Distribution, Outward FDI, and Productivity Heterogeneity: China and Cross-countries' Evidence *

Wei Tian† Miaojie Yu‡

June 6, 2019

Abstract

This paper examines distribution-oriented outward FDI using *Chinese* multinational firm–level data. Distribution outward FDI refers to Chinese parent firms in manufacturing that penetrate foreign markets through wholesale trade affiliates that resell exportable goods. Our estimations correct for rare-events bias and show that distribution FDI are more productive than non-FDI firms but less productive than non-distribution FDI firms. As cross-border communications costs (transportation costs) increase, there is a higher the probability that firms engage in distribution FDI (non-distribution FDI). Our endogenous income-threshold estimates show that high-productivity Chinese firms invest more in high-income countries, but not necessarily in low-income countries.

JEL: F13, O11, P51

Keywords: Distribution FDI, Firm Productivity, Rare-Events Corrections, Threshold Estimates

*We thank Cheng Chen, Guangzhou Li, Zhenyu Wu, two anonymous referees, seminar and conference participants of HKUST, Keio, Gakushuin, Nankai, Leuven, Hitotsubashi, the 17 NBER-CCER Annual Conference and the 1st Comparative Economics World Congress for their helpful comments and suggestions. Tian thanks University of International Business and Economics for financial support (No. 16PY33-71503043). Yu thanks China National Natural Science Foundation and China National Social Science Foundation for financial support (Grant No. 71573006, 71625007, 16AZD003). However, all errors are ours.

†School of International Trade and Economics, University of International Business and Economics, Beijing, China. Email: wei.tian08@gmail.com.

‡Corresponding author. China Center for Economic Research (CCER), National School of Development, Peking University, Beijing 100871, China. Phone: +86-10-6275-3109, Email: mjyu@nsd.pku.edu.cn.

1 Introduction

Distribution-oriented outward foreign direct investment (FDI) refers to the phenomenon of home parent manufacturing firms that penetrate foreign markets through wholesale trade affiliates that resell exportable goods. Distribution-oriented outward FDI is an important phenomenon in developed countries like the United States (Hanson et al. 2001), and in developing countries like China. However, there is relatively scant research on this topic. The present paper aims to fill this gap.

Outward FDI includes two main categories: distribution-oriented and non-distribution production-oriented FDI. Distribution FDI includes the business-service foreign affiliates and the wholesale foreign affiliates. The business-service FDI mainly refers to building overseas business office to explore foreign market, to promote sales, and to serve customers in the hosting countries. Similarly, the wholesale FDI refer to oversea intermediaries of parent firms to help exporting and sales in the host countries.

The wholesale foreign affiliates in the united states accounted for over 20% of total foreign sales by multinationals even in a decade ago. The number of wholesale foreign affiliates is around 50 percent of that of production foreign affiliates. In a developing country like China, the proportion of distribution FDI is even higher. In 2017, China's outward FDI flow accounts for 11.1 percent of global FDI flow and ranks second in the world, just following the United States. The share of China's distribution outward FDI increased from around 28 percent in 2004 to more than 51 percent. Within distribution FDI, the wholesale FDI flow accounts for 17% of the total FDI flow. By comparison, production FDI only accounts for around 18%.

Previous pioneering works such as Hanson et al. (2001) and Horstmann and Markusen (1996) make significant efforts for us to understand the characteristics of distribution FDI. Horstmann and Markusen (1996) argue that firms have two options in foreign markets: export or distribution FDI. Exporters need to find a local agent, which has private information advantage on its own effort and foreign market. Thus, home exporters have to pay additional information rent. By contrast, building a wholly-owned distribution affiliate requires extra fixed costs. So firms will make decisions by considering the trade-off between the two. By contrast, Hanson et al. (2001) implicitly assume that firm export and distribution-oriented FDI are complementary, as distribution FDI is set up to promote exports. They compare the trade-off between distribution- and production-oriented FDI and find that firms operating in countries with high income tax would prefer distribution FDI rather than production FDI, to avoid paying the high corporate tax. Our paper is mostly motivated by Hanson et al. (2001), in searching for the trade-off among exporting, distribution FDI, and production FDI.

However, we are not still entirely clear why some firms choose distribution FDI while others do not, and why distribution FDI is more popular in some countries like China than in other countries. What causes some firms to engage in distribution outward FDI? Moreover, which investment characteristics in the host country matter for firms to engage in distribution FDI? The present paper seeks to answer such questions. We argue that distribution FDI plays an important auxiliary but significant role to boost China's exports. In accompany with China's fast productivity growth in the new century (Feenstra et al., 2014), distribution FDI provides a cheaper alternative for a bunch of Chinese exporting firms to realize the cost-saving effects in reducing the cross-border communication costs.

The current paper presents four main findings. First, firms with distribution outward FDI are found to be more productive than non-FDI firms, but less productive than non-distribution FDI firms. These findings imply that the popularity of distribution outward FDI may be attributed to the fact that most Chinese exporting firms are insufficiently productive to set up overseas production lines. As a compromise, they set up a service or distribution center abroad to promote exports. This finding echoes the stylized fact that China's exports have increased rapidly in the new century. In addition, we find strong sorting behavior between production FDI, distribution FDI, and non-FDI exports. To explain these findings, inspired by Oldenski (2012), we extend the model of Helpman et al. (2004) to understand this sorting behavior. Our estimates are based on a comprehensive FDI decision data set covering all Chinese FDI manufacturing firms during 2000-08. However, it is important to stress that only a very small proportion of firms in our large sample engaged in FDI activity. Thus, the standard nonlinear binary estimates would have downward estimation bias (King and Zeng 2001). We thus correct for such rare-events estimation bias in the paper.

Second, we distinguish the cross-country (i.e., cross-border) communications costs that occur during distribution and sales (like the costs of import procedures, promoting goods, and services before and after sales) from the usual transportation costs (i.e., iceberg transportation costs and tariffs) to demonstrate the importance of distribution outward FDI for exporting firms. We find that the higher are the cross-border communications costs, the higher is the probability that firms engage in distribution outward FDI. By contrast, the higher are the iceberg transportation costs, the higher is the probability that firms engage in non-distribution (or production) outward FDI. These findings are intuitive in the sense that, by setting up a business office or wholesale and retail subsidiary, the firm can largely reduce the asymmetric rent charged by local agents (Horstmann and Markusen 1996). By contrast, firms can save on transportation costs when exporting is replaced by production FDI. These findings are also highly consistent with our theoretical predictions.

Third, by allowing for firm heterogeneity in choosing cross-country host destinations, we find that the role of a firm's productivity in its FDI flow differs by destination income. Highly productive firms are more likely to invest in rich countries, but not necessarily in poor countries. By estimating an endogenous threshold of income in host countries, our threshold regressions find that rich countries receive more FDI flows

Fourth, we find strong evidence on the intensive margin of distribution-oriented FDI. We find that firm productivity significantly boosts distribution FDI flow once firms self-select into distribution FDI. Different from previous studies on Chinese outward FDI, we were able to obtain confidential information on the outward FDI flow for total FDI flow and distribution FDI flow in Zhejiang province, one of the most important FDI provinces in China. This is a novel finding in the literature on understanding China's outward FDI, as the publicly released nationwide FDI decision data set has the substantial pitfall that data on firms' FDI flows are unavailable.

The paper makes the following three contributions to the literature. First, it enriches the understanding of distribution outward FDI. As documented by Boatman (2007), as distribution FDI does not save production costs, distribution FDI has received little attention in the literature from theoretical and empirical works, except a few exceptions, such as Horstmann and Markusen (1996), Hanson et al. (2001), and Kimura and Lee (2006). We show that distribution FDI is complementary to firm export as a type of downward vertical FDI. As illustrated in our theoretical framework, firms face a trade-off between variable cost and fixed cost. Firms engaged in exporting without FDI, regardless of distribution or production orientation, bear an additional variable cost of cross-border communications (Oldenski 2012).[1] However, firms engaged in

[1] Oldenski (2012) finds evidence that firms would prefer exporting if the activities require complex within-firm communication. Instead, firms would prefer FDI if the goods and services require direct communication with consumers. Based on Russ (2007), Ramondo et al. (2013) find that countries with less volatile fluctuations are

distribution FDI have a larger fixed cost. The trade-off between variable cost and fixed cost can be interpreted as a new form of the standard concentration-proximity trade-off. Thus, productivity heterogeneity plays an important role in understanding distribution FDI. Only highly productive firms would self-select into distribution FDI.

Second, the paper enriches the understanding of China's distribution FDI. Different from China's exports, on which there is already a fairly large micro-level literature (Qiu and Xue, 2014), few papers have investigated China's FDI. Kolstad and Wiig (2012) find that Chinese FDI is attracted to three destinations: countries with lower institutional quality, countries that are rich in natural resources, and large markets. Using the same universal nationwide FDI decision data set, Chen and Tang (2014) find that firm productivity and the probability of firm FDI are positively correlated. Wang et al. (2015) use China's firm-level data and find that access to external finance increases the probability that firms engage in outward FDI. Chen et al. (2018) explore how domestic distortions affect firms' outward FDI decision. We take one step forward to examine a large and important part of Chins's FDI–the distribution FDI. Our binary estimates find that the sorting predictions among non-FDI, distribution FDI, and production FDI work well in China. Thus, different from the mixed findings on Chinese exports and firm productivity,[2] we confirm that the sorting behaviors among domestic sales, exporting, and FDI proposed by Helpman et al. (2004) apply to Chinese FDI firms.

Third and more importantly, we explore the intensive margin of firm FDI flows (on all FDI and distribution FDI), which is almost completely absent in previous studies because of the unavailability of data. As introduced in detail in the next section, although the Ministry of Commerce of China released the list of FDI firms (henceforth, the FDI decision data set), the data set does not report each firm's FDI volume in all years. To overcome this data challenge, we accessed a confidential FDI data set compiled by the Department of Commerce in Zhejiang province, which reports firms' FDI volume in addition to all other information covered in the FDI decision data set. Thanks to this novel data set, we are able to explore the intensive margin of firm FDI in China.

served relatively more by foreign affiliates than by exporters.

[2] Lu (2010) finds that Chinese exporters are less productive. However, Dai et al. (2016) and Yu (2015) argue that that finding was because of the presence of China's processing exporters, which are less productive than non-exporters and non-processing exporters. Once processing exporters are excluded, Chinese exporters are more productive than non-exporters, in line with the theoretical predictions of Melitz (2003).

The rest of the paper is organized as follows. Section 2 extends Melitz et al. (2004) to show sorting equilibrium by productivity heterogeneity. Section 3 describes our data sample, followed by a careful scrutiny of measures of firm productivity. Section 4 examines the role of firm productivity in the firm's FDI decision. Section 5 explores the intensive margin of FDI flows. Section 6 discusses the firm's investment destination and Section 7 concludes.

2 Model

We construct a cross-country theoretical framework by extending Helpman et al. (2004) to capture the behavior of distribution FDI. We assume that each country has a representative constant elasticity of substitution utility function as follows:

$$U = (\int_{\Omega} x(\varphi)^{\frac{\sigma-1}{\sigma}} d\varphi)^{\frac{\sigma}{\sigma-1}}$$

where $x(\varphi)$ is the consumption of product φ, and $\sigma > 1$.

Each firm in country i produces one product using labor as the only input, and the firm has a random labor productivity φ following Pareto distribution, where $\Pr(\varphi > x) = (\frac{b}{x})^k, k > \sigma - 1, b > 1$. So $\frac{1}{\varphi}$ is the variable production cost for each unit of goods produced. The firm first decides whether to enter the market. If entry, a sunk cost of f_E is required. After the entry, the firm observes his productivity φ. to set up the production plant. If the firm would like to serve foreign countries, there are three possible ways: (1) export without any foreign investment, (2) export and also set up a foreign affiliate to promote exports, and (3) set up a foreign plant to produce and sell overseas. The firm must pay a fixed cost f_X for the first choice; a fixed cost $f_X + f_S$ for the second choice, where f_S is the up-front cost to set up a foreign affiliate; and a fixed cost f_M for the third choice to build a foreign plant. Here we assume that the fixed costs satisfy the following ranking $f_M > f_X + f_S > f_X > f_D$.[3] We will validate these assumptions in the empirical part of the paper as well.

An iceberg transportation cost $\tau_{ij} > 1$ is needed for export, which means τ_{ij} units of product are required for one unit sold in country j. But if the firm builds a distribution affiliation, the transportation cost may be reduced to $\mu\tau_{ij}$, $0 < \mu < 1$, $\mu\tau_{ij} > 1$. The discount factor μ

[3] Note that fixed costs for production FDI can be decomposed into two components: fixed cost for production (f_M^P) and fixed cost for setting up the firm's own distribution center (f_S) which is similar to the fixed cost of distribution FDI. As an usual assumption in the literature, fixed cost for production in production-type FDI is assumed to be higher than its counterpart for exports: $f_M^P > f_X$. We thus have $f_M = f_M^P + f_S > f_X + f_S$.

captures the cost reduction of investing in a trading subsidiary, which allows firms to distribute their products independently.

Oldenski (2012) points out that the expenses incurred during communications between the domestic firm and foreign customers are crucial when firms are making the decision whether to export or build an overseas plant. It is important to distinguish cross-border communications costs from transportation costs. Most of cross-border communications costs are incurred after the goods are transported to the destination and can be reduced by setting up a local business office, that is, distribution FDI, which makes the import procedure and service more effective.[4] However, transportation costs can be only phased out when the goods are no longer imported but produced locally, that is, via production FDI. Another difference is that most of cross-border communications costs are irrelevant to firm productivity, since those costs are incurred after the transportation. Transportation costs are iceberg costs, which vary across firms with different productivity.

To capture these aspects, similar to Berman et al. (2012), we introduce a linear cross-border communication cost in our model. We assume that firms that only export have to pay η_j units of labor for the communications costs additional to production costs, but those who build an overseas distribution foreign affiliate do not. The value of η_j captures the cost-saving effects from establishing a business office, which helps firms to serve foreign customers by promoting sales and improving after-sales services. In this way, a destination country with a poor doing-business environment may be associated with a poor record in enforcing contracts, which would generate more communications costs. Different from production FDI, which saves transportation costs, distribution FDI mainly reduces the cross-border communications costs incurred.

As in Helpman et al. (2004), wages (w) are equal to unity across countries by introducing a homogenous good sector in which one unit of labor is used to produce one unit of output. The homogenous good can be traded freely and an exogenous fraction of income is spent on it. The marginal cost for each product sold, $MC^d = \frac{w}{\varphi}, MC^e = \frac{\tau_{ij}w}{\varphi} + \eta_j w, MC^s = \frac{\mu\tau_{ij}w}{\varphi}, MC^m = \frac{w_j}{\varphi}$, represents the marginal cost for selling in the domestic market, exporting without foreign investment, exporting as well as distribution investment, and building a foreign production plant.

[4]In practice, some communication costs such as consulting and negotiation could occur even before a trade deal. We thank a referee for pointing this out.

