price of bundled goods where k is a component of. For $k \in \mathcal{B}_{ft}$, $Q^D_k(\mathbf{P}_{ft})$ should be a function of its price and the price of all of the individual products of which bundle k is composed. After deriving expressions for the quantity demanded, plug it into the profit function to derive J_{ft} number of first-order conditions with joint markups. Denote it as $\mathbf{\Gamma}(P_{fj \in \mathcal{G}_{ft} \cup \mathcal{B}_{ft} t}, Q_{fj \in \mathcal{G}_{ft} \cup \mathcal{B}_{ft} t}, \psi_{ft}(\mathbf{v}_{ft}); \mu_{fj \in \mathcal{G}_{ft} t}) = 0$ and recover joint markups numerically as $G_{ft} = 2$ case.

2.2 *Recovering Markups for Separate Pricing*

Firms with separate pricing decisions maximizes firm-level profit by maximizing profits from each product independently.[13] Thus, their first-order conditions from each product-level profit entail information about *independent* product markups. The profit maximization problem for product j is

$$\underset{P_{fjt}}{\operatorname{argmax}}\, \Pi_{fjt}(P_{fjt}) = \underset{P_{fjt}}{\operatorname{argmax}} (P_{fjt} - c_{fjt})Q^D_{fjt}(P_{fjt}) \tag{6}$$

where the quantity demanded for separate pricing firms is $Q^D_{fjt}(P_{fjt}) = \int_{P_{f1t}}^{\infty} \psi_{fjt}(x)dx$. Then the first-order condition for price gives the following equation for markups in terms of consumer valuation $\psi_{fjt}(v_{fjt})$ and prices.

$$\mu_{fjt} = \frac{P_{fjt}\psi_{fjt}(P_{fjt})}{P_{fjt}\psi_{fjt}(P_{fjt}) - (1 - \Psi_{fjt}(P_{fjt}))}$$

Once again, note that after estimating consumer's valuation distribution $\psi_{ft}(x, y)$, $\psi_{fjt}(x)$ and markups can be calculated.

[13] In this paper, separate pricing firms include single- and multi-product firms without bundling practices. Markups for multi-product firms without bundling could be recovered using the previous method of joint pricing with $d_{fbt} = 0$ or with this separate pricing method.

2.3 *Consumer Valuation Estimation*

This section describes how the consumer valuation distribution ψ is estimated using following the approached proposed by Letham et al. (2014). Those who are not interested can skip this section. The joint probability density function, $\psi_{ft}(\mathbf{vt})$, describes how the consumer's valuation for each product is distributed as well as how it is correlated with their valuations of other products at the firm-year level. Because it differs at the firm level, it can capture cross-sectional differences in firms, such as consumer types and quality (hence price). While making specific assumptions on the correlation structure across goods for $\psi_{ft}(\mathbf{vt})$ is possible[14], if the relationship between goods affects markups in a meaningful way, specific assumptions will likely distort the estimation of markups.[15] Using transaction data, Letham et al. (2014) propose a statistically consistent inference procedure using copulas to recover correlated consumer valuations. The key intuition is to put a parametric assumption on the joint density function's marginal distributions and choose a specific copula function that will fit the overall correlation structure well. The marginal distribution will contain information on the valuation's marginal structure; hence, the demand for each product can be recovered from the marginal distribution afterward. After parameters for the marginal distributions are estimated, the copula parameter is estimated using these marginal parameters to fit the data in a maximum likelihood sense.

Define a transaction as a deal between seller and buyer during a certain period of time. In a retail setting where consumers buy goods often in small amounts, each day would be a good choice for the period of time. In trade, where buyer firms purchase goods in large amounts from specific sellers, monthly or yearly may be an aqueduct period of time depending on goods of interest. Consider a set of transaction data that consists of two components. One component is purchase data, $\mathbf{y}^s = [y_1^s, ..., y_{G_{ft}}^s]$, where y_j^s is 1 if item j is sold in transaction s and 0 otherwise.[16] The other component is the price data for individual products at

[14] See Letham et al. (2014) for a survey of studies that made either independent or perfectly correlated assumptions on the correlation structure across goods.

[15] While it is described in terms of profits rather than markups, Letham et al. (2014) shows how imposing an independent correlation assumption could lead to very different predictions on possible profit when bundled product is introduced.

[16] The item j here is a unit product with a quantity equal to one. The unit demand assumption is relaxed by treating q units of a product sold as 1 unit of a product sold q times during the estimation procedure.

transaction s, $\mathbf{P}^s = [P_1^s, ..., P_{G_{ft}}^s]$. Let S denote the total number of transactions. Since consumers maximize utility, $y_j^s = 1$ if and only if $v_j^s \geq P_j^s$. This relationship provides a model for the relationship between the latent variable valuations v_j^s and transaction data (y_j^s, P_j^s).

The copula $\mathbb{C}_{ft}(\cdot)$ for $\Psi_{ft}(\cdot)$ is a distribution function over $[0,1]^{G_{ft}}$ with uniform margins such that $\Psi_{ft}(v_{f1t}, ..., v_{fG_{ft}t}) = \mathbb{C}_{ft}(\Psi_{f1t}(v_{f1t}), ..., \Psi_{fG_{ft}t}(v_{fG_{ft}t}))$. The copula $\mathbb{C}_{ft}$ contains all information on the dependence structure between the components of $(v_{f1t}, ..., v_{fG_{ft}t})$ and combines each marginal distribution Ψ_{fkt} in a way to return the joint distribution Ψ_{ft}. Suppose each marginal distribution is a function of parameters $\boldsymbol{\theta}_{fjt}$, i.e, $\Psi_{fjt}(v_{fjt}; \boldsymbol{\theta}_{fjt})$ and the copula distribution belongs to a family with parameters $\boldsymbol{\phi}_{ft}$, i.e, $\Psi_{ft}(\boldsymbol{v}_{ft}; \boldsymbol{\theta}_{ft}, \boldsymbol{\phi}_{ft}) = \mathbb{C}_{ft}(\Psi_{f1t}(v_{f1t}; \boldsymbol{\theta}_{f1t}), ...\Psi_{fG_{ft}t}(v_{fG_{ft}t}; \boldsymbol{\theta}_{fG_{ft}t}); \boldsymbol{\phi}_{ft})$. Letham et al. (2014) propose an inference functions for margins (*IFM*) procedure that is similar to pseudo-maximum likelihood estimation where we choose parametric forms for the margins $\Psi_{fjt}(\cdot)$ and copula $\mathbb{C}_{ft}$, then find the parameters for which $\mathbb{C}_{ft}(\Psi_{f1t}(v_{f1t}), ..., \Psi_{fG_{ft}t}(v_{fG_{ft}t}))$ is the closest to $\Psi_{ft}(v_{f1t}, ...v_{G_{ft}t})$ in a likelihood sense.

