Returns to Scale, Productivity, and Markup: Revisit the Export Premium*

Xing HU† Yating JIANG‡ Hongsong ZHANG§

November 14, 2024

Abstract

The productivity effect of exports has been the foundation for many trade-related policies. However, empirical studies usually find a mixed or limited effect of exports on productivity. We solve this puzzle and show that increasing returns to scale and markups are two important sources of gains from exports, in addition to productivity. Because output prices are typically unavailable at the firm level, we developed a new method to estimate firm-level markups, productivity, and returns to scale consistently, using the widely available revenue data. We find that exports generate substantial efficiency gains, half contributed by increasing returns to scale and the other half by exports' productivity premium. Improved efficiency allows exporters to charge higher markups while offering lower prices. Taken together, exports increase firms' profits by about a quarter in the Chinese manufacturing industry and benefit consumers by reducing prices. Increasing returns to scale also explains why TFPR may fail to capture the export premium even when exports increase markup.

Keywords: Export premium, returns to scale, productivity, markup, production function estimation

*The authors thank Jan De Loecker, Jiandong Ju, Bingjing Li, Shengyu Li, Yao Amber Li, Jiawei Mo, Larry Dongxiao Qiu, Mark Roberts, Chang Sun, Heiwai Tang, James R. Tybout, Frederic Warzynski, Miaojie Yu, Xiaodong Zhu, and Yifan Zhang for their insightful comments. The authors also benefited from discussions with participants in the 2024 Econometric Society European Meeting, the 2024 Asia Meeting of the Econometric Society, 2024 North American Summer Meeting of the Econometric Society, 2024 China Industrial Organization conference, 2024 International Association for Applied Econometrics, 2023 Hong Kong Economic Association Biennial Conference, 2023 HKUST Conference on International Economics, 2023 Danish International Economics Workshop, and the 13th Annual Meeting of China Trade Research Group.

†HKU Business School, The University of Hong Kong, huxing@connect.hku.hk

‡School of Economics, Renmin University of China. Beijing, China. ytjiang@ruc.edu.cn

§HKU Business School, The University of Hong Kong, hszhang@hku.hk

1 Introduction

The productivity effect of exports has been the foundation for many trade-related policies. However, while the export premium on productivity is often documented in the theoretical literature and case studies[1], limited or even mixed evidence has been detected in empirical research (e.g., Bernard and Jensen, 1999; Keller, 2004) typically using revenue data.[2] Recently, Garcia-Marin and Voigtländer (2019) showed that this might be because increased productivity after export caused a reduction in output prices, which offset each other in the revenue productivity measure (TFPR). This paper examines the effects of exports on production efficiency and markup, emphasizing increasing returns to scale as an important source of efficiency gains from exports in addition to productivity.

This study is motivated by two facts observed from a large dataset from the Chinese manufacturing sector. First, exports substantially increase firms' markup, as proxied by the revenue-to-variable costs ratio. Second, export reduces prices substantially, as shown in a smaller dataset in which output prices are available, consistent with De Loecker et al. (2016) and Garcia-Marin and Voigtländer (2019). The increase in markup and decline in prices directly imply a gain in production efficiency from exports. In contrast to Garcia-Marin and Voigtländer (2019), who attribute the gain in production efficiency to productivity improvement under the assumption of constant returns to scale, we show that increasing returns to scale is equally important as the productivity channel. Moreover, with constant returns to scale, TFPR grows if and only if markup increases, contradicting the fact that in our data, TFPR has no obvious change, although markup increases substantially. Increasing returns to scale resolves this puzzle.

This paper explores the multiple sources of export premium, including markup, increasing returns to scale, and productivity. The empirical analysis uses a large dataset of the production and export of Chinese manufacturing firms at the firm level. Because only revenues but not output prices/quantities are recorded in this dataset[3], we develop a new method to estimate the firm-level markup and productivity jointly, using revenue and widely available variable input expenditure. The key idea is to control for firm-level markup in the production estimation using revenue-variable input expenditure ratios, which can be shown

[1]Some theoretical examples include Clerides et al. (1998) through learning by exporting, Holmes and Schmitz Jr (2001) through export-induced R&D investment, and Melitz (2003) through reallocation. Also, see Rhee et al. (1984) for a case study from South Korea.