The derived demand for product φ is

$$X_j(\varphi) = L_j P_j^{\sigma-1}[p_j^c(\varphi)]^{-\sigma}$$

where L_j is labor income in country j, $p_j^c(\varphi) = \frac{\sigma}{\sigma-1} MC^c, c = d, e, s, m$ is the price of product φ if it is sold domestically, exported without a foreign distribution affiliate, exported with a distribution affiliate, and exported with a production affiliate, respectively. P_j is the aggregate price level in which its exact expression is shown in Appendix A. Inspired by Berman et al. (2012), the profits for domestic sales, exports, distribution FDI, and production FDI are as follows:

$$\begin{aligned}
\pi_i^d &= (\frac{1}{\varphi})^{1-\sigma} B_i - f_D && (1)\\
\pi_{ij}^e &= (\frac{\tau_{ij}}{\varphi} + \eta_j)^{1-\sigma} B_j - f_X && (2)\\
\pi_{ij}^s &= (\frac{\mu_j \tau_{ij}}{\varphi})^{1-\sigma} B_j - f_X - f_S && (3)\\
\pi_{ij}^m &= (\frac{1}{\varphi})^{1-\sigma} B_j - f_M\ , && (4)
\end{aligned}$$

where $B_j \equiv \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} L_j P_j^{\sigma-1}$. The productivity cut-off points satisfy $\pi_i^d = 0, \pi_{ij}^e = 0, \pi_{ij}^s = \pi_{ij}^e, \pi_{ij}^m = \pi_{ij}^s$ explicitly:

$$\begin{array}{ll}
(\frac{1}{\widehat{\varphi_{dj}}})^{1-\sigma} = \frac{f_D}{B_i}; & (\frac{\mu_j \tau_{ij}}{\widehat{\varphi_{sj}}})^{1-\sigma} - (\frac{\tau_{ij}}{\widehat{\varphi_{sj}}} + \eta_j)^{1-\sigma} = \frac{f_S}{B_j} \\
(\frac{\tau_{ij}}{\widehat{\varphi_{ej}}} + \eta_j)^{1-\sigma} = \frac{f_X}{B_j}; & (\frac{1}{\widehat{\varphi_{mj}}})^{1-\sigma} - (\frac{\mu_j \tau_{ij}}{\widehat{\varphi_{mj}}})^{1-\sigma} = \frac{f_M - f_S - f_X}{B_j}
\end{array}$$

where $\widehat{\varphi_{dj}}, \widehat{\varphi_{ej}}, \widehat{\varphi_{sj}}, \widehat{\varphi_{mj}}$ is the productivity cut-off point for each mode, respectively. As free entry, the expected profit of firm entry is zero. The expected profit after entry equals the entry cost f_E:

$$\int_{\widehat{\varphi_{di}}}^{\infty} \pi_i^d dG(\varphi) + \sum_{j=1, j\neq i}^{N} \left[\int_{\widehat{\varphi_{eij}}}^{\widehat{\varphi_{sij}}} \pi_{ij}^e dG(\varphi) + \int_{\widehat{\varphi_{sij}}}^{\widehat{\varphi_{mij}}} \pi_{ij}^s dG(\varphi) + \int_{\widehat{\varphi_{mij}}}^{\infty} \pi_{ij}^m dG(\varphi) \right] = f_E \qquad (5)$$

Equations (1) to (5) jointly solve the equilibrium $\widehat{\varphi_{dj}}$, $\widehat{\varphi_{eij}}$, $\widehat{\varphi_{sij}}$, $\widehat{\varphi_{mij}}$, and B_j for each country i, j. Note that the equilibrium is irrelevant to the market size (L_j). For simplicity, countries are assumed to be symmetric following Melitz (2003). As $\eta_j = \eta$, $\tau_{ij} = \tau, \mu_j = \mu$, every country has the same productivity cut-off points $\widehat{\varphi_d}$, $\widehat{\varphi_e}$, $\widehat{\varphi_s}$, $\widehat{\varphi_m}$, and B. We thus have following findings.

Proposition 1 *When every country is symmetric,* $\frac{f_X}{f_D} > \tau^{1-\sigma}$, $f_X^{\frac{1}{1-\sigma}} - \frac{1}{\mu}(f_X + f_S)^{\frac{1}{1-\sigma}} > \eta\Delta$, $f_M > f_X + f_S \frac{\mu^{\sigma-1}}{1-\mu^{\sigma-1}}(\tau^{\sigma-1} - 1)$ *where* Δ *is any upper bound of* $B^{\frac{1}{1-\sigma}}$, *we have* $\widehat{\varphi_d} < \widehat{\varphi_e} < \widehat{\varphi_s} < \widehat{\varphi_m}$.

Proof. See Appendix A for details. ■

Proposition 1 suggests that the most productive firms engage in production FDI, the next most productive firms engage in distribution FDI and export, the even next most productive firms only export, the further next productive firms do not export but only sell in the domestic market, and the least productive firms exit. The intuition is straightforward: only the most productive firms can overcome the highest fixed costs to build an overseas production plant and benefit from the cost-saving effect of cross-border communications costs and transportation costs. Less productive firms, like most of the Chinese FDI firms, can only afford the fixed costs of building international business services or distribution centers to reduce cross-border communications costs to promote their exports. The sorting equilibria for different cutoff points are shown in Figure 1.

[Insert Figure 1 Here]

Proposition 2 *(i) An increase in export-specific communication cost* η *raises* $\widehat{\varphi_e}$, *lowers* $\widehat{\varphi_s}$, *but does not affect* $\widehat{\varphi_m}$.

(ii) An increase in iceberg transportation cost τ *increases* $\widehat{\varphi_e}$ *and* $\widehat{\varphi_s}$, *and decreases* $\widehat{\varphi_m}$.

Proof. *See Appendix B for details.* ■

Proposition 2 implies that higher cross-country cross-border communications costs η and lower foreign tariffs (lower τ) increase the probability of distribution foreign investment. This is because most of the cross-border communications costs can be reduced via distribution FDI. Thus, a higher increases the attractiveness of distribution FDI compared with exporting only, but does not alter the benefit of production FDI. However, the transportation costs still exist as long as goods are exported. So a higher tariff imposed by importing countries promotes production FDI and hampers export and distribution FDI. We now turn to test these theoretical predictions.

3 Data and Measures

To investigate the impact of firm productivity on distribution FDI, we rely on three disaggregated data sets. The first data set provides the list of FDI firms in China. This data set is crucial for understanding firms' FDI decision. However, the data set does not report any FDI values. To examine the role of the intensive margin, we rely on another firm-level FDI data set, which contains information on the universal firm-level FDI activity in Zhejiang province of China. Finally, we merge the firm-level manufacturing production data with the two FDI data sets to explore the nexus between FDI and firm productivity.

3.1 FDI Decision Data

The nationwide data set of Chinese firms' FDI decisions was obtained from the Ministry of Commerce of China (MOC). MOC requires every Chinese FDI firm to report its detailed investment activity since 1980. To invest abroad, every Chinese firm is required by the government to apply to the MOC and its former counterpart, the Ministry of Foreign Trade and Economic Cooperation of China, for approval and registration. MOC requires such firms to provide the following information: the firm's name, the names of the firm's foreign subsidiaries, the type of ownership (i.e., state-owned enterprise (SOE) or private firm), the investment mode (e.g., trading-oriented affiliates, mining-oriented affiliates), and the amount of foreign investment (in U.S. dollars). Once a firm's application is approved by MOC, MOC will release the information mentioned above, as well as other information, such as the date of approval and the date of registration abroad, to the public. All such information is available except the amount of the firm's investment, which is considered to be confidential information to the firms.

Since 1980, MOC has released information on new FDI firms every year. Thus, the nationwide FDI decision data indeed report FDI starters by year. The database even reports specific modes of investment: trading office, wholesale center, production affiliate, foreign resource utilization, processing trade, consulting service, real estate, research and development center, and other unspecified types. Here trading offices and wholesale centers are classified as distribution FDI, whereas the rest are referred to as non-distribution FDI. However, since this data set does not report firms' FDI flows, researchers are not able to explore the intensive margin of firm FDI with this data set.

3.2 FDI Flow Data

To explore the intensive margin, we use another data set, which is compiled by the Department of Commerce of Zhejiang province. The most novel aspect of this data set is that it includes data on firms' FDI flows (in current U.S. dollars). The data set covers all firms with headquarters located (and registered) in Zhejiang and is a short, unbalanced panel from 2006 to 2008. In addition to the variables covered in the nationwide FDI data set, the Zhejiang data set provides each firm's name, city where it has its headquarters, type of ownership, industry classification, investment destination countries, and stock share from its Chinese parent company.

Although this data set seems ideal for examining the role of the intensive margin of firm FDI, the disadvantage is also obvious: the data set is for only one province in China.[5] Regrettably, as is the case for many other researchers, we cannot access similar databases from other provinces. Still, as discussed in Appendix C, we believe that Zhejiang's firm-level FDI flow data are a good proxy for understanding the universal Chinese firm's FDI flows. In particular, the FDI flows from Zhejiang province are outstanding in the whole of China; the distribution of both types of ownership and that of Zhejiang's FDI firms' destinations and industrial distributions are similar to those for the whole of China.

3.3 Firm-Level Production Data

Our last database is the firm-level production data compiled by China's National Bureau of Statistics in an annual survey of manufacturing enterprises. The data set covers around 162,885 firms in 2000 and 410 thousand firms in 2008 and, on average, accounts for 95 percent of China's total annual output in all manufacturing sectors. The data set includes two types of manufacturing firms: universal SOEs and non-SOEs whose annual sales are more than RMB 5 million (or equivalently around $746,000 under the current exchange rate). The data set is particularly useful for calculating measured total factor productivity (TFP), since the data set provides more than 100 firm-level variables listed in the main accounting statements, such as sales, capital, labor, and intermediate inputs.

As highlighted by Yu (2015) and Chen et al. (2018), some samples in this firm-level produc-

[5] To our knowledge, almost all previous work was not able to access nationwide universal outward FDI flow data. An outstanding exception is Wang et al. (2012), who use nationwide firm-level outward FDI data to investigate the driving force of outward FDI of Chinese firms. However, the study uses data only from 2006 to 2007; hence, it cannot explore the possible effects of the financial crisis in 2008.

tion data set are noisy and somewhat misleading, largely because of mis-reporting by some firms. To guarantee that our estimation sample is reliable and accurate, we screen the sample and omit outliers by adopting the following criteria (Feenstra et al., 2014). First, we eliminate a firm if its number of employees is less than eight workers, since otherwise such an entity would be identified as self-employed. Second, a firm is included only if its key financial variables (e.g., gross value of industrial output, sales, total assets, and net value of fixed assets) are present. Third, we include firms based on the requirements of the Generally Accepted Accounting Principles.[6]

3.4 Data Merge

We then merge the two firm-level FDI data sets (i.e., nationwide FDI decision data and Zhejiang's FDI flow data) with the manufacturing production database. Although the two data sets share a common variable—the firm's identification number—their coding systems are completely different. Hence, we use alternative methods to merge the three data sets. The matching procedure involves three steps. First, we match the three data sets (i.e., firm production data, nationwide FDI decision data, and Zhejiang FDI flow data) by using each firm's Chinese name and year. If a firm has an exact Chinese name in a particular year in all three data sets, it is considered an identical firm. Still, this method could miss some firms since the Chinese name for an identical company may not have the exact Chinese characters in the two data sets, although they share some common strings.[7] Our second step is to decompose a firm name into several strings referring to its location, industry, business type, and specific name, respectively. If a company has all identical strings, such a firm in the three data sets is classified as an identical firm.[8] Finally, to avoid possible mistakes, all approximate string-matching procedures are done manually.

Row (1) of Table 1 reports the number of manufacturing firms and row (2) reports the number of FDI starting firms by year during 2000-08. Row (3) reports the number of matching

[6] In particular, an observation is included in the sample only if the following observations hold: (1) total assets are greater than liquid assets; (2) total assets are greater than the total fixed assets and the net value of fixed assets; (3) the established time is valid (i.e., the opening month should be between January and December); and (4) the firm's sales must be higher than the required threshold of RMB 5 million.

[7] For example, "Ningbo Hangyuan communication equipment trading company" shown in the FDI data set and "(Zhejiang) Ningbo Hangyuan communication equipment trading company" shown in the National Bureau of Statistics of China production data set are the same company but do not have exactly the same Chinese characters.

[8] In the example above, the location fragment is "Ningbo," the industry is "communication equipment," the business type is "trading company," and the specific name is "Hangyuan."

FDI manufacturing firms.[9] The share of FDI manufacturing firms over total manufacturing firms shown in row (5) suggests that FDI indeed is a rare event—the share is less than 1 percent each year. The number of FDI manufacturing firms increased dramatically after 2004. More importantly, row (6) shows that the share of distribution FDI manufacturing firms over total FDI manufacturing firms increased from around 14 percent in 2000 to 55 percent in 2008, suggesting that distribution FDI has become more and more important over time.

[Insert Table 1 Here]

By using these two methods, we match Zhejiang's manufacturing firms with Zhejiang's FDI flow firms. As shown in the lower module of Table 1, of 1,270 FDI firm-years in Zhejiang province from 2006 to 2008, 407 FDI firms are engaging in manufacturing sectors, suggesting that around two-thirds of Zhejiang FDI parent firms are from service sectors or are trading intermediates (Ahn et al. 2010). Table 2 reports the summary statistics of firm characteristics for nationwide manufacturing firms and Zhejiang's manufacturing firms, respectively. The small mean of FDI indicator in both samples ascertains that FDI is a rare event during the sample periods.

Finally, as the main interest of this paper is how firm productivity affects distribution FDI, we carefully measure TFP. The augmented Olley-Pakes TFP is constructed following Brandt et al. (2012) and Yu (2015). Appendix D provides the detailed steps of our measured TFP. In particular, we estimate the production function for exporting and non-exporting firms separately in each industry. The idea is that different industries may use different technology; hence, firm TFP must be estimated for each industry. Equally important, even within an industry, exporting firms may use completely different technology than non-exporting firms. For example, some exporters, like processing exporters, only receive imported material passively (Feenstra and Hanson 2005) and hence do not have their own technology choice. We hence estimate TFP for exporters and non-exporters separately.

[Insert Table 2 Here]

We now turn to describe distribution FDI in our merged data set. In both FDI data sets, there is a variable used to describe the type of firm FDI, which includes mining, construction,

[9] Note that we merge FDI data and manufacturing production data by firm name rather than by name-year. Number of FDI manufacturing firms in row (3) reports not only FDI starting firms, but also FDI continuing firms. Thus, it is possible that there are fewer FDI starters than *matched* FDI manufacturing firms, as shown in 2007 and 2008.

R&D, production, processing trade, market seeking, wholesale, business service, and product design. As our main interest is of distribution FDI, both wholesale FDI and business-service FDI are classified to distribution FDI, following the official definition of MOC of China. Appendix Table 1 reports the proportion of distribution FDI in our sample. In the nationwide FDI data, the number of distribution FDI firms accounts for roughly half of whole FDI firm. Such a proportion even increases to 60% after merging with the production data set. Similarly, nearly 76% samples are distribution FDI in Zhejiang FDI data. The percentage also rises to 80% after merging with production data. All these suggest that distribution FDI is important in China today.

4 Extensive Margin of FDI

This section discusses how a firm's productivity affects the firm's decision to engage in FDI (i.e., the extensive margin). Before running the regressions, we provide several preliminary statistical tests to enrich our understanding of the difference in productivity between *distribution* FDI and non-FDI firms (and non-distribution FDI firms), following a careful scrutiny of the effect of firm productivity on the decision to engage in (distribution) FDI.

4.1 Descriptive Analysis on Productivity Differences

Proposition 1 suggests that firms' sales decision can be sorted by their productivity. Low-productivity firms serve in domestic markets, high-productivity firms export, higher-productivity firms engage in distribution FDI, and even higher-productivity firms participate in non-distribution FDI. Figure 2 exhibits the productivity distributions for non-FDI firms, distribution FDI firms, and non-distribution FDI firms, respectively. Overall, firm productivity for distribution FDI is clearly higher than for non-FDI firms, but lower than (though not obvious) for non-distribution FDI.

[Insert Figure 2 Here]

Eaton *et al.* (2011) find that higher-productivity firms are usually larger. If so, we would observe that, compared with non-FDI firms, FDI firms on average are larger, more productive, and export more. Table 3 checks the difference between non-FDI and FDI firms on their TFP, labor, sales, and exports. Compared with non-FDI firms, distribution FDI firms are found to be more productive, hire more workers, sell more, and export more. By sharp contrast, compared with non-distribution FDI firms, distribution FDI firms are found to be less productive, hire fewer workers, sell less, and export less. The t-values for these variables are strongly significant at the conventional statistical level.[10]

[Insert Table 3 Here]

However, the simple t-test comparisons may not be sufficient to conclude that distribution FDI firms are more productive than non-FDI firms, since FDI firms are very different from non-FDI firms in terms of size (number of employees and sales) and experience in foreign markets, as already seen.

We thus follow Imbens (2004) and perform propensity score matching (PSM) by choosing the number of firm employees, firm sales, and firm exports as covariates. Each FDI firm is matched to its most similar non- FDI firm. Since there are observations with identical propensity score values, the sort order of the data could affect the results. We thus perform a random sort before adopting the PSM approach. Column (3) in Table 3 reports the estimates for average treatment for the treated (ATT). The coefficient of ATT for distribution FDI manufacturing firms is 0.442 (compared with non-FDI firms) and highly statistically significant, suggesting that, overall, productivity for distribution FDI firms is higher than that for similar non-FDI firms during the period 2000-08. Strikingly, compared with non-distribution FDI firms, the coefficient of ATT for distribution FDI is insignificant.

