TI 2019-080/VI
Tinbergen Institute Discussion Paper

Made and Created in China: Super Processors and Two-way Heterogeneity

Zhiyuan Chen[1]
Aksel Erbahar[2]
Yuan Zi[3]

[1] Pennsylvania State University
[2] Erasmus School of Economics (ESE), Erasmus University Rotterdam, the Netherlands, and Tinbergen Institute,
[3] University of Oslo

Made and Created in China: Super Processors and Two-way Heterogeneity*

Zhiyuan Chen† Aksel Erbahar‡ Yuan Zi§

November 2019

Abstract

In this paper, we show that there exists a special breed of firms that are active in both ordinary and processing exports. Contrary to the existing literature that describes processing firms as inferior, these mixed firms are superior to other firms in multiple dimensions, and hence we call them "super processors." We build on Antràs et al. (2017) and Bernard et al. (2019) to develop a model in which firms are heterogeneous in multiple stages of production. Firms endogenously choose to become suppliers or final good producers, and those that excel in both manufacturing ability and blueprint quality choose to engage in both activities. We test our model's central prediction by exploiting China's pilot "paperless" processing trade supervision program that lowered the cost of processing trade but left ordinary trade costs unchanged. We find that facilitating processing exports induces productive domestic downstream firms to establish their own trademarks. Our results highlight that processing trade not only leads goods to be "Made in China," but also "Created in China."

JEL codes: F1, F12, F13, F14, L11, L21

Keywords: heterogeneous firms, production networks, trade policy, processing trade, China

*We are grateful to Andrew Bernard, Maarten Bosker, Jonathan Eaton, Julian Emami Namini, Julia Cajal Grossi, 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, Zi Wang, 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 with us. 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).

†Department of Economics, Pennsylvania State University, United States, *email*: chenzhiyuan1224@gmail.com

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

§Department of Economics, University of Oslo, Oslo, Norway, *email*: yuanzi.economics@gmail.com (corresponding author)

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 1978, when the country began to liberalize its economy, to over 60% in 2007, just before the Great Recession (World Bank, 2018). The literature has so far attributed China's export boom to be driven largely by firms supplying relatively low value-added tasks to foreign multinationals, as epitomized by the "Made in China" tag. However, this phenomenon is changing as Chinese firms strive to *create* their own brands and contribute to China's exports by engaging in high value-added activities (Balmer and Chen, 2017). One aspect that has been increasingly pointed out by policymakers and business people but seldom investigated in the academic literature is that after decades of efforts to become 'the factory of the world' with the help of various industrial policies, including processing trade policy, China's advanced manufacturing base is now attracting firms with innovative ideas.

For example, Shenzhen, a city where processing exports made up over 80% of total exports and 100% of GDP in the late 1990s,[1] has now become China's technology and innovation hub. A recent news article by Wang (2019) describes how the city's fast turnover of supply chains, proximity of its factories to the city-center, and the wide availability of competitively priced parts and components in its famous electronics market Huaqiangbei makes Shenzhen a boon for startups.

In this paper, we examine China's processing trade regime to make three contributions to our understanding of how firms' attributes determine their specialization within a value chain. First, we provide novel stylized facts on exporters' performance, choice of trade mode, and brand ownership. Then, we develop a model with endogenous production networks where firms are heterogeneous in both manufacturing and non-manufacturing abilities to rationalize these stylized facts. Our model provides a new source of gains from promoting processing trade, which we test by evaluating the impact of China's "paperless" processing trade policy program. We find that facilitating processing trade induces domestic downstream firms to intensify their branding activities.

While processing exports account for the majority of China's total exports in 2000-2006, existing studies describe processing firms as smaller and less productive.[2] In this paper, we present five stylized facts revealing that this is not the case if we focus on mixed exporters: firms that are active in both ordinary and processing exports. We show that mixed exporters, who contributed to over 60% of China's total processing exports in 2005, are larger and have higher labor, revenue, and physical productivity compared to other exporters. Importantly, we document that these "super

[1] We compiled these statistics using data from the Shenzhen Statistical Yearbooks.

[2] These studies include, but are not limited to, Fernandes and Tang (2015), Yu (2015), Dai et al. (2016), and Kee and Tang (2016).

processors" are 'mixed' not because they sell different products under different trade modes: for most mixed exporters, the majority of exports consists of the same product being sold to the same destination under both processing and ordinary trade modes.

Interestingly, even though being highly processing-oriented, mixed exporters' superior performance does not generalize to pure processing exporters. Compared to pure ordinary exporters, pure processors have significantly lower labor productivity and revenue productivity but greater employment and physical productivity. We empirically rule out the conjecture that the low labor and revenue productivity of pure processors are due to preferential processing policies or transfer pricing. Using a rich sample of *transaction-level* 2018 customs data with detailed product and brand information, we find (i) a link between selling one's own branded product and the use of ordinary trade mode, and (ii) a price premium associated with selling one's own branded product. This result suggests that a firm's choice of trade mode not only reflects its position inside a value chain but also its efficiency across stages of production, which ultimately affect its measured performance along various margins.

We hypothesize that pure processing firms are competent in manufacturing, but they lack the attributes necessary for non-production activities such as R&D, marketing, and branding, while ordinary firms are the opposite. This type of two-way heterogeneity rationalizes the coexistence of mixed, pure ordinary, and pure processing firms alongside their efficiency ranks, and why processing firms specialize in low value-added stages of production, supplying tasks to other firms. In line with this hypothesis, we also show that conditional on employment, pure ordinary firms are the most R&D- and advertising-intensive firms, followed by mixed exporters and then by pure processors.

Motivated by the above stylized facts, we develop a parsimonious model building on the frameworks of Antràs et al. (2017) and Bernard et al. (2019). Specifically, we allow firms to differ in two dimensions: (i) blueprint quality, which determines how good a firm is in developing its own branded final good, and (ii) manufacturing ability, which determines a firm's productivity in producing tasks, both for itself and for other firms. We let firms compete monopolistically in the final goods market and à la Bertrand in the tasks market as in Bernard et al. (2003) (BEJK hereafter), and thus they charge positive markups in both stages of production. In equilibrium, firms with good blueprints self-select into the final goods market, firms with high manufacturing ability self-select into the tasks market, and firms that have high attributes in both dimensions choose to be active in both markets. With international trade, only 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 we endogenise firms' specialization within a production network, the mass of potential suppliers is no longer exogenous as in Antràs et al. (2017) or Bernard et al. (2019). In this environment, our model generates a new positive externality of processing trade policy: facilitating processing trade raises the *ex-ante* expected profits from task production and hence encourages entry, leading to a greater mass of potential suppliers in equilibrium. When the mass of potential suppliers increases, final good producers' sourcing capacity improves, which in turn leads to lower marginal costs for those firms. Ordinary firms with high blueprint quality but low manufacturing ability (in our model, these firms also have the highest revenue productivity) benefit the most from the policy, as they rely heavily on task suppliers.

To empirically examine the model's prediction, we use China's pilot "paperless" processing supervision program in 2000-2006 as a quasi-natural experiment. As is well known, Chinese authorities closely supervise processing trade because of the special duty drawbacks and tax rebates provided to processing firms, who are required to fill out burdensome paperwork for each processing contract. The paperless program was aimed to make the supervision more efficient by eliminating paperwork through connecting firms' computer management systems to the customs' online administration system. This policy change is highly suitable for our identification strategy as it affects *only* the cost of processing exports. Moreover, the pilot program is experimental in nature and is adopted only by a few regional customs authorities at different times, limiting the scope of anticipation effects.

Before testing the model's central prediction on downstream spillovers, we first examine whether the policy was effective in increasing processing firms' exports. We find that processing exports of firms just above the qualification threshold (\$10-11 million) increased by 27% compared to the processing exports of firms just below the threshold (\$9-10 million) due to the policy. We then turn to the externality on ordinary firms and find that the policy induced downstream firms to intensify their branding activities. Consistent with the model's prediction, our results indicate that the pilot "paperless" program increased the number of trademarks for above-median productive ordinary firms by 0.24 on average, which explains about 10% of the average number of trademarks in the sample. Our results are robust to alternative and more restrictive empirical specifications and various falsification tests, and highlight that processing trade not only leads goods to be "Made in China," but also "Created in China" by providing a breeding ground of potential task suppliers for firms with good ideas.

This paper is related to several strands of the international trade literature. First, the stylized facts on mixed exporters documented in this paper are related to a large body of work on the characteristics of processing exporters in China (Fernandes and Tang, 2015; Yu, 2015; Dai et al., 2016; Li et al., 2018).[3] Different from these studies which focus more on pure processing firms, we document the dominant role of mixed exporters that engage in both ordinary and processing exports.[4] By revisiting some of the earlier findings but focusing on mixed exporters, and linking for

[3]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. However, emphasizing that revenue-based total factor productivity (TFP) calculations are confounded by price effects, and thus do not reflect "true" productivity, Li et al. (2018) calculate TFP based on quantity data and find that processing exporters are significantly more productive than non-exporters.

[4]Yu (2015), Dai et al. (2016), and Li et al. (2018) do include mixed firms (referred to as "hybrid" firms) in their analysis but do not focus on them.

the first time Chinese firms' own brand availability with their choice of trade mode, we rationalize the coexistence of pure processing, pure ordinary, and mixed exporters as well as their performance ranks across different margins.[5]

Regarding theory, our model extends the works of Antràs et al. (2017) and Bernard et al. (2019) and thus is related to the literature on firms' sourcing decisions in international and regional trade (Antràs et al., 2017; Tintelnot et al., 2018; Bernard et al., 2019; Kikkawa et al., 2019).[6] We contribute to the literature by considering a setting where upstream and downstream efficiencies are both embedded in firms, who self-select into different stages of the value chain. This specialization pattern explains well the observed firm-level performance across different exporters and yields a new source of gains from processing trade. In terms of modeling firms with multiple heterogeneities in the context of international trade, our paper is also related to Antràs and Helpman (2004), Hallak and Sivadasan (2013), Harrigan and Reshef (2015), and Bernard et al. (2018), among others.

Our empirical examination of the impact of "paperless" processing supervision program is related to the literature on trade facilitation. Focusing on how behind-the-border procedures influence trade flows, researchers find effects comparable or even higher than the ones found for border measures such as tariffs (Freund and Rocha, 2011; Hoekman and Nicita, 2011; Portugal-Perez and Wilson, 2012; Beverelli et al., 2015; Umana-Dajud, 2019).[7] More related to our study, the literature has recently began to examine the impact of the internet on international trade, finding substantial effects on firm-level exports (Kneller and Timmis, 2016; Fernandes et al., 2019) and imports (Malgouyres et al., 2019).[8] We contribute to this literature by providing the first look at the effect of digitization on processing trade, alongside its downstream spillovers.

Lastly, the evaluation of processing policy in this paper is broadly linked to the quantitative literature on the welfare implications of processing trade (Defever and Riaño, 2017; Deng, 2017;

[5]The literature has also studied the factors which determine the selection between ordinary and processing trade. Factors that are emphasized include preferential policy for processing exports (Dai et al., 2016; Brandt and Morrow, 2017; Defever and Riaño, 2017; Deng, 2017), foreign firms' outsourcing decisions (Feenstra and Hanson, 2005; Fernandes and Tang, 2012), and credit constraints (Manova and Yu, 2016). However, none of these frameworks generate the coexistence of the three types of exporters (ordinary, processing, and mixed) alongside the performance ranks we observe in the data.

[6]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, Tintelnot et al. (2018), Bernard et al. (2019), and Kikkawa et al. (2019) adapt the framework of Antràs et al. (2017) to the context of domestic production networks and study how endogenous firm-to-firm connections, geography, and markups affect shock transmissions and firm performance, respectively. Chaney (2016), Bernard and Moxnes (2018), and Johnson (2018) provide excellent reviews of the network models in international trade.

[7]Freund and Rocha (2011) find that transit delays have substantially hindered Africa's exports; Hoekman and Nicita (2011) show that behind-the-border measures such as logistics performance are important determinants for developing country exports; Portugal-Perez and Wilson (2012) show that the quality of both physical infrastructure and business environment play a role in developing country exports; Beverelli et al. (2015) find that the WTO's trade facilitation agreement has induced developing countries to diversity their export portfolio, both in terms of products and destination markets; and Umana-Dajud (2019) shows that visa-requirements reduce bilateral flows.

[8]Kneller and Timmis (2016) find a causal effect of broadband use on business service exports of firms in the UK; Fernandes et al. (2019) show that the expansion of internet in Chinese provinces caused firm-level exports to rise in 1999-2007; and Malgouyres et al. (2019) find that the roll-out of broadband in France has substantially increased firm-level imports and thus reduced the consumer price index in 1997-2007.

Brandt et al., 2019). Models in this literature characterize processing policy as either input-tariff exemptions or export subsidies, and thus the welfare effect of promoting processing trade depends on the specific choice model and parameter values.[9] Instead of quantifying the aggregate welfare by relying on a particular model, we analyze the impact of an *observed* policy change (i.e., "paperless" supervision of processing trade). Our empirical study highlights the positive spillover effect of processing trade on their domestic downstream customers, as suggested by our model.

