THE
ECONOMIC
JOURNAL

The Economic Journal, **125** (*June*), 943–988. Doi: 10.1111/ecoj.12127 © 2014 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

PROCESSING TRADE, TARIFF REDUCTIONS AND FIRM PRODUCTIVITY: EVIDENCE FROM CHINESE FIRMS*

Miaojie Yu

This article explores how reductions in tariffs on imported inputs and final goods affect the productivity of large Chinese trading firms, with the special tariff treatment that processing firms receive on imported inputs. Firm-level input and output tariffs are constructed. Both types of tariff reductions have positive impacts on productivity that are weaker as firms' share of processing imports grows. The impact of input tariff reductions on productivity improvement, overall, is weaker than that of output tariff reductions, although the opposite is true for non-processing firms only. Both tariff reductions are found to contribute at least 14.5% to economy-wide productivity growth.

The effect of trade liberalisation on firm productivity is one of the most important topics in empirical trade research. Initially, trade economists primarily focused on the effect of cutting tariffs on final goods. At present, research interest has shifted to exploration of the effect of tariff reductions on imported intermediate inputs, which is usually greater than the effect on final goods (Amiti and Konings, 2007; Goldberg *et al.*, 2010; Topalova and Khandelwal, 2011). Amiti and Konings (2007) analyse Indonesian firm-level data and find that firms' gains from reduction of input tariffs are at least twice as much as those from reduction of output tariffs. Furthermore, Topalova and Khandelwal (2011) find that Indian firms' gains from input tariff reduction could be ten times greater than those from output tariff reduction in several industries. They forcefully argue that the primary reason for this result is that access to better intermediate inputs through the reduction of input tariffs is more important than the pro-competitive effect of the reduction of output tariffs.

Different from such findings, the present article shows that reducing output tariffs has had a greater effect on productivity improvement than reducing input tariffs for large Chinese trading firms in the new century. A 10 percentage point fall in output (input) tariffs leads to a productivity gain of 9.2 (5.1)%. The positive impact of both types of tariff reductions on productivity improvement is weaker as the firm's share of processing imports grows. Such results are primarily attributable to the special tariff treatment afforded to imported inputs by processing firms as opposed to non-processing firms in China. Processing imports, which account for half of total imports in China, have zero tariffs. Further tariff reductions on imported intermediate inputs have no impact on firms that entirely engage in processing trade but still have some impact on firms that engage in both processing and non-processing trade. As the

* Corresponding author: Miaojie Yu, China Centre for Economic Research (CCER), National School of Development, Peking University, Beijing 100871, China. Email: mjyu@ccer.pku.edu.cn.

I thank Robert Feenstra, Gordon Hanson, Samuel Kortum, Kala Krishna, Zhiyuan Li, Justin Lin, Devashish Mitra, Larry Qiu, Jose Antonio Rodriguez Lopez, Robert Staiger, Joaquim Silvestre, Wei Tian, Heiwai Tang, Yang Yao and Zhihong Yu for their very helpful comments and constructive suggestions. Financial support from China's Natural Science Foundation Grant (No. 71003010) is gratefully acknowledged. I thank the editor, Rachel Griffith, and two anonymous referees for their very insightful suggestions. All errors are mine.

firm's processing share grows, input tariff reductions have a smaller impact on productivity gains. Similarly, as firms' processing share increases, the share of domestic sales decreases accordingly; and the pro-competition effects from the reductions in output tariffs are hence weaker.

The current article contributes to the literature in at least three important ways. First, it enriches the understanding of the economic growth of China, the second largest economy and the largest exporter of goods in the world. It is widely believed that China's huge foreign trade volume, a 10% of world trade, is a fundamental cause of the country's rapid economic growth. However, this conjecture is rarely supported by using Chinese micro firm-level data.[1] This study aims to fill in this gap. Using highly disaggregated transaction-level customs data and firm-level production data from 2000–6, the article thoroughly explores the nexus between foreign trade and firm productivity.

Second, processing trade is an important type of trade in many developing countries, such as Indonesia, Mexico and Vietnam. Processing trade is the process by which a domestic firm initially obtains raw materials or intermediate inputs from abroad and, after local processing, exports the value-added final goods (Feenstra and Hanson, 2005). Governments typically encourage processing trade by offering tariff reductions or even exemptions on the processing of intermediate goods. Although there are some studies on trade reform in both developed and developing countries,[2] the interaction between trade reform and processing trade is rarely explored. Hence, understanding the productivity gains from trade reform under the special tariff treatments afforded to processing trade is essential.

Last but not least, aside from adopting the widely accepted method of measuring tariffs at the sector level, I take a step forward to measure both output tariffs and input tariffs at the firm level. Perhaps because of data restrictions, previous studies have usually measured tariffs at the industrial level by using input–output tables, as in Amiti and Konings (2007), or by measuring effective tariff protection as in Topalova and Khandelwal (2011). However, such a convenient approach might face a possible pitfall because input–output tables mix up both imported intermediate inputs and domestic intermediate inputs that are not directly relevant to tariff reductions. Using input–output tables may not accurately measure the level of trade protection faced by firms. Thanks to the rich information covered by both Chinese firm-level production data and transaction-level trade data, I am able to construct novel measures of firm-specific input and output tariffs to estimate the effect of trade reforms on firm productivity. To my knowledge, this is the first attempt to measure tariffs at the firm level in the literature, although it is worthwhile to stress that my estimation results remain robust when using conventional industry-level measures of tariffs.

[1] Brandt *et al.* (2012) is an outstanding exception.

[2] The studies focusing on developed countries, among others, include Bernard *et al.* (2003) for the US and Trefler (2004) for Canada. However, more evidence has been found for developing countries, such as Bustos (2011) for Argentina, Schor (2004) for Brazil, Pavcnik (2002) for Chile, Fernandes (2007) for Colombia, Harrison (1994) for Côte d'Ivoire, Krishna and Mitra (1999) and Topalova and Khandelwal (2011) for India, Amiti and Konings (2007) for Indonesia and Levinsohn (1993) for Turkey. Other research, such as that of Lu *et al.* (2010), Lu (2011) and Ma *et al.* (2011), also explores the nexus between export growth and productivity improvement in China.

I also carefully control for two sets of endogeneity issues of firm-level tariffs and firms' self-selection to processing activities. Several endogeneity problems plague the firm-level input and output tariffs. The first one results from tariff measures themselves. Because a firm may import multiple products, it is useful to construct an import-based weight to reflect the importance of products for the firm. However, imports and tariffs are negatively correlated. In the extreme case, imports and their associated import shares are zero for prohibitive tariffs. As a result, the measure of input tariffs faces a downward bias. To address this endogeneity problem, throughout all the estimation, firm-level tariffs are constructed using time-invariant weights based on the firm's imports in the first year it appears in the sample. The second endogeneity problem relates to a possible reverse causality of tariffs with respect to productivity. Tariffs may be granted in response to domestic special interest groups, the pressure of which could be significant in countries such as India (Topalova and Khandelwal, 2011) or low in countries such as Indonesia (Amiti and Konings, 2007). Given that China acceded to the WTO in 2001, domestic pressure might not have played a key role during 2000–6. However, for the sake of completeness, an (IV) approach is adopted to control for possible reverse causality.

Another set of endogeneity issues is of firms' self-selection to processing activities. Observing that some Chinese firms are involved in both processing and ordinary trade, whereas others are only involved in one type of trade, I measure the processing variable in two ways. First, I use a processing indicator to identify whether a firm engages in processing trade. If a firm imports any products for processing purposes, as revealed in the customs data, such a firm is defined as a processing firm. However, the firm's processing share is endogenous. A firm would first decide whether to engage in processing trade and, if so, the extent to which it will engage in processing imports. To address such self-selection behaviour, I rely on a type-2 Tobit model. In the first-step probit estimates, I find that low-productivity firms self-select to engage in processing trade, possibly to enjoy the free duty on imported intermediate inputs. After obtaining the firm's fitted extent of processing imports from the second-step Heckman estimates, I use it as a measure of the processing indicator in the main estimates of the effects of tariffs on firm productivity to control for the endogeneity of the firm's processing decision. All else being constant, a high degree of engagement in processing trade is shown to reduce firm productivity.

To explore the relationship between firm productivity and output and input tariffs, I follow the standard procedure to investigate the nexus in two steps. First, the firm's total factor productivity (TFP) is measured based on a production function using the methodology of Olley and Pakes (1996), with a number of necessary modifications and extensions to fit the Chinese context. As processing firms and non-processing firms could use different technologies to produce products even within an industry, I estimate firm TFP for processing firms and non-processing firms separately within an industry. I also take the firm's learning from processing trade into account (De Loecker, 2013). Although the augmented Olley–Pakes approach is capable of controlling for the possible simultaneity bias and selection bias caused by regular OLS estimates, it relies on the important assumption that capital is more actively responsive to unobserved productivity. However, China is a labour-abundant country and hence has relatively low labour costs. In the face of a productivity shock, Chinese firms usually adjust their labour input to re-optimise production behaviour (Blomström and Kokko, 1996). Therefore, I adopt three alternative approaches to measure firm TFP:

(*i*) labour productivity;
(*ii*) the Levinsohn–Petrin (2003) TFP; and
(*iii*) the Blundell and Bond (1998) system-GMM TFP.

Given that the system-GMM TFP has an additional advantage in controlling for the role of lagged firm productivity to avoid possible serial correlation in the TFP estimation (Fernandes, 2007), I use it as the main measure of firm TFP.

It is also important to understand the mechanisms through which firm productivity improves in response to trade reforms. Inspired by previous studies, such as Amiti and Konings (2007), Goldberg *et al.* (2010) and Bustos (2011), the impact of input tariffs on productivity is straightforward, as lower tariffs induce a larger variety of inputs. By contrast, the impact of output tariffs on productivity could work directly by pressuring firms to be more productive, and/or indirectly by weeding out less-productive firms. This article finds that the pro-competition effect is mostly through the channels that pressure firms to be more productive, which is in line with the findings of Horn *et al.* (1995). Several possible channels – such as import scope and research and development (R&D) – are also discussed. Unlike Amiti and Konings (2007), my data set includes information that allows the firm's product scope (in export markets) to be directly measured as in Goldberg *et al.* (2010). In addition, similar to Bustos (2011), the analysis takes into consideration information on R&D expenses.

Finally, as economy-wide productivity is an essential measure of a country's welfare, my final step is to add firm productivity to economy-wide productivity by using Domar's (1961) weight, which corrects for possible aggregation bias due to the ignorance of vertical integration in an open economy. In brief, I find that both output and input tariff reductions contribute at least 14.5% to economy-wide productivity growth during the sample period.

The remainder of the article is organised as follows. Section 1 introduces the special tariff treatment on Chinese processing trade. Section 2 describes the unique data used in the analysis. Section 3 discusses key variables and the econometric method. Section 4 presents the empirical evidence. Finally, Section 5 concludes.

1. Special Tariff Treatment on Processing Trade

Processing trade in China began in the early 1980s. As an important means of trade liberalisation, the government encourages Chinese firms to import all or part of the raw materials and intermediate inputs, and re-export final value-added goods after local processing or assembly. As of 2012, the General Administration of Customs reports 16 specific types of processing trade in China.[3]

[3] Such types of processing trade include, among others, foreign aid (code: 12), compensation trade (13), assembly (14), processing with inputs (15), goods on consignment (16), goods on lease (17), border trade (19), contracting projects (20), outward processing (22), barter trade (30), customs warehouse trade (33) and entrepôt trade by bonded area (34).

Among these types of trade, two are the most important, namely, processing with assembly and processing with inputs.[4] Both types of processing trade are duty-free but they are characterised by an important difference. For processing with assembly, a domestic Chinese firm obtains raw materials and parts from its foreign trading partners without any payment. However, after local processing, the firm has to sell its products to the same foreign trading partner by charging an assembly fee. By contrast, for processing with inputs, a domestic Chinese firm pays for raw materials from a foreign seller. After local processing, the Chinese firm can then sell its final goods to other foreign countries.

Figure 1 shows that, compared with ordinary imports, processing imports in China accounted for just a small proportion of total imports in the early 1980s. However, China's processing imports dramatically increased in the early 1990s and began to dominate ordinary imports in 1992, when China officially announced the adoption of a market economy. Going forward, processing imports accounted for more than 50% of the country's total imports. Interestingly, processing imports with assembly were more popular in the 1980s because most Chinese firms lacked the capital needed to import. Since the 1990s, processing imports with inputs have been more prevalent. This trend can be seen clearly in Figure 2: within processing imports, the ratio of processing with assembly over processing with inputs declined from 0.41 in 2000 to 0.32 in 2006.

The primary objective of the current article is to determine how a firm's TFP reacts to output and input tariff reductions in the presence of special tariff treatments on processing trade. Therefore, understanding whether a firm engages in processing activities is important. All Chinese firms are classified into four types, namely, non-importing firms and three types of importing firms: ordinary importers, hybrid processing importers and pure processing importers. As shown in Figure 3,

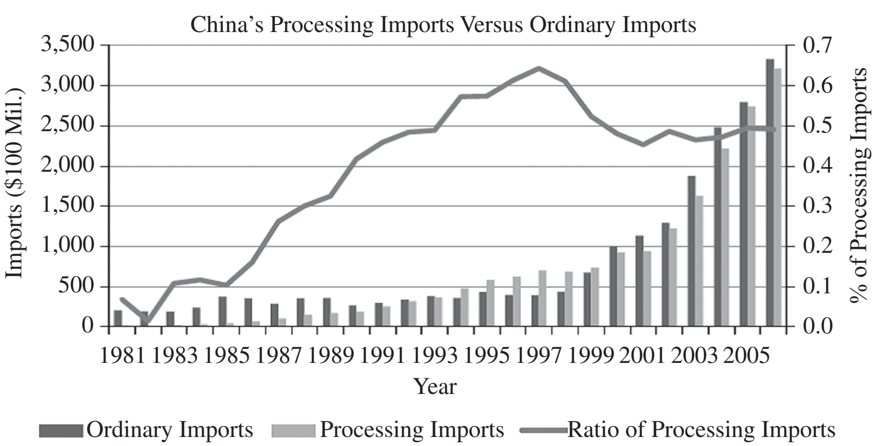

Fig. 1. *China's Processing Imports Versus Ordinary Imports*

[4] Processing with assembly is also referred to as 'processing with supplied materials', as stated in the official customs reports, or 'pure assembly' as adopted in Feenstra and Hanson (2005). Correspondingly, processing with inputs is also referred to as 'processing with imported materials' or 'input and assembly'.

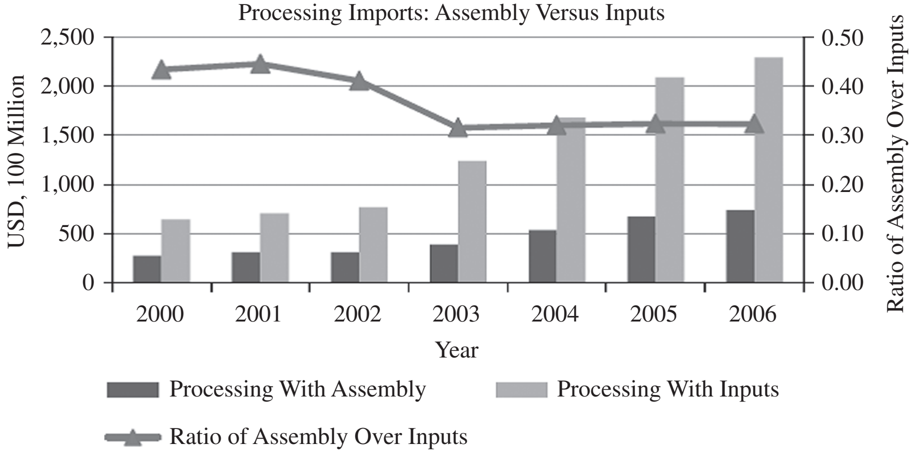

Fig. 2. *China's Processing Imports: Assembly Versus Inputs*
Sources. Customs trade data (2000–6), author's own compilation.

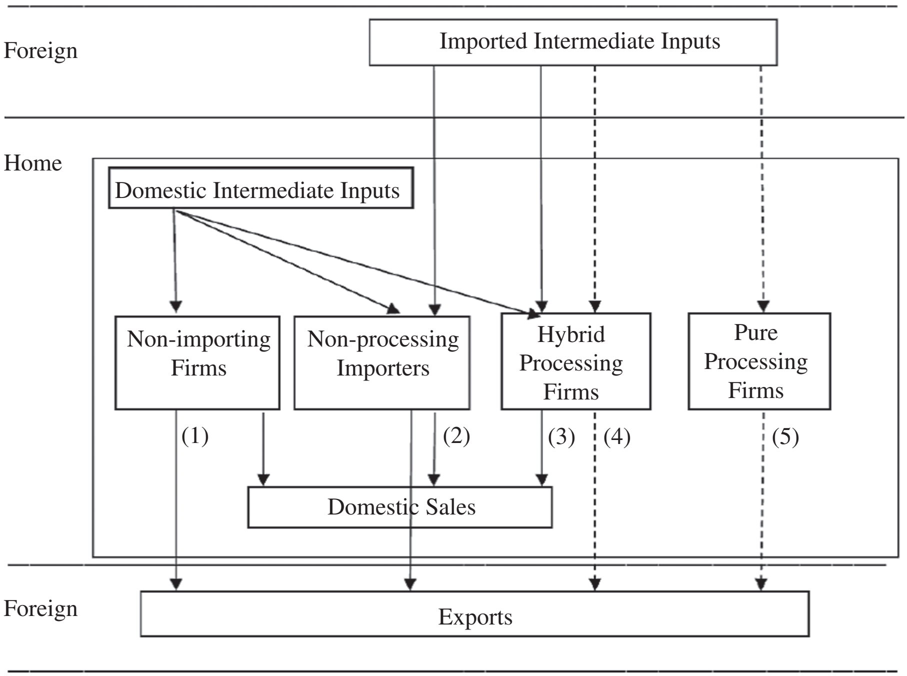

Fig. 3. *Four Types of Chinese Firms*
Note. Dotted lines denote firms' processing imports/exports; solid lines represent firms' non-processing imports/exports.

non-importing firms do not have any imports; all raw materials and intermediate inputs are locally acquired. However, non-importing firms can sell their final goods domestically and internationally (as shown by arrow (1)).