[2]See Greenaway et al. (2005) and Greenaway and Kneller (2007) for a review of micro evidence.

[3]We use this data set to include as many observations as possible in the Chinese manufacturing industry. In contrast, the output quantity/price data used in the motivational facts are smaller and only cover about two-thirds of firms, which can be used as a robustness check.

as a function of markup and returns to scale. This idea inherits the critical insight of Hall (1988) and De Loecker and Warzynski (2012) in that both use the first-order conditions of cost minimization to establish the relationship between markup and revenue-variable input expenditure ratios. We differ in that while they use this relationship to calculate firm-level markup after estimating output elasticities using output quantity, our approach uses this relationship to control for markup in the production estimation and estimate markup and productivity jointly. Using this additional information, we address the unobserved output price problem by extending Klette and Griliches (1996) to allow firm-level markup. The extended method allows us to estimate the firm-level markup, productivity, and returns to scale consistently using revenue data under the same assumptions as in De Loecker and Warzynski (2012).

The new method has several advantages. First, it only requires widely available revenue and input expenditures data, not output quantity or prices. This largely expands the breadth of its potential applications. Second, the method estimates firm-level markup and productivity jointly, allowing us to explore the export premium through both productivity and markup. Finally, because the method controlled for firm heterogeneity in markup, unlike that in the literature, the estimated productivity and output elasticities (and hence returns to scale) are not contaminated by markup. Hence, it allows us to empirically explore the export premium through an enlarged market and (potentially increasing) returns to scale. These advantages allow us to dissect the multiple sources of gains from exports and evaluate the export premium more accurately.

There are three major findings. First, export has a significant physical productivity premium, echoing Garcia-Marin and Voigtländer (2019) and Li et al. (2017) that use output quantity data. Export raises exporters' productivity by about 1.5% after correcting for unobserved output prices and markup. In contrast to the markup-adjusted estimates, export has a negligible effect on the traditionally estimated revenue-based productivity (0.2-0.3%), although export increases markup substantially. This is different with Garcia-Marin and Voigtländer (2019), which demonstrates that TFPR increases in markup proportionally under constant returns to scale. We show that the no-export effect on TFPR (even with increased markup) is due to increasing returns to scale: with increasing returns to scale, producing more following export reduces the marginal production costs, which offsets the positive effect of productivity and markup, leading to an ambiguous effect of export on TFPR.

Second, exports increase firms' total market demand in terms of both market size and markup. In all of our specifications, export increases the demand shifter faced by the firm by 19.4-22.0% on average. The increased market demand allows the firm to produce more and potentially charge a higher markup. Indeed, export increases firms' markup by 1.0-1.3% on average, which serves as one crucial channel for exporters to increase their profitability.

Third, after correcting for markup, production shows substantially increasing returns to scale. In all of our specifications, returns to scale range between 1.09-1.11, and all estimates are significantly above 1. Correcting markup is crucial for consistently estimating returns to scale. When estimating the revenue function without correcting for markup, the estimation displays decreasing returns to scale in all specifications as in the literature (e.g., Yu, 2015). The intuition is straightforward: the revenue-based returns to scale are normalized by markup that is usually greater than 1. If the revenue-based returns to scale are between 0.9-1 as estimated in the literature and the markup is around 1.2, the implied true returns to scale are already greater than 1. Our result echos Klette and Griliches (1996), who also find increasing returns to scale after controlling for industry-level average markup. We improve it by allowing for firm heterogeneity in markup. This result is also consistent with De Loecker et al. (2016), who use output quantity data to estimate the (translog) production function and thus are unaffected by the unobserved markup problem. They find that 68% of their sample observations exhibit increasing returns to scale. Increasing returns to scale provides a new channel of export premium: the increased output after export reduces production costs, improving the exporters' efficiency. It also explains why detecting export premiums on TFPR is difficult.