The rest of the paper is organized as follows. In Section 2, we discuss the data, demonstrate the prevalence of mixed exporters, and present stylized facts on the relationship between firm characteristics, trade mode, and brand ownership. Section 3 presents our benchmark model, and Section 4 describes the open-economy version and links it to the data. In Section 5, we empirically test our model by studying the impact of the pilot "paperless" supervision program for processing trade. Finally, we conclude in Section 6.

2 Data and Stylized Facts

To develop a complete understanding of the relationship between trade mode and characteristics across firms and transactions, we use six datasets in this paper. 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.[10] The 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 ordinary and processing exports.

The second and the third datasets we use are the annual industry survey (AIS) and production survey compiled by China's National Bureau of Statistics (NBS) for 2000-2006. The AIS dataset 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 for all state-owned enterprises (SOEs) and private firms that have annual sales of at least five million RMB.[11] The production survey contains firm-product level information on output quantity, which enables us to compute firm-level quantity-based (i.e., physical) TFP.[12] We follow the common methodology

[9] In Deng (2017), the welfare loss is due to the decreased number of varieties for consumers. Defever and Riaño (2017) analyze the welfare implications of subsidies with export share requirements in a quantitative trade model. They evaluate a policy that requires firms to export more than a certain share of their output to be able to qualify for the subsidy. While this policy greatly promotes exports, it exacerbates the welfare loss associated with greater protection to low-profitability firms. Brandt et al. (2019), on the other hand, find that tariff exemptions on imported inputs for processing has a positive welfare impact but they do not allow processing firms to sell domestically.

[10] 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."

[11] 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.

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

used in the literature, which includes Wang and Yu (2012), Fan et al. (2015), and Manova and Yu (2016), to merge the three datasets based on firm names, telephone numbers, and zip codes. Our matching procedure results in covering about 58% of aggregate exports, which is similar to the aforementioned papers' match rates.[13]

The fourth dataset we use is a rich sample of confidential *transaction-level* customs data for 2018. This unique dataset contains 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.[14] 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 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. Of the 34 products, 30 are from March and the rest are from January and April 2018. The wide variety of products, which are listed in Appendix Table C.1, includes goods that make up a large share of exports such as car tires, refrigerators, and mobile phones.

To empirically test our model's prediction in Section 5, we utilize a fifth dataset on firm-level 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. (2018) to obtain the number of effective trademarks per firm-year.[15] Finally, using China's publically available official customs notices, we constructed a novel dataset on the dates when each Chinese regional customs authority adopted the pilot "paperless" processing trade program, which is discussed in more detail in Section 5.

2.1 Mixed Exporters in China

We define mixed exporters as firms that engage in both processing and ordinary exports. These firms, which are also called "hybrid," are considered to be "perhaps the most interesting type of firm[s]" (Yu, 2015), but were never investigated carefully in the literature. In this section, we add to the existing stylized facts on China's exporters and begin by unpacking the "black box" of mixed exporters.

Similar to the figures presented by Dai et al. (2016), we find that even though the number of mixed exporters was only 21% of the total number of exporters, they made up 54% of exports in

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

[14] This policy change was issued as 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.

[15] We are grateful to Ran Jing for sharing the data. See Deng et al. (2018) for a detailed description of the trademarks dataset.

2005. Pure processors and pure ordinary exporters, on the other hand, made up 24% and 19% of exports in 2005 respectively.[16]

In Table 1, we present firm-level statistics for mixed exporters where we have all 50,952 mixed exporters in panel (a), and 24,470 merged (with the AIS data) mixed exporters in panel (b). First, note that the figures in both panels are similar, indicating that the merged exporter sample is representative, 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 are similarly high, suggesting that mixed exporters' main activity is often 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, mixed firms are prevalent in almost all sectors; 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, we find that there was at least one mixed exporter in 66 sectors. This result indicates that many "superstar" firms are mixed exporters.

Given that mixed exporters are the major players in Chinese exports, perhaps the literature has largely ignored them since researchers conjectured 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. Even though most mixed firms do export multiple products, they tend to sell their *core product*(s) under both trade regimes.[17] Consistent with Dai et al. (2016), Table 1 panel (b) rows 4 and 5 show that the *number* of products that are exported under both trade regimes, on average, make up only about 37% and 24% of mixed firms' total number of products and product-destinations respectively. However, the median (mean) *value* share of HS8 products that are exported through both ordinary and processing modes (mixed HS8) in a mixed firm's exports is 89% (71%), as reported in row 6 of panel (b). Row 7 analyzes the most disaggregate HS8-country level and finds a median (mean) share of 62% (55%), suggesting that the majority of mixed exporters' exports are due to selling the same product to the same destination via both trade modes.[18]

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,

[16]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 these figures exclude intermediaires and wholesalers, which made up 18% of exports in 2005. A closer look at the customs data reveals that for every year in 2000-2006, mixed exporters made up the majority of China's exports, followed by pure processors and pure ordinary firms in terms of contribution to aggregate exports, excluding intermediaires and wholesalers.

[17]The customs data reveals that mixed firms are more likely to be multi-product firms relative to pure ordinary or pure processing firms. However, as will be made clear, exporting multiple products is not a necessary nor a sufficient condition for being mixed.

[18]Also, the median (mean) share of exports destined for the same destination within products that are sold under both trade regimes (mixed HS8-country) is 98% (78%) for the merged mixed exporters.

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). Similarly, under HS6 841112 *Turbojets, of a thrust* $>$ *25KN*, there are three HS8 varieties: *turbofan engines*, *turbojets with propulsive force* $\geq$ *90KN*, and *turbojets with propulsive force* $>$ *25KN and* $<$ *90KN*. This level of detail mitigates the concern that an exporter is mixed due to its multi-product nature. We summarize the above findings in our first stylized fact:

Fact 1: The majority of most mixed firms' exports consists of the same product being sold under both ordinary and processing modes to the same destination.

We find it intriguing to think about the non-trivial existence of mixed exporters. The theoretical literature typically assumes either that processing is a different sector (Deng, 2017; Brandt et al., 2019) or that Melitz-type firms sort themselves into processing versus 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. However, in that case, mixed exporters would never sell the same product to a given destination via both trade modes, which is in contrast to our *Fact 1*. Moreover, both types of models would predict that mixed firm attributes should lie between that of processing and ordinary firms. However, as detailed in the next subsection, this is not what we find in the data.

2.2 Trade Mode and Firm Characteristics

Following the well-established literature on exporter premia pioneered by Bernard and Jensen (1995, 1999, 2004), we start by investigating whether firms that engage in different export modes have significantly different characteristics. Our exercise in this section is similar to Dai et al. (2016), but while they mainly focus on comparing exporters to non-exporters, we focus on differences *between* exporters. Note that 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 *Fact 4*, 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. Panel (b) excludes firms with foreign ownership since these firms are larger and more likely to be processors (Yu, 2015; Dai et al., 2016).[19]

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. This gives us the second fact:

Fact 2: Mixed exporters are larger than other exporters in terms of employment.

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 ($TFPR$). Based on *Fact 2*, 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.[20] Row 3 shows that the ranking we obtained based on labor productivity remains when we consider $TFPR$ calculated using the Olley-Pakes (1996) methodology.[21] The results are similar when we exclude foreign firms in panel (b) row 2.

As is well known, $TFPR$ reflects not only production efficiency but also firms' pricing behavior. In particular, focusing mainly on the leather shoes industry, Li et al. (2018) estimate quantity-based TFP ($TFPQ$) and find that exporters' efficiency is higher than non-exporters' when one uses $TFPQ$ instead of $TFPR$. They argue that processing exporters' low $TFPR$ can be explained by their relatively lower average export prices. Does this empirical regularity hold for other sectors? What is the place of mixed exporters in the $TFPQ$ rank? With these two questions in mind, we replicate their exercise focusing on the 36 of the 693 manufacturing 5-digit products in the dataset for which we can obtain reliable estimates based on data availability. The estimation methodology that uses the subsample of single-product firms and the list of our 36 products can be found in Appendix Section B and Table B.1 respectively.[22] Consistent with Li et al. (2018), we find that compared to pure ordinary exporters, pure processors have higher $TFPQ$ on average as shown in row 4 of Table 2 panel (a). Our results also show that mixed exporters have the highest physical productivity on average (though not statistically significantly different from that of pure processors). We summarize these findings in the following stylized fact:

[19] About 27% of exporters are foreign-owned. Not surprisingly, this share is larger (41%) for processing exporters.

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

[21] The Olley-Pakes methodology allows us to control for sample selection as well. Our results are robust to using the Levinsohn-Petrin (2003) approach to calculating $TFPR$.

[22] Our methodology is similar to the one used by Li et al. (2018) but slightly differs 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.

Fact 3: Mixed exporters have higher labor and revenue productivity than pure ordinary and pure processing exporters. However, processing exporters, regardless of being mixed or not, have higher physical productivity than pure ordinary exporters.

As mentioned before, existing theoretical frameworks would predict that mixed firm characteristics lie between that of pure processing and pure ordinary firms, which is not 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 $TFPR$. Nevertheless, the fact that production efficiency ($TFPQ$) ranking follows a (weakly) decreasing order of mixed exporters, pure processors, and pure ordinary exporters remains to be explained.

An alternative hypothesis is that processing firms are contracted by foreign firms to 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 profits from their economic activities beyond production, and hence can have higher $TFPR$ even with relatively low $TFPQ$. This view also gives a natural explanation to the existence of mixed exporters: firms that excel in both manufacturing and associated non-manufacturing activities.[23]

While the above two explanations are both plausible, which one plays a dominant role is an empirical question. We address this by using the 2018 customs sample and reach the following stylized fact, which we explain subsequently:

Fact 4: Ordinary exporters tend to sell their own branded products, whereas processing exporters tend to sell their customers' branded products. There is a price premium associated with selling one's own branded product.

The 2018 customs dataset allows us to extract the brand and ownership information for each transaction from the reported product specification, and label it as no brand, foreign brand, or domestic (own) brand.[24] As shown in the last row of Appendix Table A.1, 12.4%, 56.4%, and 32.7% of export value are due to transactions that have no brand, foreign brand, and domestic

[23]Related to the second hypothesis, one prevalent yet little documented fact is that many prominent Chinese firms produce their own branded products while at the same time manufacture goods for other firms (Deng, 2017). 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.

[24]To do this classification, we write a simple algorithm that breaks down the detailed string product specification into brand name and ownership information. More specifically, we split the product specification according to certain delimiters such as tabs and commas to determine brand ownership status (the Chinese customs uses '0' for no brand, '1' for domestic brand, '2' for purchased brand, '3' for original equipment manufacturer brand, and '4' for foreign brand). For transactions with missing brand ownership indicator, we use the reported brand name and compare it to the firm's name to infer brand ownership, in some cases manually. As in our previous analysis, we exclude intermediaries and wholesalers; this reduces the number of transactions from 862,567 to 591,270.

brand respectively. Regarding trade modes, the dataset reports 45 export modes which we classify into three groups: ordinary exports, processing exports, and "other" exports.[25] Based on our classification, ordinary and processing exports account for 27% and 43% of total exports in the sample.

While processing transactions are typically viewed as local manufacturers supplying customized tasks to their "branded" buyers (Manova and Yu, 2016), the link between processing trade and brand ownership has not been established empirically. Appendix Table A.1 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., FDI policy), and ϵ_{ifhc} is the error term which we cluster at the firm level. Table 3 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 (mixed) firm in January, March, or April 2018.[26] Column 2 shows that the coefficient remains negative and significant at the 10% level, albeit with a lower magnitude (-0.032).

In column 3, we regress the log unit value of transactions on brand ownership, controlling for trade 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 owned product of a firm is about 9% more expensive than that same firm's sales of the same product destined for the same destination but under a customer's brand (significant at the 5% level).

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

[25] We treat the export mode as processing exports whenever "processing" appears in the string variable. Following this rule, we group seven export modes into processing exports. We treat the export mode named as "ordinary exports" as ordinary exports. We lump the categories that we were not able to classify into "other" exports. Our results are robust if we include other potential ordinary exports "temporary exports," "foreign contracted exports," "goods for exhibition," and "samples for advertisement" in the ordinary exports category. This group of exports make up 0.01% of total exports in the customs sample.

[26] 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 percentage of mixed firm-product-country flows make up 15% of total flows, with the rest consisting of pure ordinary (51%) and pure processing flows (34%).

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 3. If transfer pricing is driving the results (i.e., processing exporters artificially depressing the price of export transactions between enterprises under common ownership or control), then we would expect to see a relatively less stark difference on $TFPQ$ between processing and ordinary firms once we exclude foreign firms—the results in Table 2 suggest the opposite. Therefore, our results indicate that the higher average price of exporters' own products is indeed a brand premium instead of reflecting input tariff exemptions or transfer pricing.

Finally, we provide some suggestive evidence that a firm's choice on trade 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.[27] 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 stay qualitatively the same.[28] This gives us our final stylized fact:

Fact 5: Pure ordinary exporters spend more on R&D and advertising than mixed exporters, who in turn spend more than pure processing exporters.

To sum up, these five stylized facts lead us to view mixed exporters as "super processors," and motivates our investigation of how firms' efficiency in manufacturing versus non-manufacturing activities could determine their specialization inside a production network. In the next section, we develop a theoretical model that could explain these stylized facts, and in which "super processors" naturally emerge.

