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

Zhiyuan Chen† Aksel Erbahar‡ Yuan Zi§

April 2021

Abstract

This paper proposes that processing trade, which played an important role in China's export miracle, not only leads goods to be "Made in China," but also "Created in China." Using unique transaction-level trade data on firms' branding information, we document four main findings. First, there exists a significant share of exporters that engages in both ordinary and processing export activities, and they exhibit superior performance in various margins. Second, even within firms, there is a tight link between firms' export mode choice and brand ownership—own branded products are typically exported under ordinary trade regime while products under other firms' brands are exported under processing trade regime. Third, there is a price premium associated with own-branded products. Fourth, Chinese firms intensify their branding activities when faced with favorable processing trade policies upstream. To rationalize these findings, we present a simple theoretical framework where firms with multi-attributes endogenously determine their specialization 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, 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

§Department of Economics, University of Zurich, Switzerland, and CEPR, *email*: yuan.zi@econ.uzh.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 supplied relatively low value-added tasks to 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. 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 "super processors" are 'mixed' not 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. But 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 cost 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.

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. Firms compete monopolistically in the final goods market and à la Bertrand in the tasks market, and thus charge positive markups in both stages of production. In equilibrium, firms with good blueprints charge higher markups and self-select to become an ordinary exporter, firms with higher manufacturing ability 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, the model rationalizes the observed performance rankings at various margins between mixed, pure ordinary, and pure processing exporters.

Since we endogenise firms' specialization within a production network, the mass of potential suppliers is no longer exogenous. Thus, our framework differs from the existing sourcing models such as the ones developed by Antràs et al. (2017) and Bernard et al. (2019), and generates a new positive externality of processing policy: facilitating processing trade raises the ex-ante expected profits from task production, leading to a greater mass of potential suppliers in equilibrium. This in turn benefits sourcing firms with good ideas. Our model thus predicts that facilitating processing exports stimulates downstream firms' branding activity, especially more so for firms with higher measured revenue productivity as they source a greater share of tasks from suppliers, which is in line with our empirical evidence. Overall, 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.

This paper 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

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

China (Fernandes and Tang, 2015; Yu, 2015; Dai et al., 2016; Kee and Tang, 2016; Li et al., 2018).[2] 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 task 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 work 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 implication of processing policy, and thus also complement the recent work that studies 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 and Wang (2021), and Deng (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. (2019), 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.

[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 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 and Wang (2021) introduce increasing returns to scale in input production in a similar framework and quantify the processing-trade-induced Dutch disease. Deng (2021) further quantifies the welfare implications of processing policy with the presence of learning-by-processing.

[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. (2019), 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.

Finally, our paper connects to the literature that studies firms with multiple heterogeneities, including but not limited to Antràs and Helpman (2004), Hallak and Sivadasan (2013), Harrigan and Reshef (2015), Bernard et al. (2018), and Huang et al. (2021).[6] None of these papers, however, emphasize the role of heterogeneities that enable firms to self-select into upstream versus downstream segments of the production network. We show that this intuitive set-up 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 exporter's 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

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 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

[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. Bernard et al. (2018) study how productivity and relationship capability can explain the matching between buyers and sellers in Belgium. Huang et al. (2021) adopt a similar framework to 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, and became effective on January 1, 2018.

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.

3 Stylized Facts

3.1 Mixed Exporters in China

We begin by unpacking the "black box" of mixed exporters in this section. Mixed exporters are defined as firms that engage in both processing and ordinary exports. The customs data shows 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

[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.

hand, made up 24% and 19% of exports in 2005 respectively.[14] 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%).

We present firm-level statistics for mixed exporters in Table 1, with the full sample in panel (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 (hence we label them as "super processors"). 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. This result indicates that many "superstar" firms are mixed exporters.

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 1 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'

[14]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 intermediaires and wholesalers, which made up 18% of exports in 2005. These figures are similar to those reported by Dai et al. (2016).

Table 1: 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.

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. Moreover, both types of models would predict that mixed firm attributes lie between that of processing and ordinary firms. However, as shown in the next subsection, this is not what we find in the data.

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. Our exercise in this section is similar to Dai et al. (2016), but while the previous literature mainly analyzed the differences between two types of exporters or non-exporters, we focus on the three types of exporters: mixed exporters, ordinary exporters, and pure processors. 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 (29 clusters). Each row of Table 2 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 2 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 2 panel (a) row 2 shows that mixed firms have 14% higher labor productivity than pure ordinary firms, whereas pure processors have 22% lower labor productivity than pure ordinary firms.[15] Row 3 shows that the ranking we obtained based on labor productivity remains when we consider $TFPR$ calculated using the Olley-Pakes (1996) methodology.[16]

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

[15] 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.

[16] Our results are robust to using the Levinsohn-Petrin (2003) methodology.

Table 2: 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
(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

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($R\&D\ exp.$)$_{it}$ and ln($advert.\ exp.$)$_{it}$ 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 5-6 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 (29 clusters) are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

the list of products can be found in Appendix A and Table A.1 respectively.[17] 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 2 panel (a)). In addition, mixed exporters have the highest physical productivity on average (though not statistically significantly different from that of pure processors).[18]

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.

Existing theoretical frameworks would predict that mixed firm characteristics lie between that

[17] 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.

[18] 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.

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 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).[19]

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 3, 12.4%, 56.4%, and 32.7% of export value are due to transactions that have no brand, foreign brand, and domestic brand respectively. Ordinary and processing exports account for 27% and 43% of total exports in this sample. While processing transactions are typically viewed as local manufacturers supplying customized tasks to their buyers (Manova and Yu, 2016), our data enables us to confirm this conjecture empirically: Table 3 shows that 52% of ordinary exports in the customs sample consists of goods with Chinese domestic brands, while 84% of processing exports consists of foreign branded products.

More formally, 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.,

[19] 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, supplies tasks to brands such as *De'Longhi*, *General Electric*, and *Sanyo* alongside exporting its own branded microwaves and air conditioners.

Table 3: 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%
Other exports	3.2%	92.8%	4.0%
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.

FDI policy), and ϵ_{ifhc} is the error term. We cluster standard errors at the firm level. Table 4 column 1 shows that processing transactions are 13 percentage points less likely to involve products with 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.[20] 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:

[20] 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 4: 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.

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 firm-product-destination, processing exports should have a lower unit value, which contradicts our finding in Table 4. If transfer pricing is driving the results (i.e., processing exporters artificially lowering 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 2 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 2 panel (a) rows 5 and 6 reveal that R&D and advertisement expenditures 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. This is in line with anecdotal evidence that pure processors tend to specialize in providing specific tasks for other firms, and thus do not need to invest in R&D or spend on advertisement, which are ultimately done by their customers. In panel (b) rows 5 and 6, we exclude foreign firms since the majority of their R&D and advertising expenses are likely to be done in their headquarter-countries, and thus are not perfectly observed in our data—the results are similar.[21]

[21] A closer look at the AIS data confirms that foreign-owned firms' China operations are significantly less R&D- and advertising-intensive when compared to Chinese exporters.

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 and gives us a clean identification, 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 processing promoting policy.

4.1 China's "Paperless" Processing Trade

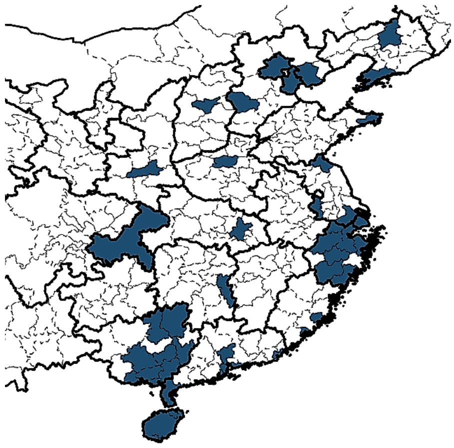
We first present the shock and the empirical context. China's customs authorities closely monitor the supply chain for processing trade because of special duty drawbacks granted to processing exporters. Thus, to organize processing trade, firms have to fill in grueling paperwork that details their financial 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."*[22]

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.[23] This policy experiment had a staggered introduction

[22]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.

[23]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

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.

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 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

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.

checks in Table A.3 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.3 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.4 column 1 suggests that the pilot program increased firm-level processing exports by around 28%. Additional estimations in Tables A.4 and A.5 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, 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 making tasks, 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., 2019), 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 share" for each prefecture-sector-year (cst):

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

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.[24] After adjusting for the one-to-many and many-to-many matches, we end up working with a slightly more aggregated set of 74 IO industries. For prefecture-sector-years with no processing exports, we set the treated processing share to zero. This share, which proxies for the intensity of the processing cost shock for each prefecture-sector-year, ranges from 0% to 100% with a mean of 8% (standard deviation: 23%).[25]

Then, we create a time-varying input shock as follows:

$$\text{Input shock}_{cnt} = \sum_s \omega_{ns} * \text{Treated processing share}_{cst},$$

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

[25]Our results are qualitatively similar if we define the shock to be simply the level of affected processing exports (the numerator of the Treated processing share$_{cst}$). These results are available on request.

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:

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

where Y_{icnt} is the number of effective trademarks a firm has,[26] Productive$_i$ indicates whether the firm's initial log labor productivity is above the median value. 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 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.[27] 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.[28]

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 5 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 0.428, significant at the 1% level. The sum of the main and interaction coefficients, 0.276, is also significant at the 1% level, indicating that a one standard deviation (0.033) increase in Input shock$_{cnt}$ raises the number of trademarks of a productive firm by 0.009 (0.276×0.033), which is 1% 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, $\ln{(labor\ prod.)}_i$, and the result stays robust.

In column 4 of Table 5, we directly control for Treated processing share$_{cnt}$ 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 barely changes when compared to column 2. Column 5 excludes SOEs

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

[27]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.