The estimation results imply a large efficiency gain from export, by reducing the marginal cost of production by an average of 2.68%. This result echoes De Loecker et al. (2016) and Garcia-Marin and Voigtländer (2019) which rely on output quantity data. Increasing returns to scale plays an important role in firms' efficiency gains from exports. Our decomposition shows that increasing returns to scale contributes to about half of the efficiency gains, and the other half is due to improved productivity after export.

The improved production efficiency and the larger market improve exporters' profitability and consumer welfare. In the data, exporters' profit rate increases by 25.33%. This results from the increased markup and reduced production costs arising from export premium on productivity and increasing returns to scale. We calculate the changes in the exporters' output prices based on the estimated markup and changes in marginal costs. We find the exporters' price declined by 1.68% on average, implying an improvement in consumer welfare.

We validate our results by using a smaller sub-sample with firm-level output quantity. The unobserved markup is no longer a problem when estimating the physical output production function using traditional approaches. Based on the physical production function, exports increase the firm's productivity by 0.5% and production shows increasing returns to scale (1.079). When estimating the model using our method in this smaller sample (pretending no output quantity data), we find consistent results (0.6% for the productivity effect and 1.035 for the RTS) after correcting for the potential positive relationship between input and output prices (as proxies of input and output quality). In contrast, the traditional revenue-based approach shows decreasing returns to scale and fails to capture the productivity gains from exports. The results show that our results based on the new estimation methods are robust.

Our results are robust when using alternative estimation approaches such as Ackerberg et al. (2015). We also checked for a more flexible production function. In the Cobb-Douglas case, all variation in markup is driven by expenditure shares because output elasticities are constant. We relax this assumption to consider the more flexible translog production function, in which case both expenditure shares and the flexible output elasticities contribute to the changes in markup. In this case, we find an even higher export premium on productivity (3.0%). Returns to scale are increasing for most firms and increase in firm size, which echos De Loecker et al. (2016)'s finding in the relationship between marginal cost and output quantity. All other results are qualitatively and quantitatively similar to our main results. We also estimate the model industry by industry and the export effect and IRS appear in most industries. Finally, it is possible that the current productivity innovation and export status are positively correlated, which can result in an endogeneity problem. Our results are consistent when using an IV approach.

The paper first contributes to the literature on gains from exports. Although the export productivity premium is often used as the basis for many trade policies and is documented in the theoretical literature and case studies (e.g., Rhee et al., 1984), little and generally mixed evidence has been detected in empirical research (e.g., Bernard and Jensen, 1999; Keller, 2004; Grieco et al., 2022). We show that this is because exports may decrease output prices, which offsets the physical productivity gains as components of revenue productivity. This is consistent with De Loecker et al. (2016), who also documents that export causes a decline in output prices using data on output prices. The paper is closest to Garcia-Marin and Voigtländer (2019), who find an export productivity premium after controlling for the negative effect of exports on prices. We contribute in three ways. First, we find that increasing returns to scale is an important channel through which export improves efficiency. Second, increasing returns to scale also explains why a limited productivity premium is found using TFPR, even when export increases markup. Finally, we show that although exports reduce firms' output prices, their markup increases substantially due to efficiency gains. This improves exporters' profitability, serving as an important channel for firms to gain from exports.

The paper contributes to the large literature on production estimation (e.g., Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015) and markup estimation. The unobserved heterogeneity in output prices and markup biases productivity estimates (e.g., Klette and Griliches, 1996; Foster et al., 2008). Klette and Griliches (1996) addresses this problem partially by replacing the output prices in the revenue function with the demand function and estimating the revenue production function, assuming common markups for all firms. De Loecker (2011) and De Loecker and Warzynski (2012) propose an approach to estimate the firm-level markup and productivity with the knowledge of output quantity data(see Bond et al., 2021, for a review), which are not widely available in most firm-level datasets. We provide a new method for consistently estimating firm-level markup, productivity, and returns to scale jointly using revenue data, exploiting the widely available data on revenue-to-variable cost ratios.