3 Theoretical Framework

To our knowledge, the phenomenon that manufacturing firms produce and export both their own and their customers' branded products is not presented in either the theoretical or the empirical literature on international trade. In the existing theoretical literature, brand differentiation has been modeled through nested constant elasticity of substitution (CES) preferences where the first nest is defined over brands and the second nest over multiple products within each brand (see, for example, Allanson and Montagna, 2005). However, these models restrict firms to produce for their own brand. Our model differs from these studies by allowing firms to be suppliers and thus sell

[27] This ranking for R&D can also be inferred from Table 7 column 5 of Dai et al. (2016).

[28] 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.

their products under their customers' brands.

Based on stylized fact 4, we let the choice of trade mode to implicitly determine firms' value chain position, and aim to present a model that also explains facts 1-3 and 5. Note that there is nothing intrinsically international about our model—a non-exporting manufacturer could also produce its own branded products and at the same time serve other firms. Hence we start from a closed economy setting and focus on firms' choice between *making* or *creating* (i.e., branding). Then, we turn to an open economy setting to discuss how our model fits the data and also to explore the implications of processing trade policy.

3.1 Basic Environment

Our model is based on Antràs et al. (2017) and Bernard et al. (2019), where we extend their framework by allowing heterogeneous firms to differ in both blueprint quality and manufacturing ability. In this model, firms endogenously determine the set of tasks they produce for other firms, the optimal production of their own final good, and the related sourcing strategy.

Preferences of a representative consumer 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 given by $\sigma > 1$.

There is a continuum of firms, and each owns a blueprint to produce a single differentiated variety. Production of 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 between the task bundle and final good production: the higher the z_j, the more productive firm j is in producing the final good.

Task production requires only labor, which is inelastically supplied at the country level. 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. Here, t_j governs the firm's average manufacturing ability in producing tasks, and θ the (inverse) dispersion of its manufacturing efficiency across tasks.

As blueprint-holders, firms can produce some tasks in-house and source some from other firms. Analogously, firms as task suppliers 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 efficiency in producing tasks tailored for its blueprint.

We assume Bertrand competition in task production following BEJK (2003). As a result, 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 crucial for our analysis on processing policy, but the rest of our results hold if we assume perfect competition in the tasks market instead. Under Bertrand competition, conditional on the set of suppliers that firm j has established relationships with, each tailored task is supplied by the lowest-cost supplier. This supplier is constrained to charge not more than the second-lowest cost supplier of that task. In order for prices and markups to be analytically tractable at the aggregate level, we impose an additional constraint: if supplier i produces task κ for firm j at an efficiency level $\phi_{1i}(\kappa)$, then this same supplied task can be 'mimicked' by j with $\phi_{2i}(\kappa) \leq \phi_{1i}(\kappa)$, with the joint distribution of $\phi_{1ij}(\kappa)$ and $\phi_{2ij}(\kappa)$ given by: $F_i(\phi_1, \phi_2) \equiv Pr[\phi_{1ij}(\kappa) \leq \phi_1, \phi_{2ij}(\kappa) \leq \phi_2] = [1 + t_i(\phi_2^{-\theta} - \phi_1^{-\theta})]e^{-t_i\phi_2^{-\theta}}$.[29]

There is an unbounded pool of prospective entrants. Firms 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.[30] Once firms make their draws, they decide to (i) exit, (ii) engage in blueprint production, (iii) engage in task production, or (iv) do both (ii) and (iii). Being active in blueprint production requires an additional fixed cost f_B. An active firm faces a constant probability δ of an adverse shock that would force it to exit every period.

In our model, firms differ in two dimensions: blueprint quality z, which indicates how good their *idea* or *brand* is, and manufacturing ability t, which determines how good they are in *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*.

3.2 Optimal Sourcing

Note that manufacturing ability t varies across firms while the relationship-specific investment f does not. As a result, a firm's optimal sourcing decision is simplified to choosing the least productive supplier it is willing to reach. As there is no fixed cost of task production, all firms are potential suppliers. Conditional on firm j being connected with i, the probability that i is the lowest-cost supplier to j for a particular task is:

$$\lambda_{ij} \equiv \lambda(z_j, t_j, t_i) = \frac{t_i}{\Theta(z_j, t_j)}, \tag{3}$$

where $\Theta(z_j, t_j)$ measures firm j's "sourcing capacity."[31] Specifically:

$$\Theta(z_j, t_j) = t_j + N \int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota), \tag{4}$$

[29] In principle, the price of an in-house produced task should equal its marginal cost. However, for tractability, we allow for positive markups for both in-house and outsourced production. One can view in-house production as equivalent to firms sourcing tasks from their quasi-independent manufacturing subsidiaries.

[30] 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.

[31] This is synonymous with "sourcing capability" in Antras et al. (2017).

where $\underline{t}_j$ is the least productive supplier that firm j sources from, and N is the endogenously determined mass of entrants.[32] As in BEJK, firm j's share of total purchases from firm i equals λ_{ij}. The price of the task bundle used by firm j, P_j^T, is given by $P_j^T = \Theta_j^{\frac{-1}{\theta}}\gamma$, where γ is a constant.[33] Thus, a firm's marginal cost of producing its own final good, c_j, is simply $\frac{P_j^T}{z_j}$. Conditional on firm j's own manufacturing ability t_j, sourcing from a larger number of suppliers leads to a lower marginal cost. Conditional on the number of suppliers, having a higher manufacturing ability also enables the firm to produce at a lower marginal cost.

Conditional on its pricing strategy, the final good producer with blueprint z_j and manufacturing ability t_j chooses the set of suppliers to maximize its profits from final good production:

$$\max_{\underline{t}_j} \pi^B(z_j, t_j) - fn(z_j, t_j) - f_B, \tag{5}$$

where $\pi^B(z_j, t_j) = Ak_1\Theta^{\frac{\sigma-1}{\theta}} z_j^{\sigma-1}$ is the operating profits, with $A = \beta L P^{\sigma-1}$ being the demand shifter, L the country's labor endowment, P the aggregate price index, and k_1 a constant.[34] The number of suppliers, $n(z_j, t_j)$, is given by $N \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota)$. Solving this maximization problem yields the optimal $\underline{t}_j$, which satisfies:

$$\underline{t}(z_j, t_j) = f\left(Ak_1 z_j^{\sigma-1}\right)^{-1} \Theta(z_j, t_j)^{1-\frac{\sigma-1}{\theta}} \frac{\theta}{\sigma-1}. \tag{6}$$

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. Intuitively, firms who have good ideas reach a greater number of suppliers, while firms who are efficient in producing tasks themselves reach fewer suppliers.

The blueprint-owner receives the profits generated by 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 can be derived à la Melitz (2003), respectively:

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

Firms also receive profits by producing tasks. From BEJK, the profits from task production are given by:

$$\pi^T(t_i) = \frac{1}{1+\theta} \sum_{j\in\Omega_i} x_{ij}(z_j, t_j, t_i), \tag{7}$$

[32] As there is no fixed cost for in-house production, a firm, conditional on having decided to bring its blueprint to production, will always produce some tasks in-house. 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.

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

[34] $k_1 = \frac{\gamma^{1-\sigma}}{\sigma}(1-\frac{1}{\sigma})^{\sigma-1}$.

where $x_{ij} = \dfrac{\sigma - 1}{\sigma} \lambda_{ij} r_j^B$ is firm j's task purchases from firm i, and Ω_i is the set of firms that source from i.[35]

3.3 Equilibrium in the Closed Economy

In equilibrium, the zero-profit condition for final good production is given by:

$$\pi^B(z, t) - fn(z, t) = f_B, \tag{8}$$

where π^B and n now stand for the optimal operating profits and the number of suppliers of a firm with blueprint z and manufacturing ability t, respectively. Rewriting t as a function of z, equation (8) gives the cutoff *curve* above which firms choose to bring their blueprint to production, which we denote as:

$$t = \Xi(z). \tag{9}$$

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 have a good blueprint. Denoting the 'worst' blueprint that is brought to production as $\underline{\mathrm{z}}$, we have:

$$\underline{\mathrm{z}} = \Xi^{-1}(\bar{t}). \tag{10}$$

As the number of task suppliers that firm j matches with increases in z_j and decreases in t_j, 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 (6), we obtain the cutoff $\underline{\mathrm{T}}$, above which firms are active in supplying tasks:

$$\underline{\mathrm{T}} = \underline{\mathrm{T}}(\bar{z}, \Xi(\bar{z})) = [Ak_1 \bar{z}^{\sigma-1}]^{-1} \Theta(\bar{z}, \Xi(\bar{z}))^{1 - \frac{\sigma-1}{\theta}}. \tag{11}$$

Therefore, the mass of active task suppliers and final good producers are respectively given by:

$$N^T = N \int_{\underline{\mathrm{T}}}^{\bar{t}} dG_t(\iota). \tag{12}$$

$$N^B = \int_{\underline{\mathrm{z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} dG_t(\iota) dG_z(\zeta). \tag{13}$$

The aggregate price index equals:

$$P = N^{\frac{1}{1-\sigma}} \left[\int_{\underline{\mathrm{z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p^B(c(\zeta, \iota))^{1-\sigma} dG_t(\iota) dG_z(\zeta) \right]^{\frac{1}{1-\sigma}}. \tag{14}$$

[35] Note that π_i^T does not depend on firm i's blueprint quality—the only interaction between z_i and t_i in our model is through in-house production.

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

$$\int_{\underline{\mathrm{z}}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} (\pi^B(\zeta,\iota) - fn(\zeta,\iota) - f_B) dG_t(\iota) dG_z(\zeta) + \int_{\underline{\mathrm{T}}}^{\bar{t}} \pi^T(\iota) dG_t(\iota) = \delta f_E. \tag{15}$$

Given optimal sourcing strategies, cutoff conditions (9), (10), and (11), firm mass equations (12) and (13), the price index (14), and the free entry condition (15) give us seven equations and seven unknowns: P, N, N^B, N^T, $\Xi(z)$, $\underline{\mathrm{z}}$, and $\underline{\mathrm{T}}$. We now formally define the equilibrium of the model:

DEFINITION 1. *The closed economy equilibrium consists of an aggregate price index P, the number of firms N, N^B, N^T, and the cutoffs $\Xi(z)$, $\underline{z}$, and $\underline{T}$ that satisfy the equilibrium conditions (9)-(15).*

3.4 Analysis of the Equilibrium

Comparative Statics

To facilitate our later analysis, we first present four sets of comparative statics regarding firms' sourcing choice, marginal costs, and profits. We relegate proofs to Appendix D.1 and D.2.[36]

- $\frac{\partial c(z_j,t_j)}{\partial z_j} < 0$, $\frac{\partial o(z_j,t_j)}{\partial z_j} > 0$, $\frac{\partial n(z_j,t_j)}{\partial z_j} > 0$, $\frac{\partial \underline{\mathrm{t}}(z_j,t_j)}{\partial z_j} < 0$. Here, $o(z_j,t_j)$ is the share of tasks a firm sources from suppliers. From the perspective of blueprint producers, firms with better blueprints have lower marginal costs, and source a larger share of tasks from a larger number of suppliers. They also incur more fixed costs and reach less efficient suppliers.

- $\frac{\partial c(z_j,t_j)}{\partial t_j} < 0$, $\frac{\partial o(z_j,t_j)}{\partial t_j} < 0$, $\frac{\partial n(z_j,t_j)}{\partial t_j} < 0$, $\frac{\partial \underline{\mathrm{t}}(z_j,t_j)}{\partial t_j} > 0$. Blueprint producers that are more efficient in task production have lower marginal costs, and source a lower share of tasks from from fewer suppliers. They are also less likely to incur fixed costs to reach less efficient suppliers.

- $\frac{\partial c(z_j,t_j)}{\partial N} < 0$, $\frac{\partial o(z_j,t_j)}{\partial N} > 0$, $\frac{\partial n(z_j,t_j)}{\partial N} \lesseqgtr 0$, $\frac{\partial \underline{\mathrm{t}}(z_j,t_j)}{\partial N} > 0$. When the number of potential suppliers increases, firms source more tasks and their marginal costs decrease. Moreover, when N increases, the mass of productive suppliers increases and thus the cutoff manufacturing ability rises. The more productive suppliers supply more tasks than less productive ones, and therefore the number of total suppliers a firm has might increase or decrease.

- If $t_i > t_{i'}$, then $x_{ij}(z_j,t_j;t_i) > x_{i'j}(z_j,t_j;t_{i'})$, $\Omega_{i'} \subset \Omega_i$, $\pi^T(t_i) > \pi^T(t_{i'})$. From the perspective of task producers, firms 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

[36] We assume throughout the paper that $\theta > \sigma - 1$: the efficiency dispersion across tasks is relatively higher than the elasticity of substitution between varieties.

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.

Selection into Operating Status

In equilibrium, equation (9) determines the zero-profit curve (ZPC), which gives the $\{z, t\}$ combinations above which firms are active in final good production; similarly, equation (11) determines the cutoff $\underline{\mathrm{T}}$ above which firms are active in task production. As such, we can visualize firms' selection into four different activities in Figure 1: (i) firms with both low z and t exit, (ii) firms with high z but low t engage in final good production and become *pure ordinary* firms, (iii) firms with low z but high t engage in task production and become *pure processing* firms, and (iv) firms with both high z and t do both activities and become *mixed* firms.

