

Pure Exporter: Theory and Evidence from China

Jiangyong Lu^a, Yi Lu^b, and Zhigang Tao^{c*}

^a Peking University

^b National University of Singapore

^c University of Hong Kong

First Draft: October 2009

This Version: April 2013

Abstract

This paper provides the first evidence about pure exporters (i.e., firms exporting all of their output to the foreign market) – a phenomenon overlooked and cannot be explained in the existing literature. It then offers a theoretical analysis of the existence and behavior of pure exporters. In particular, pure exporters arise when the export market is sufficiently large – a situation more likely to hold in developing countries as opposed to large developed countries; and their productivity levels are above those of non-exporters, but below those of firms having both domestic sales and export. These theoretical predictions are borne out in a data of Chinese manufacturing firms for the period of 1998-2005.

Keywords: Pure Exporter, Firm Heterogeneity, Exporting Behavior

JEL Codes: F12, F23, L22, D24

*We thank the editor and two anonymous referees for comments and suggests, Amber Li for sharing us the ASIF-Customs matched data.

Pure Exporter: Theory and Evidence

1 Introduction

In the past decade, there has been a growing literature on firm heterogeneity and exporting behavior. A dominant theoretical explanation is that more productive firms self-select to become exporters (e.g., Melitz, 2003). Specifically, in Melitz's framework, there is a fixed cost of production, and a fixed cost of exporting but no fixed cost of selling in the domestic market. As a result, in equilibrium there are only two types of firms: less productive firms sell only in the domestic market, while more productive firms have both domestic sales and export.¹

In reality, however, there are firms exporting all of their output (called *pure exporters*). For example, McMillan and Woodruff (1999) report that in their sample of firms in Vietnam, as high as 9% of them exports all of their production. Meanwhile, from a sample of Chinese manufacturing firms for the period of 1998-2005, we find that nearly 3% of firms (17% of exporters) are pure exporters. How to explain the existence of pure exporters? And what kinds of firms choose to become pure exporters? In this paper, we offer a theoretical explanation for pure exporters, and then test the theoretical predictions using a data of Chinese manufacturing firms.

In our theoretical analysis, we build upon Melitz (2003)'s framework by relaxing its key assumption about fixed cost of selling. Instead of assuming that there is no fixed cost of selling in the domestic market but a fixed cost of exporting, we assume that there is also a fixed cost for domestic sales albeit lower than that of exporting. Under such framework, we have the same results as in Melitz (2003) when the export market is not sufficiently large, that is, there are just two types of firms in equilibrium with the more productive firms having both domestic sales and export while the less productive firms selling only in the domestic market. However, when the export market is sufficiently large, there are three types of firms in equilibrium: firms having both domestic sales and export are the most productive, followed by pure exporters, and finally by firms with domestic sales only.

The intuition for our theoretical results is as follows. The profit functions for the three types of firms are similar, that is, a linear function of firm productivity, and the two key variables generating different profits across the three types of firms are the intercept (which is determined by the fixed costs

¹Existing empirical studies only include a dummy indicating whether or not a firm exports without distinguishing pure exporters from exporters that also have domestic sales (see for example, Clerides, Lach, and Tybout, 1998; Bernard and Jensen, 1999).

of operations) and the slope (which is determined by wage rate and market size). First, the three types of firms differ in the fixed costs of operations (including both production and sales): firms with both domestic sales and export have the highest fixed costs, followed by pure exporters, and finally by firms with domestic sales only. Second, as wage rate is the same across the three types of firms (because production takes place in the same country), the variations in the slope variable come from the differences in market size, with a bigger market size leading to a steeper slope.

It is clear that firms with both domestic sales and export have the steepest slope among the three. However, the ranking in the slope between firms with domestic sales only and firms with export only depends on the relative size of the domestic market vis-à-vis the export market. When the export market is sufficiently larger than the domestic market, firms with both domestic sales and export have the steepest slope, followed by pure exporters, and finally by firms with domestic sales only. Combined with the ranking in fixed costs of operations, it follows that firms having both domestic sales and export are the most productive, followed by pure exporters, and finally by firms with domestic sales only. However, when the export market is not sufficiently large, firms with domestic sales have steeper slope than pure exporters, and they dominate pure exporters as they also enjoy lower fixed costs of operations. Hence, in equilibrium, there are only two types of firms, with the more productive firms having both domestic sales and export while the less productive firms selling only in the domestic market.

Next, using the data set of Chinese manufacturing firms, we compare the three types of firms in terms of productivity. Preliminary statistics reported in Table 1 show that firms having both domestic sales and export always have the highest rank among the three types of firms in terms of employment, fixed assets, output, and productivity.² On the other hand, firms with domestic sales have the lowest ranking except in the category of fixed assets. For further empirical analysis, we regress firm productivity on a dummy variable for domestic sales only, and a dummy variable for domestic sales and export, together with a list of industry, region and year dummies. Regression results show that firms with domestic sales and export have the highest productivity, followed by pure exporters, and finally by firms with domestic sales only, consistent with our theoretical predictions.

These results are robust to the exclusion of outlying observations, to the inclusion of other firm characteristics (such as firm size, capital intensive,

²Here we estimate the total factor productivity using four different methodologies, that is, OLS, fixed-effect, GMM, and Levinsohn and Petrin (2003)'s. Details of these estimations are presented in Section 3.

and skill labor intensity), an alternative classification of Chinese domestic firms, two sub-samples (i.e., before and after WTO accession in 2002), and two alternative estimation methods (i.e., ordered Probit and multinomial logit). In particular, there are concerns that our finding of the sorting pattern among different types of firms may be driven by the special trading regime in China. Specifically, in China there are many firms that are allowed to import their inputs freely but have to sell all their outputs to foreign markets (namely *processing trader*). Meanwhile, many exporters are located in export processing zones (like Shenzhen, Zhuhai, Shantou), which have different policies and regulations about importing and exporting. Moreover, pure exporters may specialize in different foreign markets than those operating in both domestic and foreign markets, and hence our productivity measure may reflect rather the differences in export market conditions (Mayer, Melitz, and Ottaviano, 2012). To see whether our results are driven by these alternative explanations, we use the ASIF-Customs matched data, from which we observe whether an exporter is a processing trader, whether it is located in an export processing zone and how many foreign market it exports to, and find that our findings regarding the sorting pattern among firms remain robust to the direct control for these alternative explanations.

This paper contributes to the literature by being the first one documenting and then offering a theory to explain the existence and behavior of pure exporters. In particular, pure exporters arise when the export market is sufficiently large vis-a-vis the domestic market, a situation more likely to hold in developing countries than in large developed countries. This also explains why we are able to identify pure exporters that are overlooked in the existing literature, because our empirical work utilizes the data from China in contrast to most of the existing work that use data from large developed countries.