The paper also provides empirical support for the new growth theory (Romer, 1986) and new trade theory (Krugman, 1980, 1995), whose results are based on the assumption of increasing returns to scale. Empirically, Romer (1986) presented evidence on increasing returns to scale at the macro level and used it as the fundamental to explain long-term growth. However, this result is questioned by Basu and Fernald (1997), who show that the estimated returns to scale vary largely depending on different levels of aggregation and could even be decreasing using gross output value or value-added data. Moreover, most of the estimates using microdata find decreasing returns to scale without correcting for markup (e.g., Yu, 2015; Dai et al., 2016, using the same data), except a few exceptions (Klette and Griliches, 1996). We show that this might be due to the existence of unobserved large markup, which contaminates the estimation of output elasticity and thus returns to scale using revenue data. We extend Klette and Griliches (1996) and provide an approach to estimate firm-level markup and productivity simultaneously using revenue data. We provide strong evidence that, after correcting the impact of markup, production shows strong increasing returns to scale, even at the firm level. This result supports the argument of new growth theory and new trade theory.

In the rest of the paper, Section 2 introduces the data and motivational facts. Section 3 develops a framework to jointly estimate firm-level markup, productivity, and returns to scale consistently based on revenue data. Section 4 presents the estimation results. Section 5 validates our results in a smaller sample with rare output quantity data at the firm level. In Section 6, we calculate the gains from export in firm profitability and consumer welfare and evaluate the relative importance of the multiple sources of export premium. Section 7 checks the robustness of our results, and Section 8 concludes.

2 Data and Motivation

2.1 Data and Summary Statistics

Our empirical analysis focuses on the impact of direct ordinary exporting in the Chinese manufacturing industry. The analysis mainly uses two datasets.

The first dataset is the Annual Survey of Industrial Enterprises (ASIE) from 1998 to 2008, collected by the National Bureau of Statistics of China (NBS). The dataset contains detailed input and output information of all State Owned Enterprises (SOE) and non-SOE firms whose annual sales are greater than or equal to 5 million RMB (around 0.604 million U.S. dollars according to the exchange rate of 2000). The variables include total sales, material expenditure, capital stock, wage expenditure, labor employed, and other production information at the firm level. However, as in many other firm-level surveys, the dataset does not record these firms' output quantities or prices.

We clean the data using the following rules. First, we drop observations that contain negative or missing values for the following variables: total sales, total revenue, total employment, fixed capital, material expenditure, the cost of goods sold, city code, and holdings, following Feenstra et al. (2014) and Dai et al. (2016). Second, we only keep observations whose liquid assets are less than or equal to total assets, whose total fixed assets are less than or equal to total assets, whose net value of fixed assets is less than or equal to total assets, and whose labor employed is greater than or equal to eight, as in Mo et al. (2021). Moreover, we winsorize the data using P1 and P99 as the cutoff for major variables used in the analysis (i.e., total sales, labor employed, fixed capital, material expenditure, labor expenditure, and depreciation of the year).

The second dataset is China's custom records from 2000 to 2006 collected by China's General Administration of Customs (GAC). This dataset contains transaction-level import and export information, including trade type, price, quantity, etc. The trade type allows us to separate ordinary trade from processing trade, and we focus on ordinary exports.

We merge these two datasets following Yu (2015). The matched sample covers 15.28% of the observations (17.02% of firms) from the NBS dataset. Finally, our sample contains 1,234,292 observations from 401,020 firms, among which 188,633 observations from 68,267 firms were involved in direct ordinary export. This is highly consistent with those papers focusing on the same datasets and adopting similar merging methods. [4]

[4] Dai et al. (2016) reported 1,244,382 observations from 424,546 firms with 688,65 direct exporters, and Ge et al. (2015) reported 77,087 direct exporters in their merged datasets respectively.

The matched sample is our main dataset for the empirical analysis. Table 1 reports the summary statistics of the main variables used in our empirical analysis. Besides the directly reported variables, we also calculate a measure of "raw markup ($\tilde{\mu}_{jt}$)", defined as the ratio of total sales to the sum of material expenditure and wage expenditure. We show in Appendix A.1 that the true markup (μ_{jt}), defined as the price-marginal cost ratio, equals the raw markup ($\tilde{\mu}_{jt}$) normalized by returns to scale. Hence, the raw markup is an important source of variation for true markup. In the special case with constant output elasticities (e.g., Cobb-Douglas production function), the raw markup is the only source of variation in the true markup. The insight is similar to De Loecker (2011) and De Loecker and Warzynski (2012).

