Made and Created in China: The Role of Processing Trade*

Zhiyuan Chen† Aksel Erbahar‡ Yuan Zi§

February 2022

Abstract

This paper proposes that processing trade not only leads goods to be "Made in China," but also "Created in China." Using unique transaction-level trade data, we find that a significant share of exporters engages in both ordinary and processing export activities, and they exhibit superior performance. There is a close link between firms' export mode choice and brand ownership, and there is a price premium associated with own-branded products. Lastly, firms intensify their branding activities when facing favorable processing policies upstream. To rationalize these findings, we present a simple framework where multi-attributes firms endogenously specialize within the production network.

JEL codes: F12, F13, F14

Keywords: heterogeneous firms, production networks, processing trade

*We are grateful to Suzanne Bijkerk, Maarten Bosker, Julia Cajal Grossi, Mi Dai, Jonathan Eaton, Julian Emami Namini, Inga Heiland, Laura Hering, Sacha Kapoor, Vladimir Karamychev, Kala Krishna, Franceso Lippi, Andreas Moxnes, Monika Mrázová, Ezra Oberfield, Vincent Rebeyrol, Frédéric Robert-Nicoud, Yu Shi, Felix Tintelnot, Lorenzo Trimarchi, Bauke Visser, Yikai Wang, Zi Wang, Weisi Xie, Stephen Yeaple, and Zhihong Yu for helpful comments. We thank Ran Jing from the University of International Business and Economics in Beijing for sharing her Chinese trademark data. We also thank seminar participants at the China International Conference in Economics, Erasmus University Rotterdam, Graduate Institute of Geneva, Norwegian Business School, Pennsylvania State University, University of Namur, University of Nottingham, and University of Oslo for useful feedback. Yanlin Zhou provided excellent research assistance. Yuan Zi acknowledges funding received from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 715147).

†School of Business, Renmin University of China, China, *email*: chenzhiyuan@rmbs.ruc.edu.cn

‡Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands, and Tinbergen Institute, *email*: erbahar@ese.eur.nl

§The Graduate Institute of International and Development Studies, Geneva, Switzerland, and CEPR, *email*: yuan.zi@graduateinstitute.ch

1 Introduction

"[W]hereas during the later part of the twentieth century and early twenty-first century, the world became used to reading the Made in China *label on every conceivable type of product, mankind is increasingly getting used to a ubiquitous* Branded in China *tag. What is clear is that China has fallen in love with brands."*

-John M.T. Balmer and Weifeng Chen, *Advances in Chinese Brand Management,* 2017

China's trade as percentage of its GDP rose from below 10% in late 1970s to over 60% just before the Great Recession (World Bank, 2018). During this period, Chinese firms specialized in relatively low value-added stages of the global value chain and supplied foreign multinationals largely through processing trade, as epitomized by the "Made in China" tag. However, this phenomenon is changing. After decades of efforts to become 'the factory of the world,' China's large manufacturing base is now a breeding ground for firms with innovative ideas. Between 2000 and 2014, Chinese firms' share of technology improvement budget dedicated to in-house R&D rose from 78% to 84% (Wei et al., 2017); Chinese firms' domestic invention patent filings and trademark applications grew, on average, by over 30% each year, with an even faster growth since 2008 (Eberhardt et al., 2016; Deng et al., 2020).

An unexplored angle of this switch from "Made in China" to "Created in China" is the role of processing trade, which lets firms to forego paying tariffs on imports that they process to export. While processing trade accounted for the majority of China's total exports and was the key driver of China's export boom, relatively little attention has been paid to its main participants—exporters that engaged in both ordinary and processing exports. These mixed firms made up about a fifth of processing exporters, and contributed to over 60% of total Chinese processing exports, explaining about half of China's export surge during 2000-2006. Even though they are considered to be "perhaps the most interesting type of firm[s]" (Yu, 2015), they were never carefully investigated in the literature.

In this paper, we start by unpacking the "black box" of mixed firms to examine firm performance and their specialization within a production network. We find that mixed exporters are larger and have higher revenue and physical productivity compared to firms that engage in only ordinary (i.e., pure ordinary exporters) or only processing (i.e., pure processors) activities. Importantly, unlike what is suggested in the literature, these firms are not 'mixed' because they sell different products under different export modes: the majority of their exports consists of the same product being sold to the same destination under both processing and ordinary trade modes.

Even though being highly processing-oriented, mixed exporters' superior labor and revenue productivity does not generalize to pure processing exporters. On the other hand, pure processing exporters have significantly higher physical productivity when compared to pure ordinary exporters. In addition, using a novel transaction-level customs data with detailed product and brand information, we find that firms tend to export their own branded products using ordinary trade mode, and that there is a price premium associated with selling one's own branded product. This finding suggests that a firm's export mode not only reflects its position inside a production network, but is also closely related to its efficiency across different stages of production (i.e., manufacturing versus branding), which ultimately determines its measured performance at various margins.

Having established the set of stylized facts on exporters' performance, export mode, and brand ownership, we next examine the impact of processing trade policy. To this end, we use China's pilot "paperless" processing supervision program implemented in 2000-2006 as a quasi-natural experiment. The paperless program significantly reduced the burden of red-tape on processing business by replacing processing-related paperwork with the customs' automatic, online administration system.[1] This policy shock is highly suitable for our study and gives us a clean identification, as it affects only the costs of processing trade, leaving other costs of a firm unchanged. By exploiting the staggered introduction of the policy to different regions in China, and by comparing firms around the qualification cutoff, we document that the paperless processing program increased firm-level processing exports by 28%. We also find that the policy induced downstream firms to intensify their branding activities: the number of trademarks for above-median productive domestic firms increased by about 1% on average. From a development perspective, this result indicates that processing trade policies help domestic firms to create new products, suggesting that there are some unexplored gains from trade.

In the last part of the paper, we build a parsimonious model to rationalize our findings in a unified yet intuitive framework. Our model features an endogenous production network in which firms are heterogeneous in both manufacturing and branding abilities. In equilibrium, firms with good blueprints but low manufacturing ability outsource production and become downstream firms, and those with intermediate manufacturing ability and blueprint quality become ordinary exporters. Firms with higher manufacturing ability but low blueprint quality become pure processing exporters, and firms with exceptional blueprint quality and manufacturing ability become mixed exporters, i.e., firms that both export their own brands and serve as manufacturing suppliers for foreign firms. As such, our model rationalizes the observed ranks at various margins between mixed, pure ordinary, and pure processing exporters. The model also yields the prediction that is consistent with our empirical finding: facilitating processing trade raises the ex-ante expected profits from manufacturing, leading to a greater mass of potential suppliers, which benefits sourcing firms with good ideas. Thus our results highlight that processing trade not only led goods to be "Made in China," but also "Created in China" by providing a breeding ground of suppliers for firms with good ideas.

Our work is related to several strands of the trade literature. First, our stylized facts on mixed exporters are related to a large body of work on the characteristics of processing exporters in China (Fernandes and Tang, 2015; Yu, 2015; Dai et al., 2016; Kee and Tang, 2016; Li et al., 2018).[2]

[1]The details of this policy are given in Section 4.1.

[2]Fernandes and Tang (2015) find that processing firms are less diversified in products and destinations when compared to ordinary exporters, and Yu (2015) shows that their productivity does not change considerably with

Different from these studies which focus on pure processing firms, we document the dominant role of mixed exporters that engage in both ordinary and processing exports. We also provide novel empirical facts that shed light on firms in supply-chain trade by relating for the first time the characteristics of different types of exporters with their brand ownership and choice of trade modes, using a unique transaction-level trade data on firms' branding information.

This paper does not intend to disentangle all the mechanisms behind processing trade. Rather, we highlight the key feature of processing firms, i.e, they are typically contract-taking suppliers to foreign downstream firms. Relatedly, we view policies such as duty exemptions or paperless supervision as factors that increase a firm' propensity to engage in processing activities. By doing so, we complement the works of Feenstra and Hanson (2005), Fernandes and Tang (2012), Dai et al. (2016), Manova and Yu (2016), Brandt and Morrow (2017), Defever and Riaño (2017), and Deng (2021) who emphasize the role of different policy factors that shape firms' export mode choice.[3] We rely on rich transaction- and firm-level Chinese data and a unique quasi-natural experiment, the pilot "paperless" processing supervision program, to shed light on the implications of processing policy, and thus also complement the work that examines the welfare implications of processing trade through the lens of various quantitative trade models, e.g., Defever and Riaño (2017), Brandt et al. (2019), Deng (2021), and Deng and Wang (2021).[4]

Our paper also contributes to the literature on firms' sourcing decisions in international and regional trade, e.g., Antràs et al. (2017), Lim (2018), Bernard et al. (2019b), Kikkawa et al. (2019), and Dhyne et al. (2021).[5] These papers emphasize that sourcing decisions are important in explaining firms' performance, shock transmissions, aggregate gains from trade, and business cycle fluctuations. Our paper shows that it is also useful to take the network feature into account to explain exporters' performance under processing trade.

Finally, our paper connects to the literature that study firms with multiple heterogeneities, including Antràs and Helpman (2004), Hallak and Sivadasan (2013), Harrigan and Reshef (2015),

trade liberalization. Dai et al. (2016) find that compared to non-exporters and ordinary exporters, processing firms have lower revenue productivity, skill intensity, and profitability, and they pay lower wages and spend little on R&D. Kee and Tang (2016) show that China's processing exporters began to use domestic inputs instead of imported materials during 2000-2007. Li et al. (2018) calculate physical total factor productivity (TFP) based on quantity data and find that processing exporters are significantly more productive than non-exporters.

[3]Dai et al. (2016), Brandt and Morrow (2017), Defever and Riaño (2017), and Deng (2021) emphasize the role of special duty drawbacks; Feenstra and Hanson (2005) and Fernandes and Tang (2012) emphasize foreign firms' outsourcing decisions; Manova and Yu (2016) highlight the importance of credit constraints.

[4]Defever and Riaño (2017) analyze the welfare implications of subsidies with export share requirements in a quantitative export model. Brandt et al. (2019) quantify the welfare effects of duty exemptions under China's processing trade based on a multi-industry Ricardian model. Deng (2021) quantifies the welfare implications of processing policy with the presence of learning-by-processing. Deng and Wang (2021) introduce increasing returns to scale in input production in a similar framework and quantify the processing-trade-induced Dutch disease.

[5]Building on Tintelnot (2017), Antràs et al. (2017) study firms' optimal sourcing decisions across countries, and predict that the intensive and extensive margins of sourcing are positively related to firm productivity. Redefining countries as locations within a country, Bernard et al. (2019b), Kikkawa et al. (2019), and Dhyne et al. (2021) adapt the framework of Antràs et al. (2017) to the context of domestic production networks and study how geography, markups, and endogenous firm-to-firm connections affect shock transmissions and firm performance, respectively. Lim (2018) quantifies the importance of endogenous network adjustment for business cycles. Chaney (2016), Bernard and Moxnes (2018), and Johnson (2018) provide excellent reviews of the network models in international trade.

Manova and Yu (2017), Ariu et al. (2019), and Huang et al. (2021).[6] None of these papers, however, emphasize the role of heterogeneities that enable firms to self-select into different stages of the production network. Combining rich Chinese firm-level trade and production data with a novel transaction-level data with branding information, we show that the intuitive set-up of our model rationalizes a rich set of stylized facts on Chinese firms, and provides new insights on processing promoting policies.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents the set of stylized facts regarding exporters' performance, export mode, and brand ownership. Section 4 examines the spillovers of processing trade to firms' branding activities by exploiting China's "paperless" processing trade program. Section 5 develops a model that rationalizes the empirical findings. Finally, Section 6 concludes.

2 Data and Processing Trade Regimes in China

2.1 Data

We use four main datasets in this paper. The first is China's 2000-2006 customs data that shows firms' monthly transactions of exports and imports at the product-country level, where products are defined at the 8-digit Harmonized Schedule (HS8) level. Since our analysis is focused on manufacturing firms, we remove intermediaries and wholesalers from the dataset.[7] The customs data allows us to observe each firm's ordinary and processing exports at the product-country level. Thus, we are able to divide firms into three mutually exclusive groups: pure processing exporters, pure ordinary exporters, and mixed exporters who are engaged in both processing and ordinary exports.

Our second dataset is a rich sample of transaction-level customs data for 2018. Unlike the commonly used 2000-2006 customs data, this sample is directly obtained from the Chinese customs without any aggregation, therefore we are able to observe all the information in firms' customs clearance records. In particular, these records contain highly detailed product and brand information for each export transaction, as the Chinese government began to require firms to report the brand information in customs declaration forms in 2018.[8] In this database, we observe firm ID, firm

[6] Antràs and Helpman (2004) study how firm-level productivity and sector-level headquarter-intensity affect firms' choices of ownership structure and supplier locations. Hallak and Sivadasan (2013) explore how differences in firms' process versus product productivity can explain the empirical observation that exporters produce higher-quality products. Harrigan and Reshef (2015) let firms differ in productivity and skill-intensity to explain the positive correlation with globalization and wage inequality. Manova and Yu (2017) focus on multi-product firms with different productivity and scope for quality, and studies firms allocate activity across products in line with a product hierarchy based on quality. Bernard et al. (2018) study how productivity and relationship capability can explain the matching between buyers and sellers in Belgium.Bernard et al. (2019a) document carry-along trade and emphasize demand-scope complementarities. Ariu et al. (2019) study the complementarity between trade in goods and services, and finally, Huang et al. (2021) study how upstream market structure affects downstream sourcing behavior.

[7] To remove intermediaries, we follow the approach taken by Ahn et al. (2011) and exclude firms whose names include words such as "import," "export," "trading," "business," "supply chain," "warehousing," and/or "investment."

[8] This policy change was issued in the No. 69 General Administration of Customs Announcement on Amending the "Regulations on the Customs Declaration of Imports and Exports of the People's Republic of China" in 2017,

name, value and quantity of exports, export destination, product specification (both in 10-digit HS code and description), and export mode. The product specification is a long string variable that provides detailed information on the type of product, and its brand name and brand ownership, which we group into three categories: no brand, domestic brands (domestically created or purchased), and foreign brands (including original equipment manufacturers). The dataset consists of 862,567 daily transactions which make up around $38 billion worth of exports in 34 HS8 products by 29,138 firms, covering product categories from 13 out of 68 HS2 manufacturing sectors.[9] The wide variety of products, which are listed in Table A.1, includes goods that make up a large share of exports such as car tires, refrigerators, and mobile phones.

The third and fourth datasets we use are the annual industry survey (AIS) and the production survey compiled by China's National Bureau of Statistics (NBS) for 2000-2006. The AIS data reports firm-level balance sheet information such as sales, value-added, number of employees, capital stock, R&D expenses, advertisement expenses, material costs, and ownership structure, which allows us to examine firms' performance along various margins.[10] The production survey contains firm-product level information on output quantity, which enables us to compute firm-level quantity-based (i.e., physical) TFP.[11] Both datasets cover all state-owned enterprises (SOEs) and private firms that have annual sales of at least five million RMB. We merge both datasets with the 2000-2006 customs data based on firm names, telephone numbers, and zip codes like in other studies using matched Chinese firm-level data. Our matching procedure results in covering about 58% of aggregate exports, which is similar to the match rate of existing studies.[12]

We utilize two additional datasets for our empirical analysis. The first is the yearly firm-level effective trademarks collected by the State Administration for Industry and Commerce in China, which we merge with the AIS data using unique firm IDs provided by Deng et al. (2020).[13] The second is the dates when each Chinese regional customs authority adopted the pilot paperless processing trade program, which we constructed using China's publicly available official customs notices. We discuss the policy in more detail in Section 4.

2.2 Processing Trade Regimes

In this section, we briefly describe the institutional details of processing trade based on our rounds of interviews with senior officials at Chinese customs and various processing-firm owners.[14] These

and became effective on January 1, 2018.

[9]Of the 34 products, 30 are from March and the rest are from January and April 2018.

[10]We follow the data cleaning procedures proposed by Brandt et al. (2012) and exclude firms with missing or negative (or zero) capital stock, value-added, or employment data, and ones that have less than 8 employees.

[11]See Li et al. (2018) for a more detailed description of the production survey and its link with the AIS survey.

[12]See the Appendix of Chen et al. (2017) for a more detailed explanation of the matching procedure.

[13]We are grateful to Ran Jing for sharing the data. See Deng et al. (2020) for a detailed description of the trademarks dataset.

[14]We are in particularly grateful to Jie Zhang and Li Liang from the research department of the statistical division of Chinese Customs, Jianming Gao and Tommy Yu from Fujian Business Association, and Chunmei Wu for their valuable inputs.

details and their reflection in data are going to help put our empirical findings in context.

Processing trade generally refers to the business activity of importing all, or part of, the raw materials from abroad and re-exporting the finished products after manufacturing within a country. Processing trade widely exists in international commerce, although many countries' customs do not distinguish it from other trade types. China separately classifies processing trade in customs data and treats these transactions with different policies as a consequence of the country's gradual opening-up and dual-track reforms. Viewed as a great way to create jobs, China provides numerous preferential conditions for processing trade such as tax rebates and tariff waivers on intermediate goods and capital equipment that are used exclusively in the production of goods for export.