Among the three types of importers, ordinary importers are firms that do not use any processing of imported intermediate inputs, although they import non-processing intermediate inputs and could sell their final goods in both domestic and foreign markets (arrow (2)).[5] In sharp contrast, pure processing importers are firms engaged only in processing activities, shown by the dotted lines in the Figure. Pure processing importers purchase 100% of their raw materials and intermediate inputs abroad and re-export their final value-added goods (arrow (5)). Such firms clearly enjoy the privilege of duty-free imports. Finally, and perhaps the most interesting type of firm, hybrid processing importers engage in both ordinary imports (arrow (3)) and processing imports (arrow (4)). Such firms enjoy free duties for their processing imports, but still pay duties for ordinary imports. Here it is important to stress that the processing trade of both hybrid and pure processing importers could include any processing type, such as assembly and processing with inputs.

2. Data

To investigate the impact of trade liberalisation on firm productivity, I rely on the following three disaggregated, large panel data sets: tariff data, firm-level production data and product-level trade data.

Tariff data can be accessed directly from the WTO and the trade analysis and information system (TRAINS).[6] China's tariff data are available at the Harmonised System (HS) six-digit disaggregated level for 2000–6. Given that the product-level trade data are at the HS eight-digit level, the product-level trade data are aggregated to the HS six-digit level to correspond with the tariff data. As I am interested in measuring the average effect of trade liberalisation on firm productivity, I use the *ad valorem* duty at the six-digit level to measure trade liberalisation.

2.1. *Firm-level Production Data*

The sample is derived from a rich firm-level panel data set that covers between 162,885 firms (in 2000) and 301,961 firms (in 2006). The data are collected and maintained by China's National Bureau of Statistics (NBS) in an annual survey of manufacturing enterprises. Complete information on the three major accounting statements (i.e. balance sheet, profit and loss account, and cash flow statement) is available. In brief, the data set covers two types of manufacturing firms – all state-owned enterprises (SOEs) and non-SOEs whose annual sales exceed RMB 5 million ($770,000).[7] The

[5] Different from processing importers, non-processing importers have to pay import tariffs for their imported intermediate inputs, although such imported goods are possibly used as inputs to produce final exportable goods. The key difference is that non-processing firms cannot show processing contracts/licences to the customs to enjoy the privilege of free duty.

[6] The data are from WTO webpage http://tariffdata.wto.org/ReportersAndProducts.aspx. Note that TRAINS data generally suffer from missing values, particularly regarding the tariffs imposed by other countries for Chinese exports. The product-destination-year combinations that have missing tariffs are hence dropped. All data sets and programmes that allow the replication of the results in the article are available online.

[7] Aggregated data on the industrial sector in the annual *China's Statistical Yearbook* by the NBS are compiled from this data set.

data set includes more than 100 financial variables listed in the main accounting statements of these firms.

Although the data set contains rich information, some samples are still noisy and are therefore misleading, largely because of misreporting by some firms.[8] Following Cai and Liu (2009), I clean the sample and omit outliers by using the following criteria. First, observations with missing key financial variables (such as total assets, net value of fixed assets, sales and gross value of the firm's output productivity) are excluded. Second, I drop firms with fewer than eight workers as they fall under a different legal regime, as mentioned in Brandt *et al.* (2012).

Following Feenstra *et al.* (2013*a*), I delete observations according to the basic rules of the Generally Accepted Accounting Principles (GAAP) if any of the following are true:

(*i*) liquid assets are greater than total assets;
(*ii*) total fixed assets are greater than total assets;
(*iii*) the net value of fixed assets is greater than total assets,
(*iv*) the firm's identification number is missing; or
(*v*) an invalid established time exists (e.g. the opening month is later than December or earlier than January).

After applying such a stringent filter to guarantee the quality of the production data, the filtered firm data are reduced by about 50% in each year, as shown in columns (3) and (4) of Appendix Table A1.

Note that, in China's customs data set, some Chinese firms do not have their own production activity but only export goods collected from other domestic firms or import goods from abroad and then sell them to other domestic companies (Ahn *et al.*, 2010).[9] To ensure the preciseness of the estimates, I exclude such trading companies from the sample in all the estimates. In particular, firms with names including any Chinese characters for Trading Company or Importing and Exporting Company are excluded from the sample.[10]

2.2. *Product-level Trade Data*

The extremely disaggregated product-level trade transaction data are obtained from China's General Administration of Customs. It records a variety of information for each trading firm's product list, including trading price, quantity and value at the HS eight-digit level. More importantly, this rich data set not only includes both import and export data but also breaks down the data into several specific types of processing trade, such as processing with assembly and processing with inputs.

Table 1 reports a simple statistical summary for Chinese product-level trade data by shipment and year for 2000–6. Overall, when focusing on the highly disaggregated HS

[8] For example, information on some family-based firms, which usually have no formal accounting system in place, is based on a unit of one RMB, whereas the official requirement is a unit of RMB 1,000.
[9] Note that in the firm-level production data, a firm's sales to trade intermediaries are accounted for as domestic sales but not exports, following the requirement of the GAAP.
[10] In China, pure trading companies are required to register with a name containing Chinese characters for 'trading company' or 'importing and exporting company'.

Table 1

Chinese Transaction-level Trade Data by Shipment and Year

Imports by shipment	2000	2001	2002	2003	2004	2005	2006	Total
Percentage of number of observations (HS eight-digit)								
Ordinary imports	2.57	3.54	3.77	5.17	6.04	6.80	7.30	35.19
Processing imports with assembly	2.46	2.72	2.37	2.59	2.77	2.79	2.77	18.47
Processing imports with inputs	3.90	4.14	3.57	4.67	5.33	5.74	5.61	32.95
Other types of processing imports	1.42	1.55	1.70	1.71	2.03	2.24	2.77	13.40
Total	10.34	11.95	11.41	14.13	16.16	17.57	18.44	100
Percentage of import value								
Ordinary imports	3.12	3.87	3.71	5.87	7.74	8.86	10.46	43.64
Processing imports with assembly	0.87	0.98	0.98	1.22	1.68	2.11	2.31	10.16
Processing imports with inputs	2.02	2.21	2.39	3.87	5.24	6.52	7.15	29.40
Other types of processing imports	1.01	1.24	1.43	1.93	2.85	3.35	4.99	16.80
Total	7.02	8.30	8.52	12.89	17.51	20.85	24.91	100

eight-digit level, approximately 35% of the 18,599,507 transaction-level observations are ordinary trade and 65% refer to processing trade. Similar proportions are obtained when measuring by trade volume: around 43% of trade volume comprises ordinary trade. Processing with inputs accounts for around 30%, whereas processing with assembly only is around 10%. The remaining 17% represents other types of processing trade, aside from assembly and processing with inputs.

2.3. *Merged Data Set*

Firm-level production data are crucial in measuring TFP, whereas product-level trade transaction data are non-substitutable in identifying a processing firm. However, researchers face some technical challenges in merging the two data sets. Although the data sets share a common variable (i.e. the firm's identification number), the coding system in each data set is completely different.[11] Hence, the firm's identification number cannot serve as a bridge to match the two data sets.

To address this challenge, following Yu and Tian (2012), I use two methods to match the two data sets by using other common variables. First, I match the two data sets by using each firm's Chinese name and year. That is, if a firm has an exact Chinese name in both data sets in a particular year, it should be the same firm.[12] As described carefully in Appendix A, I obtain 83,679 matched firms in total by using the raw production data set and the number is reduced to 69,623 in total by using the more accurate filtered production data set as described above. To increase the number of qualified matching firms as much as possible, I then use another matching technique to serve as a supplement. Namely, I rely on two other common variables to identify the firms: postal code and the last seven digits of the firm's phone number. The rationale is that firms should have a unique phone number within a postal

[11] In particular, the firm's codes in the product-level trade data are at the ten-digit level, whereas those in the firm-level production data are at the nine-digit level, with no common elements inside.

[12] The year variable is necessary as an auxiliary identification variable as some firms could change their name in different years and newcomers could possibly take their original name.

district. Although this method seems straightforward, there are subtle technical and practical difficulties.[13] The detailed merging procedures are explained in Appendix A. After merging both product-level trade data and firm-level production data, I finally obtain 76,823 common trading firms, including both importers and exporters.[14] Briefly, the merged data set accounts for around 40% of the filtered full-sample, firm-level production data set in terms of the number of exporters, and around 53% in terms of export value. By way of comparison, my matching success rate is highly comparable to that in other studies that use the same data sets, such as Ge *et al.* (2011) and Wang and Yu (2012).

How successful is the matching using this technique? Table 2 first compares the merged data and the full-sample customs trade data sets. Of the total 56,459 importing firms in the merged data, ordinary importers account for 38.1% whereas processing importers account for 61.9%. These numbers are close to their counterparts from the full-sample customs data – 27.3% for ordinary importers and 72.7% for processing importers – as shown in the last column of Table 2.[15] The proportions of hybrid processing importers and pure processing importers by year in both the merged data and the full-sample data sets are also reported in the bottom two rows of Table 2.

Given that the firm-level production data set is crucial for the construction of the regressand (i.e. firm TFP), Table 3 shows how much of total sales and total employment are accounted for by the merged data set each year during 2000–6. In particular, the proportion of exports in the merged sample over exports in the full-sample production data varies from 50% to around 58% during the sample period, suggesting that some firms enter and exit in the merged sample that is used for the

Table 2

Merged Importers by Firm Type

	Merged sample								
Percentage	2000	2001	2002	2003	2004	2005	2006	Total	Full sample
Total importers	8.8	9.9	10.6	12.4	19.4	18.0	21.0	100.0	100.0
Ordinary importers	2.4	3.0	3.7	5.0	7.5	7.3	9.1	38.1	27.3
Processing importers	6.4	6.9	6.9	7.4	12.0	10.7	11.8	61.9	72.7
Hybrid processing importers	3.0	3.2	3.5	3.9	5.8	5.3	6.0	30.7	53.0
Pure processing importers	3.4	3.6	3.4	3.5	6.2	5.4	5.9	31.2	19.7

Notes. There are 56,459 importers in total in the matched data whereas 217,372 firm importers are included in the full-sample trade data.

[13] For example, the phone numbers in the product-level trade data include both area phone codes and a hyphen, whereas those in the firm-level production data do not.

[14] Note that in the merged sample shown in column (7) of Appendix Table A1, exports for some firms reported from the customs trade data set are larger than total sales reported from the NBS production data set. I also drop such firms from the sample in column (8) of Appendix Table A1 to guarantee the quality of my merged data set.

[15] Note that the percentages for ordinary importing firms and processing firms in Table 2 are different from the import volumes for ordinary imports and processing imports shown in Table 1, as a processing importing firm (except pure processing firms) usually also has both processing imports and ordinary imports.

Table 3

Firm-level Production Information in Merged Versus Full-sample Data by Year

Types of firms (%)	2000	2001	2002	2003	2004	2005	2006	Average
Sales	23.7	24.0	23.8	24.6	27.8	25.8	28.3	25.5
Exports	51.9	50.1	52.9	50.0	55.2	51.6	57.9	52.8
Number of employees	20.2	20.9	21.6	23.0	26.5	25.5	28.7	23.8

Notes. The values in this panel are the proportions that were obtained by dividing sales/exports/number of employees in the matched data by their counterparts in the full-sample data, respectively. The last column reports the year-average percentage over 2000–6.

Table 4

Comparison of the Merged Data Set and the Full-sample Production Data Set

	Merged data			Full-sample data		
Variables	Mean	Min.	Max.	Mean	Min.	Max.
Sales (RMB 1,000)	150,053	5000	1.57e+08	85,065	5000	1.57e+08
Exports (RMB 1,000)	53,308	0	1.52e+08	16,544	0	1.52e+08
Number of employees	478	8	157,213	274	8	165,878

estimation. The merged data set includes both exporters and importers.[16] Moreover, Table 4 compares the differences between the merged data set and the full-sample firm-level data set. The merged sample has clearly higher means of sales, exports and number of employees than those in the full-sample firm-level data set. These findings suggest that the merged sample is skewed towards large firms. Thus, my findings are valid for large Chinese trading firms.

3. Measures and Empirics

In this Section, I first introduce the measures of the three key variables: firm TFP, firm-specific output tariffs and firm-specific input tariffs. For comparison, I also introduce the measure of industry-specific output and input tariffs. Finally, I discuss my empirical investigation of the effect of tariff reductions on productivity.

3.1. *TFP Measures*

I use the augmented Olley and Pakes (1996) approach to construct measures of Chinese firm-level TFP following Amiti and Konings (2007). Assuming a Cobb–Douglas production function, the usual estimation equation is as follows:

$$\ln Y_{it}^{j} = \beta_0^j + \beta_m^j \ln M_{it}^j + \beta_k^j \ln K_{it}^j + \beta_l^j \ln L_{it}^j + \epsilon_{it}, \tag{1}$$

[16] Around 60% of firms are exporters whereas the other 40% are importers. The merged sample also includes entry and exit of firms. The last paragraph of Appendix A provides more detailed descriptions on this.

where Y_{it}^j, M_{it}^j, K_{it}^j and L_{it}^j refer to firm i's output, materials, capital and labour in industry j in year t, respectively. Traditionally, TFP is measured by the estimated Solow residual which is the difference between the true data on output and the fitted value using the OLS approach. However, the OLS approach suffers from two problems: simultaneity bias and selection bias. At least some shocks to TFP changes could be observed by the firm early enough for it to change its input decisions to maximise profit. Thus, firm TFP could have a reverse endogeneity on firm input choices. Moreover, firms with low productivity that have collapsed and exited the market are excluded from the data set, indicating that the samples used for the regression are not randomly selected, which, in turn, results in estimation bias. Olley and Pakes (1996) successfully provide a semi-parametric approach to address those two biases. Subsequently, numerous studies, such as those by De Loecker (2011, 2013) and De Loecker *et al.* (2012), among others, have modified and tailored their approaches to calculating TFP. In the present article, I adopt the Olley–Pakes approach to estimate and calculate a firm's TFP with some extensions. Appendix B provides the detailed estimation procedure.

First and foremost, I estimate the production function for processing and non-processing firms separately in each industry. The idea is that different industries may use different technologies; hence, firm TFP (denoted TFP^{OP1}) is estimated separately for each industry. Equally important, even within an industry, processing firms (especially those firms engaged in processing with assembly) may use completely different technologies than non-processing firms, given that processing firms with assembly receive only imported material passively without making any profit-maximising input choices (Feenstra and Hanson, 2005). For the non-processing firm TFP estimates, since a non-processing importing firm may or may not export its final goods, I also include an export dummy to allow different TFP realisation between exporting non-processing firms and non-exporting non-processing firms. By the same token, I include an import dummy in the control function to allow different TFP realisation between non-processing importers and non-processing non-importers (but exporters). Note that two such dummies are not necessary for processing firms as, by definition, processing firms must import inputs and sell their products abroad.

Possibly, firms could learn by processing imports. If productivity gains from processing imports occur simultaneously with investment, TFP^{OP1} may have a bias on the estimated capital coefficient. Thus, ignorance of controlling for the effect of the previous period's processing activity on firm productivity may cause another bias of measured productivity. Inspired by De Loecker (2013), as an alternative approach to estimate TFP (denoted by TFP^{OP2}), I consider another control function in which both processing and non-processing firms are pooled together. More importantly, a processing dummy (i.e. a dummy that takes the value one if a firm has any processing imports and zero otherwise) is also incorporated in the control function (see Appendix B for details). This is done because processing imports may affect firm productivity and, accordingly, the TFP trajectory of a processing firm is endogenously different compared with the trajectory of a non-processing firm.

Second, I use deflated prices at the industry level to measure TFP. The measured TFP is expected to capture the firm's true technical efficiency only. However, here the measured TFP is also likely to pick up differences in price, price-cost markups and even input usage across firms (De Loecker, 2011; De Loecker and Warzynski, 2012). Admittedly, an ideal way to remove price differences across firms would be to adopt firm-specific price deflators (Foster *et al.*, 2007). However, as in many other studies, such price data are unavailable.[17] Following De Loecker *et al.* (2012), I use the industrial price to deflate the firm's output.[18] Turning to the issue of price-cost markups, as stressed by Bernard *et al.* (2003), once the price-cost markup is positively associated with true efficiency, even revenue-based productivity can work well to capture the true efficiency, as is done with physical efficiency.

Third, I take China's WTO accession in 2001 into account, as such a positive demand shock would push Chinese firms to expand their economic scales, which, in turn, would exaggerate the simultaneous bias of their measured TFP. In particular, a WTO dummy (i.e. equal to one after 2001 and zero otherwise) is included in the estimation of the capital coefficient, as discussed in Appendix B.

Fourth, the prevalence of SOEs also affects firm productivity. SOEs in China are usually accompanied by state intervention and do not necessarily make profit-maximising choices (Hsieh and Klenow, 2009). Therefore, it is important to construct an SOE indicator and add it to the control function in the first-step Olley–Pakes estimates.[19]

Finally, it is necessary to construct a real investment variable when using the Olley and Pakes (1996) approach. I adopt the perpetual inventory method as the law of motion for real capital and real investment. Nominal and real capital stocks are constructed as in Brandt *et al.* (2012). Rather than assigning an arbitrary number for the depreciation ratio, I use the exact firm's real depreciation provided by the Chinese firm-level data set. Appendix Table B1 presents the estimated coefficients for the production function and the associated log of TFP by industry for processing firms and non-processing firms, respectively. The implied scale elasticities are quite close to constant returns-to-scale elasticities for both processing firms and non-processing firms within each industry.

The augmented Olley–Pakes approach assumes that capital responds to the unobserved productivity shock with a Markov process, whereas other input factors respond without any dynamic effects. However, labour may also be correlated with an unobserved productivity shock. As highlighted by Ackerberg *et al.* (2006), it is unlikely that there is enough variation left to identify the labour coefficient by using the Olley–Pakes approach. This consideration may fit China's case more closely, given that the country is labour abundant. When facing an unobserved productivity shock, firms might re-optimise their production behaviour by adjusting their labour rather than

[17] The customs trade data provide information on unit-value, which could serve as a proxy for the price for each imported good. However, the prices of imported intermediate inputs could be much different from those of domestic intermediate inputs (Helpern *et al.*, 2010). Using the imported intermediate inputs as a proxy for all intermediate inputs may generate another unnecessary estimation bias. This bias may be exaggerated when the scope of domestic inputs is much different from the scope of foreign inputs.

