Import competition, dynamic resource allocation and productivity dispersion: micro-level evidence from China

By Sai Ding [a], Wei Jiang [b] , Puyang Sun [c]

[a] Department of Economics, Adam Smith Business School, University of Glasgow, Gilbert Scott Building, Glasgow, UK, G12 8QQ . E-mail: sai.ding@glasgow.ac.uk .
[b] School of International Business, Southwestern University of Finance and Economics, Sichuan, P. R. China, 611130; Email: weijiang923@gmail.com .
[c] Department of International Economics and Trade, Nankai University, Tianjin, P. R. China, 300071. Email: puyangsun@nankai.edu.cn .

Abstract

Can trade liberalization shape sector dynamics by inducing reallocation of resources towards more efficient use? This paper explores whether and how import competition affects productivity dispersion within 425 4-digit manufacturing industries in China. Using a number of comprehensive micro-level datasets over the period of 2000-06, we find that import penetration reduces the productivity dispersion in general and the main channel is through the competition-induced dynamic resource allocation within industries by driving the least efficient firms out of the market. The trade-induced productivity truncation is more evident for imports under the ordinary trade regime and for final imported goods and intermediate goods imported by the same industries. The results are robust to various model specifications and estimation methods.

JEL Classification

F14; L1; D24; O12

1. Introduction

Cross-sectional dispersion in firm-level productivity is not only commonly observed even within narrowly defined industries, but also varies considerably across countries at different stages of economic development. For instance, Syverson (2004) discovers that within four-digit SIC industries in the US manufacturing sector, the average total factor productivity (TFP) ratio between an industry's 90th and 10th percentile plants is 1.92[1], indicating that the plant at the 90th percentile of the productivity distribution makes almost *twice* as much output with the same measured inputs as the 10th percentile plant. Hsieh and Klenow (2009) find even bigger within-industry productivity dispersion in two large developing countries, i.e. the corresponding figures of average 90-10 TFP ratios in China and India are over 5:1.

Productivity dispersion is economically relevant, to the extent that it reflects movements away from an optimal feasible resource allocation (Asker *et al*., 2014). The large and persistent productivity dispersion across firms may suggest certain market distortion that impedes the efficient allocation of resources, and therefore has significant economic and welfare implications. For instance, Hsieh and Klenow (2009) conduct a static cross-country comparison and claim that if manufacturing firms in China and India are able to achieve the same efficiency in allocating capital and labour across production units as does the US (by moving China and India to the US dispersion of marginal products), then the corresponding aggregate TFP gains can be up to 30-50% in China and 40-60% in India.

On the other hand, looking at China itself from a dynamic view, the picture is less pessimistic. Figure 1 plots the distribution of TFP of Chinese manufacturing sectors for the years of 1998, 2003 and 2007[2]. We can see that over time there is not only a trend of rising central tendency in Chinese industry's productivity distribution, but also a lower degree of dispersion, i.e. the left tail of TFP distribution in 1998 is far thicker than that in 2003 and 2007. The increase in the mean or median of the productivity distribution and the decrease in the dispersion of productivity (as indicated by the truncation from the lower end of the productivity distribution) imply that a significant restructuring process has been at work in the Chinese manufacturing sector. Thus, a proper understanding of productivity dispersion and its determinants, and the relevant policies directed at reducing distortions or reallocating

[1] In Syverson (2004), the average difference in the natural logarithm of TFP between an industry's 90th and 10th percentile plants is 0.651, which corresponds to a TFP ratio of $e^{0.651}$=1.92.

[2] The data covers all the above-threshold manufacturing firms in China over the period 1998-2007. See detailed data description in Section 4.1.

resources is of paramount importance for such a large and transition economy like China.

In this paper, we examine *whether* and *how* import competition affects the productivity dispersion within 425 4-digit Chinese manufacturing industries. Competition in output markets in general and trade-induced competition in particular are argued to be one of the strongest drivers that contribute to resource reallocation, and thereby affecting the within-industry productivity dispersion. This argument is well supported by economic theories in international trade. According to Melitz's (2003) model with heterogeneous firms, international trade can be viewed as a catalyst for inter-firm reallocations within an industry, i.e. the exposure to trade will induce the more productive firms to enter the export market and force the least productive firms to exit, so that the aggregate productivity increases due to selection and market share reallocation. Such mechanism works for imports as well. Melitz and Ottaviano (2008) find that in the short run import competition increases competition in the domestic product market, shifting up residual demand price elasticities for all firms at any given demand level, and thereby forcing the least productive firms to exit[3]. Surprisingly, empirical evidence in this literature is rather limited. Our research tends to fill this important gap and examines the distributional effects of trade openness on aggregate productivity, i.e. whether and how import competition helps to drive the least efficient firms out of the market, to induce dynamic resource allocation towards more productive firms, and therefore to reduce productivity dispersion within industry? To the best of our knowledge, this is the first study in the trade literature addressing the effect of import competition on within-industry productivity dispersion in China.

There are at least three other important contributions of the paper. First, we aim to explicitly address the causality issue between import competition and productivity dispersion by using both the lagged level and the reductions in China's import tariffs, and the lagged level of industry average of tariffs as instruments. Compared with the existing literature (see for instance, Syverson, 2004; Balasubramanian and Sivadasan, 2009) which assumes the exogeneity when modelling dispersion, our instrumental variable (IV) approach enables us to isolate the effect of import competition from other channels, and to tackle the potential problem of reverse causality running productivity dispersion to import penetration at the industry level.

[3] In the original Melitz (2003) model, import competition does not play a role in the reallocation process due to the CES specification for demand, i.e. residual demand price elasticities are exogenously fixed and unaffected by import competition. It is also worth noting that the short-run welfare gain from import competition in Melitz and Ottaviano (2008) may be overturned in the long run due to shifts in the pattern of entry.

Second, when modelling the impact of import competition on productivity dispersion, we capture the factors from both the supply side (such as the fixed costs and sunk costs of production), and the demand side (such as product substitutability). Besides, we take into account a number of China-specific factors, which play an important role in shaping the productivity distribution of Chinese industries (such as firm ownership, market structure and government subsidy). We are not aware of any existing study which attempts to capture all these three dimensions of factors when empirically modelling productivity dispersion.

Third, although it is known in the literature that the Melitz-type mechanism would lead to reduction in standard deviation of productivity in general, the contribution of our work lies in identifying some of the potential mechanisms and channels through which imports affect productivity dispersion in China. For instance, we investigate the heterogeneous effects of imported goods according to their nature (final versus intermediate goods) and to the trade regime (ordinary versus processing trade). We also directly test whether the import-induced reduction in productivity dispersion is due to truncation from the bottom (by driving the least efficient firms out of the market) or convergence (by spurring the least efficient firms to improve their productivity). None of the existing work tends to explore the productivity implication of Chinese imports in such a systemic and comprehensive way.

The structure of the paper is as follows. Section 2 briefly reviews the relevant literature. Section 3 explains our empirical methodology. Section 4 discusses the data and sample and presents some basic summary statistics. Section 5 interprets the results of our baseline model and of various heterogeneous effects and channels in order to shed light on whether and how imports affect productivity dispersion. Section 6 conducts further tests to examine the robustness of our results. Section 7 concludes the paper.

2. Related literature

We draw significantly on three strands of literature on productivity, i.e. the effect of trade liberalization on productivity, the productivity dispersion literature, and the effect of resource misallocation on aggregate productivity.

First, there is a large literature showing that trade liberalization increases firm- and industry-level productivity (for instance, Melitz, 2003; Amiti and Konings, 2007; Fernandes, 2007; Topalova and Kandelwal, 2011; Kasahara and Lapham, 2013; Yu, 2015). For instance, Kasahara and Lapham (2013) extend Melitz (2003)'s model to incorporate imports and claim that there are substantial gains in aggregate productivity and welfare due to trade. Based on the Chilean plant-level data, their structural estimation highlights the role of importing intermediates for use in production in explaining differences in plant performance. Some recent empirical research examines the productivity gains from removing trade barriers and protections. Using Indonesian data, Amiti and Konings (2007) argue that reducing output tariffs can produce productivity gains by inducing tougher import competition, whereas cheaper imported inputs can raise productivity through learning, variety, and quality effects. Similar results are found for India and China by Topalova and Kandelwal (2011) and Yu (2015) respectively.

Second, the fact that firms differ in performance or productivity ignites another interesting literature on productivity dispersion which focuses on the entire distribution of sector productivity and within-industry firm dynamics. Syverson (2004) examines the effect of a demand side factor, product substitutability, on productivity dispersion using US industry-level data. He argues that imperfect product substitutability impedes resource reallocation so that low-substitutability industries exhibit high productivity dispersion. Balasubramanian and Sivadasan (2009) focus on the effect of sunk costs on productivity dispersion, and find that increases in capital resalability are associated with a reduction in productivity dispersion.

Third, there is a fast-growing literature which links the micro-level resource misallocation to aggregate productivity. Most works argue that the low aggregate TFP is a result of firm-/plant-level resource misallocation especially in developing countries, i.e. the most efficient firms fail to attract the large share of productive resources that efficiency would dictate (see, Olley and Pakes, 1996; Banerjee and Duflo, 2005; Foster *et al*., 2008; Heish and Klenow, 2009; Asker *et al*., 2014; Midrigan and Xu, 2014). Therefore reallocation of labour and capital across manufacturing firms is a key resource of productivity growth. In the case of China, apart from the seminal work by Heish and Klenow (2009) discussed earlier, Song *et al*. (2011) regard the initial misallocation as a pre-condition for China's sustained growth because efficient firms can count on a highly elastic supply of factors attracted from the less productive firms. Khandelwal *et al*. (2013) empirically link trade and productivity through the resource reallocation channel and claim that trade barriers such as tariffs and quotas can distort resource allocation along the intensive and extensive margins. Focusing on Chinese textile and clothing exports, they find that quota removal coincides with substantial reallocation of export activity from incumbents to entrants, as well as a productivity gain by 28%. Our paper is along these lines, but with some important new contributions as highlighted in the Section 1.

3. Empirical methodology

3.1 Measures of TFP and productivity dispersion

We construct the measure of firm-level TFP using the semi-parametric Olley and Pakes (1996) approach which alleviates both the selection bias and simultaneity bias (between input choices and productivity shocks). Another advantage of Olley-Pakes method is the flexible characterization of productivity, only assuming that it evolves according to a Markov process (Van Biesebroeck, 2007). Thus, assuming a Cobb-Douglas production function, the production function is

$$y_{ijt} = \beta_0 + \beta_k k_{ijt} + \beta_l l_{ijt} + w_{ijt} + \varepsilon_{ijt} \tag{1}$$

where y_{ijt} is the natural logarithm of value added of firm *i* in industry *j* at time *t*, defined as sales minus intermediate inputs plus value added tax; k_{ijt} is the natural logarithm of firm's capital input, which is computed using the perpetual inventory method following Brandt *et al*. (2012); l_{ijt} is the natural logarithm of firm's labour input as measured by total employment; w_{ijt} represents a productivity difference known to the firm, but unobservable to us; and ε_{ijt} is either measurement error or a shock to productivity which is not forecastable during the period in which labour can be adjusted.

Our approach is based on the recent development in the application of the Olley-Pakes method (for instance, Amiti and Konings, 2007; Brandt *et al*., 2012; Feenstra *et al*., 2014). First, we use different price deflators for inputs, outputs and investment. It is known in the productivity literature that ideally one would use firm-specific price deflators when constructing TFP, otherwise the 'omitted price bias' may occur (see Van Beveren, 2012). Since such information is not available in the data, we use different industry-specific price deflators for inputs, outputs and investment, which are directly drawn from Brandt *et al*. (2012). This implies that our TFP measure is a revenue-based productivity measure (TFPR) as introduced by Foster *et al*. (2008), which may capture both technical efficiency and price-cost markups. Second, we use the perpetual inventory method to compute the real investment variable, where the depreciation rate of physical capital is based on firms' reported actual depreciation figure rather than arbitrary assumptions. Appendix Table A1 reports the estimated coefficients of the production function and the associated log of TFP by industry. The estimated output and labour elasticities with respect to output are all positive and significant, but the magnitude exhibits significant heterogeneity among industries. The estimated TFP varies across industries too, with a 11 times difference between the most efficient sector (petroleum processing and coking) and the least efficient one (tobacco processing). We also find that most industries (19 out of 29) have exhibited decreasing returns to scale, which is consistent with the common findings in the literature (see, Brandt *et al*., 2012; Yu, 2015).