The optimization can be performed in two steps. First, each marginal distribution is fit independently to recover $\hat{\theta}_{fjt}$. In the second step, estimated marginal distributions are used to fit the correlation structure ϕ_{ft}.

$$\hat{\boldsymbol{\theta}}_{fjt} \in \underset{\boldsymbol{\theta}_{fjt}}{\operatorname{argmax}}\, l_{fjt}(\boldsymbol{\theta}_{fjt}) \quad j = 1, ..., G_{ft} \tag{7}$$

$$\hat{\boldsymbol{\phi}}_{ft} \in \underset{\boldsymbol{\phi}_{ft}}{\operatorname{argmax}}\, l_{ft}(\hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) \tag{8}$$

The likelihood function for each marginal distribution in equation (7) is derived from the observed purchase patterns of the utility-maximizing consumer. Let $\mathfrak{p}_{fj}(P_j^s)$ be the purchase probability for item j at price P_j^s which is equivalent to demand model for item j. Then the demand and inverse marginal valuation distribution functions have the following relationship.

$$\mathfrak{p}_{fj}(P_j^s) = \mathbb{P}(y_j^s = 1) = \mathbb{P}(v_j^s > P_j^s) = 1 - \Psi_{fjt}(P_j^s; \boldsymbol{\theta}_{fjt})$$

Therefore, the likelihood function can be constructed with Bernoulli distribution for y_j^s such that $y_j^s \sim Bernoulli(1 - \Psi_{fjt}(P_j^s; \boldsymbol{\theta}_{fjt}))$ and resulting in following likelihood function for given data $\{P_j^s, y_j^s\}_{s=1}^S$.

$$l_{fjt}(\boldsymbol{\theta}_{fjt}) = \sum_{s=1}^{S} (y_j^s \log(1 - \Psi_{fjt}(P_j^s; \boldsymbol{\theta}_{ft}))) + (1 - y_j^s) \log(\Psi_{fjt}(P_j^s; \boldsymbol{\theta}_{ft})) \tag{9}$$

The relationship between the marginal distribution and the demand model provides a natural selection criterion for the marginal distributions. For example, as Letham et al. (2014) has stated, if the demand model is linear, the corresponding valuation distribution is the uniform distribution. If the demand model follows the normal distribution function, the corresponding marginal valuation distribution also follows the normal distribution. For empirical analysis, I follow Letham et al. (2014) in using uniform distributions for the marginal distributions and Gaussian copula function.

Once the marginal parameters $\boldsymbol{\theta}_{ft}$ are estimated by maximizing equation (9), these estimators are used to obtain an estimate of the copula parameters $\boldsymbol{\phi}_{ft}$ along with data.

$$\begin{align}
l_{ft}(\hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) &= \sum_{s=1}^{S} \log \mathfrak{p}_f(\mathbf{y}^s | \mathbf{P}^s_{G_{ft}}, \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) \tag{10}\\
&= \sum_{s=1}^{S} \log \int \mathfrak{p}_f(\mathbf{y}^s | \mathbf{v}^s, \mathbf{P}^s_{G_{ft}}, \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) \mathfrak{p}_f(\mathbf{v}^s | \mathbf{P}^s_{G_{ft}}, \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) d\mathbf{v}^s \tag{11}\\
&= \sum_{s=1}^{S} \log \int_{v^{s,l}_{G_{ft}}}^{v^{s,u}_{G_{ft}}} \cdots \int_{v_1^{s,l}}^{v_1^{s,u}} \psi_{ft}(v_1^s, ..., v^s_{G_{ft}}; \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) dv_1^s ... dv^s_{G_{ft}} \tag{12}\\
&= \sum_{s=1}^{S} \log \sum_{k=0}^{G_{ft}} (-1)^k \sum_{I \subseteq \{1,...,G_{ft}\}, |I|=k} \Psi_{ft}(\mathbf{v}^s; \hat{\boldsymbol{\theta}}, \boldsymbol{\phi}) \tag{13}
\end{align}$$

where the equality from equations (11) to (12) use $\mathfrak{p}_f(\mathbf{v}^s | \mathbf{P}^s_{G_{ft}}, \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) = \mathfrak{p}_f(\mathbf{v}^s | \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft}) = \psi_{ft}(\cdot; \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft})$ and make use of the lower and upper limits of the integration as follows.

$$v_j^{s,l} = \begin{cases} -\infty & \text{if} \quad y_j^s = 0 \\ P_j^s & \text{if} \quad y_j^s = 1 \end{cases} \qquad v_j^{s,u} = \begin{cases} P_j^s & \text{if} \quad y_j^s = 0 \\ \infty & \text{for} \quad y_j^s = 1 \end{cases}$$

The representation of the likelihood formula in equation (12) is intractable due to multiple integrals. Letham et al. (2014) employed the rectangular integral of the probability density function to derive equation (13) where,

$$\tilde{v}_j^s(I) = \begin{cases} v_j^{s,l} & \text{if} \quad j \in I \\ v_j^{s,u} & \text{if} \quad j \notin I \end{cases}$$

Thus, the complete, statistically consistent inference procedure to estimate the consumer valuation distribution $\psi_{ft}(\mathbf{v})$ is,

$$\hat{\boldsymbol{\theta}}_{fjt} \in \underset{\boldsymbol{\theta}_{fjt}}{\operatorname{argmax}} \sum_{s=1}^{S} (y_j^s \log(1 - \Psi_{fjt}(P_j^s; \boldsymbol{\theta}_f))) + (1 - y_j^s) \log(\Psi_{fjt}(P_j^s; \boldsymbol{\theta}_f))$$

$$\hat{\boldsymbol{\phi}}_{ft} \in \underset{\boldsymbol{\phi}}{\operatorname{argmax}}_f \sum_{s=1}^{S} \log \sum_{k=0}^{G_{ft}} (-1)^k \sum_{I \subseteq \{1,\dots,G_{ft}\}, |I|=k} \Psi_{ft}(\tilde{\mathbf{v}}^s; \hat{\boldsymbol{\theta}}_{ft}, \boldsymbol{\phi}_{ft})$$

3 Data and Trade Policy Background

This section gives general description of the data and the policy background. Section 3.1 describes the Chinese Customs data that is use to explore the markup behavior across firms, time, and international markets. This subsection also describes how firms are classified into separate pricing and bundling firms and how multi-product transactions are classified into either bundling or simply a transaction with multiple products. This classification, as well as recovering markups and estimating the consumer valuation for multi-product firms, require all the price information available to the consumer at the moment of transaction. That is, while the price data is available for only products sold at a given transaction, price data for other products (including the bundle) is needed. Section 3.1.2 explains how these unobserved price data are imputed using the Customs data's monthly feature and then how firms and transactions are categorized. Basic features of the Chinese WTO accession, such as tariff changes and changes in a firm's market access, are described briefly in section 3.2.

3.1 *Data*

3.1.1 Transaction Data, Firm Classifications, and Product Choice

I take advantage of the Chinese Customs data(CCD) that the Chinese Customs Office collects to recover joint markups for multi-product firms and examine the relationship between markups and firm heterogeneity. The CCD records Chinese firm-level exports between 2000 and 2006 at the market-monthly level with corresponding HS6 codes, quantities, values, and firm characteristics such as firm names, firm types, addresses, and cities. There are a few things to note about this data set. Firstly, because this is customs data, all the empirical analysis is focused on exporter firms and their export transactions. De Loecker and Warzynski (2012) show that exporter firms, on average, have higher markups compared to their counterpart domestic firms. However, if the difference is not systematically different across firms with different pricing strategies, then focusing on exporters will not create significantly different results compared to domestic firms. Secondly, estimating the consumer valuation requires transaction data where ideally, one will have transactions at the seller firm to buyer firm in a given month. Unfortunately, the basic transaction unit in the CCD is seller firm to market in a given month. The aggregation at the buyer side could lead to misclassifying multiple single goods transactions into bundle goods transactions. To address this issue, a basic transaction is defined at the monthly level rather than aggregated to the yearly level to reduce the misclassifying error. Also, because a transaction with multiple goods needs to have a positive discount to be considered a 'bundle', misclassification errors will create downward pressure on the markup for the bundling firm. This means the empirical analysis will be carried out conservatively. Lastly, as mentioned in section 2, the empirical analysis is carried out at the market because CCD specifies the destination market at the monthly level. Hence, in the main empirical analysis, the markups are recovered at the firm-market-product-monthly level and aggregated to the firm-yearly level.

To examine how product bundling affects firm's markups, firms need to be classified into bundling and non-bundling firms. Firstly, firms are categorized into single-, or multi-product firms depending on how many goods they sell to each destination market at a give year. This classification consider firm types from the sales side rather than production side. For example, if a firm f produced multiple products but sold only a single HS6 code product to destination market d at a given year t, then firm f is classified as single-product firm in market d at year t. After firms are divided into single-, and multi-product firms, unobserved prices are imputed for multi-product firms that sold both products of interest to further divide them into bundling and non-bundling firms. With imputed price, a multi-product firm is categorized as a bundling firm if it offered both products at a discount.