To check this out, we examine productivity difference by year for each type of firm: non-FDI, distribution FDI, and non-distribution FDI firms. Table 4 shows that FDI firms are more productive than non-FDI firms by year during the sample period 2000-08.[11] The productivity difference between distribution FDI firms and non-FDI firms is significantly positive before 2003. However, this might be purely due to the fact that only few FDI firms engage in distribution in the early years. Interesing, the gap roughly declines over the period (especially after 2004), also

[10] Note that the sample size of non-FDI is 1,137,907 which is lower than the number of observation (i.e., 1,553,740) shown in Table 1 due to the fact that some firms' TFP are missing if any of following data of firm's capital, labor, intermediate inputs are unavailble.

[11] Note that TFP in 2008 is calculated and estimated differently. As in Feenstra et al. (2014), we use deflated firm value added to measure production and exclude intermediate inputs (materials) as one kind of factor input. However, we are not able to use value added to estimate firm TFP in 2008, since it is absent in the data set. We instead use industrial output to replace value added in 2008. Thus, we have to be cautious in comparing TFP in 2008 with TFP in previous years.

suggesting that distribution FDI firms indeed enjoy much productivity gain via learning from investing.

[Insert Table 4 Here]

4.2 Extensive Margin of FDI

To examine whether firm productivity plays a key role in the firm's decision to engage in distribution FDI, we start by checking whether productivity affects the firm's FDI decision, as distribution FDI is a type of FDI. In particular, we consider the following empirical specification:

$$\Pr(OFDI_{ijt} = 1) = \beta_0 + \beta_1 \ln TFP_{it} + \boldsymbol{\theta}\mathbf{X} + \varpi_j + \eta_t + \varepsilon_{it}, \tag{1}$$

where $OFDI_{ijt}$ and $\ln TFP_{ijt}$ represents FDI indicator and the log productivity of firm i in industry j in year t, respectively. $\mathbf{X}$ denotes other firm characteristics, such as firm size (produced by firm's log of employment) and types of ownership (i.e., foreign invested firms or SOEs).[12] For instance, private firms are more (equivalently, SOEs might be less) likely to invest abroad because of domestic input distortion in China (Hsieh and Klenow 2009, Chen et. al, 2018). In addition, larger firms are more likely to invest abroad because they may have an additional advantage to realize increasing returns to scale (Helpman et al, 2004). Inspired by Oldenski (2012), we also include a firm export indicator in the estimations, since an exporting firm could find it easier to invest abroad, given that it would have an information advantage on foreign markets compared with non-exporting firms. Moreover, as the measured TFP cannot be compared over industries, we normalize TFP in each industry to a range between zero and one, following Arkolakis and Muendler (2011) and Groizard et al. (2014).

Finally, as stressed by Ishikawa et al. (2010) and Ishikawa and Morita (2015), a host country has some regulations for foreign investment. Such a concern may be relevant and important for Chinese FDI, in particular in the mining industries. Although Chinese parent firms in mining industries are not covered in our data set, foreign investment regulation may be present for some manufacturing industries. To this purpose, the error term is decomposed into three components: (1) industry-specific fixed effects, (2) year-specific fixed effects η_t to control for

[12] Here, a firm that has investment from foreign countries or Hong Kong/Macao/Taiwan is defined as a foreign firm, following Feenstra et al. (2014).

firm-invariant factors such as Chinese RMB appreciation, and (3) an idiosyncratic effect ε_{it} with normal distribution $\varepsilon_{it} \sim N(0, \sigma_i^2)$ to control for other unspecified factors. Industry and year fixed effects are used to capture possible industry heterogeneity due to foreign regulations and other possible industry-variant and year-variant factors.

We start from a simple linear probability model (LPM) to conduct our empirical analysis. It is worthwhile to stress that it is inappropriate to perform firm-specific fixed effects here, given that our nationwide outward FDI data are pooled cross-section data, as we only know the year that firms start to engage in FDI but do not know the year that firms continue or cease FDI. Table 5 (except the last column) thus only includes observations with FDI starters and non-FDI firms. We include the two-digit Chinese industry classification (CIC) level industry-specific fixed effects in the LPM estimates in column (1) in Table 5. The key coefficient of firm TFP is positive and significant, although its magnitude seems very small.[13] We suspect that this is due to the well-known drawback of using the linear probability model, which is that there is no justification for why the specification is linear. In addition, the predicted probability could be less than zero or greater than one, which does not make sense. We therefore perform the probit and logit estimations using two-digit CIC-level fixed effects in columns (2) and (3), respectively, and the result is confirmed.[14]

[Insert Table 5 Here]

4.3 Estimates with Rare Events Corrections

Our estimations above may still face some bias. As observed from Tables 1 and 2, of the total 1,138,450 observations, on average only 0.44 percent of firms engage in FDI. Thus, our sample exhibits the features of rare events that occur infrequently but may have important economic implications. As highlighted by King and Zeng (2001, 2002), standard econometric methods such as logit and probit would underestimate the probability of rare events, although maximum likelihood estimators are still consistent. To see this, consider a simplified logit regression of the

[13] Note that one must be cautious on the statistical significance in the LPM estimation since the LPM model has serious problem of heteroscedasticity. We thank a referee to point this out.

[14] Note that the coefficients shown in the Probit estimates are not marginal effects, which are not reported here given that it is straightforward to calculate the marginal effects in the rare event Logit estimates in column (4).

FDI dummy on firm TFP.

$$\Pr(OFDI_{it} = 1) = \Lambda(\beta_1 \ln TFP_{it}) = \frac{\exp(\beta_1 \ln TFP_{it})}{1 + \exp(\beta_1 \ln TFP_{it})}, \tag{2}$$

where $\Lambda(\cdot)$ is the logistic cumulative density function (henceforth CDF). Since $\hat{\beta}_1 > 0$, as shown in columns (1)-(3) of Table 5, the probability of $OFDI_{it} = 1$ is positively associated with firm TFP; most of the zero-FDI observations will be to the left and the observation with $OFDI_{it} = 1$ will be to the right with little overlap. Since there are around 1.5 million observations with zero FDI, the standard binary estimates can easily estimate the illustrated probability density function curve without error, as shown by the solid line in Figure 3.[15] However, since only 0.44 percent of the observations have positive FDI, any standard binary estimates of the dashed density line for firm TFP when $OFDI_{it} = 1$ will be poor. Because the minimum of the observed rare FDI sample is larger than that of the unobserved FDI population, the cutoff point that best classifies non-FDI and FDI would be too far from the density of observations with $OFDI_{it} = 1$. This will cause a systematic bias toward the left tail and result in an underestimation of the rare events with $OFDI_{it} = 1$ (See King and Zeng (2001, 2002) for a detailed discussion).

As recommended by King and Zeng (2001, 2002), the rare-events estimation bias can be corrected as follows. We first estimate the finite sample bias of the coefficients, $bias(\hat{\boldsymbol{\beta}})$, to obtain the bias-corrected estimates $\hat{\boldsymbol{\beta}} - bias(\hat{\boldsymbol{\beta}})$, where $\hat{\boldsymbol{\beta}}$ denotes the coefficients obtained from the conventional logistic estimates.[16] Column (4) in Table 5 reports the logit estimates with rare-events corrections. The coefficient of firm TFP is slightly larger than its counterpart in column (5), suggesting that the estimation bias is not so severe.

An alternative approach to correct possible rare-events estimation errors is to use the complementary log-log model.[17] The idea is that the distributions of standard binary nonlinear models, such as probit and logit, are symmetric to the original point. So the speed of convergence toward the probability that $OFDI_{it} = 1$ is the same as that for $OFDI_{it} = 0$. This violates the feature of the rare events, which exhibit faster convergence toward the probability that $OFDI_{it} = 1$. The

[15] To illustrate the idea in a simple way, the distribution curves are drawn to be normal, although this need not be the case.

[16] Chen (2014) also adopts this method to explore how negative climate shocks (e.g., severe drought, locust plagues) affected peasant uprisings.

[17] The CDF of the complementary log-log model is $C(\mathbf{X}'\boldsymbol{\beta}) = 1 - \exp(-\exp(\mathbf{X}'\boldsymbol{\beta}))$ with margin effect $\exp(-\exp(\mathbf{X}'\boldsymbol{\beta}))\exp(\mathbf{X}'\boldsymbol{\beta})\boldsymbol{\beta}$. The complementary log-log model also have an additional advantage to avoid a strong assumption of normal distribution in the Probit model, which seems unlikely in our data.

complementary log-log model can address this issue, since the model has a left-skewed extreme value distribution, which also exhibits a faster convergence speed toward the probability that $OFDI_{it} = 1$ (Cameron and Trivedi, 2005). The complementary log-log model in column (5) in Table 5 shows that the coefficient of firm TFP is fairly close to its counterparts in conventional logit estimates and rare-events logit estimates, suggesting that the estimation bias caused by the property of "rare events" is not so severe in our estimates. One possible reason is that we still do not control for possible reverse causality of FDI on firm productivity, which will be addressed shortly.

So far we include foreign multinational firms in the regressions. But there may be a concern that such foreign firms do not really fit with our analysis for two reasons. First, we only observe a selected sample of foreign firms that have already chosen to be present in China. Second, it is possible that some Chinese domestic firms invest in Hong Kong and Macao and hence should be treated as "multinational" firms, which in turn invest back in China. To avoid such bound-back behavior, we drop foreign firms in column (6) and still find similar results.

Another issue is about our FDI decision data per se. As we only observe firms that engage in new FDI, it is good enough for us to examine firms that transition from non-FDI to (any type of) FDI. However, as we do not have information on firms exiting FDI, we are not able to control for this. A possible concern is that some firms were SOEs but then were privatized. Since these firms may have made FDI decisions in the past that were not profit maximizing, once privatized, the firms may decide to unload assets that are not profitable. We indeed observe some indirect evidence from the regressions. The coefficients of SOEs in our previous tables are negative and significant, suggesting that SOEs are less likely to engage in FDI activity.

To address this concern, we run two experiments. First, we drop SOEs from the sample to see whether our main result is affected by SOEs. The estimates in column (7) in Table 5 show that our main results are not changed by doing so. Finally, for the sake of completeness, we include firms that ever employed FDI and non-FDI firms in the last column in Table 5.[18] In any case, our benchmark findings are insensitive to such robustness checks. Highly productive firms are more likely to engage in FDI.

[18] If a firm ever engaged in FDI, we assume that it always engages in FDI afterward during the sample.

4.4 Multinomial Logit Estimates with Distribution FDI

We now examine whether this finding—that highly productive firms are more likely to engage in FDI—applies to distribution FDI firms. Table 6 is our first key table. The regressands in columns (1)-(2) are FDI mode, in which zero refers to non-FDI, one is distribution FDI, and two is non-distribution FDI. We again use firm relative TFP to measure firm productivity in all estimates. As firms' decisions to engage in non-FDI or distribution FDI or non-distribution FDI are made simultaneously, we adopt the multinomial logit model in which the regressand in column (1) is distribution FDI, whereas that in column (2) is non-distribution FDI. The positive and significant sign of firm productivity in column (1) suggests that highly productive firms are more likely to engage in distribution FDI than non-FDI. The coefficient of firm productivity in column (2) is again positive and significant. More importantly, its magnitude is larger than its counterpart in column (1), suggesting that even higher productive firms are more likely to engage in non-distribution FDI. Finally, we find that larger firms are more likely to invest abroad, whereas SOEs are less likely to do so. Exporting firms are more likely to engage in distribution FDI by employing their information advantage (Oldenski 2012), which is in line with the intuition that distribution FDI serves trade.

There are four important caveats for these key findings. First, our theoretical model and empirical regressions discuss three options of firm choice: non-FDI, distribution FDI, and non-distribution FDI. However, it is possible that some firms do not directly export their products and hence have no incentive to set up their own distribution center abroad. Instead, they may rely on domestic trade intermediaries to sell their products abroad (Ahn et al. 2010).[19] As we have already dropped those firms, our current estimates would not suffer from such a concern.

Second, another possible option buried in firms' non-FDI choice is that firms contract with outside firms to undertake distribution for them. This is particularly true when firms export intermediate inputs.[20] Of course, in an Antras-Helpman (2004) type of setting like ours, such firms will be less productive than firms that undertake distribution themselves. However, our ranking of firm productivity will only hold where there are incomplete contracts in distribution that make integration an attractive option. There may be a concern about whether the ranking is still valid if some industries are more or less perfectly contractible (Feenstra and Hanson 2005).

[19] As in Ahn et al. (2010), about 20% of observations are identified and dropped in the customs data.

[20] By contrast, firms that export final goods and have no own distribution center, by default, have to find local agents to distribute their products.

To address this concern, we first identify the 20 most contract-intensive three-digit-level Chinese industries strictly following Nunn (2007).[21] Such industries mainly concentrate on equipment manufacturing and electronic components. By dropping from the sample industries in which firms almost always contract out distribution, the estimates in columns (3) and (4) with our restricted sample confirm that our previous findings are still strongly robust.

The third caveat is the striking finding (in columns (1)-(4)) that foreign (i.e., non-Chinese) invested firms are less likely to engage in FDI activity. One possible reason is that most foreign firms engage in processing trade, as found in Dai et al. (2016). Usually, processing exporters are less productive and enjoy special tariff treatment in China (Yu 2015). Such firms do not fit with our story and need to be dropped. Since the firm-level production data do not include firms' processing status, we instead drop from the sample pure exporters, that is, firms that sell all their products abroad, by taking advantage of the fact that processing firms have to export all their products by law. The multinomial logit estimates in columns (5)-(6) without foreign firms and pure exporters show robust evidence. Another possiblity that foreign firms are less likely to engage in distribution FDI is due to the fact that foreign invested firms actually have clear foreign customers and hence they do not need foreign distribution subsidiaries. By directly dropping foreign firms, the estimates in columns (7) and (8) yield similar results.

The last caveat is on mergers and acquisitions (henceforth M&A). There may be a concern that non-FDI firms may acquire a (domestic FDI or foreign) firm to use its distribution center as well. If so, our previous regressions may suffer estimation bias as even low-productive non-FDI firms can have their own foreign distribution network. However, this is not a problem if a non-FDI firm acquires a domestic FDI firm. In this case, the firm indeed has to report such an activity to the MOC and is classified as an FDI firm.[22] By contrast, there would be some estimation bias if a firm directly acquires foreign firms. Even in this case, the firm is identified as FDI firms in next year. To rule out this situation, we use the nationwide M&A data compiled by Bloomberg to identify Chinese non-FDI manufacturing firms with complete foreign acquisition deals.[23] Columns (7) and (8) drop foreign-invested firms and non-FDI firms with foreign acquisitions and still find robust results.

[21] We first make concordance between North American IO six-digit and HS eight-digit codes, following another concordance between HS eight-digit and Chinese Industries Classification (CIC) three-digit codes.

[22] Our FDI decision data set includes M&A activities, although it does not have a variable to stand this out.

[23] We thank Cheng Chen of HKU to kindly share us with such data.

Finally, to ensure that the findings above are not driven by the mass of non-FDI firms, in column (9) we also drop non-FDI firms and perform the logit estimates in which the regressand is the distribution FDI indicator (i.e., zero refers to non-distribution FDI and one refers to distribution FDI). It turns out that firm TFP has negative and significant coefficients, which is consistent with the sorting behavior illustrated by our theoretical model above.

[Insert Table 6 Here]

4.5 Endogeneity of Firm Productivity

Table 4 shows that the productivity mean of firms engaging in (distribution) FDI is increasing over time, suggesting that firms may have learning effects from investing. Firms that engage in investment may be able to absorb better technology or gain managerial efficiency from host countries (Oldenski 2012), which in turn boosts firm productivity. To exclude this effect, the sample we use only includes manufacturing non-FDI firms and manufacturing FDI starters, which means as long as the firm starts to invest abroad, it will no longer appear in the sample the next year. But the potential spillover effect of existing FDI firms may also lead to a possible endogeneity problem.

To mitigate the endogeneity issue, we adopt an instrumental variable approach. Admittedly, it is an empirical challenge to find an ideal instrument. Here we use the lag of firms' on-the-job training expenses as the instrument of firm productivity. The economic rationale is straightforward. As highlighted by Acemoglu and Pischke (1998) and Yeaple (2005), firms with more on-the-job training expenses usually are more productive. However, firms with more training expenses will not necessarily have more FDI. A one-year time lag is also helpful to avoid that possibility that firms' FDI decision reversely affects last year's on-the-job training. The simple correlation between firms' FDI decision and firm's lagged training expenses is close to nil (0.06), as shown in the sample. Note that we only have training data for 2004-2007. Thus, our IV estimates cover observations during 2005-2008 only.