Having obtained the firm-level TFP, we compute our measures of productivity dispersion. The primary productivity dispersion measure is the interquartile range (IQ range), i.e. the interquartile productivity difference divided by the industry's median productivity level. One key advantage of IQ range is that it is less sensitive to the outliers. Alternative measures such as standard deviations in TFP (scaled by industry mean productivity), the difference between TFP at the 90^{th} and 10^{th} percentile of the distribution, and at the 95^{th} and 5^{th} percentile of the distribution are also computed. To save space, we only report the results based on IQ range and standard deviations, and other results (which are quite similar) are available upon request.

3.2 Baseline model specification and hypotheses

The sources of productivity dispersion lie in both the supply-side-production factors such as technology shocks, management skill, R&D or investment patterns, and the demand-side conditions such as product differentiation and substitutability. In particular, Syverson (2004) argues that when consumers can easily switch between producers, inefficient (high costs) producers cannot operate profitably. Hence, an increase in product substitutability raises the cutoff productivity level, thus lowering productivity dispersion. We follow this line of thinking and take into account both the demand- and supply-side factors affecting productivity dispersion in China. Besides, as a large developing country, China has a number of institutional features which need to be captured when examining the impact of import competition on productivity dispersion of Chinese industries.

Our baseline model is specified as follows:

$$Dispersion_{jpt} = \alpha_0 + \alpha_1 IMP_{jpt} + \theta X_{jpt} + \eta_t + \zeta_p + \xi_j + \mu_{jpt} \quad (2)$$

where the dependent variable is the productivity dispersion measure of industry j and province p at year t, which is defined by either the IQ range or standard deviations of TFP; IMP_{jpt} is the import penetration ratio which is defined as follows:

$$IMP_{jpt} = \frac{Import_{jpt}}{Import_{jpt}+Output_{jpt}-Export_{jpt}} \quad (3)$$

where $Import_{jpt}$, $Export_{jpt}$ and $Output_{jpt}$ are total imports, exports and outputs of industry j and province p at year t. Import penetration ratio is argued to be a better proxy for trade liberalization than tariffs, as the latter does not take into account any non-tariff barriers of trade (Levinsohn, 1993). It is also worth noting that the province dimension is included in both the dependent variable and the key independent variable in order to reflect the geographic features that affect productivity distribution and international trade among industries in China[4]. This is consistent with the literature that local protectionism is prevalent in China which impedes the free flow of goods and services across provinces (Bai *et al.*, 2004). The relationship between trade liberalization and within-industry productivity dispersion is well predicted by Melitz (2003) and Melitz and Ottaviano (2008), i.e. the benefits of exposure to foreign competition/markets enjoyed by the more productive domestic firms should drive the least efficient domestic producers out of business, thereby decreasing productivity dispersion. We therefore expect α_1 to be significant and negative in equation (2).

X_{jpt} consists of three groups of control variables, i.e. the demand-side factors, supply-side factors, and China-specific factors. On the demand-side factors, following Syverson (2004), we use a vector of measurable proxies for substitution elasticities among the outputs of industry producers. The first measure, $VALUELB_j$, represents a geographic barrier to substitution, which is the natural logarithm of the weighted sum of the dollar-value-to-weight ratios of all product classes in a given 4-digit industry j, where the weights are the product classes' shares of the total industry tonnage shipped[5]. Geographic barriers to substitution arise when transport costs hinder producers from practically selling their output beyond certain distances. Therefore goods valuable in relation to their weight are more economical to ship. Industries with high values of $VALUELB_j$ are expected to have less geographically segmented output market and greater substitutability. We therefore expect a significantly negative relationship between $VALUELB_j$ and productivity dispersion.

The other substitutability measure is advertising intensity (ADV_j), which is defined as

[4] Our data contains the detailed information on the geographic location of trading firms, which is used to define the province of trade flows.

[5] The transport data is from the US Bureau of the Census. We convert the SIC industry codes to corresponding GB (2002) industry level when merging it to the Chinese dataset.

total advertising expenditure in an industry j divided by total revenue[6]. The effect of branding and advertising on product substitutability is argued to be ambiguous. On the one hand, advertising may create artificial product differentiation so that industries with higher advertising intensities exhibit more product differentiation and less product substitutability; on the other hand, advertising is argued to be informative and serves to educate consumers about superior product, which allows more productive firms to take market share away from less efficient competitors. Hence we keep an open view on the coefficient of advertising intensity in the productivity dispersion equation.

We employ two variables to capture the supply-side factors, i.e. fixed operating costs, and sunk entry costs, both of which are expected to affect the critical productivity cutoff level and therefore the industry-level productivity dispersion. First, following Syverson (2004), we define the industry fixed cost index ($Fixed\ Cost_j$) as the share of nonproduction workers in total employment in each Chinese industry j[7]. This measure is to proxy for the amount of overhead labour required by the industry technology and therefore the relative size of production-related fixed costs. It is argued that higher fixed costs make it difficult for inefficient firms to be profitable, leading them to exit in equilibrium. Thus we expect a significantly negative relationship between fixed costs and productivity dispersion at the industry level.

Second, we adopt the method of Balasubramanian and Sivadasan (2009) to measure sunk entry costs ($Sunk\ Cost_j$), which is a capital resalability index defined as the share of used capital investment in total capital investment at each 4-digit industry j[8]. This measure of capital resalability is to capture recoverability of investments, which is an inverse proxy for the extent of sunkenness of capital investments. Compared with the standard method of Sutton (1991), where investments in physical capital (usually in the median plant size) are used to proxy sunk costs, the capital resalability index better accords with the theoretical definition of sunk costs where the resale value of investment should be strictly excluded. According to Hopenhayn (1992), sunk costs act as a barrier to entry and exit, and protect incumbent firms. Thus, an increase in sunk costs (as reflected by a decrease in capital resalability) leads to a reduction in the cutoff productivity, implying an increase in the productivity dispersion.

[6] The data is from Compustat, a database that has financial statement data on all listed US firms. We convert the 3-digit SIC industry codes to corresponding GB (2002) industry level when merging it to the Chinese dataset.

[7] Data come from various issues of China statistical yearbook.

[8] The used capital expenditure data is from the US Bureau of the Census. We convert the SIC industry codes to corresponding GB (2002) industry level when merging it to the Chinese dataset.

We also include a number of China-specific factors which may affect industry-level productivity dispersion in the Chinese context. First, we include two ownership variables, SOE_{jpt} and FIE_{jpt}, which are defined as the share of state-owned capital and foreign capital in total capital in 4-digit industry j in province p at year t respectively. It is widely believed that despite decades of economic reform, state-owned enterprises (SOEs) remain the least efficient sector in the economy with an average return on capital well below that in the private sector (Dougherty and Herd, 2005; Ding *et al*., 2012). On the other hand, foreign ownership is associated with both higher levels of TFP and fewer financial constraints (Manova *et al*., 2015). We hypothesize that both SOE_{jpt} and FIE_{jpt} may increase productivity dispersion in the industry but from two different directions, i.e. the state ownership hinders the exit of least efficient firms therefore increasing the dispersion from the lower end of the distribution, whereas foreign ownership increases the top end of the productivity distribution and enlarges the dispersion from the right.

Second, we construct the Herfindahl–Hirschman Index (HHI_{jpt}) to capture the market structure or competition status in 4-digit industry j in province p at year t, where a lower HHI indicates higher degree of competition in the industry. The threat from competitors both intra- and inter-industry will affect resource allocation and then productivity dispersion (Syverson, 2011). Tougher domestic competition is argued to lower productivity dispersion, i.e. inefficient firms are hard to survive in a very competitive market.

Third, government subsidy may affect the entry and exit of firms in the market, and therefore influence productivity dispersion in industries. Our subsidy measure ($Subsidy_{jpt}$) is defined as the ratio of subsidy to the value added of firms in 4-digit industry j in province p at year t. We expect a positive relationship between subsidy and productivity dispersion, as the government subsidy may keep the least efficient producers viable.

Lastly, the error term in equation (2) comprises four components: (*i*) the time-specific fixed effect, η_t, accounting for possible business cycles and macroeconomic shocks such as an appreciation of the Chinese yuan; (*ii*) the province-specific fixed effect, ζ_p, which captures geographic factors that influence productivity such as transportation costs, financial market development, tax treatment and so on; (*iii*) the industry-specific fixed effect, ξ_j, reflecting time-invariant industrial features affecting productivity such as factor costs and factor intensities; and (*iv*) an idiosyncratic error term, μ_{jpt}, controlling for other unspecified factors.

3.3 Estimation methods

The endogeneity problem is argued to be less serious when modelling productivity dispersion as firms do not observe the industry-level distribution information when making decisions. Nevertheless, it is still possible that the industry-province pairs with lower allocative efficiency or higher productivity dispersion may experience larger increases in import penetration due to the presence of lower-productivity domestic producers. Alternatively changes in demand conditions may influence both import penetration ratio and the productivity dispersion. We therefore adopt the IV approach in order to shed light on the causality between import competition and productivity dispersion in China.

Three sets of instruments are used in our analysis. First, inspired by Yu (2015), we use the one-year lag of product-level output tariffs obtained from WTO as instruments for the import penetration ratio. The rationale is that past tariffs may be highly correlated with current tariffs and therefore affecting industries' import penetration level, however they do not have any direct impact on the industries' productivity dispersion. Second, the changes in tariff rates are used as another set of instruments since China was required to lower tariff rates systematically when it joined the WTO and the reductions were roughly proportional across products. Lastly, we use the one-year lag of similar industry tariffs, defined as the average of the rest of the 4-digit industry tariffs in the same 2-digit industry[9], as the third group of instruments. The justification is that tariffs among similar industries are highly correlated, but lagged tariffs of other industries do not affect the productivity dispersion of the industry of interest. This method is inspired by Cassiman and Veugelers (2002) and Lin *et al*. (2010) where industry average is used to instrument endogenous firm-specific variables. A number of diagnostic tests are then conducted to verify the quality of the three sets of instruments.

4. Data and summary statistics

4.1 Data and sample

Our industry-level study is based on the aggregation of information from a number of comprehensive microeconomic datasets, including the firm-level production data drawn from

[9] For instance, one 2-digit industry '13' has three 4-digit industries ('1311', '1312' and '1313'). Then the instrument for 4-digit industry '1311' is the lagged average of tariffs of the rest 4-digit industries '1312' and '1313'.

the annual survey of Chinese industrial firms by National Bureau of Statistics (NBS), the transaction-level trade data from Chinese General Administration of Customs (GAC), the product-level tariff information published by World Trade Organization (WTO), and a number of US datasets (such as Compustat and US Bureau of the Census).

The first firm-level dataset is drawn from the annual accounting reports filed by industrial firms with the NBS over the period of 1998-2007. This dataset includes all SOEs and other types of enterprises with annual sales of five million yuan (about $817,000) or more. These firms operate in the manufacturing sectors[10] and are located in all 30 Chinese provinces or province-equivalent municipal cities[11]. Following the literature, we drop observations with negative total assets minus total fixed assets, negative total assets minus liquid assets, and negative sales, as well as negative accumulated depreciation minus current depreciation. Firms with less than eight employees are also excluded as they fall under a different legal regime (see, Brandt *et al*., 2012). Lastly, to isolate our results from potential outliers, we exclude observations in the one percent tails of each of the regression variables.