Uniqueness

Because of Cobb-Douglas preferences and the constant markups on both task and final good production, total profits from task production is exogenously given by $\Pi^T = \frac{(\sigma-1)\beta L}{\sigma(\theta+1)}$. Denoting the average net profits of active blueprint producers as $\tilde{v}^B$, the free entry condition can be rewritten as:

$$\tilde{v}^B(P, N) = \frac{N\delta f_E - \frac{(\sigma-1)\beta L}{\sigma(\theta+1)}}{N^B(P, N)}. \tag{16}$$

The increase in N is associated with a decrease in N^B/N. The intuition is that when N increases, competition in the final goods market intensifies. This forces the least productive firms to exit, and therefore lowers the share of active final good producers.

In equilibrium, given P and N, firms' operating decisions are uniquely determined. Hence the system of equations that characterize the equilibrium can be simplified to two equations linking P and N—the free entry (FE) condition (16), and the aggregate price (AP) equation (14). 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 2 and formally prove in Appendix Section D.3.

4 Open Economy

We now turn to the open economy case where we have two countries: Home and Foreign (denoted with asterisk). The differentiated sector is subject to iceberg trade costs such that $\tau_B, \tau_T > 1$ units have to be shipped for one unit of final goods and tasks to reach the destination, respectively.

Exporting final goods also requires a fixed cost f_X, and we assume that the homogeneous sector is freely traded.

The equilibrium can be solved analogous to the closed economy case. Consider first the final goods market. The blueprint producers at Home face $N + N^*$ potential task suppliers, whose manufacturing abilities are distributed as $\hat{G}_t(\iota) = \frac{N}{N+N^*}G_t(\iota) + \frac{N^*}{N+N^*}G_t^*(\iota\tau_T^{*\theta})$.[37] Here, we use subscripts D and X to denote firms' export decisions: D stands for selling domestically and X stands for exporting. Then, a blueprint producer's optimization problem can be written as:

$$v^B(z_j, t_j) = \max_{\{D,X,Exit\}} \left\{v_D^B(z_j, t_j), v_X^B(z_j, t_j), 0\right\}, \tag{17}$$

where:

$$v_D^B(z_j, t_j) = \max_{\underline{t}} \left\{\pi_D^B(z_j, t_j) - fn(z_j, t_j) - f_B\right\},$$

$$v_X^B(z_j, t_j) = \max_{\underline{t}} \left\{\pi_D^B(z_j, t_j)(1 + \frac{A_F}{A}\tau_B^{1-\sigma}) - fn(z_j, t_j) - f_B - f_X\right\}.$$

It is easy to verify that v_D^B and v_X^B are both upward sloping in z, with v_X^B being steeper since: (i) market access is greater, and (ii) improved market access leads to a lower marginal cost of production via increased optimal sourcing. Therefore, among two firms that have the same manufacturing ability, the one with the better blueprint quality is more likely to export.

Firms are indifferent between selling domestically and exporting when $v_D^B = v_X^B$. This yields the export cutoff curve $t_X = \Xi_X(z)$. Expressions for the zero-profit curve and the domestic task production cutoff remain the same as in the closed economy equilibrium. The only change is that the sourcing capacity of a given firm becomes $\Theta(z_j, t_j) = t_j + (N+N^*)\int_{\underline{t}(z_j,t_j)}^{\bar{t}} \iota d\hat{G}_t(\iota)$ as it can now reach task suppliers in both countries. The number of active final good producers and exporters in Home can then be expressed as:

$$N^B = \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} dG_t(\iota)dG_z(\zeta), \qquad N_X^B = \int_{\underline{z}_X}^{\bar{z}} \int_{\Xi_X(\zeta)}^{\bar{t}} dG_t(\iota)dG_z(\zeta),$$

where $\underline{z} = \Xi^{-1}(\bar{t})$ and $\underline{z}_X = \Xi_X^{-1}(\bar{t})$. The aggregate price index therefore satisfies:

$$P^{1-\sigma} = N \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p^B(\zeta, \iota)^{1-\sigma} dG_t(\iota)dG_z(\zeta) + N_X{}^* \int_{\underline{z}_X^*}^{\bar{z}} \int_{\Xi_X^*(\zeta)}^{\bar{t}} \left(\tau_B^* p^{B*}(\zeta, \iota)\right)^{1-\sigma} dG_t^*(\iota)dG_z^*(\zeta). \tag{18}$$

Next we turn to the tasks market. The supply of tasks mirrors the demand from the final goods market. With the presence of export costs, the task supplier with the lowest manufacturing ability

[37] Note that having iceberg trade costs is equivalent to increasing the average production cost of foreign task suppliers by $\tau_T^{*\theta}$.

in Home that exports to Foreign satisfies: $\underline{T}_X = \tau_T^{\theta}\underline{T}^*$. The free entry condition becomes:

$$\int_{\underline{Z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} v^B(\zeta, \iota) dG_t(\iota) dG_z(\zeta) + \int_{\underline{T}}^{\bar{t}} \pi^T(\iota) dG_t(\iota) = \delta f_E. \tag{19}$$

Given the mass of entrants and the aggregate price indices in both countries, firms' optimal sourcing and operating decisions can be determined; equations (18) and (19) provide four equations, from which we can uniquely identify N, N^*, P, and P^*, and hence solve the equilibrium.

4.1 Selection into Export Mode

Introducing international trade yields two additional cutoffs compared to the closed economy equilibrium. As presented graphically in Figure 3, 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 has both high manufacturing ability and blueprint quality, it becomes a mixed exporter. The following proposition summarizes this result:

PROPOSITION 1. *In equilibrium, firms with both low z and t exit, and firms with intermediate z or t operate solely domestically. Firms with high z but low t become pure ordinary exporters; firms with low z but high t become pure processing exporters; and firms with both high t and z engage in both activities and become* mixed *exporters.*

4.2 Linking the Model to Data

We now examine how the model can help us explain the stylized facts in Section 2. For the remainder of this section, we refer to pure processing exporters (PP) as processing exporters, pure ordinary exporters (PO) as ordinary exporters, and exporters that engage in both activities (Mix) as mixed exporters. 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:*

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

Physical TFP Physical TFP measures the efficiency of a firm in turning inputs into outputs in terms of quantities. In our model this is best captured by t_j, which reflects the average efficiency of a firm in task production. We first compare the t of mixed exporters with that of processing exporters. We use G_t^s to denote the cumulative distribution function (*cdf*) of t in equilibrium for

$s \in \{PP, PO, Mix\}$. As the export cutoff curve is downward sloping, and z and t are distributed independently, there are relatively more firms with lower t among processing exporters. This in turn implies that G_t^{Mix} FSD G_t^{PP}, a result we formally prove in Appendix Section D.4. Next, we compare processing exporters to ordinary exporters. The export selection cutoff ensures that t_{PO} is always lower than t_{PP}. Therefore G_t^{PP} FSD G_t^{PO}. Then, by applying the ranking theorem, we immediately have that $E_{Mix}[t] > E_{PP}[t] > E_{PO}[t]$.[38]

R&D and advertisement expenses Similarly, we use G_z^s to denote the *cdf* of z for $s \in \{PP, PO, Mix\}$. We first prove that G_z^{PO} FSD G_z^{Mix} and G_z^{Mix} FSD G_z^{PP} in Appendix Section D.4. As Figure 3 intuitively suggests, 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.

We then rationalize stylized fact 3 with a simple twist in our model. 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.[39] Note that 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. However, $f_{RD}(a)$ is increasing in z in equilibrium; thus by the ranking theorem, we immediately have that $E_{PO}[f_{RD}] > E_{Mix}[f_{RD}] > E_{PP}[f_{RD}]$.

Employment We first compare the employment of mixed and processing exporters. As labor is the only input, the associated employment increases in t. We assume that f is sufficiently small such that $\underline{t}_j < t_j$ always holds for exporters. As a result, we have that $E_{Mix}[l \mid t] = E_{PP}[l \mid t]$.[40] From the comparative statics, we know that the employment of a firm is increasing in t. Applying the ranking theorem, it is immediate that $E_{Mix}[l] > E_{PP}[l]$.

Next, we compare the employment of processing and ordinary exporters. Consider a processing exporter j and an ordinary exporter j'. As $t_j > t_{j'}$, for any final good producer that both j and j' supply tasks to, firm j has higher sales. Firm j also reaches a larger number of final good producers. Therefore, $x_j > x'_j$, which in turn implies that $l_j > l'_j$. As this inequality holds for any processing exporter j and ordinary exporter j', $E_{PP}[l] > E_{PO}[l]$ holds as well.

Labor productivity 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:

[38] Moreover, the comparative statics result suggests that sales to each customer, the number of customers, and the total profits from task production are increasing in t. The ranking theorem therefore implies that, on average, mixed firms have greater processing exports, reach more customers, and earn higher total profits from processing when compared to processing exporters.

[39] We assume that $f'_{RD} > 0$ and $f''_{RD} > 0$.

[40] This assumption greatly simplifies our analysis, resulting in no 'additional' labor for in-house production, as exporter j sources from all suppliers (including itself) with manufacturing ability t_j anyways.

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

We first compare processing and ordinary exporters. The export cutoff ensures that $v^B_{PO} > v^B_{PP}$ for any pair of firms. We also showed that $l_{PP} > l_{PO}$ always holds. Therefore, for any processing exporter j' and ordinary exporter j, $LP_j > LP_{j'}$, and thus $E_{PO}(LP) > E_{PP}(LP)$. The comparison between ordinary and mixed exporters is less obvious; in Appendix Section D.4, we show that $\frac{v_j^B}{l_j}$ is an increasing function of z_j, and *can* be an increasing function of t_j. If the latter holds, firms with the highest labor productivity will be mixed. When their share is sufficiently large, we have that $E_{Mix}(LP) > E_{PO}(LP)$.

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. The share parameter on labor is α and the revenue TFP of firm j is then given by:

$$TFPR(z_j, t_j) = ln\left(\frac{v_j^B + \pi_j^T + l_j}{{l_j}^{\alpha}{k_j}^{1-\alpha}}\right) \propto ln\left(\frac{v_j^B + \pi_j^T + l_j}{l_j}\right) = LP_j.$$ [41]

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

4.3 Processing Trade Policy

There is widespread belief among policymakers that exporting is beneficial for a country's economic development. As a result, many emerging countries, most notably China, adopted policies that encourage exporting such as processing trade policy that exempts exporters from paying input tariffs. Another processing trade policy championed mostly by East and Southeast Asian economies is the establishment of export processing zones that provide various incentives to processing exporters (Radelet and Sachs, 1997).[42] However, existing work typically suggests that promoting processing trade can also crowd out ordinary firms and thus reduce welfare (Defever and Riaño, 2017; Deng, 2017).[43] Does our framework provide any new insights regarding processing trade policy?

To highlight the model's novel prediction, we focus on a small open economy setting such that changes at Home does not change any aggregate variables of Foreign. Our model predicts that when Home introduces a processing policy that lowers τ_T, firms' task exporting opportunities increase. These opportunities raise the *ex-ante* expected value from task production, and thus firms' expected profits from final good production must decrease for the free entry condition to hold. Therefore, 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

[41] 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.

[42] Nevertheless, the Chinese customs data shows that 86% of processing exporters in 2000-2006 were located outside of special economic zones, which include export processing zones.

[43] The only exception is Brandt et al. (2019), where tariff exemptions on imported inputs for processing has a positive welfare impact.

N, the aggregate price index remains unchanged. As illustrated in Figure 4, these together imply that the equilibrium N increases while P decreases.

With the increase in both the mass of potential suppliers and the final goods market competition, firm-level heterogeneities remain important determinants of profitability. In Appendix Section D.5, we show that in this case, the rise in net profits from final good production, v^B, is an increasing function of z but a decreasing function of t, which we summarize in the following proposition:

PROPOSITION 2. *When τ_T decreases, the rise in net profits from final good production for a given firm j increases in z_j and decreases in t_j.*

Intuitively, firms with high blueprint quality but low manufacturing ability are firms that rely more on suppliers, and thus they benefit more from the increase in N. 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 they could source from. Another implication is that when the number of potential suppliers increase in equilibrium, firms will be more specialized in what they are relatively good at. Firms with good blueprints are less constrained by their manufacturing ability, and thus are more likely to thrive in the final goods market. Analogously, firms with good manufacturing ability but low blueprint quality are less likely to produce their own branded products and more likely to specialize in processing.