[28]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 5: 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
Input shock$_{cnt-1}$	0.084	-0.152**	0.047	-0.164**	-0.136**	-0.792**	-0.053	-0.099*
	(0.084)	(0.073)	(0.081)	(0.072)	(0.062)	(0.305)	(0.069)	(0.056)
× Productive$_i$		0.428***		0.437***	0.361***	1.368***	0.136*	0.231***
		(0.087)		(0.084)	(0.091)	(0.311)	(0.075)	(0.050)
× ln (*labor prod.*)$_i$			0.213***					
			(0.055)					
Treated processing share$_{cnt-1}$				-0.010				
				(0.009)				
× Productive$_i$				0.007				
				(0.012)				
ln (*empl.*)$_{it}$	0.102***	0.101***	0.101***	0.101***	0.097***	0.275***	0.042***	0.058***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.016)	(0.004)	(0.003)
ln (*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.

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 number of employees and the capital stock have a positive and significant effect on trademarks in all regressions, as expected. Overall, results in Table 5 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 Model of Multi-Attributes Firms and Endogenous Sourcing

In this section, we provide a parsimonious model of multi-attributes firms organizing production and innovation activities. We show that the model not only rationalizes our empirical findings, but also provides new insights on the welfare implications of processing trade polices.

Our model generalizes the work of Antràs et al. (2017) in two important yet intuitive ways. First, instead of taking firms' position at the upstream or downstream of the production process as fixed, we let it be endogenously determined. Second, we introduce imperfect competition in the tasks market, so that firms in both stages of production could charge a positive markup. We focus on intuitions in the paper and relegate formal proofs to Appendix C.

5.1 Environment

We consider an economy where consumer preferences are Cobb-Douglas over two sectors. The numeraire sector produces a homogeneous product with one unit of labor, while the other sector produces differentiated products and is the focus of our analysis. An exogenous fraction β of income is spent on differentiated products. Preferences across differentiated products exhibit CES, with the elasticity of substitution being $\sigma > 1$.

There is a continuum of firms; each owns a blueprint to produce a single differentiated variety. Producing a variety requires the assembly of a bundle of tasks $t \in [0, 1]$ under a CES production function with an elasticity of substitution $\rho > 1$. The quality of the blueprint owned by firm j is denoted by z_j, which governs the mapping from the task bundle to final good production: a higher z_j indicates that firm j is more productive in producing the final good.

Task production requires only labor, which is inelastically supplied. All tasks are blueprint-specific, and firm j's efficiency in producing a task is drawn from a Fréchet distribution with a firm-specific level parameter t_j and a shape parameter θ, with $\theta > \sigma - 1$. Here, t_j governs the firm's average manufacturing ability in producing tasks, and θ the (inverse) dispersion of its manufacturing efficiency across tasks.

Firms differ in their blueprint quality z, which indicates how good their *brand* is, and manufacturing ability t, which determines how good they are at *production.* In the rest of the paper, we refer to firms with high z as *firms with good blueprints*, and firms with high t as *firms with high manufacturing ability.* Also, we refer to firms that bring their blueprints to production as *blueprint (or final good) producers*, and firms that only supply tasks to others as *task producers.*

As blueprint holders, firms organize their production for the bundle of tasks. For each task, they can either produce it in-house or source it from other firms. Analogously, as task suppliers, firms can produce both for their own and other firms' final goods. A firm observes the average manufacturing ability of a supplier, but needs to pay a fixed cost f to establish a production relationship and discover the supplier's actual efficiency in producing tasks tailored for its blueprint.

We assume Bertrand competition in task production à la Bernard et al. (2003) (BEJK hereafter). Under this setting, even firms that do not bring their blueprint to production can earn positive profits by supplying tasks for other firms. Allowing for positive profits in task production is essential for our spillover analysis of processing policy, but the rest of our results hold if we assume perfect competition in the tasks market instead.

There is an unbounded pool of prospective entrants who learn about their blueprint quality and manufacturing ability after incurring a fixed entry cost f_E, measured in homogeneous inputs. We let z and t be drawn independently from two distributions $g_z(z)$ and $g_t(t)$ with support in $(0, \bar{z}]$ and $(0, \bar{t}]$, respectively.[29] Once firms make their draws, they decide to (i) exit, (ii) engage in blueprint production, (iii) participate in task production, or (iv) do both (ii) and (iii). Being active in blueprint production requires an additional fixed cost f_B. An operating firm faces a constant probability δ of an adverse shock that would force it to exit every period.

International trade is costly. The differentiated sector is subject to iceberg trade costs such that $\tau_B, \tau_T > 1$ units are required to be shipped for one unit of final goods and tasks, respectively, to reach the foreign destination (denoted with asterisk). Exporting final goods also requires a fixed cost f_X. The homogeneous sector is assumed to be freely traded.

5.2 Analysis of the Equilibrium

Optimal Sourcing

By assuming that manufacturing ability t varies across firms while the relationship-specific investment f does not, we simplify the firm's optimal sourcing decision to be choosing the least

[29] Additionally, we assume that $g_t'(t) < 0$, $\lim_{t \to \bar{t}} g(t) = 0$, and $-tg_t'(t) < g_t(t)$. The latter assumption can also be written as $-\frac{\partial \ln g(t)}{\partial \ln(t)} < 1$, which guarantees that the marginal reduction in firms' marginal cost decreases as they reach less efficient suppliers.

productive supplier as in Bernard et al. (2019). Conditional on firm j being connected with i, the probability that i is the lowest-cost supplier to j for a particular task is:

$$\lambda_{ij} \equiv \lambda(z_j, t_j, t_i) = \frac{t_i}{\Theta(z_j, t_j)}, \tag{4}$$

where $\Theta(z_j, t_j) = t_j + N \int_{\underline{t}_j}^{\bar{t}} s dG_s(\iota)$ measures firm j's "sourcing capacity." The variable $\underline{t}_j$ is the least productive supplier that firm j sources from, and N is the endogenously determined mass of entrants.[30] As in BEJK, the price of the task bundle used by firm j is given by $P_j^T = \Theta_j^{\frac{-1}{\theta}} \gamma$,[31] hence the firm's marginal cost of producing its own final good is $c_j = \frac{P_j^T}{z_j}$. Intuitively, conditional on firm j's own manufacturing ability t_j, sourcing from a larger number of suppliers lowers the marginal cost. Conditional on the number of suppliers, a firm with manufacturing ability produces at a lower marginal cost.

Final good producers choose the set of suppliers to maximize their profits from final good production. This yields the optimal cutoff $\underline{t}_j$ that satisfies:

$$\underline{t}_j \equiv \underline{t}(z_j, t_j) = f \left(A k_1 z_j^{\sigma-1} \right)^{-1} \Theta(z_j, t_j)^{1-\frac{\sigma-1}{\theta}} \frac{\theta}{\sigma - 1}. \tag{5}$$

where $A = \beta L P^{\sigma-1}$ is the demand shifter and k_1 is a constant.[32] The measure of suppliers $n(z_j, t_j)$ is given by $N \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota)$. With a slight abuse of notation, we also use $\underline{t}_j$ to refer to the least efficient supplier that firm j matches with in equilibrium. It is easy to show that $\underline{t}_j \equiv \underline{t}(z_j, t_j)$ increases in z_j and decreases in t_j. Naturally, firms who have better ideas reach a greater number of suppliers, while firms that are efficient in producing tasks themselves source from fewer suppliers.

The blueprint-owner receives the profits from selling the final product. For firm j with marginal cost $c_j \equiv c(z_j, t_j)$, its price, quantity, revenues, and operating profits of blueprint production are:

$$p_j^B = \frac{\sigma}{\sigma - 1} c_j, \qquad q_j^B = A(\frac{\sigma}{\sigma - 1})^{-\sigma} c_j^{-\sigma}, \qquad r_j^B = A(\frac{\sigma}{\sigma - 1})^{1-\sigma} c_j^{1-\sigma}, \qquad \pi_j^B = \frac{r_j^B}{\sigma},$$

respectively. Firms also generate profits by producing tasks, which are given by:

$$\pi^T(t_i) = \frac{1}{1+\theta} \sum_{j \in \Omega_i} x_{ij}(z_j, t_j, t_i), \tag{6}$$

where $x_{ij} = \dfrac{\sigma - 1}{\sigma} \lambda_{ij} r_j^B$ is firm j's purchases from firm i, and Ω_i is the set of buyers from i.

Given firms' optimal sourcing strategies, the model yields four sets of sharp yet intuitive com-

[30] Including a fixed cost for in-house production would create an additional set of "factoryless" firms that do not engage in manufacturing as identified by Bernard and Fort (2015). Since our analysis does not include these firms, we refrain from adding such a fixed cost.

[31] $\gamma^{1-\rho} = \frac{1+\theta-\rho+(\rho-1)(\frac{\rho}{\rho-1})^{-\theta}}{1+\theta-\rho} \Gamma(\frac{2\theta-\rho+1}{\theta})$, and Γ is the gamma function.

[32] L is the country's labor endowment, P is the aggregate price index, and $k_1 = \frac{\gamma^{1-\sigma}}{\sigma}(1 - \frac{1}{\sigma})^{\sigma-1}$.

parative statics regarding firms' sourcing choices, marginal costs, and profits.[33] First, producers with better blueprints have lower marginal costs, source a larger share of tasks from a larger number of suppliers, and also incur higher fixed costs and reach less efficient suppliers. Second, producers that are more efficient in task production have lower marginal costs, source a lower share of tasks from fewer suppliers, and are less likely to incur fixed costs to reach less efficient suppliers. Third, when the number of potential suppliers increases, firms outsource more tasks and their marginal costs decrease. Moreover, when N increases, the measure of productive suppliers increases and thus the cutoff manufacturing ability rises. Because more productive suppliers also tend to supply more tasks, the number of total suppliers a firm has might increase or decrease. Finally, task producers with better manufacturing ability supply tasks to more firms, and also supply a larger number of tasks to a given firm. Since the expected profit margin for each connection is $\frac{1}{1+\theta}$, a firm with higher manufacturing ability also earns more profits for each business connection, and thus in total earns more profits from task production. These results will be useful to characterize firms' specialization and the associated performance differences in Section 5.3.