A central feature of these preferential processing policies is that they aim to help firms integrate into global value chains and manufacture goods for foreign firms. This aim was reflected on the choice of policy instruments and because most Chinese firms were not competitive enough to directly export to the global market when the country began to open-up in the 1980s.[15] Coupled with the relatively cheap labor force of China that attracted firms in developed countries to outsource manufacturing to China, processing trade helped China became an export powerhouse.

The preferential access to processing trade also has a cost. In order to deter other types of firms from evading taxes and tariffs, processing trade is subject to much tighter governmental supervision compared to ordinary trade: every processing contract with detailed information on inputs, outputs, and production processes has to be registered and approved in advance by the Chinese customs before any transaction takes place, which are then subject to stricter customs checks.[16] These policies effectively helped to select businesses that the Chinese government targeted: 84% of processing exports in our transaction sample can be explained by firms making products for foreign brands, as we show in the next subsection. That is, the majority of processing contracts are for Chinese firms "making" goods for foreign contractors, which we take as the "de facto" definition for processing trade throughout the paper.

A key feature of processing trade is that it is defined by *contracts*, not by firms (see order No. 113 of the General Administration of Customs of PRC). This reflects a form of governmental supervision: the Chinese customs approves a firm's filing of a processing transaction if it satisfies certain requirements; then, this transaction will be subject to the relevant policies.[17] A firm can, for example, engage in processing trade and sell domestically at the same time, but only its processing transactions will be subject to processing-specific benefits and regulations. Thus, while we define

[15] For example, when import duty exemption policy was introduced in 1988, China' total trade counted for less than 1% of global trade and over 50% of it were in agriculture and primary goods. From 1978 to 2000, processing trade increased over 64 times while ordinary trade increased by only three times. In 1981, processing trade counted for only 6% of China's total trade, but by 1996 it exceeded 50% of China's total trade.

[16] One way to avoid complicated customs procedure is to operate in export processing zones. However, these zones are highly exclusive and only fit for firms working for extremely stable contractors with fixed inputs and outputs. In 2000-2006, there were 74,184 unique processing exporters, of which only 0.9% of them were located in export processing zones, and 96% of these firms were either foreign-owned or joint ventures.

[17] We thank to Jie Zhang and Li Liang from the research department of the statistical division of Chinese Customs for this clarification.

Table 1: Transition Matrix

Type	PO_{it+1}	PP_{it+1}	Mix_{it+1}
PO_{it}	93.5	0.27	6.23
PP_{it}	1.32	84.1	14.58
Mix_{it}	11.3	6.59	82.11

Notes: PO_{it}, PP_{it}, and Mix_{it} that indicate whether the firm i is a pure ordinary exporter, pure processor, or a mixed exporter in year t respectively. The matrix shows the probability of switching from one type to another in China during 2000-2006.

exporters that export solely through the processing regime as pure processors, we identify mixed exporters as firms that report both ordinary and processing trade to the Chinese customs.

3 Stylized Facts

3.1 Mixed Exporters in China

We begin by unpacking the "black box" of mixed exporters in this subsection. Mixed exporters are defined as firms that engage in both processing and ordinary exports. We find that mixed firms made up about a fifth of processing exporters, and contributed to over 60% of total Chinese processing exports, explaining about a half of China's export surge during 2000-2006. In particular, two findings stand in contrast to the existing literature. First, we do not find evidence that would support the view that there is a linear upgrade from processing to hybrid and to ordinary trade: most industries' top exporters are mixed firms. Second, mixed firms are not 'mixed' because they sell different products under different export modes: the majority of their exports consists of the same product being sold to the same destination under both processing and ordinary trade modes. In what follows we present these findings in steps, which lead us to further examine various firm characteristics across exporter types in the next subsection.

The customs data show that even though the number of mixed exporters was only 21% of the total number of exporters, they made up 54% of exports in 2005. Pure processors and pure ordinary exporters, on the other hand, made up 24% and 19% of exports in 2005 respectively.[18] Mixed firms' exports also made up the bulk (48%) of China's export boom in 2000-2006, with the rest of the growth explained almost equally by exports of pure ordinary firms (21%) and pure processors (24%). As shown in Table 1, firms rarely change their type. Pure ordinary exporters change their type less than 7% of the time, whereas pure processors and mixed firms change their type less than 20% of the time. Firms usually do not switch directly between pure ordinary and pure processing, whereas other types of switches are observed with a similar level of magnitude.

We present firm-level statistics for mixed exporters in Table 2, with the full sample in panel

[18] The rest is made by firms that did not fit into one of the three groups as they engaged in other export modes such as re-exporting, and made up about 3% of exports. Note that we exclude intermediaries and wholesalers, which made up 18% of exports in 2005. These figures are similar to those reported by Dai et al. (2016).

Table 2: Mixed Exporters

	(a) All mixed exp.			(b) Merged mixed exp.		
	Median	Mean	Sd.	Median	Mean	Sd.
(1) Processing share	0.64	0.58	0.36	0.66	0.58	0.36
(2) Processing share, mixed HS8	0.71	0.62	0.34	0.74	0.63	0.34
(3) Processing share, mixed HS8-country	0.68	0.62	0.32	0.70	0.63	0.32
(4) Pure-assembly share	0.00	0.26	0.42	0.00	0.22	0.39
(5) Share of mixed HS8	0.29	0.37	0.31	0.31	0.37	0.30
(6) Share of mixed HS8-country	0.19	0.25	0.24	0.20	0.24	0.23
(7) Value share of mixed HS8	0.87	0.68	0.37	0.89	0.71	0.35
(8) Value share of mixed HS8-country	0.59	0.53	0.37	0.62	0.55	0.36

Notes: This table shows the processing intensity (processing exports/total exports) of mixed exporters in rows 1-3, the share of their processing exports done via the *pure-assembly* (as opposed to *import-and-assembly*) regime in row 4, and their composition of exports (mixed exports/total exports) in rows 5-8, at different levels of aggregation. Panel (a) reports figures for the entire sample of 50,952 mixed exporters, whereas panel (b) reports figures for the subsample of 24,470 mixed exporters that can be matched to the AIS data (merged) for 2000-2006.

(a) and the merged sample in panel (b). The figures in both panels are similar, and thus we refer to statistics in panel (b) from here on. Row 1 shows that the median (mean) share of processing exports in a mixed firm's total exports is 66% (58%). Corresponding shares at the firm-HS8 and firm-HS8-country levels in rows 2 and 3 are similarly high, suggesting that mixed exporters' main activity is processing trade. Nevertheless, mixed exporters contribute substantially to China's ordinary trade as well—in 2005, they made up 63% and 42% of China's processing and ordinary exports, respectively. Moreover, in 51 of the 68 HS2 manufacturing sectors, the top firm in terms of export value was a mixed exporter. Looking at the top three firms in each sector, there was at least one mixed exporter in 66 sectors.

One may conjecture that these firms are 'mixed' because they export multiple products, some under processing trade and others under ordinary trade, potentially due to differences in input tariff schemes. Surprisingly, a careful look at the data reveals that this is not the main explanation. In Table 2 panel (b), we show that the *number* of products exported under both trade regimes, on average, accounts for 37% of mixed firms' total number of exported products (row 5). In terms of values, the median (mean) *value* share of products that are exported through both ordinary and processing modes (mixed HS8) in a mixed firm's exports is as high as 89% (71%) (row 7). In other words, mixed exporters tend to sell their core product(s) under both trade regimes.

One can argue that there might still be different kinds of products within an HS8 code. This is less of a concern since China's product classification at the HS8 level is highly detailed: for example, there are seven different HS8 under the internationally-standardized HS6 code 520811 *Plain weave, unbleached, weighing not more than* $100g/m^2$, that specify the type of cotton used (e.g., medical gauze). This level of detail mitigates the concern that an exporter is mixed due to its multi-product nature. We find similar results even when we look at the more disaggregate product-country level (panel (b) rows 6 and 8).

The fact that firms serve the same products or the same product-destinations under both trade regimes also suggest that their choice of trade mode cannot be primarily driven by trade policies that ex-ante are only different across products, firms, or destinations. For example, if input tariff exemptions for processing trade makes it cheaper for a firm to export a certain product under processing trade regime, it should export this product only via the processing trade regime. These findings do not change even when we consider 'pure assembly' and 'import-and-assembly' separately; the data shows that mixed firms' and pure processors' average share of 'pure-assembly' in their processing exports were very similar in 2000-2006 (22% versus 16%). Also, the government is seldom directly involved with mixed firms: the data shows that only 7% of mixed firms are state-owned enterprises. The top-5 HS2 sectors that mixed exporters engage in are the same as the top-5 sectors for pure ordinary and pure processing firms (HS: 62, 61, 85, 84, 39), suggesting mixed exporters are also not an special phenomena of some specific sectors or industries.

The non-trivial existence of mixed exporters is intriguing. The theoretical literature typically assumes either that processing is a different sector (Brandt et al., 2019; Deng, 2021) or that heterogeneous firms as in Melitz (2003) sort themselves into processing or ordinary trade based on productivity differences combined with a variable-fixed cost trade-off (Brandt and Morrow, 2017; Defever and Riaño, 2017). Mixed exporters, if mentioned, are generated by bringing in some product- or destination-specific shock to fixed costs. In that case, mixed exporters would never sell the same product to a given destination via both export modes.

3.2 Export Mode and Firm Characteristics

Following the well-established literature on exporter premia pioneered by Bernard and Jensen (1995, 1999, 2004), we investigate whether firms that engage in different export modes have significantly different characteristics. Lu (2010) showed that China was exceptional in the sense that it did not have the exporter premia that was found for virtually all other countries. Dai et al. (2016) showed that this lack of exporter premia was due to processing exporters, whose productivity lagged behind that of non-exporters. Several other papers including Fernandes and Tang (2015), Li et al. (2018), and Brandt et al. (2019) focused largely on the differences between ordinary and processing exporters. Instead, our focus is on mixed firms and their comparison to other types of exporters. Specifically, we bring in production and novel transaction-level trade data with brand information to better understand the source of performance differences between firms.

From here on, we use the merged exporters database, and use the two-digit Chinese Industry Classification (CIC) reported in the AIS data for our definition of sectors (except for Facts 2 and 3, for which we use the 2018 customs sample). We run the following regression:

$$Y_{it} = \beta_1 PP_{it} + \beta_2 Mix_{it} + \delta_{ht} + \epsilon_{it}, \tag{1}$$

where Y_{it} is an outcome variable (e.g., $\ln(empl.)_{it}$, where $empl.$ is for employment) for firm i in year t, PP_{it} and Mix_{it} are dummies for pure processing and mixed exporters respectively (pure ordinary exporters is the omitted group), δ_{ht} are sector-year fixed effects, and ϵ_{it} is the error term which we cluster at the sector level to allow for correlated sector-level shocks.[19] Each row of Table 3 shows results from a separate regression, and coefficients can be interpreted as relative to pure ordinary exporters. All regressions except for row 1 include $\ln(empl.)$ as a control variable for firm size. Panel (b) excludes firms with foreign ownership.

Table 3 panel (a) row 1 shows that compared to pure ordinary firms, pure processors and mixed firms have, on average, 30% and 38% more employment respectively. The statistical difference between the two coefficients (Prob.$> F = 0.07$) reveals that mixed exporters are also larger than pure processors. This size premium remains when we exclude foreign firms in panel (b): pure processors and mixed exporters are 21% and 38% larger than pure ordinary exporters respectively.

The existing empirical research, including Mayer and Ottaviano (2008) and Bernard et al. (2012) for European and US firms respectively, finds that larger firms tend to have higher labor productivity and revenue TFP ($TFPR$). Does this result hold for mixed exporters? Table 3 panel (a) row 2 shows that mixed firms have 14% higher labor productivity (i.e., value added per employee) than pure ordinary firms, whereas pure processors have 22% lower labor productivity than pure ordinary firms.[20] Row 3 shows that the ranking we obtained based on labor productivity remains when we consider $TFPR$ calculated using the Olley-Pakes (1996) methodology.[21]

As is well documented in the literature, $TFPR$ reflects not only firms' technical (or manufacturing) efficiency (quantity-based TFP, or $TFPQ$), but also their prices. In particular, focusing on the Chinese leather shoes industry, Li et al. (2018) find that exporters' $TFPQ$ is higher than non-exporters', while their $TFPR$ is lower than non-exporters'. Does this empirical regularity hold for other sectors? What is the rank of mixed firms' $TFPQ$ among exporters? With these two questions in mind, we compute $TFPQ$ focusing on the 36 of the 693 manufacturing 5-digit products for which we can obtain reliable quantity information. The estimation methodology and the list of products can be found in Appendix A and Table A.1 respectively.[22] Consistent with Li et al. (2018), we find that compared to pure ordinary exporters, pure processors have higher $TFPQ$ on average (row 4 of Table 3 panel (a)). In addition, mixed exporters have the highest physical productivity on average (though not statistically significantly different from that of pure

[19] Clustering at the firm level produces significantly lower standard errors.

[20] In a similar vein, Dai et al. (2016) show that pure processing exporters are less productive than non-exporters, who are less productive than non-processing and "hybrid" exporters.

[21] As explained in Appendix A, we use only single-product firms to compute $TFPQ$, and thus the regressions for $TFPR$ and $TFPQ$ consist of single-product producers only and include product-year fixed effects. Our $TFPR$ results are robust to using the Levinsohn-Petrin (2003) methodology.

[22] Our methodology is similar to the one used by Li et al. (2018) but differs slightly since instead of following De Loecker et al. (2016) and use a translog production function, we use the Olley-Pakes (1996) methodology with a Cobb-Douglas production function to control for selection. This difference, and our larger coverage of sectors, can explain the discrepancy that while we find mixed exporters and pure processors to have the highest $TFPQ$, they find that pure processors' $TFPQ$ is higher than that of "hybrid" firms.

Table 3: Mixed Exporter Premia

(a) All exporters	PP_{it}		Mix_{it}		Obs.
(1) ln($empl.$)$_{it}$	0.30***	(0.07)	0.38***	(0.04)	208,514
(2) ln($labor\ prod.$)$_{it}$	-0.22***	(0.03)	0.14***	(0.03)	197,661
(3) $TFPR_{it}$	-0.14**	(0.07)	0.12***	(0.04)	9,297
(4) $TFPQ_{it}$	0.02*	(0.01)	0.03***	(0.01)	9,297
(5) ln($R\&D\ exp.$)$_{it}$	-0.81***	(0.15)	-0.27***	(0.05)	208,514
(6) ln($advert.\ exp.$)$_{it}$	-1.00***	(0.13)	-0.37***	(0.06)	193,919
(7) ln($trademarks$)$_{it}$	-0.47***	(0.05)	-0.18***	(0.03)	208,514
(b) Excl. foreign firms	PP_{it}		Mix_{it}		Obs.
(1) ln($empl.$)$_{it}$	0.21***	(0.06)	0.38***	(0.04)	159,938
(2) ln($labor\ prod.$)$_{it}$	-0.05	(0.04)	0.21***	(0.03)	152,073
(3) $TFPR_{it}$	-0.02	(0.06)	0.14***	(0.04)	7,037
(4) $TFPQ_{it}$	0.04**	(0.02)	0.04***	(0.01)	7,037
(5) ln($R\&D\ exp.$)$_{it}$	-0.78***	(0.17)	-0.24***	(0.06)	159,938
(6) ln($advert.\ exp.$)$_{it}$	-0.95***	(0.14)	-0.33**	(0.06)	149,466
(7) ln($trademarks$)$_{it}$	-0.46***	(0.06)	-0.19***	(0.04)	159,938

Notes: This table reports the results of running specification (1). Each row is a separate OLS regression of the dependent variable shown in column 1 on dummy variables PP_{it} and Mix_{it} that indicate whether the firm i is a pure processor or a mixed exporter in year t respectively (pure ordinary is the omitted group). ln($R\&D\ exp.$)$_{it}$, ln($advert.\ exp.$)$_{it}$, and ln($trademarks$)$_{it}$ are calculated by $\ln(x+1)$ to avoid dropping zeros. $TFPR_{it}$ and $TFPQ_{it}$ refer to TFP calculated using revenue and quantity data respectively (see the text for details). Rows 1-2 and 5-7 include sector-year fixed effects, and all except those in the first row control for firm size. Rows 3-4 focus on single-product producers only and thus include product-year fixed effects. Coefficients for the two dummy variables are significantly different from each other in all rows except for row 4 in both panels. Standard errors clustered by 2-digit CIC industries are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

processors).[23] Note that processing intensity as captured by the share of processing in total exports varies across mixed exporters with a mean of 58% and standard deviation 36%. In Appendix Table A.3 we focus on mixed exporters, and find qualitatively similar results for processing intensity.

We summarize our findings regarding firms' performance in the following stylized fact:

Fact 1: Mixed exporters are larger than pure processors, who are larger than pure ordinary exporters in terms of employment. Mixed exporters have higher labor and revenue productivity than pure ordinary exporters, who have higher labor and revenue productivity than pure processors. However, mixed exporters and pure processors have higher physical productivity than pure ordinary exporters.