[18] As in Brandt *et al.* (2012), the output deflators are constructed using 'reference price' information from China's Statistical Yearbooks, whereas input deflators are constructed based on output deflators and China's national input–output table (2002).

[19] By the official definition reported in the *China City Statistical Yearbook* (2006), SOEs include firms such as domestic SOEs (code: 110), state-owned joint venture enterprises (141), and state-owned and collective joint venture enterprises (143) but exclude state-owned limited corporations (151). Appendix Table C3 presents the transitional probability for all SOEs.

their capital. I use the Blundell and Bond (1998) system-GMM approach to capture the dynamic effects of other input factors. By assuming that the unobserved productivity shock depends on a firm's previous periods realisations, the system-GMM approach models TFP as affected by all types of inputs in both current and past realisations.

In particular, this model has the following dynamic representation:

$$
\begin{aligned}
\ln y_{it}^{j} = & \gamma_0^j + \gamma_1^j \ln L_{it}^j + \gamma_2^j \ln L_{i,t-1}^j + (\gamma_3^j \ln L_{it}^j + \gamma_4^j \ln L_{i,t-1}^j) PE_{it} \\
& + \gamma_5^j \ln K_{it}^j + \gamma_6^j \ln K_{i,t-1}^j + (\gamma_7^j \ln K_{it}^j + \gamma_8^j \ln K_{i,t-1}^j) PE_{it} \\
& + \gamma_9^j \ln M_{it}^j + \gamma_{10}^j \ln M_{i,t-1}^j + (\gamma_{11}^j \ln M_{it}^j + \gamma_{12}^j \ln M_{i,t-1}^j) PE_{it} \\
& + \gamma_{13}^j \ln y_{i,t-1}^j + \gamma_{14}^j \ln y_{i,t-1}^j PE_{it} + \gamma_{15} PE_{it} + \varsigma_i + \zeta_t + \omega_{it},
\end{aligned} \tag{2}
$$

where ς_i is firm i's fixed effect, ζ_t is the year-specific fixed effect, and PE_{it} is a processing indicator that takes the value one if a firm has any processing imports and zero otherwise. The idiosyncratic term ω_{it} is serially uncorrelated if no measurement error exists.[20] Consistent estimates of the coefficients in the model can be obtained by using a system-GMM approach. The idea is that labour and material inputs are not taken as exogenously given but are instead allowed to change over time as capital grows. Appendix Table C1 presents the estimated coefficients for system-GMM firm TFP by industry.[21] Overall, the estimated log TFP increases 0.17 log points (from 2.28 in 2001 to 2.45 in 2006), registering a 2.62% annual growth rate, which is very close to the findings in Brandt *et al.* (2012).

3.2. *Firm-specific Tariffs*

A firm could produce multiple products and, thus, its productivity could be affected by multiple tariff lines. Hence, it is important to properly measure the input tariff level faced by firms. As mentioned above, processing imports are duty-free in China. Given that a firm could engage in both processing imports (P) and non-processing imports (O), I construct a firm-specific input tariff index (FIT_{it}) as follows:

$$
FIT_{it} = \sum_{k \in O} \frac{m_{i,initial\_year}^k}{\sum_{k \in M} m_{i,initial\_year}^k} \tau_t^k, \tag{3}
$$

where $m_{i,initial\_year}^k$ is firm i's imports of product k in the first year the firm appears in the sample. Note that $O \cup P = M$ where M is the set of the firm's total imports. The set of processing imports does not appear in (3) because processing imports, again, are duty-free. The firm's input tariffs are constructed by using time-invariant weights to avoid the well-known endogeneity of weighted tariffs: imports are negatively associated with tariffs. For products with prohibitive tariffs, their imports and the associated import share would be zero. Accordingly, if the import weight is measured in the

[20] As discussed by Blundell and Bond (1998), even if transient measurement error exists in some of the series (i.e. $\omega_{it} \sim$ MA(1)), the system-GMM approach can still provide consistent estimates of the coefficients in (2).

[21] Appendix Table C1 reports the associated specification tests for system-GMM estimates including AR(1) and AR(2) tests and Hansen over-identification tests. For most Chinese two-digit level industries, the system-GMM estimates have first-order serial autocorrelation but not second-order serial autocorrelation. The Hansen over-identification tests also suggest that the instruments are valid for most industries.

current period, the measure of firm tariffs would face a downward bias. Therefore, following Topalova and Khandelwal (2011), I measure the import weight for each product using data for the firm's first year in the sample.

Turning to the construction of firm-level output tariffs, product-level domestic sales would be an ideal proxy for capturing the role of each product within a firm. However, such data are unavailable. Hence, I rely on an index to circumvent this data restriction. As a more productive firm is not only capable of selling its products domestically, but also internationally (Melitz, 2003), a product would, in general, be sold domestically if it is sold abroad. Assuming a product is sold domestically and internationally in the same proportions, I consider a following weighted output tariff index (FOT_{it}) for firm i in year t:

$$FOT_{it} = \sum_k \left(\frac{X^k_{i,initial\_year}}{\sum_k X^k_{i,initial\_year}} \right) \tau^k_t, \tag{4}$$

where τ^k_t is the *ad valorem* tariff of product k in year t. The ratio in the parentheses is the value weight of product k, measured by the firm's exports of product k in its initial year in the sample, $X^k_{i,initial\_year}$, over the firm's total exports in the initial year.[22] Inspired by Topalova and Khandelwal (2011), exports for each product are fixed at the initial period to avoid possible reverse causality in firm productivity with respect to measured output tariffs.

This measure suffers from two important caveats. First, a firm may sell a product at home but not abroad (i.e. it is a pure domestic firm), which could be fairly reasonable as recent studies show that multi-product firms often sell different products at home and abroad (Bernard *et al.* 2011; Arkolakis and Muendler, 2012). In this case, the export weight for such a product in (4) is zero and the firm's output tariff measure fails to capture any pro-competition effects. This argument also holds for pure exporting firms that sell their products abroad only (around 12.2% of firms are pure exporters in my matched data). To ensure that my main estimation results are not biased by such firms, I drop pure domestic firms and pure exporting firms from the sample in all regressions.

Second, the exported and domestic shares of a product are assumed to be equal. Note that this is a strong assumption indeed as the product composition of exports may be very different from the composition of domestic sales. This is especially true for China, which holds an important position in global supply chains (GSCs) and produces some intermediates that cannot be used in the domestic production sector.[23]

[22] Alternatively, the weighted output tariff index can be written as

$$FOT_{it} = \sum_k \left[v^k_{i,initial\_year} \Big/ \left(\sum_k v^k_{i,initial\_year} \right) \right] \tau^k_t$$

and the domestic value of product k for firm i is

$$v^k_{i,initial\_year} = \left(X^k_{i^k_{i,initial\_year}} \Big/ \sum_k X^k_{i,^k_{i,initial\_year}} \right) \left(Y_i - \sum_k X^k_{i,^k_{i,initial\_year}} \right),$$

where Y_i is firm i's total sales in its initial year. Therefore, the difference enclosed by the second parentheses measures firm i's total domestic sales.

[23] Besides, when firms sell in both the domestic and export markets, the quality of the products is likely to be different, with better quality products sold to the export markets. As data on unit-price, a common proxy of product quality, are unavailable for domestic products, here I am not able to distinguish the quality difference between domestic products and exportable products, which is a future research topic once data are available. I thank a referee for correctly pointing this out.

Table 5

China's Output Tariffs and Input Tariffs by Year

	Firm output tariffs		Firm input tariffs		Industry output tariffs		Industry input tariffs	
Year	Mean (1)	SD (2)	Mean (3)	SD (4)	Mean (5)	SD (6)	Mean (7)	SD (8)
2000	15.57	12.03	2.54	4.90	21.43	8.78	3.00	3.63
2001	12.39	9.40	2.37	5.06	17.77	6.07	2.98	3.78
2002	9.63	8.22	1.68	3.53	14.28	6.05	1.41	1.66
2003	8.82	7.51	1.94	3.70	12.46	5.21	0.41	0.27
2004	7.59	7.08	1.87	3.59	11.27	4.60	0.36	0.25
2005	7.00	6.78	1.71	3.53	10.49	4.46	0.34	0.21
2006	7.46	6.46	2.18	3.72	10.27	4.20	0.35	0.18
All years	8.29	7.65	1.98	3.82	11.88	5.63	0.69	0.15

Notes. Columns (1)–(4) report the mean and standard deviation of firm output tariffs and firm input tariffs with initial time-invariant weights as described in (4) and (3), in the text. Columns (5) and (6) report the mean and standard deviation of industry-level output tariffs and columns (7)–(8) report the mean and standard deviation of industry-level input tariffs that are constructed using the 2002 input–output table for China.

Because of data restrictions, I am not able to check this out directly. However, as this problem would bias the measure of firm output tariffs differently depending on the industry and depending on the intensity of the sector of processing firms, I run further regressions by distinguishing more integrated industries from less integrated industries and by separating the sample by the intensity of the sector in processing firms. As shown in the text later, all such robustness checks suggest that my main results are still valid even considering such within-firm differences in product composition.

Columns (1)–(4) in Table 5 report firm-specific input and output tariffs computed using (3) and (4), respectively. The average firm-specific output tariffs were cut in half from around 15.6% in 2000 to 7.4% in 2006, and their standard deviation also dropped by around 50% over the same period. Firm-specific input tariffs are much lower than output tariffs. Input tariffs also exhibit a sharp declining trend during the sample period.

3.3. *Industry-specific Tariffs*

Similar to Amiti and Konings (2007), the sector output tariffs at the two-digit Chinese industry classification (CIC) level are obtained by taking a simple average of the HS six-digit codes within each two-digit CIC industry code.[24] The industry-level input tariff index is measured by

$$IIT_{ft} = \sum_n \left(\frac{input_{nf}^{2002}}{\sum_n input_{nf}^{2002}} \right) \tau_{nt}, \tag{5}$$

[24] The reason for not using weighted import tariffs, again, is to avoid the endogeneity of tariffs: imports are negatively correlated to tariffs.

Table 6
Simple Correlations of China's Output Tariffs and Input Tariffs

	Firm output tariffs	Firm input tariffs	Industry output tariffs	Industry input tariffs
Firm output tariffs	1.000			
Firm input tariffs	0.092	1.000		
Industry output tariffs	0.477	−0.073	1.000	
Industry input tariffs	0.328	−0.062	0.578	1.000

where IIT_{ft} denotes the industry-level input tariffs facing firms in industry f in year t. τ_{nt} is the import tariff of input n in year t. The weight in parentheses is measured as the cost share of input n in the production of industry f, for which data can be obtained from by China's input–output table for 2002.[25]

As shown in columns (5)–(8) in Table 5, the information in these columns is in line with that obtained by using the firm-level tariffs in columns (1)–(4): both output and input tariffs dramatically fell over the sample period. Similar patterns can be found from their standard deviations. Firm-specific output tariffs seem to be lower than industrial output tariffs. In sharp contrast, firm-specific input tariffs are higher than industry-specific input tariffs. One possible reason for the under-measurement of industrial input tariffs is that the inclusion of non-importing firms in intermediate input industries biases the industrial input weight in (5) which does not show up in the corresponding firm-specific input tariffs.[26] The simple correlations reported in Table 6 confirm this point: industry-specific input tariffs are only weakly correlated to firm-specific input tariffs ($|corr.| = 0.06$), whereas industry-specific output tariffs are strongly correlated to firm-specific output tariffs, as expected ($|corr.| = 0.48$).

3.4. *Empirical Specification*

To investigate the effects of input and output tariff reductions on firm productivity, I consider the following empirical framework:

$$\begin{aligned}\ln TFP_{it} = &\beta_0 + \beta_1 FOT_{it} + \beta_2 FOT_{it} \times PE_{it} + \beta_3 FIT_{it} + \beta_4 FIT_{it} \times PE_{it} \\ &+ \beta_5 PE_{it} + \boldsymbol{\theta}\mathbf{X}_{it} + \varpi_i + \eta_t + \mu_{it,}\end{aligned} \tag{6}$$

[25] China's input–output table is compiled every five years; the most recent updates were in 2007. As my data sample is between 2000 and 2006, I adopt the input–output table from 2002. In particular, I proceed with the following steps to calculate the industry-specific tariffs. As there are 71 manufacturing sectors reported in China's input–output table (2002) and only 40 manufacturing sectors reported in the CIC, the first step is to find the correspondence between sectors in the input–output table and the CIC. The second step matches the CIC sectors with the International Standard Industrial Classification (ISIC, rev. 3). Note that China's government adjusted its CIC in 2003. I make the same adjustment in the sample. The third step is to link the ISIC and the HS six-digit classification to find the corresponding tariffs from the WTO. The final step calculates the average industry-level tariffs, which are aggregated to the CIC sector level.

[26] For example, if firm i in industry f uses 50% lumber with 1% tariffs and 50% steel with 10% tariffs, then the firm-specific input tariff is 5.5%. However, if industry f uses more domestic lumber, the industrial weight of lumber increases to 70%. Accordingly, the industry-specific input tariffs are reduced to $0.7 \times 1\% + 0.3 \times 10\% = 3.7\%$, which is significantly lower than its counterpart of firm-specific input tariffs.

where $\ln TFP_{it}$ is the logarithm of firm i's measured TFP in industry j in year t, whereas FIT_{it} and FOT_{it} denote firm-level input tariffs and output tariffs as measured in (3) and (4), respectively. The augmented Olley–Pakes TFP is adopted for the baseline estimates, but the system-GMM TFP is adopted as the main measure, given that it enjoys rich, measured flexibility. PE_{it} is a processing indicator that equals one if firms import any processing products in year t, and zero otherwise. An interaction term between the firm's output (input) tariff and the processing indicator is also included to capture a possible heterogeneous effect of output (input) tariff reductions on firm productivity between processing and ordinary firms.

In addition, β_5 in (6) measures other possible gains from processing trade not caused by trade liberalisation. $\mathbf{X}_{it}$ denotes other firm characteristics, such as type of ownership (i.e. SOEs or multinational firms). SOEs are traditionally believed to have relatively low economic efficiency and, hence, low productivity (Hsieh and Klenow, 2009). By contrast, multinational firms have higher productivity in part because of international technology spillovers (Keller and Yeaple, 2009) or fewer financial constraints (Manova *et al.*, 2009). Therefore, I construct two indicators to measure the roles of SOEs and multinational firms. In particular, a firm is classified as a foreign firm if it has any investments from other countries (regimes). A large proportion of the inflow of foreign investment comes from Hong Kong/Macao/Taiwan, so these investments are considered in the construction of such an indicator.[27] As a result, 77% of trading firms are classified as multinational affiliates.[28] Similarly, I construct an indicator for SOEs, which is one if a firm has any investment from the government, and zero otherwise.

Finally, the error term is divided into three components:

(*i*) firm-specific fixed effects ϖ_i to control for time-invariant but unobservable factors such as managerial ability;
(*ii*) year-specific fixed effects η_t to control for firm-invariant factors such as an appreciation of the *renminbi* (RMB); and
(*iii*) an idiosyncratic effect μ_{it} with normal distribution $\mu_{it} \sim \mathrm{N}(0, \sigma_i^2)$ to control for other unspecified factors.

However, the empirical specification above faces an identification challenge. The processing indicator in (6) is a relatively crude measure of processing activity, which may overestimate the role of processing firms. For example, if a firm has only a very small proportion of processing imports over total imports, it is still classified as a processing firm, yet its primary operation remains in ordinary trade. To overcome this challenge, I consider a continuous measure of the extent to which a firm is engaged in

[27] Specifically, foreign-invested enterprises (FIEs) include the following firms: foreign-invested joint-stock corporations (code: 310), foreign-invested joint venture enterprises (320), fully FIEs (330), foreign-invested limited corporations (340), Hong Kong/Macao/Taiwan (henceforth, H/M/T) joint-stock corporations (210), H/M/T joint venture enterprises (220), fully H/M/T-invested enterprises (230) and H/M/T-invested limited corporations (240). Appendix Table C4 presents the transitional probability for such foreign firms.

[28] At first glance, these ratios are significantly higher than their counterparts reported in other studies, such as Feenstra *et al.* (2013*a*). However, this finding simply reflects the fact that the present article covers only large trading firms. Large, non-trading firms have been excluded.

processing trade to replace the processing indicator, and the extent of processing engagement ($Pext_{it}$) is measured through firm i's total processing imports over total imports in year t. In particular, I consider the following specification for my main estimation:

$$\begin{aligned}\ln TFP_{it} = \beta_0 + \beta_1 FOT_{it} + \beta_2 FOT_{it} \times Pext_{it} + \beta_3 FIT_{it} + \beta_4 FIT_{it} \times Pext_{it} \\ + \beta_5 Pext_{it} + \boldsymbol{\theta}\mathbf{X}_{it} + \varpi_i + \eta_t + \mu_{it}.\end{aligned} \tag{7}$$

Yet, a new identification challenge arises from the coefficients of the variable $Pext_{it}$ itself and its interaction terms: β_2, β_4 and β_5. These coefficients differ across industries as different industries use different technologies (Pavcnik, 2002). More importantly, even within an industry, the decision to engage in processing trade is endogenous to firms. Previous works, such as Dai *et al.* (2012), find that less-productive firms self-select to engage in processing trade. If so, a firm's extent of processing engagement is also endogenous as firms with a high extent of processing engagement may be less productive. That is, β_2, β_4 and β_5 vary across firms. My estimating equation thus has random coefficients that are correlated with the endogenous extent of processing engagement, so it is a correlated random coefficients (CRC) model (Wooldridge, 2008).