2 Data, Variables, and Descriptive Analysis

2.1 Data and Variables

Our empirical analysis uses data from annual surveys of industrial firms (ASIF) conducted by the National Bureau of Statistics of China for the period of 1998 to 2005. These annual surveys covered all state-owned enterprises, and those non-state-owned enterprises with annual sales of five million Chinese currency (about US\$650,000) or more. The data provides detailed information on firms' identification, operations and performance, including firm ownership, output and export, which are of special interest to this study.

The number of manufacturing firms varies from over 140,000 in the late 1990s to over 243,000 in 2005. The percentage of China's total exports contributed by firms in our dataset was just below 70% in late 1990s, and was as high as 76% in 2005, indicating that our data set is highly comprehensive.

According to the classification of the National Bureau of Statistics of China, firms with more than 25% equity shares held by foreign multinationals are classified as foreign affiliates, and the rest is classified as China's domestic firms. As the literature almost exclusively examines the exporting behavior of domestic firms, in this study we also focus on the exporting behavior of domestic firms, and hence simply refer to domestic firms as firms.³ The number of firms in China with valid information on export, output, employment, fixed assets and intermediate inputs ranges from 112,246 in 1998 to 192,234 in 2005. And totally, we have 1,039,792 firm-year observations

To estimate total factor productivity (TFP), we first use the OLS regression method. Specifically, we use the constant value of output, deflate the fixed assets by the fixed-assets investment price index and intermediate inputs by the producer price index, and estimate for firms in each 2-digit industry and each year (see also Bernard and Jensen, 1999). The OLS estimation of TFP, however, may suffer from the simultaneity problem, specifically, input choices could be endogenously determined by unobservable productivity shocks. This may lead to an upward bias in the estimation coefficients of more variable inputs such as capital (Van Biesebroeck, 2007). We therefore use three alternative estimation methods, that is, panel fixed-effect estimation, the instrumental estimation (i.e., GMM), and semi-parametric estimation⁴ (i.e., Levinsohn and Petrin (2003)'s TFP estimation method).⁵ Table 2 provides the correlation among these four different measures of TFP.

³In addition, as shown in Lu, Lu, and Tao (2010), the exporting behavior of foreign affiliates is rather different and complicated, as multinationals can choose the location of production as well as the location of market.

⁴Another semi-parametric estimation method for dealing with the endogeneity problem is Olley and Pakes method (1996), which uses investment as a proxy for unobservable productivity shocks. However, there is substantial missing information on investment in our data. Therefore Olley and Pakes method is not econometrically efficient in our case.

⁵However, these three alternative estimation methods may also have their own estimation concerns, for example, the semi-parametric estimation may lead to larger biases than the OLS estimates if unobservable productivity shocks are mostly transitory and the firm fixed effects are significant. For detailed discussion on the differences among various methods for estimating TFP, please see Van Biesebroeck (2007).

2.2 Pure Exporters

Based on their export figures, we can potentially classify firms into three types: firms only sell in the Chinese market (referred to as *nonexporters*), firms sell in both the Chinese and foreign markets (referred to as *regular exporters*), and firms sell only in the foreign market (referred to as *pure exporters*). As shown in Panel A of Table 1, for the period of 1998-2005, 80.96% of firm-year observations belongs to the nonexporters group, 15.75% belongs to the regular exporters group, and finally 3.29% belongs to the pure exporters group. In Panels B-E, we look at the ratios for different years, i.e., 1998, 2000, 2002, and 2005. Clearly, similar patterns hold for each of these years: a majority of firms is nonexporter and around 18% of exporters is pure exporter.

The interesting finding of pure exporters presents a challenge to the existing firm heterogeneity literature, as in the Melitz (2003)'s model, firms either sell only in the domestic market or sell in both the domestic and foreign markets. Why are there pure exporters? And what kinds of firms choose to become pure exporters?

To have a rough idea of the behavior of pure exporters, we calculate the mean values of their several key characteristics (i.e., output, employment, fixed assets and TFP), and compare them to those of nonexporters and regular exporters. As shown in Panel A of Table 1, we find that for the whole sample period (i.e., 1998-2005), *regular exporters* always have larger output, employment and fixed assets, and higher TFP than *pure exporters*. Meanwhile, *pure exporters* are better than *non-exporters* in terms of output, employment, and TFP. The same pattern holds when we compare these three types of firms at different years (i.e., 1998, 2000, 2002, and 2005) in Panels B-D of Table 1.

To further corroborate our finding of the ranking order among these three types of firms, we plot the three distributions of TFP corresponding to each of the three types of firms (i.e., nonexporters, regular exporters, and pure exporters) in Figure 1. Clearly, we find that the TFP distribution of regular exporters dominates that of pure exporters, which in turn dominates that of non-exporters.

To explain such ranking order, we, in the next two sections, first provide a theory built upon Melitz (2003)'s framework, and then test our theoretical predictions using more rigorous empirical analysis.

3 Theoretical Analysis

In this section, we build upon Melitz (2003)'s framework to explain the existence of pure exporters and analyze its differences from other two types of firms (exporters and non-exporters).

Consider a world of two countries (i.e., home (H) and foreign (F)), two sectors (i.e., a homogeneous good (X) produced with a constant returns to scale technology and a continuum of differentiated goods (Y) produced with an increasing returns to scale technology), and one production factor (labor; mobile across sectors within a country but immobile across countries).

Following the literature, we take the homogeneous good (X) as a numéraire and assume the utility function for the differentiated goods (Y) to be a constant elasticity of substitution (CES) function. Then the demand function for variety ω of the differentiated goods Y in country l can be derived as:

$$y^l(\omega) = \alpha^{\frac{-\alpha}{1-\alpha}} I^l (p^l(\omega))^{\frac{-1}{1-\alpha}} \quad (1)$$

where $l \in \{H, F\}$ is the index for the country; $y^l(\omega)$ is the consumption of variety ω of the differentiated goods Y in country l ; $I^l \equiv M^l (Y^l)^{\frac{\mu-\alpha}{1-\alpha}}$ is the measure for the size of market in country l , where M^l is the number of consumers and Y^l is the index of aggregate consumption of differentiated goods in country l ; and $p^l(\omega)$ is the price of variety ω in country l . The elasticity of substitution between any two differentiated goods is $\sigma \equiv 1/(1 - \alpha) > 1$. The variety parameter ω is left out hereon as all the cases are symmetric.

The production of the differentiated goods (y) takes place in the home country. The unit production cost is given by c/θ , where θ is the firm-specific productivity measure drawn from a common distribution. Meanwhile, the fixed cost of production is same across all firms and given by f_p . Moreover, the transport cost of differentiated goods to the foreign market takes the form of an iceberg cost, i.e., one needs $t > 1$ units of final product in order to ship 1 unit to an abroad market.