If we view a mixed firm as a combination of a pure processing and a pure ordinary firm, we would expect that mixed firm characteristics lie between that of pure processing and pure ordinary firms, which stands in contrast with what we find in the data. One obvious rationalization would be that processing transactions have lower prices due to, for example, input tariff exemptions or

[23] In unreported results, we regress productivity on the processing share of exports, and find a linear and positive relationship with $TFPQ$ and a non-linear inverted-U relationship with $TFPR$. These results confirm the ones above with exporter-type dummies.

transfer pricing (Li et al., 2018), which would disproportionately distort the average export price of pure processors, and hence render the lowest $TFPR$. This could explain why the production efficiency ($TFPQ$) is greater for pure processors compared to ordinary exporters, but not why mixed exporters have the highest $TFPQ$.

An alternative hypothesis is that processing firms contribute to relatively less value-added stages of production (e.g., manufacturing), and thus get a lower share of profits when compared to their foreign buyers (Feenstra and Hanson, 2005; Dai et al., 2016; Manova and Yu, 2016). Given that most value-added comes from firms' non-manufacturing activities such as innovation and marketing, processing firms can be efficient in production yet have low $TFPR$. On the contrary, ordinary producers can claim more profits thanks to their branding activities, and hence can survive even with a relatively low $TFPQ$. This view also gives a natural explanation to the existence of mixed exporters: they are firms that excel in both manufacturing and non-manufacturing activities. The second hypothesis is also consistent with the fact that many prominent Chinese firms produce their own branded products while at the same time manufacture goods for other firms (Deng, 2021).[24]

To identify the dominant explanation among the two hypotheses, we use the 2018 customs sample to examine the relationship between product trade mode, price, and brand ownership of firms. As described in the data section, the 2018 customs dataset allows us to extract the brand ownership information for each export transaction, and label it as no brand, foreign brand, or domestic (own) brand. As shown in the last row of Table 4, 12.4%, 56.4%, and 32.7% of export value are due to transactions that have no brand, foreign brand, and domestic brand, respectively in our sample. Importantly, we find a tight link between the choice of processing trade mode, and the production of foreign branded goods. Table 4 shows that 84% of processing exports in this customs sample consists of foreign branded products, while only 33% of ordinary exports consists of foreign branded products. While processing transactions are typically viewed as local manufacturers supplying customized productions to their buyers (Manova and Yu, 2016), our data enable us to confirm this conjecture empirically.

We run the following transaction-level regression:

$$D_{ifhc} = \beta P_{ifhc} + \delta_{hc} + \epsilon_{ifhc}, \tag{2}$$

where D_{ifhc} is a dummy indicating whether firm f's export transaction i of product h (at the HS10 level) to country c is for its own Chinese domestic brand (as opposed to foreign or no brand), P_{ifhc} is a dummy for processing trade (as opposed to ordinary trade), δ_{hc} are HS10-country fixed effects to control for product-destination determinants of processing trade policy and brand ownership (e.g., FDI policy), and ϵ_{ifhc} is the error term. We cluster standard errors at the firm level. Table 5 column 1 shows that processing transactions are 13 percentage points less likely to involve products with

[24]For instance, *Shenzhou International*, a large Chinese textile manufacturer with its own brand, does processing for world-renowned brands such as *Adidas*, *Nike*, and *Uniqlo*. *Galanz*, a prominent home appliance producer to brands such as *De'Longhi*, *General Electric*, and *Sanyo* alongside exporting its own branded microwaves and air conditioners.

Table 4: Export Mode and Brand Ownership: Summary Statistics

	(1) No brand	(2) Foreign brand	(3) Domestic brand
Ordinary exports	14.3%	33.5%	52.2%
Processing exports	7.0%	83.9%	9.1%
Total	12.4%	56.4%	32.7%

Notes: This table reports the share of export modes in no brand, foreign brand, and domestic brand categories in columns 1, 2, and 3 respectively, using the 591,270 manufacturing export transactions in the 2018 customs data sample (after excluding the 271,297 transactions made by wholesalers and intermediary firms). We extract brand ownership information for each transaction from the reported string product specification using an algorithm (see the text for details), which we then classify as no brand, foreign brand, or domestic (own) brand. We classify the 45 export modes reported in the dataset into three broader groups: ordinary exports, processing exports, and other exports.

domestic brands when compared to ordinary transactions (significant at the 1% level). In column 2, we include firm-product-country fixed effects which implies that we are comparing transactions of the same HS10 sold to the same destination by the same firm.[25] Column 2 shows that the coefficient remains negative and significant at the 10% level: mixed firms' processing exports are 3.2 percentage points less likely to include their own branded products when compared to their ordinary exports of the same product to the same destination. Hence we arrive at the following stylized fact:

Fact 2: Ordinary transactions tend to involve firms' exports of their own branded products, whereas processing transactions tend to involve firms' exports of their customers' branded products.

In column 3, we regress the log unit value of transactions on brand ownership, controlling for export mode, and including product-country fixed effects. We find a positive relationship between brand ownership and unit values, even when we include firm-product-country fixed effects in column 4. The estimated coefficient indicates that a domestically branded product of a firm is about 9% more expensive than that same firm's sales of the same product to the same destination but under a different brand (significant at the 5% level). The positive correlations between non-processing export mode and brand ownership, as well as between brand ownership and brand premium support the hypothesis that price differences between processing and ordinary exporters can be explained by their specialization within a value chain. This results in the following stylized fact:

Fact 3: There is a price premium associated with selling one's own branded product.

Now let us turn to the first explanation that emphasized input price differences among exporters. If the observed $TFPR$ and $TFPQ$ differences between firms are due to processing exports being subject to lower input tariffs or preferential tax policies, then the export price for processing goods might be mechanically lower. However, the above conjecture would imply that within a

[25]There is enough variation even at this level as the average (median) number of transactions for each firm-product-country in our regression sample is 9.7 (2). Note also that 7% of the 15,078 firms in our regression sample are mixed, with the rest consisting of pure ordinary (82%) and pure processing firms (11%). The mixed firm-product-country flows make up 15% of total flows, with the rest consisting of pure ordinary (51%) and pure processing flows (34%).

Table 5: Export Mode and Brand Ownership: Regressions

Dependent var.:	D_{ifhc}		$\ln uv_{ifhc}$	
	(1)	(2)	(3)	(4)
P_{ifhc}	-0.126***	-0.032*	-0.072	0.092**
	(0.039)	(0.016)	(0.162)	(0.044)
D_{ifhc}			0.197*	0.088**
			(0.110)	(0.038)
Product-country FE	Yes	No	Yes	No
Firm-product-country FE	No	Yes	No	Yes
R^2	0.30	0.85	0.81	0.92
Obs.	445,437	427,567	419,009	402,169

Notes: This table reports the results of running specification (2). D_{ifhc} indicates whether transaction i of firm f in product h (at the HS10 level) to destination c is a domestic own brand transaction, P_{ifhc} indicates whether this transaction is classified under processing trade, and $\ln uv_{ifhc}$ is the log unit value of this transaction. Standard errors clustered by firms are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

firm-product-destination, processing exports should have a lower unit value, which contradicts our finding in Table 5. If transfer pricing is driving the results (i.e., processing exporters artificially lower the price of export transactions between enterprises under common ownership or control), then we would expect to see a less stark difference in $TFPQ$ between processing and ordinary firms once we exclude foreign firms—the results in Table 3 suggest the opposite. Therefore, we conclude that the higher average price of exporters' own products is more likely due to brand premium instead of input tariff exemptions or transfer pricing.

Finally, we provide some suggestive evidence that a firm's choice on export mode is indeed associated with its branding activities. Table 3 panel (a) rows 5, 6, and 7 reveal that R&D investment, advertisement expenditures, and number of trademarks across firms are in the following decreasing order: pure ordinary exporters, mixed exporters, and pure processors. In fact, 85% of pure processors did not have any R&D or advertising expenses in 2005. Whereas the average number of trademarks for pure ordinary and mixed exporters is 2.8, this figure is only 0.8 for pure processors. This is in line with anecdotal evidence that pure processors tend to specialize in manufacturing for other firms, and thus do not need to invest in R&D and trademarks or spend on advertisement, which are ultimately done by their customers. In panel (b) rows 5, 6, and 7, we exclude foreign firms since the majority of their R&D, advertising, and trademark expenses are likely to be done in their headquarter-countries, and thus are not perfectly observed in our data—the results are similar. In Appendix Table A.3, we find that for mixed exporters, as processing intensity increases, R&D and advertisement expenditures as well as the number of trademarks decrease as expected.

4 Processing Promoting Policy

If a firm's choice on export mode is indeed associated with its branding activities and reflect its position in the supply chain, the next question naturally arises: does encouraging "making" generate any positive effect on "creating"? Anecdotal evidence suggests that such a positive spillover is not rare: the success of *Xiaomi*, now the world's fourth-largest smartphone company, crucially relied on its world-leading suppliers such as *Inventec* and *Zepp*—companies that predominantly engaged in processing trade. *LifEase*, the "Chinese *Muji*" developed by *NetEase*, works directly with the suppliers for brands such as *Burberry*, *Gucci*, and *Rimowa* to produce its items.

We examine whether promoting processing trade helps downstream firms to eventually come up with their own branded products by exploiting China's experimentation with "paperless" processing trade in 2000-2006. This policy shock is highly suitable for our study as it affects only the cost of processing exports, leaving other exporting costs of a firm unchanged. To the best of our knowledge, this paper is the first to examine the effect of this policy.

4.1 China's "Paperless" Processing Trade

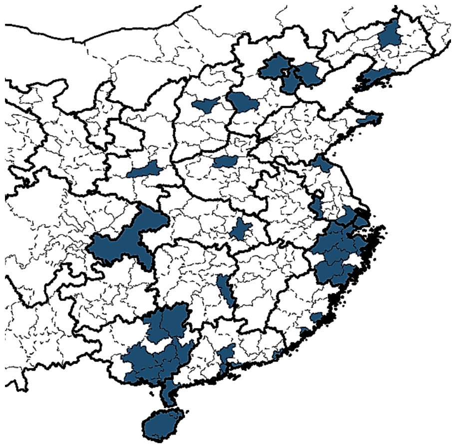
We first present the shock and the empirical context. China's customs authorities closely monitor the supply chain for processing trade because of special duty drawbacks granted to processing exporters. Thus, to organize processing trade, firms have to fill in grueling paperwork that details their financial condition and upstream and downstream connections for each contract, and then wait to get approved by the customs authority. In order to make processing trade less costly for firms, China began to experiment with an online supervision system in 2000. By connecting firms' computer management systems to the customs' online administration system, it made the processing trade application paperless, and thus significantly reduced the burden of red-tape on processing firms. As quoted from a news article by *International Business Daily*: *"...the traditional methods, from preparing the contract to getting approval, takes at least two weeks—sometimes one needs to visit several governmental offices hundreds of times. After adopting online supervision, the application takes less than an hour. As a result, the company's customs clearance costs are reduced by more than 20%, and the clearance speed is greatly improved."*[26]

The pilot program for paperless processing trade targeted Class A firms: firms that had at least \$10 million worth of exports. Favorable to our setting, this threshold of \$10 million was set by the Chinese authorities in 1999 as a way to classify firms for administrative purposes and is unrelated to the paperless processing trade program.[27] This policy experiment had a staggered introduction to different prefectures: between 2000 and 2006, customs authorities of 50 (out of 334) prefectures in 18 (out of 34) provinces of China adopted the pilot program, as illustrated in Figure 1. By the

[26]The original article is in Chinese and can be found at: `http://jm.ec.com.cn/article/jmzx/jmzxdfjm/jmzxguangzhou/200409/498189_1.html`; translated by the authors.

[27]As paperless supervision requires firms to have an Enterprise Resource Planning (ERP) system (a computer software for business management), customs authorities naturally targeted large firms for the pilot since most of them had already installed an ERP system. Hence, the threshold of \$10m provides a simple yet established selection criteria. See `http://www.people.com.cn/zixun/flfgk/item/dwjjf/falv/6/6-1-50.html` (Chinese) for the official firm classification notice, and `http://www.fdi.gov.cn/1800000121_39_1919_0_7.html` (Chinese) for the official notice that explains the pilot program that targets Class A firms. We observe firms' eligibility, but not whether they actually adopt the program. We exclude the electronics sector from our analysis since firms in this industry had a lower threshold (\$5m) to qualify for the pilot program.

Figure 1: Adoption of the Pilot Paperless Processing Trade Program



Notes: This map shows the 50 Chinese prefectures that adopted the pilot online supervision system during 2000-2006.

end of 2006, inspired by the success of the pilot program, the policy rolled over nation-wide and was made available to all processing firms, regardless of size.

4.2 The Direct Impact of "Paperless" Processing Trade

We first show that the pilot paperless program has been highly effective in increasing processing exports. In particular, we compare firms within a $1m bandwidth at the right and left side of the $10m threshold before and after the introduction of the "paperless" program. By incorporating a bandwidth, our approach resembles a regression discontinuity (RD) design with difference-in-differences (DD), similar to Bøler et al. (2015) who examine the effect of R&D policy in Norway using a difference-in-differences approach, and Jia (2014) who analyzes the effect of treaty ports on Chinese prefectures by selecting a control group based on balancing checks. As emphasized by Lemieux and Milligan (2008), selecting an appropriate control group in DD and thus having a DD-RD type of estimation is crucial to get unbiased treatment effect estimates given that the pre-treatment processing export trends of the treatment and control groups are parallel. This approach also allows us to take advantage of our panel data structure, using several years before and after the policy adoption, which enables us to estimate lagged effects. Moreover, our use of firm fixed effects allows us to focus strictly on within-firm variation, making DD-RD more robust to confounders when compared to a simple RD.

The direct impact of processing policy on processing trade is not our main interest, hence we relegate our detailed empirical analysis and robustness checks to Appendix B. The balancing checks in Table A.4 panel (b) reveal that our selected treatment and control group of firms are similar in almost all key aspects, while the full sample of firms are not (Table A.4 panel (a)). Figure A.1 panel (b) shows that the pre-trends between our treatment and control groups are similar, with the $10-11m firms increasing their processing exports sharply in $t+1$. In contrast, the pre-trends between firms below and above $10m when using the full sample are very different

(Figure A.1 panel (a)). Our baseline estimation in Table A.5 column 1 suggests that the pilot program increased firm-level processing exports by around 28%. Additional estimations in Tables A.5 and A.6 show that the result is robust to including a rich set of fixed effects, controlling for lagged processing shares, excluding foreign-owned firms, and using alternative bandwidths. Most importantly, our falsification tests with 'false' thresholds yield point estimates that are insignificant and close to zero. Similarly, when focusing on ordinary instead of processing exports of mixed firms, the coefficient of interest is insignificant.

4.3 Downstream Spillovers and Trademarks

We now turn to the downstream spillovers of the pilot paperless processing trade program. We hypothesize that by promoting firms that are good at manufacturing, the policy will in turn benefit downstream firms that are good at "creating" to develop their own brands. Existing empirical research suggests that supplier-buyer relationships are highly localized (Bernard et al., 2019b), and thus we expect that downstream firms in the same prefecture as the affected suppliers would be more likely to benefit from the spillover and thus apply for new trademarks.

We first define the "treated processing exports" for each prefecture-sector-year (cst):

$$\text{Treated processing exports}_{cst} = \sum_{i \in A} \text{processing exports}_{icst},$$

where $i \in A$ are processing firms that are above the \$10m threshold. Here sector s is defined based on the industry classification used in China's 2002 Input-Output (IO) table. To compute treated processing exports, we first concord HS8 from the customs data to the IO industry classification.[28] After adjusting for the one-to-many and many-to-many matches, we end up with a slightly more aggregated set of 74 IO industries. Then, we create a time-varying input shock as follows:

$$\text{Input shock}_{cnt} = \sum_{s} \omega_{ns} * \text{Treated processing exports}_{cst},$$

where ω_{ns} are cost share of upstream industry s in downstream industry n, which we calculate based on the Chinese 2002 IO table. We then run the following specification:

$$Y_{icnt} = \exp\Big(\beta \ \ln\left(\text{Input shock}\right)_{cnt} \times \text{Productive}_i + \lambda \ln\left(empl.\right)_{it} + \psi \ln\left(capital\right)_{it} + \gamma_i + \delta_{nt} + \phi_{ct}\Big) \times \epsilon_{icnt}, \quad (3)$$

where Y_{icnt} is the number of effective trademarks a firm has,[29] and Productive_i indicates whether the firm's initial log labor productivity is above the median value.[30] We include $\ln(empl.)_{it}$ and $\ln(capital)_{it}$ to control for firm-level employment and capital stock, firm fixed effects γ_i to control

[28] We thank Yu Shi for providing us with the HS8-IO industry correspondance table.

[29] Trademarks are the legal basis for brands and thus we are using the number of effective trademarks as a proxy for firms' branding activity.

[30] To make sure that we retain zeros, we add 1 to Input shock_{cnt} before taking the natural log and including it in our regressions.

for unobserved firm-level characteristics, sector-year fixed effects δ_{nt} to control for sector-specific supply and demand shocks, and prefecture-year fixed effects ϕ_{ct} to control prefecture-wide policy changes that might affect trademark applications.[31] Standard errors are clustered two-way at the prefecture and sector level. Due to the count nature of our dependent variable, we estimate specification (3) using a Poisson pseudo-maximum likelihood (PPML) model.[32]

Our identification relies on the plausible assumption that the timing of introducing the pilot paperless processing experiment by a prefecture's customs is exogenous to the branding activities of non-processing firms in the same region. To achieve clean identification, we exclude pure processing firms since the timing of the policy might be correlated with unobserved productivity shocks to local processing exporters, which at the same time could also affect these firms' branding activity. These sample modifications, however, do not change our results qualitatively.