Heckman and Vytlacil (1998) recommend replacing the endogenous variable in a CRC model – or the extent of processing engagement in my case – with its predicted value.[29] In the next Section, I estimate the extent of processing engagement with a Heckman procedure, or type-2 Tobit model, using the exogenous variables $\mathbf{Z}_{it}$ which is be specified in the next Section. In particular, I have

$$Pext_{it} = \mathrm{E}(Pext_{it}|\mathbf{Z}_{it}) + \epsilon_{it}, \text{ with } \mathrm{E}(\epsilon_{it}|\mathbf{Z}_{it}) = 0. \tag{8}$$

By substituting (8) into (7), I obtain:

$$\begin{aligned}\ln TFP_{it} = \; & \beta_0 + \beta_1 FOT_{it} + \beta_2 FOT_{it} \times \mathrm{E}(Pext_{it}|\mathbf{Z}_{it}) + \beta_3 FIT_{it} \\ & + \beta_4 FIT_{it} \times \mathrm{E}(Pext_{it}|\mathbf{Z}_{it}) + \beta_5 \mathrm{E}(Pext_{it}|\mathbf{Z}_{it}) \\ & + \boldsymbol{\theta}\mathbf{X}_{it} + \varpi_i + \eta_t + \varepsilon_{it},\end{aligned} \tag{9}$$

where the error term is $\varepsilon_{it} = (\beta_2 FOT_{it} + \beta_4 FIT_{it} + \beta_5)\varepsilon_{it} + \mu_{it}$.[30] All the terms appearing within this error have zero expected value conditional on $\mathbf{Z}_{it}$, so that ε_{it} is conditionally uncorrelated with these exogenous variables and they can be used for estimation. Finally, as suggested by Wooldridge (2008), a correction to the standard errors must be made to reflect the use of estimated regressors in (9), which I implement by bootstrapping.

[29] Feenstra *et al.* (2013*a*) also apply this method to estimate the impact of credit constraints on firm's exports.

[30] Similar to Heckman and Vytlacil (1998), the conditional homoscedasticity of covariance assumption for the term $\varepsilon_{it}\mu_{it}$ is needed to ensure that it would not bias the estimates.

4. Estimation Results

4.1. *Baseline Results*

As described above, the merged data set is skewed towards large trading firms, which are the main focus of the present article. Still, it is worthwhile checking whether the relatively high attrition rate of the merged data set affects the estimation results. Hence, my estimation begins with a comparison between the full-sample data set and the merged data set.

I start off the estimation in Table 7 by using conventional industry-level tariffs, as introduced in subsection 4.3. Columns (1) and (2) first run regressions using full-sample firm data. As processing information is not included in the full-sample firm data, it is ignored in the estimation. As firms in different industries would adopt different technologies, it would be inappropriate to combine firms across all industries without controlling for industrial differences (Pavcnik, 2002). Therefore, I control for industry-level fixed effects at the two-digit CIC level in the estimates in column (1). It turns out that both industrial output tariffs and input tariffs are negatively and statistically significantly correlated with firm productivity, which is consistent with the findings of many other studies. Column (2) takes a step forward to control for firm-specific fixed effects and year-specific fixed effects. The coefficient of industry output tariffs is still negative and significant. Strikingly enough, the coefficient of industry input tariffs is positive. However, this is not a worry as the coefficient is statistically insignificant. One possible reason for such an unanticipated finding is the inclusion of non-importing firms that appeared in the full-sample firm data set but did not directly benefit from reductions in tariffs on the imported intermediate inputs.

The rest of the regressions reported in Table 7 use the merged data set, which only includes large trading firms. For a close comparison with columns (1) and (2), the

Table 7
Benchmark Estimates for Comparisons

	Full-sample data set		Merged data set	
Regressand: $\ln TFP_{ijt}^{OP}$	(1)	(2)	(3)	(4)
Industry output tariffs	−0.563**	−0.264***	−0.601***	−0.154*
	(−2.77)	(−8.42)	(−5.09)	(−1.91)
Industry input tariffs	−2.54**	0.133	−1.46***	−1.45***
	(−4.97)	(0.93)	(−4.08)	(−3.53)
Industry-specific fixed effects	Yes	No	Yes	No
Firm-specific fixed effects	No	Yes	No	Yes
Year-specific fixed effects	No	Yes	No	Yes
Observations	315,416	315,416	82,570	82,570
Prob. > F	0.000	0.000	0.000	0.000
R^2	0.21	0.13	0.34	0.02

Notes. t-values are in parentheses. Significant at *10%, **5% and ***1%. Regressions in columns (1) and (2) use the entire sample for Chinese firms (2000–6), whereas those in columns (3) and (4) use the matched sample for Chinese trading firms (2000–6). Regressions in columns (1) and (3) are clustered at the two-digit Chinese industry level. Industry input tariffs are calculated by using the 2002 time-invariant input–output matrix for China as described in (5) in the text. Regressions in columns (1) and (3) are clustered at the one-digit industry level.

estimates in column (3) control for industry-level fixed effects, whereas those in column (4) control for firm-specific and year-specific fixed effects. The coefficients of both industry output tariffs and input tariffs are found to be negative and significant.[31]

I include the processing indicator (i.e. one if a firm has any processing imports and zero otherwise) in the first three columns of Table 8, given that processing information is available in the merged data set. To check whether the estimation results are sensitive to different TFP measures, column (1) uses TFP^{OP1} in which the productivities of processing firms and non-processing firms are estimated using different control functions, whereas column (2) uses TFP^{OP2} in which productivities of processing firms and non-processing firms are jointly estimated as the regressand. In addition, columns (1) and (2) abstract from the interaction term between output (input) tariffs and the processing indicator. After controlling for firm-specific and year-specific fixed effects, both industry output tariffs and industry input tariffs are negatively correlated with firm productivity. Their coefficients are statistically significant. Meanwhile, the coefficient of the processing indicator is negative and significant, indicating that processing firms have low productivity.

However, the Olley–Pakes TFP measure that is used in columns (1) and (2) of Table 8 still suffers from three possible pitfalls. First, the Olley–Pakes approach does not allow output to exhibit any serial correlation, which is likely. Second, it assumes that firms will mostly adjust their capital usage when facing an exogenous shock. However, this may not be the case for China, given that Chinese firms are able to access relatively cheap labour. Finally, there are many missing values for investment in the Chinese firm data, which are essential for computing the Olley–Pakes TFP.[32] By way of comparison, the system-GMM TFP measure is better at overcoming such pitfalls: It has enough flexibility to allow for possible serial autocorrelation and to allow firms to adjust all inputs including not only capital, but also labour and materials. In addition, the computation of system-GMM TFP no longer relies on investment as a proxy variable. I therefore use the system-GMM TFP as the main measure of firm productivity from column (3) of Table 8 to the rest estimates in the article.

To examine the possibly heterogenous impact of tariff reductions on firm productivity, column (3) of Table 8 includes interaction terms for the processing indicator and industry output and input tariffs. The coefficients of output tariffs and input tariffs themselves and their interaction with the processing indicator are still statistically significant. However, the processing indicator exhibits an erratic sign, although it is insignificant. I suspect this is because the processing indicator is a relatively crude measure of processing activity, which may overestimate the role of processing firms. For example, if a firm has only a very small proportion of processing imports over total imports, it is still classified as a processing firm, yet its primary operation remains in ordinary trade. I then consider a continuous measure of the extent to which a firm is engaged in processing trade to replace the processing indicator in the rest of Table 8; the extent of processing engagement is measured by the firm's total processing imports over total imports each year.

[31] As in common, the R^2 in all estimates with firm-specific and year-specific fixed effects in the artcle is exclusive of both firm-specific and year-specific dummies.

[32] Around 40% of the observations are missing investment data.

Table 8

Preliminary Estimates

Tariffs measure:	Industry tariffs			Firm tariffs		
Processing measure:	Processing dummy			Extent of processing imports		
Regressand:	ln TFP_{ijt}^{OP1} (1)	ln TFP_{ijt}^{OP2} (2)	ln TFP_{ijt}^{GMM} (3)	ln TFP_{ijt}^{GMM} (4)	ln TFP_{ijt}^{GMM} (5)	ln TFP_{ijt}^{GMM} (6)
Output tariffs	−0.161**	−0.715***	−1.010***	−1.074***	−1.069***	−0.315***
	(−1.98)	(−12.53)	(−25.17)	(−11.20)	(−9.92)	(−4.61)
Output tariffs × processing variable			−0.099*	−0.614***	−0.604***	−0.234***
			(−1.79)	(−4.92)	(−4.21)	(−2.69)
Input tariffs	−1.468***	−1.332***	−0.656***	−1.667***	−1.379**	−0.572***
	(−3.57)	(−5.19)	(−5.13)	(−2.90)	(−2.26)	(−5.37)
Input tariffs × processing variable			0.561***	2.233***	2.251***	2.409***
			(3.26)	(3.56)	(3.33)	(8.01)
Processing variable	−0.010*	−0.011**	0.001	−0.097***	−0.077***	−0.180***
	(−1.76)	(−2.53)	(0.14)	(−6.66)	(−4.62)	(−17.54)
Year-specific fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm-specific fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Pure domestic firms	Yes	Yes	Yes	Yes	No	No
Pure exporting firms	Yes	Yes	Yes	Yes	No	No
Observations	82,558	82,314	97,299	35,172	24,457	27,679
R^2	0.02	0.01	0.03	0.12	0.12	0.09

Notes. Robust t-values are in parentheses. Significant at *10%, **5% and ***1%. Regressions in columns (1)–(5) use industry-level output tariffs and input tariffs, which are calculated using the 2002 time-invariant input–output matrix for China as described in (5) in the text. Regressions in column (6) use firm-specific output tariffs and input tariffs which are computed using the time-invariant weight in the initial period that the firm first appears in the data set. Columns (1)–(3) use a processing dummy (one if a firm has any processing imports and zero otherwise), whereas columns (4)–(6) use the extent of processing imports as a proxy for the processing variable. Regressands in columns (1)–(2) are Olley–Pakes TFP with different first-step control functions as introduced in Appendix B, whereas those in columns (3)–(6) are system-GMM TFP.

Column (4) of Table 8 gives the results of a regression of system-GMM firm TFP on industry-level input and output tariffs. The coefficients of the output and input tariffs are still negative and statistically significant. The variable for the extent of processing imports turns out to be negative and significant. As one of the novel measures of the present article is firm-specific output and input tariffs, I now turn to compare the estimation results using industry-level tariffs and firm-level tariffs. Because firm-specific output tariffs, as introduced in (4), cannot apply to pure domestic firms or pure exporting firms, I drop such firms in column (5) with measures of industry-level output and input tariffs and in column (6) with measures of firm-specific output and input tariffs for comparison.

The coefficients of output (input) tariffs in columns (5) and (6) are all negative and statistically significant. In terms of economic magnitudes, the differences in the coefficients of output (input) tariffs between the two columns are sizable. When moving from the industry-level measure of output tariffs in column (5) to the firm-specific measure of input tariffs in column (6), the coefficient is reduced from −1.07 to −0.32. Likewise, the point estimate of the input tariffs is reduced more than half moving from the measure of industrial input tariffs to the measure of firm-specific input tariffs.

Such sizable differences indicate the pitfalls of using industry-level measures of tariffs. First, output tariff reductions for some products in an industry are not directly relevant to a firm in the same industry if the firm never produces such products. Thus, the pro-competitive effects would be overestimated if output tariffs were measured at the industry level. By the same token, the cost-saving effects of cutting input tariffs are also overstated with the industry measure of input tariffs. Second, compared with output tariffs, the estimation bias for input tariffs could be more severe as the industry measure of input tariffs is also contaminated by the use of an input–output matrix, which also mixed up both imported intermediate inputs and domestic intermediate inputs that are not directly relevant to the cut in tariffs. Finally, ignorance of the 'free-duty' phenomenon for processing imports generates an additional measurement error in industrial input tariffs for Chinese firms. To avoid such possible estimation bias, I use a firm-specific measure of tariffs in the rest of the article.

4.2. *Self-selection to Processing*

Columns (4) and (5) of Table 8 use the extent of processing imports and its interaction with output and input tariffs, but the processing imports variable is endogenous. As shown in column (1) of Table 8, processing firms are associated with low productivity. Thus, it is interesting to compare the TFP trajectories of processing firms with those of non-processing firms. As shown in the last column of Table 9, processing firms, overall, are less productive than non-processing firms. Interestingly, the productivity difference between processing and non-processing firms roughly decreases over the years, suggesting that a catching-up process of processing firms may take place.[33] Such

[33] Appendix Table C5 also reports the transitional probability for processing firms. The switching of processing firms is an interesting topic for future research, although it is beyond the scope of the present article.

Table 9

TFP Trajectories of Processing Versus Non-processing Firms by Year

Firm productivity ($\ln TFP_{ijt}^{GMM}$)	2001	2002	2003	2004	2005	2006	Overall
Non-processing firms	2.458	2.465	2.518	2.544	2.585	2.625	2.576
Processing firms	2.416	2.432	2.462	2.539	2.575	2.629	2.551
Difference	0.042***	0.033***	0.056***	0.005	0.010*	−0.003	0.025***
	(2.90)	(2.57)	(4.98)	(0.64)	(1.74)	(−0.58)	(7.63)
Comparisons using nearest-neighbour matching							
Average treatment	0.040***	0.032***	0.014	0.034***	0.032***	0.051***	0.031***
on the treated	(3.64)	(3.08)	(1.30)	(5.08)	(5.88)	(9.24)	(10.13)
Average treatment	0.031***	0.018***	0.004	0.037***	0.027***	0.041***	0.027***
on the control	(2.60)	(2.18)	(0.46)	(4.92)	(5.57)	(7.86)	(9.60)

Notes. t-values corrected for clustering at the firm level are in parentheses. Significant at *10%, **5% and ***1%. Estimates for both average treatment on the treated (i.e. processing firms) and average treatment on the control (i.e. non-processing firms) are obtained by using the nearest-neighbour matching approach in which firm size and firm sales are chosen as covariates.

comparisons are straightforward. However, they bear a cost because processing firms may be very different from non-processing firms in terms of size. To overcome such a pitfall, as suggested by Imbens (2004), I perform the nearest-neighbour matching between the treatment group (i.e. processing firms) and the control group (i.e. non-processing firms) by choosing the number of firm employees and firm sales as covariates. Each processing firm would find its most similar non-processing firm. Table 9 reports both the estimates for average treatment for the treated (ATT) and average treatment for the control (ATC). For instance, the coefficient of ATT for all processing firms is 0.037 and highly statistically significant, suggesting that, overall, productivity for processing firms is lower than that for similar non-processing firms.

The estimates in Table 9 hint that low-productivity firms may self-select to engage in processing trade. To control for this, I introduce a type-2 Tobit model or, equivalently, a bivariate sample selection model (Cameron and Trivedi, 2005). The type-2 Tobit specification includes:

(*i*) a processing participation equation,

$$\text{Processing}_{it} = \begin{cases} 0 \;\; if \;\; V_{it} < 0 \\ 1 \;\; if \;\; V_{it} \geq 0 \end{cases}, \tag{10}$$

where V_{it} denotes a latent variable faced by firm *i*; and

(*ii*) an 'outcome' equation whereby the firm's extent of processing imports is modelled as a linear function of other variables.

In particular, I estimate the following selection equation using a probit model:

$$\begin{aligned} \Pr(\text{Processing}_{it} = 1) = \Pr(V_{it} \geq 0) = \Phi(&\alpha_0 + \alpha_1 \ln TFP_{it-1} + \alpha_2 SOE_{it-1} \\ &+ \alpha_3 FIE_{it-1} + \alpha_4 \ln L_{it-1} + \alpha_5 Tenure_{it-1} + \lambda_j + \varsigma_t), \end{aligned} \tag{11}$$

where $\Phi(.)$ is the cumulative density function of the normal distribution. In addition to the logarithm of the firm's TFP, a firm's decision to engage in processing trade is also

Table 10

The Heckman Two-step Estimates of Bivariate Selection Model

Heckman two-step:	1st step		2nd step	
Regressand:	Processing indicator		Extent of processing	
One-period lag of log TFP (ln TFP_{ijt}^{GMM})	−0.126***	(−7.23)	−0.176***	(−15.17)
One-period lag of log Labour	0.152***	(25.55)	0.031***	(3.23)
One-period lag of SOEs indicator	−0.160***	(−2.82)	−0.039	(−1.47)
One-period lag of foreign indicator	0.978***	(68.97)	0.299***	(5.05)
One-period lag of firm tenure	0.004***	(5.02)	–	
Inverse Mills ratio	–		0.172**	(2.10)
Year-specific fixed effects	Yes		Yes	
Industry-specific fixed effects	Yes		Yes	
Observations	58,629		21,232	

Notes. t-values corrected for clustering at the firm level are in parentheses. Significant at *10%, **5% and ***1%. The sample selection model is presented in (10) and (11) in the text. The regressand in the first-step is the firm's processing dummy, whereas that in the second step is the firm's extent of processing imports. Firm-level system-GMM TFP is adopted as a measure of firm productivity. Firm tenure is used as an exclusion variable that appeared in the first step but not the second step. The three-digit Chinese industry-specific fixed effects and year-specific fixed effects are also included in the estimation.

affected by other factors, such as its ownership (whether it is an SOE or a multinational firm) and size (measured by the logarithm of the number of employees). Note that the bivariate sample selection estimation require an excluded variable that affects the firm's processing decision but does not appear in the extent of processing equation (Cameron and Trivedi, 2005). Here the firm's age ($Tenure_{it-1}$) serves this purpose, as previous studies have found that a firm's export probability is higher for older firms (Amiti and Davis, 2011). By contrast, my sample also reveals that the simple correlation between a firm's extent of processing imports and the firm's age is close to nil (−0.04), suggesting that the firm's age can be excluded in the second-step Heckman estimates.[34] All regressors in the type-2 Tobit selection model are of a one-period lag as it usually takes time for such factors to affect a firm's processing choice. Finally, I include the three-digit CIC industrial dummies, λ_j, and year dummies, ς_t, to control for other unspecified factors.

Table 10 reports the estimation results for the type-2 Tobit selection model. From the first-step probit estimates (11), low-productivity firms are more likely to engage in processing trade. Similarly, large and foreign firms are more likely to engage in processing trade. However, SOEs are less likely to become processing firms. Finally, as predicted, firms that were established earlier are more likely to engage in processing trade. I then include the computed inverse Mills ratio obtained in the first-step probit estimates in the second-step Heckman estimation as an additional regressor. It turns out that the estimated coefficients have exactly identical signs as obtained in the first-step estimates. Thus, after controlling for the endogenous selection of processing, I obtain the fitted value of the firm's extent of processing, which is used to replace the firm's actual extent of processing in the rest of estimates, as discussed above.