The second database from the Chinese Customs contains detailed transaction-level information of all imports and exports in China during the period of 2000-06, which includes 243 trading partners and 7526 different products in the 8-digit Harmonized System (HS). A feature of this dataset is its rich information on trade transactions. For instance, for each transaction it reports the transaction date, 8-digit HS product code, trade volume, trading partner, unit price, shipment method, trade regime and so on. To ensure the accuracy of the estimates, we eliminate the trading firms which do not engage in manufacturing but act as intermediaries between domestic producers/suppliers and foreign trade partners (see, Ahn *et al*., 2011; Manova and Zhang, 2012; and Yu, 2015). Merging the two datasets above gives us a final sample for the period of 2000-06, which has a good representativeness of exporting sector in China but is skewed toward large manufacturing firms[12].

Our tariff data is from WTO, which provides product-level tariffs at the 6-digit HS level of all WTO member countries/regions. Following Yu (2015), we use the average ad valorem (AV) duty in our empirical regression[13]. Lastly, when computing our measures of product substitutability, we use the US data for 3-digit SIC sectors from Syverson (2004) and

[10] We exclude utilities and mining sectors for our research purpose in this paper.
[11] Our dataset does not contain any firm in Tibet.
[12] The detailed merging technique and outcome of the first two datasets are available in Appendix 2.
[13] China’s tariffs from 1998 to 2000 are missing from WTO, but we manage to get the 2000 tariff data from the Chinese Customs. We thank one referee for raising this point and Professor Miaojie Yu for sharing the data.

then match them to our GB (2002) industry level. Similarly, our measure of sunk costs is from US Bureau of the Census as in Balasubramanian and Sivadasan (2009). One benefit of using the US industry information is their strict exogeneity in our regressions.

4.2 Summary statistics

Table 1 provides the summary statistics of key variables in the baseline model. The average productivity dispersion measure based on the IQ range is 0.591 and the corresponding figure based on the standard deviation is 0.452. The import penetration ratio is averaged at 0.104 among all industries during the sample period. The two demand-side factors ($VALUELB$ and ADV) and two supply-side factors ($Fixed\ Cost$ and $Sunk\ Cost$) are industry-specific and time-invariant variables so that the sample size is 425 4-digit industries in China. The proportion of state-ownership (SOE) and foreign-ownership (FIE) is 17.0% and 11.9% respectively in the sample. The market structure measure (HHI) is averaged at 19.4% among Chinese industries. Lastly, government subsidy ($Subsidy$) is averged about 0.4% of value added of firms.

We provide further summary statistics at the industry level. Appendix Table A3(a) presents the productivity dispersion of 2-digit Chinese industries, where the dispersion measure is based on the IQ range[14]. There is significant cross-sectional heterogeneity of productivity dispersion among 2-digit industries. For instance, some oligopolistic sectors such as tobacco processing (1.159 in 2000) have much higher dispersion than the more competitive sectors such as textile (0.529 in 2000). In terms of time dynamics, we find that the productivity dispersion shows a decreasing trend for most industries over the sample period of 2000-06, indicating that the reallocation process plays a substantial role in the data[15]. Appendix Table A3(b) reports the import penetration ratio in 2-digit Chinese industrial sectors during the period of 2000-06. It is interesting to see that there exists vast heterogeneity among industries, where import penetration shows a rising trend in some industries (such as electronic machinery) and a decreasing trend in others (such as textile).

In order to have a general idea on the relationship between import penetration and productivity dispersion, we aggregate the data and plot the relationship of these two variables during the period of 2000-06 in Figure 2. We find that the aggregate within-industry

---

[14] To save space, the productivity dispersion based on the standard deviation is not reported but available upon request.

[15] Exception holds for four industries of leather, educational goods, petroleum processing and other manufacturing where productivity dispersion displays no significant change or a non-linear trend.

productivity dispersion in China decreases over time, whereas import penetration rises steadily during the sample period. Thus, an interesting research question arises: whether and how imports contribute to the reduction of productivity dispersion in China?

5. Empirical results

5.1 Baseline results

Table 2 presents the results of our baseline model of equation (2) in order to shed light on *whether* import competition has an effect on industry-level productivity dispersion. The first stage IV results show that all three sets of instruments (lagged tariffs, first-differenced tariffs and lagged similar industry tariffs) have a significant and negative effect on the import penetration ratio. The second stage results confirm the exogenous role of imports in reducing within-industry productivity dispersion, i.e. a one percentage point increase in import penetration reduces the productivity dispersion by 0.3 percentage points (as measured by the IQ range, Column 4) or 0.2 percentage points (as measured by standard deviation, Column 8). This is consistent with the theoretical prediction that increased competition from trade could result in the truncation of within-industry productivity dispersion by inducing the dynamic resource allocation towards more efficient firms (see, Bernard *et al*., 2003; Melitz, 2003; Melitz and Ottaviano, 2008). Our new results for China also correct the puzzle in Syverson (2004) where the effect of trade openness on productivity dispersion is absent.

In terms of the measures of product substitutability, the coefficient of the dollar-value-to-weight ratio ($VALUELB$) is significant and negative, which is in line with the theoretical hypothesis that higher geographic barrier to substitution reduces the cutoff productivity level, and thus increasing productivity dispersion. On the other hand, the advertising intensity (ADV) has a significant and positive effect on productivity dispersion, indicating that greater artificial product differentiation reduces product substitutability and increases productivity dispersion.

Fixed cost ($Fixed\ Cost$) is found to reduce productivity dispersion and to improve resource allocation, supporting the theoretical prediction that higher fixed costs can help to drive the inefficient firms out of the market, thus contributing to the reduction of productivity dispersion. The coefficient of sunk cost ($Sunk\ Cost$) is also negative and significant. This is because the capital resalability index is an inverse proxy for the extent of sunkenness of capital investments. Sunk costs can impede competitive forces and prevent the attainment of both technical efficiency and allocative efficiency, as they make the act of exit costly and affect the discipline on incumbents. Our result confirms this argument.

The results of all China-specific variables are in line with our expectation, where both state- and foreign-ownership (SOE and FIE) are found to have positive and significant effect on dispersion. We find that a more competitive market (lower HHI) is associated with better resource allocation and lower productivity dispersion. And the effect of government subsidy ($Subsidy$) on productivity dispersion appears to be positive and significant, indicating the adverse effect of government intervention on resource allocation and productivity distribution within industries.

A number of diagnostic tests are conducted in order to verify the quality of the three sets of instruments. First, we use the under-identification test based on the Kleibergen-Paap rk LM statistics to check whether the excluded instruments are correlated with the endogeneous regressors. As shown in Table 2, the null hypothesis that the model is under-identified is rejected at the 1 percent significance level. Second, the weak-identification test based on the Cragg-Donald Wald F statistics provide strong evidence for rejecting the null hypothesis that the first stage regression is weakly identified at the 1 percent significance level. Third, we conduct the Durbin–Wu–Hausman test for endogeneity and our results reject the null hypothesis that the endogenous regressors in the model are in fact exogenous at either the 5 or 10 percent significant level, which justifies the use of IV approach for the estimation. Thus, our baseline model proves that import competition is an important means to reduce productivity dispersion within industries, and all control variables including supply-side, demand-side and China-specific factors are important to determine the productivity dispersion in the way suggested by various theories discussed in Section 3.2.

5.2 Heterogeneous effect: the role of trade regime

We are interested in exploring some heterogeneous effects on *how* import competition affects productivity dispersion in China. One feature of China's trade pattern is the sheer magnitude of processing trade[16]. According to Yu (2015), processing imports account for more than 50 percent of China's total imports. There is a rising literature on the effect of different trade regimes on firm performance in China, which indicates that generally speaking, firms conducting processing trade have inferior performance than their counterparts who are

[16] Processing trade is officially defined as business activities in which the operating enterprise imports all or part of the raw or ancillary materials, spare parts, components, and packaging materials, and re-exports finished products after processing or assembling these materials/parts (Manova and Yu, 2012).

engaged in ordinary trade business (see, Jarreau and Poncet, 2012; Wang and Yu, 2012; Yu, 2015; Manova and Yu, 2016). Following this line of thinking, we distinguish the heterogeneous effect of imports in various trade regimes and expect that the role of imports in reducing productivity dispersion in domestic industries is only evident for ordinary trade but not for processing trade.

We use the trade regime information from the Customs dataset to classify all Chinese imports into two categories, i.e. ordinary- versus processing-trade imports. Then we compute the corresponding import penetration ratio for ordinary-trade imports ($OT-IMP$) and that for processing-trade imports ($PT-IMP$). The summary statistics (Appendix 6) indicates that the processing-trade import penetration ratio (0.046) is more than 4 times than the ordinary-trade import penetration ratio (0.010), indicating the important role of processing trade in Chinese imports. The econometric results in Table 3 Panel A show that imports under the ordinary-trade regime have a significantly negative effect on productivity dispersion when both measures of productivity dispersion are used (Columns 2 and 5). However, the results on processing-trade imports are not statistically significant (Columns 1 and 4). Including both types of import penetration ratio into the regression simultaneously, we find that although both variables are significantly negative, the magnitude is much bigger for ordinary-trade imports than for processing-trade imports for both measures of dispersion[17](Columns 3 and 6). The robust effect of ordinary-trade imports in our productivity dispersion regressions implies that only imports aiming for domestic markets are conductive to the reduction of productivity dispersion in domestic industries, whereas those imports aiming for exporting after local processing have no or very little resource reallocation effect on the domestic industries.

5.3 Heterogeneous effect: the nature of imported products

The effect of imports on productivity dispersion may depend on the nature of imported goods. On the one hand, imports of final goods[18] lead to tougher competition in the domestic market, which forces firms to increase their efficiency, drives the least efficient domestic producers out of market and thereby reducing the productivity dispersion. This is often referred to as the *pro-competitive effect* of trade liberalization (see, for instance, Topalova and Khandelwal, 2011; De Loecker and Goldberg, 2013). On the other hand, openness to foreign

[17] The difference is statistically significant at the 5% significance level.
[18] Final goods are the goods that are ultimately consumed by the consumers rather than used in the production of another good.

supply markets increases the availability of intermediate goods[19] that may be cheaper or with a higher quality and technological content than domestic products (Maggioni, 2013; Halpern *et al*., 2015). In other words, trade liberalization brings in more and cheaper imported inputs, which can raise domestic firms' productivity via learning, variety, and quality effect. This is referred to as the *input effect* which drives the productivity gains of individual firms, but not necessarily contributes to resource reallocation or reduction of productivity dispersion within industries.

In order to test for such hypotheses, we firstly make the distinction between the final goods and intermediate goods by judging whether the imported goods are purchased by domestic manufacturing firms. Since manufacturing firms seldom engage in the retail business in China, the goods imported by manufacturing firms are intermediate goods for the production of final goods. On the other hand, we treat imports which are not purchased by manufacturing firms as final goods, which are directly purchased and consumed by consumers[20]. Then among the imported intermediate goods, we further distinguish between the intermediate goods which are imported by upstream industries[21] and those imported by the same industries. The rationale for this further classification is that the same import flow of intermediate goods may represent a competitive threat for firms operating in that sector, but an opportunity for the downstream firms. Lastly, we compute the import penetration ratio for the final imported goods ($FIN-IMP$), for the intermediate goods imported by the upstream industries ($UP-IMP$), and for the intermediate goods imported by the same industries ($SAME-IMP$). We expect that both the final imported goods and the intermediated goods imported by the same industries may reveal the pro-competitive effect, whereas the input effect can be dominant for the intermediate goods imported by upstream industries.

Table 3 Panel B reports the results. We find the effect of import competition on productivity dispersion is significant and negative for both the final imported goods and the intermediate goods imported by the same industries, and the magnitude is greater for the latter. On the other hand, the effect becomes significantly positive for the intermediate goods imported by the upstream industries. This is in line with our expectation that the pro-competitive effect of imports is well captured by the final imported goods, where the tough

---

[19] Intermediate goods are goods that are used as inputs in the production of final goods, such as partly finished goods.