Thus far the setup is the same as in Melitz (2003). The departure of our model from his lies in the assumption about the fixed cost of selling the differentiated goods. In Melitz (2003), it is assumed that there is zero fixed cost of selling in the home market, but a positive fixed cost of selling in the foreign market. In contrast, we assume that there is also a positive fixed cost of selling the home market (denoted by f_s^H), though it is lower than the fixed cost of selling in the foreign market (denoted by f_s^F), which is lower than the fixed cost of selling in both markets (denoted by f_s^{HF}), i.e., $0 < f_s^H < f_s^F < f_s^{HF}$.

A firm needs to decide where to sell its products. There are three possible choices: selling only in the home market (non-exporters), selling only in the foreign market (pure exporters), and selling in both home and foreign markets. For ease of exposition, we denote these three choices by $\sigma(H)$, $\sigma(F)$, and $\sigma(HF)$, respectively.

Given the above setup, we can derive the equilibrium profit function for these three choices as:

$$\left\{ \begin{array}{l} \pi_{\sigma(H)} = \frac{(1-\alpha)I^H}{C} \Theta - (f_p + f_s^H) \\ \pi_{\sigma(F)} = \frac{(1-\alpha)I^F}{C} \Theta - (f_p + f_s^F) \\ \pi_{\sigma(HF)} = \frac{(1-\alpha)(I^H + \frac{I^F}{T})}{C} \Theta - (f_p + f_s^{HF}) \end{array} \right. , \quad (2)$$

where $\Theta \equiv \theta^{\frac{\alpha}{1-\alpha}}$ is a monotonic transform of productivity θ ; $C \equiv c^{\frac{\alpha}{1-\alpha}}$ is a monotonic transform of unit production cost c ; $T \equiv t^{\frac{\alpha}{1-\alpha}}$ is a monotonic transform of transport cost t ; and I^l is the market size in country l , $l \in \{H, F\}$.

Note that the profit function for each of these three choices is a linear function of Θ , and it just differs in the slope term (denoted by ϕ) and the intercept term (the negative of all the fixed costs, denoted by F). The comparison of the fixed costs across the three choices is straightforward, in which:

$$F_{\sigma(H)} < F_{\sigma(F)} < F_{\sigma(HF)}, \quad (3)$$

where $F_{\sigma(H)} = f_p + f_s^H$; $F_{\sigma(F)} = f_p + f_s^F$; and $F_{\sigma(HF)} = f_p + f_s^{HF}$.

The slope term (ϕ) is determined by the unit cost of production (the denominator, C) and the size of the markets (the nominator, $\sum I^l$). As the production takes place only in the home market, the three choices have the same unit cost of production, and they only differ in the size of the markets. The choice $\sigma(HF)$ involves the selling in both the home and the foreign markets, and thus it has the largest market coverage or the steepest slope term. The comparison of the slope term between the choice $\sigma(H)$ and the choice $\sigma(F)$ hinges upon the relative size of the home market and the foreign market (adjusted by the transport cost). When the (transport-cost-adjusted) foreign market is smaller than the home market (that is, $\frac{I^F}{T} < I^H$), the slope term of the choice $\sigma(H)$ is steeper than that of the choice $\sigma(F)$. When the (transport-cost-adjusted) foreign market is larger than the home market (that is, $\frac{I^F}{T} \geq I^H$), the slope term of the choice $\sigma(F)$ is steeper than that of the choice $\sigma(H)$. So we have the following ranking of the slope term for these three choices:

$$\begin{cases} \phi_{\sigma(F)} < \phi_{\sigma(H)} < \phi_{\sigma(HF)} & \text{when } \frac{I^F}{T} < I^H \\ \phi_{\sigma(H)} < \phi_{\sigma(F)} < \phi_{\sigma(HF)} & \text{when } \frac{I^F}{T} > I^H \end{cases}, \quad (4)$$

where $\phi_{\sigma(H)} = \frac{(1-\alpha)}{C} I^H$; $\phi_{\sigma(F)} = \frac{(1-\alpha)}{C} \frac{I^F}{T}$; and $\phi_{\sigma(HF)} = \frac{(1-\alpha)}{C} \left(I^H + \frac{I^F}{T} \right)$.

Figure 2 shows the optimal choice when the foreign market is not sufficiently large, in which in the equilibrium there are only two types of firms, regular exporter and nonexporters, and the former has higher productivity than the latter.

Figure 3 shows the optimal choice when the foreign market is sufficiently large. In the equilibrium, all three types of firms exist, and the most productive firms choose to sell both in the home and foreign markets, the least productive firms sell only in the home market, and those in the middle sell only in the foreign market.

We conclude our theoretical findings in the following Proposition:

Proposition: *When the foreign market is not sufficiently large, in equilibrium there are only two types of firms: the more productive firms sell in both the home and foreign markets, while the less productive firms sell only in the home market (the non-exporters). When foreign markets are sufficiently large, in equilibrium there are three types of firms: the most productive firms sell in both the home and foreign markets, the least productive ones sell only in the home market (the non-exporters), and those in the middle sell only in foreign markets (the pure exporters).*

Proof: See the Appendix.

The intuition for the proposition is as follows. For the case where the foreign market is not sufficiently large, the choice of selling only in the foreign market ($\sigma(F)$) is always dominated by the choice of selling only in the home market ($\sigma(H)$). This is because the former has a higher fixed costs but a smaller market coverage than the latter. Meanwhile, compared with the choice of selling in both the home and the foreign markets ($\sigma(HF)$), $\sigma(H)$ has a lower fixed costs but a smaller market coverage. Thus, the equilibrium choice depends on firm productivity as elucidated in the literature on firm heterogeneity and exporting behavior, with the more productive firms choosing $\sigma(HF)$ while the less productive ones choosing $\sigma(H)$.

For the case where the foreign market is sufficiently large, none of these three choices is always dominated by others. As we move from the choice of selling only in the home market ($\sigma(H)$), to the choice of selling only in the

foreign market ($\sigma(F)$), and finally to the choice of selling in both the home and foreign markets ($\sigma(HF)$), the fixed costs are increasing (i.e., $F_{\sigma(H)} < F_{\sigma(F)} < F_{\sigma(HF)}$), but so are the market coverage (i.e., $\phi_{\sigma(H)} < \phi_{\sigma(F)} < \phi_{\sigma(HF)}$). The equilibrium choice depends on firm productivity, namely, the most productive firms choose $\sigma(HF)$, the least productive ones choose $\sigma(H)$, and those in the middle choose $\sigma(F)$.