An obvious difficulty in testing the above proposition is that we do not observe z and t.[44] However, we do observe firm-level outcomes that are functions of z and t, such as employment or labor productivity. Therefore, given aggregate variables and parameters, we can back out a firm's blueprint quality and manufacturing ability using information on observables. This means that we can translate Proposition 2 to ask: how does the reaction of net profits from blueprint production to trade costs ($\partial v^B / \partial \tau_T$) change with respect to firm-level characteristics that are directly measurable? Following this line of thought, we prove in Appendix Section D.6 that when τ_T decreases, the rise in net profits from final good production for a given firm j increases in labor productivity, conditional on employment. At the extensive margin, this implies that when τ_T decreases, firms with higher labor productivity are more likely to enter the final goods market. We summarize this in Corollary 1:

COROLLARY 1. *When τ_T decreases, conditional on employment, firms with higher labor productivity are more likely to bring their blueprints to production.*

We formally test Corollary 1 in Section 5. We do not observe the time when a firm starts to produce its own branded product, and thus we use a close proxy: firms' registration of trademarks. Trademarks are often symbols that identify goods as manufactured by a particular person or company and confer an exclusive right to use a specific brand (Baroncelli et al., 2005); hence we

[44] Ideally, we would test the effect of a change in processing trade policy on the mass of *potential* suppliers, which is unobservable.

can view them as registered brands. If we simply extend our model by allowing firms to register their brands via costly trademark applications to avoid potential piracy, we reach the prediction that when τ_T decreases, firms with higher labor productivity would be more likely to register their trademarks.

5 Empirical Analysis

In this section, we test Corollary 1 using China's "paperless" processing trade policy in 2000-2006.[45] Our choice of policy shock, the "paperless" processing supervision program, is highly suitable for our identification strategy as it affects *only* the cost of processing exports, leaving other exporting costs as well as inherent marginal costs unchanged. Moreover, to the best of our knowledge, the literature has not empirically examined the effects of such a trade facilitation program on firms, let alone its downstream spillovers.[46]

5.1 China's "Paperless" Processing Trade

As is well known, China's customs authorities closely monitor the supply chain for processing trade due to special duty drawbacks and tax rebates 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.*"[47]

Favorable to our setting, the pilot program for paperless processing trade targeted Class A

[45] Note that we derived Corollary 1 under the assumption of a small open economy. It is hard to think of China as a small country today, but it is a reasonable assumption to make for our sample period 2000-2006. China's total exports made only 4% of world exports in 2000; its processing exports, even if we focus on the top five destinations, made up 3% of those countries' total imports on average. In contrast, China's total exports was 21% of its GDP in 2000, of which 55% was done by processing firms (these statistics are compiled using data from the Chinese customs, UN Comtrade, and the World Bank). Hence for a given policy shock on processing trade, its direct impact on Chinese firms, in relative terms, should be much larger than its impact on the foreign market.

[46] There are a number of policy evaluations that focus on the digitization of trade: Duval and Mengjing (2017) document the proliferation of paperless trade provisions in regional trade agreements, Duval et al. (2018) estimate the reduction in trade costs due to digital trade, and the UN (2017) policy report "Trade Facilitation and Paperless Trade Implementation" describes a survey on the implementation of paperless trade measures in the world.

[47] 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.

firms: firms that had at least \$10 million worth of exports. 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.[48] This policy experiment was gradually introduced 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 5.[49] By the end of 2006, due to the success of the pilot program, the policy rolled over nation-wide and was made available to *all* processing firms, regardless of size.

We start by running the following specification at the firm-level to test the direct effect of the policy:[50]

$$\ln(\text{proc. exp.})_{icst} = \alpha + \beta\ OS_{ict-1} + \gamma_i + \delta_{st} + \phi_{ct} + \epsilon_{icst}, \tag{20}$$

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.[51] 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 which we cluster 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 (20) should be interpreted as an intention-to-treat effect.

Our identification strategy relies on the assumption that treatment is assigned randomly to firms. This is not the case since only Class A firms are allowed to use the pilot program. The \$10 million threshold is a relatively high threshold since around 90% of processing firms export less than this amount in a given year. In addition, more than half of processing firms in our sample export less than \$1 million worth of goods annually. This makes the treatment and control groups highly different from each other. Thus, in our benchmark specification, we restrict our sample to firms that have between \$9 to \$11 million worth of processing exports, where the treatment and control groups include firms that exported \$10-11 and \$9-10 million worth of processing goods in the year prior to policy adoption, respectively. This is the most relevant and restrictive bandwidth for processing exports that still allows some variation for our independent variable.

Appendix Table A.2 shows balancing checks and reveals that using the entire sample of processors does not generate a suitable control group. Columns 1-3 show that firms that are above the

[48] As paperless supervision requires firms to have the 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 the ERP system. Hence, the threshold of \$10 million provides a simple yet established selection criteria. See `http://www.people.com.cn/zixun/flfgk/item/dwjjf/falv/6/6-1-50.html` (Chinese) for the official firm classification notice, and `http://www.fdi.gov.cn/1800000121_39_1919_0_7.html` (Chinese) for the official notice that explains the pilot program that targets Class A firms.

[49] Suzhou was the first prefecture to adopt the pilot program in 2000, followed by four prefectures in 2001, one prefecture each in 2002 and 2003, six in 2004, 28 in 2005, and nine in 2006.

[50] Our empirical methodology is similar in its approach to the one used by Bøler et al. (2015).

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

$10 million threshold are significantly more processing-oriented than the ones below the threshold (89% vs. 72%). They are also significantly less likely to be exiters (2% vs. 7%) or entrants (5% vs. 12%), with higher average log annual growth rates (15% vs. 8%). Columns 4-6 show that, for the restricted regression sample, the treatment and control groups are similarly less likely to exit and enter, and they have average log annual growth rates that are not statistically different from each other. The treatment group has a higher processing share of exports, but the magnitude of three percentage points is economically small.[52] Appendix Figure A.1 illustrates the processing export trends, where the implementation time 0 indicates the year the pilot program was adopted. The figure shows that the pre-trends between the chosen treatment and control groups are similar, with the firms having $10-11 million increasing their processing exports sharply in $t + 1$ and $t + 2$.

We report the estimation results of (20) with robustness checks in Table 4. The first column in panel (a) shows the benchmark result: firms that are in the treatment group in year $t - 1$ increase their processing exports by 27% in year t, relative to the control group of firms with $9-10 million 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 OS_{ict-1} and OS_{ict-3} have the expected positive signs and are significant at the 5% level.[53] In column 3, we control for whether the firm has just started exporting (entrant) and whether the firm has stopped exporting (exiter) in that year since firms' adoption of the paperless processing program might be linked to their age and future prospects. This results in a smaller sample size, but the coefficient remains positive and significant at the 1% level.[54]

In column 4 of Table 4, to check the sensitivity of our treatment classification, we allocate firms into treatment and control groups based on the *maximum* of their last two years' (prior to policy adoption) processing exports instead of our benchmark allocation based on one-year lagged processing exports: the coefficient remains significant at the 5% level and similar in magnitude. Column 5 uses a first-difference (FD) specification and reveals that the program increased the log growth rate of treated firms' processing exports by 8.5 log points (significant at the 10% level) relative to the control group. In column 6, we do a falsification analysis by looking at the effect on the ordinary exports of mixed exporters. Mixed exporters that are above the $10 million threshold are eligible to adopt the paperless system, which should affect their processing exports but not directly their ordinary exports. Consistent with this conjecture, the coefficient in column 6 is not statistically different from zero. On the contrary, column 7 shows that those mixed exporters increase their processing exports: the estimated coefficient has a similar magnitude to

[52] The number of observations is highly skewed towards small processors in the entire sample, whereas they are roughly equally distributed in the restricted regression sample.

[53] Note that we use two-year intervals instead of one to avoid collinearity among the dummies.

[54] We lose the first (2000) and last (2006) years of our sample to correctly identify entrants and exiters respectively.

our benchmark case and remains significant at the 10% level.

The bottom panel of Table 4 provides further robustness checks. In column 8, we do a falsification analysis by setting the threshold to \$9 million, and the bandwidth (bw) to \$8-10 million, and find a statistically insignificant coefficient. In columns 9 and 10, we use wider bandwidths of \$7.5-12.5 and \$5-15 million respectively and find positive and statistically significant coefficients, albeit with lower magnitudes. Column 11 restricts the sample to firms that export continuously in 2000-2006 to rule out entry/exit dynamics. In column 12, we exclude the electronics sector since processing firms in this industry have a lower threshold (\$5 million) to qualify for the pilot program. In columns 13 and 14, we exclude firms with foreign ownership and SOEs respectively. None of these changes affect our coefficient qualitatively.[55]

5.2 Downstream Spillovers and Trademarks

Now that we have established that the pilot program did increase firm-level processing exports, we turn to its downstream spillovers. Corollary 1 states that *productive* downstream firms would be more likely to establish their own brands/trademarks thanks to the larger mass of potential suppliers when processing cost τ_T falls. Existing empirical research suggests that supplier-buyer relationships are highly localized (Bernard et al., 2019), and thus we expect that geographically close firms are more likely to be affected. In other words, we hypothesize that downstream firms that are in the same prefecture as the affected suppliers would be more likely to benefit from the spillover and thus apply for new trademarks.

To examine the effect of the pilot program on firm-level branding activity, we focus on the sample of *non-processing domestic* firms. We exclude processing exporters as this allows us to cleanly examine the spillover to firms that engage in ordinary activities; we also exclude foreign-owned firms since their trademark applications more likely reflect protecting their existing brands rather than bringing in new blueprints to production. We then run the following specification:

$$Y_{icst} = \alpha + \beta OS_{ct-1} \times \text{Productive}_i + \lambda \ln{(empl.)}_{it} + \psi \ln{(capital)}_{it} + \gamma_i + \delta_{st} + \phi_{ct} + \epsilon_{icst}, \quad (21)$$

where Y_{icst} is the number of effective trademarks a firm has,[56] OS_{ct-1} is the policy adoption indicator as before, and Productive_i indicates whether the firm's initial log labor productivity is above the median value. We focus on the interaction coefficient as our model predicts that only productive firms will be more likely to bring their blueprint to production when faced with a positive upstream shock. 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

[55] In unreported results, we use a specification with the most restrictive prefecture-sector-year fixed effects. This takes much of meaningful variation away but we continue to find a positive coefficient (0.206), significant at the 10% level.

[56] As mentioned before, note that trademarks are the legal basis for brands and thus we are using the number of effective trademarks as a proxy for firms' branding activity. The discussion on how to link trademarks to brands under our theoretical framework is provided at the end of Section 4.3.

fixed effects δ_{st} to control for sector-specific supply and demand shocks (sectors are at the 4-digit CIC level), and prefecture-year fixed effects ϕ_{ct} to control province-wide policy changes that might affect trademark applications.[57] We cluster the errors ϵ_{icst} two-way at the prefecture and sector level as before. Due to the large number of fixed effects, we estimate specification (21) linearly to avoid the incidental parameters problem. We provide various robustness checks with alternative measures.

Our identification relies on the plausible assumption that the timing of introducing the pilot paperless *processing* program by a prefecture's customs is exogenous to the branding activities of the *non-processing* firms in the same region. We also rely on the fact that processing exporters can sell domestically. The literature has largely ignored this possibility, but processing firms do sell domestically if they pay the required taxes.[58] The matched customs-AIS data indicates that 76% of processing exporters in 2005 also sold domestically.[59]

Table 5 panel (a) has our benchmark results. In column 1, in order to focus on the main effect, we use a less restrictive specification with province-year (instead of prefecture-year) fixed effects, and find that the pilot program does not have a significant effect on the number of trademarks for the average firm. In column 2, we interact OS_{ct-1} with Productive$_i$, and find that the pilot paperless processing trade program increased the number of trademarks of a productive firm by 0.241 (0.346-0.105), which is 10.2% of the average number of effective trademarks (2.37) for firms with above-median productivity. In columns 3 and 4, we add prefecture-year and the most restrictive prefecture-sector-year fixed effects respectively, and in both cases the interaction coefficient barely changes in magnitude and stays significant at the 1% level.

In panel (b) of Table 5, we do several robustness checks. In column 5, instead of the Productive$_i$ indicator, we interact OS_{ct-1} with the firm's demeaned initial labor productivity, $\ln(\textit{labor prod.})_i$, and the result stays robust. In column 6, the dependent variable is a dummy that indicates whether the firm has at least one effective trademark. In column 7, we focus on the log number of trademarks, which results in a smaller sample size due to dropping firms with no trademarks. The coefficients show that the processing trade shock has positive effects on both the extensive and the intensive margins of trademark activity. In column 8, we remove the 2,560 firms that have larger than 25 trademarks (outliers at the 99th percentile).[60] Column 9 excludes SOEs from the sample as these firms' trademark activities might be subject to government controls.[61] Neither of these robustness checks change the qualitative result. We also find that the number of employees and the

[57] 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. The distribution is highly skewed to the right even if we zoom in on firms that have at least one trademark. Thus, we do a robustness check by excluding outlier firms that have more than 25 trademarks.

[58] The official customs regulatory document that explains how processing exporters can sell domestically can be found in http://www.customs.gov.cn/customs/302249/302266/302267/356603/index.html.

[59] The median (mean) exports share (exports/sales) for these processing firms was 81% (63%).

[60] Defining outliers at the 95th percentile and thus excluding firms that have larger than five trademarks produces qualitatively similar results.

[61] Out of the 235,456 firms in our sample, 22,405 are SOEs.

capital stock have a positive and significant effect on trademarks in all regressions, as expected.