[20] Note that we exclude the trading firms that do not engage in manufacturing but act as intermediaries between domestic producers/suppliers and foreign trade partners in the Customs dataset. This is because it is not clear whether the imports purchased by them are for the production of final goods or for consumers.

[21] Upstream industries are those processing the basic or raw material into intermediary products which are converted into finished products by the downstream industries.

foreign competition improves resource allocation and reduces productivity dispersion in domestic industries. The competition effect is also found for intermediate goods imported by the same industries, as such goods can compete with products in the same industry and generate learning effects from the foreign technology embodied in the imported intermediate inputs, which helps to reduce productivity dispersion. No competition effect is found for the intermediate goods imported by upstream industries, where international integration offers domestic downstream firms the opportunity to exploit an increase variety of intermediates with cheaper price or higher quality than domestic ones. Such input effect is not conducive to better resource allocation or reducing productivity dispersion.

5.4 Channels: firm exit or convergence?

In order to further shed light on how the restructuring progress has been at work in the Chinese manufacturing sector following the growing inflows of foreign imported goods, we directly test the following two hypotheses. First, is the import-induced reduction in productivity dispersion due to the truncation from the bottom by driving the least efficient firms out of the market and thus facilitating the dynamic resource allocation towards more productive firms? Second, is the effect due to convergence by spurring the least efficient firms to improve their productivity? Both mechanisms could potentially lead to the reduction of within-industry productivity dispersion but the way of achieving the aggregate trade-driven productivity gains is different, i.e. the former can be viewed as the *between-firm channel* whereas the latter can be viewed as the *within-firm channel*.

We first estimate an IV probit model to examine the factors that determine the probability of firm exit as follows.

$$Prob(exit_{ijt+1}) = \beta_0 + \beta_1 TFP_{ijt} + \beta_2 IMP_{jpt} + \beta_3 TFP_{ijt} * IMP_{jpt} + \theta X_{ijt} + \mu_{ijt} \qquad (4)$$

where the dependant variable, $exit_{ijt}$, is a dummy variable equal to one if firm i in industry *j* exits at year *t+1*; TFP_{ijt} and IMP_{jpt} are the firm-level TFP and industry-level import penetration ratio as defined in Sections 3.1 and 3.2; and X_{ijt} include some firm-specific features such as firm age (Age_{ijt})[22] and firm size ($Size_{ijt}$)[23], and some industry-specific features such as ownerships (SOE_{jpt} and FIE_{jpt}) and market structure (HHI_{jpt}) as defined in Section 3.2. The instruments used for the import penetration ratio are the same as those

[22] Firm age is defined as the difference between the current year t and the opening year of the firm i.
[23] Firm size is the natural logarithm of total assets of firm i at year t.

defined in Section 3.3.

To test for the second hypothesis, we estimate the following model on firm-level TFP growth.

$$TFPG_{ijt} = \gamma_0 + \gamma_1 TFP_{ijt-1} + \gamma_2 IMP_{jpt} + \gamma_3 IMP_{jpt} * TFP_{ijt-1} + \theta X_{ijt} + \eta_t + \xi_i + \zeta_j + \mu_{ijt} \tag{5}$$

where the dependant variable, $TFPG_{ijt}$, is the TFP growth rate of firm i in industry j at year t; $TFP_{ijt-1}$is the lagged TFP level of firm i to capture the convergence or catch up effect; IMP_{jpt} and X_{ijt} are exactly the same as those employed in equation (4). The error term in equation (5) comprises the time-specific fixed effect (η_t), the firm-specific fixed effect (ξ_i), the industry-specific fixed effect (ζ_j) and an idiosyncratic error term (μ_{ijt}). The two-stage IV approach is used to address the potential endogeneity of import penetration, using the same sets of instruments discussed in Section 3.3.

Table 4 reports the results. In columns (1)-(3)[24], we find that the effect of TFP on firm exit is significantly negative, i.e. the least efficient firms are more likely to exit the market. The industry-level import penetration tends to increase the probability of firm exit, but the effect is mitigated by firm's TFP level as indicated by the negative and significant interaction term. This proves our first hypothesis that tougher import competition drives the least efficient firms out of the market. In columns (4)-(6), the negative and significant lagged TFP indicates that in general firms with initial lower TFP level tends to have a faster subsequent TFP growth due to the convergence or catch up effect. However, interestingly, import competition does not contribute to this process, i.e. import penetration itself has a significant and negative effect on firm-level TFP, and the negative effect is mitigated by firms' TFP level as indicated by the positive interaction term. This rejects our second hypothesis that import competition spurs the least efficient firms to improve their productivity. Such evidence is also consistent with our earlier findings in Ding *et al*. (2016) that import competition stimulates domestic firms' productivity growth only if firms and their industries are close to the world technology frontier, but discourages such growth for laggard firms and industries.

Thus, our exercise provides some direct evidence for the resource reallocation effect of trade liberalization as suggested by Melitz (2003) and Melitz and Ottaviano (2008). Import competition forces the least productive firms to exit, induces dynamic resource allocation

[24] The coefficients reported in columns (1)-(3) are the marginal effects from the IV probit estimation.

towards more productive firms, and thereby reducing the productivity dispersion within the industry. Our results, however, do not support the argument that import competition increases the productivity of the least efficient firms in China.

6. Further robustness tests

We adopt a number of empirical methods to examine the robustness of our findings. To save space, we will only discuss the main findings of these tests, and all detailed results and summary statistics of new variables used in these tests are reported in Appendices 5 and 6.

First, the quality of the productivity dispersion index is likely to be affected by the size of industries (Syverson, 2004; Balasubramanian and Sivadasan, 2009). For instance, industries with a small number of firms and monopolistic market structure tend to have lower productivity dispersion, which however cannot be interpreted as better resource allocation. To correct for this potential bias, we set the reciprocal of the number of firms as weights and run weighted regressions as a robustness test (Appendix Table A5(a)). Our main finding that import competition reduces the within-industry productivity dispersion remains intact.

Second, one justification of our use of industry-province as the unit of observation, instead of just industry, is the prevalence of local protectionism in China, as found in Bai *et al*. (2004). To directly test for this hypothesis, we include the import penetration ratio of neighbouring provinces (Neighbour-IMP)[25], along with the original IMP, into the regression. In Appendix Table A5(b), we find that the Neighbour-IMP has no impact on the within industry productivity dispersion of the province of interest, whereas its own IMP remains negative and significant. This provides further justification for our empirical methodology.

Thirdly, we adopt a number of alternative methods to construct TFP in order to examine the robustness of our results on productivity dispersion. For instance, the productivity measure in Appendix Table A5(c) is based on Levinsohn and Petrin (2003) approach, where intermediate inputs are used to proxy unobserved productivity in order to address Olley and Pakes (1996)'s problem of lumpy investment. Secondly, Wooldridge (2009)'s approach is used in Appendix Table A5(d), which is a unified approach allowing for the possibility that the first stage of Olley and Pakes (1996) or Levinsohn and Petrin (2003)

[25] We define the neighbouring provinces as provinces sharing common borders. We then take the average of import penetration ratios of all neighbouring provinces for each province. One exception is Hainan province, which is an island and does not share any common borders with other provinces. We use Guangdong province as its neighbouring province which is geographically closest to Hainan province.

approach actually contains identifying information for parameters on the variable inputs. Lastly, TFP is estimated using a system GMM estimator in Appendix Table A5(e), where fixed effects are allowed to take into account firms' (unmeasured) productivity advantages that persist over time. Our results remain robust when productivity dispersion is computed based on all these alternative TFP measures.

7. Conclusion

According to De Loecker and Goldberg (2013), the effect of trade liberalization on individual firms is important; however, what we care more about is how an industry, country or group of countries is affected by trade. Reallocation of economic resources from less towards more productive firms is one way in which industry- (or country) performance can increase even in the *absence* of any effects on individual firms. Thus, we follow this line of thinking by exploring the aggregate market reallocation effect of import competition rather than the efficiency improvement effect of individual firms in this study.

Using a number of comprehensive microeconomic datasets, we examined the effect of competition pressure from foreign imports on productivity dispersion of Chinese industries. Controlling for the supply-side, demand-side and China-specific factors, we found that the import-driven truncation of productivity distribution is indeed present in 425 narrowly defined manufacturing industries in China. We then distinguished the heterogeneous effects for various types of imports and found that the role of imports in reducing productivity dispersion is more evident for ordinary-trade imports and for final imported goods and intermediate goods imported by the same industries. A direct test on the channels through which import competition affects industry-level productivity dispersion shows that the trade-induced dynamic resource allocation within industries is achieved by driving the least efficient firms out of the market.

Our research has some policy implications. By virtue of its market size and growth momentum, China is an important trade partner of most of the economies in the world. The robust demand from China on manufacturing products has contributed significantly to the global recovery from the recent financial crisis. Our research calls for the further reduction of trade barriers, which is not only conducive to the reduction of resource misallocation in China, but also has significant economic and policy impacts to the rest of the world.

Supplementary material

Supplementary material- the Appendix and the Data files - are available online at the OUP website.

Funding paragraph

This work was supported by the Chinese National Social Science Foundation Grant on the project of ‘Chinese firms’ upgrading and industry quality’ [grant number: 12BJL049].

Acknowledgement

We are grateful for the constructive comments from two anonymous referees and the journal editor, Francis Teal. We thank Daniel Yi Xu, Miaojie Yu, Zhihong Yu, Johannes Van Biesebroeck, Beata Javorcik, Sandra Poncet and the participants at the international workshop on ‘Globalization of Chinese industrial sector: productivity, trade and finance’ at Glasgow in September 2013, the GEP China/Ifo/CEPII Conference on ‘Structural change and trade efficiency’ at the University of Nottingham, Ningbo in November 2013, the seminar at CCER, Peking University in November 2013, the Royal Economic Society Conference at Manchester in April 2014, and the China research seminar in Birmingham in May 2015 for constructive comments.

Reference

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

Amiti, M., and Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review*, 97, 1611-1638.

Asker, J., Collard-Wexler, A., and De Loecker, J. (2014). Dynamic inputs and resource (mis) allocation. *Journal of Political Economy*, 122, 1013-1063.

Bai, C. E., Du, Y. J., Tao, Z. G. and Tong, S. Y. (2004). Local protectionism and regional specialization: evidence from China’s industries. *Journal of International Economics*, 63, 397-417.

Balasubramanian, N., and Sivadasan, J. (2009). Capital resalability, productivity dispersion, and market structure. *The Review of Economics and Statistics*, 91, 547-557.

Banerjee, A., and Duflo, E. (2005). Growth theory through the lens of development economics. In P. Aghion, and S. Durlauf, *Handbook of Economic Growth*. Amsterdam: Elsevier.

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

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

Cassiman, B. and Veugelers, R. (2002). R&D cooperation and spillovers: some empirical evidence from Belgium. *American Economic Review*, 92, 1169-1184.

De Loecker, J., and Goldberg, P. K. (2013). Firm performance in a global market. NBER Working Paper No. 19308. Cambridge, MA.

Ding, S., Guariglia, A. and Knight, J. (2012). Negative Investment in China: Financing Constraints and Restructuring versus Growth. Leverhulme Centre for Research on Globalization and Economic Policy, Research Paper 12/01. Nottingham, UK.

Ding, S., Sun, P. and Jiang, W. (2016). The effect of import competition on firm productivity and innovation: does the distance to technology frontier matter? *Oxford Bulletin of Economics and Statistics*, 78, 197-227.

Dougherty, S., and Herd, R. (2005). Fast falling barriers and growing concentration: the emergence of a private economy in China. OECD Economics Department Working Paper No. 471. Paris, France.

Feenstra, R. C., Li, Z., and Yu, M. (2014). Exports and credit constraints under incomplete information: theory and evidence from China. *Review of Economics and Statistics*, 96, 729–744.