In this paper, CPU (central processing unit) and GPU (graphics processing unit) are selected for the analysis.[17] The products are selected based on the following criteria. Firstly, there have to be enough observations. Electrical machines are one of the most exported goods from China during the sample period. Also, the relationship between goods must be considered. Goods are chosen where firms are likely to produce both and sell them as a bundle. CPU and GPU are both processing units that are essential parts of a computer. A basic description of the data for CPU and GPU is summarized in Table 1.

Table 1: Summary Statistics

	Observation	(%)
Firm Characteristics(fdt)		
Number of Firms	1,726	100.0%
Single-product Firms	178	10.31%
Multi-product Firms without Product bundling	1,660	96.18%
Multi-product Firms with Product bundling	85	4.93%
Transaction Characteristics($fdjm$)		
Number of Transactions	51,882	100.0%
$MPT_{fdm} = 0$	49,787	97.06%
$MPT_{fdm} = 1$	2,095	2.94%
$Bundling_{fdm} = 1$ in $MPT_{fdm} = 1$	932	30.79%

Note: Following are explanations for the subscripts: f is for firms, d is for destination markets, j is for products(CPU and GPU), and m and t are time subscripts that each stand for month and year. MPT_{fdm} is a dummy that refers to multi-product transactions that consist of both CPU and GPU between firm f and market d for month m.

Table 1 shows that multi-product firms are the majority firm types in the data sample. Specifically, out of 1,763 firms, 1,660 firms, at one point, have been a multi-product firms. Among those multi-product firms, 85 firms engage in product bundling with CPU and GPU goods. Only 9.4% of firms have changed their status during the sample period and majority of firms remained with their original type. Table 1 also shows that while majority of transactions are single-product transactions, for multi-product transactions 30.79% of them are bundling transaction.

[17]The products are classified at the HS6 code level. Specifically, CPU is {847130, 847141, 847149, 847150, 847160, 847170} and GPU is {847180}.

3.1.2 Price Imputation

This subsection describes how unobserved price data are imputed using the monthly feature of the Customs data to classify firms into bundling and non-bundling firms and estimate markups. The key intuition is to impute the missing price data using the observed price data from the closest period and the definition of the bundle price. Those who are not interested in details can skip this part.

For a firm f at a market d, let (y_1, y_2) denote the dummy variables for selling product 1 and product 2 and (p_1, p_2) be the corresponding *observed* price for a transaction. Let (x_1, x_2, x_b, d) be the final *imputed* prices for product 1, product 2, product bundle, and bundling discount for the transaction that will be used to estimate consumer valuations and recover markups. The bundling discount is calculated as $d = (x_1 + x_2) - x_b$. If a transaction has $(y_1, y_2) = (1, 1)$ and $d > 0$, then that transaction is transaction with product bundling. Transactions with either $(y_1, y_2) \neq (1, 1)$ or $d = 0$ is not classified as bundling. Thus, transactions with $(y_1, y_2) = (1, 1)$ but $d = 0$ is simply transaction that happened to have multiple products and are not classified as bundling transactions.

Firstly, for a given transaction, if price is observed, then imputed price is simply observed price itself, i.e., $x_j = p_j$. For example, in the case of $(y_1, y_2) = (1, 0)$, $x_1 = p_1$ and for $(y_1, y_2) = (1, 1)$, $x_b = p_1 + p_2$. If price is not observed for product j, this means $y_j = 0$ for $j = 1, 2$. Then for a given firm-market-year, find the closest transaction where only $y_j = 1$ and $y_{-j} = 0$ where $-j$ denote the other good. If there is no such transaction, find the closest transaction with $(y_1, y_2) = (1, 1)$. Use the price from found transaction, i.e, $x_j = p_j$ where p_j is from found transaction. For $j = b$, find the closest transaction with $(y_1, y_2) = (1, 1)$ and impute it as $x_b = p_1 + p_2$ where prices are from found transaction. If there are no transactions with $(y_1, y_2) = (1, 1)$, simply put it as $x_b = x_1 + x_2$.

Table 2 presents a basic example of how the procedure looks like. To show the process more clearly, multi-product firms that sold both CPU and GPU to market d at a given year t are divided into five groups. Firms that have sold only good j and and both goods are classified into category j. That is firms in group 1 have the following; $(y_1, y_2) = \{(1, 0), (1, 1)\}$. Firms that have only single-product transactions, i.e.,

Table 2: Price Imputation Example

	y_1	y_2	p_1	p_2	x_1	x_2	x_b	d	Bundle?
Group 1	1	0	80	-	80	100	170	10	No
	1	1	70	100	80	100	170	10	Yes
Group 2	0	1	-	120	70	120	170	20	No
	1	1	70	100	70	120	170	20	Yes
Group 3	1	0	80	-	80	120	200	0	No
	0	1	-	120	80	120	200	0	No
Group 4	1	1	60	100	60	100	160	0	No
	1	1	70	120	70	120	190	0	No
Group 5	1	0	80	-	80	120	170	30	No
	0	1	-	120	80	120	170	30	No
	1	1	70	100	80	120	170	30	Yes

Note: This table shows how price imputation for unobserved prices are carried out with a simple example. A transaction is classified as a bundling if it have $(y_1, y_2) = (1, 1)$ and $d > 0$. Note that by construction, firms in group 3 and 4 can never be classified into bundling firms by design.

$(y_1, y_2) = \{(1, 0), (0, 1)\}$ are in group 3. Firms that only sells in multi-product transactions are in group 4. Lastly, firms that have sold all composition of goods are classified into group 5, that is, they have $(y_1, y_2) = \{(1, 0), (0, 1), (1, 1)\}$. Note that by construction, firms in group 3 and 4 can never be classified into bundling firms by design.[18]

3.2 *Background*

China's WTO accession, which took place in 2001, has included substantial tariff reductions (see Lu et al. (2015)) and improved market access. In this section, I document the impact of China's trade liberalization on CPU and GPU using tariff data from WITS TRAINS and trade values from UN COMTRADE. To examine the tariff reductions and improved overall market access for CPU and GPU, I focus on the top 30 markets where China had the most transactions for CPU and GPU. They account for 95.5% of quantities and 96.6% of trade value.[19]

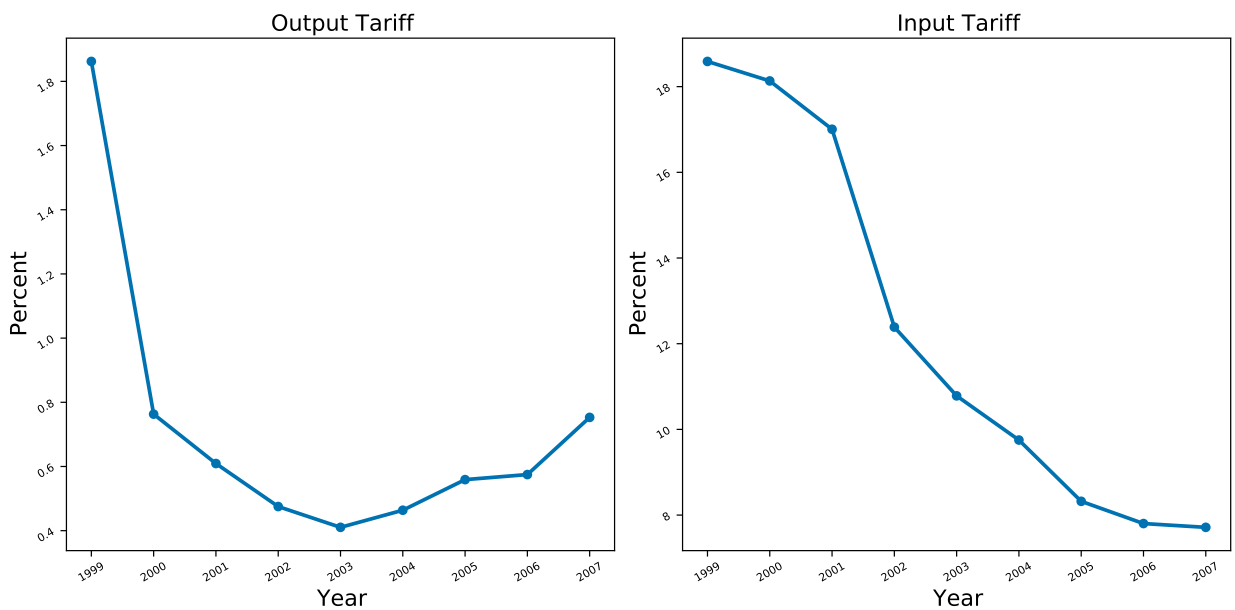
[18]This is to reduce misclassifying non-bundling firms into bundling firms.