[34] Note that even when the firm's age is included, its coefficient in the second-step Heckman estimate is also statistically insignificant.

4.3. *Endogeneity Issues*

The specifications in Tables 7 and 8 face three possible endogeneity problems. The first one relates to the measure of firm input tariffs, because imports and tariffs are strongly correlated. This problem is essentially solved by using measures of tariffs based on time-invariant weights. The second relates to the possible reverse causality between firm productivity and exports. As the firm's productivity improves, its exports may grow faster for some products than for others. The disproportional growth in exports of some products would challenge the validity of a time-variant measure of firm output tariffs. To avoid this possibility, measures of tariffs based on time-invariant weights, as in (4), have been used in all specifications.

However, there is still another possible reverse causality problem. Although tariff reductions are regulated by the GATT/WTO agreements, they are still, to some extent, endogenous because firms in low-productivity sectors would lobby the government for protection (Grossman and Helpman, 1994), that is, to maintain related internationally negotiated tariffs at a relatively high level. I control for such reverse causality by using an IV approach.

Identifying a good instrument for tariffs is challenging. Inspired by Amiti and Konings (2007), here I construct a one-year lag of firm-specific output tariffs and input tariffs as instruments.[35] The economic rationale is as follows. The government generally has difficulty in removing the high protection status quo from an industry with high tariffs, possibly because of domestic pressure from special interest groups. Hence, compared with other sectors, industries with high tariffs one year ago would still be expected to have relatively high tariffs at present.

Column (1) of Table 11 presents 2SLS fixed-effects estimates using the previous tariffs with time-invariant weights as instruments.[36] After controlling for reverse causality, reductions in both firm input tariffs and firm output tariffs lead to firm productivity growth. As noted before, the measure of firm output tariffs may suffer from a pitfall because of the assumption of equal shares between domestic sales and exports for each product produced, as the product composition of exports may be different from that of domestic sales by the sector integration of GSCs and by the intensity of the sectors in processing firms. To address this concern, besides dropping pure domestic firms and pure exporters from the sample, I run two sets of auxiliary regressions. First, all industries are classified into two groups (more integrated and less integrated) according to their 'production depth' of engaging (GSCs) which is measured by the value-added ratio to gross industrial output (OECD, 2010). By taking the mean of such ratios across two-digit level industries as a cut-off, columns (2) and (3) regress the impact of tariff reductions on firm productivity by the extent of GSCs integrating. Second, columns (4) and (5) run regressions for sectors with high (low)

[35] Accordingly, the interaction between the firm's input and output one-period tariff with the time-invariant weight and the fitted extent of processing trade are adopted as additional instruments in all IV estimates.

[36] Note that adopting firm-specific fixed effects here would cause a huge loss of observations as most of the firms do not have a continuous panel in the sample. Such a pattern is more pronounced in the 2SLS estimates when using the one-year lagged tariffs as instruments. I therefore include the disaggregated three-digit CIC industry-specific fixed effects and year-specific fixed effects in all 2SLS estimates.

Table 11

IV Estimates with Measure of System-GMM TFP

Regressand: ln TFP_{ijt}^{GMM}	All sample (1)	GSCs integrated: Less (2)	GSCs integrated: More (3)	Processing intensity: Low (4)	Processing intensity: High (5)
Firm output tariffs	−1.319***	−0.825***	−1.962***	−1.657**	−1.941***
	(−4.60)	(−2.13)	(−3.66)	(−3.98)	(−4.47)
Firm output tariffs × fitted extent of processing	0.817*	0.802	1.184	1.321	1.765***
	(1.72)	(1.18)	(1.41)	(1.53)	(2.67)
Firm input tariffs	−1.712***	−2.821***	−1.519***	−1.883**	−3.447**
	(−3.46)	(−3.57)	(−2.76)	(−3.50)	(−2.32)
Firm input tariffs × fitted extent of processing	2.460***	2.497*	2.818**	3.478**	3.546*
	(2.54)	(1.75)	(2.71)	(2.65)	(1.72)
Fitted extent of processing	−0.740***	−1.005***	−0.778***	−0.944***	−0.833***
	(−17.66)	(−15.99)	(−10.28)	(−12.28)	(−11.95)
Kleibergen–Paap rank LM χ^2 statistic	87.75†	428.5†	961.9†	883.6†	639.1†
Kleibergen–Paap rank Wald F statistic	95.94†	112.0†	257.1†	234.2†	171.3†
Year-specific fixed effects	Yes	Yes	Yes	Yes	Yes
Industry-specific fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	22,812	8,374	14,438	13,633	9,179
R^2	0.17	0.18	0.16	0.16	0.23
First-stage regressions					
IV1: Firm output tariffs with a lag	0.004***	0.005***	0.003***	0.003***	0.004***
	(12.03)	(9.91)	(9.38)	(8.40)	(4.19)
IV2: Firm output tariffs with a lag × fitted extent of processing	0.004***	0.004***	0.004***	0.005***	0.004***
	(19.15)	(12.67)	(5.92)	(11.72)	(7.69)
IV3: Firm input tariffs with a lag	0.005***	0.004***	0.005***	0.005***	0.005***
	(8.89)	(19.62)	(4.22)	(7.95)	(3.82)
IV4: Firm input tariffs with a lag × fitted extent of processing	0.008***	0.008***	0.008***	0.007***	0.010***
	(14.31)	(9.02)	(7.85)	(10.33)	(9.01)

Notes. t-values in parentheses are obtained using bootstrapped standard errors. Significant at *10%, **5% and ***1%. Column (1) includes the entire sample in the regression. Columns (2) and (3) include sectors that are less (more) integrated in global supply chains (GSCs), respectively, using the industrial average ratio of value-added to gross industrial output as cut-offs. Columns (4) and (5) include the sectors with low (high) intensity of the sectors in processing firms, respectively, in which the intensity is measured by share of number of processing firms over number of total firms in each industry. †(‡) indicates significance of the p-value at the 1 (5)% level. In the first-stage regressions, IV1 reports the coefficient of the firm output tariffs with initial time-invariant weight and one-period lag of tariffs, using firm output tariffs with initial time-invariant weight and current tariffs as the regressand. IV2 reports the coefficient of the interaction between fitted extent of processing obtained from the second-step Heckman estimates in Table 8 and firm output tariffs with initial time-invariant weight and one-period lag of tariffs, using the interaction between fitted extent to processing and current tariffs as the regressand. Similarly, IV3 reports the coefficient of the firm input tariffs with initial time-invariant weight and one-period lag of tariffs using firm input tariffs with initial time-invariant weight and current tariffs as the regressand. IV4 reports the coefficient of the interaction between fitted extent of processing and firm input tariffs with initial time-invariant weight and one-period lag of tariffs, using the interaction between fitted extent of processing and firm output tariffs with initial time-invariant weight and current tariffs as the regressand. Pure domestic firms and pure exporters are dropped from the sample in all estimates.

intensity of the sectors in processing firms, respectively, in which the intensity is measured by share of number of processing firms over number of total firms in each industry and the mean of the ratios across industries is taken as the cut-off. In all cases, the coefficients of output and input tariffs are significant and in line with my previous findings.

Several tests were performed to verify the quality of the instruments. First, I use the Kleibergen–Paap LM χ^2 statistic to check whether the excluded instruments are correlated with the endogenous regressors. As shown in Table 11, the null hypothesis that the model is under-identified is rejected at the 1% significance level. Second, the Kleibergen–Paap (2006) F-statistics provide strong evidence for rejecting the null hypothesis that the first stage is weakly identified at a highly significant level.[37] Finally, the first-stage estimates reported in the lower module of Table 11 offer strong evidence to justify such instruments. In particular, all the t-values of the instruments are significant. Finally, standard errors are corrected for the use of the estimated regressors by bootstrapping.[38]

4.4. *Further Robustness Checks of 2SLS Estimates*

It is also worthwhile checking whether the effects of firm-level input and output tariffs on firm productivity pick up only the role of firm size, given that large firms usually have high productivity, or whether the effects are sensitive to the inclusion of the firm's type of ownership. I therefore include an SOE indicator, a foreign indicator, and the log of labour (i.e. a measure of firm size) in all the 2SLS estimates in Table 12.

Because measured TFP may also pick up the difference in prices and price-cost mark-ups across firms, column (1) of Table 12 performs the 2SLS estimates using the logarithm of the firm's labour productivity as the regressand. As the log of firm labour is already used as the denominator of the regressand, it is no longer appropriate to include it as a control variable for firm size in the regression. I instead use the log of the firm's capital-labour ratio as a proxy.

To check further whether my main findings are sensitive to the measure of firm TFP and the empirical specifications, column (2) also uses the Levinsohn–Petrin (2003) TFP as the regressand while controlling for other variables as in column (1). Column (3) still uses the system-GMM as the regressand but includes the above-mentioned controlling variables. Overall, the main findings of the estimates in these columns are highly consistent with those in Table 11: the impact of input tariff reductions on productivity improvement, overall, is weaker than that of output tariff reductions. The firm's gains from tariff reductions are diminishing as the firm's processing imports share increases.

Thus far, the effect of China's import tariff reductions on firm efficiency has been carefully investigated. However, although China has substantially reduced its import

[37] Note that the Cragg and Donald (1993) F-statistic is no longer valid because it only works under the i.i.d. assumption. As here I have four (more than three) endogenous variables, STATA does not report the critical values for the Kleibergen–Paap (2006) weak instruments test. In this case, Baum *et al.* (2007) suggest that one can safely adopt 10 as a critical value as initiated by Staiger and Stock (1997). As all my Kleibergen–Paap (2006) F-statistics are one-order much higher than 10, it is safe to reject the null hypothesis of weak instruments in all estimates.

[38] There are in fact four steps to my estimation: the selection (11); the second-step Heckman equation used to obtain the predicted extent of processing; the first-step of 2SLS where the predicted extent of processing is a regressor; and the second-step of 2SLS estimates. Panel bootstrapping by randomly drawing firms is done in the last two steps.

Table 12

More Robust IV Estimates

Regressand:	$\ln LP_{ijt}$ (1)	$\ln TFP_{ijt}^{LevP}$ (2)	$\ln TFP_{ijt}^{GMM}$ (3)	$\ln TFP_{ijt}^{GMM}$ (4)	Weighted $\ln TFP_{ijt}^{GMM}$ (5)
Firm output tariffs	−1.980***	−1.217**	−1.100***	−1.096***	−1.159***
	(−3.49)	(−2.02)	(−4.51)	(−4.62)	(−4.47)
Firm output tariffs × fitted extent of processing	2.260**	−0.106	0.677	0.675	0.812**
	(2.03)	(−0.08)	(1.63)	(1.47)	(1.96)
Firm input tariffs	−3.866**	−5.069***	−1.380***	−1.378***	−1.589***
	(−2.30)	(−2.69)	(−2.66)	(−2.47)	(−2.57)
Firm input tariffs × fitted extent of processing	8.610***	10.309***	2.448**	2.435**	2.664**
	(2.36)	(2.59)	(2.12)	(2.09)	(2.06)
Fitted extent of processing	−2.737***	−2.901***	−1.251***	−1.251***	−1.311***
	(−22.42)	(−23.00)	(−26.78)	(−23.61)	(−27.83)
SOEs indicator	−0.619***	−0.369***	−0.187***	−0.187***	−0.188***
	(−11.60)	(−5.15)	(−7.71)	(−7.51)	(−7.81)
Foreign ownership indicator	0.493***	0.475***	0.220***	0.220***	0.229***
	(19.38)	(24.15)	(27.24)	(32.40)	(28.84)
Firm size	0.325***	0.559***	0.068***	0.068***	0.072***
	(51.51)	(81.26)	(34.23)	(29.81)	(24.59)
Firm external tariffs				0.001	0.001
				(1.09)	(1.22)
Kleibergen–Paap rank LM χ^2 statistic	106.5†	92.00†	105.4†	105.4†	105.5†
Kleibergen–Paap rank Wald F statistic	54.98†	47.78†	55.18†	55.10†	55.10†
Year-specific fixed effects	Yes	Yes	Yes	Yes	Yes
Industry-specific fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	19,296	15,759	19,283	19,283	19,283
R^2	0.40	0.53	0.30	0.30	0.65

Notes. t-values in parentheses are obtained using bootstrapped standard errors. Significant at *10%, **5% and ***1%. †indicates significance of p-value at the 1% level. The regressand is log of value-added labour productivity ($\ln LP_{ijt}$) in column (1) and Levinsohn–Petrin (2003) TFP ($\ln TFP_{ijt}^{LevP}$) in column (2), and conventional measure of system-GMM TFP ($\ln TFP_{ijt}^{GMM}$) in columns (3) and (4). The regressand in column (5) is weighted system-GMM TFP which is calculated by multiplying $\ln TFP_{ijt}^{GMM}$ with their relative standard deviations across firms within an industry at the two-digit level. In all IV estimates, I control for year-specific fixed effects and time-invariant two-digit level Chinese industry fixed-effects. Firm size in columns (2)–(5) is proxied by log of firm labour, whereas in column (1) it is proxied by firm's capital-labour ratio. All instruments used are the same as those in Table 9. Pure domestic firms and pure exporters are dropped from the sample.

tariffs in the new century, Chinese exporters have also enjoyed large tariff reductions in their export destinations. Access to large foreign markets could possibly create incentives for productivity upgrading, especially if such investments require substantial fixed costs. Thus, controlling for tariff reductions in China's export destinations is also worthwhile to obtain a precise estimate of the effect of import tariff reductions on firm TFP.

To measure tariff reductions in a firm's export destination markets, I construct an index of firm-specific external tariffs (FET_{it}) as follows:[39]

[39] Note that all the main findings are not changed if firm external tariffs are measured using time-invariant export weights. The reason for choosing a time-variant export weight is to allow a dynamic response of the firm's exports to a reduction in foreign tariffs.

$$FET_{it} = \sum_k \left[\left(\frac{X_{it}^k}{\sum_k X_{it}^k} \right) \sum_c \left(\frac{X_{ikt}^c}{\sum_c X_{ikt}^c} \right) \tau_{kt}^c \right], \tag{12}$$

where τ_{kt}^c is product k's *ad valorem* tariff imposed by export destination country c in year t. A firm may export multiple types of products to multiple countries. The ratio in the second set of parentheses in (12), $X_{ikt}^c / \sum_c X_{ikt}^c$, measures the export ratio of product k produced by firm i but consumed in country c, yielding a weighted external tariff across Chinese firms' export destinations. Similarly, the first term in parentheses in (12), $X_{it}^k / \sum_k X_{it}^k$, measures the proportion of product k's exports over firm i's total exports. The mean of the firm-specific external tariff is only 0.9%, which is significantly lower than its counterpart for firm-specific import tariffs on final goods (8.3%). This makes good economic sense. The most important export destinations for Chinese firms are developed countries, such as the US and the countries of the EU, which usually set substantially lower import tariffs on exporters from developing countries like China. Column (4) of Table 12 presents the estimation results including a variable for the firm's external tariffs in the regressions. The coefficient of firm external tariffs is statistically insignificant. One possible reason for this is that Chinese firms had already entered foreign markets before 2000. Thus, tariff reductions in Chinese firms' export destinations have no statistically significant effect in reducing the fixed costs of exports.

Still, the regressand used in all the estimation is a measure of TFP, estimated in various ways. As the observations are estimated but not observed, it is worthwhile controlling for the fact that some observations are estimated more precisely than the others. Therefore, I compute the standard deviation of system-GMM TFP both across firms within an industry and across all firms and divide its sector average by the total average to multiply the firm's system-GMM TFP as the regressand in the last column of Table 12.[40] I obtain similar results as before: the effect of firm tariffs on productivity declines as the firm's processing imports grow. The overall impact of output tariff reductions is stronger than that of input tariff reductions.

Finally, the great flexibility of the system-GMM estimation method indeed provides a unique opportunity to obtain the effects of tariff reduction on firm productivity using a one-step approach. That is, the coefficients of both input coefficients for the production function and tariffs are obtained simultaneously. I hence experiment with this in Appendix Table C2, as additional robustness checks.[41]

[40] See columns (5) and (6) of Appendix Table C1. I thank a referee for suggesting this point.

[41] Using the log of firm output as the regressand, both the current period and a one-period lag realisation of firm inputs – labour, capital and materials – are included as regressors. Simultaneously, firm output and input tariffs based on time-invariant weights, the extent of processing imports and its interaction with tariffs are included as another set of regressors. To control for possible endogeneity, I adopt a one-period lag of firm output (input) tariffs with time-invariant weights as instruments as before. Appendix Table C2 reports the 2SLS fixed-effects estimates using the one-step system-GMM approach. All estimation results are highly consistent with the previous findings: the impact of tariff reductions on productivity improvement shrinks as the firm's processing imports grow. Overall, firm output tariff reduction leads to stronger productivity gains than firm input tariff reductions.

4.5. *Discussion of Channels*

The article has presented rich evidence that both output and input tariff reductions boost firm productivity. However, we still have little understanding about the channels through which these effects occur. The impact of input tariffs on productivity is relatively direct, as lower tariffs induce access to a larger variety of imported intermediate inputs (Helpern *et al.*, 2010).[42] Reductions in output tariffs are found to have a pro-competitive effect. However, it is less clear whether such a pro-competitive effect is realised through improvement in the efficiency of firms that are present in the market, or through weeding out the less-productive firms from the market.

To test these two possible channels, I first include an always-present firms indicator (i.e. it equals one if the firm is present in all years during 2000–6 and otherwise zero) in column (1) of Table 13. The always-present indicator has a positive and significant sign, suggesting that always-present firms are more productive. To check whether low-productivity firms collapse and exit from the market, column (2) includes an exit indicator that takes the value one if firms exit from the market in the next year and zero otherwise. The insignificant sign of the exiting dummy suggests that exiters do not have a significant productivity difference compared with non-exiting firms. This finding is different from the predictions in Melitz (2003).