Fernandes, A. (2007). Trade policy, trade volumes and plant-level productivity in Colombian manufacturing industries. *Journal of International Economics*, 71, 52-71.

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

Halpern, L., Koren, M., and Szeidl, A. (2015). Imported inputs and productivity. *The American Economic Review*, 105, 3660-3703.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long-run equilibrium. *Econometrica*, 60, 1127-1150.

Hsieh, C.-T., and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, CXXIV, 1403-1447.

Jarreau, J., and Poncet, S. (2012). Export sophistication and economic growth: evidence from China. *Journal of Development Economics*, 97, 281-292.

Kasahara, H., and Lapham, B. (2013). Productivity and the decision to import and export: theory and evidence. *Journal of International Economics*, 89, 297-316.

Khandelwal, A. K., Schott, P. K., and Wei, S.-J. (2013). Trade liberalization and embedded institutional reform: evidence from Chinese exporters. *American Economic Review*, 103, 2169-2195.

Levinsohn, J. (1993). Testing the imports-as-market-discipline hypothesis. *Journal of International Economics*, 35, 1-22.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservable, *Review of Economic Studies*, 70, 317-41.

Lin, C., Lin P., and Song, F. (2010). Property rights protection and corporate R&D: evidence from China, *Journal of Development Economics*, 93, 49–62.

Maggioni, D. (2013). Productivity dispersion and its determinants: the role of import penetration. *Journal of Industry, Competition and Trade*, 13, 537-561.

Manova, K., and Zhang, Z. (2012). Export prices across firms and destinations. *Quarterly Journal of Economics*, 127, 379-436.

Manova, K., Wei, S.-J., and Zhang, Z. (2015). Firm exports and multinational activity under credit constraints. *Review of Economics and Statistics*, 97, 574-588.

Manova, K., and Yu, Z. (2016). How firms export: processing vs. ordinary trade with financial frictions. Journal of International Economics, forthcoming.

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

Melitz, M. and Ottaviano, G. (2008). Market size, trade and productivity. *Review of Economic Studies*, 75, 295-316.

Midrigan, V., and Xu, Y. D. (2014). Finance and misallocation: evidence from plant-level data. *American Economic Review*, 104, 422-458.

Olley, S., and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64, 1263-1297.

Song, Z., Storesletten K., and Zilibotti, F., (2011) Growing like China. *American Economic Review*, 101, 202-241.

Sutton, J. (1991). *Sunk Costs and Market Structure*. Cambridge, MA: MIT Press.

Syverson, C. (2004). Product substitutability and product dispersion. *The Review of Economics and Statistics*, 86, 534-550.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49, 326-365.

Topalova, P., and Klandelwal, A. (2011). Trade liberalization and firm productivity: the case of India. *The Review of Economics and Statistics*, 93, 995-1009.

Van Beveren, I. (2012) Total factor productivity estimation: A practical review, *Journal of Economic Surveys*, 26, 98-128.

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *Journal of Industrial Economics*, 55, 529-569.

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

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104, 112-114.

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

Table 1. Summary statistics of key variables in the baseline regression

	Observation	Mean	Std. Dev.	Min	Max
Dependent variable (productivity dispersion)					
TFP Dispesion (IQ Range)	16985	0.590	0.240	0.042	2.817
TFP Dispersion (Std. Dev)	16985	0.452	0.137	0.033	1.783
Independent variables					
IMP	16985	0.104	0.183	0.000	0.988
VALUELB	425	0.052	0.132	0.000	1.200
ADV	425	0.012	0.018	0.000	0.184
Fixed Cost	425	0.298	0.104	0.098	0.654
Sunk Cost	425	0.094	0.059	0.003	0.438
SOE	16985	0.170	0.232	0.000	1.000
FIE	16985	0.119	0.176	0.000	1.000
HHI	16985	0.194	0.158	0.002	0.985
Subsidy	16985	0.004	0.013	0.000	0.932
Instruments					
Output tariffs	16985	14.629	10.315	0.000	69.500
Changes of output tariffs	16985	14.348	7.942	0.455	52.500
Similar industry tariffs	16985	11.116	2.756	0.000	45.000

Table 2. Baseline results: the effects of import competition on productivity dispersion

Dispersion Measure	IQ Range				Standard Deviation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMP	-0.244***	-0.288***	-0.295***	-0.300***	-0.087*	-0.132***	-0.138***	-0.154***
	(0.078)	(0.076)	(0.074)	(0.074)	(0.046)	(0.044)	(0.040)	(0.040)
VALUELB	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	0.007***	0.004**	0.003	0.003	0.008***	0.006***	0.004***	0.004***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Cost	-0.151***	-0.162***	-0.121***	-0.122***	-0.065***	-0.079***	-0.037**	-0.038**
	(0.031)	(0.030)	(0.029)	(0.029)	(0.018)	(0.017)	(0.015)	(0.015)
Sunk Cost	-0.221***	-0.221***	-0.191***	-0.190***	-0.105***	-0.103***	-0.073***	-0.072***
	(0.031)	(0.031)	(0.030)	(0.030)	(0.017)	(0.017)	(0.016)	(0.016)
SOE		0.235***	0.204***	0.204***		0.183***	0.152***	0.151***
		(0.012)	(0.012)	(0.012)		(0.006)	(0.006)	(0.006)
FIE		0.086***	0.086***	0.086***		0.035***	0.034***	0.035***
		(0.013)	(0.013)	(0.013)		(0.007)	(0.006)	(0.006)
HHI			0.297***	0.297***			0.301***	0.301***
			(0.014)	(0.014)			(0.007)	(0.007)
Subsidy				0.209*				0.254***
				(0.124)				(0.077)
Under-identification test	728.678 [a]	712.990 [a]	710.401 [a]	703.147 [a]	728.678 [a]	712.990[a]	710.401[a]	703.147[a]
Weak identification test	282.000 [a]	280.469 [a]	280.186 [a]	277.110 [a]	282.000 [a]	176.202 [a]	280.186 [a]	277.110[a]
Durbin Wu-Hausman test	6.320 [b]	8.851 [b]	9.641 [b]	10.030 [b]	3.838 [b]	2.940 [c]	3.872 [b]	4.502 [b]
Adj-R^2	0.051	0.084	0.116	0.115	0.082	0.149	0.256	0.255
Observation	16985	16985	16985	16985	16985	16985	16985	16985
First-Stage Regressions								
IV1: Lagged tariff level	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
IV2: First-differenced tariff	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
IV3: Similar-industry tariff	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. In the first-stage regressions, IV1 reports the coefficients of the one-period lag of output tariffs, IV2 reports the coefficients of the changes of output tariffs, and IV3 reports the coefficients of the lagged similar industry tariffs. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Heterogeneous effects: the role of import competition on productivity dispersion

Dispersion Measure		IQ Range			Standard Deviation	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. The role of trade regimes						
PT-IMP	-0.012		-0.574*	-0.009		-0.420***
	(0.010)		(0.340)	(0.007)		(0.151)
OT-IMP		-6.555***	-4.549*		-2.244**	-2.374**
		(2.293)	(2.453)		(0.931)	(0.914)
Panel B. The nature of imported goods						
FIN-IMP	-0.371***			-0.181***		
	(0.077)			(0.040)		
SAME-IMP		-0.554***			-0.283***	
		(0.137)			(0.071)	
UP-IMP			5.522***			2.242**
			(2.072)			(0.896)
Observation	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. OT-IMP is the import penetration ratio for ordinary-trade imports; PT-IMP is the import penetration ratio for processing-trade imports; FIN-IMP is the import penetration ratio for imported final goods; SAME-IMP is the import penetration ratio for intermediate goods imported by the same industries; UP-IMP is the import penetration ratio for intermediate goods imported by the upstream industries; All other control variables and province, year, and industry fixed effects are included but not reported. See Appendix 4 for the full results of all the estimations. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Channels: the effects of import competition on firm exit and TFP convergence

Dependent variables	Firm Exit			TFP Growth		
	(1)	(2)	(3)	(4)	(5)	(6)
TFP	-0.089***	-0.025***	-0.024***			
	(0.002)	(0.003)	(0.003)			
TFP×IMP	-0.082***	-0.097***	-0.104***			
	(0.017)	(0.019)	(0.020)			
LTFP				-0.759***	-0.894***	-0.894***
				(0.006)	(0.005)	(0.006)
LTFP×IMP				0.280***	0.288***	0.293***
				(0.041)	(0.040)	(0.041)
IMP	0.337***	0.408***	0.438***	-1.260***	-1.310***	-1.332***
	(0.079)	(0.085)	(0.094)	(0.191)	(0.184)	(0.188)
Size		-0.056***	-0.057***		0.482***	0.482***
		(0.001)	(0.001)		(0.002)	(0.002)
Age		0.014***	0.015***		-0.002	-0.002
		(0.001)	(0.001)		(0.002)	(0.002)
SOE			-0.003**			0.000
			(0.001)			(0.003)
FIE			-0.009*			0.047***
			(0.005)			(0.009)
HHI			0.020***			0.010
			(0.005)			(0.009)
Under-identification test	60.050 [a]	63.974 [a]	59.001 [a]	54.056 [a]	55.648 [a]	54.853 [a]
Weak identification test	20.018 [a]	21.326 [a]	19.668 [a]	18.020 [a]	18.551 [a]	18.286 [a]
Durbin Wu-Hausman test	15.312 [a]	16.108 [a]	22.454 [a]	17.367 [a]	17.776 [a]	18.399 [a]
Adj-R^2	0.127	0.139	0.142	0.098	0.101	0.113
Observation	1329569	1284787	1284653	918285	912841	912756

Notes: Heteroskedasticity-robust standard errors are reported in parentheses; TFP is the firm-level TFP; LTFP is the lagged firm-level TFP; IMP is the industry-level import penetration ratio; Size is the natural logarithm of firm's total assets; Age is the firm age; SOE, FIE and HHI are defined in the same way as the baseline model. The coefficients reported in columns (1)-(3) are the marginal effect from the IV probit estimation. Firm, year and industry fixed effects are included in columns (4)-(6) but not reported to save space. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. TFP distribution in Chinese manufacturing industries (1998, 2003, 2007)