[19]These markets are Hong Kong, USA, Japan, Taiwan, Netherlands, Singapore, Germany, UK, South Korea, Australia, Malaysia, France, India, Thailand, UAE, Canada, Italy, Spain, Philippines, Brazil, Mexico, Belgium, South Africa, Israel, Turkey, Finland, New Zealand, Ireland, Indonesia, Poland in order of frequencies.

3.2.1 Tariff

Figure 2 displays how China's aggregated output and input tariffs have evolved. For each market, output tariffs are at the HS2 code level. These output tariffs for each market are aggregated using each market's trade value as weights to construct the aggregated output tariffs. For input tariffs, I follow De Loecker et al. (2016) and construct input tariffs for each market by passing the tariff data at ISIC Rev3 level to China's input-output matrix table for 1995-2010 and use value as weights to create the aggregated input tariff.[20][21] Figure 2 shows that trade liberalization brought a sharp decline in both output and input tariffs for CPU and GPU.

Figure 2: Tariffs for CPU and GPU from 1998 to 2007



Note: This figure plots aggregated output and input tariffs for China from 1998 to 2007 for HS2 level 84, which contains CPU and GPU.

3.2.2 Market Access

Another way to examine how China's WTO accession impacted the computer parts for Chinese exporters is to evaluate the market access improvement. I borrow intuition from

[20] The formal definition of the input tariff is as follows. $\tau_{idt}^{input} = \sum_k a_{ki}\tau_{kdt}^{output}$ where τ_{kdt}^{output} is the export tariff for market d to China on industry k at time t and a_{ki} is the share of industry k in the value of industry i from the input-output table.

[21] Output and input tariffs for each market are displayed in the appendix.

Fugazza and Nicita (2011) and define two terms for market access at the firm level. Let $DMAC_{fdt}$ and $RMAC_{fdt}$ refer to direct market access conditions and relative market access conditions for Chinese firm f for market d at time t. These indices are defined as follows.

$$DMAC_{fdt} = \frac{\sum_j exp_{fjdt}\eta_{fjdt}\tau_{jdt}}{\sum_j exp_{fjdt}\eta_{fjdt}}, \qquad RMAC_{fdt} = \frac{\sum_j exp_{fjdt}\eta_{fjdt}(\tau_{jdt}^{others} - \tau_{jdt})}{\sum_j exp_{fjdt}\eta_{fjdt}} \tag{14}$$

where exp_{fjdt} denotes the amount Chinese firm f export to market d for product j at time t, η_{fjdt} is firm f's import elasticity for good j for market d, τ_{jdt} is a tariff that Chinese firms face in market d for product j at time t and $\tau_{jdt}^{others} = \frac{\sum_o exp_{ojdt}\tau_{jt}^o}{\sum_o exp_{ojdt}}$ captures weighted average tariffs faced by competitor firms from other country o that have access to market d. Thus, $DMAC_{fdt}$ captures the overall tariff faced by an exporter while $RMAC_{fdt}$ captures the overall tariff faced by an exporter relative to that faced by competitors from different countries.

[[MARKET ACCESS FIGURES GO HERE]]

Figure **??** shows that average $DMAC_{fdt}$ decreased considerably during 1999 when China undertook many tariff cuts to meet the WTO criteria. While the average $DMAC_{fdt}$ increased slightly afterward, the general trend indicates that China's market access improved during the sample period. Similarly, the average $RMAC_{fdt}$ increased and remained higher, indicating that relative market access disadvantages that Chinese exporter firms faced improved.

4 Empirical Analysis

4.1 Markup Descriptions

As described in section 2, markups for separate pricing firms are calculated from the first-order condition derived from the profit maximization problem while the joint markups for the bundling firms are recovered numerically from the expressions from the first-order conditions. These markups are all recovered at the firm-market-product-monthly level and aggregated into various levels using values as weights.

Table 3 presents recovered markups across firm types at firm-market-yearly level. Column

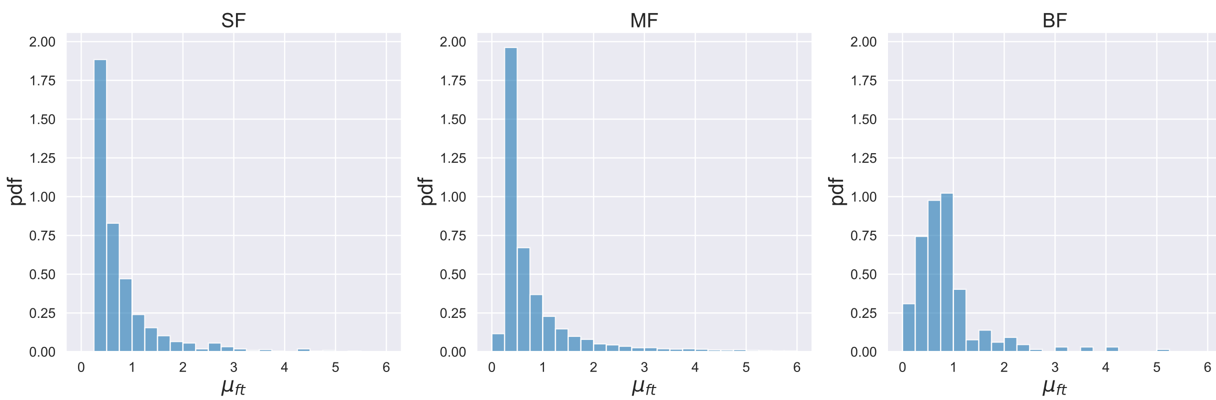
Table 3: Markup Results

	Total		Single-		Multi- No Bundling		Multi- Bundling	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CPU	0.74	0.64	0.73	0.57	0.74	0.65	0.82	0.49
GPU	0.80	0.67	0.78	0.65	0.80	0.69	0.87	0.62

Note: This table reports the average and median value of recovered markups by firm types. Markups are at the firm-market-product-year level. Here, the top and bottom 3% values are trimmed.

(1) and (2) reports mean and standard deviations of markups for all firm types. On average, CPU and GPU products are priced about 70 to 80 % higher than their original cost. Within the sample, mean values of recovered markups increase as we move from a single-product firm to a multi-product firm without bundling to a multi-product firm with bundling. This could indicate firms are utilizing their status, i.e., being a multi-product firm and/or engaging in a bundling practice to increase their market power.

Figure 3: Markups of CPU and GPU by firm-types



Note:This figure plots the histogram of firm-market-year level markups for CPU and GPU. 'SF' stands for single-product firms, 'MF' for multi-product firms without bundling practices and 'BF' for multi-product firms with bundling practices. Here, the top and bottom 3% values are trimmed.

Figure 3 visualizes the recovered markups for single-, multi-product firms with and without bundling. There are a few things to be noted. First, regardless of the firm type, most markups are clustered below one.[22] Also, there seems to be a systematic difference across firm types. That is, while the markup values of non-bundling firms present high-peaked distribution, the markup values of bundling firms are more dispersed,

[22] Note that a markup value of 0.7 means the firms in enjoying 70% of marginal cost as profit for each unit.

indicating product bundling could affect their markup distribution.