One might be concerned that the above result does not specifically identify the *downstream* firms that are affected by the processing trade shock. In order to dispel this concern, we use an alternative strategy that uses China's official 2002 IO table. Since our shock is based on the customs data which is in HS classification, we concord the shock to the industry classification used in China's IO table, which we then concord to the 4-digit CIC level used in the AIS firm-level data. For this, we use crosswalks from the HS8-IO industry concordance to the CIC-IO industry concordance to create 74 tradable industries.[62] Once we redefine our shock at this new sector level, we aggregate the IO table to the level of the 74 IO industries, labeling the unmatched non-tradable industries as "other." We also assign each firm an IO industry based on the CIC-IO industry concordance table.

We define the "treated processing share" for each prefecture-sector-year (cst) in the following way:

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

where $i \in A$ are processing firms that are above the \$10 million threshold,[63] and sector s is defined at the IO level. 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%).[64]

Then, we create a time-varying input shock using the treated processing share for each output sector n and prefecture c as follows:

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

where ω_{ns} are cost shares from the redefined IO table. We then run the following specification:

$$Y_{icnt} = \alpha + \beta \text{ Input shock}_{cnt} \times \text{Productive}_i + \lambda \ln{(empl.)}_{it} + \psi \ln{(capital)}_{it} + \gamma_i + \delta_{nt} + \phi_{ct} + \epsilon_{icnt}, \quad (22)$$

where Y_{icnt} is the number of trademarks as before, δ_{nt} are sector-year fixed effects, now at the IO level, ϕ_{ct} are prefecture-year fixed effects, and ϵ_{icnt} is the error term which we cluster two-way at the prefecture and sector level. Compared to (21), specification (22) allows us to focus directly on downstream firms at the cost of some measurement error created by sector aggregation.

Table 6 column 1 shows that the input shock does not have a significant effect on the number of trademarks for the average firm. In column 2, we interact the input shock variable with the

[62] We thank Yu Shi for providing the HS8-IO industry correspondance table. There are 7,428 HS8 matched to 85 distinct IO industries. We adjust for the one-to-many and many-to-many matches using the aggregation algorithm provided by Van Beveren et al. (2012). The CIC-IO industry concordance table is from Zi (2019). Since the CIC codes changed after 2002, we first adjust the CIC industries overtime to create uniform CIC industry codes using pre- to post-2002 CIC concordance tables.

[63] We adjust for the electronics sector's lower threshold of \$5 million.

[64] 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.

Productive$_i$ dummy, and find an interaction coefficient of 2.222, significant at the 1% level. This result indicates that a one standard deviation (0.033) increase in Input shock$_{cnt}$ raises the number of trademarks of a productive firm by 0.035 ($(2.222 - 1.167) \times 0.033$), which is 1.5% of the average number of trademarks (2.34) for firms with above-median productivity. In column 3, we directly control for the treated processing share as well as its interaction with Productive$_i$ (i.e., control for Output shock$_{cnt}$ = Treated processing share$_{cnt}$). We include this control since promoting processing policy might crowd out ordinary firms and hence directly affect their branding activities. The estimated coefficient in column 3 barely changes when compared to column 2. Finally, in column 4, we use the strictest prefecture-sector-year fixed effects, and find an interaction coefficient of 2.487, significant at the 1% level. Overall, results in Table 6 confirm the findings in Table 5 that the pilot paperless processing trade program has induced downstream firms to increase their branding activity as predicted by our model.

6 Conclusion

In this paper, we 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 these mixed firms, who engage predominantly in processing, are superior to other firms in multiple dimensions. We revisited some of the earlier findings in the literature by focusing on these "super processors," and provided a set of novel stylized facts on firms' performance, brand ownership, and choice of trade mode.

We then formalized a parsimonious model based on the frameworks of Antràs et al. (2017) and Bernard et al. (2019). In the model, 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 rationalized the ranking among the different types of exporters that we observe in the data, and provided a new source of gains from promoting processing trade: facilitating processing trade raises the *ex-ante* expected profits from task production and hence encourages entry, leading a greater mass of potential suppliers, which eventually benefits downstream ordinary firms, especially the ones with good ideas but low manufacturing ability who rely heavily on suppliers.

In the last part of the paper, we tested our model's prediction using China's pilot "paperless" processing supervision program in 2000-2006 as a quasi-natural experiment. Consistent with the model's prediction, we found that promoting processing trade not only increased the processing exports of targeted firms, but also induced productive domestic downstream firms to establish their own trademarks. Overall, our theoretical and empirical analyses in this paper highlighted that processing trade has allowed goods to be not only "Made in China," but also "Created in China" by providing a breeding ground of potential task 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.

Allanson, P. and Montagna, C. (2005). Multiproduct firms and market structure: An explorative application to the product life cycle. *International Journal of Industrial Organization*, 23(7-8):587–597.

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.

Balmer, J. M. and Chen, W. (2017). *Advances in Chinese Brand Management*. Springer.

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

Bernard, A. B., 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.

Beverelli, C., Neumueller, S., and Teh, R. (2015). Export diversification effects of the WTO trade facilitation agreement. *World Development*, 76(C):293–310.

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.

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.

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.

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. (2017). Processing trade and global idea diffusion. *mimeo.*

Deng, X., Jing, R., and Liang, Z. (2018). Trade liberalization and domestic brands: Evidence from China's accession to the WTO. *mimeo.*

Duval, Y. and Mengjing, K. (2017). Digital trade facilitation: Paperless trade in regional trade agreements. *ADBI Working Paper*, 747.

Duval, Y., Utoktham, C., and Kravchenko, A. (2018). Impact of implementation of digital trade facilitation on trade costs. *ARTNeT Working Paper Series*, 174.

Fan, H., Li, Y. A., and Yeaple, S. R. (2015). Trade liberalization, quality, and export prices. *Review of Economics and Statistics*, 97(5):1033–1051.

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. M., Mattoo, A., Nguyen, H., and Schiffbauer, M. (2019). The internet and Chinese exports in the pre-ali baba era. *Journal of Development Economics*, 138(C):57–76.

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.

Freund, C. and Rocha, N. (2011). What constrains Africa's exports? *The World Bank Economic Review*, 25(3):361–386.

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.

Hoekman, B. and Nicita, A. (2011). Trade policy, trade costs, and developing country trade. *World Development*, 39(12):2069–2079.

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.

Kneller, R. and Timmis, J. (2016). ICT and exporting: The effects of broadband on the extensive margin of business service exports. *Review of International Economics*, 24(4):757–796.

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.

Malgouyres, C., Mayer, T., and Mazet-Sonilhac, C. (2019). Technology-induced trade shocks? Evidence from broadband expansion in France. *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.

Portugal-Perez, A. and Wilson, J. S. (2012). Export performance and trade facilitation reform: Hard and soft infrastructure. *World Development*, 40(7):1295–1307.

Radelet, S. and Sachs, J. (1997). Asia's reemergence. *Foreign Affairs*, 76(6):44–59.

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

Tintelnot, F., Kikkawa, A. K., Mogstad, M., and Dhyne, E. (2018). Trade and domestic production networks. *NBER Working Paper Series*, 25120.

Umana-Dajud, C. (2019). Do visas hinder international trade in goods? *Journal of Development Economics*, 140(C):106–126.

UN (2017). Trade facilitation and paperless trade implementation. Technical report, United Nations.

Van Beveren, I., Bernard, A. B., and Vandenbussche, H. (2012). Concording EU trade and production data over time. *NBER Working Paper Series*, 18604.

Wang, A. (2019). Team LEWIS ASIA: Shenzen's potential as the next global hub of innovation [available at: https://www.teamlewis.com/asia/magazine/shenzhen-the-next-global-innovation-hub/].

Wang, Z. and Yu, Z. (2012). Trading partners, traded products and firm performances of China's exporter-importers: Does processing trade make a difference? *The World Economy*, 35(12):1795–1824.

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.

Zi, Y. (2019). Trade liberalization and the great labor reallocation. *mimeo*.

Tables and Figures

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) Share of mixed HS8	0.29	0.37	0.31	0.31	0.37	0.30
(5) Share of mixed HS8-country	0.19	0.25	0.24	0.20	0.24	0.23
(6) Value share of mixed HS8	0.87	0.68	0.37	0.89	0.71	0.35
(7) 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, and their composition of exports (mixed exports/total exports) in rows 4-7, 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.

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.

Table 3: Export Mode and Brand Ownership

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.

Table 4: Paperless Trade and Processing Exports

Dep. var.: ln(proc. exp.)$_{icst}$	(1) Benchmark	(2) Leads & lags	(3) Entry & exit	(4) Max.	(5) FD	(6) Mixed only (ordinary)	(7) Mixed only (proc.)
OS_{ict-1}	0.274***	0.166**	0.250***	0.223**	0.085*	-0.208	0.254*
	(0.095)	(0.073)	(0.082)	(0.084)	(0.044)	(0.232)	(0.145)
OS_{ict+1}		0.053					
		(0.118)					
OS_{ict-3}		0.289***					
		(0.019)					
Entrant$_{it}$			-1.331***				
			(0.144)				
Exiter$_{it}$			-1.540***				
			(0.281)				
Firm FE	Yes	Yes	Yes	Yes	No	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	2,566	2,566	1,867	2,855	2,118	1,388	1,388
R^2	0.59	0.59	0.73	0.54	0.28	0.86	0.66

Dep. var.: ln(proc. exp.)$_{icst}$	(8) False threshold	(9) bw: $7.5-12.5m	(10) bw: $5-15m	(11) Always exporters	(12) No electronics	(13) No foreign firms	(14) No SOEs
OS_{ict-1}	0.039	0.115**	0.066**	0.296***	0.277**	0.296**	0.245***
	(0.049)	(0.054)	(0.032)	(0.077)	(0.126)	(0.119)	(0.077)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	3,053	6,911	15,818	1,714	1,718	1,128	2,151
R^2	0.58	0.54	0.53	0.61	0.62	0.65	0.60

Notes: This table reports the results of running specification (20). OS_{ict-1} indicates the implementation of the pilot paperless processing trade program in prefecture c in year $t-1$ for firm i. Sector refers to the top (core) HS2 of each firm. Column 1 shows the benchmark overall effect. In column 2, we add a lead (OS_{ict+1}) and another lag (OS_{ict-3}) to check for anticipation effects. Column 3 controls for firm entry and exit. In column 4, we allocate firms into treatment and control groups based on the maximum of their last two years' (prior to policy adoption) processing exports instead of our benchmark allocation based on one-year lagged processing exports. In column 5, we use a first-difference (FD) specification. Column 6 and 7 focus on mixed exporters' ordinary and processing exports respectively. In column 8, we do a falsification analysis by setting the threshold to $9m, and the bandwidth (bw) to $8-10m. In columns 9 and 10, we widen the bandwidth to $7.5-12.5m and $5-15m respectively. Columns 11 restricts the sample to always exporters. Columns 12, 13, and 14 exclude firms that are in the electronics sector, foreign firms, and SOEs, 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.

Table 5: Trademarks

	(a) Benchmark				(b) Robustness checks				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var.: Y_{icst}	Overall effect	Median	+CT FE	+ CST FE	Demeaned	Extensive margin	Intensive margin	No outliers	No SOEs
OS_{ct-1}	0.075	-0.105							
	(0.103)	(0.117)							
$\times$ Productive$_i$		0.346***	0.348***	0.309***		0.006*	0.029**	0.056**	0.264***
		(0.097)	(0.095)	(0.106)		(0.003)	(0.012)	(0.023)	(0.101)
$\times \ln(\textit{labor prod.})_i$					0.244***				
					(0.081)				
$\ln(\textit{empl.})_{it}$	0.355***	0.349***	0.340***	0.294***	0.288***	0.015***	0.058***	0.116***	0.297***
	(0.052)	(0.052)	(0.051)	(0.056)	(0.055)	(0.001)	(0.004)	(0.011)	(0.060)
$\ln(\textit{capital})_{it}$	0.151***	0.151***	0.150***	0.137***	0.137***	0.008***	0.030***	0.053***	0.144***
	(0.021)	(0.021)	(0.020)	(0.024)	(0.025)	(0.001)	(0.003)	(0.006)	(0.025)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	No	No	No	No	No	No
Province-year FE	Yes	Yes	No	No	No	No	No	No	No
Prefecture-year FE	No	No	Yes	No	No	No	No	No	No
Prefecture-sector-year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	942,212	942,212	942,193	831,190	831,190	831,190	281,853	821,803	755,406
R^2	0.89	0.89	0.89	0.92	0.92	0.95	0.96	0.94	0.92

Notes: This table reports the results of running specification (21). Y_{icst} is the number of trademarks of firm i in sector s residing in prefecture c in year t. Sectors refer to 425 4-digit CIC industries. Productive$_i$ indicates firms whose initial log labor productivity is larger than the median. $\ln(\textit{labor prod.})_i$ is demeaned initial labor productivity of firm i. Column 1 shows the main effect, without the interaction. Column 2 adds the Productive$_i$ interaction. Columns 3 and 4 add prefecture-year and prefecture-sector-year fixed effects respectively. Column 5 uses $\ln(\textit{labor prod.})_i$ instead of Productive$_i$. In column 6, Y_{icst} is a dummy variable that indicates whether the firm has a trademark, whereas in column 7, Y_{icst} is the log number of trademarks. Column 8 excludes firms that have more than 25 trademarks (outliers), and column 9 excludes SOEs. 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 6: Trademarks, with IO Linkages