Table 1: Summary Statistics

Statistics	Median	Mean	sd	IQR	IDR
Total Sales (million USD)	1.977	4.456	7.357	3.575	9.742
Material Expenditure (million USD)	1.546	3.506	5.866	2.806	7.663
Capital Stock (million USD)	0.442	1.392	2.927	1.066	3.296
Wage Expenditure (million USD)	0.121	0.244	0.371	0.205	0.534
Labor Employed	101	183.083	241.282	152	380
Raw Markup ($\tilde{\mu}$)	1.154	1.182	0.239	0.228	0.470
Export Probability	0	0.153	0.335	0	1
Export Share (Full sample)	0	0.067	0.215	0	0.153
Export Share (Con. on Export)	0.342	0.440	0.373	0.753	0.986
Number of Exporting Firms	68,267 (17.02%)				
Total Number of Firms	401,020				
Exporting Observations	188,633 (15.28%)				
Observations	1,234,292				

Note: All monetary values here are in millions of 2000 U.S. dollars.

Besides the main datasets, we motivate this study and validate our new methods using a smaller sample of firms that have output quantity data. The output quantity survey from 2000 to 2006 contains measures of output quantity at the firm-product level, allowing us to calculate a rough measure of firm output prices. After merging it with our main data, we have 444,475 observations with output quantity. Among them, 88,487 observations were from exporting firms, accounting for a bit over one-third of the 243,472 exporting observations in our main sample. Although we did not use the smaller dataset in our main analysis due to its incomplete coverage, it provides good motivation for our study and validation of our new estimation method.

2.2 Motivational Facts

As one of the most important events in the Chinese economy in the 21st century so far, China's accession to the WTO in 2001 led massive Chinese firms to export. Did Chinese exporters enjoy efficiency gains from exports? This subsection presents two motivational facts that strongly support large efficiency gains from export: a large export premium in markup and a substantial reduction of output prices after export.

2.2.1 Markup Premium of Export

Although firm-level markup is not directly observed in the data, we use the revenue-to-variable cost ratio (raw markup) as a proxy for markup. In the special case of the Cobb-Douglas production function, the raw markup is the only source of variation in true markup. Table 1 shows that there is a large heterogeneity in markup across firms, with an interquartile range of 0.228. Figure 1 compares the average markup of firms before and after they start exporting. It is shown that the exporters' markup is significantly higher after they start exporting than before. The margin is about 2 percentage points.

2.2.2 Price Decline after Export

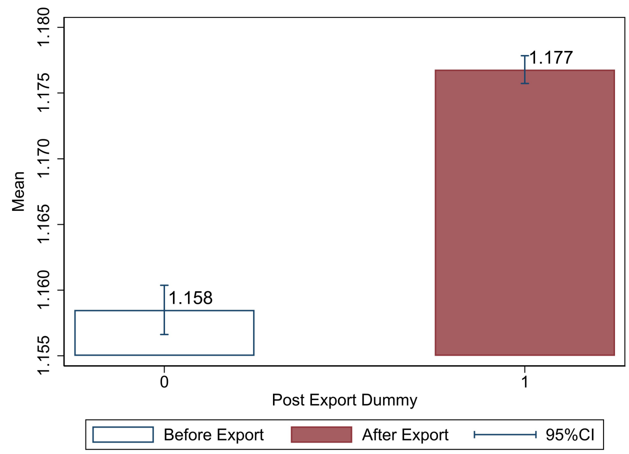
Although output quantity and prices are not available in our main data, the smaller quantity survey provides some information to calculate firms' output prices. We define the firm-level output prices as the ratio of its sales over an index of output quantity.[5] We estimate the effect of export on the prices charged by exporters, after controlling for firm FE, time FE, and a series of firm characteristics.