4.2 Markups, Firm Heterogeneity and Trade Liberalization

4.2.1 Markups and Firm Heterogeneity

In this subsection, I study how product bundling affect firm's markups and their changes in response to the trade liberalization.The following regression examines how markups differ across firm heterogeneity, such as single-, multi-, and multi-product firms with bundling.

$$\log \mu = \delta_{FE} + \delta_M^1 D_M + \delta_B^1 D_B + \varepsilon^1 \tag{15}$$

where the markup (μ) and the error term (ε^1) is from firm-market-product-monthly level to firm-yearly level. D_M is a multi-product firm dummy, and D_B is a dummy variable for multi-product firms that engage in bundling practice. Both D_M and D_B varies by market-year, indicating that firm category is based on sales side rather than production side. To capture any market specific or yearly trend as well as firm type[23] trend, market, year, and firm type fixed effects are included. In this regression, δ_M^1 measures the percentage markup premium that a multi-product firm that does not engage in bundling has relative to single-product firms(i.e., "multi-product premium"). The percentage premium that the multi-product firm with bundling has over multi-product firms that do not engage in bundling(i.e., "bundling premium") will be captured by δ_B^1. Thus, $\delta_M^1 + \delta_B^1$ measures the percentage premium of multi-product firms with bundling over single-product firms.

Table 4 shows the result of equation (15). Column (1) reports regression coefficients from baseline transaction unit, that is firm-market-product-monthly level, column (2) at firm-market-yearly level, and finally, column (3) at firm-yearly level. The results from Table 4 align with economic intuition. The multi-product firm dummy coefficients are significant at a 1% level and positive for baseline transaction and firm-market-yearly level. While this coefficient is negative for firm-yearly level, it is insignificant. Specifically, for baseline transactions, multi-product firms without bundling practice have 9.65% higher markups than single-product firms on average and 8.97% higher markups at

[23]Such as SOEs and private companies.

Table 4: Markups and Firm Heterogeneity

	(1)	(2)	(3)
	$log\mu_{fdjm}$	$log\mu_{fdt}$	$log\mu_{ft}$
D_M	0.0921***	0.0859***	-0.0698
	(0.0092)	(0.0228)	(0.0999)
D_B	0.0424***	0.2671***	0.1938***
	(0.0183)	(0.0522)	(0.0549)
Market FE	Yes	Yes	No
Year FE	Yes	Yes	Yes
Firm Type FE	Yes	Yes	Yes
F-statistic	27.74	11.98	4.422
Observation	36,076	7,459	3,386

Note: This table reports the coefficients from the regression (15). The dependent variable is (log) markup. Each column is an OLS regression result of log markup on firm heterogeneity for observations for CPU and GPU with various levels. Column (1) is shows the results for baseline transactions and column (2) shows the results at the firm-destination market-yearly level. Finally column (3) shows the results at the firm-yearly level. The standard errors are in parenthesis and are bootstrapped. Significance : * 10 percent, ** 5 percent, *** 1 percent.

firm-market-yearly level. Also, the bundling dummy coefficients are all significant and positive across all results. For baseline transactions, multi-product firms that engage in bundling practices have, on average, 4.34% higher markups than multi-product firms without bundling practices. Similarly, at firm-market-yearly level, multi-product firms with bundling practice enjoyed 30.6% higher markups compared to multi-product firms without bundling practice and 39.6% higher markups compared to single-product firms on average. At firm-yearly level, multi-product firms with bundling goods have 21.4% higher markups than those without bundling. In short, these results show that firms could utilize product bundling to exercise market power and retain higher markups compared to others and highlight the importance of incorporating firm heterogeneity into consideration when examining markups.

4.2.2 Markups and Trade Liberalization

Firm heterogeneity may also impact how firms and their markups react to trade policy changes. In this section, I inspect how firms' reactions to changes in market competitiveness induced by trade policy change varies across firm type. We would expect to see overall markups for Chinese exporters to increase after trade liberalization due to a decrease in tariff, but these upward trends may differ across firm heterogeneity (single- vs multi- vs multi- with bundling). To analyze this, I look at (1) simple before and after analysis of China's accession to WTO, and (2) improved market access analysis.

A simple before and after analysis of the effects of trade liberalization on markups is carried out using the following regression equation. In this regression analysis, China's accession to the WTO on 11th December 2001 is treated as trade liberalization.

$$\log \mu = \delta^2_{FE} + \delta^2_{TL} D_{TL} + \delta^2_{TL*M} D_{TL} D_M + \delta^2_{TL*B} D_{TL} D_B + \varepsilon^2 \tag{16}$$

where D_{TL} is a dummy variable for the trade liberalization and δ_{FE} is a market and firm type fixed effects to capture any trend. As before, δ^2_{TL} will capture the effect of trade liberalization on markups for single firms, $\delta^2_{TL}+\delta^2_{TL*M}$ on markups of multi-product firms without bundling, and $\delta^2_{TL} + \delta^2_{TL*M} + \delta^2_{TL*B}$ on markups on multi-product firms with bundling.

Table 5: Markups and Trade Liberaliztion

	(1)	(2)	(3)
	$log\mu_{fdjm}$	$log\mu_{fdt}$	$log\mu_{ft}$
D_{TL}	0.1278***	0.0014	0.0184
	(0.0121)	(0.0308)	(0.1194)
$D_{TL}D_M$	0.0237**	0.0558**	−0.0374
	(0.0106)	(0.0272)	(0.1155)
$D_{TL}D_B$	0.0889***	0.3561***	0.2525***
	(0.0181)	(0.0518)	(0.0553)
Market FE	Yes	Yes	No
Firm Type FE	Yes	Yes	Yes
F-statistic	86.71	15.22	5.580
Observation	36,076	7,851	3,386

Note: This table reports the coefficients from the regression (16). The dependent variable is (log) markup. Each column is an OLS regression result of log markup on firm heterogeneity for observations for CPU and GPU with various levels. Column (1) is shows the results for baseline transactions and column (2) shows the results at the firm-destination market-yearly level. Finally column (3) shows the results at the firm-yearly level. The standard errors are in parenthesis and are bootstrapped. Significance : * 10 percent, ** 5 percent, *** 1 percent.

Table 5 shows the results from the regression equation (16). As expected, the trade liberalization dummy coefficient shows positive values across markup units but is only significant at the most disaggregated level. The coefficient of the interaction term of the trade liberalization and being a multi-product firm is positive and significant until firm-market-yearly level. Lastly, the bundling dummy coefficients are all positive and significant across markup units. In terms of magnitude, in the case of transaction unit markups, trade liberalization brings 13.6% higher markups on average for single-product firms, 16.0% higher markups for multi-product firms without bundling practice, and 25.3% higher markups for multi-product firms with bundling. At firm-market-yearly level, on average, single-product firms gain 0.144% higher markups due to trade liberalization; multi-product firms without bundling gain 5.88%, while multi-product firms with bundling practice enjoy 48.7% higher markups. At firm-yearly level, bundling firms enjoy 26.9% higher markups after trade liberalization. These regression results also show how important it is to account for firm heterogeneity in analyzing trade policy changes-induced market power changes.

[[MARKET ACCESS ANALYSIS GOES HERE]]

4.2.3 Trade Liberalization and Pass-through

Furthermore, following De Loecker et al. (2016), I characterize the degree to which Chinese input tariff reductions are passed through to consumers in settings with joint markups and bundling. Denote log of estimated marginal costs as $\log \hat{c}_{fjdt} = \log c_{fjdt} + \zeta_{fjdt}$, where ζ_{fjdt} captures the deviations and measurement error. Then, the following regression equation can conduct a cost to price pass-through analysis.

$$\begin{aligned} \log P_{fjdt} &= \delta_{FE} + \delta_1^4 \log \hat{c}_{fjdt} + (-\delta_1^4 \zeta_{fjdt} + \varepsilon_{fjdt}^4) \\ &= \delta_{FE} + \delta_1^4 \log \hat{c}_{fjdt} + u_{fjdt} \end{aligned} \tag{17}$$

Due to endogeneity issues caused by unobserved $-\delta_1^4 \zeta_{fjdt}$ terms, input tariffs or lagged marginal costs, which are correlated with marginal cost but uncorrelated with measurement error ζ_{fjdt} are used as instruments. To examine how firm heterogeneity plays role in cost pass-through, interaction terms with D_M and D_B are included in the regression analysis. The result with $\delta_1^4 < 1$ implies that the price decreased less than the cost reduction, indicating imperfect pass-through.