It is interesting to point out why pure exporters do not exist in equilibrium under Melitz (2003)'s framework. In Melitz (2003), it is assumed that the fixed cost of selling in the home market is zero (i.e., $f_s^H = 0$). Under this assumption, the choice of selling only in the foreign market ($\sigma(F)$) is always dominated by the choice of selling in both the home and the foreign markets ($\sigma(HF)$). This is because the former has the same fixed costs as the latter (i.e., $F_{\sigma(F)} = F_{\sigma(HF)}$), but has a smaller market coverage than the latter (i.e., $\phi_{\sigma(F)} < \phi_{\sigma(HF)}$). Intuitively, as there is no extra fixed cost of selling in the home market, firms always have sales in the home market.

Under our framework (i.e., $0 < f_s^H < f_s^F$), however, pure exporters may exist in equilibrium, and the condition for its existence is that the foreign market is sufficiently larger than the home market. However, if this condition is not satisfied, the choice of selling in the foreign market (or the pure exporters) is dominated by the choice of selling only in the home market (or the non-exporters), and the equilibrium choice is between selling only in the home market ($\sigma(H)$) and selling in both the home and foreign markets ($\sigma(HF)$) just as in Melitz (2003).

4 Empirical Analysis

4.1 Main Results

To further investigate the exporting behavior of firms in China, we estimate the following equation:

$$TFP_{firt} = \alpha + \beta \cdot Home_{firt} + \gamma \cdot Home\ and\ Foreign_{firt} + \delta_i + \eta_r + \lambda_t + \varepsilon_{firt} \quad (5)$$

where TFP_{firt} is the TFP of firm f in industry i , region r and year t ; $Home_{firt}$ is a dummy variable having value of one if firm f sells only in the home market, and zero otherwise; $Home\ and\ Foreign_{firt}$ is a dummy variable having value of one if firm f sells in both the home and foreign markets, and zero otherwise; δ_i , η_r and λ_t are 4-digit industry dummy, region dummy,⁶ and year dummy, respectively; and ε_{firt} is the error term. To deal

⁶Region here refers to 22 provinces, 4 province-level municipalities, and 5 minority autonomous regions in China.

with the possible heteroskedasticity problem, we use the robust standard error clustered at the firm level.

Regression results for equation (5) are reported in Table 3. We use TFP estimated using Levinsohn and Petrin (2003)'e method as the dependent variable in Column (1), TFP estimated using OLS method as the dependent variable in Column (2), TFP estimated using panel fixed-effect method as the dependent variable in Column (3), and finally TFP estimated using GMM as the dependent variable in Column (4). It is clear that in all these regressions, the coefficient for $Home_{firt}$ is negative and statistically significant, whereas the coefficient for $Home$ and $Foreign_{firt}$ is positive and statistically significant. These results suggest that firms having sales in both the home and foreign markets are the most productive, followed by firms with sales only in the foreign market, and finally, by firms with sales only in the home market. The sorting pattern in terms of productivity across the three types of firms reinforces our preliminary comparison in Table 1 and leads support to our theoretical predictions in the Proposition in Section 2.

Note that our Proposition suggests that pure exporters arise only when the export market is sufficiently large vis-à-vis the domestic market. It is reasonable to argue that the condition of sufficiently large export market holds more likely for developing countries such as China as compared with large developed countries such as the United States. Indeed our identification of the existence of pure exporters among the Chinese manufacturers lends support to the above argument. To further provide corroborative evidence for such argument, we, from the Chinese Customs data (from 2000 to 2006), obtain the information about which foreign markets each 3-digit industry-province exports to and how many. Based on such information, we calculate for each industry-province a weighted average GDP of all these foreign markets with the weight being the export share as a proxy for the size of the concerned firm's potential export market (denoted as *Export Market Size*).⁷ Specifically, *Export Market Size* is constructed as follows

$$Export\ Market\ Size_{irt} = \sum_j \omega_{irjt} GDP_{jt},$$

where i , r , j , and t represent three-digit industry, region, foreign country, and year; $\omega_{irjt} \equiv \frac{export_{irjt}}{\sum_j export_{irjt}}$ is the export share of foreign country j ; and GDP_{jt} is the GDP of foreign country j at year t . To avoid the issue that the observed exporter status and the number of export markets are jointly determined, we use the lagged values in the regressions.

⁷We thank the associate editor for providing this suggestion. And we also thank Yifan Zhang for sharing with us the concordance table between HS-6 and SIC-3.

We, in Columns 1-2 of Table 4, first conduct the firm-level analysis, in which we pool all the firm-year observations and regress an indicator of *Pure Exporter* on *Export Market Size* along with several firm characteristics (such as TFP, wage, firm size, and capital labor ratio) and the dummies of industry, region and year. Clearly, we find that it is more likely to observe a pure exporter if the export market size is larger, leading further support to our theoretical predictions.

In Columns 3-4 of Table 4, we use another specification, that is, we aggregate all the relevant firm-year level variables to the industry-region-year level (using the mean values), and regress the ratio of pure exporters (in terms of number) in an industry-region cell on *Export Market Size* along with all the controls. It is found that the share of pure exporters is higher in industry-region with larger potential export markets, consistent with the firm-level analysis in Columns 1-2 of Table 4 and our theoretical predictions.

4.2 Robustness Checks

In this sub-section, we conduct a series of robustness checks on the productivity ranking of *nonexporters*, *pure exporters*, and *regular exporters*. As the results with each of these four measures of TFP are similar, we only report the estimation results using TFP estimated by Levinsohn and Petrin (2003)'s method as the dependent variable to save space.

Outlying Observations. To address the concern that our results could be driven by some outlying observations, we exclude the top and bottom 1% observations in our sample and repeat the analysis. The results shown in Column 1 of Table 5 demonstrate clearly that our findings in Table 3 remain robust.

Additional Controls. To make sure that our findings are not entirely driven by other firm characteristics, we alternatively incorporate *Firm Size* (measured as the logarithm of employment), *Capital Intensive* (measured as the logarithm of capital labor ratio), and *Skilled Labor Intensity* (measured as the logarithm of the ratio of labor with collage degree or above) in the regression analysis.⁸ The result reported in Columns 2-4 of Table 5 reveal that our findings remain robust to the inclusion of these additional firm characteristics.

Alternative Definition of Domestic Firms. We use an alternative definition of domestic firms – the official ownership type reported by firms in

⁸Note that as the information about labor's education level is only available in 2004, the regression with the inclusion of *Skilled Labor Intensity* only use the sample of year 2004, as a result of which the sample size is reduced to 117,607 from 1,039,792.

the survey – instead of that implied by equity ownership. Specifically, there are five types of ownership: state-owned firms, collectively-owned firms, joint-stock companies, privately-owned firms, and foreign-invested firms. We treat firms with the first four types of ownership as domestic firms. As shown in Column 5 of Table 5, our findings remain robust to this alternative definition of domestic firms.