Table 6 presents the estimation results. As suggested in column 1, we find that the adoption of the pilot program is positively associated with the number of trademarks of downstream firms, although the coefficient is statistically insignificant. This is intuitive, as almost 60% of below-median productive firms had at most one trademark between 2000-2006. Nevertheless, we expect that a greater input exposure to the pilot program should help productive firms to boost their trademark activity. Hence, in column 2, we interact the input shock variable with the Productive_i dummy, and find an interaction coefficient of 0.003, significant at the 1% level. The coefficient indicates that a one standard deviation (5.866) increase in $\ln(\text{Input shock})_{cnt}$ raises the number of trademarks of a productive firm by 0.012 ($(0.003 - 0.001) \times 5.866$), which is 1.2% of the median number of trademarks (1). In column 3, to allow for a more flexible effect, instead of the Productive_i dummy, we interact Input shock_{cnt} with the firm's demeaned initial labor productivity, $\ln(\textit{labor prod.})_i$, and the result stays robust.

In column 4 of Table 6, we directly control for $\text{Treated processing exports}_{cnt}$ of the firm's own industry as well as its interaction with Productive_i. We include this control since promoting processing policy might crowd out ordinary firms and hence directly affect their branding activities. The estimated coefficient remains the same, and we see that the own industry effects are nil—this is expected since pure processing firms who rarely engage in trademark activity are excluded from the sample. Column 5 excludes SOEs from the sample as these firms' trademark activities might be subject to government controls. In column 6, we estimate our specification using OLS instead of PPML. Neither of these robustness checks change the qualitative result. In column 7, the dependent variable is a dummy that indicates whether the firm has at least one effective trademark. In column 8, we use the log number of trademarks, which results in a smaller sample size due to dropping firms with no trademarks. The coefficients show that the input shock has positive effects on trademark development at both the extensive and the intensive margins. We also find that the

[31] Slightly more than a third of firms in our dataset have at least one effective trademark in 2000-2006. The average number of effective trademarks is 1.6, with standard deviation 9.6.

[32] For our PPML estimations, we use Correia et al.'s (2019) Stata package *ppmlhdfe*, which is robust to convergence issues inherent in maximum-likelihood estimation with multiple high-dimensional fixed effects.

Table 6: Trademarks with IO Linkages

Dep. var.: Y_{icnt}	(1) Overall effect	(2) Median	(3) Demeaned	(4) Output control	(5) No SOEs	(6) OLS	(7) Extensive margin	(8) Intensive margin
$\ln(\text{Input shock})_{cnt-1}$	0.001	-0.001	0.001	-0.001	-0.000	-0.008	-0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.008)	(0.002)	(0.001)
$\times$ Productive$_i$		0.003***		0.003***	0.002***	0.010***	0.001**	0.002***
		(0.001)		(0.001)	(0.001)	(0.002)	(0.000)	(0.000)
$\times \ln(\textit{labor prod.})_i$			0.001***					
			(0.000)					
$\ln(\text{Treated processing exports})_{cnt-1}$				-0.000				
				(0.000)				
$\times$ Productive$_i$				0.000				
				(0.000)				
$\ln(\textit{empl.})_{it}$	0.102***	0.100***	0.100***	0.100***	0.097***	0.274***	0.042***	0.057***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.016)	(0.004)	(0.003)
$\ln(\textit{capital})_{it}$	0.054***	0.054***	0.054***	0.054***	0.053***	0.135***	0.025***	0.029***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.012)	(0.002)	(0.002)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	408,051	408,051	408,051	408,051	366,052	408,051	408,051	370,409
pseudo-R^2	0.48	0.48	0.48	0.48	0.48	0.90 (R^2)	0.03	0.93 (R^2)

Notes: This table reports the results of running specification (3) using a PPML model. Y_{icnt} is the number of trademarks of firm i in downstream sector n residing in prefecture c in year t. Sectors refer to 57 downstream IO industries. Productive$_i$ indicates firms whose initial log labor productivity is larger than the median. Column 6 uses OLS instead of PPML. In column 7, Y_{icnt} is a dummy variable that indicates whether the firm has a trademark, whereas in column 8, Y_{icnt} is the log number of trademarks (estimated linearly). Standard errors clustered at the prefecture and sector level are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

number of employees and the capital stock have a positive and significant effect on trademarks in all regressions, as expected. Overall, results in Table 6 suggest that the pilot paperless processing trade program has induced more productive downstream firms to increase their branding activity.

We summarize our empirical finding in the following fact:

Fact 4: Chinese firms intensified their branding activities when faced with favorable processing trade policies upstream.

5 A Simple Model to Rationalize the Findings

The stylized facts and empirical results presented above lead us to view mixed exporters as firms that are superb in both manufacturing efficiency and branding ability. These two abilities jointly determine firms' export and specialization decisions and affect their observed characteristics. In this section, we provide a parsimonious model of multi-attributes firms to rationalize our empirical findings. In particular, we highlight two modeling pieces that help explain our results: (1) two-dimensional heterogeneity in making and creating, and (2) a positive but low profit margin in manufacturing. To emphasize the sufficient model structure that rationalizes empirical findings, rest of the model is as stylized as possible. Proofs of the main results are provided in Appendix C.

5.1 Model Setup

Consider an economy where consumer preferences are Cobb-Douglas over two sectors: a homogeneous sector producing one unit of product with one unit of labor, and a differentiated sector that is the focus of our analysis. A fraction β of income is spent on the differentiated sector and the preference across varieties is CES with elasticity of substitution $\sigma > 1$. The sector constitutes a continuum of firms, and each owns a blueprint to produce a single differentiated variety. The demand for variety j is:

$$q_j = A_1 z_j p_j^{-\sigma},$$

where A_1 denotes the aggregate demand shifter and z_j reflects the quality of the blueprint owned by firm j. Other things equal, varieties with better blueprints attract more demand. The price p_j refers to the price of variety j.

To link with our empirical findings, we distinguish between the "making" and "creating" of a variety. A variety's blueprint quality (i.e., creating) is associated with the firm who owns the blueprint. A variety's manufacturing efficiency (i.e., making), on the other hand, is tied to the production efficiency of the firm who makes it. We specify firms' making decision as the following. Production only requires labor, which is inelastically supplied (L).[33] As a manufacturer, firm j can produce both for its own and for other firms' blueprints. Its marginal cost of production is $1/t_j$

[33] Denote P the aggregate price index and N the endogenously determined mass of entrants, we impose one regularity condition that L is sufficiently large so that $\frac{P}{\theta N} < 1$ always holds.

when it produces its own variety, but when making for other firms, as every production contract is unique and has different manufacturing requirements, we assume production efficiency is subject to uncertainty. Specifically, for every blueprint, firm j draws a production efficiency from a Fréchet distribution with level parameter t_j and shape parameter θ, where $\theta > \sigma - 1$. That is, t reflects on average how good firm j is at manufacturing.

Analogously, as a blueprint holder, firm j can organize its production in-house or outsource production to other firms. It observes the t of all firms, but needs to pay a fixed cost f in terms of labor to discover a supplier's actual efficiency in manufacturing for its blueprint. In addition, outsourcing requires an additional f_o units of labor to coordinate production. A blueprint holder optimally chooses the number of reached suppliers, draw blueprint-specific productivity from each supplier, and choose the one with the lowest marginal cost of production as its contract manufacturer. We assume that ex-post gains are shared through Nash bargaining and the bargaining power of manufacturer is γ.[34]

There is an unbounded pool of prospective entrants who learn about their blueprint quality z and manufacturing ability t after incurring a fixed entry cost f_E. We assume that z and t are drawn from two distributions $G_z(z)$ and $G_t(t)$ with supports $(0, \bar{z}]$ and $(0, \bar{t}]$, respectively. Once firms draw their abilities, they decide whether (i) to bring own blueprint to production (in-house or outsource) and/or (ii) be active in manufacturing for other firms' blueprints. Bringing one's own blueprint to production requires an additional fixed cost f_B. Finally, there is a constant probability δ that forces a firm to exit in each period.

When it comes to international trade, it is natural to distinguish the trade costs associated with goods that are *made* domestically and exported, and goods that are *owned* by domestic firms and sold abroad. We assume that iceberg trade costs are associated with the "making" locations, i.e., $\tau_t > 1$ units are required to be shipped for one unit of domestic manufactured variety to be consumed in the foreign country, regardless of whether the blueprint is foreign or domestic. The fixed cost of exporting is typically associated with getting access to a certain market, and thus we assume that selling to foreign markets requires an additional cost f_X borne by domestic blueprint holders, regardless of where the goods are made. The homogeneous good is freely traded.

Processing promoting policies, depending on the type of instruments, can either be modeled as reducing the export costs of foreign varieties manufactured by domestic firms (such as the paperless program), or reducing fixed costs f_o (such as processing zones). As it will become clear, trade policies can only *shift* cutoffs and hence affect the share of firms engaging in certain type of exports, while two-dimensional heterogeneity is key to *generate* the three types of exporters and the observed ranks of their characteristics.

[34]The model predictions are robust to different assumptions of market structure; we refer interested reader to the previous version of the paper (Chen et al., 2020), where we consider Bertrand competition in manufacturing.

5.2 Firm Specialization

The model yields a natural specialization of firms within a value chain. From a blueprint holder's perspective, conditional on outsourcing, the least productive supplier that a profit-maximizing firm j contacts solves:

$$\underline{t}_j \equiv \underline{t}(z_j) = f(A(1-\gamma)\Gamma(\frac{\theta+1-\sigma}{\theta})z_j)^{-1}\frac{\theta}{\sigma-1}\Theta(z_j)^{1-\frac{\sigma-1}{\theta}}, \tag{4}$$

where $A = \frac{1}{\sigma}(1-\frac{1}{\sigma})^{\sigma-1}\beta L P^{\sigma-1}$, Γ stands for the gamma function, and $\Theta(z_j) = N\int_{\underline{t}_j(z_j)}^{\bar{t}} \iota dG_t(\iota)$ measures firm j's "sourcing pool." Intuitively, more suppliers that firm j contacts with, more likely it finds a manufacturer producing its variety at a low cost. Firms with better blueprints benefit more from contracting with a productive manufacturer, hence $\underline{t}_j$ decreases in z_j.

If firm j chooses to outsource (O) the production of its variety, the expected profit is given by:

$$\pi_j^O = \frac{(1-\gamma)}{\sigma}Az_jE(c_j^{1-\sigma})\left(1-\frac{\sigma-1}{\theta}\frac{\int_{\underline{t}_j}^{\bar{t}}\underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)}\right) - f_B - f_o, \tag{5}$$

where $E(c_j^{1-\sigma}) = \Gamma(\frac{\theta+1-\sigma}{\theta})\Theta(z_j)^{\frac{\sigma-1}{\theta}}$. If firm j chooses to produce in-house (I), its expected profit is:

$$\pi_j^I = Az_jt_j^{\sigma-1} - f_B.$$

Therefore, firm j will choose to produce its variety in-house if $\pi_j^I \geq \pi_j^O > 0$, outsource if $\pi_j^O > \pi_j^I > 0$, and exit otherwise. This yields three cutoff curves, and firm j would find it optimal to:

(1) outsource its variety if $z_j > z_1, z_j < \psi^{-1}(t_j)$,

(2) produce in-house if $z_j > \phi(t_j), z_j \geq \psi^{-1}(t_j)$, and

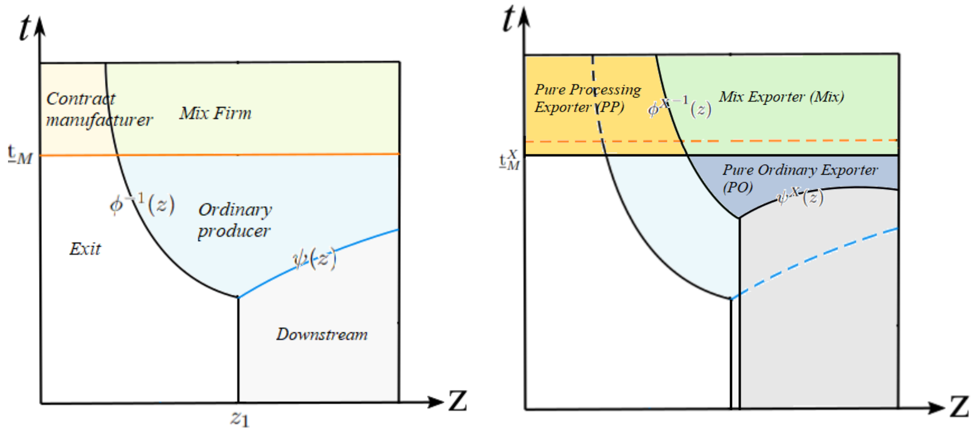
(3) exit otherwise,

where z_1 solves $\pi_j^O(z_1) = 0$, $\psi(z) = ((1-\gamma)E(c^{1-\sigma})(1-\frac{\sigma-1}{\theta}\frac{\int_{\underline{t}_j}^{\bar{t}}\underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)}) - f_o)^{\frac{1}{\sigma-1}}$ and $\phi(t) = \frac{f_B}{At^{\sigma-1}}$. As visualized in panel (a) of Figure 2, when production is outsourced, firm j's own manufacturing ability does not matter, therefore cutoff between exit and outsource is a straight vertical line. Blueprint quality and manufacturing ability are complements in production, thus the cutoff between in-house production and exit, $\phi^{-1}(z)$, is downward sloping. Lastly, since higher blueprint quality means a greater return from outsourcing, a firm with higher z is more likely to outsource given the same manufacturing ability. Therefore the cutoff $\psi(z)$ between in-house production and outsourcing is upward sloping.

As firms with better blueprints look for more potential suppliers, the active manufacturer with the least productivity has to be the firm that is reached by the blueprint holder with the best blueprint quality $\bar{z}$.[35] This yields the manufacturing cutoff $\underline{t}_M$, above which firms will be active in

[35] Because the manufacturing abilities are drawn from Fréchet, the least productive manufacturer would have a

Figure 2: Firm Specialization



(a) Closed economy specialization

(b) Open economy specialization

producing for other firms' blueprints:

$$\underline{t}_M \equiv \underline{t}(\bar{z}) = f(A(1-\gamma)\Gamma(\frac{\theta+1-\sigma}{\theta})\bar{z})^{-1}\frac{\theta}{\sigma-1}\Theta(\bar{z})^{1-\frac{\sigma-1}{\theta}}. \tag{6}$$

This is illustrated as the orange horizontal line in panel (a) of Figure 2.

Putting the decisions of contracted making and own production together, the model gives rise to firms' specialization based on their heterogeneity in two dimensions. Shown in panel (a) of Figure 2, firms with low z become pure contract manufacturers (light yellow area), and firms with high z and high t also produce their own variety and become mixed firms (light green area). Firms with intermediate z and t produce and only produce for their own blueprint, becoming ordinary producers (light blue area). Firms with high z but low t outsource production and become downstream firms (grey area). Firms with both low z and t exit (white area).

With international trade, the cutoff contract manufacturer for foreign firms satisfies $\underline{t}_M^X = \tau_t^\theta \underline{t}_M^*$. Since a large share (40%) of pure processors did not have any domestic sales in our sample period, we assume that the foreign blueprint qualities are higher such that the export processing cutoff is lower (i.e., $\underline{t}_M^X < \underline{t}_M$). For domestic varieties, the model yields three additional export cutoff curves and two new equilibrium decisions: export with in-house production and export with outsourced production. Note that cutoff between in-house production and outsource for exporting firms $\psi^X(z)$ is strictly above $\psi(z)$, since improved market access always lead to greater gains from outsourcing. As graphically presented in panel (b) of Figure 2, with international trade, a subset of entrants survive and a smaller subset of them export. Active manufacturing firms have higher manufacturing ability than firms that exit, while processing exporters have even higher manufacturing ability. If a firm has high manufacturing ability but low blueprint quality, it becomes a pure processing exporter. Similarly, those with the 'worst' blueprint quality exit, better ones operate in the domestic market,

positive chance of being the cheapest supplier. With a continuum number of firms, the supplier's production in equilibrium is positive by law of large numbers.

and the even better ones export and become pure ordinary exporters. If a firm excels at both manufacturing ability *and* blueprint quality, it becomes a mixed exporter.

5.3 Linking the Model to Empirics

Our model generates a rich set of firm types, but for the sake of empirical relevance, we focus on pure processing exporters (PP), pure ordinary exporters (PO), and exporters that engage in both activities (Mix). We now discuss how our simple framework can rationalize the empirical findings in sections 3 and 4.