$$AveragePrice_{jt} = \beta_0^P + \beta_{exp}^P D_{jt}^{exp} + \beta^X \mathbf{X}_{jt} + \gamma_j + \gamma_t + \xi_{jt}^P, \tag{1}$$

where D_{jt}^{exp} is the export dummy equaling 1 if firm j is exporting at time t and 0 otherwise. Thus, the coefficient of interest, β_{exp}^P, captures the average export effect on firm-level prices. $\mathbf{X}_{jt}$ is a set of control variables, including firm size (measured by the total sales), ownership,

[5] One caveat is that because we only observe the firm-product level output quantity and firm-level output value, in principle we cannot calculate firm-level output price index for multi-product firms. In this exercise, we tried to use different methods to calculate the average price at the firm level, and the results are robust. First, we only use single-product firms, which allows us to calculate firm-level output prices accurately. Second, we define firm-level output price as the ratio of its total sales to simple aggregation of reported quantity. Third, we define firm-level output price as the ratio of its total sales to the quantity of the major product with the highest quantity. Although each of the three methods is subject to some limitations, all of them show consistent results. Moreover, we also checked the single-product firm's price change with the quantity (price) information. The results are robust and are reported in Table B1 in the appendix.

Figure 1: Raw Markup Mean Before&After Exporting



Note: The range represents 95% confidence interval.

and capital intensity. γ_j and γ_t represent firm fixed effect and year fixed effect, respectively. ξ_{jt}^P is an i.i.d shock to the firm's average price.

The estimation results are reported in Table 2. It is shown that export reduces the prices charged by exporters significantly, which is consistent with the findings in Li et al. (2017). In the fully-fledged specification in column 4, export reduces the firm's output price by 3.3% on average.

The lowered output prices, together with the increased markup as shown above, imply an export premium on production efficiency: export reduces marginal production costs, so that the firm can charge a higher markup, but at a lower price. The gain in production efficiency may arise from an export productivity premium or increased scale returns (if any) given the larger market after export. However, if the markup is greater than 1 and unobserved, it will bias down the estimate of returns to scale and its resulting export premium based on the revenue production function. In Section 3, we develop a new method to estimate the multiple sources of gains from exports, using widely available revenue data and input expenditure.

2.3 Key Insights: Increasing Returns to Scale and TFPR

Why may TFPR fail to capture the efficiency gains from exports? To provide insight, we decompose TFPR into its contributing factors. By definition, changes in TFPR are contributed by changes in physical productivity and prices,

$$\Delta \text{TFPR} = \Delta p + \Delta\omega = \Delta \ln \mu + \Delta mc + \Delta\omega. \tag{2}$$

Table 2: Average Price Change

Parameter	(1) P_{jt}	(2) $\ln(P_{jt})$	(3) P_{jt}	(4) $\ln(P_{jt})$
D_{jt}^{exp}	-7.063**	0.010	-8.985***	-0.033***
	(2.8332)	(0.0100)	(2.8299)	(0.0100)
Firm Size (Sales)			YES	YES
Capital Intensity			YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Adjusted R^2	0.762	0.919	0.762	0.921
Observations	88,487	88,487	88,487	88,487

Note: Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Where $p, \omega, \ln \mu$, and mc represent the logarithm of output price, physical productivity, markup, and marginal costs of production, respectively. $\triangle X$ represents changes in X. The second equality in (2) holds following the definition of markup, $\ln \mu = p - mc$. In the special case of constant returns to scale, a 1-percent increase in productivity leads to a 1-percent decline in marginal costs, offsetting each other. Therefore, TFPR changes if and only if there is a change in markup. Garcia-Marin and Voigtländer (2019) provide empirical support using firm-level data from Colombia and Chile.