Table 6: Pass-through of Costs to Prices

	$\log P_{fjdm}$			
	(1)	(2)	(3)	(4)
$\log c_{fjdt}$	0.9608***	0.9649***	0.9499***	0.9537***
	(0.0011)	(1.1e-06)	(0.0015)	(1.2e-05)
$\log c_{fjdt} D_M$	-	-	0.0142***	0.0146***
	-	-	(0.0013)	(2.2e-06)
$\log c_{fjdt} D_B$	-	-	−0.0270***	−0.0265***
	-	-	(0.0035)	(2.0e-06)
IV	No	Yes	No	Yes
Market FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm Type FE	Yes	Yes	Yes	Yes
F-statistic	8.3e+05	8.2e+05	2.8e+05	8.4e+05
Observation	34,394	34,394	34,394	34,394

Note: The dependent variable is (log) prices. Column 1 is an OLS regression on log marginal costs that will suffer from endogeneity problems. Column 2 instruments the log marginal costs with lag marginal costs. Columns (3) and (4) incorporate firm heterogeneity into columns (1) and (2) by adding multi-product firm and bundling dummy interactions. The standard errors are in parenthesis and are bootstrapped. Significance : * 10 percent, ** 5 percent, *** 1 percent.

Table 6 displays the results of equation (17). The first two columns are OLS and IV regression results without firm heterogeneity. The pass-through rate is 96.08% and 96.49%, indicating it is close to a complete pass-through. While it is close to one, the remaining portion of the cost decrease is not lead to a decrease in prices but an increase in markups. Columns (3) and (4) delve deeper and examine how firm heterogeneity impacts a firm's pass-through. Column (3) displays regression results without an instrument and indicates pass-through of firms differs greatly depending on the firm type. Specifically, the pass-through rate of single-product firms is 94.99%, slightly lower than 96.08%. Firms that produce multiple goods have a higher pass-through rate of 96.41%, and firms that engage in bundling practices can decrease this pass-through rate to 93.71%. These values align with Table 5, where multi-product firms with bundling enjoyed higher markups after trade liberalization. In column (4) with instruments, the coefficient on marginal cost interacted with the bundling dummy is significant and remains negative. These results imply that bundling firms can retain their market power and increase markups more than others.

5 Conclusion

By nature, multi-product firms may have a different pricing strategy than single-product firms. Product bundling is a prominent example where multi-product firms price goods jointly. This paper advances the literature by proposing a method to estimate joint markups for multi-product firms with mixed bundling practices. Examination shows that multi-product firms with product bundling enjoy higher markups than their counterpart, which enjoys higher markups than single-product firms. Also, multi-product firms with bundling can retain their market power and markups relatively well in trade liberalization due to having the lowest pass-through rate of costs to prices. These results show how multi-product firms could utilize product bundling for retaining their prices and markups in response to trade liberalization.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. (2015) "Identification Properties of Recent Production Function Estimators." *Econometrica*, 83, 2411–2451.

Adams, William James, and Janet L. Yellen. (1976) "Commodity Bundling and the Burden of Monopoly." *The Quarterly Journal of Economics*, 90, 475–498.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. (2000) "Plants and Productivity in International Trade." *The American Economic Review*, 93, 1268–1290.

Berry, Steven, James Levinsohn, and Ariel Pakes. (1995) "Automobile Prices in Market Equilibrium." *Econometrica*, 63, 841–890.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. (2021) "Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation from Production Data." *Journal of Monetary Economics*, 121, 1–14.

Chen, Yongmin, and Michael H. Riordan. (2013) "Profitability Of Product Bundling." *International Economic Review*, 54, 35–57.

Crawford, Gregory S, and Ali Yurukoglu. (2012) "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review*, 102, 643–685.

De Loecker, Jan, Pinelopi Koujianou Goldberg, Amit Khandelwal, and Nina Pavcnik. (2016) "Prices, Markups and Trade Reform." *Econometrica*, 84, 445–510.

De Loecker, Jan, and Frederic Warzynski. (2012) "Markups and Firm-Level Export Status." *The American Economic Review*, 102, 2437–2471.

Doraszelski, Ulrich, and Jordi Jaumandreu. (2019) "Using Cost Minimization to Estimate Markups." *mimeo.*

Flynn, Zach, Amit Gandhi, and James Traina. (2019) "Measuring Markups with Production Data." *mimeo.*

Fugazza, Marco, and Alessandro Nicita. (2011) "On The Importance Of Market Access For Trade." *Policy Issues In International Trade and Commodities Study Series*, 50.

Iyoha, Ebehi, Ji Hye Heo, and Joel Rodrigue. (2022) "Product Bundling in International Markets." *mimeo.*

Jaumandreu, Jordi. (2018) "Dangerous Shortcuts: Ignoring Marginal Cost Determinants in Markup Estimation." *mimeo.*

Klette, Tor Jakob, and Zvi Griliches. (1996) "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous." *Journal of Applied Econometrics*, 11, 343–361.

Letham, Benjamin, Wei Sun, and Anshul Sheopuri. (2014) *Latent Variable Copula Inference for Bundle Pricing from Retail Transaction Data.* PMLR.

Levinsohn, James, and Amil Petrin. (2003) "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70, 317–341.

Lu, Yi, Linhui Yu, and Linhui Yu. (2015) "Trade Liberalization and Markup Dispersion: Evidence from China's WTO Accession." *American Economic Journal: Applied Economics*, 7, 221–253.

McAfee, R. Preston, John McMillan, and Michael D. Whinston. (1989) "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values." *The Quarterly Journal of Economics*, 104, 371–383.

Olley, G. Steven, and Ariel Pakes. (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64, 1263–1297.

Stigler, George J.. (1963) "United States v. Loew's Inc.: A Note on Block-Booking." *The Supreme Court Review*, 1963, 152–157.

Zhou, Jidong. (2017) "Competitive Bundling." *Econometrica*, 85, 145–172.

Zhou, Jidong. (2021) "Mixed Bundling in Oligopoly Markets." *Journal of Economic Theory*, 194, 105257.

Appendix

A. First Order Conditions from the Profit Maximization

A.1 Derivation of the expressions

Here I demonstrate steps for calculating the first order conditions in Section 2. Recall the profit function was

$$\begin{aligned}
\Pi_{ft}(\mathbf{P}_{ft}) &= (P_{f1t} - c_{f1t})Q^D_{f1t}(\mathbf{P}_{ft}) + (P_{f2t} - c_{f2t})Q^D_{f2t}(\mathbf{P}_{ft}) + (P_{fbt} - c_{f1t} - c_{f2t})Q^D_{fbt}(\mathbf{P}_{ft}) \\
&= (P_{f1t} - c_{f1t}) \int_{P_{f1t}}^{\infty} \int_0^{P_{fbt}-P_{f1t}} \psi_f(x,y)dydx + (P_{f2t} - c_{f2t}) \int_0^{P_{fbt}-P_{f2t}} \int_{P_{f2t}}^{\infty} \psi_f(x,y)dydx \\
&+ (P_{fbt} - c_{f1t} - c_{f2t}) \left[\int_{P_{f1t}}^{\infty} \int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(x,y)dydx + \int_{P_{fbt}-P_{f2t}}^{P_{f1t}} \int_{P_{fbt}-x}^{\infty} \psi_f(x,y)dydx \right]
\end{aligned}$$

Taking derivative with respect to p_{f1t} results in the following equation.