Physical Productivity *Physical TFP* measures a firm' efficiency in transforming inputs into quantity outputs, which corresponds to the manufacturing ability t in our model. The processing export cutoff $\underline{t}_M^X$ ensures that t_{PO} is always lower than t_{PP} and t_{Mix}, and the downward sloping cutoff curve ϕ^{X-1} ensures that by selection there are always more firms with greater t being mixed than pure processing exporters. Therefore our model naturally generates the $TFPQ$ ranking observed in data: mixed exporters on average have the highest physical productivity, followed by processing exporters, and then by ordinary exporters.

Revenue and Labor Productivity The *log labor productivity* of a firm j is given by $LP(z_j, t_j) = ln\left(\frac{\pi^I(z_j,t_j)+\pi^M(t_j)+l(z_j,t_j)}{l(z_j,t_j)}\right)$, which can be expressed as an employment weighted average of its labor productivity of being a blueprint producer and a contract manufacturer.[36] Manufacturing is often considered as the least value-added stage in a value chain, which translates into a low-valued γ in our model. If γ is sufficiently small, processing exporters exhibit the lowest labor productivity. Mixed exporters with superior manufacturing and making ability, naturally exhibit greater labor productivity for their ordinary part of the production compared to ordinary exporters. However, greater manufacturing ability also implies more demand from outsourcing, which reduces the aggregate labor productivity of mixed firms. When the first force dominates, our model naturally generates the labor productivity ranking observed in data: $E_{Mix}(LP) > E_{PO}(LP) > E_{PO}(LP)$.

The model can also rationalize the ranking for the *revenue TFP*. To be consistent with the Olley-Pakes estimation of TFP, we can instead assume that varieties are produced using labor and capital with a Cobb-Douglas technology, with a share parameter on labor being α. In this case, the revenue TFP of firm j is given by:

$$TFPR(z_j, t_j) = ln(\frac{\pi^I(z_j, t_j) + \pi^M(t_j)}{{l_j}^{\alpha}{k_j}^{1-\alpha}}) \propto ln(\frac{\pi^I(z_j, t_j) + \pi^M(t_j)}{l_j}) = LP_j,$$

in equilibrium, where w_K is the rental price of capital. The ranking is therefore the same as that of labor productivity, which is consistent with the data.

[36] $LP(z_j, t_j)$ can be expressed as $\ln\left(\frac{\pi^I(z_j,t_j)}{l^B(z_j,t_j)}s^B + \frac{\gamma}{(\sigma-1)}(1-s^B)+1\right)$, where l^B and s^B are the level and share of employment used for producing j's own variety, respectively.

R&D and Advertisement Expenditures In the data, we find that pure ordinary exporters spend more on R&D and advertising than mixed exporters, who spend more than pure processing exporters. Suppose that firms draw their blueprint quality and manufacturing ability sequentially. After observing its z, a firm can choose whether to incur an additional cost $f^{RD}(a)$ to improve its blueprint quality to $za^{\frac{1}{\sigma-1}}$ before observing its manufacturing ability t.[37] Standard assumption that $f^{RD\prime} > 0$ and $f^{RD\prime\prime} > 0$ applies, and hence f^{RD} is an increasing function of z in equilibrium. As Figure 2 panel (b) illustrates, there are relatively more processing exporters with lower z compared to mixed and ordinary exporters, and thus the model predicts that pure processing exporters spend the least on R&D and advertising. Compared to mixed exporters, the downward-sloping cutoff $\phi^{X^{-1}}$ selects relatively more high z firms to become pure ordinary exporters. However, the upward-sloping outsourcing cutoff ψ^X at the same time also pushes more high z firms to become downstream firms (hence out of the comparison sample). When the first effect dominates, our model also rationalizes that pure ordinary exporters spend more on R&D and advertising than mixed exporters, as observed in the data.

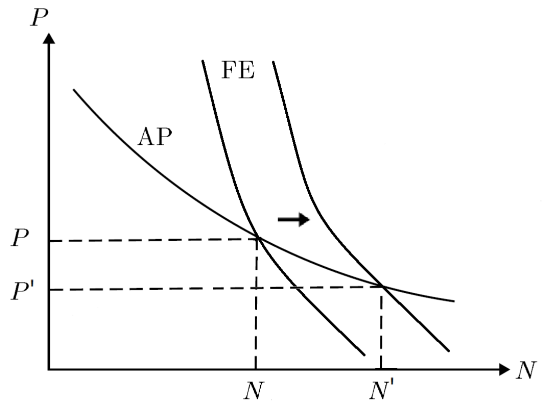
Employment On average mixed exporters have greater employment compared to pure processing exporters for two reasons. First, mixed exporters employ more labor for manufacturing other firms' varieties because they have greater t on average. Second, mixed exporters have additional labor for producing their own varieties. The employment ranking between processing and ordinary exporters is also intuitive. The key is to recognize that the production, and hence the employment of processing exporters, can be viewed as that of a "compound firm," whose production efficiency is given by manufacturers, but whose blueprint quality is given by outsourced blueprint holders. Therefore, this "compound firm" can have on average greater t and z compared to that of pure ordinary exporters, and thus greater employment. Hence our model also naturally generates the observed employment rankings in data.

Processing Trade Policy To highlight the idea, consider a small open economy setting such that changes at home do not affect any aggregate variables of foreign.[38] In Appendix C.1, we show that the equilibrium can be solved by solving the aggregate price index (AP) and free entry condition (FE) jointly as functions of P and N. Both curves are downward sloping, and the FE curve cuts the AP curve once from the above, ensuring the existence and uniqueness of the equilibrium (proved in Appendix C.2). When the processing promotion policy lowers the trade costs of foreign varieties manufactured by domestic firms, their ex-ante expected profits from manufacturing increase. Therefore, domestic firms' expected profits from bringing their blueprint into production must decrease for the free entry condition to hold: i.e., the FE curve shifts outwards. At the same time, the small open economy assumption ensures that the change in processing policy casts no direct impact on blueprint holders, therefore for a given N, the aggregate price index remains unchanged. As illustrated in Figure 3, these together imply that the equilibrium N increases while

[37]In this case, the blueprint quality distribution remains orthogonal to the distribution of t, and thus all other predictions derived from the model still hold.

[38]The small open economy assumption does not change our results qualitatively; see footnote 39.

Figure 3: Impact of promoting processing trade



P decreases.[39]

With the increase in both the mass of potential suppliers and the final goods market competition, we show in Appendix C.3 that the increase in net profits of outsourced domestic blueprint holders (i.e., the downstream firm corresponding to the data) is greater for firms with better blueprint quality, because they benefit more from having a greater pool of potential suppliers that they could source from. Intuitively, when the number of potential suppliers increases in equilibrium, firms will be more specialized in what they are relatively good at. Thus, promoting processing trade not only directly benefits suppliers with high manufacturing abilities ("Made in China") but also helps firms with good ideas ("Created in China").

We do not directly observe z in the data, but we can back out its value using some of the observables. We prove in Appendix C.3 that when processing trade costs decrease, the rise in net profits for a blueprint holder j increases in its labor productivity. In our empirical analyses, we examined firms' registration of trademarks. Trademarks are often symbols that identify goods as manufactured by a particular person or company and confer an exclusive right to use a specific brand (Baroncelli et al., 2005); hence we can view them as registered blueprints. If we extend our model by allowing firms to register their blueprints via costly trademark applications to avoid potential piracy, it is immediate that when τ_t decreases, firms with higher labor productivity would be more likely to register their trademarks, matching the empirical findings.

6 Conclusion

In this paper, we first unpacked the "black box" of mixed exporters that engage in both ordinary and processing exports. Contrary to the existing literature that describes processing firms as inferior,

[39]Without the small economy assumption, processing policy will lower the production costs of foreign varieties, which would shift the AP curve downwards, generating a competitive effect that has been the focus of some papers in the processing trade literature (e.g., Deng, 2016; Brandt et al., 2019). With the help of Figure 3, one can immediately see that this would push equilibrium N further up and P further down, but would not (qualitatively) change the model's prediction.

we showed that mixed firms, who engage predominantly in processing, are superior to other firms in multiple dimensions. Using a unique transaction-level customs dataset with branding information, we then provided novel stylized facts on the relationship between exporters' performance, export mode, and brand ownership. In particular, making and exporting products under other firms' brands are typically done via processing trade with significantly lower prices, which rationalizes the observed physical versus revenue TFP rankings between mixed, pure ordinary, and pure processing exporters. These relationships hold even within firm-product-destination level, suggesting that making and branding decisions need to be considered jointly at a disaggregate level. Using China's pilot "paperless" processing supervision program in 2000-2006 as a quasi-natural experiment, we also found that promoting processing trade induced domestic downstream firms to establish their own trademarks.

To rationalize our empirical findings, we provided a simple theoretical framework where multi-attributes firms endogenously determine their specialization. In particular, the model yields a novel positive impact of processing trade policy: facilitating processing trade leads to a greater mass of potential suppliers, which eventually benefits downstream firms with good ideas. Overall, our theoretical and empirical analyses highlighted that firms can be good at different stages of the value chain, and these heterogeneous abilities do not necessarily translate into a single measure for firm performance. Finally, our analysis showed that processing trade can lead goods to be not only "Made in China," but also "Created in China" by providing a breeding ground of suppliers for firms with good ideas.

References

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

Antràs, P., Fort, T. C., and Tintelnot, F. (2017). The margins of global sourcing: Theory and evidence from US firms. *American Economic Review*, 107(9):2514–64.

Antràs, P. and Helpman, E. (2004). Global sourcing. *Journal of Political Economy*, 112(3):552–580.

Ariu, A., Breinlich, H., Corcos, G., and Mion, G. (2019). The interconnections between services and goods trade at the firm-level. *Journal of International Economics*, 116:173–188.

Baroncelli, E., Fink, C., and Javorcik, B. S. (2005). The global distribution of trademarks: Some stylised facts. *The World Economy*, 28(6):765–782.

Bernard, A. B., Blanchard, E. J., Van Beveren, I., and Vandenbussche, H. (2019a). Carry-along trade. *The Review of Economic Studies*, 86(2):526–563.

Bernard, A. B. and Jensen, J. B. (1995). Exporters, jobs, and wages in US manufacturing: 1976-1987. *Brookings Papers on Economic Activity. Microeconomics*, 26(1995):67–119.

Bernard, A. B. and Jensen, J. B. (1999). Exceptional exporter performance: Cause, effect, or both? *Journal of International Economics*, 47(1):1–25.

Bernard, A. B. and Jensen, J. B. (2004). Why some firms export. *Review of Economics and Statistics*, 86(2):561–569.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics*, 4(1):283–313.

Bernard, A. B. and Moxnes, A. (2018). Networks and trade. *Annual Review of Economics*, 10:65–85.

Bernard, A. B., Moxnes, A., and Saito, Y. U. (2019b). Production networks, geography, and firm performance. *Journal of Political Economy*, 127(2):639–688.

Bernard, A. B., Moxnes, A., and Ulltveit-Moe, K. H. (2018). Two-sided heterogeneity and trade. *Review of Economics and Statistics*, 100(3):424–439.

Brandt, L., Biesebroeck, J. V., and Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics*, 97(2):339–351.

Brandt, L., Li, B., and Morrow, P. M. (2019). Is processing good?: Theory and evidence from China. *mimeo.*

Brandt, L. and Morrow, P. M. (2017). Tariffs and the organization of trade in China. *Journal of International Economics*, 104(C):85–103.

Bøler, E. A., Moxnes, A., and Ulltveit-Moe, K. H. (2015). R&D, international sourcing, and the joint impact on firm performance. *American Economic Review*, 105(12):3704–3739.

Chaney, T. (2016). Networks in international trade. In Bramoulle, Y., Galleoti, A., and Rogers, B., editors, *The Oxford Handbook of the Economics of Networks*, pages 754–75. Oxford University Press.

Chen, Z., Erbahar, A., and Zi, Y. (2020). Made and created in China: Super processors and two-way heterogeneity. *CEPR Discussion Paper Series DP14486.*

Chen, Z., Zhang, J., and Zheng, W. (2017). Import and innovation: Evidence from Chinese firms. *European Economic Review*, 94(C):205–220.

Correia, S., Guimarães, P., and Zylkin, T. (2019). Ppmlhdfe: Fast Poisson estimation with high-dimensional fixed effects. *arXiv preprint arXiv:1903.01690.*

Dai, M., Maitra, M., and Yu, M. (2016). Unexceptional exporter performance in China? The role of processing trade. *Journal of Development Economics*, 121(C):177–189.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

Defever, F. and Riaño, A. (2017). Subsidies with export share requirements in China. *Journal of Development Economics*, 126(C):33–51.

Deng, J. (2016). Processing trade, VAT rebate and international trade: Evidence from China's firm-level data. *mimeo.*

Deng, J. (2021). Processing trade and global idea diffusion. *mimeo.*

Deng, J. and Wang, Z. (2021). Processing-trade-induced Dutch disease. *mimeo.*

Deng, X., Jing, R., and Liang, Z. (2020). Trade liberalization and domestic brands: Evidence from China's accession to the WTO. *The World Economy*, 43(8):2237–2262.

Dhyne, E., Kikkawa, A. K., Mogstad, M., and Tintelnot, F. (2021). Trade and domestic production networks. *The Review of Economic Studies*, 88(2):643–668.

Eberhardt, M., Helmers, C., and Yu, Z. (2016). What can explain the Chinese patent explosion? *Oxford Economic Papers*, 69(1):239–262.

Feenstra, R. C. and Hanson, G. H. (2005). Ownership and control in outsourcing to China: Estimating the property-rights theory of the firm. *The Quarterly Journal of Economics*, 120(2):729–761.

Fernandes, A. P. and Tang, H. (2012). Determinants of vertical integration in export processing: Theory and evidence from China. *Journal of Development Economics*, 99(2):396–414.

Fernandes, A. P. and Tang, H. (2015). Scale, scope, and trade dynamics of export processing plants. *Economics Letters*, 133(C):68–72.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425.

Hallak, J. C. and Sivadasan, J. (2013). Product and process productivity: Implications for quality choice and conditional exporter premia. *Journal of International Economics*, 91(1):53–67.

Harrigan, J. and Reshef, A. (2015). Skill-biased heterogeneous firms, trade liberalization and the skill premium. *Canadian Journal of Economics*, 48(3):1024–1066.

Huang, H., Manova, K., and Pisch, F. (2021). Firm heterogeneity and imperfect competition in global production networks. *mimeo*.

Jia, R. (2014). The legacies of forced freedom: China's treaty ports. *Review of Economics and Statistics*, 96(4):596–608.

Johnson, R. C. (2018). Measuring global value chains. *Annual Review of Economics*, 10(1):207–236.

Kee, H. L. and Tang, H. (2016). Domestic value added in exports: Theory and firm evidence from China. *American Economic Review*, 106(6):1402–36.

Kikkawa, A. K., Magerman, G., and Dhyne, E. (2019). Imperfect competition in firm-to-firm trade. *ECARES Working Papers*, 2019-05.

Lemieux, T. and Milligan, K. (2008). Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics*, 142(2):807–828.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Li, Y. A., Smeets, V., and Warzynski, F. (2018). Processing trade, productivity and prices: Evidence from a Chinese production survey. *HKUST IEMS Working Paper*, 2018-58.

Lim, K. (2018). Endogenous production networks and the business cycle. *mimeo*.

Lu, D. (2010). Exceptional exporter performance? Evidence from Chinese manufacturing firms. *mimeo*.

Manova, K. and Yu, Z. (2016). How firms export: Processing vs. ordinary trade with financial frictions. *Journal of International Economics*, 100(C):120–137.

Manova, K. and Yu, Z. (2017). Multi-product firms and product quality. *Journal of International Economics*, 109:116–137.

Mayer, T. and Ottaviano, G. I. P. (2008). The happy few: The internationalisation of European firms. *Intereconomics*, 43(3):135–148.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):263–97.

Ornaghi, C. (2006). Assessing the effects of measurement errors on the estimation of production functions. *Journal of Applied Econometrics*, 21(6):879–891.

Tintelnot, F. (2017). Global production with export platforms. *The Quarterly Journal of Economics*, 132(1):157–209.

Wei, S.-J., Xie, Z., and Zhang, X. (2017). From "Made in China" to "Innovated in China": Necessity, prospect, and challenges. *Journal of Economic Perspectives*, 31(1):49–70.

World Bank (2018). World Bank Open Data [available at: `https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS?locations=CN`].

Yu, M. (2015). Processing trade, tariff reductions and firm productivity: Evidence from Chinese firms. *The Economic Journal*, 125(585):943–988.

Appendix A Calculating Physical TFP

To calculate physical TFP, we use the firm-product level production survey conducted by the NBS in China. This survey records information on products produced by all SOEs and private firms that have annual sales of at least five million RMB in 2000-2006.[40] To be able to assign an export mode for each firm, we merge this database with the merged Chinese customs-AIS dataset using unique firm IDs. Then, to obtain reliable productivity estimates at the firm level, we focus on single-product firms. Counting by the number of firm-product-year observations, single-product firms account for 56% of observations. Considering the relatively large amount of single-product observations, we expect that focusing on these observations will not severely bias our results. To ensure that the sample size is large enough to perform the estimation, we keep product categories with at least 2,000 firm-year observations and at least four years of existence.[41] Moreover, for each product category we require that there are at least 50 yearly observations. This results in a sample of 36 products (out of 693 manufacturing products) and 145,832 firm-year observations. Table A.2 lists the 36 products with their brief descriptions.