Different Sub-samples. We split the whole sample into two sub-samples to take care of the possible changes of exporting behavior over time. In particular, China entered into the WTO near the end of 2001, which might facilitate the export of China’s domestic firms and enlarge the foreign market vis-à-vis the domestic market. Hence, we split the sample period into two, the pre-WTO period (1998-2001) and the post-WTO period (2002-2005), and repeat the analysis. As shown in Columns 6-7 of Table 5, the estimated coefficients for $Home_{firt}$ are negative and statistically significant for both the pre- and the post-WTO periods, though the magnitude of the coefficient drops substantially from the pre- to the post-WTO period. Intuitively, with China’s entry into the WTO, entry barriers into the foreign market are lowered down (or transport cost t drops in our model), which narrows down the productivity gap between non-exporters and pure exporters. Meanwhile, the estimated coefficients for $Home$ and $Foreign_{firt}$ are positive and statistically significant for both the pre- and the post-WTO periods, with similar magnitudes. Intuitively, the productivity gap between pure exporters and firms with sales in both the home and foreign markets is driven by the fixed cost of selling in the home market as well as the size of the home market, none of which is significantly affected by China’s entry into the WTO.

Alternative Estimation Methods. To reflect the self-selection feature of exporting behavior by firms in terms of their productivity levels as stated in the Proposition, we use two alternative estimation methods, that is, ordered Probit and multinomial logistic estimation. Specifically, we construct a new variable, called $Exporting Status_{firt}$, which takes a value of 1 if a nonexporter, a value of 2 if a pure exporter, and a value of 3 if a regular exporter. According to the order of self-selection found in the Proposition, we have

$$\begin{cases} Exporting Status_{firt} = 1 & \text{if } \Theta \leq \gamma_1 \\ Exporting Status_{firt} = 2 & \text{if } \gamma_1 < \Theta \leq \gamma_2 , \\ Exporting Status_{firt} = 3 & \text{if } \gamma_2 < \Theta \end{cases} \quad (6)$$

where Θ is firm productivity. The Ordered Probit regression estimates not only the effect of firm productivity, but also the two cutoff points (i.e., γ_1 , and γ_2), from which we can test whether they are in the increasing order and statistically different. Regression results are reported in Columns 1-2 of Table

6 (without and with firm controls, respectively). It is found that in both regressions the estimated coefficients of firm productivity are positive and statistically significant. These results indicate that along with the increase in productivity, a firm is more likely to switch from nonexporter to pure exporter (i.e., those with productivity above γ_1), and to exporter (i.e., those with productivity above γ_2). Moreover, in both specifications, the estimated cutoff points of productivity (i.e., γ_1 and γ_2) display an increasing order, that is, $\gamma_1 < \gamma_2$, and the Chi2 tests show that this order is statistically significant.

One potential concern of the Ordered Probit estimation is that it already imposes an order on the outcome choice, that is, outcomes with higher values are better (or worse) than those with lower values. As a way of checking whether our findings are due to this artificial ordering problem, we also use the multinomial logit regression. Set *Exporting Status*_{*firt*} = 1 as the base outcome and the multinomial logistic estimation generates two relative risk ratios, corresponding to the other two outcomes (that is, *Exporting Status*_{*firt*} = 2 and *Exporting Status*_{*firt*} = 3). A relative risk ratio for the explanatory variable X_k measures the change in the predicted odds favoring *Exporting Status*_{*firt*} = $j \in \{2, 3\}$ relative to the base outcome *Exporting Status*_{*firt*} = 1 associated with an 1-unit increase in X_k . In other words, the relative risk ratio (rrr_{j1}) for X_k takes the following form:

$$rrr_{j1} = \frac{P(\text{Exporting Status}_{firt} = j | X_k + 1)}{P(\text{Exporting Status}_{firt} = 1 | X_k + 1)} \bigg/ \frac{P(\text{Exporting Status}_{firt} = j | X_k)}{P(\text{Exporting Status}_{firt} = 1 | X_k)}.$$

Hence, $rrr_{j1} > 1$ means that with an increase in X_k , a firm is more likely to choose outcome value j relative to the base outcome; whereas $rrr_{j1} < 1$ means that with an increase in X_k , a firm is less likely to choose outcome value j relative to the base outcome. Regression results are reported in Columns 3-4 of Table 6 (without and with firm controls, respectively). As results are consistent across these two specification, we explain our findings using the results in Column 10. The relative risk ratio for outcome *Exporting Status*_{*firt*} = 2 over basic outcome *Exporting Status*_{*firt*} = 1 (rrr_{21}) is found to be $1.392 > 1$ and statistically significant. This means that with an increase in firm productivity level in the last period, a firm is more likely to switch from selling only in the home market to selling only in foreign markets, which is consistent with the theoretical prediction in the Proposition. Given the relative risk ratio rrr_{31} , we can calculate the relative risk ratio for outcome *Exporting Status*_{*firt*} = 3 over outcome *Exporting Status*_{*firt*} = 2, that is

$$rrr_{32} = rrr_{31} / rrr_{21} = 2.268 > 1.$$

This means that an increase in firm productivity level is associated with a higher probability of selling in both the home and foreign markets than

selling only in foreign markets, which is again consistent with the theoretical prediction in the Proposition.

Other Potential Explanations. Thus far, we have documented that the productivity level of pure exporters lies in between the levels of nonexporters and regular exporters, and we explain such finding as the self-selection by firms in terms of their productivity levels. However, there are some other potential explanations, e.g., related to the special trading regime in China. Specifically, when China opened its economy after 1978, it first established some export processing zones (e.g., Shenzhen, Zhuhai, Shantou) and only allowed exports and imports within these special zones, to protect its fragile domestic economy. Given that firms located in export processing zones are export-oriented (and many are pure exporters) and these special zones have some favorable policies toward exporters, one may be concerned that our findings of pure exporters could be driven by the location of these firms, in particular, the export processing zones. To address this concern, we first match our ASIF firm-level data to the China Customs data,⁹ from which we obtain the information about whether an exporter is located in an export processing zone or not. And then we construct an additional control, namely *Export Processing Zone*, which takes a value of 1 for exporters located in the export processing zone and a value of 0 for exporters located outside the export processing zone and nonexporters. As shown in Column 1 of Table 7, our main results regarding the ranking order among the three different types of firms remain robust to this additional control, indicating that our results are not driven by the location of the exporters.