$$\begin{aligned}
&\int_{P_{f1t}}^{\infty} \int_0^{P_{fbt}-P_{f1t}} \psi_f(x,y)dydx + (P_{f1t} - c_{f1t}) \frac{\partial}{\partial P_{f1t}} \Big[\int_{P_{f1t}}^{\infty} \int_0^{P_{fbt}-P_{f1t}} \psi_f(x,y)dydx \Big] \\
&+ (P_{fbt} - c_{f1t} - c_{f2t}) \frac{\partial}{\partial P_{f1t}} \Big[\int_{P_{f1t}}^{\infty} \int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(x,y)dydx + \int_{P_{fbt}-P_{f2t}}^{P_{f1t}} \int_{P_{fbt}-x}^{\infty} \psi_f(x,y)dydx \Big] = 0
\end{aligned}$$

Note that the first term corresponds to $Q^D_{f1t}(\mathbf{P}_{ft})$. For the second term, denote $G_{f1t}(P_{f1t}, P_{fbt}, x) = \int_0^{P_{fbt}-P_{f1t}} \psi(x,y)dx$ and $H_{f1t}(P_{f1t}, P_{fbt}) = \int_{P_{f1t}}^{\infty} G_{f1t}(P_{f1t}, P_{fbt}, y)dy$. Then applying the Leibniz rule gives the following for the second term.

$$\begin{aligned}
\frac{\partial}{\partial P_{f1t}} H_{f1t}(P_{f1t}, P_{fbt}) &= -G_{f1t}(P_{f1t}, P_{fbt}, P_{f1t}) + \int_{P_{f1t}}^{\infty} \frac{\partial}{\partial P_{f1t}} G_{f1t}(P_{f1t}, P_{fbt}, x)dx \\
&= -\int_0^{P_{fbt}-P_{f1t}} \psi_f(P_{f1t}, y)dy - \int_{P_{f1t}}^{\infty} \psi_f(x, P_{fbt} - P_{f1t})dx
\end{aligned}$$

Similarly, for the third in the first order condition, let $G_{f2t}(P_{fbt}, x) = \int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(x,y)dy$ with $H_{f2t}(P_{f1t}, P_{fbt}) = \int_{P_{f1t}}^{\infty} G_{f2t}(P_{f1t}, P_{fbt}, y)dy$, and for the fourth term let $G_{f3t}(P_{fbt}, y) = \int_{P_{fbt}-y}^{\infty} \psi_f(x,y)dy$ with $H_{f3t}(P_{f1t}, P_{f2t}, P_{fbt}) = \int_{P_{fbt}-P_{f2t}}^{P_{f1t}} G_{f3t}(P_{fbt}, y)dy$.

Then taking derivative following the Leibniz rule gives the following expressions.

$$\begin{aligned}
\frac{\partial}{\partial P_{f1t}} H_{f2t}(P_{f1t}, P_{fbt}) &= -G_{f2t}(P_{f1t}, P_{fbt}, P_{f1t}) + \int_{P_{f1t}}^{\infty} \frac{\partial}{\partial P_{f1t}} G_{f2t}(P_{f1t}, P_{fbt}, y) dy \\
&= -\int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(P_{f1t}, y) dy + \int_{P_{f1t}}^{\infty} \psi_f(x, P_{fbt} - P_{f1t}) dx \\
\frac{\partial}{\partial P_{f1t}} H_{f3t}(P_{f1t}, P_{f2t}, P_{fbt}) &= G_{f3t}(P_{fbt}, P_{f1t}) \\
&= \int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(P_{f1t}, y) dy
\end{aligned}$$

Plugging these terms into the original f.o.c and using the definition of $Q_{f1t}(\mathbf{P}_{ft})$ gives,

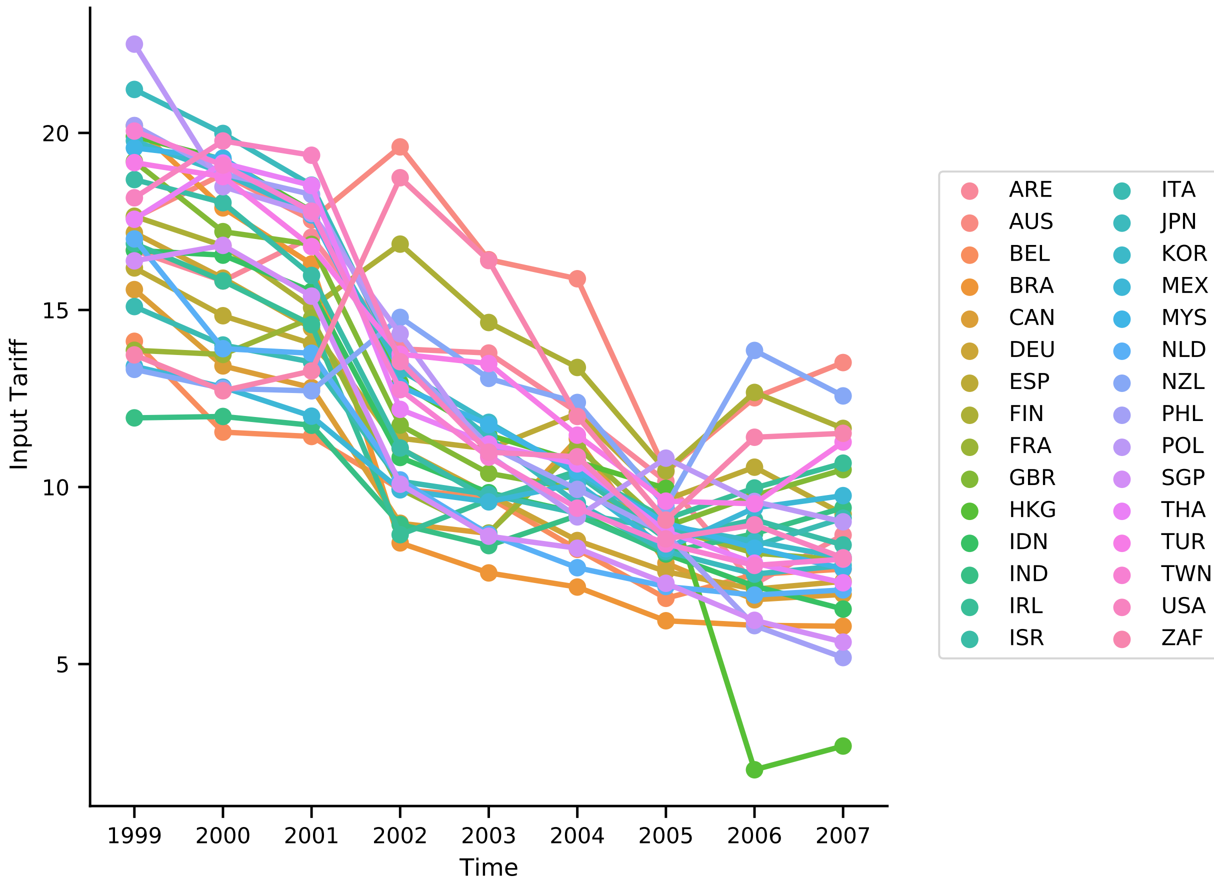
$$\begin{aligned}
&Q^D_{f1t}(\mathbf{P}_{ft}) - (P_{f1t} - c_{f1t})\Big[\int_0^{P_{fbt}-P_{f1t}} \psi(P_{f1t}, y) dy + \int_{P_{f1t}}^{\infty} \psi_f(x, P_{fbt} - P_{f1t}) dx\Big] \\
&+ (P_{fbt} - c_{f1t} - c_{f2t})\Big[-\int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(P_{f1t}, y) dy + \int_{P_{f1t}}^{\infty} \psi_f(x, P_{fbt} - P_{f1t}) dx + \int_{P_{fbt}-P_{f1t}}^{\infty} \psi_f(P_{f1t}, y) dy\Big] \\
&= Q^D_{f1t}(\mathbf{P}_{ft}) - (1 - \mu^{-1}_{f1t}) P_{f1t} \int_0^{P_{fbt}-P_{f1t}} \psi_f(P_{f1t}, y) dy + [(1 - \mu^{-1}_{f2t}) P_{f2t} - d_{fbt}] \int_{P_{f1t}}^{\infty} \psi_f(x, P_{fbt} - P_{f1t}) dx
\end{aligned}$$