Dep. var.: Y_{icnt}	(1) Overall effect	(2) Median	(3) Output control	(4) + CNT FE
Input shock$_{cnt-1}$	0.023	-1.167	-1.159	
	(0.743)	(0.882)	(0.960)	
$\times$ Productive$_i$		2.222***	2.406***	2.487***
		(0.658)	(0.693)	(0.687)
Output shock$_{cnt-1}$			0.019	
			(0.086)	
$\times$ Productive$_i$			0.143	0.204*
			(0.092)	(0.104)
$\ln(empl.)_{it}$	0.341***	0.336***	0.336***	0.314***
	(0.065)	(0.064)	(0.064)	(0.058)
$\ln(capital)_{it}$	0.155***	0.155***	0.155***	0.142***
	(0.029)	(0.029)	(0.029)	(0.027)
Firm FE	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	No
Prefecture-year FE	Yes	Yes	Yes	No
Prefecture-sector-year FE	No	No	No	Yes
Obs.	940,068	940,068	940,068	919,408
R^2	0.89	0.89	0.89	0.90

Notes: This table reports the results of running specification (22). 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 1 shows the main effect, without the interaction. Column 2 adds the Productive$_i$ interaction. In column 3, we also include Output shock$_{cnt-1}$ and its interaction with Productive$_i$ to control for own-industry effects. Column 4 adds prefecture-sector-year fixed effects. 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 1: Selection into Operating Status

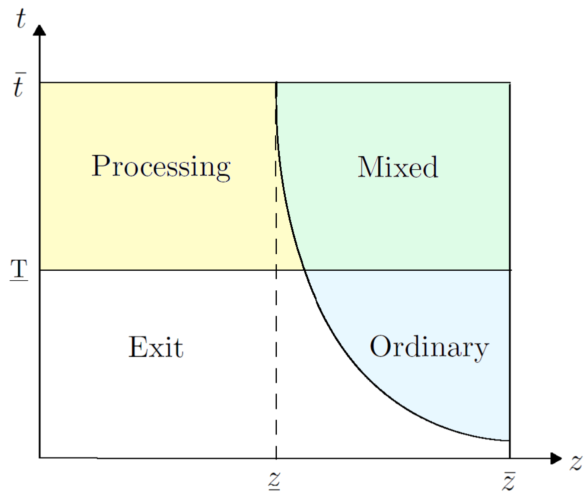


Figure 2: Determination of the Equilibrium

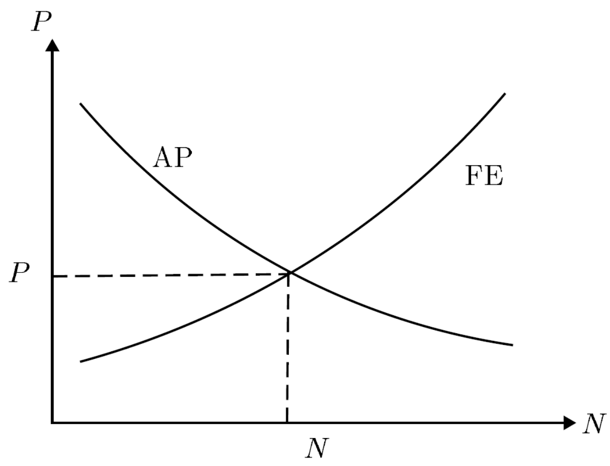


Figure 3: Selection with International Trade

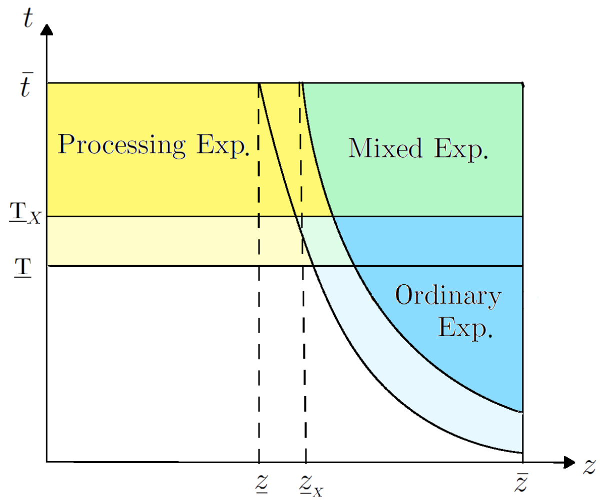


Figure 4: Impact of Processing-promoting Policy

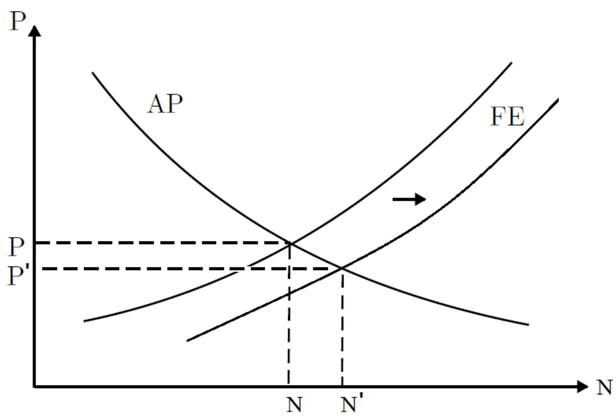
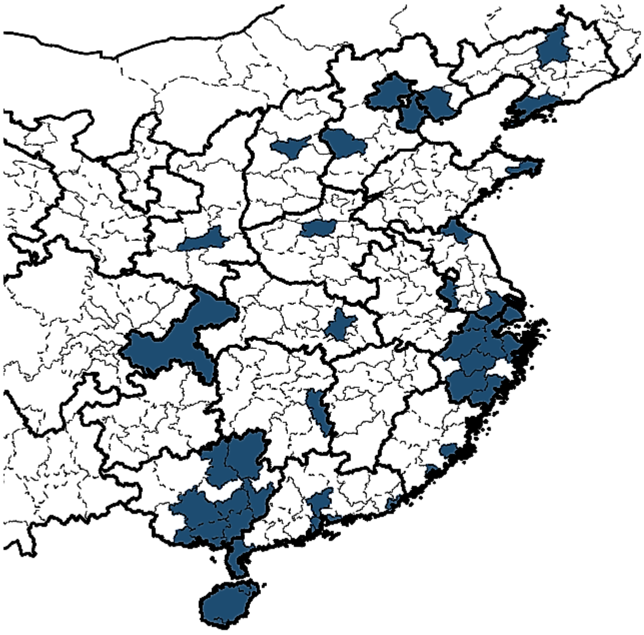


Figure 5: Adoption of the Pilot Paperless Processing Trade Program



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

A Appendix Tables and Figures

Table A.1: 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%

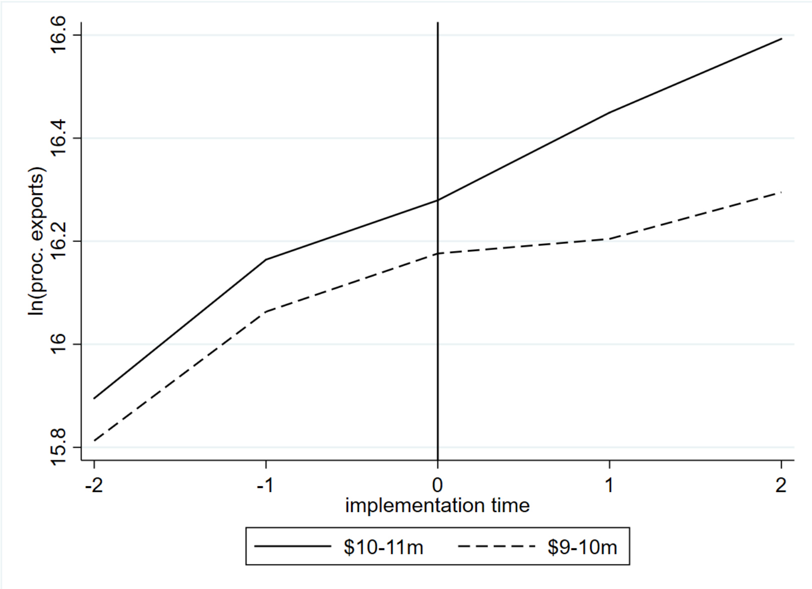
Notes: This table reports the share of export modes in no brand, foreign brand, and domestic brand categories in columns 1, 2, and 3 respectively, using the 591,270 manufacturing export transactions in the 2018 customs data sample (after excluding the 271,297 transactions made by wholesalers and intermediary 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.

Table A.2: Comparisons of Firms

Sample	(1) All >\$10m proces- sors	(2) All <\$10m proces- sors	(3) All *t*-value	(4) \$10-11m proces- sors	(5) \$9-10m proces- sors	(6) \$9-11m *t*-value
Proc. share of exports	0.89	0.72	-56.74***	0.89	0.86	-3.70***
Exiter	0.02	0.07	20.61***	0.02	0.02	-0.07
Entrant	0.05	0.12	25.22***	0.05	0.05	-0.06
Avg. log annual growth	0.15	0.08	-5.79***	0.19	0.14	-1.20
Obs.	14,756	238,412		1,151	1,415	

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 above and below the \$10m threshold respectively (entire sample). Columns 4 and 5 represent the means of the variables for exporters that have \$10-11m and \$9-10m processing exports respectively (regression sample). Columns 3 and 6 show the *t*-statistics that indicate whether the means are statistically different from each other in the entire and regression samples respectively; *** denote statistical significance at the 1% level.

Figure A.1: Processing Export Trends



Notes: This figure plots the level of processing exports for exporters that had \$10-11m and \$9-10m worth of processing exports in the year prior to policy adoption. Implementation time 0 indicates the year the prefecture's customs authority adopted the pilot paperless processing trade program.

B 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.[65] 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.[66] 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 B.1 lists the 36 products with their brief descriptions.

B.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.[67] 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:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma m_{it} + \omega_{it} + \varepsilon_{it}, \tag{23}$$

where q_{it} is output quantity, 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 firms' prices within an industry. As a consequence, the estimated productivity contains information

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

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

[67] 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.

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:

$$m_{it} = \tilde{m}_{it} - p_{Mit}. \tag{24}$$

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:

$$\bar{m}_{it} = \tilde{m}_{it} - p_{Mjt}. \tag{25}$$

Plugging (25) into (24), 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:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma \bar{m}_{it} + \omega^*_{it} + \varepsilon_{it}, \tag{26}$$

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. We find this to be 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:

$$p_{Mit} = f(p_{it}, \mathbf{X}_{it}). \tag{27}$$

This allows us to express physical material input as:

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

Thus, the production function we estimate is given by:

$$q_{it} = \alpha k_{it} + \beta l_{it} + \gamma \tilde{m}_{it} + \gamma f(p_{it}, \mathbf{X}_{it}) + \omega_{it} + \varepsilon_{it}. \tag{28}$$

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

$$p_{it} = \log\left(\frac{R_{it}}{Q_{it}}\right), \tag{29}$$

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

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

Table B.1: 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 Section B. The English product specifications are translated from `http://www.i5a6.com/hscode/`.

C List of Products in the Brand Data Sample

C.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/.

D Theory Appendix

D.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:

$$\underline{t}(z_j, t_j) = \frac{\theta f}{\sigma - 1}\left(Ak_1 z_j^{\sigma-1}\right)^{-1}\Theta(z_j, t_j)^{1-\frac{\sigma-1}{\theta}}. \tag{30}$$

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 $\Theta_j \equiv \Theta(z_j, t_j)$ 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:

$$\frac{\partial \underline{t}_j}{\partial z_j} = \frac{\theta f}{(\sigma-1)Ak_1}\left[(1-\sigma)z_j^{-\sigma}\Theta_j^{1-\frac{\sigma-1}{\theta}} + z_j^{1-\sigma}\frac{\partial \Theta_j^{1-\frac{\sigma-1}{\theta}}}{\partial z_j}\right], \tag{31}$$

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

$$\frac{\partial n(z_j, t_j)}{\partial z_j} = \frac{\partial n(z_j, t_j)}{\partial \underline{t}_j}\frac{\partial \underline{t}_j}{\partial z_j} > 0. \tag{32}$$

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

$$o(z_j, t_j) = 1 - \frac{t_j}{\Theta_j}. \tag{33}$$

It immediately follows that:

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

Lastly, the unit cost is expressed as:

$$c(z_j, t_j) = \frac{\Theta_j^{-\frac{1}{\theta}}\gamma^{\frac{1}{1-\rho}}}{z_j}. \tag{34}$$

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 (30) with respect to t_j, we get:

$$\frac{\partial \underline{t}_j}{\partial t_j} \propto \left(1 - \frac{\sigma - 1}{\theta}\right) \frac{\partial \Theta_j}{\partial t_j}. \tag{35}$$

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 (35), 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:

$$\frac{\partial n\left(z_j, t_j\right)}{\partial t_j} = \frac{\partial n\left(z_j, t_j\right)}{\partial \underline{t}_j} \frac{\partial \underline{t}_j}{\partial t_j} < 0. \tag{36}$$