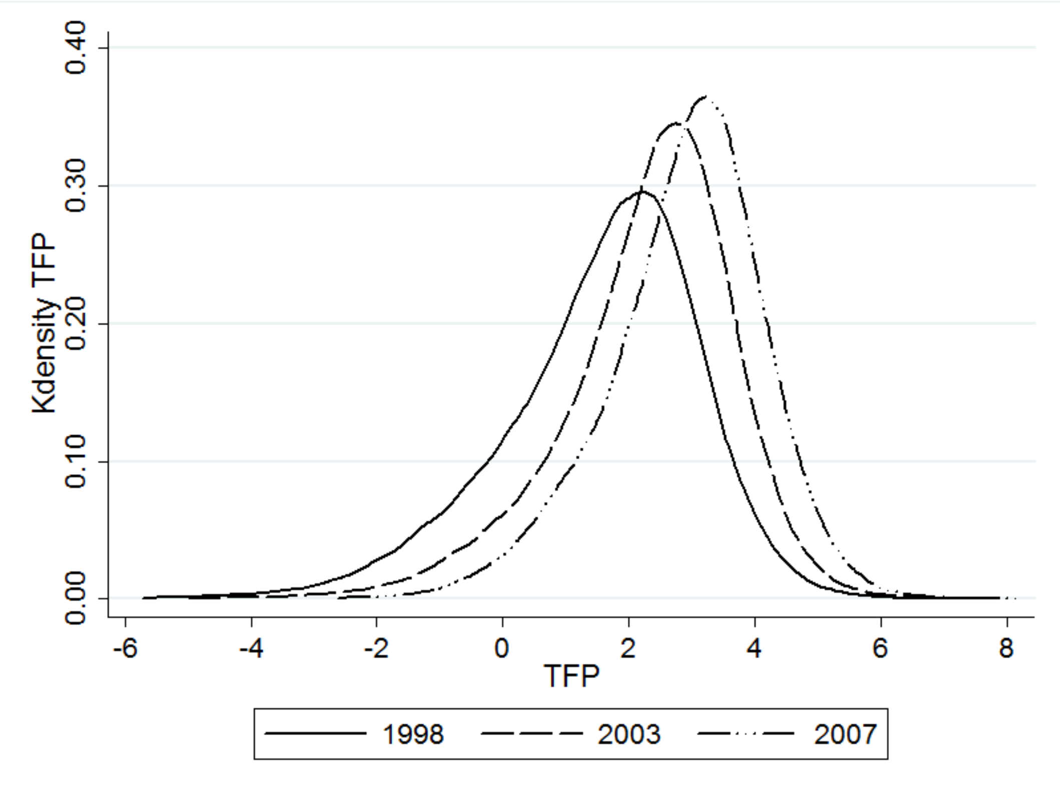
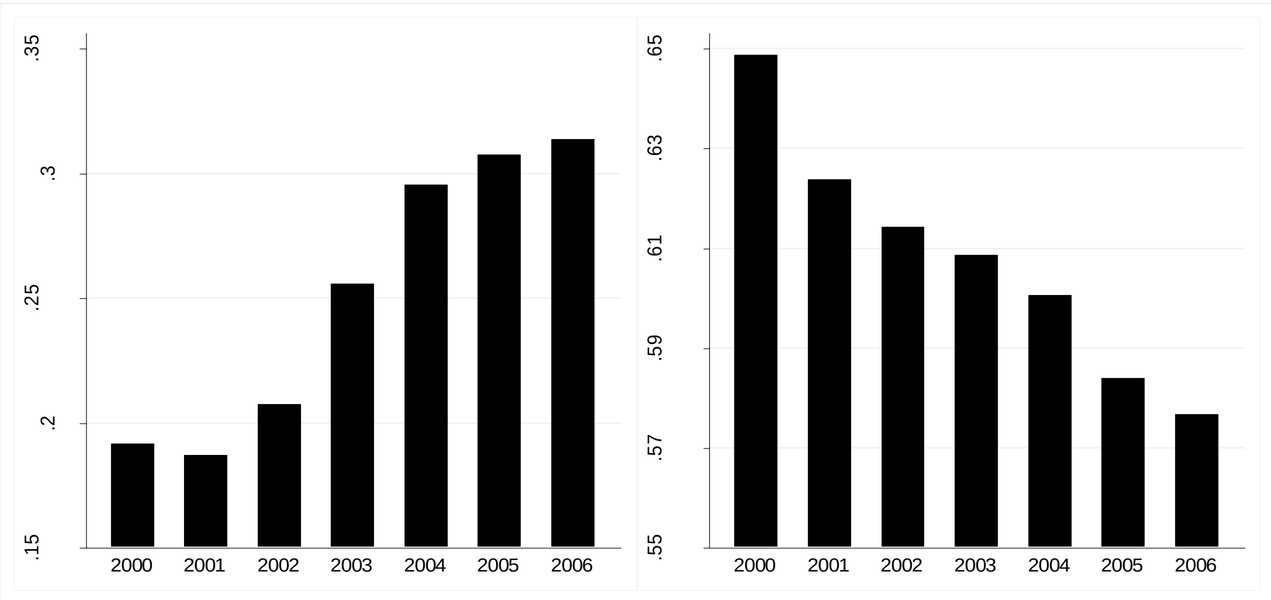


Figure 2. Import penetration and productivity dispersion in Chinese industries (2000-06)



Notes: import penetration is defined as the ratio of imports to the sum of GDP plus exports minus imports using the data from China Statistical Yearbook from 2000 to 2006; productivity dispersion is defined as the Olley-Pakes TFP interquartile (IQ) range using the NBS firm-level data from 2000 to 2006.

Import competition, dynamic resource allocation and productivity dispersion: micro-level evidence from China

By Sai Ding, Wei Jiang , Puyang Sun

Online Appendix 1. TFP estimation results

Table A1. TFP Estimates using Olley-Pakes approach

Industry code	Industry name	Capital	Labour	Returns-to-Scale	TFP
13	Food processing industry	0.464***	0.497***	-0.039	2.390
14	Food manufacturing industry	0.641***	0.593***	0.234***	1.203
15	Beverage manufacturing industry	0.628***	0.481***	0.109***	1.616
16	Tobacco processing industry	0.625***	0.386***	0.011	1.075
17	Textile industry	0.403***	0.415***	-0.182***	3.201
18	Clothing and other fiber products manufacturing	0.357***	0.542***	-0.101***	3.003
19	Leather, fur, down and down products industry	0.344***	0.504***	-0.152***	3.251
20	Timber processing, bamboo, cane, palm fiber and straw products industry	0.451***	0.455***	-0.094***	3.449
21	Furniture manufacturing industry	0.429***	0.681***	0.11***	2.166
22	Paper and paper products industry	0.503***	0.376***	-0.121***	2.574
23	Printing and record medium reproduction industry	0.814***	0.396***	0.21***	1.453
24	Educational and sports goods industry	0.259***	0.523***	-0.218***	3.189
25	Petroleum processing and coking industry	0.351***	0.355***	-0.294***	3.500
26	Chemical materials and chemical products manufacturing industry	0.458***	0.362***	-0.18***	2.638
27	Pharmaceutical manufacturing industry	0.534***	0.370***	-0.096***	2.591
28	Manufacture of chemical fibers industry	0.505***	0.361***	-0.134***	2.413
29	Rubber product industry	0.472***	0.392***	-0.136***	2.492
30	Plastic products industry	0.475***	0.389***	-0.136***	3.282
31	Non-metallic mineral products industry	0.618***	0.284***	-0.098***	1.961
32	Ferrous metal smelting and rolling processing industry	0.466***	0.459***	-0.075**	2.344

33	Non-ferrous metal smelting and rolling processing industry	0.382***	0.434***	-0.184***	3.332
34	Fabricated metal products industry	0.425***	0.440***	-0.135***	2.751
35	General machinery manufacturing industry	0.492***	0.391***	-0.117***	2.213
36	Special equipment manufacturing industry	0.674***	0.400***	0.074**	1.902
37	Transportation equipment manufacturing industry	0.631***	0.507***	0.138***	1.775
39	Weapons and ammunition industry	0.467***	0.452***	-0.081**	2.215
40	Electrical machinery and equipment manufacturing	0.453***	0.509***	-0.038	2.944
41	Electronic and communication equipment manufacturing industry	0.528***	0.418***	-0.054*	2.139
42	Instrumentation and culture, office machinery manufacturing industry	0.353***	0.469***	-0.178***	3.346

Notes: This table reports the estimated coefficients of the production function and the associated log of TFP by industry; the Returns-to-Scale column reports the sum of the capital and labour elasticities minus 1, and the result of Wald test of constant returns to scale is indicated by the significance level, i.e. if the value is statistically insignificant, then the null hypothesis of constant return to scale cannot be rejected; and if the value is significant and positive (negative), it indicates increasing (decreasing) returns to scale; industry code is the 2-digit Chinese Standard Industrial Classification (CSIC) code; *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix 2. Dataset merging information

The difficulty of merging the NBS dataset and the Customs dataset lies in the absence of a common firm identifier shared by both datasets. We therefore rely on other firm characteristics such as firm name, telephone number, zip code, and firm address to achieve the best possible match of two datasets. Table A2(a) presents a brief summary of the datasets. We find that the number of exporting firms in the NBS dataset is much smaller than that in the Customs dataset[1]. There are two explanations for this discrepancy. First, most trading firms are quite small, so that they are not included in the 'above-scale' NBS dataset (Yu, 2015). Second, the NBS dataset covers manufacturing firms only, whereas the Customs dataset consists of trading firms in all sectors in China such as manufacturing, agriculture, service, and so on. During the period of 2000-06, the number of exporting firms in our merged dataset accounts for 58.5% of total exporting firms in the NBS dataset on average.

Table A2(b) and Table A2(c) provide the representativeness of our merged sample compared with the full-sample NBS data. Table A2(b) shows how much of total sales, exports and employment are accounted for by the merged data each year during 2000-06. On average, our merged data covers 44% of total sales, 74% of total exports and 36.4% of total employment in the firm-level NBS data. Table A2(c) shows that our merged sample has higher means of sales, exports and number of employees than the corresponding figures in the full-sample NBS data. These findings suggest that the merged sample is skewed toward large manufacturing firms in China.

[1] Note that although Customs dataset includes both imports and exports information, the NBS dataset contains firms' exporting information only.

Table A2(a). Basic summary of datasets

	2000	2001	2002	2003	2004	2005	2006
NBS data [a]	119,444	131,437	145,464	163,332	238,078	237,116	263,158
##Export	36,908	40,128	45,040	50,616	76,607	74,395	77,723
Customs data [b]	81,995	89,660	104,245	124,299	153,779	179,666	208,425
##Export	62,746	68,487	78,612	95,688	120,590	144,030	171,205
##Import	62,750	67,588	77,303	87,934	102,242	113,456	121,835
Merged data [c]	25,712	29,615	33,918	39,020	56,937	57,058	60,999
##Export	19,104	21,914	25,683	30,611	44,790	46,372	50,211
##Import	18,094	20,041	22,700	25,787	36,943	36,332	38,102
Merge Ratio [d]	51.76%	54.61%	57.02%	60.48%	58.47%	62.33%	64.60%

Notes: (a) the NBS firm-level dataset includes above-scale firms in the manufacturing sectors in China; it also reports firms' export sales, but there is no information on imports; (b) Customs dataset contains detailed product-level information of international trade (both exports and imports) at the monthly level; we therefore aggregate such information to the firm-year level in order to merge it with the NBS dataset; (c) The merge of the two dataset is mainly based on the firm name, and other firm characteristics such as telephone number, zip code and firm address; (d) The merge ratio is computed as the number of exporting firms in the merged dataset in relation to the number of total exporting firms in the NBS dataset.

Table A2(b). Firm-level production information in merged vs. full-sample NBS data by year

Variables	2000	2001	2002	2003	2004	2005	2006	Average
Sales (%)	41.19	43.08	43.88	45.54	45.98	44.95	43.47	44.01
Exports (%)	68.55	71.05	72.94	74.47	77.10	76.67	77.42	74.03
Number of Employees (%)	30.31	33.05	35.15	37.29	39.14	39.94	39.90	36.40

Notes: the value in this table represents the percentages of total sales, exports and employment of the merged data in the full-sample NBS data.