We perform IV probit estimates in column (1) in Table 7. In column (2) we once again use the rare-events logit estimates with endogenous TFP. This is done in two steps. In the first-stage estimation, we regress the lag of firms' training expenses as an excluded variable on firm TFP, as well as other included variables such as indicators of SOE, foreign, exporter, and log labor. The standard errors of all the coefficients are bootstrapped with 100 replicates. The bottom module of Table 7 shows that the coefficient of log firm training expenses is positively correlated with firm TFP and strongly significant at the conventional statistical level. The F-statistic is greater than 10, which suggests that the IV is not weak in the statistical sense. After correcting for rare-events estimates bias, the coefficient of fitted firm TFP in the rare-events logit in column (2) is found to be *much* larger than the regular logit estimates, suggesting that regular binary estimates face a severe downward bias once correcting for endogeneity bias.

Columns (3) and (4) report the IV multinomial logit estimation results. Once the fitted firm TFP is obtained from the first-step IV estimates, we regress the multinomial logit estimates in which the regressand is one for distribution FDI and two for non-distribution FDI. Again, the coefficients of firm TFP for distribution FDI and non-distribution FDI are positive and significant. The magnitude of firm TFP for non-distribution FDI is even larger, which confirms our sorting equilibrium.

[Insert Table 7 Here]

4.6 Discussions of Fixed Costs Ordering

Our theoretical model is built on the assumptions on the ordering of fixed costs for non-FDI firms, distribution FDI firms, and non-distribution FDI firms. Although such assumptions are standard and used in other research, such as Helpman et al. (2004), it is still curious whether the ordering of various fixed costs can be validated by the data pattern. Table 8 picks up this task.

In general, it is challenging to check directly the validity of the fixed-costs ordering, as data on the fixed costs for non-FDI firms and (non-)distribution FDI firms, to our best knowledge, are unavailable. Still, Table 8 attempts to offer some indirect evidence to validate the ordering assumption. As suggested by Dai et al. (2016), we use firms' log advertising expenses to proxy for firms' fixed costs.[24] The idea is that FDI firms spend more on advertising fees to understand the environment in foreign markets and market penetration.

We thus construct two indicators: (1) a non-FDI indicator that equals one if a firm has no FDI and zero otherwise, and (2) a non-distribution FDI indicator that equals one if a firm has non-distribution FDI and zero otherwise. The default omitted group is distribution FDI firms.

[24] Note that the Chinese manufacturing firm-level production data set only provides firm advertising expenses during 2004-2007.

Our underlying assumption is that distribution FDI firms have higher fixed costs than non-FDI firms but lower fixed costs than non-distribution FDI firms. If this ordering is supported by the data, it should be observed that the non-FDI indicator has a negative coefficient, whereas the non-distribution FDI indicator has a positive coefficient.

These outcomes are exactly what we observe in Table 8. The estimates in column (1) start from a simple regression with two indicators as well as year-specific and two-digit industry fixed effects. Column (2) includes several firm-characteristic control variables to control for firm size (proxied by log firm labor), firm type of ownership (foreign firms or SOE), and firm export status.

Column (3) drops foreign firms from the sample and, more importantly, includes an additional export dummy to distinguish the difference between domestic advertising and foreign advertising, as our data only report firms' whole advertising expenses but do not report market-specific advertising expenses. It is also possible that a firm's advertising share in foreign countries would increase with the number of countries that it served. If so, it is possible that the firm's export intensity would increase with the number of investing destinations. We thus include an additional control variable of firm export intensity interacted with industries in column (4) and still find similar results, though the coefficient of non-distribution FDI indicator is no longer statistically significant. In any case, the anticipated signs of the non-FDI indicator and non-distribution FDI indicator strongly validate our assumption of firms' fixed-cost ordering discussed in the theoretical framework.

[Insert Table 8 Here]

5 Type-2 Tobit Estimates of Intensive Margin

Thus far, we can safely conclude that high-productivity Chinese manufacturing firms are more likely to engage in distribution FDI. We now turn to explore the role of firm productivity in FDI flow. Since we only have Zhejiang province's FDI flow data, we start by examining whether our previous findings based on nationwide FDI decision data hold for Zhejiang's FDI manufacturing firms, as discussed carefully in Appendix C. The estimates in Appendix Table 3 and their associated discussions in Appendix E clearly suggest that all our previous findings on the extensive margin of FDI hold well for the Zhejiang subsample.

To examine the *intensive* margin of firm productivity in FDI flow, we start from the simple OLS estimates in columns (1)-(2) in Table 9 by using different measures of firm productivity. We see that highly productive firms have more FDI flow regardless of the measure of firm productivity. Replacing the regressand with log FDI of distribution FDI firms yields similar results as shown in columns (3) and (4).

However, there may still be a concern that the FDI decision and FDI flow are strongly correlated. To address this question, we appeal to a bivariate sample selection model, or equivalently, a Type-2 Tobit model (Cameron and Trivedi 2005). The Type-2 Tobit specification includes: (i) an FDI participation equation where $OFDI_{it}^{D}$ denotes distribution FDI:

$$OFDI_{it}^{D} = \begin{cases} 0\, if\ U_{it} < 0 \\ 1\, if\ U_{it} \geq 0 \end{cases}, \tag{3}$$

where U_{it} denotes a latent variable faced by firm i; and (ii) an "outcome" equation whereby the firm's distribution FDI flow is modeled as a linear function of other variables. In particular, we use a logit model to estimate the following selection equation:

$$\begin{aligned} \Pr(OFDI_{it}^{D} &= 1) = \Pr(U_{it} \geq 0) = \Lambda(\gamma_0 + \gamma_1 \ln TFP_{it} + \gamma_2 SOE_{it} \\ &+\gamma_3 FIE_{it} + \gamma_4 FX_{it} + \gamma_5 \ln L_{it} + \gamma_6 Tenure_{it} + \xi_j + \lambda_t) \end{aligned} \tag{4}$$

where $\Lambda(.)$ is the logistic CDF. In addition to the logarithm of firm productivity, a firm's FDI decision is also affected by other factors, such as the firm's ownership (whether it is an SOE or a foreign firm), export status (FX equals one if a firm exports and zero otherwise), and size (measured by the logarithm of the number of employees).

Our estimations here include three steps. Because FDI firms may improve their productivity via investment abroad, in the first step, firm TFP is instrumented by the lag of log training expenses, as introduced above.[25] In the second step, our Type-2 Tobit model requires an excluded variable that affects the firm's FDI decision but does not affect its FDI flow. Here the firm's age ($Tenure_{it}$) serves this purpose, since the literature finds that a firm's tenure is highly correlated with the firm's export decision (Amiti and Davis 2011). It was shown in our previous estimates that the export decision and the FDI decision are highly correlated. By contrast, the simple correlation between FDI *flow* and firm tenure is close to nil (0.07), which confirms that tenure can serve as an excluded variable in the third-step Heckman estimates. For the third

[25] Note that standard errors in Table 9 are bootstrapped with 100 replicates.

step, we include the two-digit CIC industrial dummies ξ_j and year dummies λ_t to control for other unspecified factors.

Table 9 reports the estimation results for the bivariate sample selection model. As shown in column (5), high-productivity firms are more likely to engage in distribution FDI. We then include the computed inverse Mills ratio obtained in the third-step Heckman estimates in column (6) with the log distribution of FDI flow. The positive and significant coefficient of firm TFP suggests that high-productivity firms have more distribution FDI. Finally, columns (7)-(9) perform another robustness check of the Heckman estimates in which the regressand in the first step is the indicator of total FDI and that in the second step is log total FDI flow in column (8) and log distribution FDI flow in column (9). It turns out that our previous findings are not changed at all in such robustness checks.

[Insert Table 9 Here]

6 Cross-Country Investment Destination

Thus far, we have found evidence that high-productivity firms are more likely to invest abroad. Once a firm invests, the higher is its productivity, the more the firm invests abroad. The firm's investment decision follows the sorting behavior predicted by Proposition 1. High-productivity firms engage in distribution FDI and even higher productive firms participate in non-distribution FDI. As argued before, the importance of distribution FDI is that it can reduce the cross-border communications costs of exporting firms for service and distribution overseas. We now check whether the investment environment and income in the destination country affect the firm's distribution FDI decision.

6.1 Communication Costs in Destination Markets

Proposition 2 of our theoretical model states that an increase in cross-border communications costs (iceberg transport costs) would increase the probability of distribution (non-distribution) foreign investment. We now turn to examine whether this theoretical prediction is supported by the data.

To measure cross-border communications costs, we use data from the World Bank's Doing Business project. We first use the host country's *days* of import document preparation as a proxy for cross-border communications costs. It is important to stress that these import costs are destination-country-specific, independent of industries (or firms), but depend on the import volume. For each unit of a given product exported to a given country, such costs are roughly the same across different exporting firms, regardless of firm productivity. These features are consistent with the characteristics of cross-border communications costs sketched in our theoretical model. To make a further distinction between communications costs and transportation costs, we include the destination country's simple average import tariffs as a proxy for transportation costs. In addition, we control for log bilateral distance. These data are all publicly available from the World Bank.[26]

Table 10 is our second key table. Columns (1) and (2) present the multinomial logit estimates; the regressand in column (1) is distribution FDI and that in column (2) is non-distribution FDI. Several interesting findings merit special attention. First, the coefficients of firm relative productivity in columns (1) and (2) are all positive and significant. The magnitude of firm TFP in column (2) is higher than its counterpart in column (1). These findings are similar to our above findings and consistent with our theoretical predictions.

Second and more importantly, the coefficient of days of import document preparation in column (1) is positive and significant, whereas its counterpart in column (2) is *insignificant*, indicating that an increase in cross-border communications costs raises the probability of distribution FDI but not necessary that of non-distribution FDI, since higher cross-border communications costs attract more exporting firms to establish a foreign business office to reduce such costs, exactly as predicted by our theoretical model.

Third, our theoretical model also predicts that an increase in iceberg transportation costs would increase the probability of firms engaging in non-distribution FDI but is ambiguous on the probability of firms participating in distribution FDI, since distribution FDI does not reduce iceberg transportation costs as long as the firm exports. If this prediction is supported by the data, the iceberg transportation costs variable should exhibit a positive coefficient in column (2). We hence use the import country's simple-average tariffs as a proxy for iceberg transportation costs. The coefficient of import tariffs has a positive and significant sign in column (2) in Table

[26]Note that, in all regressions in Tables 10-12, we drop all tax-haven destinations, such as Hong Kong and Virgin Islands, from the sample, as Chinese FDI firms usually do not really invest in such regions but only use them as exprôt instead. Similarly, it is very likely that firms will switch their FDI type from distribution FDI this year to non-distribution FDI next year, as shown in Appendix Table 2.

10.

There may be curiosity about whether these results are driven by the income level of the destination country, as high-income countries usually have more transparent and efficient customs processes. And the probability of firms engaging in outward FDI would decrease as countries are further apart. We hence include per capita gross domestic product (GDP) and log bilateral distance as control variables in the multinomial logit estimates in all estimates. To control for other unspecified factors, in addition to the standard year-specific fixed effects and two-digit industry-specific fixed effects, we include the interaction between the year dummy and log bilateral distance, given that bilateral distance is time-invariant, in columns (3) and (4) in Table 10. Finally, as a robustness check, the last two columns in Table 10 drop China's multinational firms (i.e., firms with foreign indicator equal to one). The results are similar to those found earlier.

[Insert Table 10 Here]

6.2 Investment Decision by Destination Income

As seen in Table 10, destination countries' income plays an important role in FDI decisions. It is worthwhile to take a step forward to consider whether firm productivity matters for host countries' income. Interestingly, the literature offers divergent answers to this question. Head and Ries (2003) use Japanese data and find that firms investing in poor countries have even lower productivity than do non-FDI firms. However, studies like Damijan et al. (2007) find that the income level of the host country has no significant effect on Slovenian firms' FDI decision.

We consider a multinomial logit model to explore the role of firm productivity in the decision to engage in FDI in different income destinations. We first split our Zhejiang sample into two groups, low-income destination countries and high-income destination countries, by using a *predetermined* income threshold suggested by the World Bank. The base category of our multinomial logit regression is non-FDI firms, so the probability that firm i chooses to invest in country j (poor or rich) is as follows:[27]

$$\Pr(OFDI_{it}^{D}=j|\mathbf{X}_{it})=\begin{cases}\frac{1}{1+\sum_{k=2}^{3}\exp(\mathbf{X}_{it}\beta_k)} & (j\text{ is without FDI})\\ \frac{\exp(\mathbf{X}_{it}\beta_j)}{1+\sum_{k=2}^{3}\exp(\mathbf{X}_{it}\beta_k)} & (j\text{ is distribution FDI to poor }\mathit{or}\text{ to rich countries})\end{cases}, \tag{5}$$

where the regressors $\mathbf{X}_{it}$ include firm productivity and other control variables, such as export status, firm size (i.e., log firm labor) and firm ownership.

We start our regressions with a predetermined income threshold in Table 11. According to the World Bank's classifications in 2008, a country with per capita GDP less than $3,855 is classified as a lower-middle-income country, whereas a country with per capita GDP greater than $10,000 is classified as a high-income country. We hence start our multinomial logit estimates by defining FDI destination countries with income less than $3,855 as poor countries. After controlling for year-specific fixed effects and industry-specific fixed effects, the coefficient of firm productivity is positive and statistically significant for firms investing in rich countries and negative and insignificant for firms investing in poor countries. These findings hold when we increase the income threshold to $10,000, as shown in columns (3) and (4) in Table 11.

The economic rationale is straightforward. Chinese FDI firms have different motivations for their FDI behavior. Some firms seek foreign markets, whereas some seek foreign sourcing for natural resources (Huang and Wang 2011). As high-income foreign markets are usually highly competitive, only high-productivity firms are able to invest and seek local markets there. By contrast, firms investing in poor destinations are not mainly seeking foreign markets; instead, the firms may be interested in natural resources or cheaper labor in the host countries. The latter is especially true for firms in labor-intensive industries, such as textiles and garments. For instance, Chinese FDI firms that invest in Africa (e.g., Ethiopia and Madagascar) mostly are low-productivity firms in labor-intensive industries.

[Insert Table 11 Here]

However, it is important to stress that the benchmark estimations shown in Table 11 armed with predetermined cross-country income level suffer from two pitfalls. First, the threshold varies by year. There are no clear and time-invariant cutoffs for the income groups. Second, even if

[27] Note that only two firms invest in both rich countries and poor countries. For simplicity we drop those two firms from our sample.

the cutoffs are fixed, the effect of firm productivity on firm FDI may not exactly correspond to the predetermined income cutoffs. That is, host countries' per capita GDP is an endogenous threshold for FDI firms in response to productivity movement.

To overcome these empirical challenges, Hansen (1999, 2000) provides an econometric approach that considers endogenous threshold regressions. To motivate this, consider an empirical specification with a country's per capita GDP (*pcgdp*) as a threshold variable:

$$\begin{cases} OFDI_{it} = \boldsymbol{\beta}\mathbf{X}_{it} + \epsilon_{it} \; if \; pcgdp_{it} < T \\ OFDI_{it} = \boldsymbol{\theta}\mathbf{X}_{it} + \epsilon_{it} \; if \; pcgdp_{it} \geq T \end{cases}, \tag{6}$$

where T is the threshold parameter to be estimated. $OFDI_{it}$ is firm i's FDI flow in year t. $\mathbf{X}_{it}$ refers to all regressors, including firm productivity. Without loss of generality, ϵ_{it} is i.i.d with normal distribution: $\epsilon_{it} \tilde{} N(0, \sigma_i^2)$. By using an indicator function $I(\cdot)$, we can re-express Equ. (6) as:

$$OFDI_{it} = \boldsymbol{\beta}\mathbf{X}_{it} \cdot I(pcgdp_{it} < T) + \boldsymbol{\theta}\mathbf{X}_{it} \cdot I(pcgdp_{it} \geq T) + \epsilon_{it}.$$

As this is a nonlinear regression, we can use the nonlinear squares approach to minimize the sum of squared residuals. Since the estimators also include the threshold parameter $\hat{T}$, the most convenient computational method to obtain the linear squares estimate is via concentration. Thus, the optimal threshold parameter $\hat{T}$ is chosen to minimize the concentrated sum of squared errors function so that $\hat{T} = \arg\min \; SSR(\beta(T), \theta(T), T)$. Based on this, Hansen (1999, 2000) developed an asymptotic distribution theory for the threshold regression estimates.