Meanwhile, as another strategy to open but also protect the domestic economy, foreign firms were allowed to import their inputs freely but had to sell all their outputs to foreign markets, so called *processing trade*.¹⁰ Given the pure exporter nature of these processing traders and the technology advancement of foreign firms, it could be possible that our results are driven by these processing traders. To address this concern, we again use the ASIF-Customs matched data, from which we collect the information whether an exporter conduct processing trade, ordinary trade, or both. We then construct an additional control, namely *Processing Trader*, which takes a value of 1 if an exporter conducts any processing trade and 0 otherwise. Regression results are reported in Column 2 of Table 7. Clearly, our main results still remain robust.¹¹

⁹For the use of the ASIF-Customs matched data, see Fan, Li and Lai (2012), Ma, Tang, and Zhang (2012), Wang and Yu (2012)

¹⁰For papers looking at processing trade in China, see Manova and Yu (2012), Wang and Yu (2012), etc.

¹¹Alternatively, we have experimented with excluding processing traders or firms located

Finally, there is a concern that pure exporters may specialize in different foreign markets than those operating in both domestic and foreign markets, and hence our productivity measure may reflect rather the differences in export market conditions (Mayer, Melitz, and Ottaviano, 2012). To address this concern, we use the *Export Market Size* variable constructed in Section 4.1 to capture the potential exposure of different foreign markets to different firms and include it as an additional control in the analysis. Regression results are reported in Column 3 of Table 7. Again, we find that our main results remain robust to this additional control.

In summary, the findings of our results robust to the control for export processing zones, processing trader or the size of export markets imply that our results cannot be mainly explained by these alternative theories.

5 Conclusion

This paper studies the existence and behavior of pure exporters, which are overlooked and cannot be explained by the existing literature. Building upon Melitz (2003)'s framework, we first identify the condition for the existence of pure exporters, that is, the sufficiently large foreign market relative to the domestic market. We then show that in the presence of pure exporters, their productivity levels are above those of non-exporters, but below those of firms having both domestic sales and export. To examine the relevance of these theoretical predictions, we use a data of manufacturing firms for the period of 1998-2005 from China, for which the foreign market is arguably much larger compared with the domestic market. From this data, we find quite a substantial number of pure exporters, and their productivity ranking vis-à-vis the other two types of firms highly consistent with our theoretical predictions.

in export processing zones, and find similar results (available upon request).

References

- [1] Bernard, A. B. and J. B. Jensen (1999). "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47(1): 1-25.
- [2] Clerides, S. K., S. Lach, and J. R. Tybout (1998). "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics* 113(3): 903-947.
- [3] Fan, H., Y. Li, and E. Lai (2012). "Credit Constraints, Quality, and Export Prices: Theory and Evidence from China." working paper.
- [4] Levinsohn, J. and A. Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies* 70(2): 317-342.
- [5] Lu, J., Y. Lu, and Z. Tao (2010). "Exporting Behavior of Foreign Affiliates: Theory and Evidence." *Journal of International Economics* 81: 197-205.
- [6] McMillan, J. and C. Woodruff (1999). "Interfirm Relationships And Informal Credit In Vietnam." *Quarterly Journal of Economics* 114(4): 1285-1320.
- [7] Ma, Y., H. Tang, and Y. Zhang (2012). "Productivity, Factor Intensity, and Product Switching: Evidence from Chinese Exporters." working paper.
- [8] Manova, K. and Z. Yu (2012). "Firms and Credit Constraints along the Value-Added Chain: Processing Trade in China." NBER Working Paper 18561.
- [9] Melitz, M. J. (2003). "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71(6): 1695-1725.
- [10] Melitz, M.J., T. Mayer, and G. Ottaviano (2012). "Market Size, Competition, and the Product Mix of Exporters." working paper.
- [11] Olley, G.S. and A. Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64: 1263-1297.

- [12] Van Biesebroeck, J. (2007). "Robustness of Productivity Estimates." *Journal of Industrial Economics* 55(3): 529-569.
- [13] Wang, Z. and Z. Yu (2012). "Trading Partners, Traded Products, and Firm Performances of China's Exporter-Importers: Does Processing Trade Make a Difference?" *World Economy*.35: 1795-1824.

Appendix

Proof of Proposition: Let us first consider the case that the foreign market is not sufficiently large, i.e., $\frac{I^F}{T} < I^H$. In this case, the choice $\sigma(F)$ (selling only in the foreign market or pure exporting) is dominated by the choice $\sigma(H)$ (selling only in the home market or domestic sale only), as

$$\begin{aligned}\pi_{\sigma(F)} &= \frac{(1-\alpha)\frac{I^F}{T}}{C}\Theta - (f_p + f_s^F) \\ &< \frac{(1-\alpha)I^H}{C}\Theta - (f_p + f_s^F) \\ &< \frac{(1-\alpha)I^H}{C}\Theta - (f_p + f_s^H) = \pi_{\sigma(H)}.\end{aligned}$$

The first inequality comes as $\frac{I^F}{T} < I^H$, while the second inequality is due to the assumption $0 < f_s^H < f_s^F$. Hence, in the equilibrium, there are only two available choices: $\sigma(H)$ and $\sigma(HF)$ (selling in both home and foreign markets).

Denote the cutoff point Θ_1 as $\pi_{\sigma(H)}(\Theta_1) = 0$, i.e.,

$$\Theta_1 = \frac{f_p + f_s^H}{(1-\alpha)I^H}C$$

and the cutoff point Θ_2 as $\pi_{\sigma(H)}(\Theta_2) = \pi_{\sigma(HF)}(\Theta_2)$, i.e.,

$$\Theta_2 = \frac{f_s^{HF} - f_s^H}{(1-\alpha)\frac{I^F}{T}}C$$

When $\frac{I^F}{T} < \frac{f_s^{HF} - f_s^H}{f_p + f_s^H}I^H$, we have

$$0 < \Theta_1 < \Theta_2.$$

Thus, we have a clear dichotomy that when the foreign market is not sufficiently large, more productive firms sell in both the home and the foreign markets and less productive firms sell only in the home market.

Next, let us consider the case that foreign market is sufficiently large, i.e., $\frac{I^F}{T} > I^H$. In this case, in the equilibrium, there are three available choices: $\sigma(H)$, $\sigma(F)$ and $\sigma(HF)$.

Denote the cutoff point Θ'_1 as $\pi_{\sigma(H)}(\Theta'_1) = 0$ i.e.,

$$\Theta'_1 = \frac{f_p + f_s^H}{(1-\alpha)I^H}C$$

and the cutoff point Θ'_2 as $\pi_{\sigma(H)}(\Theta'_2) = \pi_{\sigma(F)}(\Theta'_2)$, i.e.,

$$\Theta'_2 = \frac{f_s^F - f_s^H}{(1 - \alpha) \left(\frac{I^F}{T} - I^H \right)} C$$

and the cutoff point Θ'_3 as $\pi_{\sigma(F)}(\Theta'_3) = \pi_{\sigma(HF)}(\Theta'_3)$, i.e.,

$$\Theta'_3 = \frac{f_s^{HF} - f_s^F}{(1 - \alpha) I^H} C.$$

When $\frac{f_p + f_s^F}{f_p + f_s^H} I^H > \frac{I^F}{T} > \frac{f_s^{HF} - f_s^H}{f_s^{HF} - f_s^F} I^H$, we have

$$0 < \Theta'_1 < \Theta'_2 < \Theta'_3.$$

Thus, the most productive firms sell in both the home and foreign markets, followed by those selling in the foreign market only, and then by those selling in the home market only.