Closed Economy

We first discuss the closed economy equilibrium–a relatively simple setting that demonstrates the intuition. Most of the results preserve when we bring in international trade. In this setting, the zero-profit condition for final-goods production yields the cutoff curve: a combination of z's and t's above which firms choose to bring their blueprint to production, which we denote as:

$$t = \Xi(z). \tag{7}$$

It is easy to verify that $\Xi(z)$ is decreasing in z. Intuitively, if a firm is competitive in the final goods market despite having a low manufacturing ability, it must possess a good blueprint. The worst blueprint brought to production, $\underline{z}$, must satisfy that $\underline{z} = \Xi^{-1}(\bar{t})$. On the other hand, the active task supplier with the least manufacturing ability can only be reached by the active blueprint-holder j who has the best blueprint but the lowest manufacturing ability, i.e., $(z_j = \bar{z}, t_j = \Xi(\bar{z}))$. Plugging this into equation (5), we obtain the cutoff $\underline{\mathrm{T}}$ above which firms are active in supplying tasks:

$$\underline{\mathrm{T}} = \underline{\mathrm{T}}(\bar{z}, \Xi(\bar{z})) = [Ak_1 \bar{z}^{\sigma-1}]^{-1} \Theta(\bar{z}, \Xi(\bar{z}))^{1-\frac{\sigma-1}{\theta}}. \tag{8}$$

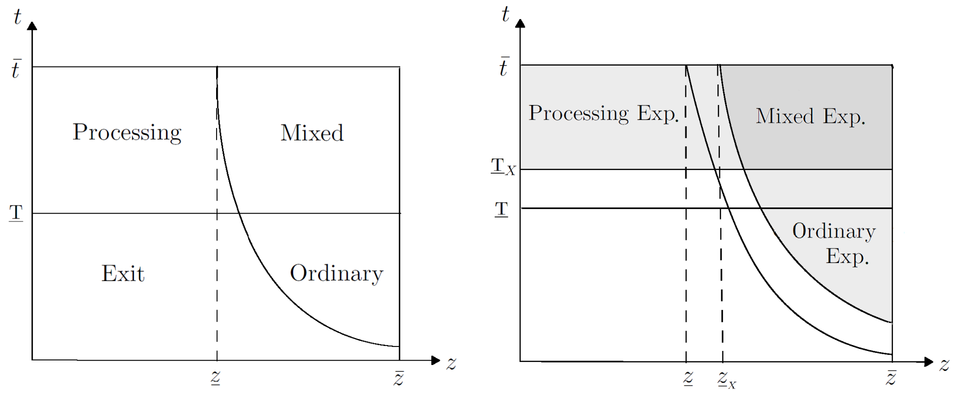
The aggregate final-good price index equals:

$$P = N^{\frac{1}{1-\sigma}} \left[\int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p^B(c(\zeta, \iota))^{1-\sigma} dG_t(\iota) dG_z(\zeta) \right]^{\frac{1}{1-\sigma}}. \tag{9}$$

Lastly, free entry implies that in equilibrium the expected profits must equal the sunk entry cost.

[33] Appendix C.1 contains formal proofs for the comparative statics for blueprint producers, and Appendix C.2 for task producers.

Figure 2: Firm Specialization



(a) Closed economy specialization

(b) Open economy specialization

Letting v^B be the profits generated from blueprint production, i.e., $v^B \equiv \pi^B(\zeta,\iota) - fn(\zeta,\iota) - f_B$, the free entry condition can be written as:

$$\int_{\underline{Z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} (\pi^B(\zeta,\iota) - fn(\zeta,\iota) - f_B) dG_t(\iota) dG_z(\zeta) + \int_{\underline{T}}^{\bar{t}} \pi^T(\iota) dG_t(\iota) = \delta f_E. \tag{10}$$

Conditional on optimal sourcing, the closed economy equilibrium therefore consists of the aggregate price index P, the number of entrants N, the cutoffs $\Xi(z)$, and $\underline{T}$ that satisfy equations (7)-(10).

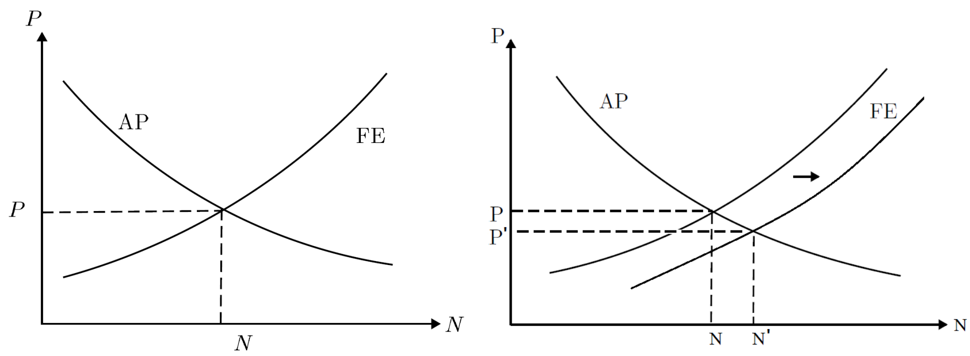
What is the specialization of firms in equilibrium? Equation (7) determines the zero-profit curve (ZPC) for final goods production and equation (8) yields cutoff $\underline{T}$ above which firms are active in task production. As visualized in Figure 2 panel (a), firms with both low z and t exit, with high z but low t engage in final good production and become *pure ordinary* firms, and those with low z but high t engage in task production and become *pure processing* firms. Finally, firms with both high z and t are active in both production and become *mixed* firms.

The system of equations that characterize the equilibrium can be simplified to two equations linking P and N—the free entry (FE) condition (10) and the aggregate price (AP) equation (9). When N increases, the marginal cost of blueprint production decreases, competition in the final goods market intensifies, and thus the AP curve is downward sloping. In contrast, a higher N implies a decrease in expected profits from task production, and a lower N^B/N implies that a smaller fraction of entrants will be active in blueprint production—both require P to rise to make firms indifferent to enter, and thus the FE curve is increasing and cut by the AP curve only once from above in the (P, N) space. This ensures the uniqueness of the equilibrium, which we present graphically in Figure 3 panel (a). The formal proof for the uniqueness of the equilibrium is in Appendix C.3.

Open Economy

In the presence of international trade and trade costs, the task supplier with the lowest manu-

Figure 3: Equilibrium Analysis



(a) Determination of the equilibrium

(b) Impact of promoting processing trade

facturing ability in home that exports to foreign satisfies: $\underline{\mathrm{T}}_X = \tau_T^{\theta}\underline{\mathrm{T}}^*$. On the final good market, firms are indifferent between selling domestically and exporting given the same net profit. This yields the export cutoff curve $t_X = \Xi_X(z)$. The net profit of selling domestically and exporting both increase in z and t, with the latter being steeper because: (i) market access is greater, and (ii) improved market access leads to a lower marginal cost of production via increased optimal sourcing. Thus other things equal, firms with better blueprint/manufacturing ability are more likely to export.

Introducing international trade yields two additional cutoffs compared to the closed economy equilibrium. As presented graphically in Figure 2 panel (b), a subset of entrants survive in each country and a smaller subset of those export. On the task production margin, active task producers have higher manufacturing ability than firms who exit, while task exporters have even higher manufacturing ability. Similarly, firms with the 'worst' blueprint quality exit, the better ones operate only in the domestic market, and the ones with the highest blueprint quality export. If a firm is good at both manufacturing and designing blueprints, it becomes a mixed exporter.

In the open economy equilibrium, expressions for the zero-profit curve and the domestic task production cutoff remain the same as before. The only change is that the sourcing capacity of a given firm becomes $\Theta(z_j, t_j) = t_j + (N + N^*) \int_{\underline{t}(z_j,t_j)}^{\bar{t}} \iota d\hat{G}_t(\iota)$, where $\hat{G}_t(\iota) = \frac{N}{N+N^*} G_t(\iota) + \frac{N^*}{N+N^*} G_t^*(\iota \tau_T^{*\theta})$. In words, firms now face greater sourcing potential as they can reach both domestic and international suppliers, although the production efficiency of foreign suppliers is discounted by $\tau_T^{*\theta}$ because of the iceberg trade costs. Similar to the case of closed economy, 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.

5.3 Linking the Model to Empirics

In this section we show that our framework is consistent with the empirical findings in sections 3 and 4. We implicitly let the choice of export mode to determine firms' specialization, motivated by *Fact 2* and aim to explain the remaining facts. We refer to pure processing exporters (PP) as processing exporters, pure ordinary exporters (PO) as ordinary exporters, and exporters that engage in both activities as mixed exporters (Mix). Formal proofs are delegated to Appendix C.4 and C.5.

Productivity We first link the equilibrium outcome to the empirical facts on different types of productivity measures. *Physical TFP* measures the efficiency of a firm in transforming inputs into outputs in terms of quantities. In our model this is captured the best by t_j, which reflects the average efficiency of a firm in task production. As the export cutoff curve is downward sloping, by selection there are relatively more firms with lower t among processing exporters compared to mixed exporters—thus, mixed exporters exhibit higher physical productivity on average. In addition, the task selection cutoff ensures that t_{PO} is always lower than t_{PP}. Therefore our model implies that mixed exporters on average have the highest physical productivity, followed by processing exporters, and then by ordinary exporters.

The *log labor productivity* of firm j is measured as $LP(z_j, t_j) = ln\left(\frac{v^B(z_j,t_j)+\pi^T(z_j,t_j)+l(z_j,t_j)}{l(z_j,t_j)}\right)$, where v^B and π^T are net profits from blueprint and task production respectively. Given $\pi^T = \frac{1}{\theta}l$, the above equation can be simplified to:

$$LP(z_j, t_j) = ln\left(\frac{v^B(z_j, t_j)}{l(z_j, t_j)} + \frac{\theta + 1}{\theta}\right).$$

We first compare processing and ordinary exporters. The export cutoff ensures that $v^B_{PO} > v^B_{PP}$ for any pair of firms. As we show later, $l_{PP} > l_{PO}$ always holds. Therefore, for any processing exporter j' and ordinary exporter j, $LP_j > LP_{j'}$, and thus $E_{PO}(LP) > E_{PP}(LP)$. To compare ordinary and mixed exporters, we first show that $\frac{v^B_j}{l_j}$ is an increasing function of z_j, and under mild regularity conditions, also an increasing function of t_j. In this case, we have that $E_{Mix}(LP) > E_{PO}(LP)$.