Amiti and Konings (2007) argue that tariff reductions could result in firms switching their scope from low to high-productivity products. However, they do not have information on firm scope because of Indonesian data restrictions. Thus, they use a switching dummy as a compromise. However, my merged data set includes information on exporters' scope. Many Chinese firms export multiple products, with the maximum reaching 745 export products. The logarithm of the firm's export scope is included in column (3) of Table 13, and its coefficient is positive and significant, suggesting that firms exporting more products have higher productivity. In column (4), the log of the firm's scope is then interacted with firm-specific input and output tariffs. The interaction of output tariffs and log scope is found to be significant, whereas that of input tariffs and log scope is insignificant, indicating that at least a few gains from output tariff reductions are attributable to product switching, as also found by Amiti and Konings (2007) with their more limited data. However, this channel is not important for input tariff reductions.

Last but not least, firms' productivity gains from trade reform may also result from the channel of investing in new technologies (Bustos, 2011). Firms with higher R&D expenses are expected to have higher productivity. This conjecture is verified in column (5) of Table 13 by including a variable for the firm's log R&D. In the last column, the logarithm of R&D is also interacted with the firm-specific input and output tariffs. Interestingly, the interaction coefficients of the output and input tariffs and R&D are insignificant, showing that the gains from both output and input tariff reductions do not result from investing in new technologies. One reason is the limited firm R&D data in my sample: around 80% of the observations do not contain valid

[42] Besides variety, Amiti and Konings (2007) highlight two other possible channels through which cheaper imported inputs can raise productivity: learning and quality effects.

Table 13

IV Estimates for Channels

	Firm's selection		Multi-product firms		R&D expenses	
Regressand: $\ln TFP_{ijt}^{GMM}$	(1)	(2)	(3)	(4)	(5)	(6)
Firm output tariffs	−1.081***	−1.086***	−0.838***	−0.468	−1.119***	−1.628
	(−4.10)	(−3.44)	(−3.51)	(−1.54)	(−2.16)	(−1.17)
Firm output tariffs × fitted extent of processing	0.934**	0.934*	1.026**	1.139***	0.421	0.785
	(2.03)	(1.82)	(2.30)	(2.38)	(0.37)	(0.51)
Firm output tariffs × log of firm's scope				−0.263***		
				(−3.45)		
Firm output tariffs × log of firm's R&D						0.061
						(0.40)
Firm input tariffs	−1.671***	−1.672***	−1.267***	−1.199***	−2.060*	−0.899
	(−4.10)	(−2.88)	(−3.36)	(−3.31)	(−1.73)	(−0.52)
Firm input tariffs × fitted extent of processing	3.557***	3.575***	4.065***	3.486***	4.711	3.889
	(4.07)	(2.94)	(4.33)	(4.29)	(1.53)	(1.35)
Firm input tariffs × log of firm's scope				0.224		
				(1.08)		
Firm input tariffs × log of firm's R&D						−0.150
						(−0.73)
Fitted extent of processing	−1.500***	−1.501***	−1.467***	−1.461***	−1.471***	−1.476***
	(−40.41)	(−29.71)	(−35.02)	(−32.76)	(−10.87)	(−9.16)
SOEs indicator	−0.249***	−0.238***	−0.216***	−0.217***	−0.245***	−0.244***
	(−12.94)	(−12.89)	(−9.87)	(−8.38)	(−8.56)	(−6.80)
Foreign ownership indicator	0.281***	0.282***	0.228***	0.229***	0.310***	0.309***
	(40.84)	(39.83)	(28.95)	(29.88)	(18.64)	(19.74)
Log of labour	0.079***	0.079***	0.061***	0.061***	0.078***	0.078***
	(34.27)	(31.90)	(35.40)	(27.95)	(14.85)	(13.77)
Log of capital-labour ratio	0.033***	0.033***	0.021***	0.019***	0.044***	0.045***
	(18.21)	(15.31)	(8.82)	(8.68)	(6.97)	(8.74)
Firm exits next year		0.009				
		(0.92)				
Always-present firm indicator	0.013*					
	(1.89)					
Log of firm's scope			0.042***	0.059***		
			(19.57)	(8.16)		
Log of R&D					0.028***	0.028***
					(11.33)	(2.18)
Year-specific fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry-specific fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19,190	19,190	19,190	19,190	3,331	3,331
R^2	0.38	0.38	0.40	0.40	0.47	0.47

Notes. t-values in parentheses are obtained using bootstrapped standard errors. Significant at *10%, **5% and ***1%. The two-digit Chinese industry-specific fixed effects are included in the estimation.

R&D expenses,[43] thus the effect of R&D is under-estimated for firms to realise gains from tariff reductions.

4.6. *Economic Magnitudes and Welfare Contributions*

This subsection discusses the economic magnitudes of tariff reductions. As shown in the IV estimates in column (1) of Table 11, the regressand is in logarithms whereas the regressors are in levels. Thus, the estimated key coefficients can be interpreted as semi-elasticities. With tariffs as natural numbers used in the regressions (e.g. the mean of firm output tariffs is 0.083, as reported in Table 5), the own coefficient of the firm output (input) tariffs is −1.32 (−1.71). Measuring tariffs in percentage points (so the mean of firm output tariffs in the sample is 8.3 percentage points), such coefficients are changed to −0.0132 (−0.0171), implying that a 10 percentage point fall in output tariffs for non-processing firms leads to a 0.132(0.171) increase in log TFP, or equivalently, a productivity gain of 13.2 (17.1)%.[44]

Equally important, the firm's productivity gains from cutting input and output tariffs become smaller as the firm's processing imports share grows. On average, the impact of the output tariff reductions on productivity improvement is $-0.013 + 0.008 \times 0.49 = -0.0092$, given that the mean of the fitted extent of processing is 0.49, implying that a 10 percentage point fall in output tariffs leads to a productivity gain of 9.2%. Analogously, the average impact of a reduction in input tariffs is $-0.017 + 0.025 \times 0.49 = -0.0051$, indicating that a 10 percentage point fall in input tariffs leads to a productivity gain of 5.1%, almost 56% as high as the gains from reducing output tariffs.[45] Average firm output tariffs were cut 8.2 percentage points (from 15.6% in 2000 to 7.4% in 2006), which thus predicts $0.009 \times 8.2 = 7.4\%$ productivity gain and contributes 44.4% of the 0.17 log point increase in firm productivity covered in the sample. By the same token, the average firm input tariffs were cut 0.36 percentage points (from 2.54% in 2000 to 2.18% in 2006), which thus predict $0.005 \times 0.36 = 0.18\%$ productivity gain and contributes 1.1% of the 0.17 log point increase in log of TFP. Adding these numbers, tariff reductions, overall, contribute around 45.5% to productivity growth for the firms covered in the sample.

As economy-wide productivity growth is one of the best measures of a country's standard of living, my final step is to offer a more intuitive economic interpretation for the contribution of tariff reductions to China's aggregated productivity growth. The adding-up of firm productivity to economy-wide productivity is non-trivial as, because

[43] In particular, R&D in 2004 is completely missing. Moreover, around 50% of firms report negative or zero R&D expenses in my sample.

[44] My estimates are also close to other studies such as Amiti and Konings (2007), who find that a 10 percentage point fall in output (input) tariffs leads to a productivity gain of 6.4 (12.7)% using data on Indonesian firms.

[45] It is also interesting to check the productivity gains from tariff reductions for pure processing firms, for which the ratio of processing imports to total imports equals one. As firm input tariffs for pure processing firms reduce to zero, given that processing imports are duty-free, one cannot directly calculate such productivity gains from column (1) of Table 11. However, as the impact of the input tariff reductions is given by $-0.0171 + 0.0246 \times \mathrm{E}(Pext_{it}|\mathbf{Z}_{it})$, by using a sufficiently high value for the extent of processing (e.g. the 90th percentile of $\mathrm{E}(Pext_{it}|\mathbf{Z}_{it}) = 0.69$) as a proxy of pure processing firms, the impact of input tariff reductions is close to zero, confirming that heavy processing firms rarely gain from input tariff reductions.

of the presence of vertical integration, intermediate inputs across firms (sectors) contribute to aggregated productivity by allowing productivity gains in successive firms (sectors) to augment one another (OECD, 2001).[46] As initiated by Domar (1961) and later elaborated by Hulten (1978) and Feenstra *et al.* (2013*b*), the economy-wide TFP can be aggregated by using the 'Domar weight' which is defined by each firm's gross output relative to economy-wide absorption (i.e. total gross output minus trade surplus). I then calculate the aggregated TFP using Domar weights for each year. It turns out that aggregated log of TFP increases around 0.53 log points (from 0.56 in 2000 to 1.09 in 2006).[47] As described before, both output and input tariff reductions, on average, lead to productivity gains of 7.54% + 0.11% = 0.076, and thus contribute to 14.5% of the 0.53 log point increase in economy-wide log productivity. A final remark is that the calculation here presumes that tariff cuts have no impact on firm productivity beyond the sample. As tariff reductions still, in reality, have beneficial ripple effects beyond the set of firms in the sample, the calculated contribution to the whole economy should be interpreted as a lower-bound number.

5. Concluding Remarks

To explore how reductions in tariffs on imported inputs and final goods affect firm productivity, the article has exploited the special tariff treatment afforded to imported inputs by processing firms as opposed to non-processing firms in China. As a popular trade pattern in a large number of developing countries, including China, processing trade plays an important role in the realisation of productivity gains. Overall, I find that the impact of output tariff reduction is greater than that of input tariff reduction for large Chinese trading firms. More interestingly, the positive impact of reduction in input (output) tariffs on firm productivity is weaker as firms' processing import share grows.

This article is one of the first to explore the role of processing trade in Chinese firms' productivity gains. The rich data set enables the determination of whether a firm engages in processing trade and the examination of the effect of the firm's extent of processing trade engagement on productivity. With such information, firm-level input and output tariffs were also constructed, as one of the first attempts in the literature, which, in turn, enriches the understanding of the economic effects of China's special tariff reforms in processing trade.

Appendix A. Matching Production and Trade Data Sets

My discussion on matching the two data sets (i.e. firm-level production data and firm-customs data) here draws heavily from Yu and Tian (2012). As mentioned in the text, I go through two

[46] For example, if TFP growth for both shoe and rubber firms is 1%, the simple average of such firms' TFP growth will be 1%. However, productivity growth of the integrated rubber and shoe industry will be more than 1%, as the shoe firms' productivity gains cumulate with those of the rubber firms as the latter sells inputs to the former.

[47] To calculate Domar-weight TFP, the Domar weight is multiplied by four since the gross output of my merged sample only accounts for a quarter of total gross output in the full-sample data set, as shown in Table 3. See also Appendix C for a careful derivation of the Domar-weight aggregate productivity.

steps to merge transaction-level trade data with firm-level production data. In the first step, I match the two data sets by firm name and year. The year variable is necessarily an auxiliary identifier because some firms could have different names across years and newcomers could possibly take their original names. Using the raw (i.e. unfiltered) production data set, I come up with 83,679 merged firms; this number is further reduced to 69,623 with the more accurately filtered production data set.

In the second step, I use another matching technique as a supplement. In particular, I adopt two other common variables to identify firms: postal code and the last seven digits of a firm's phone number. The rationale is that firms should have different and unique phone numbers within a postal district. Although this method seems straightforward, subtle technical and practical difficulties still exist. For instance, the production-level trade data set includes both area codes and a hyphen in the phone numbers, whereas the firm-level production data set does not. Therefore, I use the last seven digits of the phone number to serve as the proxy for firm identification for two reasons. First, in 2000–6, some large Chinese cities (e.g. Shantou in Guangdong province) added one more digit at the start of their seven-digit phone numbers. Therefore, using the last seven digits of the number will not confuse firm identification. Second, in the original data set, phone numbers are defined as a string of characters with the phone postal code; however, it is inappropriate to de-string such characters to numerals because a hyphen is used to connect the postal code and phone number. Using the last seven-digit sub-string neatly solves this problem.

A firm might not include information on its name in either the trade or the production data set. Similarly, a firm could lose its phone and/or postal code information. To be sure that the merged data set can cover as many common firms as possible, I then include observations in the matched data set if a firm occurs in either the name-adopted matched data set or the phone-and-post-adopted matched data set.

As shown in Appendix Table A1, column (1) reports the number of observations of HS eight-digit monthly transaction-level trade data from China's General Administration of Customs by year. As shown at the bottom of column (1), there are more than 118 million monthly trade transactions conducted by 286,819 firms during the seven years, as shown in column (2). Meanwhile, if no further data cleaning and stringent filter criteria are adopted as introduced in the text, column (3) shows that there are 615,591 large manufacturing firms in China. However, after stringent filtering according to GAAP requirements, around 70% of them survive – number of the filtered firms is 438,165 as seen at the bottom of column (4). Accordingly, column (5) reports the number of matched firms using exactly identical company names in both trade data set and raw production data set. By contrast, column (6) reports number of matched firms using exactly identical company names in both the trade data set and the filtered production data set, which results in 69,623 matched firms.

Column (7) reports the number of matched firms using exactly identical company names and exactly identical postal codes and phone numbers in both the trade and raw production data sets. The number of merged firms increases to 91,299. By way of comparison, my matching performance is highly comparable with that of other similar studies. For example, Ge *et al.* (2011) use the same data sets and similar matching techniques and end up with 86,336 merged firms. Finally, if I match the more stringent filtered production data set with the firm-level data set using exactly identical company names and postal–phone code numbers but drop firms whose customs-reported exports are higher than NBS-reported firm sales, I end up with 76,823 firms in total, as shown in the last column of Appendix Table A1. I use these firms to run the regressions because they are the most reliable firms that can pass various stringent filtering processes in the firm production data.

After merging both the product-level trade data and the firm-level production data, the 76,823 common trading firms account for approximately 27% of the 286,819 firms in the product-level trade data set and approximately 17% of the 438,146 valid firms in the firm-level production data set (11% of the valid firms are exporters, whereas 6% of them are importers).

Table A1

Matched Statistics – Number of Firms

	Trade data		Production data		Matched data			
Year	Transactions (1)	Firms (2)	Raw firms (3)	Filtered firms (4)	w/Raw firms (5)	w/Filtered firms (6)	w/Raw firms (7)	w/Filtered firms (8)
2000	10,586,696	80,232	162,883	83,628	18,580	12,842	21,425	15,748
2001	12,667,685	87,404	169,031	100,100	21,583	15,645	24,959	19,091
2002	14,032,675	95,579	181,557	110,530	24,696	18,140	28,759	22,291
2003	18,069,404	113,147	196,222	129,508	28,898	21,837	33,901	26,930
2004	21,402,355	134,895	277,004	199,927	44,338	35,007	49,891	40,711
2005	24,889,639	136,604	271,835	198,302	44,387	34,958	49,891	40,387
2006	16,685,377	197,806	301,960	224,854	53,748	42,833	49,680	47,591
All years	118,333,831	286,819	615,951	438,165	83,679	69,623	91,299	76,823

Notes. Column (1) reports number of observations of HS eight-digit monthly transaction-level trade data from China's General Administration of Customs by year. Column (2) reports number of firms covered in the transaction-level trade data by year. Column (3) reports number of firms covered in the firm-level production data set compiled by China's National Bureau of Statistics without any filter and cleaning. By contrast, column (4) presents number of firms covered in the firm-level production data set with careful filtering according to GAAP requirements. Accordingly, column (5) reports number of matched firms using exactly identical company names in both the trade data set and the raw production data set. By contrast, column (6) reports number of matched firms using exactly identical company names in both the trade data set and the filtered production data set. Column (7) reports number of matched firms using exactly identical company names and exactly identical postal codes and phone numbers in both the trade data set and the raw production data set. By contrast, column (8) reports number of matched firms using exactly identical company names and exactly identical postal codes and phone numbers in both the trade data set and the filtered production data set.

Given that only 27% of firms are exporters in the firm-level production data set (Feenstra *et al.*, 2013*b*), the merged data set hence accounts for around 40% of the filtered full-sample firm-level production data set in terms of number of exporters, and around 53% of exports in terms of export value.

Appendix B. The Augmented Olley–Pakes TFP Measures

In this Appendix, I estimate the measured Olley–Pakes TFP by taking the role of processing trade into account. In the article, the Olley–Pakes TFP is estimated in three ways:

(*i*) TFP^{OP} which is used in the full-sample estimates in columns (1) and (2) in Table 7;
(*ii*) TFP^{OP1} which separates processing firms and non-processing firms into two groups and uses different control function approaches, as discussed below, and is used in columns (3) and (4) in Table 7 and column (1) in Table 8; and
(*iii*) TFP^{OP2} which pools processing firms and non-processing firms together for estimation and is used in column (2) in Table 8.

It is important to stress that different versions of Olley–Pakes TFP do not qualitatively change my estimation results.

By assuming that the expectation of future realisation of the unobserved productivity shock, v_{it}, relies on its contemporaneous value, firm *i*'s investment is modelled as an increasing function of both unobserved productivity and log capital, $k_{it} \equiv \ln K_{it}$. Following previous works, such as Amiti and Konings (2007), the Olley–Pakes approach was revised by adding other control variables as extra arguments of the investment function as follows:

$$I_{it} = \tilde{I}(\ln K_{it}, \upsilon_{it}, FX_{it}, WTO_t, SOE_{it}), \tag{B.1}$$

where FX_{it} is a dummy to measure whether firm i exports in year t as firm's export decision may affect firm investment. As my firm data set is from 2000 to 2006, I include a WTO dummy (i.e. one for a year after 2001 and zero for before) in the investment function. Finally, given the importance of state intervention, SOEs would have different decision behaviour than non-SOEs. I therefore include an SOE dummy in the investment function as well. Therefore, the inverse function of (B.1) is $\upsilon_{it} = \tilde{I}^{-1}(\ln K_{it}, I_{it}, FX_{it}, WTO_t, SOE_{it})$. The unobserved productivity also depends on log capital and other arguments. The estimation specification (M.1) in the text can now be written as follows:

$$\ln Y_{it} = \beta_0 + \beta_m \ln M_{it} + \beta_l \ln L_{it} + g(\ln K_{it}, I_{it}, FX_{it}, WTO_t, SOE_{it}) + \epsilon_{it}, \tag{B.2}$$

where $g(\cdot)$ is defined as $\beta_k \ln K_{it} + \tilde{I}^{-1}(\ln K_{it}, I_{it}, FX_{it}, WTO_t, SOE_{it})$. Following Olley and Pakes (1996), fourth-order polynomials in log-capital, log-investment, firm's export dummy and import dummy are used to approximate $g(\cdot)$.[48] With this specification, the coefficient of labour β_l and that of materials β_m can be estimated as the first-step procedure.