A.1 Methodology and Estimation

Our goal is to compare the production efficiency of exporters with different export modes. Following Foster et al. (2008), we use quantity data to get rid of the estimation bias caused by the heterogeneity in output pricing. Because we do not have information on firms' inputs, the input price dispersion may also bias our productivity estimates. To deal with this concern, we follow De Loecker et al. (2016) and use output prices to control for the input price dispersion. Note that for the final sample with single-product firms, 19% of firms exit before the end of sample period. This attrition rate can potentially cause a selection bias as first pointed out by Olley and Pakes (1996). To deal with this concern, we also control for firm exit.[42] We outline the estimation framework below.

The log-linearized Cobb-Douglas production technology for firm i in period t is assumed to be in the form of:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma m_{it} + \omega_{it} + \varepsilon_{it}, \tag{7}$$

where q_{it} is output quantity of firm i in year t, k_{it} is fixed assets, l_{it} is the number of employees, m_{it} is materials, ω_{it} is physical productivity, and ε_{it} is the productivity shock that is exogenous to the firm's production decision. We aim to estimate ω_{it}, which is observable to the firm but not to the econometrician.

Most of the existing literature has estimated TFP using deflated revenue data. However, these output price deflators are usually at the industry level, and thus they ignore the heterogeneity in

[40] See Li et al. (2018) for a detailed description of the production survey.

[41] As a robustness check, we change the threshold to 1,000 and results are qualitatively the same.

[42] In addition to using a Cobb-Douglas instead of a translog production function, our methodology slightly differs from Li et al. (2018) as we control for selection using the Olley-Pakes method.

firms' prices within an industry. As a consequence, the estimated productivity contains information on output prices, causing revenue productivity ($TFPR$) to be systematically different than physical productivity ($TFPQ$). The quantity data helps us to control for the output price dispersion if we can observe firms' input usage. Unfortunately, like in most other production survey datasets, we do not have information on the amount (in quantities) of each input used for production. However, we do observe the total expenditure on materials, denoted by $\tilde{m}_{it}$. Letting p_{Mit} be the log of material prices, we immediately have:

$$m_{it} = \tilde{m}_{it} - p_{Mit}. \tag{8}$$

If we use the industry-level material price index p_{Mjt} to deflate material expenditures, the material input used in the production function can be written as:

$$\bar{m}_{it} = \tilde{m}_{it} - p_{Mjt}. \tag{9}$$

Plugging (9) into (8), we can express the quantity of materials as:

$$m_{it} = \bar{m}_{it} + p_{Mjt} - p_{Mit}.$$

Therefore, we can rewrite the production function as:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma \bar{m}_{it} + \omega^*_{it} + \varepsilon_{it}, \tag{10}$$

where:

$$\omega^*_{it} = \omega_{it} + \gamma(p_{Mjt} - p_{Mit}).$$

This implies that the productivity obtained will contain information on input prices: $p_{Mjt} - p_{Mit}$. This input price bias can potentially create misleading results about the productivity differences for different types of exporters, especially if this input price is also correlated with export mode. This is of particular concern because processing exporters can use imported materials duty-free (as long as the output that uses these materials is exported).

The existing literature has also documented the necessity of controlling for input prices in estimating production functions (Ornaghi, 2006). Taking advantage of the quantity and revenue data, we control for the firm's input price using its output price. The underlying assumption is that the output price contains information on the firm's input price within a narrowly defined product category. Specifically, denoting p_{it} as the output price, the input price is assumed to be a non-parametric function of p_{it} and other firm characteristics:

$$p_{Mit} = f(p_{it}, \mathbf{X}_{it}). \tag{11}$$

This allows us to express physical material input as:

$$m_{it} = \tilde{m}_{it} - f(p_{it}, \mathbf{X}_{it}).$$

Thus, the production function we estimate is given by:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma \tilde{m}_{it} + \gamma f(p_{it}, \mathbf{X}_{it}) + \omega_{it} + \varepsilon_{it}. \tag{12}$$

In our estimations, we use sales and quantity data to construct output price in the following way:

$$p_{it} = \log\left(\frac{R_{it}}{Q_{it}}\right), \tag{13}$$

where R_{it} and Q_{it} are firm i's sales in values and quantities respectively in year t. We follow the Olley-Pakes methodology except that in the first-stage estimation, in addition to k_{it}, l_{it}, and $\tilde{m}_{it}$, we add polynomials of logged output prices to control for material prices. We also control for firm exit as a function of polynomials of capital stock, investment, and year dummies. This allows us to address the potential selection bias caused by less productive firms exiting the sample. To account for heterogeneity in production technology, we perform the estimation product by product.[43] Once we estimate the production function coefficients, we then compute our physical productivity ($TFPQ$) estimates, which are used in the regressions in Table 3.

[43]The production function estimation results are available upon request.

Appendix B The Direct Impact of China's "Paperless" Program

In this section, we test the direct impact of the paperless program. Note that the \$10 million threshold is a high bar: around 90% of processing firms export less than \$10m in a given year, and more than half of processing firms in the sample export less than \$1m worth of goods annually. As shown in Table A.4 panel (a) columns 1-3, compared to the rest of firms in our data, firms that are above the \$10m threshold are more likely to be mixed, more likely to be importers, less likely to be exiters or entrants, less likely to be foreign-owned, and more likely to be SOEs. They are also more processing-oriented and grow faster on average. The last two rows use variables from the merged AIS-customs data and reveal, expectedly, that the above-threshold firms are significantly larger both in terms of employment and capital. Moreover, Figure A.1 panel (a) suggests that the processing export pre-trends of the two groups are not parallel, which would threaten the identification strategy in a simple difference-in-differences (DD) framework.

To address this, we compare firms that exported between \$10-11m worth of processing goods with firms that exported between \$9-10m before the policy was introduced. By incorporating this bandwidth, our approach resembles a regression discontinuity (RD) design with difference-in-differences (DD-RD). As emphasized by Lemieux and Milligan (2008), selecting an appropriate control group in DD and thus have a DD-RD type of estimation is crucial to get unbiased treatment effect estimates given that the pre-treatment processing export trends of the treatment and control groups are parallel. This approach also allows us to take full advantage of our panel data structure, using several years before and after the policy adoption, which enables us to estimate lagged effects. Moreover, our use of firm fixed effects allows us to focus strictly on within-firm variation, making DD-RD more robust to confounders when compared to a simple RD.

The balancing checks in Table A.4 panel (b) reveal that our selected treatment and control group of firms are similar in almost all key aspects. There are two statistically significant discrepancies between the two groups: \$10-11m firms are slightly more processing oriented (89% versus 84%) and they are less likely to be foreign-owned (45% versus 51%). With firm fixed effects, we control for ownership and partially for the difference in processing shares, but we do two further robustness checks: we restrict the sample to non-foreign firms, and include lagged processing share as a control. Most importantly, Figure A.1 panel (b) shows that the pre-trends between the chosen treatment and control groups are similar, with the \$10-11m firms increasing their processing exports sharply in $t + 1$. Note that even though our choice of bandwidth is a relevant and restrictive bandwidth for processing exports that still allows some variation for our independent variable, our results are qualitatively insensitive to alternative bandwidths as shown in our robustness checks.

We start by running the following DD-RD specification at the firm-level to test the direct effect of the policy:

$$\ln(\text{proc. exp.})_{icst} = \alpha + \beta \; OS_{ict-1} + \gamma_i + \delta_{st} + \phi_{ct} + \epsilon_{icst}, \tag{14}$$

where $\ln(\text{proc. exp.})_{icst}$ is the processing exports of firm i that resides in prefecture c, with its core

HS2 sector s, in year t.[44] OS_{ict-1} is a dummy variable that indicates the adoption of the pilot paperless processing trade program in prefecture c in year $t-1$ that targeted firm i, γ_i are firm fixed effects, δ_{st} are sector-year fixed effects to control for overall supply and demand shocks, ϕ_{ct} are prefecture-year fixed effects to capture aggregate prefecture shocks, and ϵ_{icst} is the error term. We cluster standard errors two-way at the prefecture and sector level to allow for correlated shocks. Our main independent variable OS_{ict-1} is lagged by one year to allow some time for firms to adapt to the new declaration system. Since we do not observe whether the firm is actually using the paperless system, the estimate of β in (14) should be interpreted as an intention-to-treat effect.

We report the estimation results of (14) in Table A.5. The first column shows the benchmark result: firms that are in the treatment group in year $t-1$ increase their processing exports by 28% in year t, relative to the control group of firms with \$9-10m of exports in the year prior to policy adoption. An important identification concern is that the exact implementation time of the pilot program may be known to firms beforehand, making the timing of the policy adoption correlated to firms' strategic decisions. In column 2, we use a leads and lags strategy to rule out anticipation effects, and find that the lead variable OS_{ict+1} is not statistically different from zero, while the coefficient of OS_{ict-1} barely changes when compared to column 1. In column 3, we control for lagged processing share since our balancing checks in Table A.4 indicate that the \$10-11m firms are slightly more processing-oriented than the \$9-10m firms—the coefficient remains identical. Similarly, In column 4, we exclude foreign firms since our balancing checks show that there were more foreign-owned firms in the \$9-10m sample when compared to the \$10-11m sample. This results in a larger and more precisely estimated coefficient. In column 5, we change our dependent variable to processing intensity (i.e., share of exports that are processing) and find that the treatment causes firms to focus more on processing instead of ordinary exports.

In Table A.6, we show that our results are not sensitive to controlling for entry and exit in column 1, using a first-difference specification in column 2, using alternative bandwidths of \$9.5-10.5m and \$8.5-11.5m respectively in columns 3 and 4, or restricting the sample to always exporters or non-SOEs respectively in columns 5 and 6. Column 7 does a falsification analysis by focusing on the ordinary exports of mixed exporters, which shows a coefficient that is not statistically different than zero. On the contrary, column 8 shows that mixed exporters do increase their processing exports as expected. In columns 9 and 10, we do falsification analyses by setting the threshold to \$9m and \$11m, and the bandwidth to \$8-10m and \$10-12m respectively—coefficients in both columns are not statistically different than zero. These robustness checks support our finding that the pilot program increased firm-level processing exports.

[44]We assign a core HS2 sector to each exporter based on the ranked value of exports in its initial export year.

Appendix C Theory Appendix

C.1 Solving the Model and Comparative Statics

Conditional on blueprint holder j being connected with i, the probability that a manufacturer i is the lowest-cost supplier is:

$$\lambda_{ij} \equiv \lambda(z_j, t_i) = \frac{t_i}{\Theta(z_j)}, \tag{15}$$

where $\Theta(z_j) \equiv \Theta_j = N \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota)$. Given the well-known properties of the Fréchet distribution, the probability that the least-cost supplier's marginal cost of production is smaller than c is given by $Pr(c_j \leq c) = 1 - e^{-\Theta_j c^\theta}$ and expected marginal cost of production of j is therefore $E(c_j) = \Theta_j^{\frac{-1}{\theta}} \Gamma(\frac{\theta+1}{\theta})$ and $E(c_j^{1-\sigma}) = \Gamma(\frac{\theta+1-\sigma}{\theta})\Theta_j^{\frac{\sigma-1}{\theta}}$. The measure of reached suppliers equals:

$$n_j = N \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota).$$

Firm j's net profits (from its own blueprint) when production is outsourced is given by:

$$\pi_j^O = (1-\gamma) A z_j E(c_j^{1-\sigma}) - n_j f - f_B - f_o, \tag{16}$$

where $A = \frac{1}{\sigma}(1-\frac{1}{\sigma})^{\sigma-1}\beta L P^{\sigma-1}$. Taking the first derivative with respect to $\underline{t}_j$ of the above equation, we get:

$$\frac{\sigma-1}{\theta}(1-\gamma) A z_j \Gamma(\frac{\theta+1-\sigma}{\theta}) N^{\frac{\sigma-1}{\theta}} \left(\int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota) \right)^{\frac{\sigma-1}{\theta}-1} \underline{t}_j = Nf. \tag{17}$$

This yields the optimal cutoff $\underline{t}_j$ that satisfies:

$$\underline{t}_j \equiv \underline{t}(z_j) = f((1-\gamma) A z_j E(c_j^{1-\sigma}))^{-1} \frac{\theta}{\sigma-1} \Theta(z_j). \tag{18}$$

Note that the first order condition (17) implies:

$$(1-\gamma) A z_j E(c_j^{1-\sigma}) \times \frac{\sigma-1}{\theta} \frac{\underline{t}_j}{\Theta(z_j)} = f, \tag{19}$$

at the equilibrium. Using equation (19) to substitute f in (16), we get the expected profit of j if the firm chooses to outsource the production of its variety:

$$\pi_j^O = (1-\gamma) A z_j E(c_j^{1-\sigma}) \left(1 - \frac{\sigma-1}{\theta} \frac{\int_{\underline{t}_j}^{\bar{t}} \underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)} \right) - f_B - f_o. \tag{20}$$

On the other hand, if firm j chooses to produce in house (I), its expected profit is:

$$\pi_j^I = A z_j t_j^{\sigma-1} - f_B.$$

Clearly, from the blueprint holder's perspective, firm j chooses to produce its variety in-house iff $\pi_j^I \geq \pi_j^O > 0$, outsource iff $\pi_j^O > \pi_j^I > 0$, and exit otherwise. This yields three cutoffs:

$$\pi^I(z,t) = 0 \quad \Rightarrow \quad z = \frac{f_B}{At^{\sigma-1}},$$

$$\pi^O(z_1) = 0,$$

$$\pi^O(z) = \pi^I(z,t) \quad \Rightarrow \quad t = ((1-\gamma)E(c^{1-\sigma})(1 - \frac{\sigma-1}{\theta}\frac{\int_{\underline{t}_j}^{\bar{t}} \underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)}) - f_o)^{\frac{1}{\sigma-1}}.$$

Therefore, firm j as a blueprint holder would find it optimal to outsource its variety if:

$$z_j > z_1, z_j < \psi^{-1}(t_j), \tag{21}$$

and produce in house if:

$$z_j > \phi(t_j), z_j \geq \psi^{-1}(t_j), \tag{22}$$

where z_1 solves $\pi_j^O(z_1) = 0$, $\psi(z) \equiv ((1-\gamma)E(c^{1-\sigma})(1 - \frac{\sigma-1}{\theta}\frac{\int_{\underline{t}_j}^{\bar{t}} \underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)}) - f_o)^{\frac{1}{\sigma-1}}$ and $\phi(t) \equiv \frac{f_B}{At^{\sigma-1}}$. On the other hand, the actively producing firm with the least manufacturing ability can only be reached by firms with the best blueprint quality $\bar{z}$ (in comparative statics below we show formally $\frac{\partial \underline{t}_j}{\partial z_j} < 0$). This yields the manufacturing cutoff $\underline{t}_M$, above which firms will be active in producing for other firms' blueprints:

$$\underline{t}_M \equiv \underline{t}(\bar{z}) = f(A(1-\gamma)\Gamma(\frac{\theta+1-\sigma}{\theta})\bar{z})^{-1}\frac{\theta}{\sigma-1}\Theta(\bar{z})^{1-\frac{\sigma-1}{\theta}}. \tag{23}$$

Firm i's expected profit from manufacturing other firms' blueprints is given by:

$$\pi_i^M \equiv \pi^M(t_i) = \sum_{\{j:\underline{t}_j \leq t_i\}} \gamma \lambda_{ij} A z_j E(c_j^{1-\sigma}).$$

Intuitively, π_i^M equals the sum of expected profits from contracting with potential blueprint holders times the probability that the firm actually matches with each of these blueprint holders.

From the above analyses we solved for firms' decisions given the aggregate price index P and the mass of entrants N. Additionally, the aggregate price (AP) index is given by:

$$P^{1-\sigma} = N\left(\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} z_j \tilde{p}(z_j)^{1-\sigma} g(z,t)dtdz + \int_{z_1}^{\bar{z}}\int_{\psi(z)}^{\bar{t}} z_j p_j^{1-\sigma} g(z,t)dtdz + \int_{z_2}^{z_1}\int_{\phi^{-1}(z)}^{\bar{t}} z_j p_j^{1-\sigma} g(z,t)dtdz\right), \tag{24}$$

where $p_j = ((1-\frac{1}{\sigma})t_j)^{-1}$ and $\tilde{p}(z_j)^{1-\sigma} = (1-\frac{1}{\sigma})^{\sigma-1}E(c(z_j)^{1-\sigma})$. Note that although the actual marginal cost of production for an outsourced variety is a random variable, since there is a continuum of varieties the law of iterated expectations applies, and thus we can write the aggregate price index as such.