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 tasks 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:

$$TFPR(z_j, t_j) = ln(\frac{v^B_j + \pi^T_j + l_j}{{l_j}^{\alpha}{k_j}^{1-\alpha}}) \propto ln(\frac{v^B_j + \pi^T_j + l_j}{l_j}) = LP_j.^{34}$$

The ranking is therefore the same as that of labor productivity.

R&D and advertisement expenditures In the data, we find that pure ordinary exporters

[34] This is because ${l_j}^{\alpha}{k_j}^{1-\alpha} = {l_j}^{\alpha}(\frac{w}{w_K}l_j)^{1-\alpha} \propto l_j$ in equilibrium, where w_K is the rental price of capital.

spend more on R&D and advertising than mixed exporters, who spend more than pure processing exporters. A simple twist in our baseline model can rationalize this finding. 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.[35] We assume $f'_{RD} > 0$ and $f''_{RD} > 0$. Given the setting, $f_{RD}(a)$ is increasing in z in equilibrium. As Figure 2 panel (b) illustrates, there are relatively more firms with lower z among processing exporters compared to mixed exporters, and more firms with lower z among mixed exporters compared to ordinary exporters. Mathematically, this implies that the z distribution of pure ordinary exporters first-order stochastically dominates (FSD) the corresponding distribution of mixed exporters, which FSD the one of pure processing exporters. Applying the ranking theorem, we then have that pure ordinary exporters on average have the highest expenditure, followed by mixed exporters, then by pure processing exporters.[36]

Employment We first compare the employment of mixed and processing exporters. As labor is the only input, the associated employment increases in t. Intuitively, on average mixed exporters have greater employment compared to pure processing firms, as (1) mixed exporters employ more labor for task production, and (2) mixed exporters have additional labor for ordinary production. The comparison between processing and ordinary exporters is slightly less intuitive. Consider a processing exporter j and an ordinary exporter j'. We then compare the employment of pure processing and pure ordinary exporters. By selection pure processing exporters always have the greater manufacturing ability; this, under mild regularity condition that pure ordinary exporters always have positive sourcing, implies that ordinary exporters will always have less in-house production than any processing firm they source from, and hence employ fewer number of workers for production.

Processing Trade Policy Lastly, we investigate our model's new implication regarding processing trade policy. To highlight the idea, consider a small open economy setting such that changes at home does not affect any aggregate variables of foreign. When a processing promotion policy lowers τ_T, firm's ex-ante expected profit from task production increases. Therefore, their expected profits from final good production must decrease for the free entry condition to hold: i.e., the FE curve shifts downwards. On the other hand, the small open economy assumption ensures that the change in τ_T casts no direct impact on the final goods market, and therefore for a given N, the aggregate price index remains unchanged. As illustrated in Figure 3 panel (b), these together imply that the equilibrium N increases while P decreases. This welfare improvement comes from the fact that promoting processing trade not only directly benefits task suppliers with high manufacturing abilities ("Made in China"), but also helps firms with good ideas ("Created in China") by increasing the pool of suppliers that they could source from.

When τ_T decreases, the rise of net profits from final good production for a given firm j increases

[35] In 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.

[36] The ranking theorem is presented in Appendix C.4.

in labor productivity conditional on employment, a result we prove in Appendix C.5. The intuition of this result is straightforward: conditional on employment, higher labor productivity implies that the firm has greater branding ability compared to its manufacturing skills, and hence it will benefit more from having a greater pool of firms that are good at making compared to a firm that is already very good at manufacturing. At the extensive margin, this means that when τ_T decreases, firms with higher labor productivity are more likely to enter the final goods market, which is consistent with our last fact.

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, we showed that mixed firms, who engage predominantly in processing, are superior to other firms in multiple dimensions. Using an 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.

Conventional wisdom or workhorse trade models are not able to explain our stylized facts in a single framework. To rationalize our findings, we provided a simple theoretical framework where multi-attributes firms endogenously determine their specialization within the production network. We allowed for markups in both stages of production and introduced two dimensions of firm heterogeneity: manufacturing ability, which determines how efficient a firm is in producing tasks, and blueprint quality, which determines how good a firm is in selling its own branded products. Our framework provided a new source of gains: facilitating processing trade raises the ex-ante expected profits from task production and hence encourages entry, leading to a greater mass of potential suppliers, which eventually benefits downstream ordinary firms. Overall, our theoretical and empirical analyses highlighted that processing trade led 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.

Bernard, A. B., Eaton, J., Jensen, J. B., and Kortum, S. (2003). Plants and productivity in international trade. *American Economic Review*, 93(4):1268–1290.

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. (2019). 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., 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. (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.

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.[37] 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.[38] 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.[39] 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:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma m_{it} + \omega_{it} + \varepsilon_{it}, \tag{11}$$

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

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

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

[39] 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:

$$m_{it} = \tilde{m}_{it} - p_{Mit}. \tag{12}$$

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:

$$\bar{m}_{it} = \tilde{m}_{it} - p_{Mjt}. \tag{13}$$

Plugging (13) into (12), 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:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma \bar{m}_{it} + \omega^*_{it} + \varepsilon_{it}, \tag{14}$$

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:

$$p_{Mit} = f(p_{it}, \mathbf{X}_{it}). \tag{15}$$

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:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma \tilde{m}_{it} + \gamma f(p_{it}, \mathbf{X}_{it}) + \omega_{it} + \varepsilon_{it}. \quad (16)$$

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

$$p_{it} = \log\left(\frac{R_{it}}{Q_{it}}\right), \quad (17)$$

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.[40] Once we estimate the production function coefficients, we then compute our physical productivity ($TFPQ$) estimates, which are used in the regressions in Table 2.

[40]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.3 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.3 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:

$$\ln(\text{proc. exp.})_{icst} = \alpha + \beta\ OS_{ict-1} + \gamma_i + \delta_{st} + \phi_{ct} + \epsilon_{icst}, \tag{18}$$

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.[41] 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 (18) should be interpreted as an intention-to-treat effect.

We report the estimation results of (18) in Table A.4. 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.3 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 Table A.5, 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.

[41] 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 Comparative Statics for Blueprint Producers

Comparative statics for z_j 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:

$$\underline{t}\left(z_j,\, t_j\right) = \frac{\theta f}{\sigma - 1}\left(Ak_1 z_j^{\sigma-1}\right)^{-1}\Theta(z_j, t_j)^{1-\frac{\sigma-1}{\theta}}. \tag{19}$$

Since A and $z_j^{\sigma-1}$ enter the expression of $\underline{t}$ multiplicatively, they should affect other choice variables similarly. To save space, we only show the comparative statics for z_j. For clarity, we denote $\Theta_j \equiv \Theta\left(z_j,\, t_j\right)$ and $\underline{t}_j \equiv \underline{t}\left(z_j,\, t_j\right)$. Taking the derivative of $\underline{t}\left(z_j,\, t_j\right)$ with respect to z_j, we obtain:

$$\frac{\partial \underline{t}_j}{\partial z_j} = \frac{\theta f}{\left(\sigma - 1\right) Ak_1}\left[\left(1-\sigma\right) z_j^{-\sigma}\Theta_j^{1-\frac{\sigma-1}{\theta}} + z_j^{1-\sigma}\frac{\partial \Theta_j^{1-\frac{\sigma-1}{\theta}}}{\partial z_j}\right], \tag{20}$$

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 (20) 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 $\frac{\partial \underline{t}_j}{\partial z_j} < 0$. Note that $n\left(z_j,\, t_j\right) = N\int_{\underline{t}_j}^{\bar{t}} dG_t\left(\iota\right)$, and thus $\frac{\partial n}{\partial \underline{t}_j} < 0$. By the chain rule, we have:

$$\frac{\partial n\left(z_j,\, t_j\right)}{\partial z_j} = \frac{\partial n\left(z_j,\, t_j\right)}{\partial \underline{t}_j}\frac{\partial \underline{t}_j}{\partial z_j} > 0. \tag{21}$$

Our model implies that the share of tasks outsourced by firm j, $o\left(z_j,\, t_j\right)$, is given by:

$$o\left(z_j,\, t_j\right) = 1 - \frac{t_j}{\Theta_j}. \tag{22}$$

It immediately follows that:

$$\frac{\partial o\left(z_j,\, t_j\right)}{\partial z_j} \propto \frac{\partial \Theta_j}{\partial z_j} = \frac{\partial \Theta_j}{\partial \underline{t}_j}\frac{\partial \underline{t}_j}{\partial z_j} > 0.$$

Lastly, the unit cost is expressed as:

$$c\left(z_j,\, t_j\right) = \frac{\Theta_j^{-\frac{1}{\theta}}\gamma^{\frac{1}{1-\rho}}}{z_j}. \tag{23}$$

Note that $\Theta_j^{-\frac{1}{\theta}}$ is decreasing in z_j since $\frac{\partial \Theta_j^{-\frac{1}{\theta}}}{\partial z_j} \propto -\frac{\partial \Theta_j}{\partial \underline{t}_j}\frac{\partial \underline{t}_j}{\partial z_j} < 0$ and z_j^{-1} is also decreasing in z_j. This implies that $\frac{\partial c(z_j, t_j)}{\partial z_j} < 0$.