Table 1, Comparison of three types of Chinese manufacturers

	Panel A: Whole Sample (1998-2005)		
	Domestic Sales Only (nonexporters)	Domestic Sales and Export (regular exporters)	Export Only (pure exporters)
Number of Observations	841,818	163,730	34,244
Share of Total Sample	80.96%	15.75%	3.29%
Logarithm of Employment	4.656	5.452	4.988
Logarithm of Output	9.363	10.316	9.553
Logarithm of Fixed Assets	8.158	8.925	7.487
TFP LP	3.747	4.121	3.834
	Panel B: Year 1998		
	Domestic Sales Only (nonexporters)	Domestic Sales and Export (regular exporters)	Export Only (pure exporters)
Number of Observations	92,377	16,500	3,369
Share of Total Sample	82.30%	14.70%	3.00%
Logarithm of Employment	4.833	5.839	5.083
Logarithm of Output	8.916	10.171	9.411
Logarithm of Fixed Assets	8.117	9.227	7.580
TFP LP	3.515	4.013	3.762
	Panel C: Year 2000		
	Domestic Sales Only (nonexporters)	Domestic Sales and Export (regular exporters)	Export Only (pure exporters)
Number of Observations	93,876	16,652	3,859
Share of Total Sample	82.07%	14.56%	3.37%
Logarithm of Employment	4.755	5.690	5.082

Logarithm of Output	9.149	10.342	9.521
Logarithm of Fixed Assets	8.157	9.153	7.550
TFP LP	3.649	4.119	3.829
	Panel D: Year 2002		
	Domestic Sales Only (nonexporters)	Domestic Sales and Export (regular exporters)	Export Only (pure exporters)
Number of Observations	104,454	19,865	5,248
Share of Total Sample	80.62%	15.33%	4.05%
Logarithm of Employment	4.642	5.439	4.991
Logarithm of Output	9.386	10.332	9.513
Logarithm of Fixed Assets	8.103	8.886	7.387
TFP LP	3.754	4.123	3.818
	Panel E: Year 2005		
	Domestic Sales Only (nonexporters)	Domestic Sales and Export (regular exporters)	Export Only (pure exporters)
Number of Observations	151,811	31,771	8,652
Share of Total Sample	78.97%	16.56%	4.50%
Logarithm of Employment	4.453	5.153	4.875
Logarithm of Output	9.670	10.377	9.647
Logarithm of Fixed Assets	8.073	8.692	7.508
TFP LP	3.899	4.158	3.863

Table 2, Correlations among different measures of TFP

	TFP LP	TFP OLS	TFP FE	TFP GMM
TFP LP	1.0000			
TFP OLS	0.6681	1.0000		
TFP FE	0.9118	0.9060	1.0000	
TFP GMM	0.9222	0.8907	0.9971	1.0000

Table 3, Main results

	1	2	3	4
Dependent Variable	TFP LP	TFP OLS	TFP FE	TFP GMM
Domestic Sales Only	-0.139** [0.004]	-0.014** [0.002]	-0.054** [0.003]	-0.044** [0.003]
Domestic Sales and Export	0.190** [0.004]	0.021** [0.002]	0.078** [0.003]	0.096** [0.003]
Controls				
Industry Dummy	X	X	X	X
Region Dummy	X	X	X	X
Year Dummy	X	X	X	X
Number of Observations	1,039,792	1,039,792	1,039,792	1,039,792
R-squared	0.2033	0.1159	0.1581	0.1636

Note: Standard errors, clustered at the firm-level, are reported in the bracket. ** represents statistical significance at the 1% level.

Table 4, Export market size and the existence of pure exporters

Dependent Variable Estimation Specification	1 Pure Exporter Firm-Year Level, OLS	2 Pure Exporter Firm-Year Level, OLS	3 Ratio of Pure Exporters Industry-Region-Year Level, Tobit	4 Ratio of Pure Exporters Industry-Region-Year Level, Tobit
Export Market Size (Lagged)	0.053*** [0.010]	0.054*** [0.010]	0.025*** [0.005]	0.024*** [0.005]
TFP LP (Lagged)		-0.042*** [0.014]		0.041*** [0.013]
Ln Wage (Lagged)		0.121*** [0.012]		0.009 [0.009]
Ln Labor (Lagged)		0.071*** [0.008]		0.009 [0.012]
Ln Capital Labor Ratio (Lagged)		-0.175*** [0.005]		-0.052*** [0.006]
Controls				
Industry Dummy	X	X	X	X
Region Dummy	X	X	X	X
Year Dummy	X	X	X	X
Number of Observations	340,647	339,417	36,617	36,568
R-squared	0.3731	0.3869	0.5400	0.5479

Note: Standard errors, clustered at the firm-level, are reported in the bracket in Columns 1-3; whereas standard errors clustered at the industry-region level are reported in the bracket in Columns 4-5. * and ** represent statistical significance at the 5% and 1% levels, respectively.

Table 5, Robustness checks

	1	2	3	4	5	6	7
Dependent Variable	TFP LP						
Estimation Specification	Excl. Outliers	Incl. Controls		Alternative Definition		1998-2001	2002-2005
Domestic Sales Only	-0.116** [0.004]	-0.014** [0.003]	-0.162** [0.004]	-0.114** [0.008]	-0.137** [0.004]	-0.172** [0.006]	-0.103** [0.004]
Domestic Sales and Export	0.167** [0.004]	0.098** [0.003]	0.159** [0.004]	0.176** [0.008]	0.207** [0.004]	0.202** [0.006]	0.192** [0.004]
Controls							
Industry Dummy	X	X	X	X	X	X	X
Region Dummy	X	X	X	X	X	X	X
Year Dummy	X	X	X		X	X	X
Firm Size		X					
Capital Intensity			X				
Skilled Labor Intensity				X			
Number of Observations	1,018,996	1,039,792	1,039,792	117,607	1,025,030	462,642	577,150
R-squared	0.1964	0.3533	0.2133	0.1856	0.2070	0.2191	0.1650

Note: Standard errors, clustered at the firm-level, are reported in the bracket. ** represents statistical significance at the 1% level.