The free entry condition (FE) is given by:

$$\int_{\underline{t}_M}^{\bar{t}} \pi^M(t)dG_t(t) + \int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \pi^O(z)g(z,t)dtdz + \int_{z_1}^{\bar{z}}\int_{\psi(z)}^{\bar{t}} \pi^I(z,t)g(z,t)dtdz + \int_{z_2}^{z_1}\int_{\phi^{-1}(z)}^{\bar{t}} \pi^I(z,t)g(z,t)dtdz = \delta f_E, \tag{25}$$

where $\pi^M(t)$ is the profit from being a contract manufacturer, and z_2 is the blueprint quality cutoff of firms that produce in-house, i.e., z_2 solves $\pi_j^I(z_2,\bar{t}) = 0$. Therefore, we have two equations to solve for two unknowns and hence reach the equilibrium. Similar to the case with international trade, given the mass of entrants and the aggregate price indices in both countries, firms' optimal sourcing and operating decisions can be determined. By plugging the associated variables as functions of N, N^*, P, P^* into the aggregate price equations and the free entry conditions for home and foreign, we can solve for the equilibrium.

Comparative statics for z_j, P, and A It is easy to show that the second-order condition of the optimization problem requires that $\theta > \sigma - 1$. Recall that optimal cut-off for sourcing is:

$$\underline{t}_j \equiv \underline{t}(z_j) = f((1-\gamma)Az_j\Gamma(\frac{\theta+1-\sigma}{\theta}))^{-1}\frac{\theta}{\sigma-1}\Theta(z_j)^{1-\frac{\sigma-1}{\theta}}. \tag{26}$$

Note that A and z_j always show up multiplicatively, and hence it is sufficient to do comparative statics for one of them. Without loss of generality we focus on z_j. We first examine how the cutoff $\underline{t}_j$ changes with respect to changes in z_j:

$$\frac{\partial \underline{t}_j}{\partial z_j} \propto \frac{\partial({z_j}^{-1}\Theta(z_j)^{1-\frac{\sigma-1}{\theta}})}{\partial z_j} \propto \left[-z_j^{-1}{\Theta_j}^{1-\frac{\sigma-1}{\theta}} + \frac{\partial \Theta_j^{1-\frac{\sigma-1}{\theta}}}{\partial z_j}\right], \tag{27}$$

where:

$$\frac{\partial \Theta_j^{1-\frac{\sigma-1}{\theta}}}{\partial z_j} = \left(1-\frac{\sigma-1}{\theta}\right)\Theta_j^{-\frac{\sigma-1}{\theta}}\frac{\partial \Theta_j}{\partial \underline{t}_j}\frac{\partial \underline{t}_j}{\partial z_j}.$$

Now suppose $\frac{\partial \underline{t}_j}{\partial z_j} > 0$, then the right-hand side of equation (27) will be negative because $\frac{\partial \Theta_j}{\partial \underline{t}_j} < 0$ and $\theta > \sigma - 1$. This leads to a contradiction, which implies that $\frac{\partial \underline{t}_j}{\partial z_j} < 0$. Then it is straightforward to show that $\frac{\partial \Theta_j}{\partial z_j} > 0$. As $E(c_j) = \Theta_j^{\frac{-1}{\theta}}\Gamma(\frac{\theta+1}{\theta})$ and $E(c_j^{1-\sigma}) = \Gamma(\frac{\theta+1-\sigma}{\theta})\Theta_j^{\frac{\sigma-1}{\theta}}$, it is immediate that $\frac{\partial E(c_j)}{\partial z_j} < 0$, $\frac{\partial E(c_j^{1-\sigma})}{\partial z_j} > 0$. Finally, by the envelope theorem we know that $\frac{\partial \pi_j^O}{\partial z_j} > 0$. As A and z_j enter the function multiplicatively, we immediately know that $\frac{\partial \pi_j^O}{\partial A} > 0$, $\frac{\partial \underline{t}_j}{\partial A} < 0$. As A=$\beta LP^{\sigma-1}$and $\sigma > 1$, applying the chain rule we get $\frac{\partial \pi_j^O}{\partial P} > 0$, $\frac{\partial \underline{t}_j}{\partial P} < 0$.

Comparative statics for N Taking the derivative of equation (26) with respect to N, we obtain:

$$\frac{\partial \underline{t}_j}{\partial N} \propto \left(1-\frac{\sigma-1}{\theta}\right)\Theta_j^{-\frac{\sigma-1}{\theta}}\frac{\partial \Theta_j}{\partial N}. \tag{28}$$

From the expression of Θ_j and that $\theta > \sigma - 1$, we obtain:

$$\frac{\partial \Theta_j}{\partial N} \propto \left(\int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota) - N \underline{t}_j g_t(\underline{t}_j) \frac{\partial \underline{t}_j}{\partial N} \right). \tag{29}$$

Now suppose that $\frac{\partial \underline{t}_j}{\partial N} \leq 0$, then expression (29) implies that $\frac{\partial \Theta_j}{\partial N} > 0$. By (28), this in turn means that $\frac{\partial \underline{t}_j}{\partial N} > 0$, which is a contradiction. Therefore, $\frac{\partial \underline{t}_j}{\partial N}$ has to be positive. This also implies that $\frac{\partial \Theta_j}{\partial N} > 0$ by inspecting (28). As before, from the expressions of $E(c_j)$ and $E(c_j^{1-\sigma})$ we know immediately that $\frac{\partial E(c_j)}{\partial N} < 0$, $\frac{\partial E(c_j^{1-\sigma})}{\partial N} > 0$. Finally, by the envelope theorem it is also straightforward to show that $\frac{\partial \pi_j^O}{\partial N} > 0$.

C.2 Proof of Existence and Uniqueness

We decompose the proof of uniqueness into three parts. In the first part, we show that the aggregate price index is decreasing in N. In the second part, we prove that the FE curve is increasing in N. In the last part, we prove that the FE curve cuts AP curve once and only once from above.

Part I: The AP curve is downward sloping.

Denote $(1-\frac{1}{\sigma})^{\sigma-1}E(c(z_j)^{1-\sigma})$ by $\tilde{p}(z)^{1-\sigma}$ and let $p(t) = ((1-\frac{1}{\sigma})t)^{-1}$, the aggregate price index can be written as:

$$F_{AP} = \left(\int_{z_1}^{\bar{z}} \int_0^{\psi(z)} z\tilde{p}(z)^{1-\sigma} g(z,t) dt dz + \int_{z_1}^{\bar{z}} \int_{\psi(z)}^{\bar{t}} z p(t)^{1-\sigma} g(z,t) dt dz + \int_{z_2}^{z_1} \int_{\phi^{-1}(z)}^{\bar{t}} z p(t)^{1-\sigma} g(z,t) dt dz \right) - \frac{P^{1-\sigma}}{N}.$$

Taking the partial derivative of F_{AP} with respect to P and N and applying the Leibniz rule, we get:

$$\frac{\partial F_{AP}}{\partial P} = \int_{z_1}^{\bar{z}} \int_0^{\psi(z)} z \frac{\partial \tilde{p}(z)^{1-\sigma}}{\partial P} g(z,t) dt dz - \int_0^{\psi(z_1)} z_1 \tilde{p}(z_1)^{1-\sigma} g(t) dt \frac{\partial z_1}{\partial P} + (\sigma - 1)\frac{P^{-\sigma}}{N},$$

$$\frac{\partial F_{AP}}{\partial N} = \int_{z_1}^{\bar{z}} \int_0^{\psi(z)} z \frac{\partial \tilde{p}(z)^{1-\sigma}}{\partial N} g(z,t) dt dz - \int_0^{\psi(z_1)} z_1 \tilde{p}(z_1)^{1-\sigma} g(t) dt \frac{\partial z_1}{\partial N} + \frac{P^{1-\sigma}}{N^2}.$$

From the comparative static analyses, we know that $\frac{\partial \tilde{p}(z)^{1-\sigma}}{\partial P} = \frac{\partial E(c_j^{1-\sigma})}{\partial P} > 0$, $\frac{\partial \tilde{p}(z)^{1-\sigma}}{\partial N} = \frac{\partial E(c_j^{1-\sigma})}{\partial N} > 0$. Recall that z_1 solves $\pi_j^O(z_1) = 0$. The net profit π_j^O increases in z, P, and N, and therefore by the implicit function theorem $\frac{\partial z_1}{\partial P} < 0$, $\frac{\partial z_1}{\partial N} < 0$. Thus, for both $\frac{\partial F_{AP}}{\partial P}$ and $\frac{\partial F_{AP}}{\partial N}$, the first term is positive, the second term is negative, and the last term is positive. Therefore, $\frac{\partial F_{AP}}{\partial P} > 0$ and $\frac{\partial F_{AP}}{\partial N} > 0$. Applying the implicit function theorem we get:

$$\frac{\partial P}{\partial N} \mid_{AP} = -\frac{\partial F_{AP}/\partial N}{\partial F_{AP}/\partial P} < 0.$$

Therefore, the AP curve is downward sloping.

Part II: The FE curve is downward sloping.

Note that because firms charge constant markups, the total amount of operating profit is fixed. The free entry condition therefore can also be written as:

$$\frac{\beta L}{\sigma N}-\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}(nf+f_B+f_E)g(z,t)dtdz-\int_{z_1}^{\bar{z}}\int_{\psi(z)}^{\bar{t}}f_B g(z,t)dtdz-\int_{z_2}^{z_1}\int_{\phi^{-1}(z)}^{\bar{t}}f_B g(z,t)dtdz=\delta f_E,.$$

Let:

$$F_{FE}=\delta f_E-\frac{\beta L}{\sigma N}+\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}(nf+f_B+f_E)g(z,t)dtdz+\int_{z_1}^{\bar{z}}\int_{\psi(z)}^{\bar{t}}f_B g(z,t)dtdz+\int_{z_2}^{z_1}\int_{\phi^{-1}(z)}^{\bar{t}}f_B g(z,t)dtdz.$$

Taking the partial derivative of F_{FE} with respect to P and applying the Leibniz rule, we get:

$$\begin{aligned}\frac{\partial F_{FE}}{\partial P}&=\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial nf}{\partial P}g(z,t)dtdz-\left(\int_0^{\psi(z)}(n(z_1)f+f_B+f_E)g_t(t)dt+\int_{\psi(z_1)}^{\bar{t}}f_B g_t(t)dt\right)\frac{\partial z_1}{\partial P}\\&+\int_{\phi^{-1}(z_1)}^{\bar{t}}f_B g_t(t)dt\frac{\partial z_1}{\partial P}-\int_{\phi^{-1}(z_2)}^{\bar{t}}f_B g_t(t)dt\frac{\partial z_2}{\partial P}.\end{aligned}$$

As $\phi^{-1}(z_1)=\psi(z_1)$ and $\phi^{-1}(z_2)=\bar{t}$ always hold in equilibrium, we can simplify the above expression to:

$$\frac{\partial F_{FE}}{\partial P}=\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial nf}{\partial P}g(z,t)dtdz-\left(\int_0^{\psi(z)}(n(z_1)f+f_B+f_E)g_t(t)dt\right)\frac{\partial z_1}{\partial P}.$$

From the comparative static analyses we know that $\frac{\partial t_j}{\partial P}<0$, which implies $\frac{\partial n}{\partial P}>0$. Therefore, the first term of the above equation is positive. For the second term, recall that z_1 solves $\pi_j^O(z_1)=0$. The net profit π_j^O increases in both z and P, and therefore by the implicit function theorem $\frac{\partial z_1}{\partial P}<0$. Therefore the second term is negative, and hence $\frac{\partial F_{FE}}{\partial P}>0$.

Similarly, we can combine the net profit of outsourced firms and contracted manufacturers, and write F_{FE} as the following:

$$\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}(\pi^O(z)+\theta Az_jE(c_j^{1-\sigma}))g(z,t)dtdz+\int_{z_1}^{\bar{z}}\int_{\psi(z)}^{\bar{t}}\pi^I(z)g(z,t)dtdz+\int_{z_2}^{z_1}\int_{\phi^{-1}(z)}^{\bar{t}}\pi^I(z)g(z,t)dtdz=\delta f_E.$$

Taking the partial derivative of F_{FE} with respect to N and applying the Leibniz rule:

$$\frac{\partial F_{FE}}{\partial N}=\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial\pi^O(z)+\partial\theta Az_jE(c_j^{1-\sigma})}{\partial N}g(z,t)dtdz-\left(\int_0^{\psi(z)}(\pi^O(z)+\theta Az_jE(c_j^{1-\sigma}))g_t(t)dt\right)\frac{\partial z_1}{\partial N}.$$

From our analysis before we know that the first term is positive, as $\frac{\partial\pi^O(z)}{\partial N}>0,\frac{\partial AE(c_j^{1-\sigma})}{\partial N}>0$. We proved already that z_1 decreases in N, so the second term is negative. Therefore we conclude $\frac{\partial F_{FE}}{\partial N}>0$ and:

$$\frac{\partial P}{\partial N}\mid_{FE}=-\frac{\partial F_{FE}/\partial N}{\partial F_{FE}/\partial P}<0.$$

That is, the FE curve is downward sloping.

Part III: The AP curve is flatter compared to the FE curve for any given (P, N), and FE curve cuts AP curve once from above.

Note that since both the FE and AP curves are downward sloping, to prove that the AP curve is flatter compared to the FE curve, it is equivalent to show:

$$\frac{\partial F_{AP}/\partial N}{\partial F_{AP}/\partial P} < \frac{\partial F_{FE}/\partial N}{\partial F_{FE}/\partial P}.$$

After some algebra one can show that the expression of $\frac{\frac{\partial F_{FE}}{\partial N}}{\frac{\partial F_{FE}}{\partial P}}$ can be written as the following:

$$\frac{\frac{\partial F_{FE}}{\partial N}}{\frac{\partial F_{FE}}{\partial P}} = \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial nf}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial nf}{\partial P} g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial P}}.$$

From Part I of the analysis, we know:

$$\frac{\frac{\partial F_{AP}}{\partial N}}{\frac{\partial F_{AP}}{\partial P}} = \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} z\frac{\partial \tilde{p}(z)^{1-\sigma}}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)} z_1\tilde{p}(z_1)^{1-\sigma}g(t)dt\frac{\partial z_1}{\partial N} + \frac{P^{1-\sigma}}{N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} z\frac{\partial \tilde{p}(z)^{1-\sigma}}{\partial P} g(z,t)dtdz - \int_0^{\psi(z_1)} z_1\tilde{p}(z_1)^{1-\sigma}g(t)dt\frac{\partial z_1}{\partial P} + (\sigma-1)\frac{P^{-\sigma}}{N}},$$

which can be rewritten as:

$$\frac{\frac{\partial F_{AP}}{\partial N}}{\frac{\partial F_{AP}}{\partial P}} = \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} z\frac{\partial \tilde{p}(z)^{1-\sigma}}{P^{1-\sigma}\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)} \frac{z_1\tilde{p}(z_1)^{1-\sigma}}{P^{1-\sigma}}g(t)dt\frac{\partial z_1}{\partial N} + \frac{1}{N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} z\frac{\partial \tilde{p}(z)^{1-\sigma}}{P^{1-\sigma}\partial P} g(z,t)dtdz - \int_0^{\psi(z_1)} \frac{z_1\tilde{p}(z_1)^{1-\sigma}}{P^{1-\sigma}}g(t)dt\frac{\partial z_1}{\partial P} + (\sigma-1)\frac{1}{PN}}$$

$$= \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial(\pi^O(z)+nf)}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)}(\pi^O(z_1) + n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{(1-\gamma)\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial(\pi^O(z)+nf)}{\partial P} g(z,t)dtdz - \int_0^{\psi(z_1)}(\pi^O(z) + n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial P} + (\sigma-1)\frac{(1-\gamma)\beta L}{\sigma PN}}.$$

Note that $\pi^O(z_1) = 0$ in equilibrium, therefore:

$$\frac{\frac{\partial F_{AP}}{\partial N}}{\frac{\partial F_{AP}}{\partial P}} = \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial(\pi^O(z)+nf)}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{(1-\gamma)\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial(\pi^O(z)+nf)}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)}(nf + f_B)g(t)dt\frac{\partial z_1}{\partial P} + (\sigma-1)\frac{(1-\gamma)\beta L}{\sigma PN}}$$

$$< \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial(\pi^O(z)+nf)}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial(\pi^O(z)+nf)}{\partial P} g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial P} + (\sigma-1)\frac{(1-\gamma)\beta L}{\sigma PN}},$$

$$< \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial\pi^O(z)}{\partial N} g(z,t)dtdz + \int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial nf}{\partial N} g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial\pi^O(z)}{\partial P} g(z,t)dtdz + \int_{z_1}^{\bar{z}}\int_0^{\psi(z)} \frac{\partial nf}{\partial P} - \int_0^{\psi(z_1)}(n(z_1)f + f_B)g(t)dt\frac{\partial z_1}{\partial P}}.$$

Applying the envelope theorem, we can show that $\frac{\partial\pi^O(z)}{\partial N} = \frac{P}{\theta N}\frac{\partial\pi^O(z)}{\partial P} - \frac{nf}{N}$. As we imposed the regularity assumption that L is sufficiently large so that $\frac{P}{\theta N} < 1$ always holds, $\frac{\partial \pi^O(z)}{\partial N} < \frac{\partial \pi^O(z)}{\partial P}$. Therefore:

$$\frac{\frac{\partial F_{AP}}{\partial N}}{\frac{\partial F_{AP}}{\partial P}} < \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial \pi^O(z)}{\partial P}g(z,t)dtdz + \int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial nf}{\partial N}g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f+f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial \pi^O(z)}{\partial P}g(z,t)dtdz + \int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial nf}{\partial P} - \int_0^{\psi(z_1)}(n(z_1)f+f_B)g(t)dt\frac{\partial z_1}{\partial P}}.$$

$$< \frac{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial nf}{\partial N}g(z,t)dtdz - \int_0^{\psi(z_1)}(n(z_1)f+f_B)g(t)dt\frac{\partial z_1}{\partial N} + \frac{\beta L}{\sigma N^2}}{\int_{z_1}^{\bar{z}}\int_0^{\psi(z)}\frac{\partial nf}{\partial P} - \int_0^{\psi(z_1)}(n(z_1)f+f_B)g(t)dt\frac{\partial z_1}{\partial P}} = \frac{\frac{\partial F_{FE}}{\partial N}}{\frac{\partial F_{FE}}{\partial P}}.$$

Hence we conclude the proof of Part III. Finally, note that when $N \to 0$, $F_{FE} > 0$ and $N \to \infty$, $F_{FE} < 0$ regardless of the value of P. Therefore for $F_{FE} = 0$, the range of N is bounded both from below and above. Thus, there must be some N' that for $N < N'$, $P_{AP} <P_{FE}$ and some N'' that for $N > N''$, $P_{AP} >P_{FE}$. Combined with the result of Part III, we can conclude that the equilibrium exists and is unique.