Comparative statics for t_j Taking the derivative of Equation (19) with respect to t_j, we get:

$$\frac{\partial \underline{t}_j}{\partial t_j} \propto \left(1 - \frac{\sigma - 1}{\theta}\right) \frac{\partial \Theta_j}{\partial t_j}. \tag{24}$$

Recall that $\Theta_j = t_j + N \int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota)$, which implies:

$$\frac{\partial \Theta_j}{\partial t_j} = 1 - N \underline{t}_j g_t\left(\underline{t}_j\right) \frac{\partial \underline{t}_j}{\partial t_j}.$$

If $\frac{\partial \underline{t}_j}{\partial t_j} \leq 0$, we must have that $\frac{\partial \Theta_j}{\partial t_j} > 0$. By (24), this in turn implies that $\frac{\partial \underline{t}_j}{\partial t_j} > 0$, which is a contradiction. Therefore it has to be the case that $\frac{\partial \underline{t}_j}{\partial t_j} > 0$. Using the expression of $n\left(z_j, t_j\right)$ and applying the chain rule, we have:

$$\frac{\partial n\left(z_j, t_j\right)}{\partial t_j} = \frac{\partial n\left(z_j, t_j\right)}{\partial \underline{t}_j} \frac{\partial \underline{t}_j}{\partial t_j} < 0. \tag{25}$$

From (22), we know that $\frac{\partial o(z_j, t_j)}{\partial t_j} = -\frac{1}{\Theta_j} + \frac{t_j}{\Theta_j^2} \frac{\partial \Theta_j}{\partial t_j}$. Since $\frac{\partial \Theta_j}{\partial t_j} = \frac{\partial \Theta_j}{\partial \underline{t}_j} \frac{\partial \partial \underline{t}_j}{\partial t_j} < 0$, it follows that $\frac{\partial o(z_j, t_j)}{\partial t_j} < 0$. Using the expression for the unit cost as defined in (23), we know that $\frac{\partial c(z_j, t_j)}{\partial t_j} \propto -\frac{\partial \Theta_j}{\partial t_j} \propto -\frac{\partial \underline{t}_j}{\partial t_j} < 0$.

Comparative statics for N Taking the derivative of Equation (19) with respect to N, we obtain:

$$\frac{\partial \underline{t}_j}{\partial N} = \frac{\theta f}{\sigma - 1} \left(A k_1 z_j^{\sigma - 1}\right)^{-1} \left(1 - \frac{\sigma - 1}{\theta}\right) \Theta_j^{-\frac{\sigma - 1}{\theta}} \frac{\partial \Theta_j}{\partial N}. \tag{26}$$

From the expression of Θ_j, we obtain:

$$\frac{\partial \Theta_j}{\partial N} = \left(\int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota) - N \underline{t}_j g_t\left(\underline{t}_j\right) \frac{\partial \underline{t}_j}{\partial N}\right). \tag{27}$$

Now suppose that $\partial \underline{t}_j / \partial N \leq 0$, then expression (27) implies that $\partial \Theta_j / \partial N > 0$. By (26), this in turn means that $\partial \underline{t}_j / \partial N > 0$, which is a contradiction. Therefore, $\partial \underline{t}_j / \partial N$ has to be positive. This also implies that $\partial \Theta_j / \partial N > 0$ by inspection of (26). After some algebra, one can show that:

$$\frac{\partial \underline{t}_j}{\partial N} = \frac{(1 - \frac{\sigma - 1}{\theta}) \frac{\underline{t}_j}{\Theta_j} \int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota)}{1 + (1 - \frac{\sigma - 1}{\theta}) N \underline{t}_j g_t(\underline{t}_j) \frac{\underline{t}_j}{\Theta_j}}.$$

Taking the derivative of n with respect to N, we have:

$$\frac{\partial n}{\partial N} = \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota) - N g_t\left(\underline{t}_j\right) \frac{\partial \underline{t}_j}{\partial N} = \frac{\int_{\underline{t}_j}^{\bar{t}} dG_t + N(1 - \frac{\sigma - 1}{\theta}) \frac{\underline{t}_j}{\Theta_j} \int_{\underline{t}_j}^{\bar{t}} (\underline{t}_j - \iota) dG_t(\iota)}{1 + (1 - \frac{\sigma - 1}{\theta}) N \underline{t}_j g_t(\underline{t}_j) \frac{\underline{t}_j}{\Theta_j}}. \tag{28}$$

Inspecting the right-hand side, the first term is positive and the second term is negative. As a result, $\partial n(z_j, t_j)/\partial N$ can either be positive or negative. By expression (22), we have:

$$\frac{\partial o}{\partial N} \propto \frac{\partial \Theta_j}{\partial N} > 0.$$

Lastly, the change in unit cost with respect to N is:

$$\frac{\partial c(z_j, t_j)}{\partial N} \propto -\frac{\partial \Theta_j}{\partial N} < 0.$$

C.2 Comparative Statics for Task Producers

Now we consider two task producers denoted by i and i'. For any given blueprint producer j, its purchase of tasks from i and i' are x_{ij} and $x_{i'j}$, respectively. Without loss of generality, we assume that j has established business relations with both suppliers, i.e., $\min\{T_i, T_{i'}\} \geq \underline{t}(z_j, t_j)$. In this case, recall that the bilateral trade between two firms is given by:

$$x_{ij} = \lambda_{ij} x_j = \frac{T_i}{\Theta_j} x_j, \tag{29}$$

$$x_{i'j} = \lambda_{i'j} x_j = \frac{T_{i'}}{\Theta_j} x_j. \tag{30}$$

This implies that $x_{i'j} > x_{ij}$. Since Ω_i represents the set of firms that source from i, we can express it as:

$$\Omega_i = \left\{ j |\, \underline{t}_j \leq T_i \right\}.$$

When $T_i < T_{i'}$, for any $j \in \Omega_i$, $\underline{t}_j \leq T_i < T_{i'}$, which implies that $j \in \Omega_{i'}$. This indicates that $\Omega_i \subseteq \Omega_{i'}$. Because $\underline{t}_j$ is a continuous and monotone function with respect to z_j or t_j, and there is a continuum of firms, there exists a j' such that $T_i < \underline{T}^*_{j'} < T_{i'}$. Therefore $\Omega_i \subset \Omega_{i'}$. Lastly, note that profits of the task producer is given by $\pi^T(T_i) = \frac{1}{1+\theta} \sum_{j \in \Omega_i} x_{ij}$, and thus $\pi^T(T_i) < \frac{1}{1+\theta} \sum_{j \in \Omega_i} x_{i'j} < \frac{1}{1+\theta} \sum_{j \in \Omega_{i'}} x_{i'j} = \pi^T(T_{i'})$.

C.3 Proof of Uniqueness

We decompose the proof of uniqueness into two parts. In the first part, we show that the aggregate price index is increasing in N. In the second part, we prove that FE curve is increasing in N.

Part I: $P(N)$ is decreasing in N.

We prove that $P(N)$ is decreasing in N by contradiction. Recall that the aggregate price index is:

$$P = N^{\frac{1}{1-\sigma}} \left[\int_{\underline{Z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p^B (c(\zeta, \iota))^{1-\sigma} dG_t(\iota) dG_z(\zeta) \right]^{\frac{1}{1-\sigma}}. \tag{31}$$

Consider $N' > N$ and $P' \geq P$. As we showed in the comparative statics, $\frac{\partial c_j}{\partial N} < 0$, $\frac{\partial c_j}{\partial P} < 0$, and hence $p'_j < p_j$. As v_j increases in P and decreases in p_j, $v'_j > v_j \geq 0$ for any firm j active in blueprint production at the old equilibrium. Therefore:

$$P' < N'^{\frac{1}{1-\sigma}} \left[\int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p'^B(c(\zeta,\iota))^{1-\sigma} dG_t(\iota) dG_z(\zeta) \right]^{\frac{1}{1-\sigma}}$$
$$< N^{\frac{1}{1-\sigma}} \left[\int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p^B(c(\zeta,\iota))^{1-\sigma} dG_t(\iota) dG_z(\zeta) \right]^{\frac{1}{1-\sigma}} = P,$$

which contradicts $P' \geq P$. Hence it must be that $P' < P$, which concludes the proof.

Part II: FE curve is upward sloping.

Let $F_{FE}(P,N) = \int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} v^B(\zeta,\iota) dG_t(\iota) dG_z(\zeta) + \int_{\underline{\mathrm{T}}}^{\bar{t}} \pi^T(\iota) dG_t(\iota) - \delta f_E$. The proof proceeds in three steps.

Step 1: $\frac{\partial F_{FE}}{\partial P} > 0$. Note that $\int_{\underline{\mathrm{T}}}^{\bar{t}} \pi^T(\iota) dG_t(\iota) = \frac{(\sigma-1)\beta L}{N\sigma(\theta+1)}$. Applying the Leibniz rule,

$$\frac{\partial F_{FE}}{\partial P} = \int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \frac{\partial v^B(\zeta,\iota)}{\partial P} dG_t(\iota) dG_z(\zeta)$$
$$- \int_{\Xi(\underline{\mathrm{Z}})}^{\bar{t}} g_z(\underline{\mathrm{z}}) v^B(\underline{\mathrm{z}},\iota) \frac{\partial \underline{\mathrm{Z}}}{\partial P} dG_t(\iota) - \int_{\underline{\mathrm{Z}}}^{\bar{z}} g_t(\Xi(\zeta)) v^B(\zeta,\Xi(\zeta)) \frac{\partial \Xi(\zeta)}{\partial P}(\iota) dG_z(\zeta) \ .$$

As $v^B(\zeta,\Xi(\zeta)) = 0$, $\Xi(\underline{\mathrm{z}}) = \bar{t}$, the last two terms of above equation are zero; hence:

$$\frac{\partial F_{FE}}{\partial P} = \int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \frac{\partial v^B(\zeta,\iota)}{\partial P} dG_t(\iota) dG_z(\zeta).$$

By the envelope theorem, we know that $\frac{\partial v^B}{\partial P} = (\sigma-1)\frac{v^B}{P} \geq 0$, which holds with equality when $v = 0$. Hence:

$$\frac{\partial F_{FE}}{\partial P} = \int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} (\sigma-1)\frac{v^B}{P} dG_t(\iota) dG_z(\zeta) > 0.$$

Step 2: $\frac{\partial F_{FE}}{\partial N} < 0$. Applying the Leibniz rule, we get:

$$\frac{\partial F_{FE}}{\partial N} = \int_{\underline{\mathrm{Z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \frac{\partial v^B(\zeta,\iota)}{\partial N} dG_t(\iota) dG_z(\zeta) - \frac{(\sigma-1)\beta L}{N^2\sigma(\theta+1)}.$$