Table 12 presents the threshold regression results. By comparison, we start from a regression without considering the threshold effect in columns (1) and (2). By abstracting away all other variables, we see that firm TFP is positively correlated to firm log FDI flow, as shown in column (1). The specification in column (2) yields similar results by controlling industry-specific fixed effects and year-specific fixed effects. The threshold regression results are reported in columns (3) and (4). The estimated threshold parameter of host countries' log per capita GDP is 10.73 (or equivalently, per capita GDP is $45,524). As before, the coefficient of firm productivity is positive and statistically significant for high-income FDI destinations. However, for low-income host countries, where per capita GDP is lower than the estimated threshold, the effect of firm productivity on firm FDI flow is statistically insignificant, suggesting that firm productivity is not a crucial determinant of firm FDI flow. This finding is robust even when we control for year-specific fixed effects and industry-specific fixed effects in columns (5) and (6), suggesting that the investment pattern described above may not apply to low-income countries. This result confirms that Chinese firms' investment in poor countries may not be labeled as "horizontal" FDI: tariff-jumping motivation or seeking foreign markets may not be a top priority for these firms (Kolstad and Wiig 2012).

[Insert Table 12 Here]

7 Concluding Remarks

This paper is one of the first to explore how firm heterogeneity influences the volume of Chinese distribution FDI. The rich data set enabled us to examine not only the extensive margin, but also the intensive margin of Chinese outward FDI.

We found that firms with distribution FDI are more productive than non-FDI firms, but less productive than non-distribution FDI firms. These findings reflect the fact that many Chinese exporters are insufficiently productive to build up production lines in foreign markets. As a compromise, such firms set up foreign distribution centers to promote their exports. We also found that with higher cross-border communications costs (iceberg transport costs), there is a higher probability that firms engage in distribution (non-distribution) FDI. Finally, high-productivity firms invest more in high-income countries but not necessarily in low-income countries.

References

[1] Acemoglu, Daron and Jörn-Steffen Pischke (1998), "Why do Firms Train? Theory and Evidence," *Quarterly Journal of Economics*, pp.79-119.

[2] Ahn, JaeBin, Amit Khandelwal, and Shang-jin Wei (2010), "The Role of Intermediaries in Facilitating Trade," *Journal of International Economics* 84(1), pp. 73-85.

[3] Amiti, Mary and Donald Davis (2011), "Trade, Firms, and Wages: Theory and Evidence," *Review of Economic Studies* 79, pp. 1-36.

[4] Amiti, Mary, and Jozef Konings (2007), "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *American Economic Review* 97(5), pp. 1611-1638.

[5] Antràs, Pol and Elhanan, Helpman (2004), "Global Sourcing," *Journal of Political Economy* 112, pp. 552-580.

[6] Arkolakis, Costas and Marc-Andreas Muendler (2011), "The Extensive Margin of Exporting Products: A Firm-level Analysis," mimeo, Yale University.

[7] Berman, Nicolas, Philippe Martin and Thierry Mayer, (2012), "How Do Different Exporters React to Exchange Rate Changes?" *Quarterly Journal of Economics*, 127, pp. 437-492.

[8] Bhattacharya Rudrani, Ila Patnaik, and Ajay Shah (2010), "Productivity and Outbound FDI in Software Services: a Reversal of the HMY Model," working paper, National Institute of Public Finance and Policy, New Delhi.

[9] Blonigen, Bruce (2001), "In Search of Substitution between Foreign Production and Exports," *Journal of International Economics* 53(1), pp. 81-104.

[10] Blonigen, Bruce (2005), "A Review of the Empirical Literature on FDI Determinants," *Atlantic Economic Journal* 33, pp. 383–403.

[11] Boatman, Kara (2007), "What Determines Distribution FDI?" *Journal of Business and Public Affairs*, 1(1).

[12] Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang (2012), "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing?" *Journal of Development Economics*, 97, pp. 339-351.

[13] Cameron, Colin and Pravin Trivedi (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press.

[14] Chen, Qiang (2014), "Climate Shocks, State Capacity, and Peasant Uprisings in North China during 25-1911 CE.," *Economica*, forthcoming.

[15] Chen, Cheng, Wei Tian, and Miaojie Yu (2018), "Outward FDI and Domestic Input Distortions: Evidence from Chinese Firms," *Economic Journal*, forthcoming

[16] Chen, Wenjie and Heiwai Tang (2014), "The Dragon is Flying West: Micro-Level Evidence of Chinese Direct Investment," *Asian Development Review* 31(2), pp. 109-140.

[17] Conconi, Paola, André Sapir, and Maurizio Zanardi (2014), "The Internationalization Process of Firms: from Exports to FDI", mimeo, Lancaster University Management School.

[18] Dai, Mi, Madhura Maitra, and Miaojie Yu (2016), "Unexceptional Exporter Performance in China? The Role of Processing Trade," *Journal of Development Economics* 121, pp.177-189.

[19] Damijan, Joze P, Saso Polanec and Janez Prasnikar (2007), "FDI and Productivity: Micro-evidence from Slovenia," *The World Economy*, 30 (1), pp. 135–155.

[20] De Loecker, Jan (2011), "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79(5), pp. 1407–1451.

[21] Eaton Jonathan, Samuel Kortum, and Francis Kramarz (2011), "An Anatomy of International Trade: Evidence from French Firms," *Econometrica* 79(5), pp. 1453-1498.

[22] Fajgelbaum Pablo, Gene Grossman, and Elhanan Helpman (forthcoming), "A Linder Hypothesis for Foreign Direct Investment," *Review of Economic Studies*.

[23] Federico, Stefano (2009), "Outsourcing versus Integration at Home and Abroad," working paper, Bank of Italy.

[24] Feenstra, Robert and Gordon Hanson (2005),"Ownership and Control in Outsourcing to China: Estimating the Property-Rights Theory of the Firm," *Quarterly Journal of Economics* 120(2), pp. 729-762.

[25] Feenstra, Robert, Zhiyuan Li, and Miaojie Yu (2014), "Exports and Credit Constraints under Incomplete Information: Theory and Evidence from China," *Review of Economics and Statistics*, 96(4), pp.729-744.

[26] Froot, Kenneth and Jeremy Stein (1991), "Exchange Rates and Foreign Direct Investment: An Imperfect Capital Markets Approach", *Quarterly Journal of Economics*, pp.1191-1217.

[27] Groizard, Luis, Priya Ranjan, Antonio Rodriguez-Lopez (2015), "Trade Costs and Job Flows: Evidence from Establishment-Level Data," *Economic Inquiry* 53(1), pp. 173-204.

[28] Goldberg, Pinelopi, Amit Khandelwal, Nina Pavcnik, and Petia Topalova (2010), "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *Quarterly Journal of Economics* 125 (4), pp. 1727-1767.

[29] Hansen, Bruce (1999), "Threshold effects in non-dynamic panels: Estimation, testing and inference," *Journal of Econometrics* 93, 345-368.

[30] Hansen, Bruce (2000), "Sample Splitting and Threshold Estimation," *Econometrica* 68(3), pp. 575-603.

[31] Hanson, Gordon H., Mataloni, Raymond J., Slaughter, Mathhew J. (2001), "Expansion Strategies of U.S. Multinational Firms," National Bureau of Economic Research (Cambridge, MA) Working Paper No. 8433.

[32] Head, Keith and John Ries (2003), "Heterogeneity and the FDI versus Export Decision of Japanese Manufacturers," *Journal of the Japanese International Economies* 17(4), pp. 448–67.

[33] Halpern, Laszlo, Miklos Koren, and Adam Szeidl (2010), "Imported Inputs and Productivity," University of California, Berkeley, Mimeo.

[34] Helpman, Elhanan, Marc J. Melitz and Steven R. Yeaple (2004), "Export vs. FDI," *American Economic Review* 94(1), pp. 300–16.

[35] Horstmann, Ignatius J. and James R. Markusen (1996), "Exploring New Markets: Direct Investment, Contractual Relations and the Multinational Enterprise," *International Economic Review*, 37(1), pp. 1-19.

[36] Hsieh, Chang-Tai and Peter Klenow (2009), "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics* 124(4), pp. 1403-48.

[37] Huang, Yiping and Bijun Wang (2011), "Chinese Direct Investment: Is There a China Model?" *China & World Economy* 19(4), pp. 1-21.

[38] Imbens, Guido (2004), "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics* 86(1), pp. 4-29.

[39] Ishikawa, Jota, Hodaka Morita, and Hiroshi Mukunnoki (2010), "FDI in Post-Production Services and Product Market Competition," *Journal of International Economics* 82(1), pp. 73-84.

[40] Ishikawa, Jota and Hodaka Morita (2015), "Trade Liberalization and Aftermarket Services for Imports," mimeo, Hitotsubashi University.

[41] Jovanovic, Boyan J.(1982), "Selection and the Evolution of Industry," Econometrica 50, pp. 649-670.

[42] Kimura, Fuku and Hyun-Hoon Lee (2006), "The Gravity Equation in International Trade in Services," *Review of International Economics*, 142(1), pp 92-121.

[43] King, Gary and Langche Zeng (2001), "Logistic Regression in Rare Events Data," *Political Analysis* 9(2), pp. 137-163.

[44] King, Gary and Langche Zeng (2002), "Estimating Absolute, Relative, and Attributable Risks in Case-Control Studies," *Statistics in Medicine* 21, pp. 1409-1427.

[45] Kolstad Ivar and Arne Wilg (2012), "What determines Chinese FDI?" *Journal of World Business*, 47(1), pp. 26-34.

[46] Levinsohn, James and Amil Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservable," *Review of Economic Studies* 70(2), pp. 317-341.

[47] Linder, Staffen Burenstam (1961), *An Essay on Trade and Transformation*, Stockholm: Almqvist and Wiksell.

[48] Lu Dan (2010), "Exceptional Exporter Performance? Evidence from Chinese Manufacturing Firms," mimeo, University of Chicago.

[49] Markusen, James R. (1984), "Multinationals, Multi-Plant Economies, and the Gains from Trade," *Journal of International Economics* 16 (3-4), pp. 205-26.

[50] Melitz, Marc (2003), "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71(6), pp. 1695–725.

[51] Nunn, Nathan (2007), "Relationship-specificity, Incomplete Contracts, and the Pattern of Trade," *Quarterly Journal of Economics*, May 2007, pp. 569-600.

[52] Oldenski, Lindsay (2012), "Export Versus FDI and the Communication of Complex Information,"*Journal of International Economics* 87, 312-322.

[53] Olley, Steven and Ariel Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica* 64(6), pp. 1263-1297.

[54] Qiu, Larry, D. and Ying Xue (2014), "Understanding China's Foreign Trade: a Literature Review (I)," *China Economic Journal*, 7(2), pp. 168-186.

[55] Ramondo, Natalia, Veronica Rapporport, Kim J. Ruhl (2014), "The Proximity-Concentration Tradeoff under Uncertainty," *Review of Economic Studies*, forthcoming.

[56] Russ, Katheryn N.(2007), "The Endogeneity of the Exchange Rate as a Determinant of FDI: A Model of Money, Entry, and Multinational Firms," *Journal of International Economics* 71(2), pp. 344-372.

[57] Wang, Bijun, Yuyan Tan, Miaojie Yu, and Yiping Huang (2015), " Direct Investment, Firm Productivity, and Credit Constraints: Evidence from Chinese Firms," *Pacific Economic Review*, forthcoming.

[58] Wang, Chengqi, Junjie Hong, Mario Kafouros, Agyenim Boateng (2012), "What Drives FDI of Chinese Firms? Testing the Explanatory Power of Three Theoretical Frameworks," *International Business Review*, 21(3). pp. 425-438.

[59] Yeaple, Stephen (2005), "A Simple Model of Firm Heterogeneity, International Trade, and Wages," *Journal of International Economics* 65(1), pp.1-20.

[60] Yu, Miaojie (2015), "Processing Trade, Tariff Reductions, and Firm Productivity: Evidence from Chinese Firms," *Economic Journal*, 125 (June), pp. 943–988.

Table 1: FDI Share in Number of Manufacturing Firms (2000-08)

Firm type	2000	2001	2002	2003	2004	2005	2006	2007	2008
Nationwide FDI decision data									
(1) Mfg. firms	84,974	100,091	110,522	129,720	200,989	198,285	248,601	258,246	222,312
(2) FDI starting firms	197	340	444	587	972	984	1,081	1,140	1,018
(3) FDI mfg. firms	14	17	20	30	103	431	761	1,168	1,183
(4) Distribution FDI mfg. firms	2	2	3	4	40	224	407	616	656
(5) FDI share (%)	0.23	0.34	0.40	0.45	0.48	0.49	0.43	0.44	0.46
(6) Distribution FDI share (%)	14.2	11.7	15.0	13.3	38.8	51.9	53.5	52.7	55.4
Zhejiang's FDI flow data									
(7) Mfg. firms							35,887	39,465	27,433
(8) FDI firms							424	419	427
(9) FDI mfg. Firms							113	163	131
(10) FDI share (%)							0.31	0.41	0.48

Note: Data are from the Ministry of Commerce of China and the authors' calculations. FDI share (%=(3)/(1)) is obtained by dividing the number of FDI starting firms by the number of manufacturing firms nationwide and in Zhejiang province, respectively. Distribution FDI share (=(4)/(3)) is obtained by dividing the number of distribution FDI manufacturing firms by the number of total FDI manufacturing firms. For Zhejiang firms, the FDI decision data are available every year during the sample, but the FDI flow data are available only for 2006-08, for which there are 1,270 FDI firms in Zhejiang and 407 of them are manufacturing firms.

Table 2: Summary Statistics of Key Variables

Sample Covered:	Nationwide (2000-08)		Zhejiang (2006-08)	
Variable	Mean	Std. dev.	Mean	Std. dev.
Firm log FDI			3.27	1.53
Firm FDI indicator	0.004	0.066	0.003	0.05
Firm TFP (Olley-Pakes)	3.61	1.18	4.08	0.94
Firm export indicator	0.29	0.451	0.42	0.49
SOE indicator	0.05	0.219	0.002	0.047
Foreign indicator	0.20	0.402	0.16	0.366
Firm log labor	4.78	1.115	4.45	0.983

Note: Number of observations nationwide is 1,140,834 whereas that of Zhejiang province is 100,847.

Table 3: Difference between Non- FDI and FDI Firms

Category	# of obs.	Mean				
Variable		TFP		Log Employees	Sales	Export
Propensity Score Matching		Unmatched	Matched			
	(1)	(2)	(3)	(4)	(5)	(6)
FDI Firms vs. Non- FDI Firms						
(i) Non-FDI firms	1,137,097	3.604	3.811	4.742	96,560	18,615
(ii) FDI firms	3,727	4.289	4.289	5.844	3,334,672	1,285,675
Difference=(ii)-(i)		0.685***	0.478***	1.102***	3,238,112***	1,267,060***
		(35.41)	(9.54)	(61.34)	(150.3)	(120.1)
(iii) Distribution FDI firms	1,954	4.180	4.180	5.604	1,245,375	460,793
Difference=(iii)-(i)		0.576***	0.442***	0.862***	1,148,815***	442,178***
		(21.55)	(8.44)	(34.83)	(50.37)	(39.22)
(iv) Non-Distribution FDI firms	1,773	4.411	4.411	6.109	5,637,258	2,194,766
Difference=(iii)-(iv)		-0.231***	-0.086	-0.505***	-4,391,884***	-1,733,972***
		(-5.73)	(-1.39)	(-8.45)	(-9.19)	(-7.63)

Note: Numbers in parentheses are t-values. *** denotes significance at the 1% level. Columns (2) and (3) report the TFP comparison by using the nearest-matching propensity score matching (PSM) approach. Column (2) is unmatched whereas column (3) is the average treatment on the treated (ATT) approach. The treated group is FDI firms, whereas the control group is non-FDI firms. Firm size (in log labor), exports, and sales are used as covariates to obtain the propensity score. Since there are observations with identical propensity score values, the sort order of the data could affect the results. The sort order is made to be random before adopting the PSM approach.