C.3 Proof that as when τ_t decreases, firms with relatively higher labor productivity will bring their blueprints to production

We decompose the proof of our model's testable prediction into three parts. In the first part, we show that when τ_M decreases, net profits from final good production increases in z_j and decreases in t_j. In the second part, we show that conditional on employment, firms' labor productivity increases as z increases. In the third part, we prove that conditional on employment, we get $\frac{\partial^2 v_j^B}{\partial \tau_M \partial LP_j} > 0$.

Part I: When τ_M decreases, net profits from final good production increases in z_j and decreases in t_j.

We decompose the proof of our model's testable prediction into two steps. In the first step, we show that when τ_t decreases, net profits from outsourcing increases in z_j. Then we show that an outsourced firm' labor productivity increases as z increases, hence $\frac{\partial^2 \pi_j^O}{\partial \tau_t \partial LP_j} > 0$.

Define changes due to a reduction of τ_t in N and P as dN and dP, respectively. By the envelope theorem, the change in profits for outsoucred downstream firm j equals:

$$d\pi_j^O = \frac{\partial \pi_j^O}{\partial N}dN + \frac{\partial \pi_j^O}{\partial P}dP = \frac{\sigma-1}{\theta}\frac{(1-\gamma)Az_jE(c_j^{1-\sigma})}{N}dN - \frac{nf}{N}dN + (\sigma-1)(1-\gamma)Az_jE(c_j^{1-\sigma})\frac{dP}{P}.$$

Recall that when firms optimize their sourcing decisions, we have that $n_j f = \frac{\sigma-1}{\theta}\frac{\int_{\underline{t}_j}^{\bar{t}} \underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)} * (1-$

$\gamma)Az_jE(c_j^{1-\sigma})$. Hence, we can rewrite $d\pi_j^O$ as:

$$
\begin{aligned}
d\pi_j^O &= \frac{\sigma-1}{\theta}(1-\gamma)Az_jE(c_j^{1-\sigma})(1-\frac{\int_{\underline{t}_j}^{\bar{t}}\underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)})\frac{dN}{N} + (\sigma-1)(1-\gamma)Az_jE(c_j^{1-\sigma})\frac{dP}{P} \\
&\propto z_jE(c_j^{1-\sigma})\left((1-\frac{\int_{\underline{t}_j}^{\bar{t}}\underline{t}_j dG_t(\iota)}{\int_{\underline{t}_j}^{\bar{t}} t dG_t(\iota)})\frac{dN}{N} + \theta\frac{dP}{P}\right). \qquad (30)
\end{aligned}
$$

Therefore:

$$
d\pi_j^O \propto z_jE(c_j^{1-\sigma})\left(\frac{N\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota)}{\Theta_j} + \theta\frac{\partial dlnP}{\partial dlnN}\right). \qquad (31)
$$

The term $z_jE(c_j^{1-\sigma})$ increases in z_j from the comparative statics. Let $F_{dv} \equiv \frac{N\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota)}{\Theta_j}$. We now have:

$$
\frac{\partial F_{dv}}{\partial z_j} = \frac{N}{\Theta_j^2}\int_{\underline{t}_j}^{\bar{t}}(-1)dG_t(\iota)\frac{\partial \underline{t}_j}{\partial z_j}\Theta_j - \frac{N}{\Theta_j^2}\frac{\partial\Theta_j}{\partial z_j}\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota).
$$

Note that $\frac{\partial\Theta_j}{\partial z_j} = \frac{\partial\Theta_j}{\partial\underline{t}_j}\frac{\partial\underline{t}_j}{\partial z_j}$. As $\frac{\partial\underline{t}_j}{\partial z_j} < 0$, the term $\frac{\partial F_{dv}}{\partial z_j} > 0$, therefore $d\pi_j^O$ must increase in z_j. Then it is straightforward to show that conditional on outsourcing, firms' labor productivity increases in z_j. If a firm self-selects to become a downstream firm, its log labor productivity is given by:

$$
LP(z_j) = ln\left(\frac{\pi^O(z_j)}{l(z_j)}+1\right) = ln\left(\frac{\pi^O(z_j)}{f_o}+1\right),
$$

since we assumed that f_o is paid in terms of unit labor, i.e., this is the only employment downstream outsourced firms own. It is immediate that $LP(z_j)$ increases in z_j as $\pi^O(z_j)$ increases in z_j as well. Finally, consider two outsourced firms with $LP_j > LP_{j'}$, then we know that $z_j > z_{j'}$ must hold. As we already proved that $\frac{\partial^2\pi_j^O}{\partial\tau_T\partial z_j} > 0$, then it is immediate that $\frac{\partial^2\pi_j^O}{\partial\tau_T\partial LP_j} > 0$, or equivalently the downstream firm j faces a relative increase in net profits compared to firm j'.

Appendix Tables and Figures

Table A.1: List of Products in the 2018 Customs Sample

HS code	Product specification
39232100	Ethylene polymer bags and bags (for transport or packaging of goods)
40112000	Tires for passenger cars or trucks
42022200	Handbags made of plastic or textile materials (with or without straps)
54075200	Dyed other polyester textured filament woven fabric
61099090	T-shirts
61102000	Pullovers
62019390	Cold weather clothes
62034290	Trousers, breeches
62043200	Cotton-made women's tops
63014000	Blankets and traveling rugs of synthetic fibers
73239300	Table, kitchen or other household articles and parts made of stainless steel
84151021	Air conditioners
84181020	Refrigerators (200 to 500 liters)
84183029	Cabinet freezers (temperature>-40 degree Celsius)
84714140	Microcomputers
84715040	Other microprocessor processing components
84717010	Hard disk drivers for automatic data processing machines
84717030	Optical drive for automatic data processing equipment
85030090	Motor stator and other motor (set) parts
85164000	Electric irons
85165000	Microwaves
85171100	Cordless telephones
85171210	GSM & CDMA digital wireless phones
85177060	Laser transceiver modules for optical communication equipment
85183000	Headphones
85219012	DVD players
85299090	High frequency tuner for satellite television reception and other purposes
85340090	Printed circuit with four layers or less
85366900	Plugs and sockets with voltage $\leq$ 1000 volts
85414020	Solar batteries
85416000	Assembled piezoelectric crystals
87120030	Mountain bikes
90138030	LCD panels
94051000	Chandeliers

Notes: This table lists the 34 products used in the 2018 customs sample. The original customs data is at the 10-digit HS (HS10) level; we report the product specification at the 8-digit level (HS8) to save space. Even at the HS8 level, the product specification is highly disaggregated and clearly defined. The English product specifications are translated from `http://www.i5a6.com/hscode/`.

Table A.2: Products in the Estimation Sample

Product code	Product name	Obs.
01567	Rice	3,777
01623	Wheat flour	6,373
01765	Refined edible vegetable oil	5,039
01994	Fresh, frozen meat	2,493
02079	Aquatic products	2,311
02305	Mixed feed	8,797
02517	Cans	2,227
03796	Yarn	9,675
04166	Printed and dyed cloth	4,206
05036	Silk	2,802
05098	Silk products	4,096
05883	Light leather	2,032
05901	Leather shoes	7,322
06982	Machine made paper	2,865
07307	Machine made cardboard	2,437
07432	Paper products	4,198
08364	Toys	2,333
13989	Paint	2,672
16866	Chemical raw material	2,723
20122	Chinese-patented drugs	5,280
21696	Plastic products	16,323
22108	Cement	4,477
22559	Folded standard brick	2,432
23245	Glass products	3,045
23325	Ceramics	3,922
23936	Refractory products	2,437
26035	Pig iron	3,775
26719	Ferroalloy	2,949
27092	Copper (copper processed material)	3,027
28677	Aluminum	2,128
31438	Stainless steel products	2,608
31872	Pump (liquid pump)	3,025
31969	Bearings	2,868
32426	Casting	3,974
41305	Power supply cable	2,052
44497	Sub-assemblies & parts	3,132

Notes: This table lists the 36 products used in our $TFPQ$ estimation. This set is a subsample of the 693 manufacturing products in the dataset, selected according to the criteria described in Appendix A. The English product specifications are translated from `http://www.i5a6.com/hscode/`.

Table A.3: Mixed Exporter Premia - Intensive Margin

(a) All exporters	proc.share$_{it}$		Obs.
(1) ln(*empl.*)$_{it}$	0.260***	(0.067)	66,326
(2) ln(*labor prod.*)$_{it}$	0.034	(0.078)	62,505
(3) $TFPR_{it}$	-0.030	(0.056)	2,697
(4) $TFPQ_{it}$	0.024	(0.016)	2,697
(5) ln(*R&D exp.*)$_{it}$	-0.431***	(0.083)	66,326
(6) ln(*advert. exp.*)$_{it}$	-0.613***	(0.105)	60,645
(7) ln(*trademarks*)$_{it}$	-0.357***	(0.053)	66,326
(b) Excl. foreign firms	proc.share$_{it}$		Obs.
(1) ln(*empl.*)$_{it}$	0.189***	(0.060)	48,869
(2) ln(*labor prod.*)$_{it}$	0.095	(0.085)	46,032
(3) $TFPR_{it}$	-0.026	(0.065)	1,969
(4) $TFPQ_{it}$	0.023	(0.019)	1,969
(5) ln(*R&D exp.*)$_{it}$	-0.463***	(0.100)	48,869
(6) ln(*advert. exp.*)$_{it}$	-0.640***	(0.123)	44,773
(7) ln(*trademarks*)$_{it}$	-0.385***	(0.062)	48,869

Notes: Each row is a separate OLS regression of the dependent variable shown in column 1 on proc. share$_{it}$: the share of processing of mixed firm i in year t. ln($R\&D\ exp.$)$_{it}$, ln($advert.\ exp.$)$_{it}$, and ln($trademarks$)$_{it}$ are calculated by $\ln(x+1)$ to avoid dropping zeros. $TFPR_{it}$ and $TFPQ_{it}$ refer to TFP calculated using revenue and quantity data respectively (see the text for details). Rows 1-2 and 5-7 include sector-year fixed effects, and all except those in the first row control for firm size. Rows 3-4 focus on single-product producers only and thus include product-year fixed effects. Standard errors clustered by 2-digit CIC industries are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

Table A.4: Comparisons of Firms

Sample	(1) All <$10m processors	(2) All >$10m processors	(3) Difference	(4) $9-10m processors	(5) $10-11m processors	(6) Difference
Mixed	0.63	0.67	-0.04***	0.62	0.62	-0.00
Importer	0.73	0.76	-0.03***	0.75	0.78	-0.03
Exiter	0.07	0.02	0.05***	0.02	0.02	-0.00
Entrant	0.11	0.04	0.07***	0.04	0.05	-0.01
Foreign	0.49	0.47	0.02***	0.51	0.45	0.06**
SOE	0.12	0.20	-0.08***	0.16	0.13	0.02
Proc. share of exports	0.70	0.86	-0.17***	0.84	0.89	-0.05***
Avg. log annual growth	0.05	0.14	-0.09***	0.12	0.18	-0.06
$\ln(empl.)$	5.43	6.82	-1.38***	6.17	6.22	-0.06
$\ln(capital)$	8.83	10.57	-1.74***	9.92	9.83	0.09
Obs.	189,195	8,818		1,019	736	

Notes: This table reports balancing checks between the treatment and control groups. Columns 1 and 2 represent the means of the variables for exporters that are below and above the $10m threshold respectively (entire sample). Columns 4 and 5 represent the means of the variables for exporters that have $9-10m and $10-11m processing exports respectively (restricted sample). Columns 3 and 6 show the differences in the means across the groups. The number of observations reported in the last row corresponds to the variable in the first row, and might deviate across variables depending on data availability. The electronics sector is excluded due to its lower $5m threshold. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

Table A.5: Paperless Trade and Processing Exports

Dep. var.: ln(proc. exp.)$_{icst}$	(1) Benchmark	(2) Leads & lags	(3) Proc. share	(4) No foreign firms	(5) Processing intensity
OS_{ict-1}	0.277**	0.281**	0.277**	0.454***	0.028*
	(0.126)	(0.119)	(0.112)	(0.101)	(0.015)
OS_{ict+1}		0.033			
		(0.161)			
Proc. share$_{ict-1}$			1.168***		
			(0.192)		
Firm FE	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes
Prefecture-year FE	Yes	Yes	Yes	Yes	Yes
Obs.	1,718	1,452	1,418	779	1,718
R^2	0.62	0.65	0.68	0.65	0.90

Notes: This table reports the results of running specification (14). OS_{ict-1} indicates the implementation of the pilot paperless processing trade programme in prefecture c in year $t-1$ for firm i (i.e., Class A firms). In column 5, the dependent variable is processing intensity captured by Proc. share$_{ict}$. Sector s refers to the top (core) HS2 of each firm. Standard errors clustered at the prefecture and sector level are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

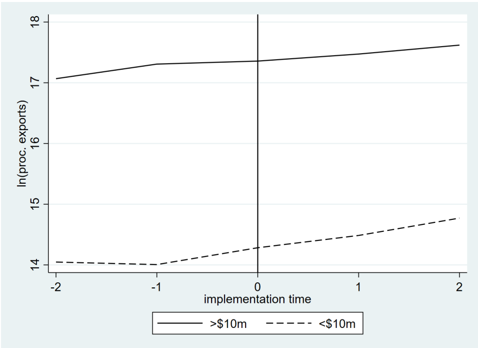
Table A.6: Paperless Trade and Processing Exports - Robustness Checks

Dep. var.: ln(proc. exp.)$_{icst}$	(1) Entry & exit	(2) FD	(3) bw: \$9.5-10.5m	(4) bw: \$8.5-11.5m	(5) Always exporters	(6) No SOEs	(7) Mixed only (ordinary)	(8) Mixed only (proc.)	(9) False threshold (\$9m)	(10) False threshold (\$11m)
OS_{ict-1}	0.238*	0.140***	0.207*	0.179**	0.239**	0.241**	-0.256	0.354*	0.019	-0.056
	(0.117)	(0.045)	(0.117)	(0.072)	(0.101)	(0.114)	(0.313)	(0.197)	(0.072)	(0.092)
Entrant$_{it}$	-1.341***									
	(0.172)									
Exiter$_{it}$	-1.125***									
	(0.212)									
Firm FE	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,254	1,417	716	2,690	1,120	1,414	936	936	2,223	1,379
R^2	0.74	0.37	0.75	0.58	0.64	0.64	0.87	0.67	0.59	0.62

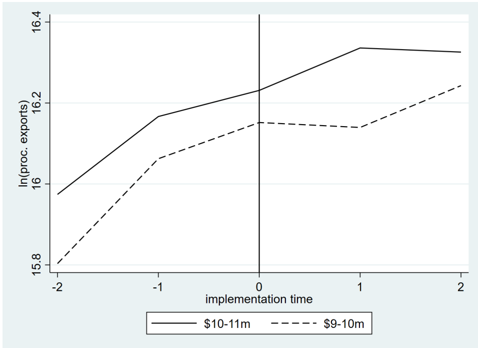
Notes: This table reports further robustness checks for the results in Table A.5. OS_{ict-1} indicates the implementation of the pilot paperless processing trade programme in prefecture c in year $t-1$ for firm i (i.e., Class A firms). Sector s refers to the top (core) HS2 of each firm. In column 2, we use a first-difference (FD) specification. In columns 3 and 4, we change the bandwidth to \$9.5-10.5m and \$8.5-11.5m respectively. In columns 9 and 10, we do falsification analyses by setting the threshold to \$9m and \$11m, and the bandwidth (bw) to \$8-10m and \$10-12m respectively. Standard errors clustered at the prefecture and sector level are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

Figure A.1: Processing Export Trends

(a) All firms



(b) $9-11m firms



Notes: The figure plots the level of processing exports for all exporters in panel (a) and exporters that had $9-11m worth of processing exports in the year prior to policy adoption in panel (b). Implementation time 0 indicates the year the prefecture's customs authority adopted the pilot paperless processing trade program.