By the envelope theorem:

$$\frac{\partial v^B}{\partial N} = \frac{\sigma-1}{\theta}\pi^B \frac{\Theta - t}{\Theta}\frac{1}{N} - \frac{fn}{N}$$
$$= \frac{\sigma-1}{\theta}\frac{\pi^B}{\Theta} \int_{\underline{\mathrm{t}}}^{\bar{t}} \iota dG_t(\iota) - f \int_{\underline{\mathrm{t}}}^{\bar{t}} dG_t(\iota),$$

and thus we get:

$$\begin{aligned}\frac{\partial F_{FE}}{\partial N} &= \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \left(\frac{\sigma-1}{\theta}\pi^B \frac{\Theta - t}{\Theta}\frac{1}{N}\right) dG_t(\iota)dG_z(\zeta) \\ &- \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \frac{fn}{N} dG_t(\iota)dG_z(\zeta) - \frac{(\sigma-1)\beta L}{N^2\sigma(\theta+1)}. \end{aligned} \tag{32}$$

Because $N \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \pi^B dG_t(\iota)dG_z(\zeta) = \frac{\beta L}{\sigma}$, equation (32) can then be simplified to:

$$\frac{\partial F_{FE}}{\partial N} = \frac{\sigma-1}{N\theta}\left(\int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \left(\frac{1}{\theta+1}\pi^B - \frac{t\pi^B}{\Theta} - \frac{\theta}{\sigma-1}fn\right) dG_t(\iota)dG_z(\zeta)\right).$$

Given that $\underline{t}_j = f\left(Ak_1 z_j^{\sigma-1}\right)^{-1}\Theta(z_j,t_j)^{1-\frac{\sigma-1}{\theta}}\dfrac{\theta}{\sigma-1}$, we can express f as a function of π_j^B and $\underline{t}_j$:

$$f = \underline{t}_j \frac{\pi_j^B}{\Theta_j}\frac{\sigma-1}{\theta}.$$

Hence we can show that $\frac{1}{\theta+1}\pi^B - \frac{t\pi^B}{\Theta} - \frac{\theta}{\sigma-1}fn = \frac{\pi^B}{\Theta}\left(\frac{\Theta}{\theta+1} - t - n\underline{t}\right)$.

Now focus on $\frac{\Theta}{\theta+1} - t - n\underline{t}$. Taking the partial derivative with respect to $\underline{t}$, we get:

$$\frac{\partial(\frac{\Theta}{\theta+1} - t - n\underline{t})}{\partial \underline{t}} = \frac{\theta}{\theta+1}N\underline{t}g_t(\underline{t}) - N\int_{\underline{t}}^{\bar{t}} g_t(\iota)d\iota. \tag{33}$$

Furthermore:

$$\frac{\partial^2(\frac{\Theta}{\theta+1} - t - n\underline{t})}{\partial \underline{t}^2} = \frac{\theta}{\theta+1}N\underline{t}g'_t(\underline{t}) + \frac{\theta}{\theta+1}Ng_t(\underline{t}) + Ng_t(\underline{t}). \tag{34}$$

Recall that $\iota g'_t(\iota) + g_t(\iota) > 0$, and hence $\frac{\partial^2(\frac{\Theta}{\theta+1}-t-n\underline{t})}{\partial \underline{t}^2} > 0$. Therefore, $\frac{\partial(\frac{\Theta}{\theta+1}-t-n\underline{t})}{\partial \underline{t}}$ reaches its maximum when $\underline{t} = \bar{t}$. As $\frac{\partial(\frac{\Theta}{\theta+1}-t-n\underline{t})}{\partial \underline{t}}$ approaches zero when $\underline{t}$ approaches $\bar{t}$, we have $\frac{\partial(\frac{\Theta}{\theta+1}-t-n\underline{t})}{\partial \underline{t}} \leq 0$. In other words, $\frac{\Theta}{\theta+1} - t - n\underline{t}$ reaches its highest value when $\underline{t}$ reaches its lowest. Recall that in equilibrium, the least productive suppliers are reached by firms with the best blueprint quality and the 'worst' manufacturing ability, i.e., $i = \{\bar{z}, \Xi(\bar{z})\}$. At the same time, $v^B(\bar{z}, \Xi(\bar{z})) = 0$, which implies that:

$$fn_i = \pi_i^B. \tag{35}$$

Optimal sourcing condition implies that:

$$f = \underline{t}_i \frac{\pi_i^B}{\Theta_i}\frac{\sigma-1}{\theta}. \tag{36}$$

Equations (35) and (36) together imply that $n_i \underline{t}_i = \Theta_i \frac{\theta}{\sigma-1}$. Therefore:

$$\frac{\Theta}{\theta+1} - t - n\underline{t} \leq \frac{\Theta_i}{\theta_i+1} - t_i - n_i\underline{t}_i < \frac{\Theta_i}{\theta_i+1} - n_i\underline{t}_i < \frac{\Theta_i}{\theta_i+1} - \frac{\Theta_i\theta}{\sigma-1} < 0. \tag{37}$$

As a result, $\frac{1}{\theta+1}\pi^B - \frac{t\pi^B}{\Theta} - \frac{\theta}{\sigma-1}fn < 0$, and hence $\frac{\partial F}{\partial N} < 0$. As $\frac{\partial F}{\partial P} > 0$ and $\frac{\partial F}{\partial N} < 0$, it is immediate that the FE curve is upward sloping:

$$\frac{\partial P(N)}{\partial N} = -\frac{\partial F_{FE}/\partial N}{\partial F_{FE}/\partial P} > 0.$$

C.4 Proofs of Ranks

To facilitate the analysis, we first introduce the ranking theorem which is used repetitively in this subsection:

Ranking Theorem. *For any increasing and piecewise differentiable function $u(x)$, if cumulative G first-order stochastically dominates (FSD) cumulative G', then:*

$$E_G[u(x)] > E_{G'}[u(x)].$$

Proof of $G_{Mix}(t)$ FSD $G_{PP}(t)$.

We first write down the cumulative distribution functions of mixed and processing exporters:

$$F_{Mix}(t < t') = \frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{\bar{t}} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)}, \qquad F_{PP}(t < t') = \frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\underline{\mathrm{Z}}}^{\Xi_X^{-1}(\iota)} dG_z(\zeta)dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{\bar{t}} \int_{\underline{\mathrm{Z}}}^{\Xi_X^{-1}(\iota)} dG_z(\zeta)dG_t(\iota)}.$$

Proving $F_{Mix}(t < t') < F_{PP}(t < t')$ for any $t' > \underline{\mathrm{T}}_X$ is equivalent to proving:

$$\frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{\bar{t}} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)} < \frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\underline{\mathrm{Z}}}^{\Xi_X^{-1}(\iota)} dG_z(\zeta)dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{\bar{t}} \int_{\underline{\mathrm{Z}}}^{\Xi_X^{-1}(\iota)} dG_z(\zeta)dG_t(\iota)},$$

which, after some algebra, is equivalent to:

$$\frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{\bar{t}} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)} - \frac{\int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{\bar{t}} dG_t(\iota)} < 0.$$

The left-hand side of above expression equals zero when $t' = \bar{t}$. Hence for the inequality to hold, it is sufficient to prove that $\frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota)}$ is increasing in t. Taking a partial derivative with respect to t', we get:

$$\begin{aligned}
\frac{\partial \frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota)}}{\partial t'} &= \frac{g_t(t')}{(\int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota))^2} \int_{\Xi_X^{-1}(t')}^{\bar{z}} dG_z(\zeta) \int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota) \\
&- \frac{g_t(t')}{(\int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota))^2} \int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) dG_t(\iota) \\
&\propto \int_{\Xi_X^{-1}(t')}^{\bar{z}} dG_z(\zeta) \int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota) - \int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) dG_t(\iota).
\end{aligned}$$

Note that as $\Xi_X^{-1}(t)$ is decreasing in t, $\Xi_X^{-1}(t')$ is smaller than any $\Xi_X^{-1}(\iota)$ with $\iota \in (\underline{\mathrm{T}}_X, t')$. Hence, $\int_{\Xi_X^{-1}(t')}^{\bar{z}} dG_z(\zeta) > \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta)$ for $\iota \in (\underline{\mathrm{T}}_X, t')$:

$$\begin{aligned}
&\int_{\Xi_X^{-1}(t')}^{\bar{z}} dG_z(\zeta) \int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota) - \int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) dG_t(\iota) \\
= &\int_{\underline{\mathrm{T}}_X}^{t'} \left(\int_{\Xi_X^{-1}(t')}^{\bar{z}} dG_z(\zeta) - \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) \right) dG_t(\iota) > 0.
\end{aligned}$$

Thus, $\partial \frac{\int_{\underline{\mathrm{T}}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) dG_t(\iota)}{\int_{\underline{\mathrm{T}}_X}^{t'} dG_t(\iota)} / \partial t' > 0$, which concludes the proof.

Proof of $G_{Mix}(z)$ FSD $G_{PP}(z)$.

Denote $z_1 \equiv \Xi_X^{-1}(\bar{t})$, $z_2 \equiv \Xi_X^{-1}(\underline{\mathrm{T}}_X)$. If $z' < z_1$, then $F_{Mix}(z < z') = 0, F_{PP}(z < z') > 0$; if $z' \geq z_2$, then $F_{Mix}(z < z') < 1, F_{PP}(z < z') = 1$. In these two cases, $F_{Mix}(z < z') < F_{PP}(z < z')$ always holds. When $z' \in [z_1, z_2)$, we have:

$$F_{Mix}(z < z') = \frac{\int_{z_1}^{z'} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta)}{\int_{z_1}^{z_2} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta) + (1 - G_t(\underline{\mathrm{T}}_X))(1 - G_z(z_2))} < \frac{\int_{z_1}^{z'} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta)}{\int_{z_1}^{z_2} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta)},$$

$$F_{PP}(z < z') = \frac{\int_{z_1}^{z'} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta) + (1 - G_t(\underline{\mathrm{T}}_X)) G_z(z_1)}{\int_{z_1}^{z_2} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta) + (1 - G_t(\underline{\mathrm{T}}_X)) G_z(z_1)} > \frac{\int_{z_1}^{z'} \int_{\underline{\mathrm{T}}_X}^{\Xi_X(\zeta)} dG_t(\iota) dG_z(\zeta)}{\int_{z_1}^{z_2} \int_{\underline{\mathrm{T}}_X}^{\Xi_X(\zeta)} dG_t(\iota) dG_z(\zeta)}.$$

As $\Xi_X(z)$ is decreasing in z, the proof for $G_{Mix}(t)$ FSD $G_{PP}(t)$ applies here as well. Therefore, we have:

$$F_{Mix}(z < z') < \frac{\int_{z_1}^{z'} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta)}{\int_{z_1}^{z_2} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta)} < \frac{\int_{z_1}^{z'} \int_{\underline{\mathrm{T}}_X}^{\Xi_X(\zeta)} dG_t(\iota) dG_z(\zeta)}{\int_{z_1}^{z_2} \int_{\underline{\mathrm{T}}_X}^{\Xi_X(\zeta)} dG_t(\iota) dG_z(\zeta)} < F_{PP}(z < z'),$$

when $z' \in [z_1, z_2)$. This concludes the proof.