Table 4: TFP Difference between Distribution FDI Firms and other Firms by Year

Firm TFP	2000	2001	2002	2003	2004	2005	2006	2007	2008
(1) Non-FDI Firms	3.109	3.001	3.218	3.283	3.065	3.421	3.540	3.659	4.966
(2) All FDI Firms	4.396	4.190	4.376	5.309	4.163	3.855	3.738	3.738	5.194
(3) Distribution FDI Firms	5.514	5.564	5.888	4.514	3.855	3.634	3.578	3.744	5.150
(4) Non-Distribution FDI Firms	4.210	4.007	4.109	5.432	4.358	4.093	3.923	4.026	5.248
TFP Difference=(3)-(1)	2.404***	2.562***	2.670***	1.231***	0.790***	0.213***	0.038	0.085**	0.184***
	(3.23)	(3.49)	(4.50)	(2.43)	(4.56)	(3.09)	(0.73)	(2.01)	(5.14)
TFP Difference=(3)-(4)	1.304***	1.557***	1.778***	-0.917	-0.502***	-0.458***	-0.344***	-0.282***	-0.098*
	(2.03)	(2.75)	(3.55)	(-1.44)	(-1.90)	(-4.34)	(-4.45)	(-4.35)	(-1.79)

Note: Numbers in parentheses are t-values. ***(**, *) denotes significance at the 1% (5%, 10%) level. The table reports firm TFP difference between distribution FDI and non-distribution FDI firms and firm TFP difference between distribution FDI and non-FDI firms.

Table 5: Effects of Firm Productivity on FDI Decision (2000-08)

Regressand: Indicator of	FDI Starter							Ever FDI
Econometric Method:	LPM	Probit	Logit	Rare Event Logit	Complementary Log-Log	Logit	Logit	Logit
Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Firm Relative TFP	0.001***	0.321***	1.076***	1.086***	1.084***	1.079***	0.988***	0.893***
	(4.54)	(5.11)	(5.62)	(5.68)	(5.61)	(4.66)	(5.11)	(7.69)
SOE Indicator	-0.001***	-0.310***	-0.874***	-0.876***	-0.878***	-0.963***	–	-0.613***
	(-3.91)	(-5.17)	(-4.95)	(-4.96)	(-4.85)	(-5.43)		(-5.96)
Foreign Indicator	-0.001***	-0.183***	-0.577***	-0.575***	-0.573***	–	-0.578***	-0.704***
	(-5.26)	(-7.77)	(-8.18)	(-8.16)	(-8.76)		(-8.18)	(-15.55)
Log Firm Labor	0.001***	0.208***	0.615***	0.614***	0.610***	0.642***	0.605***	0.577***
	(15.59)	(26.18)	(28.33)	(28.30)	(29.31)	(25.56)	(27.13)	(42.90)
Export Indicator	0.002***	0.431***	1.373***	1.373***	1.372***	1.460***	1.370***	1.096***
	(15.27)	(19.91)	(19.64)	(19.63)	(20.46)	(19.19)	(19.40)	(27.60)
SOE Dropped	No	No	No	No	No	No	Yes	No
Foreign Firms Dropped	No	No	No	No	No	Yes	No	No
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1,138,450	1,137,309	1,137,309	1,138,450	1,138,450	896,958	1,097,537	1,139,683

Note: The regressand is the FDI indicator. Numbers in parentheses are t-values clustered at the firm level (if applicable). *** denotes significance at the 1% level. Column (6) drops foreign firms from the sample whereas column (7) drops SOEs from the sample.

Table 6: Multinomial Logit Estimates of Firm Productivity on Distribution FDI Decision (2000-08)

Regressand: Indicator of	Distri.	Non-Distri.	Distri.	Non-Distri.	Distri.	Non-Distri.	Distri.	Non-Distri.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Firm Relative TFP	0.616***	1.625***	0.779***	1.726***	0.677***	1.798***	1.178***	2.012***
	(3.05)	(6.75)	(3.64)	(6.79)	(2.67)	(6.21)	(8.56)	(12.71)
SOE indicator	-1.588***	-0.653***	-1.594***	-0.475**	-1.631***	-0.756***	-1.324***	-0.430***
	(-4.61)	(-3.13)	(-4.10)	(-2.23)	(-4.70)	(-3.62)	(-6.81)	(-3.54)
Foreign indicator	-0.601***	-0.497***	-0.504***	-0.392***			–	–
	(-6.74)	(-4.76)	(-5.42)	(-3.71)				
Log firm labor	0.533***	0.698***	0.549***	0.695***	0.539***	0.728***	0.456***	0.666***
	(18.44)	(23.44)	(17.63)	(22.09)	(16.38)	(21.73)	(21.67)	(32.10)
Export indicator	1.824***	0.967***	1.792***	0.982***	1.961***	1.076***	1.570***	0.745***
	(18.42)	(10.27)	(17.00)	(9.86)	(17.88)	(10.35)	(26.94)	(12.13)
Foreign firms dropped	No	No	No	No	Yes	Yes	Yes	Yes
Pure exporters dropped	No	No	No	No	Yes	Yes	No	No
M&A FDI firms included	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Most contractable sectors dropped	No	No	Yes	Yes	No	No	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1,138,450		1,021,633		898,068		897,034	

Note: Numbers in parentheses are t-values. ***(**, *) denotes significance at the 1(5, 10)% level. The omitted group is non-FDI firms. Columns (1) and (2) are multinomial logit estimates in which the regressand in column (1) is distribution FDI whereas that in column (2) is non-distribution FDI. The regressands in columns (3) and (4) are multinomial logit estimates in which 20 three-digit CIC industries with high contract intensity a là Nunn (2007) are dropped. The regressands in columns (5) and (6) are multinomial logit estimates in which both foreign firms and pure exporters (i.e., firm exports equal firm sales) are dropped. Estimates in columns (7) and (8) drop both foreign firms and non-FDI firms that have ever merged or acquired foreign firms. The regressand in column (9) is the distribution FDI indicator (i.e., zero refers to non-distribution FDI and one refers to distribution FDI).

Table 7: IV Estimates of Firm Productivity on FDI Decision (2005-08)

Econometric method:	Probit	Rare events Logit	Multinomial Logit	
Regressand: FDI Indicator for	All FDI	All FDI	Distribution	Non-Distribution
	(1)	(2)	(3)	(4)
Firm TFP	0.854***	4.575***	1.376***	1.487***
	(31.24)	(7.40)	(3.54)	(3.62)
SOE Indicator	0.149***	0.799***	-0.969***	0.160
	(3.43)	(2.80)	(-2.41)	(0.62)
Foreign Indicator	-0.091***	-0.524***	-0.720***	-0.467***
	(-5.33)	(-6.34)	(-6.73)	(-3.83)
Log Firm Labor	0.094***	0.532***	0.510***	0.685***
	(8.31)	(18.40)	(14.89)	(17.52)
Export Indicator	0.260***	1.562***	1.858***	1.021***
	(10.89)	(18.92)	(15.84)	(9.07)
Year-specific Fixed Effects	Yes	Yes	Yes	Yes
Industry-specific Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	484,212	484,212	484,212	484,212
	First-Stage Regression			
IV: Log Firm Training Expenses			0.028***	
			(30.40)	

Notes: The regressands in columns (1), (2), and (5) are the FDI indicator whereas those in columns (3) and (6) are the distribution FDI indicator and those in columns (4) and (7) are the non-distribution FDI indicator. Numbers in parenthesis are bootstrapped t-values. ***(**) denotes significance at the 1(5)% level. The estimates in column (1) are probit IV estimates, whereas those in columns (2) and (5) are IV rare-events logit estimates. Columns (3)-(4) are IV multinomial logit estimates. The omitted group is non-FDI firms. The instrument is the one-year lag of firm's log training expenses from the period 2005 to 2008.

Table 8: Advertising Expenses by Different Types of Firms (2004-07)

Regressand: Log Firm Advertising Expenses	(1)	(2)	(3)	(4)
Non FDI Indicator	-1.32***	-0.56***	-0.62***	-0.77***
	(-8.33)	(-5.53)	(-4.27)	(-5.23)
Non-Distribution FDI Indicator	0.57**	0.32**	0.32	0.22
	(2.33)	(2.14)	(1.38)	(0.97)
SOE Indicator		-0.54***	-0.53***	-0.56***
		(-19.13)	(-13.69)	(-14.52)
Foreign Indicator		0.22***	–	–
		(16.32)		
Log Firm Labor		0.70***	0.69***	0.72***
		(138.07)	(88.20)	(92.69)
Export Indicator		-0.02	0.08***	
		(-1.58)	(4.75)	
Export Share $\times$ 2-digit Industry Fixed Effects	No	No	No	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
2-digit Industry Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	128,028	128,026	99,679	98,424
R-squared	0.06	0.22	0.23	0.23

Notes: The regressand is firm log advertising expenses. Numbers in parentheses are t-values clustered at the firm level. ***(**,*) denotes significance at the 1(5, 10)% level. The omitted group is distribution FDI firms. Columns (3) and (4) drop foreign firms from the sample. Column (4) includes the interaction terms between export intensity (i.e., firm exports over firm total sales) with two-digit industry fixed effects.

Table 9: Intensive Estimates of Distribution FDI vs Total FDI Flow for Zhejiang Firms (2006-08)

Econometric Method:	OLS				Heckman				
					1^{st}-step	2^{nd}-step	1^{st}-step	2^{nd}-step	
Regressand:	Log FDI of All FDI Firms		Log FDI of Dist. FDI Firms		Dist. FDI Dummy	Log Dist. FDI	FDI Dummy	Log FDI	Log Dist. FDI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Firm Relative TFP	0.922*		0.866*						
	(1.92)		(1.88)						
Firm TFP		0.204*		0.193*	0.720***	4.968***	0.703***	3.429*	4.751***
		(1.90)		(1.88)	(4.01)	(2.66)	(4.12)	(1.92)	(2.86)
SOE Indicator	2.492***	2.488***				2.689***	-0.329	1.141	
	(7.09)	(7.08)				(7.77)	(-0.85)	(1.40)	
Foreign Indicator	0.081	0.080	0.097	0.097	-0.185***	-1.158**	-0.150***	-0.643	-0.880**
	(0.39)	(0.39)	(0.46)	(0.46)	(-3.06)	(-2.15)	(-2.69)	(-1.40)	(-2.00)
Log Firm Labor	0.314***	0.315***	0.339***	0.340***	0.133**	1.162***	0.149**	0.986***	1.294***
	(4.03)	(4.04)	(4.53)	(4.54)	(2.03)	(3.44)	(2.35)	(2.61)	(3.58)
Export Indicator	-0.603**	-0.600**	-0.647**	-0.647**	0.527***	2.987**	0.502***	1.754	2.645**
	(-2.04)	(-2.03)	(-2.05)	(-2.05)	(5.45)	(2.10)	(5.71)	(1.33)	(2.15)
Firm Tenure					0.002		0.003		
					(0.39)		(1.00)		
Inverser mills ratio						7.506**		5.207*	7.372***
						(2.58)		(1.82)	(2.79)
Year Fixed Effects	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Number of Observations	251	251	210	210	60198	251	60,358	251	210
R-squared	0.13	0.13	0.15	0.15	–	0.16	–	0.15	0.18
	.First-stage Estimates								
IV: Log Firm Training	–	–	–	–	0.049***	–	0.011***	–	–
					(21.86)		(22.29)		

Note: Numbers in parentheses are (bootstrapped) t-values. ***(**,*) denotes significance at the 1% (5%, 10%) level. The regressand in the OLS estimates in columns (1)-(2) are total FDI flow whereas those in columns (3)-(4) are distribution FDI flow. The rest of the table is the three-step Heckman IV estimates: (i) The first stage regresses firm TFP on the one-year lag of firm's log training expenses. (ii) The second stage is the probit estimates, which regress the firm distribution FDI indicator on the fitted variables of TFP obtained from (i). Firm tenure serves as the exclusive variable, which is only included in the second-stage but not in the third-stage. (iii) The third-stage estimates in column (6) regress log distribution FDI on fitted TFP and other control variables. The inverse Mills ratio calculated from the second-stage is inserted in the third stage as an additional regressor. The estimates in columns (7)-(9) are similar to the estimates in columns (5) and (6) except the regressand of the second stage is firm the FDI indicator and that of the third stage is firm's log FDI flow in column (8) and firm's log distribution FDI flow in column (9).

Table 10: Multinomial Logit Estimates by Host Investment Environment (2006-2008)

Regressand: Indicator of	Dist.	Non-Dist.	Dist.	Non-Dist.	Dist.	Non-Dist.
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Relative TFP	3.41***	4.84**	3.41***	4.84**	4.24***	4.38*
	(3.17)	(2.08)	(3.17)	(2.08)	(3.02)	(1.94)
Days of Import Document Preparation	0.26*	-0.19	0.26*	0.03	0.30*	0.05
	(1.76)	(-0.83)	(1.70)	(0.12)	(1.65)	(0.20)
Import Tariffs	0.08	0.13**	0.07	0.17**	0.04	0.17*
	(1.51)	(2.04)	(1.43)	(2.18)	(0.66)	(1.95)
SOE Indicator	-10.85***	3.22***	-14.92***	2.82**	-10.86***	3.21***
	(-72.12)	(3.00)	(-75.15)	(2.44)	(-68.68)	(2.94)
Foreign Indicator	0.60**	0.43	0.52*	0.39	–	–
	(2.24)	(0.74)	(1.94)	(0.66)		
Log per-capita GDP	0.97***	0.33	0.97***	0.65	0.81**	0.60
	(3.36)	(0.92)	(3.37)	(1.56)	(2.48)	(1.19)
Foreign Firms Dropped	No	No	No	No	Yes	Yes
Log Bilateral Distance controlled	Yes	Yes	Yes	Yes	Yes	Yes
Year-specific Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy×Log Bilateral Distance	No	No	Yes	Yes	Yes	Yes
Industry-specific Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1,999,441		1,999,441		1,679,207	

Note: Numbers in parentheses are (bootstrapped) t-values. ***(**,*) denotes significance at the 1% (5%, 10%) level. The regressands in columns (1), (3), and (5) are the indicator of distribution FDI, whereas those in columns (2), (4), and (6) are the indicator of non-distribution FDI. The omitted group is non-FDI firms.

Table 11: Multinomial Logit Estimates of the Distribution FDI Decision by Destination Income (2006-08)

Predetermined income threshold	GDPPC=$3,885		GDPPC=$10,000	
Regressand: Distribution FDI decision	Dist. FDI to poor	Dist. FDI to rich	Dist. FDI to poor	Dist. FDI to rich
	(1)	(2)	(3)	(4)
Firm Relative TFP	0.058	1.371***	0.377	1.420***
	(0.04)	(3.06)	(0.36)	(3.03)
Log of labor	0.500***	0.656***	0.668***	0.637***
	(3.05)	(10.44)	(5.49)	(9.46)
Export indicator	1.569***	2.302***	1.974***	2.254***
	(3.05)	(8.47)	(4.28)	(8.00)
SOE indicator	-15.33***	-15.67***	-17.70***	-17.44***
	(-25.21)	(-55.99)	(-35.42)	(-61.60)
Foreign indicator	-0.773	-0.177	-1.016**	-0.086
	(-1.25)	(-1.04)	(-2.12)	(-0.49)
Year-specific fixed effects	Yes	Yes	Yes	Yes
Industry-specific fixed effects	Yes	Yes	Yes	Yes
Number of observations	100,806		100,806	

Note: The regressand is the distribution FDI indicator in which zero denotes not any type of FDI, one refers to distribution FDI to poor countries, and two denotes distribution FDI to rich countries. Numbers in parentheses are t-values. ***(**) denotes significance at the 1% (5%) level. The sample in this table covers Zhejiang manufacturing firms during 2006-08. Two-digit Chinese industry classification industry-specific fixed effects are included in all the estimations.

Table 12: Threshold Estimates by Income of Host Countries (2006-08)

Estimated threshold:	Without		With threshold			
Log GDPPC=10.726	threshold		Low	High	Low	High
Regressand: Firm log FDI flow	(1)	(2)	(3)	(4)	(5)	(6)
Firm TFP	0.209***	0.254***	0.067	0.395***	0.107	0.460***
	(2.29)	(2.61)	(0.67)	(2.43)	(1.00)	(2.68)
Constant	2.500***	2.303***	2.945***	1.977***	2.767***	1.588***
	(6.37)	(5.24)	(6.69)	(2.78)	(5.50)	(2.04)
Year fixed effects	No	Yes	No	No	Yes	Yes
Industry fixed effects	No	Yes	No	No	Yes	Yes
Number of observations	251	251	165	86	165	86
(Joint) R-squared	0.023	0.038	0.061		0.082	

Note: The regressand is firm log FDI flow. Numbers in parentheses are t-values clustered at the firm level. *** denotes significance at the 1% level. Estimates in this table are threshold estimates a la Hanson (2000) by using FDI destination income as the threshold. The estimates in columns (1) and (2) are standard OLS estimates without considering the heteroskedasticity of the threshold. Columns (4) to (6) are estimated by using the estimated threshold (log per capita GDP is 10.726). Joint R-squareds are reported in columns (4) to (6). Columns (5) and (6) include CIC two-digit industry-level fixed effects and year-specific fixed effects. The 95% confidence interval estimates for each variable are not reported to save space, although they are available upon request.