Table A2(c). Comparison of the merged data and the full-sample NBS data

Variables	Full-sample data			Merged data		
	Mean	Min	Max	Mean	Min	Max
Sales (RMB 1,000)	68717.90	5001	9993990	135093.9	5001	9987010
Exports (RMB 1,000)	17196.51	0	1.52E+08	57453.14	0	1.52e+08
Number of Employees	272.55	8	147722	455.8332	8	101375

Online Appendix 3. Additional summary statistics

Table A3(a). Productivity dispersion in 2-digit Chinese industrial sectors (2000-06, based on the interquartile range)

Industrial sectors	2000	2001	2002	2003	2004	2005	2006
Food processing industry	0.656	0.647	0.635	0.616	0.603	0.579	0.561
Food Manufacturing industry	0.680	0.647	0.650	0.651	0.648	0.642	0.631
Beverage Manufacturing industry	0.838	0.834	0.801	0.790	0.795	0.751	0.723
Tobacco processing industry	1.159	1.101	1.082	0.970	0.974	0.806	0.694
Textile industry	0.529	0.517	0.506	0.496	0.491	0.476	0.476
Clothing and other fiber products manufacturing	0.579	0.562	0.542	0.537	0.535	0.532	0.529
Leather, fur, down and down products industry	0.608	0.629	0.625	0.608	0.604	0.604	0.607
Timber Processing, Bamboo, Cane, Palm Fiber and Straw Products industry	0.597	0.549	0.539	0.516	0.532	0.472	0.486
Furniture Manufacturing industry	0.672	0.617	0.594	0.610	0.604	0.595	0.594
Paper and paper products industry	0.580	0.554	0.540	0.527	0.526	0.527	0.530
Printing and Record Medium Reproduction industry	0.814	0.769	0.751	0.718	0.678	0.653	0.619
Educational and Sports Goods industry	0.543	0.555	0.541	0.543	0.557	0.535	0.543
Petroleum processing and coking industry	0.708	0.684	0.665	0.657	0.710	0.713	0.751
Chemical materials and chemical products manufacturing industry	0.655	0.606	0.582	0.557	0.556	0.532	0.521
Pharmaceutical Manufacturing industry	0.638	0.614	0.606	0.616	0.649	0.616	0.607
Manufacture of Chemical Fibers industry	0.821	0.772	0.781	0.721	0.696	0.640	0.673
Rubber product industry	0.575	0.570	0.603	0.573	0.550	0.540	0.554
Plastic products industry	0.537	0.512	0.510	0.509	0.516	0.505	0.511
Non-metallic mineral products industry	0.501	0.481	0.472	0.452	0.457	0.440	0.436
Ferrous metal smelting and rolling processing industry	0.635	0.624	0.628	0.643	0.673	0.630	0.627
Non-ferrous metal smelting and rolling processing industry	0.692	0.677	0.666	0.664	0.676	0.640	0.623
Fabricated Metal Products industry	0.593	0.557	0.536	0.529	0.530	0.516	0.517
General machinery manufacturing industry	0.613	0.595	0.577	0.558	0.560	0.543	0.530
Special equipment manufacturing industry	0.534	0.516	0.504	0.471	0.467	0.456	0.452
Transportation Equipment Manufacturing industry	0.688	0.624	0.631	0.623	0.616	0.591	0.589
Electrical machinery and equipment manufacturing	0.659	0.639	0.634	0.636	0.645	0.640	0.637
Electronic and communication equipment manufacturing industry	0.746	0.702	0.730	0.752	0.730	0.734	0.727
Instrumentation and culture, office machinery manufacturing industry	0.678	0.653	0.623	0.613	0.644	0.621	0.641
Other manufacturing industry	0.523	0.524	0.507	0.546	0.545	0.514	0.534

Table A3(b). Import penetration ratio in 2-digit Chinese industrial sectors (2000-06)

	2000	2001	2002	2003	2004	2005	2006
Food processing industry	0.075	0.066	0.068	0.075	0.076	0.056	0.052
Food Manufacturing industry	0.037	0.040	0.035	0.032	0.032	0.025	0.023
Beverage Manufacturing industry	0.007	0.006	0.005	0.005	0.005	0.005	0.006
Tobacco processing industry	0.015	0.017	0.013	0.015	0.014	0.019	0.022
Textile industry	0.126	0.123	0.101	0.092	0.075	0.061	0.049
Clothing and other fiber products manufacturing	0.049	0.046	0.040	0.039	0.040	0.030	0.026
Leather, fur, down and down products industry	0.079	0.066	0.072	0.064	0.068	0.049	0.043
Timber Processing, Bamboo, Cane, Palm Fiber and Straw Products industry	0.129	0.087	0.086	0.090	0.070	0.048	0.031
Furniture Manufacturing industry	0.060	0.075	0.076	0.104	0.099	0.069	0.058
Paper and paper products industry	0.173	0.157	0.144	0.133	0.11	0.088	0.073
Printing and Record Medium Reproduction industry	0.032	0.033	0.029	0.026	0.026	0.025	0.024
Educational and Sports Goods industry	0.290	0.312	0.339	0.348	0.371	0.359	0.325
Petroleum processing and coking industry	0.033	0.03	0.138	0.148	0.161	0.158	0.184
Chemical materials and chemical products manufacturing industry	0.317	0.331	0.317	0.324	0.319	0.297	0.271
Pharmaceutical Manufacturing industry	0.061	0.073	0.068	0.070	0.072	0.068	0.065
Manufacture of Chemical Fibers industry	0.153	0.164	0.158	0.143	0.127	0.102	0.072
Rubber product industry	0.076	0.084	0.076	0.102	0.103	0.094	0.111
Plastic products industry	0.131	0.126	0.120	0.128	0.126	0.125	0.115
Non-metallic mineral products industry	0.045	0.046	0.043	0.043	0.041	0.035	0.030
Ferrous metal smelting and rolling processing industry	0.131	0.129	0.143	0.148	0.116	0.111	0.088
Non-ferrous metal smelting and rolling processing industry	0.225	0.227	0.209	0.216	0.193	0.193	0.160
Fabricated Metal Products industry	0.132	0.135	0.134	0.134	0.145	0.123	0.114
General machinery manufacturing industry	0.353	0.377	0.362	0.373	0.355	0.313	0.283
Special equipment manufacturing industry	0.371	0.452	0.421	0.426	0.404	0.368	0.329
Transportation Equipment Manufacturing industry	0.114	0.177	0.142	0.171	0.15	0.145	0.172
Electrical machinery and equipment manufacturing	0.464	0.511	0.607	0.686	0.858	0.903	1.079
Electronic and communication equipment manufacturing industry	0.494	0.555	0.573	0.702	0.746	0.737	0.705
Instrumentation and culture, office machinery manufacturing industry	0.152	0.148	0.122	0.113	0.114	0.090	0.075
Other manufacturing industry	0.672	0.782	0.733	0.830	0.693	0.678	0.646

Online Appendix 4. Full results: various heterogeneous effects on how imports affect productivity dispersion

Table A4(a). Heterogeneous effect: the role of trade regimes

Dispersion Measure	IQ Range			Standard Deviation		
	(1)	(2)	(3)	(4)	(5)	(6)
PT-IMP	-0.012		-0.574*	-0.009		-0.420***
	(0.010)		(0.340)	(0.007)		(0.151)
OT-IMP		-6.555***	-4.549*		-2.244**	-2.374**
		(2.293)	(2.453)		(0.931)	(0.914)
VALUELB	-0.002***	-0.004***	-0.004***	-0.001***	-0.002***	-0.002***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
ADV	0.008***	-0.004	-0.005	0.006***	0.003	0.002
	(0.001)	(0.005)	(0.004)	(0.001)	(0.002)	(0.002)
Fixed Cost	-0.028	-0.135***	-0.181***	0.007	-0.030	-0.063***
	(0.019)	(0.048)	(0.044)	(0.010)	(0.020)	(0.022)
Sunk Cost	-0.197***	0.359	0.189	-0.074***	0.114	-0.010
	(0.029)	(0.222)	(0.224)	(0.015)	(0.088)	(0.081)
SOE	0.195***	0.256***	0.234***	0.148***	0.167***	0.151***
	(0.012)	(0.031)	(0.030)	(0.006)	(0.012)	(0.011)
FIE	0.088***	0.004	0.068	0.037***	0.007	0.054***
	(0.013)	(0.036)	(0.051)	(0.007)	(0.014)	(0.020)
HHI	0.294***	0.430***	0.390***	0.300***	0.346***	0.317***
	(0.014)	(0.052)	(0.053)	(0.007)	(0.021)	(0.020)
Subsidy	0.174	1.098*	1.004**	0.243***	0.552**	0.483***
	(0.121)	(0.596)	(0.487)	(0.076)	(0.248)	(0.185)
Under-identification test	6.962 [c]	19.857 [a]	7.707 [a]	6.962 [c]	19.857 [a]	7.707 [a]
Weak identification test	1.791	16.603 [b]	2.543	1.791	16.603 [a]	2.543
Durbin Wu-Hausman test	3.909 [b]	20.669 [a]	12.076 [a]	7.991 [a]	8.679 [a]	7.725 [a]
Adjusted R^2	0.132	0.129	0.133	0.182	0.193	0.199
Observation	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Table A4(b). Heterogeneous effect: the nature of imported goods

Dispersion Measure	IQ Range			Standard Deviation		
	(1)	(2)	(3)	(4)	(5)	(6)
FIN-IMP	-0.371***			-0.181***		
	(0.077)			(0.040)		
IN-IMP		-0.554***			-0.283***	
		(0.137)			(0.071)	
UP-IMP			5.522***			2.242**
			(2.072)			(0.896)
VALUELB	-0.003***	-0.003***	-0.001***	-0.002***	-0.002***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	0.003	0.003	0.019***	0.004***	0.004***	0.011***
	(0.002)	(0.002)	(0.005)	(0.001)	(0.001)	(0.002)
Fixed Cost	-0.157***	-0.106***	0.092	-0.056***	-0.033**	0.056**
	(0.031)	(0.026)	(0.057)	(0.016)	(0.014)	(0.025)
Sunk Cost	-0.167***	-0.200***	-0.155**	-0.060***	-0.076***	-0.059**
	(0.032)	(0.034)	(0.067)	(0.017)	(0.018)	(0.029)
SOE	0.206***	0.188***	0.196***	0.152***	0.143***	0.147***
	(0.009)	(0.010)	(0.018)	(0.005)	(0.005)	(0.008)
FIE	0.091***	0.125***	-0.101	0.037***	0.055***	-0.041
	(0.011)	(0.016)	(0.073)	(0.006)	(0.008)	(0.032)
HHI	0.300***	0.300***	0.232***	0.303***	0.303***	0.274***
	(0.012)	(0.013)	(0.034)	(0.006)	(0.007)	(0.015)
Subsidy	0.237*	0.191	-0.163	0.268***	0.247***	0.100
	(0.131)	(0.143)	(0.297)	(0.069)	(0.075)	(0.129)
Under-identification test	588.756 [a]	297.618 [a]	24.879 [a]	588.756 [a]	297.618 [a]	24.879 [a]
Weak identification test	230.368 [a]	111.229 [a]	8.332	230.368 [a]	111.229 [a]	8.332
Durbin Wu-Hausman test	20.731 [a]	20.008 [a]	32.574 [a]	13.52 [a]	17.546 [a]	18.882 [a]
Adjusted R^2	0.101	0.105	0.111	0.124	0.119	0.125
Observation	16985	16985	16984	16985	16985	16984

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix 5. Full results: further robustness tests

Table A5(a). The weighted regression method

Dispersion Measure	IQ Range				Standard Deviation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMP	-0.298***	-0.288***	-0.316***	-0.320***	-0.162***	-0.192***	-0.180***	-0.185***
	(0.095)	(0.076)	(0.090)	(0.090)	(0.056)	(0.053)	(0.048)	(0.048)
VALUELB	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	0.005**	0.004**	0.002	0.002	0.006***	0.004***	0.003**	0.003**
	(0.003)	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Cost	-0.132***	-0.162***	-0.103***	-0.104***	-0.065***	-0.077***	-0.034*	-0.035**
	(0.036)	(0.030)	(0.034)	(0.034)	(0.020)	(0.019)	(0.018)	(0.018)
Sunk Cost	-0.244***	-0.221***	-0.215***	-0.214***	-0.119***	-0.119***	-0.083***	-0.083***
	(0.044)	(0.031)	(0.041)	(0.041)	(0.024)	(0.023)	(0.021)	(0.021)
SOE		0.235***	0.221***	0.220***		0.196***	0.163***	0.162***
		(0.012)	(0.014)	(0.014)		(0.008)	(0.007)	(0.007)
FIE		0.086***	0.081***	0.081***		0.034***	0.034***	0.034***
		(0.013)	(0.015)	(0.015)		(0.008)	(0.008)	(0.008)
HHI			0.298***	0.298***			0.325***	0.325***
			(0.017)	(0.017)			(0.009)	(0.009)
Subsidy				0.215				0.243***
				(0.137)				(0.077)
Under-identification test	728.678 [a]	712.990 [a]	710.401 [a]	703.147 [a]	728.678 [a]	712.990[a]	710.401[a]	703.147[a]
Weak identification test	282.000 [a]	280.469 [a]	280.186 [a]	277.110 [a]	282.000 [a]	176.202 [a]	280.186 [a]	277.110[a]
Durbin Wu-Hausman F test	6.320 [b]	8.851 [b]	9.641 [b]	10.030 [b]	3.838 [b]	2.940 [c]	3.872 [b]	4.502 [b]
Adj-R^2	0.034	0.084	0.107	0.106	0.055	0.130	0.251	0.250
Observation	16985	16985	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Table A5(b). The effect of import competition of neighbouring province on TFP dispersion