The three different versions of Olley–Pakes TFP use different control functions. The control function of TFP^{OP} which is used in the full-sample estimates cannot control for the firm's import status, as the full-sample production data set does not report import status. However, the import dummy is incorporated in the other two approaches (TFP^{OP1} and TFP^{OP2}) when using a matched sample to estimate. The difference between TFP^{OP1} and TFP^{OP2} is whether processing firms are separated from non-processing firms.

B.1. *TFPOP Used in the Full-sample Data Set*

In the full-sample data set, information on the firm's import status and processing status is unavailable. I hence adopt the following functional form:

$$g(\ln K_{it}, I_{it}, FX_{it}, WTO_t, SOE_{it}) = (\alpha_0 + \alpha_1 WTO_t + \alpha_2 FX_{it} + \alpha_3 SOE_{it}) \sum_{h=0}^{4} \sum_{q=0}^{4} \delta_{hq} (\ln K_{it})^h I_{it}^q. \tag{B.3}$$

In the first step, I obtain estimates of $\hat{\beta}_m$ and $\hat{\beta}_l$ for non-processing (ordinary) firms. I then calculate the residual R_{it} which is defined as $R_{it} \equiv \ln Y_{it} - \hat{\beta}_m \ln M_{it} - \hat{\beta}_l \ln L_{it}$.

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

$$R_{it} = \beta_k \ln K_{it} + \tilde{I}^{-1}(g_{i,t-1} - \beta_k \ln K_{i,t-1}, \hat{pr}_{i,t-1}) + \epsilon_{it}, \tag{B.4}$$

where $\hat{pr}_i$ denotes the fitted value for the probability of the firm 's exit in the next year. As the specific 'true' functional form of the inverse function $\tilde{I}^{-1}(\cdot)$ is unknown, it is appropriate to use fourth-order polynomials in $g_{i,t-1}$ and $\ln K_{i,t-1}$ to approximate it. In addition, (B.4) also requires the estimated coefficients of the log-capital in the first and second terms to be identical. Therefore, non-linear least squares seems to be the most desirable econometric technique. Finally, the Olley–Pakes type of TFP for ordinary firm i in industry j is obtained once the estimated coefficient $\hat{\beta}_k$ is obtained:

$$\ln TFP_{ijt}^{OP} = \ln Y_{it} - \hat{\beta}_m \ln M_{it} - \hat{\beta}_k \ln K_{it} - \hat{\beta}_l \ln L_{it}. \tag{B.5}$$

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

B.2. TFP^{OP1} *with Separate Estimates for Processing and Non-processing Firms*

By contrast, the control functions used in TFP^{OP1} for processing firms and non-processing firms are different. If a firm is engaged in any processing imports, it is defined as a processing firm; otherwise it is defined as a non-processing (ordinary) firm. I first separate all firms in the sample into two groups – non-processing (ordinary) firms and processing firms. The control function for non-processing firms in the first-step estimates takes the following form:

$$g^{ord}(\ln K_{it}, I_{it}, FX_{it}, IM_{it}, WTO_t, SOE_{it}) = (\theta_0 + \theta_1 WTO_t + \theta_2 FX_{it} + \theta_3 IM_{it} + \theta_4 SOE_{it}) \sum_{h=0}^{4}\sum_{q=0}^{4} \delta_{hq}^{ord} (\ln K_{it})^h I_{it}^q, \tag{B.6}$$

where IM_{it} denotes the import dummy that takes the value one if firm i in year t is an importer, and zero otherwise. The estimates in the second step are identical to the corresponding estimates in the first approach TFP^{OP}. The Olley–Pakes type of TFP for ordinary firm i in industry j is obtained once the estimated coefficient $\hat{\beta}_k^{ord}$ is obtained:

$$\ln TFP_{ijt}^{ord} = \ln Y_{it} - \hat{\beta}_m^{ord} \ln M_{it} - \hat{\beta}_k^{ord} \ln K_{it} - \hat{\beta}_l^{ord} \ln L_{it}. \tag{B.7}$$

The estimates for processing firms have two important differences from those for ordinary firms. First, the coefficients of all inputs are allowed to be different because processing firms could use different technologies from ordinary firms. Second, because processing firms, by definition, are both importers and exporters, I do not need to introduce the export dummy or the import dummy in their investment function or the fourth-order polynomials. That is, the polynomials for processing firms are as follows:

$$g^{proc}(\ln K_{it}, I_{it}, WTO_t, SOE_{it}) = (\gamma_0 + \gamma_1 WTO_t + \gamma_2 SOE_{it}) \sum_{h=0}^{4}\sum_{q=0}^{4} \delta_{hq}^{proc} (\ln K_{it})^h I_{it}^q. \tag{B.8}$$

The rest of the procedures for processing firm TFP are the same as their counterparts for non-processing firms. The Olley–Pakes type of TFP for processing firm i in industry j is obtained as follows:

$$\ln TFP_{ijt}^{proc} = \ln Y_{it} - \hat{\beta}_m^{proc} \ln M_{it} - \hat{\beta}_k^{proc} \ln K_{it} - \hat{\beta}_l^{proc} \ln L_{it}. \tag{B.9}$$

I hence obtain two different sets of TFP for ordinary firms and processing firms. Their estimated input coefficients and measured TFP are shown in Appendix Table B1. The series of TFP^{OP1} is obtained by stacking them together.

B.3. TFP^{OP2} *with Learning from Processing*

Following De Loecker (2013), I now allow firms to learn from processing trade. Therefore, the export dummy is endogenously correlated with firm investment.

To obtain TFP^{OP2}, the difference from standard Olley–Pakes estimates is the first-step estimation. I first insert the processing dummy, PE_{it}, into the investment function as follows:

$$I_{it} = \tilde{I}(\ln K_{it}, \upsilon_{it}, FX_{it}, IM_{it}, WTO_t, SOE_{it}, PE_{it}). \tag{B.10}$$

Therefore, the inverse function of (B.10) is $\upsilon_{it} = \tilde{I}^{-1}(\ln K_{it}, I_{it}, FX_{it}, IM_{it}, WTO_t, SOE_{it}, PE_{it})$. To capture the possible learning effects from processing, the export decision was presumed to be made prior to the realisation of firm productivity. Hence, the productivity processing function $g(\cdot)$ is defined as $\beta_k \ln K_{it} + \upsilon_{it+1}$ where the productivity realisation υ_{it+1} uses the following polynomial specification as in De Loecker (2013):

$$\upsilon_{it+1} = \sum_{s=0}^{4}\sum_{m=0}^{4} \beta_{sm} PE_{it}^s \upsilon_{it}^m + \zeta_{it+1} \tag{B.11}$$

Table B1

Estimates of Olley–Pakes TFP by Processing and Ordinary Firms Separately

Chinese industry	Ordinary firms			Processing		
	Labour	Materials	Capital	Labour	Materials	Capital
13	0.242	0.875	0.052	0.116	0.884	0.066
14	0.023	0.926	0.050	0.037	0.925	0.074
15	0.185	0.508	0.268	0.243	0.505	0.088
17	0.017	0.884	0.059	0.089	0.834	0.041
18	0.054	0.858	0.076	0.177	0.669	0.142
19	0.126	0.895	0.023	0.118	0.808	0.000
20	0.126	0.895	0.023	0.044	0.913	0.003
21	0.055	0.917	0.042	0.101	0.873	0.103
22	0.111	0.907	0.008	0.027	0.896	0.063
23	0.023	0.821	0.039	0.105	0.836	0.025
24	0.068	0.764	0.123	0.104	0.863	0.036
26	0.086	0.795	0.063	0.007	0.927	0.024
27	0.108	0.862	0.040	0.038	0.860	0.038
28	0.116	0.789	0.033	0.016	0.837	0.041
29	0.061	0.569	0.174	0.073	0.938	0.032
30	0.118	0.633	0.182	0.125	0.696	0.114
31	0.073	0.851	0.047	0.050	0.870	0.035
32	0.046	0.976	0.051	0.038	0.961	0.010
33	0.053	0.815	0.080	0.055	0.850	0.076
34	0.041	0.867	0.048	0.044	0.883	0.026
35	0.065	0.875	0.024	0.032	0.917	0.026
36	0.090	0.823	0.076	0.038	0.869	0.111
37	0.058	0.888	0.047	0.054	0.924	0.029
39	0.013	0.830	0.103	0.102	0.826	0.000
40	0.071	0.831	0.072	0.086	0.878	0.086
41	0.081	0.906	0.015	0.139	0.567	0.168
42	0.055	0.917	0.045	0.142	0.818	0.094

Notes. This Table reports the estimates of log of Olley–Pakes TFP ($\ln TFP^{OP1}$) by separating ordinary firms and processing firms. The Chinese industries and associated codes are classified as follows: processing of foods (13), manufacture of foods (14), beverages (15), textiles (17), apparel (18), leather (19), timber (20), furniture (21), paper (22), printing (23), articles for cultures and sports (24), petroleum (25), raw chemicals (26), medicines (27), chemical fibres (28), rubber (29), plastics (30), non-metallic minerals (31), smelting of ferrous metals (32), smelting of non-ferrous metals (33), metal (34), general machinery (35), special machinery (36), transport equipment (37), electrical machinery (39), communication equipment (40), measuring instruments (41) and manufacture of artwork (42). I do not report the standard errors for each estimated coefficient to save space, although they are available upon request.

with $E(\zeta_{it+1} PE_{it}) = 0$. Note that firm innovation ζ_{it+1} thus is different from the standard Olley–Pakes step where $\zeta_{it+1} = \upsilon_{it+1} - \upsilon_{it}$. Compared with other dummies, such as the exporting dummy, the processing dummy is not only used in the second-step estimates, but also in the first-step estimates. Similarly, the inverse investment function can be characterised as the following control function:

$$\upsilon_{it} = (\lambda_0 + \lambda_1 WTO_t + \lambda_2 FX_{it} + \lambda_3 IM_{it} + \lambda_4 PE_{it} + \lambda_5 SOE_{it}) \sum_{h=0}^{4} \sum_{q=0}^{4} \delta_{hq} (\ln K_{it})^h I_{it}^q.$$

The second-step estimates are standard as above. After obtaining the coefficients of capital, labour and materials, the TFP^{OP2} is calculated as follows:

$$\ln TFP_{ijt}^{OP2} = \ln Y_{it} - \hat{\beta}_m \ln M_{it} - \hat{\beta}_k \ln K_{it} - \hat{\beta}_l \ln L_{it}. \tag{B.12}$$

Appendix C. Derivation of Domar-aggregation Productivity

This Appendix interprets how to add firm productivity to economy-wide aggregate productivity using Domar's (1961) weight under an open-economy set-up. The Appendix draws heavily from OECD (2001) and Feenstra *et al.*(2013*a*). The challenging part of the aggregation comes from the fact that domestic intermediate inputs used by firms do not show up in the economy-wide production possibility frontier (PPF), as they represent intra-industry flows that are absorbed in a process of vertical integration. To concretise this idea, consider the following PPF:

$$T(FA, N, IM, \pi) = 0, \tag{C.1}$$

where FA denotes China's final absorption (or equivalently, final demand), N denotes all domestic primary inputs such as capital and labour, IM is imported intermediate inputs and π is aggregate TFP. By assuming inputs are homogenous of degree zero in FA, N, IM and π and perfectly competitive markets, the productivity change can be traced as follows:

$$\frac{\mathrm{d}\ln\pi}{\mathrm{d}t} = \frac{\mathrm{d}\ln FA}{\mathrm{d}t} - \frac{P_N N}{P_{FA}FA}\frac{\mathrm{d}\ln N}{\mathrm{d}t} - \frac{P_{IM}\,IM}{P_{FA}\,FA}\frac{\mathrm{d}\ln IM}{\mathrm{d}t}, \tag{C.2}$$

where $(P_N N)/(P_{FA}FA)$ is the share of primary inputs in total final absorption and $(P_{IM}\,IM)/(P_{FA}\,FA)$ is the share of imported intermediate inputs in total final absorption. Both terms sum to unity because of zero profit in a perfectly competitive set-up. To link the aggregate economy with firm-level economic activities, each term in (C.2) can be decomposed as follows:

$$\begin{aligned}\frac{\mathrm{d}\ln FA}{\mathrm{d}t} &= \sum_i \frac{P^i FA^i}{P_{FA}\,FA}\frac{\mathrm{d}\ln FA^i}{\mathrm{d}t}\\ \frac{\mathrm{d}\ln N}{\mathrm{d}t} &= \sum_i \frac{P_N^i N^i}{P_N N}\frac{\mathrm{d}\ln N^i}{\mathrm{d}t}\\ \frac{\mathrm{d}\ln IM}{\mathrm{d}t} &= \sum_i \frac{P_{IM}^i\,IM^i}{P_{IM}\,IM}\frac{\mathrm{d}\ln IM^i}{\mathrm{d}t}.\end{aligned} \tag{C.3}$$

That is, aggregated final demand (aggregated primary inputs, aggregated imported intermediate inputs) can be written as a weighted average of firms' demand (primary inputs, imported intermediate inputs). By inserting (C.3) back into (C.2), I obtain:

$$\frac{\mathrm{d}\ln\pi}{\mathrm{d}t} = \sum_i \frac{P^i\,FA^i}{P_{FA}\,FA}\frac{\mathrm{d}\ln FA^i}{\mathrm{d}t} - \frac{P_N\,N}{P_{FA}\,FA}\left(\sum_i \frac{P_N^i\,N^i}{P_N\,N}\frac{\mathrm{d}\ln N^i}{\mathrm{d}t}\right) - \frac{P_{IM}\,IM}{P_{FA}\,FA}\left(\sum_i \frac{P_{IM}^i\,IM^i}{P_{IM}\,IM}\frac{\mathrm{d}\ln IM^i}{\mathrm{d}t}\right). \tag{C.4}$$

Turning to measures of firm productivity, consider the following production function, which is homogenous of degree one:

$$Y^i = \pi^i f(N^i, M^i, IM^i), \tag{C.5}$$

where Y^i, N^i, M^i, and IM^i denote firm i's output, primary inputs, domestic intermediate inputs and imported intermediate inputs, respectively. π^i is the Hicks-neutral TFP. Total differentiate (C.5) to obtain the following equation:

$$\frac{\mathrm{d}\ln\pi^i}{\mathrm{d}t} = \frac{\mathrm{d}\ln Y^i}{\mathrm{d}t} - \frac{P_N^i\,N^i}{P^i\,Y^i}\frac{\mathrm{d}\ln N^i}{\mathrm{d}t} - \frac{P_M^i\,M^i}{P^i\,Y^i}\frac{\mathrm{d}\ln M^i}{\mathrm{d}t} - \frac{P_{IM}^i\,IM^i}{P^i\,Y^i}\frac{\mathrm{d}\ln IM^i}{\mathrm{d}t}. \tag{C.6}$$

Note that each firm gets zero profit as the market structure is perfect competition, which implies:

$$P^i\,Y^i = P_N^i\,N^i + P_M^i\,M^i + P_{IM}^i\,IM^i. \tag{C.7}$$

Thus, the input shares in the last three terms in (C.6) sum to unity. Meanwhile, the firm's total demand (i.e. demand for intermediate goods and final goods) is equal to its production value (i.e. supply):

Table C1

Estimates of System-GMM Firm TFP by Industry

	Estimated coefficients			TFP	SD	Weighted TFD	Tests (p-value)		
Chinese industry	Labour (1)	Materials (2)	Capital (3)	(4)	(5)	TFP (6)	AR(1) (7)	AR(2) (8)	Hansen (9)
13	0.094	0.718	0.010	2.575	0.387	2.884	0.000	0.987	0.443
14	0.089	0.828	0.003	2.528	0.380	2.776	0.000	0.396	0.603
15	0.077	0.677	0.152	2.677	0.465	3.599	0.063	0.724	1.00
17	0.065	0.748	0.002	2.523	0.298	2.175	0.007	0.389	0.569
18	0.068	0.724	0.020	2.447	0.326	2.303	0.000	0.317	0.834
19	0.050	0.868	0.029	2.488	0.323	2.320	0.015	0.858	0.676
20	0.015	0.844	0.010	2.851	0.412	3.398	0.011	0.510	0.548
21	0.114	0.795	0.001	2.650	0.309	2.367	0.000	0.051	0.808
22	0.151	0.655	0.011	2.705	0.338	2.644	0.424	0.570	1.00
23	0.178	0.474	0.051	2.618	0.341	2.578	0.036	0.059	0.846
24	0.098	0.609	0.058	2.485	0.281	2.018	0.030	0.411	0.990
25	0.017	0.700	0.173	2.865	0.498	4.127	0.156	0.744	1.00
26	0.142	0.701	0.034	2.669	0.353	2.721	0.000	0.868	0.222
27	0.014	0.748	0.054	2.764	0.350	2.797	0.008	0.988	0.712
28	0.052	0.812	0.088	2.674	0.326	2.520	0.082	0.280	1.00
29	0.165	0.633	0.025	2.593	0.348	2.606	0.015	0.691	0.899
30	0.128	0.865	0.022	2.690	0.335	2.605	0.000	0.303	0.371
31	0.105	0.769	0.019	2.626	0.343	2.600	0.000	0.936	0.034
32	0.001	0.876	0.001	2.864	0.388	3.212	0.060	0.233	0.909
33	0.068	0.805	0.057	2.592	0.386	2.888	0.914	0.682	0.896
34	0.022	0.840	0.021	2.480	0.318	2.279	0.009	0.161	0.788
35	0.108	0.782	0.003	2.527	0.313	2.286	0.000	0.473	0.726
36	0.091	0.719	0.089	2.604	0.356	2.681	0.000	0.845	0.537
37	0.103	0.813	0.034	2.637	0.359	2.737	0.090	0.893	0.393
39	0.309	0.628	0.101	2.503	0.394	2.847	0.049	–	0.743
40	0.158	0.729	0.021	2.833	0.451	3.692	0.013	–	0.368
41	0.061	0.889	0.012	2.682	0.465	3.603	0.028	0.281	0.767
42	0.088	0.667	0.012	2.450	0.295	2.090	0.041	–	0.564