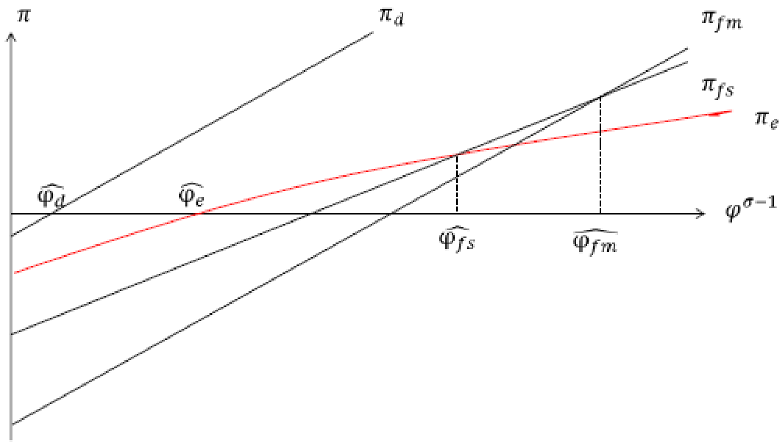


Figure 1: Firm Productivity, Export, Distribution and Non-distribution FDI

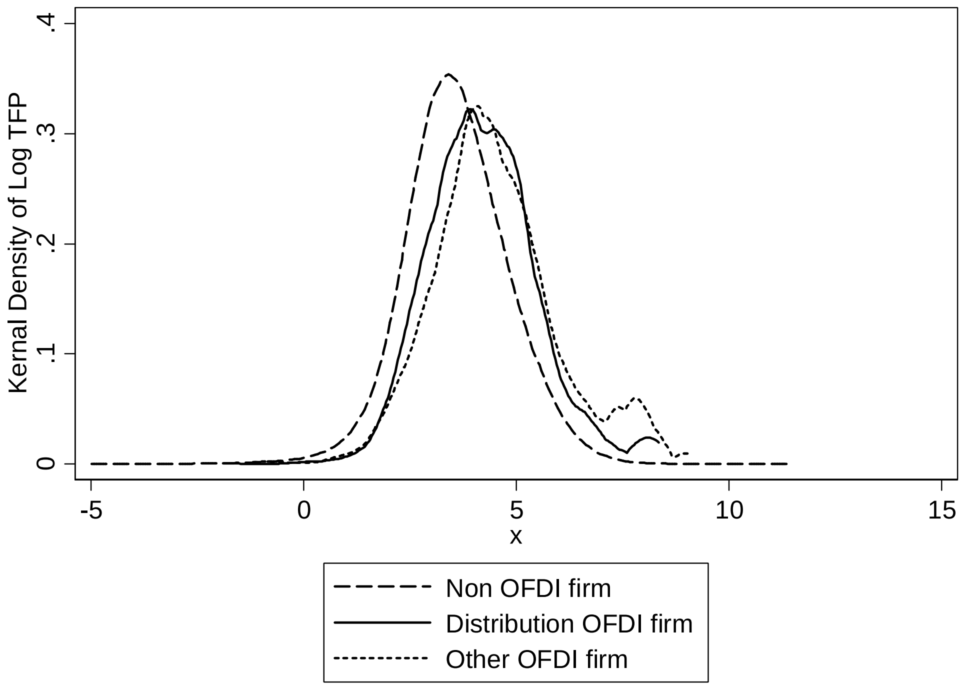
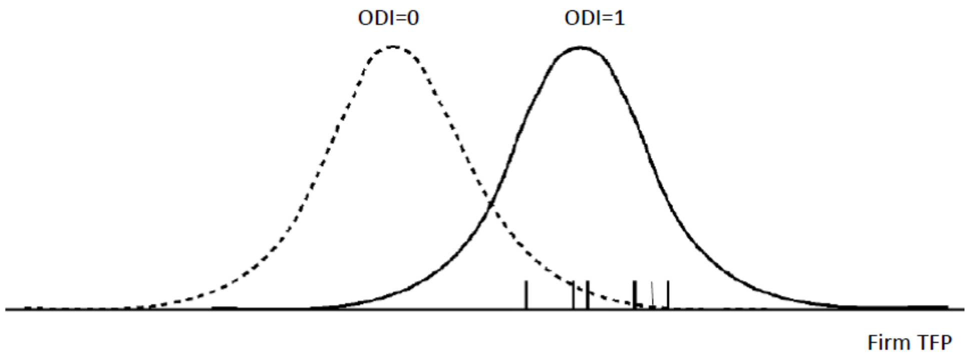


Figure 2: Firm Productivity by Firm Type



Note: Samples are sorted by firm TFP. The short vertical line represents rare observations with FDI=1 whereas the many observations with FDI=0 are not drawn. The solid (dotted) curve refers to the probability density with FDI=1(FDI=0). The cutoff points that best classify FDI=0 and FDI=1 would be too far to the right as argued in the text.

Figure 3: Rare Events of FDI Firms

8 Appendix (for online publication)

8.1 Appendix A: Proof of Proposition 1

Proof: The derived demand for product φ is

$$X_j(\varphi) = L_j P_j^{\sigma-1}[p_j^c(\varphi)]^{-\sigma}$$

where L_j is labor income in country j, $p_j^c(\varphi) = \frac{\sigma}{\sigma-1} MC^c, c = d, e, fs, fm$ is the price of product φ if it is domestically sold, exported without distribution foreign affiliate, exported with a distribution affiliate or with production affiliates respectively. P_j is the aggregate price level, where

$$P_j = \left\{ \begin{array}{c} \sum_{h=1,h\neq j}^{N} \frac{\sigma}{\sigma-1} L_h \left[\int_{\widehat{\varphi_{ehj}}}^{\widehat{\varphi_{fshj}}} (\frac{\tau_{hj}}{\varphi} + \eta_j)^{1-\sigma} dG(\varphi) + \int_{\widehat{\varphi_{fshj}}}^{\widehat{\varphi_{fmhj}}} (\frac{\mu_j \tau_{hj}}{\varphi})^{1-\sigma} dG(\varphi) + \int_{\widehat{\varphi_{fmhj}}}^{\infty} (\frac{1}{\varphi})^{1-\sigma} dG(\varphi) \right] \\ + \frac{\sigma}{\sigma-1} L_j \int_{\widehat{\varphi_{dj}}}^{\infty} (\frac{1}{\varphi})^{1-\sigma} dG(\varphi) \end{array} \right\}^{\frac{1}{1-\sigma}}$$

where $\widehat{\varphi_{dj}}, \widehat{\varphi_{ehj}}, \widehat{\varphi_{fshj}}, \widehat{\varphi_{fmhj}}$ are the productivity cut-off point for each mode.

From Equ. (1) and (2), we know that when $\frac{f_X}{f_D} > (\tau + \widehat{\varphi_e}\eta)^{1-\sigma}, \widehat{\varphi_d} < \widehat{\varphi_e}$.So when $\frac{f_X}{f_D} > \tau^{1-\sigma}, \widehat{\varphi_d} < \widehat{\varphi_e}$.

Deriving the LHS of (3) with respect to $\frac{1}{\varphi}$,we get $d\left[(\frac{\mu\tau_{ij}}{\varphi})^{1-\sigma} - (\frac{\tau_{ij}}{\varphi} + \eta_j)^{1-\sigma}\right] / d\left(\frac{1}{\varphi}\right) = (1-\sigma)\tau \left[\left(\mu^{\frac{\sigma-1}{\sigma}} \tau \frac{1}{\varphi}\right)^{-\sigma} - \left(\tau\frac{1}{\varphi} + \eta\right)^{-\sigma} \right] < 0$, which means a higher φ induces higher relative returns of distribution investment compared with export without FDI. Thus (3) has a single solution. Similarly, we derive the LHS of (4) with respect to $\frac{1}{\varphi}$ to get $d\left[(\frac{1}{\varphi})^{1-\sigma} - (\frac{\mu\tau_{ij}}{\varphi})^{1-\sigma}\right] / d\left(\frac{1}{\varphi}\right) = (1-\sigma)\left[1 - (\mu\tau)^{1-\sigma}\right]\left(\frac{1}{\varphi}\right)^{-\sigma} < 0$.So the higher is φ, the more profitable is building a production plant relative to a distribution affiliate.

Because the LHS of (3) is increasing with φ, so $\widehat{\varphi_e} < \widehat{\varphi_{fs}}$ equals

$$(\frac{\mu\tau}{\widehat{\varphi_e}})^{1-\sigma} - (\frac{\tau}{\widehat{\varphi_e}} + \eta)^{1-\sigma} < (\frac{\mu\tau}{\widehat{\varphi_{fs}}})^{1-\sigma} - (\frac{\tau}{\widehat{\varphi_{fs}}} + \eta)^{1-\sigma} = \frac{f_{IS}}{B}$$

i.e. $(\frac{\mu\tau}{\widehat{\varphi_e}})^{1-\sigma} < \frac{f_X}{B} + \frac{f_{IS}}{B}$. Solving $\frac{\tau}{\widehat{\varphi_e}}$ from (2) and inserting into the inequality, then we can get:

$$B^{\frac{1}{1-\sigma}} < \frac{1}{\eta}\left[f_X^{\frac{1}{1-\sigma}} - \frac{1}{\mu}(f_X + f_{IS})^{\frac{1}{1-\sigma}}\right]$$

Thus if Δ is one upper bound of $B^{\frac{1}{1-\sigma}}$,then $\Delta < \frac{1}{\eta}\left[f_X^{\frac{1}{1-\sigma}} - \frac{1}{\mu}(f_X + f_{IS})^{\frac{1}{1-\sigma}}\right]$ensures $\widehat{\varphi_e} < \widehat{\varphi_{fs}}$.The existence of Δ is shown below.

From (3) and (4), we get

$$\left(\frac{\widehat{\varphi_{fm}}}{\widehat{\varphi_{fs}}}\right)^{1-\sigma} = \frac{f_{IS}\left[(\mu\tau)^{\sigma-1} - 1\right]}{f_{IM} - f_{IS} - f_X} \times \frac{1}{1 - \left(\frac{1}{\mu} + \frac{\eta\widehat{\varphi_{fs}}}{\mu\tau}\right)^{1-\sigma}} < \frac{f_{IS}\left[(\mu\tau)^{\sigma-1} - 1\right]}{f_{IM} - f_{IS} - f_X}\frac{1}{1 - \mu^{\sigma-1}}$$

So when $f_{IM} > f_X + f_{IS}\frac{\mu^{\sigma-1}}{1-\mu^{\sigma-1}}(\tau^{\sigma-1}-1)$,the above equation is smaller than one, then $\widehat{\varphi_{fs}} < \widehat{\varphi_{fm}}$.

Now we turn to the proof of the existence of Δ. Since φ follows the Pareto distribution, we have $\int_{\varphi}^{\infty} dG(\varphi) = \left(\frac{b}{\varphi}\right)^k$, $\int_{\varphi}^{\infty}(\frac{1}{\varphi})^{1-\sigma}dG(\varphi) = \frac{kb^k}{k-(\sigma-1)}\left(\frac{b}{\varphi}\right)^{k-(\sigma-1)}$.

When every country is symmetric, from (5) we get

$$B = \frac{EF}{VProfit1 + (N-1)\int_{\widehat{\varphi_e}}^{\widehat{\varphi_{fs}}}\left(\frac{\tau}{\varphi}+\eta\right)^{1-\sigma}dG(\varphi)}$$

where EF is the expected fixed cost of entry

$$\begin{aligned} EF &= f_E + f_D(\frac{b}{\widehat{\varphi_d}})^k + (N-1)f_X\left((\frac{b}{\widehat{\varphi_e}})^k - (\frac{b}{\widehat{\varphi_{fs}}})^k\right) \\ &\quad +(N-1)\left(f_X+f_{IS}\right)\left((\frac{b}{\widehat{\varphi_{fs}}})^k - (\frac{b}{\widehat{\varphi_{fm}}})^k\right) + (N-1)f_{IM}(\frac{b}{\widehat{\varphi_{fm}}})^k \\ &= f_E + f_D(\frac{b}{\widehat{\varphi_d}})^k + (N-1)f_X(\frac{b}{\widehat{\varphi_e}})^k \\ &\quad +(N-1)f_{IS}(\frac{b}{\widehat{\varphi_{fs}}})^k + (N-1)\left(f_{IM}-f_X-f_{IS}\right)(\frac{b}{\widehat{\varphi_{fm}}})^k \\ &> f_E \end{aligned}$$

And $VProfit1$ is the expected of variable profit of selling domestically, export with distribution FDI and building an overseas production plant.

$$\begin{aligned} VProfit1 &= \frac{kb^k}{k-(\sigma-1)}\left[\begin{array}{c} (\frac{1}{\widehat{\varphi_d}})^{k-(\sigma-1)} + (N-1)\left(\mu\tau\right)^{1-\sigma}\left((\frac{1}{\widehat{\varphi_{fs}}})^{k-(\sigma-1)} - (\frac{1}{\widehat{\varphi_{fm}}})^{k-(\sigma-1)}\right) \\ +(N-1)(\frac{1}{\widehat{\varphi_{fm}}})^{k-(\sigma-1)} \end{array}\right] \\ &= \frac{kb^k}{k-(\sigma-1)}\left[\begin{array}{c} (\frac{1}{\widehat{\varphi_d}})^{k-(\sigma-1)} + (N-1)\left(\mu\tau\right)^{1-\sigma}(\frac{1}{\widehat{\varphi_{fs}}})^{k-(\sigma-1)} \\ +(N-1)\left(1-(\mu\tau)^{1-\sigma}\right)(\frac{1}{\widehat{\varphi_{fm}}})^{k-(\sigma-1)} \end{array}\right] \end{aligned}$$

We assume $b > 1$, so $0 < (\frac{1}{\widehat{\varphi}})^{k-(\sigma-1)} < 1$ for any $\widehat{\varphi}$.Thus, the above equation satisfies$VProfit1 < \frac{kb^k}{k-(\sigma-1)}N$.

$(N-1)\int_{\widehat{\varphi_e}}^{\widehat{\varphi_{fs}}}\left(\frac{\tau}{\varphi}+\eta\right)^{1-\sigma}dG(\varphi)$ is the expected profit from export without FDI, and

$$\begin{aligned} (N-1)\int_{\widehat{\varphi_e}}^{\widehat{\varphi_{fs}}}\left(\frac{\tau}{\varphi}+\eta\right)^{1-\sigma}dG(\varphi) &< (N-1)\int_{\widehat{\varphi_e}}^{\widehat{\varphi_{fs}}}\left(\frac{\tau}{\widehat{\varphi_{fs}}}+\eta\right)^{1-\sigma}dG(\varphi) \\ &= (N-1)b^k\left(\frac{\tau}{\widehat{\varphi_{fs}}}+\eta\right)^{1-\sigma}\left((\frac{1}{\widehat{\varphi_e}})^k - (\frac{1}{\widehat{\varphi_{fs}}})^k\right) \\ &< (N-1)b^k\left(\frac{\tau}{\widehat{\varphi_{fs}}}+\eta\right)^{1-\sigma} < (N-1)b^k\eta^{1-\sigma} \end{aligned}$$

So $B > \frac{f_E}{\frac{kb^k}{k-(\sigma-1)}N+(N-1)b^k\eta^{1-\sigma}}$,and $\Delta = \left(\frac{f_E}{\frac{kb^k}{k-(\sigma-1)}N+(N-1)b^k\eta^{1-\sigma}}\right)^{\frac{1}{1-\sigma}}$ is an upper bound of $B^{\frac{1}{1-\sigma}}$.