From (33), 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 (34), 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 (30) with respect to N, we obtain:

$$\frac{\partial \underline{t}_j}{\partial N} = \frac{\theta f}{\sigma - 1} \left(A k_1 z_j^{\sigma-1}\right)^{-1} \left(1 - \frac{\sigma - 1}{\theta}\right) \Theta_j^{-\frac{\sigma-1}{\theta}} \frac{\partial \Theta_j}{\partial N}. \tag{37}$$

From the expression of Θ_j, we obtain:

$$\frac{\partial \Theta_j}{\partial N} = \left(\int_{\underline{t}_j}^{\bar{t}} \iota dG_t(\iota) - N \underline{t}_j g_t\left(\underline{t}_j\right) \frac{\partial \underline{t}_j}{\partial N}\right). \tag{38}$$

Now suppose that $\partial \underline{t}_j / \partial N \leq 0$, then expression (38) implies that $\partial \Theta_j / \partial N > 0$. By (37), 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 (37). 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:

$$\frac{\partial n}{\partial N} = \int_{\underline{t}_j}^{\bar{t}} dG_t(\iota) - N g_t\left(\underline{t}_j\right) \frac{\partial \underline{t}_j}{\partial N} = \frac{\int_{\underline{t}_j}^{\bar{t}} dG_t + N(1 - \frac{\sigma-1}{\theta}) \frac{\underline{t}_j}{\Theta_j} \int_{\underline{t}_j}^{\bar{t}} (\underline{t}_j - \iota) dG_t(\iota)}{1 + (1 - \frac{\sigma-1}{\theta}) N \underline{t}_j g_t(\underline{t}_j) \frac{\underline{t}_j}{\Theta_j}}. \tag{39}$$

Inspecting the right-hand side, the first term is positive and the second term is negative. As a result, $\partial n\left(z_j,\, t_j\right)/\partial N$ can either be positive or negative. By expression (33), 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:

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

D.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., $\min\left\{T_i,\, T_{i'}\right\} \geq \underline{t}\left(z_j,\, t_j\right)$. In this case, recall that the bilateral trade between two firms is given by:

$$x_{ij} = \lambda_{ij} x_j = \frac{T_i}{\Theta_j} x_j, \tag{40}$$

$$x_{i'j} = \lambda_{i'j} x_j = \frac{T_{i'}}{\Theta_j} x_j. \tag{41}$$

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 $\pi^T\left(T_i\right) = \frac{1}{1+\theta}\sum_{j\in\Omega_i} x_{ij}$, and thus $\pi^T\left(T_i\right) < \frac{1}{1+\theta}\sum_{j\in\Omega_i} x_{i'j} < \frac{1}{1+\theta}\sum_{j\in\Omega_{i'}} x_{i'j} = \pi^T\left(T_{i'}\right)$.

D.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:

$$P = N^{\frac{1}{1-\sigma}} \left[\int_{\underline{Z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} p^B (c(\zeta, \iota))^{1-\sigma} dG_t(\iota) dG_z(\zeta) \right]^{\frac{1}{1-\sigma}}. \tag{42}$$

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:

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

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,

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

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:

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

and thus we get:

$$\frac{\partial F_{FE}}{\partial N} = \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \left(\frac{\sigma-1}{\theta} \pi^B \frac{\Theta - t}{\Theta} \frac{1}{N} \right) dG_t(\iota) dG_z(\zeta) \\ - \int_{\underline{z}}^{\bar{z}} \int_{\Xi(\zeta)}^{\bar{t}} \frac{fn}{N} dG_t(\iota) dG_z(\zeta) - \frac{(\sigma-1)\beta L}{N^2 \sigma(\theta+1)}. \tag{43}$$

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

$$\frac{\partial(\frac{\Theta}{\theta+1} - t - n\underline{t})}{\partial \underline{t}} = \frac{\theta}{\theta+1} N \underline{t} g_t(\underline{t}) - N \int_{\underline{t}}^{\bar{t}} g_t(\iota) d\iota. \tag{44}$$

Furthermore:

$$\frac{\partial^2(\frac{\Theta}{\theta+1} - t - n\underline{t})}{\partial \underline{t}^2} = \frac{\theta}{\theta+1} N \underline{t} {g'}_t(\underline{t}) + \frac{\theta}{\theta+1} N g_t(\underline{t}) + N g_t(\underline{t}). \tag{45}$$

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:

$$fn_i = \pi_i^B. \tag{46}$$

Optimal sourcing condition implies that:

$$f = \underline{t}_i \frac{\pi_i^B}{\Theta_i} \frac{\sigma-1}{\theta}. \tag{47}$$

Equations (46) and (47) together imply that $n_i \underline{t}_i = \Theta_i \frac{\theta}{\sigma-1}$. Therefore:

$$\frac{\Theta}{\theta+1} - t - n\underline{t} \leq \frac{\Theta_i}{\theta_i+1} - t_i - n_i\underline{t}_i < \frac{\Theta_i}{\theta_i+1} - n_i\underline{t}_i < \frac{\Theta_i}{\theta_i+1} - \frac{\Theta_i \theta}{\sigma-1} < 0. \tag{48}$$

As a result, $\frac{1}{\theta+1}\pi^B - \frac{t\pi^B}{\Theta} - \frac{\theta}{\sigma-1} f n < 0$, 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:

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

D.4 Proofs of Ranks

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

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

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

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

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

which, after some algebra, is equivalent to:

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

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 $\frac{\int_{\underline{T}_X}^{t'} \int_{\Xi_X^{-1}(\iota)}^{\bar{z}} dG_z(\zeta) dG_t(\iota)}{\int_{\underline{T}_X}^{t'} dG_t(\iota)}$ is increasing in t. Taking a partial derivative with respect to t', we get:

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

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

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

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

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

Denote $z_1 \equiv \Xi_X^{-1}(\bar{t})$, $z_2 \equiv \Xi_X^{-1}(\underline{T}_X)$. 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:

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

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

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:

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

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:

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

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

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

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:

$$\frac{\partial ln v_j^B}{\partial ln z_j} = (\sigma - 1)\frac{\pi_j^B}{v_j^B} > \sigma - 1, \tag{49}$$

$$\frac{\partial ln l_j^B}{\partial ln z_j} = \frac{(\frac{\sigma-1}{\theta} - 1)(\sigma - 1)M_j}{1 + (1 - \frac{\sigma-1}{\theta})M_j} + (\sigma - 1) = \frac{\sigma - 1}{1 + (1 - \frac{\sigma-1}{\theta})M_j} < \sigma - 1, \tag{50}$$

where $M_j \equiv N\underline{\mathrm{t}}_j g_t(\underline{\mathrm{t}}_j)\frac{\mathrm{t}_j}{\Theta}$. 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:

$$\frac{\partial ln v_j^B}{\partial ln t_j} = \frac{(\sigma - 1)t_j \pi_j^B}{\theta v_j^B \Theta_j}, \tag{51}$$

$$\frac{\partial ln l_j^B}{\partial ln t_j} = 1 - \frac{(1 - \frac{\sigma-1}{\theta})t_j}{\Theta_j} \frac{1}{1 + (1 - \frac{\sigma-1}{\theta})M_j}. \tag{52}$$

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.

D.5 Proof of Proposition 2

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:

$$\begin{aligned} dv_j^B &= \frac{\sigma - 1}{\theta}\frac{\pi_j^B}{\Theta_j}\frac{\partial \Theta_j}{\partial N}dN - \underline{t}_j\frac{\pi_j^B}{\Theta_j}\frac{\sigma - 1}{\theta}\frac{\partial n}{\partial N}dN + (\sigma - 1)\pi_j^B\frac{dP}{P} \\ &\propto \frac{1}{\Theta_j}N\frac{\partial \Theta_j}{\partial N}N - \frac{1}{\Theta_j}\underline{t}_j\frac{\partial n}{\partial N} + \theta\frac{\partial dlnP}{\partial dlnN}. \end{aligned} \tag{53}$$

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 (53), we get:

$$dv_j^B = \frac{\sigma - 1}{\theta}\pi_j^B\left(\frac{N\int_{\underline{t}_j}^{\bar{t}}(\iota - \underline{t}_j)dG_t(\iota)}{\Theta_j} + \theta\frac{\partial dlnP}{\partial dlnN}\right). \tag{54}$$

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:

$$
\begin{aligned}
\frac{1}{\Theta_j}\frac{1}{1+(1-\frac{\sigma-1}{\theta})M_j}F_{dv}+\frac{\partial F_{dv}}{\partial t_j} &= \frac{N}{\Theta_j^2}\frac{\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota)}{1+(1-\frac{\sigma-1}{\theta})M_j}+\frac{\partial F_{dv}}{\partial t_j} \\
&< \frac{N}{\Theta_j^2}\frac{\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota)}{1+(1-\frac{\sigma-1}{\theta})M_j}-\frac{N}{\Theta_j^2}\frac{\partial \Theta_j}{\partial t_j}\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota) \\
&\propto \frac{1}{1+(1-\frac{\sigma-1}{\theta})M_j}-\frac{\partial \Theta_j}{\partial t_j}=0.
\end{aligned}
\tag{55}
$$

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:

$$
\begin{aligned}
\frac{\partial f_{dv}}{\partial \underline{t}_j} &= -g_t(\underline{t}_j)+\Delta_1\frac{\partial^2\Theta_j}{\Theta_j\partial \underline{t}_j^2}\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota)+\Delta_1\frac{\partial \Theta_j}{\Theta_j\partial \underline{t}_j}\int_{\underline{t}_j}^{\bar{t}}(-1)dG_t(\iota) \\
&\quad+\frac{\Delta_2\partial^2\Theta_j}{N\partial \underline{t}_j^2}-\Delta_1\frac{1}{\Theta_j^2}(\frac{\partial \Theta_j}{\partial \underline{t}_j})^2\int_{\underline{t}_j}^{\bar{t}}(\iota-\underline{t}_j)dG_t(\iota).
\end{aligned}
\tag{56}
$$

Note that:

$$-g_t(\underline{t}_j)+\Delta_1\frac{\partial\Theta_j}{\Theta_j\partial\underline{t}_j}\int_{\underline{t}_j}^{\bar{t}}(-1)dG_t(\iota)<-g_t(\underline{t}_j)+\Delta_1\frac{\underline{t}_jg_t(\underline{t}_j)}{\int_{\underline{t}_j}^{\bar{t}}\iota dG_t(\iota)}\int_{\underline{t}_j}^{\bar{t}}(1)dG_t(\iota)$$
$$=-g_t(\underline{t}_j)+\Delta_1g_t(\underline{t}_j)\frac{\int_{\underline{t}_j}^{\bar{t}}\underline{t}_jdG_t(\iota)g_t(\underline{t}_j)}{\int_{\underline{t}_j}^{\bar{t}}\iota dG_t(\iota)}<0,$$

while the rest of the terms on the right-hand side of (56) 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.

D.6 Proof of Corollary 1

We decompose the proof of Corollary 1 into two parts. In the first part, we show that conditional on employment, firms' labor productivity increases as z increases. In the second part, we prove that conditional on employment, we get $\frac{\partial^2 v_j^B}{\partial\tau_T\partial LP_j}>0$.

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

Given (49), (50), (51), and (52), 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:

$$\frac{\partial lnt_j}{\partial lnz_j}\mid l^B=-\frac{\sigma-1}{1+\Delta_1M_j-\frac{\Delta_1t_j}{\Theta_j}}<0.$$

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

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

Therefore:

$$\frac{\partial lnt_j}{\partial lnz_j}\mid v_B-\frac{\partial lnt_j}{\partial lnz_j}\mid l^B=-\frac{1+\Delta_1M_j-\Delta_1(\sigma-1)+(\sigma-1)}{1+\Delta_1M_j-\frac{\Delta_1t_j}{\Theta_j}}. \tag{57}$$

Recall that $\Delta_1\equiv1-\frac{(\sigma-1)}{\theta}$, hence (57) can be reduced to:

$$\frac{\partial lnt_j}{\partial lnz_j}\mid v_B-\frac{\partial lnt_j}{\partial lnz_j}\mid l^B=-\frac{1+\Delta_1M_j+\frac{(\sigma-1)^2}{\theta}}{1+\Delta_1M_j-\frac{\Delta_1t_j}{\Theta_j}}<0.$$

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 D.2. Because of constant markups, this in turn implies that $\frac{\partial l_j^T}{\partial t_j}>0$. Hence:

$$\frac{\partial lnt_j}{\partial lnz_j} \mid l = -\frac{z_j}{t_j}\frac{\frac{\partial l_j^B}{\partial z}}{\frac{\partial l_j^B}{\partial t_j}+\frac{\partial l_j^T}{\partial t_j}} > -\frac{t_j}{z_j}\frac{\frac{\partial l_j^B}{\partial t_j}}{\frac{\partial l_j^B}{\partial z}} \equiv \frac{\partial lnz_j}{\partial lnt_j} \mid l_B. \tag{58}$$

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

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

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 (58), it is easy to verify that $\frac{\partial lnt_j}{\partial lnz_j} \mid l < 0$. Therefore, we have $t_j < t_{j'}$. Recall that in Section D.5 we proved Proposition 2 and 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 $\frac{\partial^2 v_j^B}{\partial \tau_T \partial LP_j} \mid l > 0$.