Proof of $G_{PO}(z)$ FSD $G_{Mix}(z)$.

When $z' \in [z_1, z_2)$, $F_{PO}(z < z') = 0, F_{Mix}(z < z') > 0$, and hence $F_{PO}(z < z') < F_{Mix}(z < z')$ holds. When $z' \geq z_2$, we have:

$$\begin{aligned} F_{Mix}(z < z') &= \frac{\int_{z_1}^{z_2} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta) + (1 - G_t(\underline{\mathrm{T}})) \int_{z_2}^{z'} dG_z(\zeta)}{\int_{z_1}^{z_2} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta) + (1 - G_t(\underline{\mathrm{T}}_X))(1 - G_z(z_2))} \\ &> \frac{(1 - G_t(\underline{\mathrm{T}}_X)) \int_{z_2}^{z'} dG_z(\zeta)}{(1 - G_t(\underline{\mathrm{T}}_X))(1 - G_z(z_2))} = \frac{\int_{z_2}^{z'} dG_z(\zeta)}{1 - G_z(z_2)}. \end{aligned}$$

Similarly, one can show that when $z' \geq z_2$:

$$F_{PO}(z < z') = \frac{\int_{z_2}^{z'} \int_{\Xi_X(\zeta)}^{\underline{\mathrm{T}}_X} dG_t(\iota) dG_z(\zeta)}{\int_{z_2}^{\bar{z}} \int_{\Xi_X(\zeta)}^{\underline{\mathrm{T}}_X} dG_t(\iota) dG_z(\zeta)} < \frac{\int_{z_2}^{z'} dG_z(\zeta)}{1 - G_z(z_2)}.$$

Therefore, $F_{Mix}(z < z') > F_{PO}(z < z')$, i.e., $G_{PO}(z)$ FSD $G_{Mix}(z)$.

Labor Productivity. The labor productivity of firm j is given by:

$$LP_j = \frac{v_j^B}{l_j} + (1 + \frac{1}{\theta}).$$

Note that:

$$\frac{\partial ln v_j^B}{\partial ln z_j} = (\sigma - 1) \frac{\pi_j^B}{v_j^B} > \sigma - 1, \tag{38}$$

$$\frac{\partial ln l_j^B}{\partial ln z_j} = \frac{(\frac{\sigma-1}{\theta} - 1)(\sigma - 1) M_j}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} + (\sigma - 1) = \frac{\sigma - 1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} < \sigma - 1, \tag{39}$$

where $M_j \equiv N \underline{\mathrm{t}}_j g_t(\underline{\mathrm{t}}_j) \frac{\underline{\mathrm{t}}_j}{\Theta}$. Hence $\frac{v_j^B}{l_j^B}$ increases in z_j. As the labor used for producing tasks for other firms does not change with z, it immediately follows that $\frac{v_j^B}{l_j}$ is increasing in z_j as well. Similarly, it is easy to verify that:

$$\frac{\partial ln v_j^B}{\partial ln t_j} = \frac{(\sigma - 1) t_j \pi_j^B}{\theta v_j^B \Theta_j}, \tag{40}$$

$$\frac{\partial ln l_j^B}{\partial ln t_j} = 1 - \frac{(1 - \frac{\sigma-1}{\theta}) t_j}{\Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j}. \tag{41}$$

Hence, we have:

$$
\begin{aligned}
\frac{\partial ln v_j^B}{\partial ln t_j} - \frac{\partial ln(l_j^B)}{\partial ln t_j} &= \frac{(\sigma-1)t_j\pi_j^B}{\theta v_j^B \Theta_j} + \frac{(1-\frac{\sigma-1}{\theta})t_j}{\Theta_j}\frac{1}{1+(1-\frac{\sigma-1}{\theta})M_j} - 1 \\
&= \frac{t_j}{\Theta_j}\left(\frac{(\sigma-1)\pi_j^B}{\theta v_j^B} + \frac{(1-\frac{\sigma-1}{\theta})}{1+(1-\frac{\sigma-1}{\theta})M_j}\right) - 1 \\
&> \frac{t_j}{\Theta_j}\frac{(\sigma-1)\pi_j^B}{\theta v_j^B} - 1 = \frac{(\sigma-1)l_j^B}{\theta v_j^B} - 1.
\end{aligned}
$$

Thus for $\frac{v_j^B}{l_j^B}$ to increase in t_j, it is necessary that $\frac{l_j^B}{v_j^B} > \frac{\theta}{\sigma-1}$. This can happen if the fixed cost of exporting is sufficiently high, so that the production employment is much larger than profits even for firms with the best manufacturing ability. If, at the same time, when t increases, the increase in production workers due to the increased task supply is not high enough to completely offset the increase in $\frac{v_j^B}{l_j^B}$, then $\frac{v_j^B}{l_j}$ will increase in t.

C.5 Proof that as the cost of processing exports (τ_T) decreases, conditional on employment, 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_T 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 $\frac{\partial^2 v_j^B}{\partial \tau_T \partial LP_j} > 0$.

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

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 from final good production for firm j, dv_j^B, equals:

$$
dv_j^B = \frac{\partial v_j^B}{\partial N}dN + \frac{\partial v_j^B}{\partial P}dP = \frac{\sigma-1}{\theta}\frac{\pi_j^B}{\Theta_j}\frac{\partial \Theta_j}{\partial N}dN - f\frac{\partial n_j}{\partial N}dN + (\sigma-1)\pi_j^B\frac{dP}{P}.
$$

Recall that when firms optimize their sourcing decisions, we have that $f = \underline{t}_j\frac{\pi_j^B}{\Theta_j}\frac{\sigma-1}{\theta}$. Hence, we can rewrite dv_j^B as:

$$
\begin{aligned}
dv_j^B &= \frac{\sigma-1}{\theta}\frac{\pi_j^B}{\Theta_j}\frac{\partial \Theta_j}{\partial N}dN - \underline{t}_j\frac{\pi_j^B}{\Theta_j}\frac{\sigma-1}{\theta}\frac{\partial n}{\partial N}dN + (\sigma-1)\pi_j^B\frac{dP}{P} \\
&\propto \frac{1}{\Theta_j}N\frac{\partial \Theta_j}{\partial N}N - \frac{1}{\Theta_j}\underline{t}_j\frac{\partial n}{\partial N} + \theta\frac{\partial dlnP}{\partial dlnN}. \qquad (42)
\end{aligned}
$$

With international trade, the expression of Θ_j is the following:

$$\Theta_j = t_j + (N + N^*) \int_{\underline{t}_j}^{\bar{t}} \iota d\hat{G}_t(\iota),$$

where $\hat{G}_t(\iota) = \frac{N}{N+N^*} G_t(\iota) + \frac{N^*}{N+N^*} G_t^*(\iota \tau_T^{*\theta})$. After some algebra, one can show that:

$$\frac{\partial \Theta_j}{\partial N} = \int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota), \qquad \frac{\partial n}{\partial N} = \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota).$$

Plugging the above two equations into (42), we get:

$$dv_j^B = \frac{\sigma - 1}{\theta} \pi_j^B \left(\frac{N \int_{\underline{t}_j}^{\bar{t}} (\iota - \underline{t}_j) dG_t(\iota)}{\Theta_j} + \theta \frac{\partial dlnP}{\partial dlnN} \right). \tag{43}$$

Let $F_{dv} \equiv \frac{N \int_{\underline{t}_j}^{\bar{t}} (\iota - \underline{t}_j) dG_t(\iota)}{\Theta_j}$. We now have:

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

Recall that:

$$\frac{\partial \pi_j^B}{\partial t_j} = \frac{(\sigma - 1)\pi_j^B}{\theta \Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j}.$$

Therefore:

$$\begin{aligned}
\frac{\partial dv_j^B}{\partial t_j} &\propto \frac{(\sigma - 1)\pi_j^B}{\theta \Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} (F_{dv} + \theta \frac{\partial dlnP}{\partial dlnN}) + \pi_j^B \frac{\partial F_{dv}}{\partial t_j} \\
&< \frac{\pi_j^B}{\Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} F_{dv} + \pi_j^B \frac{\partial F_{dv}}{\partial t_j} \propto \frac{1}{\Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} F_{dv} + \frac{\partial F_{dv}}{\partial t_j}.
\end{aligned}$$

As $\frac{\partial \underline{t}_j}{\partial t_j} > 0$ and $\frac{\partial \Theta_j}{\partial t_j} > 0$, the following inequality holds:

$$\begin{aligned}
\frac{1}{\Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} F_{dv} + \frac{\partial F_{dv}}{\partial t_j} &= \frac{N}{\Theta_j^2} \frac{\int_{\underline{t}_j}^{\bar{t}} (\iota - \underline{t}_j) dG_t(\iota)}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} + \frac{\partial F_{dv}}{\partial t_j} \\
&< \frac{N}{\Theta_j^2} \frac{\int_{\underline{t}_j}^{\bar{t}} (\iota - \underline{t}_j) dG_t(\iota)}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} - \frac{N}{\Theta_j^2} \frac{\partial \Theta_j}{\partial t_j} \int_{\underline{t}_j}^{\bar{t}} (\iota - \underline{t}_j) dG_t(\iota) \\
&\propto \frac{1}{1 + (1 - \frac{\sigma-1}{\theta}) M_j} - \frac{\partial \Theta_j}{\partial t_j} = 0.
\end{aligned} \tag{44}$$

This concludes the proof that $\frac{\partial dv_j^B}{\partial t_j} < 0$. Similarly, taking the partial derivative of F_{dv} with respect to z_j yields:

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

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 $\frac{\partial \underline{t}_j}{\partial z_j} < 0$, we have:

$$\begin{aligned}\frac{\partial dv_j^B}{\partial z_j} &\propto \frac{(\sigma-1)\pi_j^B}{\theta\Theta_j}(-\frac{\partial \Theta_j}{\partial \underline{t}_j})(F_{dv} + \theta\frac{\partial dlnP}{\partial dlnN}) + \frac{\pi_j^B N}{\Theta_j^2}\left(\int_{\underline{t}_j}^{\bar{t}} dG_t(\iota)\Theta_j + \frac{\partial \Theta_j}{\partial \underline{t}_j}\int_{\underline{t}_j}^{\bar{t}}(\iota - \underline{t}_j)dG_t(\iota)\right) \\ &\propto (1 - \frac{(\sigma-1)}{\theta})\frac{\partial \Theta_j}{\partial \underline{t}_j}F_{dv} + \frac{\partial \Theta_j}{\partial \underline{t}_j}\left(-\frac{(\sigma-1)\partial dlnP}{\partial dlnN}\right) + N\int_{\underline{t}_j}^{\bar{t}} dG_t(\iota).\end{aligned}$$

Denoting $1 - \frac{(\sigma-1)}{\theta} \equiv \Delta_1 \in (0,1)$, $\left(-\frac{(\sigma-1)\partial dlnP}{\partial dlnN}\right) \equiv \Delta_2 > 0$, we simplify the above expression to:

$$\frac{\partial dv_j^B}{\partial z_j} \propto \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota) + \Delta_1\frac{\partial \Theta_j}{\Theta_j\partial \underline{t}_j}\int_{\underline{t}_j}^{\bar{t}}(\iota - \underline{t}_j)dG_t(\iota) + \frac{\Delta_2\partial\Theta_j}{N\partial \underline{t}_j} \equiv f_{dv}.$$

Note that given $-tg_t'(t) < g_t(t)$, we have $\frac{\partial^2\Theta_j}{\partial \underline{t}_j^2} < 0$. Thus:

$$\begin{aligned}\frac{\partial f_{dv}}{\partial \underline{t}_j} &= -g_t(\underline{t}_j) + \Delta_1\frac{\partial^2\Theta_j}{\Theta_j\partial \underline{t}_j^2}\int_{\underline{t}_j}^{\bar{t}}(\iota - \underline{t}_j)dG_t(\iota) + \Delta_1\frac{\partial \Theta_j}{\Theta_j\partial \underline{t}_j}\int_{\underline{t}_j}^{\bar{t}}(-1)dG_t(\iota) \\ &\quad + \frac{\Delta_2\partial^2\Theta_j}{N\partial \underline{t}_j^2} - \Delta_1\frac{1}{\Theta_j^2}(\frac{\partial \Theta_j}{\partial \underline{t}_j})^2\int_{\underline{t}_j}^{\bar{t}}(\iota - \underline{t}_j)dG_t(\iota). \end{aligned} \tag{45}$$

Note that:

$$\begin{aligned}-g_t(\underline{t}_j) + \Delta_1\frac{\partial\Theta_j}{\Theta_j\partial \underline{t}_j}\int_{\underline{t}_j}^{\bar{t}}(-1)dG_t(\iota) &< -g_t(\underline{t}_j) + \Delta_1\frac{\underline{t}_j g_t(\underline{t}_j)}{\int_{\underline{t}_j}^{\bar{t}}\iota dG_t(\iota)}\int_{\underline{t}_j}^{\bar{t}}(1)dG_t(\iota) \\ &= -g_t(\underline{t}_j) + \Delta_1 g_t(\underline{t}_j)\frac{\int_{\underline{t}_j}^{\bar{t}}\underline{t}_j dG_t(\iota)g_t(\underline{t}_j)}{\int_{\underline{t}_j}^{\bar{t}}\iota dG_t(\iota)} < 0,\end{aligned}$$

while the rest of the terms on the right-hand side of (45) are all negative, and thus we have $\frac{\partial f_{dv}}{\partial \underline{t}_j} < 0$. As $\lim_{\underline{t}=\bar{t}} f_{dv}(\underline{t}) = 0$, we know that $\frac{\partial dv_j^B}{\partial z_j} > 0$ when $\underline{t}_j \neq \bar{t}$. Hence dv_j is increasing in z_j.

Part III: Conditional on employment, firms' labor productivity increases as z increases.

Given (38), (39), (40), and (41), we calculate how z and t change along the iso-l^B and iso-v^B curves. Along the iso-l^B curve, we have:

$$\frac{\partial lnt_j}{\partial lnz_j} \mid l^B = -\frac{\sigma - 1}{1 + \Delta_1 M_j - \frac{\Delta_1 t_j}{\Theta_j}} < 0.$$

Along the iso-v^B curve, we have:

$$\frac{\partial ln t_j}{\partial ln z_j} \mid v^B = -\frac{\theta_1 \Theta_j}{t_j} < 0.$$

Therefore:

$$\frac{\partial ln t_j}{\partial ln z_j} \mid v_B - \frac{\partial ln t_j}{\partial ln z_j} \mid l^B = -\frac{1 + \Delta_1 M_j - \Delta_1(\sigma - 1) + (\sigma - 1)}{1 + \Delta_1 M_j - \frac{\Delta_1 t_j}{\Theta_j}}. \tag{46}$$

Recall that $\Delta_1 \equiv 1 - \frac{(\sigma-1)}{\theta}$, hence (46) can be reduced to:

$$\frac{\partial ln t_j}{\partial ln z_j} \mid v_B - \frac{\partial ln t_j}{\partial ln z_j} \mid l^B = -\frac{1 + \Delta_1 M_j + \frac{(\sigma-1)^2}{\theta}}{1 + \Delta_1 M_j - \frac{\Delta_1 t_j}{\Theta_j}} < 0.$$

Denoting the number of production workers employed to produce tasks for other firms as l_j^T, we know that $\frac{\partial \pi_j^T}{\partial t_j} > 0$ from the comparative statics proof in Section C.2. Because of constant markups, this in turn implies that $\frac{\partial l_j^T}{\partial t_j} > 0$. Hence:

$$\frac{\partial ln t_j}{\partial ln z_j} \mid l = -\frac{z_j}{t_j} \frac{\frac{\partial l_j^B}{\partial z}}{\frac{\partial l_j^B}{\partial t_j} + \frac{\partial l_j^T}{\partial t_j}} > -\frac{t_j}{z_j} \frac{\frac{\partial l_j^B}{\partial t_j}}{\frac{\partial l_j^B}{\partial z}} \equiv \frac{\partial ln z_j}{\partial ln t_j} \mid l_B. \tag{47}$$

Therefore, $\frac{\partial ln t_j}{\partial ln z_j} \mid v_B - \frac{\partial ln t_j}{\partial ln z_j} \mid l < 0$ holds as well. This in turn implies that holding employment constant, with the increase in z, v^B must increase, since:

$$\frac{\partial v_j^B}{\partial z_j} \mid l = \frac{\partial v_j^B}{\partial z_j} + \frac{\partial v_j^B}{\partial t_j} \frac{\partial t_j}{\partial z_j} \mid l \propto -\frac{\partial ln t_j}{\partial ln z_j} \mid v_B + \frac{\partial ln t_j}{\partial ln z_j} \mid l > 0.$$

Recall that the labor productivity of firm j is given by:

$$LP_j = \frac{v_j^B}{l_j} + (1 + \frac{1}{\theta}).$$

Holding l_j constant, LP_j is positively associated with v_j^B. Therefore, conditional on employment, firms' labor productivity increases as z increases.

Part II: $\frac{\partial^2 v_j^B}{\partial \tau_T \partial LP_j} \mid l > 0$. Consider two firms j and j' with the same employment, but $LP_j > LP_{j'}$. From Part I, we know that $z_j > z_{j'}$ must hold. Moreover, given (47), it is easy to verify that $\frac{\partial ln t_j}{\partial ln z_j} \mid l < 0$. Therefore, we have $t_j < t_{j'}$. Recall that we showed that $\frac{\partial^2 v_j^B}{\partial \tau_T \partial z_j} > 0$ and $\frac{\partial^2 v_j^B}{\partial \tau_T \partial t_j} < 0$. Hence, it immediately follows that $\frac{\partial v_j^B}{\partial \tau_T} > \frac{\partial v_{j'}^B}{\partial \tau_T}$. This concludes the proof that $\frac{\partial^2 v_j^B}{\partial \tau_T \partial LP_j} \mid l > 0$.

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: 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.4: Paperless Trade and Processing Exports

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

Notes: This table reports the results of running specification (18). 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. Standard errors clustered at the prefecture and sector level are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

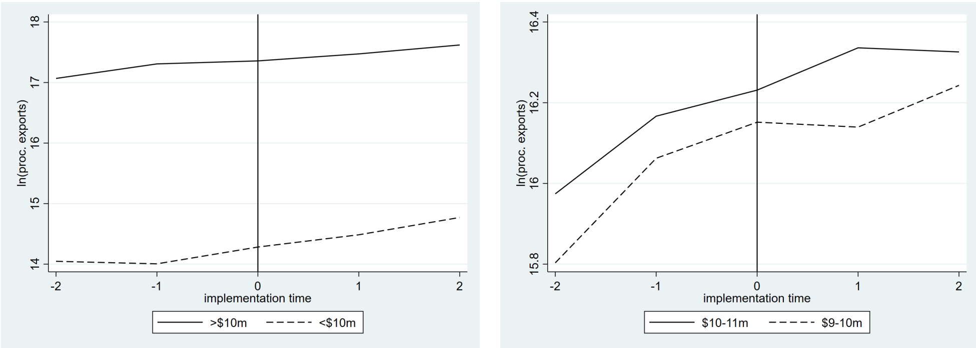
Table A.5: Paperless Trade and Processing Exports - Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep. var.: ln(proc. exp.)$_{icst}$	Entry & exit	FD	bw: \$9.5-10.5m	bw: \$8.5-11.5m	Always exporters	No SOEs	Mixed only (ordinary)	Mixed only (proc.)	False threshold (\$9m)	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.4. 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.