8.2 Appendix B: Proof of Proposition 2

It is obvious from Equ. (2), (3) and (4) that an increase in η raises $\widehat{\varphi_e}$, lowers $\widehat{\varphi_{fs}}$,and does not affect $\widehat{\varphi_{fm}}$; a decrease in μ decreases $\widehat{\varphi_e}$, does not affect $\widehat{\varphi_e}$ and increases $\widehat{\varphi_{fm}}$; and a decrease in τ decreases $\widehat{\varphi_e}$ and increases $\widehat{\varphi_{fm}}$. We only need to verify a decrease in τ decreases $\widehat{\varphi_{fs}}$. Derive the LHS of (3) with respect to τ so that we get $d\left[(\frac{\mu\tau_{ij}}{\varphi})^{1-\sigma} - (\frac{\tau_{ij}}{\varphi}+\eta_j)^{1-\sigma}\right]/d(\tau) = \frac{(1-\sigma)}{\varphi}\left[\left(\mu^{\frac{\sigma-1}{\sigma}}\frac{\tau}{\varphi}\right)^{-\sigma} - \left(\frac{\tau}{\varphi}+\eta\right)^{-\sigma}\right] < 0$.So a lower τ leads to a higher relative variable profit from engaging in distribution FDI, thus generates a lower $\widehat{\varphi_{fs}}$.

It is worthwhile to note that when μ is sufficiently small, η is large and f_{IS} is not large enough, $\widehat{\varphi_{fs}}$ could be smaller than $\widehat{\varphi_e}$.For example, suppose $\sigma = 2$,and $f_{IS} < f_X(\frac{\tau+\eta}{\mu\tau} - 1)$,then from (2) and (3) we get

$$\frac{(\frac{\tau}{\widehat{\varphi_e}}+\eta)^{-1}}{(\frac{\mu\tau}{\widehat{\varphi_{fs}}})^{-1} - (\frac{\tau}{\widehat{\varphi_{fs}}}+\eta)^{-1}} = \frac{f_X}{f_{IS}}$$

Rearrange it to get

$$\begin{aligned}\frac{\widehat{\varphi_e}}{\widehat{\varphi_{fs}}} &= \frac{f_X}{f_{IS}}(\tau+\eta\widehat{\varphi_e})\left(\frac{1}{\mu\tau} - \frac{1}{\tau+\eta\widehat{\varphi_{fs}}}\right)\\ &= \frac{f_X}{f_{IS}}\left(\frac{\tau+\eta\widehat{\varphi_e}}{\mu\tau} - \frac{\tau+\eta\widehat{\varphi_e}}{\tau+\eta\widehat{\varphi_{fs}}}\right)\end{aligned}$$

Suppose $\frac{\widehat{\varphi_e}}{\widehat{\varphi_{fs}}} < 1$,then $\frac{\tau+\eta\widehat{\varphi_e}}{\tau+\eta\widehat{\varphi_{fs}}} < 1$,then $\frac{\widehat{\varphi_e}}{\widehat{\varphi_{fs}}} > \frac{f_X}{f_{IS}}\left(\frac{\tau+\eta\widehat{\varphi_e}}{\mu\tau} - 1\right) > \frac{f_X}{f_{IS}}(\frac{\tau+\eta}{\mu\tau} - 1) > 1$,contradicts. So in this case, $\frac{\widehat{\varphi_e}}{\widehat{\varphi_{fs}}} > 1, i.e. \widehat{\varphi_e} > \widehat{\varphi_{fs}}$. (Q.E.D.)

8.3 Appendix C: Distribution of Zhejiang FDI Firms

Zhejiang's firm-level FDI flow data are a good proxy for understanding the nationwide Chinese firms' FDI flow for the following reasons. First, the FDI flow from Zhejiang province is outstanding in the whole of China. Firms in Zhejiang have engaged in FDI since 1982. Such firms were the pioneers of Chinese FDI activity. As reported by MOC, only around 10 firms began to engage in FDI before 1982. Since then, Zhejiang has maintained a fast growth rate similar to that of other large eastern provinces, such as Guangdong, Jiangsu, and Shandong. In 2008, Zhejiang had 2,809 FDI firms (including greenfield firms and M&A firms), accounting for 21 percent of all FDI firms in China, and became the largest province in the number of FDI firms. In terms of FDI flow, Zhejiang's FDI also maintained a high plateau, ranking at the very top in the entire country from 2006 to 2009. Zhejiang's FDI accounted for 16 percent of the country's FDI flow and became the largest FDI province in 2010.

Second, the distribution of type of ownership of FDI firms in Zhejiang province is consistent with that across the country. According to the *Statistical Bulletin of China's outward Foreign*

Direct Investment (MOC, 2013) of the Ministry of Commerce, 75 percent of all Chinese FDI firms are private limited liability corporations in terms of the number of firms.[28] In Zhejiang province, 70 percent of FDI firms are private firms.

Third, the distribution of Zhejiang FDI firms' destinations is similar to that of the whole country. Up to 2009, Chinese FDI firms invested in 177 countries (regimes) and 71.4 percent of FDI volume was invested in Asia. Hong Kong is the most important destination for Chinese FDI firms.[29] This observation also applies to Zhejiang's FDI firms. Most FDI firms in Zhejiang invest in Asia, Europe, and North America. Hong Kong and the United States are the two destinations with the largest investments. The most common investment mode is to set up production affiliates and create a marketing network by establishing a trade-oriented office.

Finally, the industrial distribution of Zhejiang's FDI firms is similar to that for the whole of China. According to the *Statistical Bulletin of China's outward Foreign Direct Investment* (MOC, 2013), the top sector for Chinese FDI firmss investment is retail and wholesale. This is similar to the case of Zhejiang. The lower module of Table 1 shows the number of FDI firms in 2006–08, resulting in a total of 1,270 FDI firm-year observations in the database.

8.4 Appendix D: TFP Measure

The main interest of this paper is to investigate how firm productivity affects firm FDI. Hence, it is crucial to measure firm productivity accurately. Traditionally, TFP is measured by the estimated Solow residual between the true data on output and its fitted value using the OLS approach. However, the OLS approach suffers from two problems, namely, simultaneity bias and selection bias. Following Amiti and Konings (2007) and Yu (2015) in assuming a Cobb-Douglas production function, we adopt the augmented Olley-Pakes semi-parametric approach to deal with simultaneity bias and selection bias in measured TFP. In particular, we tailor the standard Olley-Pakes approach to fit the data for China with the following extensions.

First, we use deflated prices at the industry level to measure TFP. Previous studies, such as De Loecker (2011), stressed the estimation bias of using monetary terms to measure output when estimating the production function. In that way, one actually estimates an accounting identity. Hence, we use different price deflators for inputs and outputs. Admittedly, it would be ideal to adopt firm-specific prices as the deflators. Unfortunately, the firm-level data set does not provide sufficient information to measure prices of products. Following previous studies, such as Goldberg et al. (2010), we adopt industry-level input and output deflators for TFP measures. As in Brandt et al. (2012), the output deflators are constructed using "reference price" information from China's Statistical Yearbooks, whereas input deflators are constructed based on output deflators and China's national Input-Output Table (2002).

Third, it is important to construct the real investment variable when using the Olley-Pakes (1996) approach.[30] As usual, we adopt the perpetual inventory method to investigate the law of motion for real capital and real investment. The nominal and real capital stocks are constructed

[28] In 2013, SOEs accounted for only 8 percent of the total number of outward FDI firms, although they accounted for 55.5 percent of total outward FDI volume.

[29] Note that it is possible that some Chinese ODI firms take Hong Kong as an international investment exprót since Hong Kong is a popular "tax haven." This phenomenon is beyond the scope of the present paper, although it would be interesting for future research.

[30] In the literature, the Levinsohn and Petrin (2003) approach is also popular in constructing TFP in which materials (i.e., intermediate inputs) are used as a proxy variable. This approach is appropriate for firms in countries not using a large amount of imported intermediate inputs. However, such an approach may not directly apply to China, given that Chinese firms substantially rely on imported intermediate inputs, which have prices that are significantly different from those of domestic intermediate inputs (Halpern et al., 2011).

as in Brandt et al. (2012). Rather than assigning an arbitrary number for the depreciation ratio, we use the exact firm's real depreciation provided by the Chinese firm-level data set.[31]

In particular, by assuming that the expectation of future realization of the unobserved productivity shock, υ_{it}, relies on its contemporaneous value, firm i's investment is modeled as an increasing function of unobserved productivity and log capital, $k_{it} \equiv \ln K_{it}$. Following previous works such as Amiti and Konings (2007), the Olley–Pakes approach was revised by adding the firm's export decision as an extra argument in the investment function since most firms' export decisions are determined in the previous period:

$$I_{it} = \tilde{I}(\ln K_{it}, \upsilon_{it}, EF_{it}), \tag{7}$$

where EF_{it} is a dummy to measure whether firm i exports in year t. Therefore, the inverse function of (7) is $\upsilon_{it} = \tilde{I}^{-1}(\ln K_{it}, I_{it}, EF_{it})$.[32] The unobserved productivity also depends on log capital and the firm's export decisions. Accordingly, the estimation specification can now be written as:

$$\ln Y_{it} = \beta_0 + \beta_l \ln L_{it} + g(\ln K_{it}, I_{it}, EF_{it}) + \epsilon_{it}, \tag{8}$$

where $g(\ln K_{it}, I_{it}, EF_{it})$ is defined as $\beta_k \ln K_{it} + \tilde{I}^{-1}(\ln K_{it}, I_{it}, EF_{it})$. Following Olley and Pakes (1996) and Amiti and Konings (2007), fourth-order polynomials are used in log-capital, log-investment, firm's export dummy, and import dummy to approximate $g(\cdot)$.[33] In addition, since the firm dataset is from 2000 to 2006, we include a WTO dummy (*i.e.*, one for a year after 2001 and zero for before) to characterize the function $g(\cdot)$ as follows:

$$g(k_{it}, I_{it}, EF_{it}) = (1 + EF_{it}) \sum_{h=0}^{4} \sum_{q=0}^{4} \delta_{hq} k_{it}^{h} I_{it}^{q}. \tag{9}$$

After finding the estimated coefficients $\hat{\beta}_m$ and $\hat{\beta}_l$, we calculate the residual R_{it} which is defined as $R_{it} \equiv \ln Y_{it} - \hat{\beta}_l \ln L_{it}$.

The next step is to obtain an unbiased estimated coefficient of β_k. To correct the selection bias, Amiti and Konings (2007) suggest estimating the probability of a survival indicator on a high-order polynomial in log-capital and log-investment. One can then accurately estimate the following specification:

$$R_{it} = \beta_k \ln K_{it} + \tilde{I}^{-1}(g_{i,t-1} - \beta_k \ln K_{i,t-1}, \hat{pr}_{i,t-1}) + \epsilon_{it}, \tag{10}$$

where $\hat{pr}_i$ denotes the fitted value for the probability of the firm 's exit in the next year. Since the specific "true" functional form of the inverse function $\tilde{I}^{-1}(\cdot)$ is unknown, it is appropriate to use fourth-order polynomials in $g_{i,t-1}$ and $\ln K_{i,t-1}$ to approximate that. In addition, (10) also requires the estimated coefficients of the log-capital in the first and second term to be identical. Therefore, non-linear least squares seem to be the most desirable econometric technique. Finally, the Olley–Pakes type of TFP for each firm i in industry j is obtained once the estimated coefficient $\hat{\beta}_k$ is obtained:

$$TFP_{it}^{OP} = \ln Y_{it} - \hat{\beta}_k \ln K_{it} - \hat{\beta}_l \ln L_{it}. \tag{11}$$

[31] Note that even with the presence of exporting behavior, the data still exhibit a monotonic relationship between TFP and investment.

[32] Olley and Pakes (1996) show that the investment demand function is monotonically increasing in the productivity shock υ_{ik}, by making some mild assumptions about the firm's production technology.

[33] Using higher order polynomials to approximate $g(\cdot)$ does not change the estimation results.

8.5 Appendix E: Extensive Margin Estimates of Zhejiang Sample

Appendix Table 3 examines whether our previous findings based on nationwide FDI decision data hold for Zhejiang's FDI manufacturing firms. The linear probability model estimates in column (1) confirm that Zhejiang's high-productivity manufacturing firms are more likely to engage in FDI during the sample period 2006–08. The probit estimates in column (2) yield similar findings with a slightly larger coefficient of firm TFP. Of 100,743 manufacturing firms during the sample period, there are only 407 FDI manufacturing firms, as shown in the lower module of Table 1. That is, the probability of FDI is only 0.39 percent, suggesting that firm FDI activity is also a rare event in Zhejiang province during the sample period and the standard logit estimation results may have a downward bias. In Table 8, we again correct for such bias by using rare-events logit estimates in column (3) and complementary log-log estimates in column (4). The estimated coefficients of firm productivity are *much* larger than their counterparts in columns (1) and (2), indicating that the downward bias in the regular estimates is fairly large. The increases in the odds ratio caused by firm productivity are similar to their counterparts in Table 5.

Columns (5) and (6) perform multinomial estimates in which the regressand is distribution FDI in column (5) and non-distribution FDI in column (6). The coefficient of firm productivity in column (5) is still positive and significant whereas that in column (6) is positive but insignificant. A less important but interesting finding is that the SOE control variable turns to be positive and significant. We suspect such striking findings are due to the inclusion of processing FDI in the category of non-distribution FDI. By dropping such processing FDI in the multinomial estimates in columns (7) and (8), the coefficient of firm productivity in column (8) once again is positive and significant; more importantly, its magnitude is larger than that of distribution FDI, indicating that high-productivity firms are more likely to engage in non-distribution FDI. The coefficients of SOE variable turns to be negative, though still insignificant, as in other previous estimates.

Appendix Table 1: Summary Statistics of Distribution FDI

Firm Number	Nationwide FDI data 2000-2007			Zhejiang FDI data 2006-2008		
	Distribution FDI		Other FDI	Distribution FDI		Other FDI
Before Merge	2039	48%	2205	967	76%	304
After Merge	203	59%	142	337	83%	68

Note: The FDI type in Zhejiang data is classified to 12 types including wholesale, business office, production, processing trade, R&D, construction, mining, market seeking, agriculture, housing, and product design.

Appendix Table 2: Transitional Matrix of FDI Mode for FDI Firms

	Year t+1	
Year t	Distri. FDI	non-Distri. FDI
Distri. FDI	0.91	0.09
non-Distri. FDI	0.25	0.75

Note: Each number in the table is the probability of the firm's outward FDI mode in t+1, conditional on the FDI mode in t. Non-FDI firms in both periods are dropped.

Appendix Table 3: Extensive Margin Estimates for Zhejiang Firms (2006-08)

Econometric Method:	LPM	Probit	Rare events	Comp.	Multinomial Logit		Multinomial Logit	
			logit	log-log	Distribution	Non-Dist.	Distribution	Non-Dist.
Regressand: FDI Indicator	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Firm Relative TFP	0.006***	0.292**	0.882***	0.879***	0.961***	0.500	0.962***	1.648*
	(6.56)	(2.55)	(2.73)	(2.80)	(2.77)	(0.65)	(2.77)	(1.66)
SOE Indicator	0.003***	0.244***	0.678***	0.672***	0.660***	0.702***	-12.695	-12.791
	(14.38)	(11.88)	(12.74)	(12.22)	(10.53)	(5.11)	(-0.03)	(-0.01)
Foreign Indicator	-0.003	-0.343	-0.387	-0.854	-13.078	0.627	-0.222	0.398
	(-0.77)	(-0.91)	(-0.39)	(-0.84)	(-0.03)	(0.57)	(-1.34)	(0.93)
Log Firm Labor	-0.002***	-0.065	-0.172	-0.175	-0.221	0.178	0.660***	0.666***
	(-3.92)	(-1.23)	(-1.16)	(-1.19)	(-1.33)	(0.50)	(10.53)	(3.94)
Export Indicator	0.003***	0.664***	2.111***	2.127***	2.187***	1.866***	2.188***	3.112***
	(8.60)	(10.43)	(9.67)	(9.86)	(9.09)	(3.82)	(9.09)	(3.03)
Firm Tenure					0.328**	0.502	0.004	0.034***
					(2.01)	(1.36)	(0.41)	(2.81)
Processing FDI Dropped	No	No	No	No	–	No	–	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	100,847	100,743	100,847	100,847	100,847	100,847	100,832	100,832

Notes: The regressands in columns (1)-(4) are the FDI indicator. Numbers in parentheses are t-values. ***(**,*) denotes significance at the 1(5, 10)% level. The multinomial logit estimates in columns (5)-(6) include all non-FDI and FDI firms whereas those in columns (7)-(8) include all firms except processing FDI firms. The base types in all multinomial logit estimates are non-FDI firms.