Table 6, Alternative estimation specification

	8	9	10	11
Dependent Variable	Exporting Status			
Estimation Specification	Ordered Probit		Multinomial Logit	
TFP LP (Lagged)	0.518** [0.006]	0.184** [0.006]		
Relative Risk Ratio for Pure Exporters			1.392** [0.028]	1.080** [0.025]
Relative Risk Ratio for Regular Exporters			3.157** [0.039]	1.576** [0.020]
Cutoffs				
Cutoff 1	3.668** [0.063]	4.015** [0.062]		
Cutoff 2	4.940** [0.064]	5.340** [0.063]		
Controls				
Industry Dummy	X	X	X	X
Region Dummy	X	X	X	X
Year Dummy	X	X	X	X
Firm Characteristics		X		X
Number of Observations	631,052	628,393	631,052	628,393
Pseudo R2	0.2651	0.2951	0.2172	0.2597

Note: Standard errors, clustered at the firm-level, are reported in the bracket. ** represents statistical significance at the 1% level.

Table 7, Control for other explanations

	1	2	3
Dependent Variable	TFP LP		
Domestic Sales Only	-0.232** [0.007]	-0.174** [0.008]	-0.123*** [0.006]
Domestic Sales and Export	0.215** [0.008]	0.228** [0.008]	0.193*** [0.006]
Controls			
Industry Dummy	X	X	X
Region Dummy	X	X	X
Year Dummy	X	X	X
Export Processing Zones	X		
Processing Trader		X	
Export Market Size			X
Number of Observations	888,806	888,806	397,530
R-squared	0.2024	0.2031	0.2017

Note: Standard errors, clustered at the firm-level, are reported in the bracket. ** represents statistical significance at the 1% level.

Figure 1, Destructions of TFP

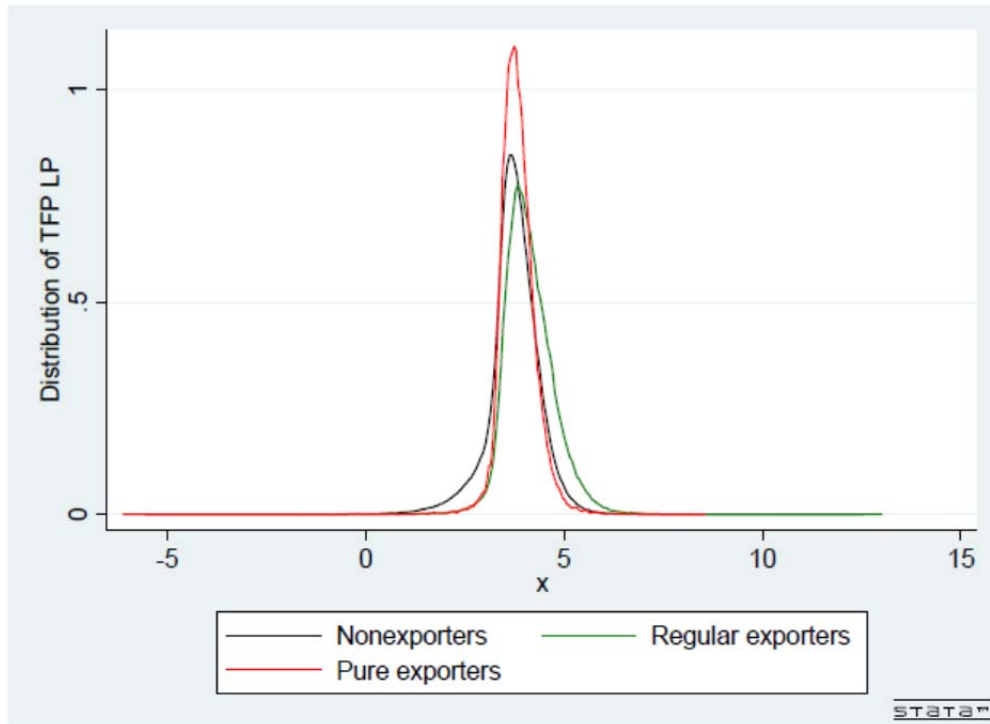


Figure 2, Optimal choice when foreign market is not sufficient large

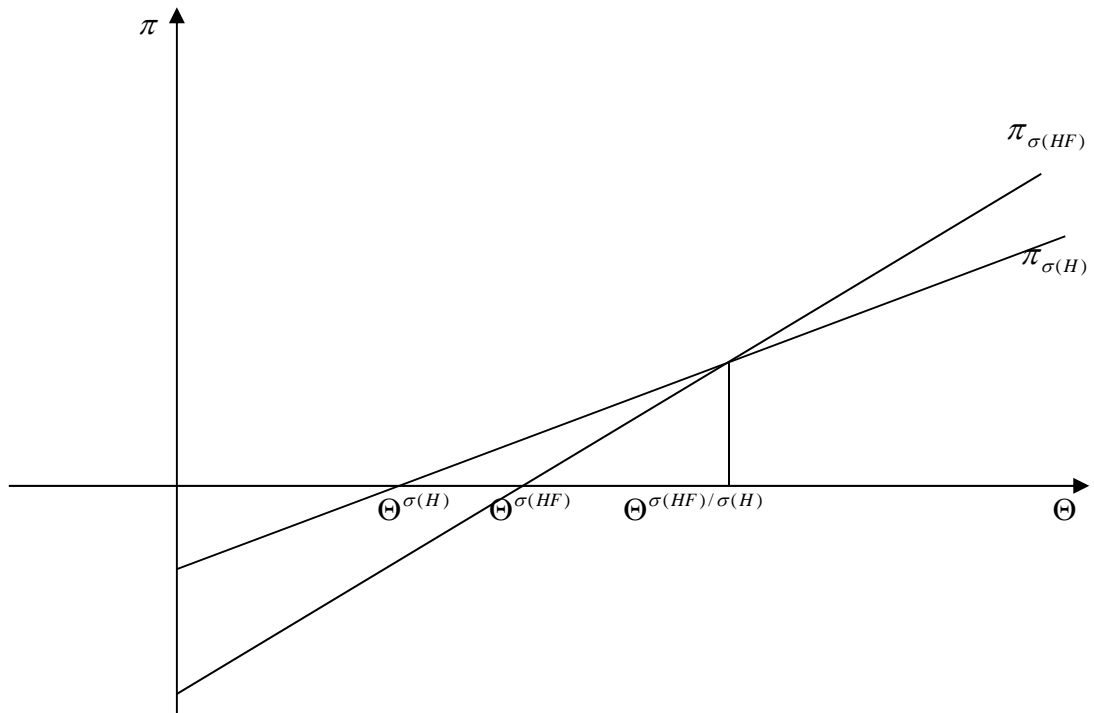


Figure 3, Optimal choice when foreign market is sufficient large