Dispersion Measure	IQ Range				Standard Deviation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMP	-0.247***	-0.291***	-0.297***	-0.302***	-0.089*	-0.132***	-0.138***	-0.145***
	(0.078)	(0.076)	(0.074)	(0.074)	(0.046)	(0.044)	(0.040)	(0.040)
Neighbour-IMP	0.027	0.021	0.021	0.021	0.007	0.004	0.003	0.003
	(0.019)	(0.019)	(0.018)	(0.018)	(0.006)	(0.006)	(0.005)	(0.005)
VALUELB	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	0.007***	0.004**	0.003	0.003	0.008***	0.006***	0.004***	0.004***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Cost	-0.152***	-0.163***	-0.121***	-0.122***	-0.065***	-0.080***	-0.037**	-0.038**
	(0.031)	(0.030)	(0.029)	(0.029)	(0.018)	(0.017)	(0.015)	(0.015)
Sunk Cost	-0.221***	-0.221***	-0.191***	-0.190***	-0.105***	-0.103***	-0.073***	-0.072***
	(0.031)	(0.031)	(0.030)	(0.030)	(0.017)	(0.017)	(0.016)	(0.016)
SOE		0.235***	0.204***	0.204***		0.183***	0.152***	0.151***
		(0.012)	(0.012)	(0.012)		(0.006)	(0.006)	(0.006)
FIE		0.086***	0.086***	0.086***		0.035***	0.034***	0.035***
		(0.013)	(0.013)	(0.013)		(0.007)	(0.006)	(0.006)
HHI			0.297***	0.297***			0.301***	0.301***
			(0.014)	(0.014)			(0.007)	(0.007)
Subsidy				0.209*				0.254***
				(0.124)				(0.077)
Under-identification test	724.371 [a]	708.776 [a]	706.208 [a]	699.005 [a]	724.371 [a]	708.776[a]	706.208[a]	699.005[a]
Weak identification test	280.096 [a]	278.588 [a]	278.302 [a]	275.249 [a]	280.096 [a]	278.588 [a]	278.302 [a]	275.249[a]
Durbin Wu-Hausman F test	6.467 [b]	8.975 [a]	9.730 [a]	10.122 [a]	0.875	2.972 [c]	3.861 [b]	4.490 [b]
Adj-R^2	0.051	0.083	0.116	0.115	0.082	0.149	0.256	0.255
Observation	16985	16985	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Table A5(c). An alternative productivity measure based on Levinsohn and Petrin (2003)

Dispersion Measure	IQ Range				Standard Deviation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMP	-0.359***	-0.382***	-0.390***	-0.393***	-0.205***	-0.223***	-0.230***	-0.234***
	(0.070)	(0.071)	(0.070)	(0.070)	(0.041)	(0.041)	(0.039)	(0.039)
VALUELB	-0.002***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	0.003*	0.002	0.001	0.001	0.002*	0.001	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Cost	-0.199***	-0.206***	-0.168***	-0.169***	-0.125***	-0.130***	-0.092***	-0.093***
	(0.027)	(0.028)	(0.027)	(0.027)	(0.017)	(0.017)	(0.016)	(0.016)
Sunk Cost	-0.127***	-0.126***	-0.098***	-0.098***	-0.099***	-0.098***	-0.070***	-0.070***
	(0.031)	(0.031)	(0.030)	(0.030)	(0.019)	(0.019)	(0.017)	(0.017)
SOE		0.102***	0.074***	0.073***		0.066***	0.038***	0.037***
		(0.011)	(0.011)	(0.011)		(0.006)	(0.006)	(0.006)
FIE		0.030**	0.029**	0.030**		0.014*	0.013*	0.013*
		(0.012)	(0.012)	(0.012)		(0.007)	(0.007)	(0.007)
HHI			0.271***	0.271***			0.275***	0.275***
			(0.014)	(0.014)			(0.008)	(0.008)
Subsidy				0.120				0.144
				(0.235)				(0.178)
Under-identification test	728.678 [a]	712.990 [a]	710.401 [a]	703.147 [a]	728.678 [a]	712.990[a]	710.401[a]	703.147[a]
Weak identification test	282.000 [a]	280.469 [a]	280.186 [a]	277.110 [a]	282.000 [a]	176.202 [a]	280.186 [a]	277.110[a]
Durbin Wu-Hausman F test	18.563 [a]	20.875 [a]	22.287 [a]	22.286 [a]	16.283 [a]	19.175 [a]	23.071 [a]	23.553 [a]
Adj-R^2	0.013	0.014	0.045	0.044	0.044	0.046	0.132	0.131
Observation	16985	16985	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Table A5(d). An alternative productivity measure based on Woodridge (2009)

Dispersion Measure	IQ Range				Standard Deviation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMP	-0.521***	-0.544***	-0.554***	-0.553***	-0.205***	-0.223***	-0.230***	-0.234***
	(0.090)	(0.091)	(0.089)	(0.090)	(0.041)	(0.041)	(0.039)	(0.039)
VALUELB	-0.003***	-0.003***	-0.002***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	-0.004	-0.005*	-0.007***	-0.007***	0.002*	0.001	-0.000	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Cost	-0.258***	-0.265***	-0.216***	-0.216***	-0.125***	-0.130***	-0.092***	-0.093***
	(0.036)	(0.036)	(0.035)	(0.036)	(0.017)	(0.017)	(0.016)	(0.016)
Sunk Cost	-0.148***	-0.147***	-0.112***	-0.112***	-0.099***	-0.098***	-0.070***	-0.070***
	(0.039)	(0.039)	(0.038)	(0.038)	(0.019)	(0.019)	(0.017)	(0.017)
SOE		0.094***	0.058***	0.058***		0.066***	0.038***	0.037***
		(0.014)	(0.014)	(0.014)		(0.006)	(0.006)	(0.006)
FIE		0.027	0.026	0.026		0.014*	0.013*	0.013*
		(0.017)	(0.017)	(0.017)		(0.007)	(0.007)	(0.007)
HHI			0.349***	0.349***			0.275***	0.275***
			(0.019)	(0.019)			(0.008)	(0.008)
Subsidy				-0.016				0.144
				(0.255)				(0.178)
Under-identification test	728.678 [a]	712.990 [a]	710.401 [a]	703.147 [a]	728.678 [a]	712.990[a]	710.401[a]	703.147[a]
Weak identification test	282.000 [a]	280.469 [a]	280.186 [a]	277.110 [a]	282.000 [a]	176.202 [a]	280.186 [a]	277.110[a]
Durbin Wu-Hausman F test	31.471 [a]	33.567 [a]	35.681 [a]	35.339 [a]	16.284 [a]	19.175 [a]	23.071 [a]	23.553 [a]
Adj-R^2	0.007	0.005	0.034	0.034	0.044	0.046	0.132	0.131
Observation	16985	16985	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Table A5(e). An alternative productivity measure based on system GMM

Dispersion Measure	IQ Range				Standard Deviation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMP	-0.359***	-0.790***	-0.798***	-0.803***	-0.205***	-0.223***	-0.230***	-0.234***
	(0.070)	(0.094)	(0.093)	(0.093)	(0.041)	(0.041)	(0.039)	(0.039)
VALUELB	-0.002***	-0.003***	-0.002***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADV	0.003*	-0.003	-0.005**	-0.005**	0.002*	0.001	-0.000	-0.000
	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Cost	-0.199***	-0.377***	-0.328***	-0.329***	-0.125***	-0.130***	-0.092***	-0.093***
	(0.027)	(0.037)	(0.037)	(0.037)	(0.017)	(0.017)	(0.016)	(0.016)
Sunk Cost	-0.127***	-0.190***	-0.155***	-0.154***	-0.099***	-0.098***	-0.070***	-0.070***
	(0.031)	(0.040)	(0.039)	(0.039)	(0.019)	(0.019)	(0.017)	(0.017)
SOE		0.099***	0.064***	0.063***		0.066***	0.038***	0.037***
		(0.014)	(0.014)	(0.014)		(0.006)	(0.006)	(0.006)
FIE		0.046***	0.046***	0.046***		0.014*	0.013*	0.013*
		(0.017)	(0.017)	(0.017)		(0.007)	(0.007)	(0.007)
HHI			0.347***	0.347***			0.275***	0.275***
			(0.019)	(0.019)			(0.008)	(0.008)
Subsidy				0.202				0.144
				(0.283)				(0.178)
Under-identification test	728.678 [a]	712.990 [a]	710.401 [a]	703.147 [a]	728.678 [a]	712.990[a]	710.401[a]	703.147[a]
Weak identification test	282.000 [a]	280.469 [a]	280.186 [a]	277.110 [a]	282.000 [a]	176.202 [a]	280.186 [a]	277.110[a]
Durbin Wu-Hausman F test	18.562 [a]	68.752 [a]	72.683 [a]	73.091 [a]	16.284[a]	19.175 [a]	23.071 [a]	23.553 [a]
Adj-R^2	0.013	0.014	0.015	0.015	0.044	0.046	0.132	0.131
Observation	16985	16985	16985	16985	16985	16985	16985	16985

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. The under-identification test is based on the Kleibergen-Paap rk LM statistic, with a null hypothesis that the model is under-identified; the weak identification test is based on the Cragg-Donald Wald F statistic, with a null hypothesis that the first stage regression is weakly identified; and the Durbin–Wu–Hausman test for endogeneity is based on the F statistics, with the null hypothesis that the endogenous regressors in the model are in fact exogenous; the superscripts [a] [b] and [c] are used to indicate that the p-value of the test statistics is below 0.01, 0.05 and 0.1 respectively, suggesting that the corresponding null hypothesis is rejected. All province, year, and industry fixed effects are included but not reported. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix 6. Summary statistics of variables in various heterogeneous effects, channels and further robustness tests

	Observation	Province*	Industry	Mean	Std. Error	Min	Max
Heterogeneous effect: the role of trade regime							
PT-IMP	16985	30	425	0.046	0.127	0.000	0.732
OT-IMP	16985	30	425	0.010	0.045	0.000	0.325
Heterogeneous effect: the nature of imported goods							
Fin-Import	16985	30	425	0.085	0.165	0.000	0.787
SAME-Import	16985	30	425	0.066	0.145	0.000	0.770
Up-Import	16985	30	425	0.013	0.036	0.000	0.233
Channels: firm exit or convergence?							
TFP	1284653	30	425	2.897	0.909	-0.715	9.101
Size	1284653	30	425	4.726	1.176	0.000	11.903
Age	1284653	30	425	1.871	0.991	0.000	5.603
Further robustness test: the effect of neighbouring province							
Neighbour-IMP	16985	30	425	0.055	6.361	0.000	0.791
Further robustness test: alternative measure of productivity dispersion based on Levinsohn and Petrin (2003)							
IQ Range	16985	30	425	0.567	0.212	0.162	1.307
Standard deviation	16985	30	425	0.432	0.133	0.141	0.793
Further robustness test: alternative measure of productivity dispersion based on Wooldridge (2009)							
IQ Range	16985	30	425	0.612	0.241	0.181	1.543
Standard deviation	16985	30	425	0.447	0.135	0.183	0.882
Further robustness test: alternative measure of productivity dispersion based on System GMM							
IQ Range	16985	30	425	0.641	0.246	0.198	1.590
Standard deviation	16985	30	425	0.445	0.141	0.154	0.763

* We do not include Tibet in our empirical estimation due to the data coverage in both NBS and Customs data.