where the equality comes from $(P_{f1t} - c_{f1t}) = (P_{f1t} - c_{f1t})\frac{P_{f1t}}{P_{f1t}} = (1 - \mu^{-1}_{f1t})P_{f1t}$ and $(P_{f1t} + P_{f2t} - c_{f1t} - c_{f2t} - d_{fbt}) = (P_{f1t} - c_{f1t})\frac{P_{f1t}}{P_{f1t}} + (P_{f2t} - c_{f2t})\frac{P_{f2t}}{P_{f2t}} - d_{fbt} = (1 - \mu^{-1}_{f1t})P_{f1t} + (1 - \mu^{-1}_{f2t})P_{f2t} - d_{fbt}$. Derivatives with respect to P_{f2t} and P_{fbt} are similar thus omitted.

A.2 Expressions with more detailed data

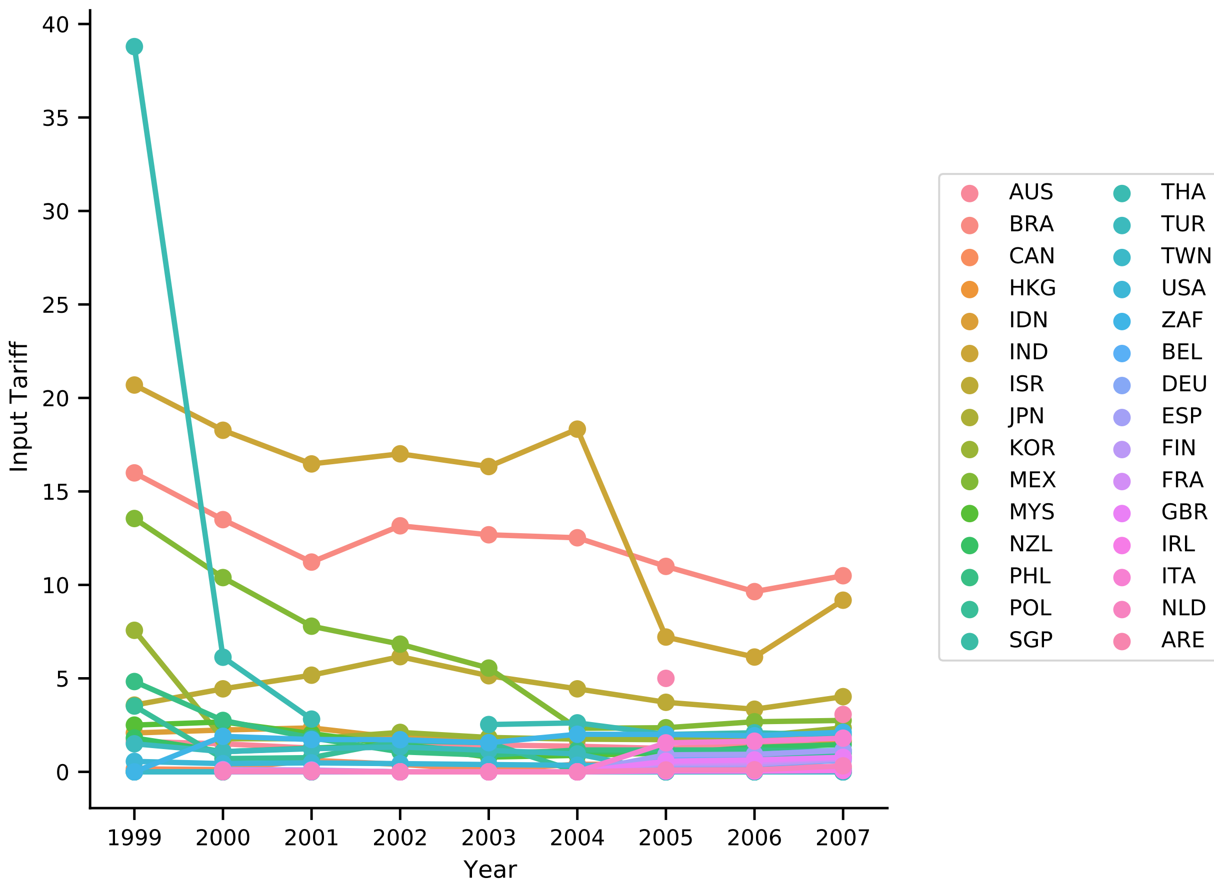
C. Others

Figure 4: Input Tariffs for CPU and GPU from 1998 to 2007



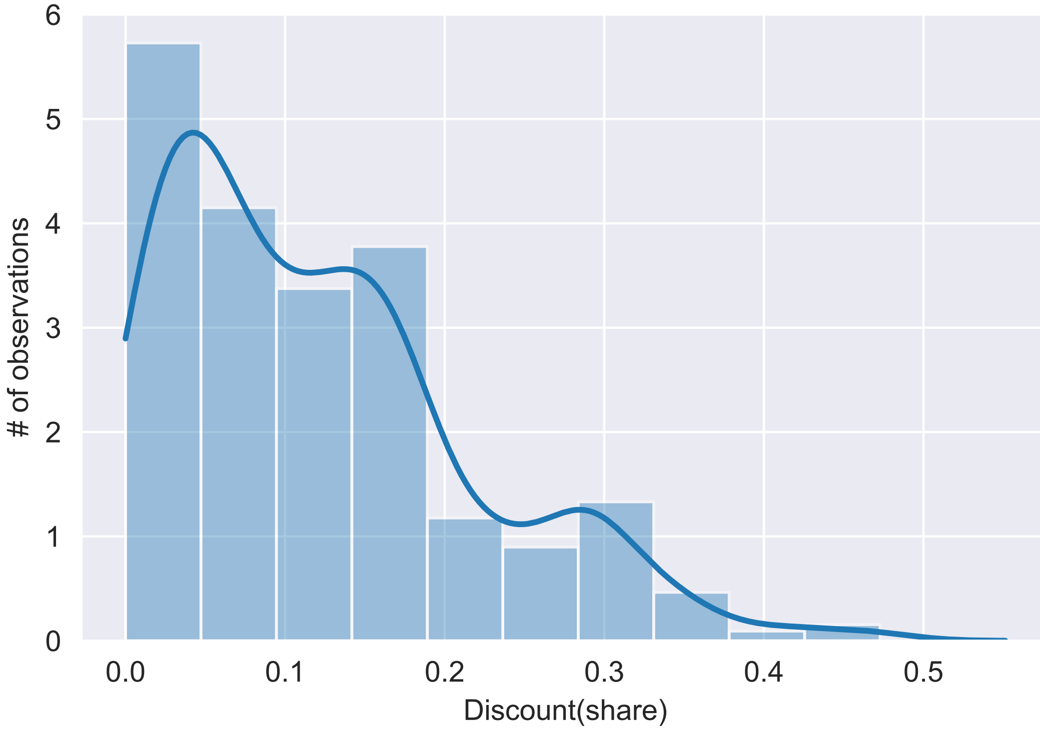
Note: This figure plots market-level input tariffs for China from 1998 to 2007 with respect to HS2 level 84 which contains CPU and GPU.

Figure 5: Output Tariffs for CPU and GPU from 1998 to 2007



Note: This figure plots market-level output tariffs for China from 1998 to 2007 with respect to HS2 level 84 which contains CPU and GPU.

Figure 6: Bundling discount share for bundled transactions



Note: This figure plots the bundling discount share which is bundling discount over the sum of each component product prices in the x-axis. The y-axis shows normalized density for the number of observations.