Notes. The Chinese industries and associated codes are classified as follows: processing of foods (13), manufacture of foods (14), beverages (15), textiles (17), apparel (18), leather (19), timber (20), furniture (21), paper (22), printing (23), articles for culture and sports (24), petroleum (25), raw chemicals (26), medicines (27), chemical fibres (28), rubber (29), plastics (30), non-metallic minerals (31), smelting of ferrous metals (32), smelting of non-ferrous metals (33), metal (34), general machinery (35), special machinery (36), transport equipment (37), electrical machinery (39), communication equipment (40), measuring instruments (41) and manufacture of artwork (42). I do not report the standard errors for each coefficient in first three columns to save space, which are available upon request. In all estimates, I include a one-period lag of capital, labour and materials. I also include a pure assembly dummy and its interaction with both current period and a one-period lag of capital, labour and materials. After obtaining system-GMM TFP in column (4), I compute the standard deviation of system-GMM TFP both across firms within an industry and across all firms, divide the industrial average to total average and multiply TFP in column (4) to obtain the weighted TFP in column (5). Numbers are p-values in Columns (6)-(8), which report various tests for the system-GMM TFP estimates.

$$P^i Y^i = \sum_k P^i\, Y^{ki} + P^i\, FA^i,$$

where prices for intermediate demand use and for final use are assumed to be equal for simplicity and Y^{ki} denotes firm i's deliveries of its product to firm k. Totally differentiate the above equation to obtain:

$$\frac{\mathrm{d}\ln FA^i}{\mathrm{d}t} = \frac{P^i\, Y^i}{P^i\, FA^i}\left(\frac{\mathrm{d}\ln Y^i}{\mathrm{d}t} - \sum_k \frac{P^i Y^{ki}}{P^i Y^i}\frac{\mathrm{d}\ln Y^{ki}}{\mathrm{d}t}\right). \tag{C.8}$$

Table C2

Additional One-step GMM Estimation with Tariffs and Production Functions

Regressand: Log of output ($\ln y_{it}$)	(1)	(2)	(3)	(4)
Firm output tariffs	−3.272**	−3.044**	−2.389*	−2.726**
	(−2.15)	(−2.11)	(−1.85)	(−2.03)
Firm output tariffs × fitted extent of processing	5.350**	5.012**	3.837	4.408*
	(2.03)	(2.03)	(1.63)	(1.87)
Firm input tariffs	−2.700***	−2.707***	−2.121**	−2.453**
	(−2.83)	(−2.78)	(−2.43)	(−2.57)
Firm input tariffs × fitted extent of processing	6.408***	6.035***	4.212**	4.826**
	(3.06)	(3.02)	(2.29)	(2.19)
Extent of processing	−1.062**	−1.055**	−0.749*	−0.933*
	(−2.03)	(−2.11)	(−1.65)	(−1.96)
Log of output at one lag ($\ln y_{it-1}$)	0.376***	0.357***	0.414***	0.358***
	(2.90)	(2.81)	(3.31)	(2.80)
Log of materials ($\ln M_{it}$)	0.553***	0.565***	0.563***	0.578***
	(15.79)	(14.60)	(15.28)	(13.91)
Log of materials at one lag ($\ln M_{it-1}$)	−0.147	−0.137	−0.161*	−0.128
	(−1.62)	(−1.50)	(−1.86)	(−1.44)
Log of labour ($\ln L_{it}$)	0.145***	0.145***	0.130***	0.129***
	(9.19)	(8.44)	(7.75)	(6.75)
Log of labour at one lag ($\ln L_{it-1}$)	−0.016	−0.014	−0.028	−0.013
	(−0.43)	(−0.41)	(−0.89)	(−0.39)
Log of capital ($\ln K_{it}$)	0.069***	0.066***	0.071***	0.065***
	(5.13)	(4.22)	(4.95)	(3.75)
Log of capital at one lag ($\ln K_{it-1}$)	−0.003	−0.002	−0.010	−0.007
	(−0.36)	(−0.26)	(−1.06)	(−0.70)
SOE indicator	−0.171***	−0.183***	−0.143***	−0.171***
	(−3.12)	(−3.15)	(−2.88)	(−2.95)
Foreign ownership indicator	0.113	0.117*	0.082	0.109*
	(1.62)	(1.73)	(1.38)	(1.73)
Year-specific fixed effects	Yes	Yes	Yes	Yes
Industry-specific fixed effects	Yes	Yes	Yes	Yes
Pure domestic firms dropped	No	Yes	No	Yes
Pure exporting firms dropped	No	No	Yes	Yes
Observations	15,308	13,675	13,383	11,750

Notes. This Table reports the one-step system-GMM dynamic panel-data estimation. t-values in parentheses are obtained using bootstrapped standard errors, corrected for clustering at the firm level. Significant at *10%, **5% and ***1%. Year-specific fixed effects and industry-level fixed effects are included. Column (1) includes the whole sample. Column (2) drops pure domestic firms. Column (3) drops pure exporting firms. Column (4) drops both pure domestic firms and pure exporting firms. As in Table 11, firm output (input) tariffs with initial time-invariant weight and one-period lag of tariffs are used as instruments for firm output (input) tariffs with initial time-invariant weight. Similarly, the interactions between fitted extent of processing obtained from the second-step Heckman estimates in Table 10 and firm output (input) tariffs with initial time-invariant weight and one-period lag of tariffs are used as instruments for the interaction between fitted extent of processing and firm output (input) tariffs.

By inserting (C.8) into (C.4), I obtain:

$$\frac{\mathrm{d}\ln\pi}{\mathrm{d}t}=\sum_i\frac{P^iY^i}{P_{FA}FA}\left(\frac{\mathrm{d}\ln Y^i}{\mathrm{d}t}-\sum_k\frac{P^iY^{ki}}{P^iY^i}\frac{\mathrm{d}\ln Y^{ki}}{\mathrm{d}t}-\frac{P^i_NN^i}{P^iY^i}\frac{\mathrm{d}\ln N^i}{\mathrm{d}t}-\frac{P^i_{IM}IM^i}{P^iY^i}\frac{\mathrm{d}\ln IM^i}{\mathrm{d}t}\right). \quad \text{(C.9)}$$

Finally, by definition, each delivery of firm k to firm i is also the intermediate input for firm i. That is, $Y^{ki}=M^{ik}$. Or equivalently, $\mathrm{d}\ln Y^{ki}/\mathrm{d}t=\mathrm{d}M^{ik}/\mathrm{d}t$. Then I have:

Table C3

Transitional Probability for State-owned Enterprises (SOEs)

Probability (%) Current period	Next period: SOEs	Next period: Non-SOEs	Total
SOEs	99.87	0.13	100
Non-SOEs	13.01	86.99	100
Total	98.21	1.79	100

Table C4

Transitional Probability for Foreign Firms

Probability (%) Current period	Next period: Foreign firms	Next period: Non-foreign firms	Total
Foreign firms	98.32	1.62	100
Non-foreign firms	0.96	99.04	100
Total	38.22	61.78	100

Table C5

Transitional Probability for Processing Firms

Probability (%) Current period	Next period: Non-processing	Next period: Processing	Total
Non-processing firms	85.90	14.10	100
Processing firms	34.14	65.86	100
Total	69.11	30.89	100

$$\sum_i \sum_k \frac{P^i Y^{ki}}{P_{FA} FA} \frac{\mathrm{d} \ln Y^{ki}}{\mathrm{d}t} = \sum_k \sum_i \frac{P^i M^{ik}}{P_{FA} FA} \frac{\mathrm{d} \ln M^{ik}}{\mathrm{d}t}. \tag{C.10}$$

The aggregated productivity measure can be readily obtained by inserting (C.10) into (C.9):

$$\frac{\mathrm{d} \ln \pi}{\mathrm{d}t} = \sum_i \frac{P^i Y^i}{P_{FA} FA} \left(\frac{\mathrm{d} \ln Y^i}{\mathrm{d}t} - \frac{P^i M^i}{P^i Y^i} \frac{\mathrm{d} \ln M^i}{\mathrm{d}t} - \frac{P_N^i N^i}{P^i Y^i} \frac{\mathrm{d} \ln N^i}{\mathrm{d}t} - \frac{P_{IM}^i IM^i}{P^i Y^i} \frac{\mathrm{d} \ln IM^i}{\mathrm{d}t} \right). \tag{C.11}$$

All terms in the parentheses of (C.11) are the change in firm productivity, as seen from (C.6). Therefore, I have:

$$\frac{\mathrm{d} \ln \pi}{\mathrm{d}t} = \sum_i \frac{P^i Y^i}{P_{FA} FA} \frac{\mathrm{d} \ln \pi^i}{\mathrm{d}t}. \tag{C.12}$$

That is, the economy-wide productivity change can be represented as a weighted sum of firm productivity change in which the weight is calculated by the firm's gross output value divided by the economy-wide total absorption (i.e. total gross output minus total trade surplus in an open economy like China). As this is initiated by Domar (1961), I hence call (C.12) the Domar-weight aggregated productivity (Tables C1–C5).

China Centre for Economic Research, National School of Development, Peking University

Submitted: 20 March 2013
Accepted: 01 November 2013

Additional Supporting Information may be found in the online version of this article:

Data S1.

References

Ackerberg, D., Caves, K. and Frazer, G. (2006). 'Structural identification of production functions', Working Paper, University of California, Los Angeles.

Ahn, J., Khandelwal, A. and Wei S. (2010). 'The role of intermediaries in facilitating trade', *Journal of International Economics*, vol. 84(1), pp. 73–85.

Arkolakis, C. and Muendler, M. (2012). 'The extensive margin of exporting products: a firm-level analysis', Working Paper 16641, NBER.

Amiti, M. and Davis, D. (2011). 'Trade, firms, and wages: theory and evidence', *Review of Economic Studies*, vol. 79(1), pp. 1–36.

Amiti, M. and Konings J. (2007). 'Trade liberalization, intermediate inputs, and productivity: evidence from Indonesia', *American Economic Review*, vol. 97(5), pp. 1611–38.

Baum, C., Schaffer, M. and Stillman S. (2007). 'Enhanced routines for instrumental variables/GMM estimation and testing', Economic Working Paper 667, Boston College.

Bernard, A. Eaton, J. Jensen, B. and Kortum S. (2003). 'Plants and productivity in international trade', *American Economic Review*, vol. 93(5), pp. 1628–90.

Bernard, A., Redding S. and Schott P. (2011). 'Multi-product firms and trade liberalization', *Quarterly Journal of Economics*, vol. 126(3), pp. 1271–318.

Blomström, M. and Kokko A. (1996). 'The impact of foreign investment on host countries: a review of the empirical evidence', Policy Research Working Paper 1745, World Bank.

Blundell, R. and Bond R. (1998). 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, vol. 87(1), pp. 11–143.

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*, vol. 97(2), pp. 339–51.

Bustos, P. (2011). 'Trade liberalization, exports and technology upgrading: evidence on the impact of MERCOSUR on Argentinean firms', *American Economic Review*, vol. 101(1), pp. 304–40.

Cai, H. and Liu Q.(2009). 'Does competition encourage unethical behaviour? The case of corporate profit hiding in China', Economic Journal, vol. 119(April), pp. 764–95.

Cameron, C. and Trivedi P. (2005), *Microeconometrics: Methods and Applications*, Cambridge: Cambridge University Press.

Cragg, J. and Donald S. (1993). 'Testing identifiability and specification in instrumental variables models', *Econometric Theory*, vol. 9(2), pp. 222–40.

Dai, M., Maitra, M. and Yu M. (2012). 'Unexceptional exporter performance in China? The role of processing trade', Working Paper, Peking University.

De Loecker, J. (2011). 'Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity', *Econometrica*, vol. 79(5), pp. 1407–51.

De Loecker, J. (2013). 'Detecting learning from exporting', *American Economic Journal: Microeconomics*, vol. 5(3), pp. 1–21.

De Loecker, J., Goldberg, P., Khandelwal, A. and Pavcnik N. (2012), 'Prices, markups and trade reform', Working Paper 17925, NBER.

De Loecker, J. and Warzynski F. (2012). 'Markups and firm-level export status', *American Economic Review*, vol. 102(6), pp. 2437–71.

Domar, E. (1961). 'On the measurement of technological change', Economic Journal, vol. 71(284), pp. 709–29.

Feenstra, R., and Hanson G. (2005). 'Ownership and control in outsourcing to China: estimating the property-rights theory of the firm', *Quarterly Journal of Economics*, vol. 120(2), pp. 729–62.

Feenstra, R., Li, Z. and Yu M. (2013*a*). 'Export and credit constraints under incomplete information: theory and empirical investigation from China', *Review of Economics and Statistics*, forthcoming.

Feenstra, R., Mandel B., Reinsdorf M., and Slaughter M. (2013*b*), 'Effects of terms of trade gains and tariff changes on the measurement of US productivity growth', *American Economic Journal: Economic Policy*, vol. 5(1), pp. 59–93.

Fernandes, A. (2007). 'Trade policy, trade volumes and plant-level productivity in Colombian manufacturing industries', *Journal of International Economics*, vol. 71(1), pp. 52–71.

Foster, L., Haltiwanger, J. and Syverson C. (2007). 'Reallocation, firm turnover, and efficiency: selection on productivity or profitability?' Working Paper, University of Chicago.

Ge, Y., Lai, H. and Zhu S. (2011). 'Intermediates imports and gains from trade liberalization', Working Paper, Michigan State University.

Goldberg, P., Khandelwal, A., Pavcnik, N. and Topalova, P. (2010), 'Imported intermediate inputs and domestic product growth: evidence from India', *Quarterly Journal of Economics*, vol. 125 (4), pp. 1727–67.

Grossman, G. and Helpman, E. (1994). 'Protection for sales', *American Economic Review*, vol. 84(4), pp. 833–50.

Harrison, A. (1994). 'Productivity, imperfect competition and trade reform', *Journal of International Economics*, vol. 36(1), pp. 53–73.

Heckman, J. and Vytlacil E. (1998). 'Instrumental variables methods for the correlated random coefficient model: estimating the average rates of return to schooling when the return is correlated with schooling', *Journal of Human Resources*, vol. 33(4), pp. 974–87.

Helpern, L., Koren, M. and Szeidl A. (2010). 'Imported inputs and productivity', Working Paper, University of California, Berkeley.

Horn, H., Lang, H. and Lundgren S. (1995). 'Managerial effort incentives, X-inefficiency and international trade', *European Economic Review*, vol. 39(1), pp. 117–38.

Hulten, C. (1978). 'Growth accounting with intermediate inputs', *Review of Economic Studies*, vol. 45(3), pp. 511–8.

Hsieh, C.-T. and Klenow P. (2009). 'Misallocation and manufacturing TFP in China and India', *Quarterly Journal of Economics*, vol. 124(4), pp. 1403–48.

Imbens, G. (2004). 'Nonparametric estimation of average treatment effects under exogeneity: a review', *Review of Economics and Statistics*, vol. 86(1), pp. 4–29.

Keller, W. and Yeaple S. (2009). 'Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States', *Review of Economics and Statistics*, vol. 91(4), pp. 821–31.

Kleibergen, F. and Paap R. (2006). 'Generalized reduced rank tests using the singular value decomposition', *Journal of Econometrics*, vol. 133(1), pp. 97–126.

Krishna, P. and Mitra D. (1999). 'Trade liberalization, market discipline and productivity growth: new evidence from India', *Journal of Development Economics*, vol. 56(2), pp. 447–62.

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

Levinsohn, J. and Petrin A. (2003). 'Estimating production functions using inputs to control for unobservable', *Review of Economic Studies*, vol. 70(2), pp. 317–41.

Lu, J., Lu, Y. and Tao, Z. (2010). 'Exporting behaviour of foreign affiliates: theory and evidence', *Journal of International Economics*, vol. 81(2), pp. 197–205.

Lu, D. (2011). 'Exceptional exporter performance? Evidence from chinese manufacturing firms', Working Paper, University of Chicago.

Ma, Y., Tang, H. and Zhang Y. (2011). 'Factor intensity, product switching, and productivity: evidence from Chinese exporters', Working Paper, Tufts University.

Manova, K., Wei, S-J. and Zhang Z. (2009). 'Firm exports and multinational activity under credit constraints', Working Paper 16905, NBER.

Melitz, M. (2003). 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica*, vol. 71(6), pp. 1695–725.

OECD (2001). 'Measuring productivity: measurement of aggregate and industry-level productivity growth', OECD, Paris, available at: http://www.oecd.org/std/productivity-stats (last accessed: 30 April 2014).

OECD (2010), 'Measuring globalisation: OECD economic globalization indicators', OECD, Paris, available at: http://www.oecd.org/sti/measuring-globalisation (last accessed: 30 April 2014).

Olley, S. and Pakes A. (1996). 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica*, vol. 64(6), pp. 1263–97.

Pavcnik, N. (2002). 'Trade liberalization, exit, and productivity improvements: evidence from Chilean plants', *Review of Economic Studies*, vol. 69(1), pp. 245–76.

Schor, A. (2004). 'Heterogenous productivity response to tariff reduction: evidence from Brazilian manufacturing firms', *Journal of Development Economics*, vol. 75(2), pp. 373–96.

Staiger, D. and Stock J. (1997) 'Instrumental variables regression with weak instruments', *Econometrica*, vol. 65 (3), pp. 557–86.

Trefler, D. (2004). 'The long and short of the Canada–US free trade agreement', *American Economic Review*, vol. 94(3), pp. 870–95.

Topalova, P. and Khandelwal A. (2011). 'Trade liberalization and firm productivity: the case of India', *Review of Economics and Statistics*, vol. 93(3), pp. 995–1009.

Wang, Z. and Yu Z. (2012). 'Trading partners, traded products, and firm performances of China's exporters and importers: does processing trade make a difference?', *World Economy*, vol. 35(12), pp. 1795–824.

Wooldridge, J. (2008). 'Instrumental variables estimation of the average treatment effect in the correlated random coefficient model', in (T. Fomby, R. Carter Hill, D.L. Millimet, J.A. Smith, E.J. Vytlacil, eds.), *Modelling and Evaluating Treatment Effects in Econometrics* (Advances in Econometrics, vol. 21), pp. 93–116, Bingley: Emerald Group Publishing Limited.

Yu, M. and Tian W. (2012). 'China's processing trade: a firm-level analysis', in (H. McMay and L. Song, eds.), *Rebalancing and Sustaining Growth in China*, pp. 111–48, Canberra: Australian National University E-press.