However, Garcia-Marin and Voigtländer (2019) may not be the whole story, because, while export increases markup as shown above, its effect on TFPR is negligible, as will be shown in Section 4 using the Chinese data. To rationalize this observation, we generalize Garcia-Marin and Voigtländer (2019) to allow for non-constant returns to scale. In this case, (2) can be written as:[6]

$$\begin{aligned}\Delta \text{TFPR} &= \Delta p + \Delta \omega = \Delta \ln \mu + \Delta mc + \Delta \omega \\ &= \Delta \ln \mu + \frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M}\Delta y - \frac{1}{\alpha_L+\alpha_M}\Delta \omega + \Delta \omega \\ &= \Delta \ln \mu + \big(1 - \frac{1}{\alpha_L+\alpha_M}\big)(\Delta \omega - \Delta y), \end{aligned} \tag{3}$$

where y represents output quantity (in logarithm) and (α_L, α_M) represent output elasticities of labor and material, respectively. (3) shows how returns to scale matter. With increasing returns to scale ($\alpha_L + \alpha_M > 1$), an increase in physical productivity (ω) will increase TFPR conditional on output level. This is because, with increasing returns to scale, a 1-percent increase in productivity reduces marginal costs by less than 1 percent, conditional on output

[6]Please refer to Appendix A.5 for the detailed derivations.

level. More importantly, producing more after starting export reduces TFPR through increasing returns to scale, which may offset the potentially positive effect of export on markup and productivity. As a result, even if exports increase markup and productivity, the effect on TFPR may still be ambiguous. This explains why the literature may fail to find an export premium based on TFPR, even when markup increases following export.

The remaining question is: Does technology show increasing returns to scale at the firm level? In the following sections, we show that traditional production function estimation approaches based on revenue data underestimate returns to scale. After controlling for firm heterogeneity in markup, firm-level technology shows strong increasing returns to scale, and it is an important source of gains from export in addition to productivity and markup.

3 Empirical Model

We develop an empirical model to explore the efficiency and demand gains from exports. Because output prices are not reported in our main data as in most micro datasets, we developed a new method to consistently estimate firm-level markup, productivity, and returns to scale jointly, using the widely available revenue data and variable input expenditure. The key idea is to use the variation in revenue-to-total variable cost ratios to control for markup, sharing a similar insight as De Loecker and Warzynski (2012). As an obvious advantage, while the latter requires output quantity data to estimate output elasticities first and then use it to construct firm-level markup, our new method does not require output quantity or prices. Instead, we estimate the markup and output elasticities jointly by expanding Klette and Griliches (1996) to allow firm heterogeneity in markup, using the additional variation in the revenue-to-total variable cost ratios to control for unobserved markup.

3.1 Setup

Firms produce heterogeneous products and compete in a monopolistic-competitive market.

Production. The gross production function of firm j at time t is Cobb-Douglas:

$$Y_{jt} = e^{(\omega_{jt}+\xi_{jt})} L_{jt}^{\alpha_L} M_{jt}^{\alpha_M} K_{jt}^{\alpha_K}, \tag{4}$$

where Y_{jt} is the output quantity. ω_{jt} is the structural productivity observed by the firm before production, and ξ_{jt} is an idiosyncratic productivity shock. The firm uses labor (L_{jt}), intermediate input (M_{jt}), and capital (K_{jt}) as inputs, as summarized in the input vector $\mathbf{X}_{jt}^{I} = (L_{jt}, M_{jt}, K_{jt})$. $\theta = (\alpha_L, \alpha_M, \alpha_K)$ is the vector of the corresponding output elasticity to be estimated. The structural productivity evolves following the AR(1) process with persistence parameter β_1^{ω} and constant β_0^{ω}:

$$\omega_{jt} = \beta_0^{\omega} + \beta_1^{\omega}\omega_{jt-1} + \beta_{exp}^{\omega} D_{jt}^{exp} + \epsilon_{jt}^{\omega}. \tag{5}$$

Where D_{jt}^{exp} is a dummy variable indicating whether firm j is involved in export business at time t, as defined in (1). β_{exp}^{ω} captures the immediate exporting premium on productivity. ϵ_{jt}^{ω} is the current innovation shock on firm j's productivity and is assumed to be i.i.d. across firms and time.

Demand and Revenue Function. The demand function is

$$y_{jt} = \eta_{jt} p_{jt} + \varphi_{jt}. \tag{6}$$

The demand function is characterized by the demand elasticity η_{jt} and demand shifter φ_{jt}, both of which vary across firms and over time. Thus, the model allows for firm heterogeneity in markup and market size. Using (6) to replace p_{jt} yields the (logarithm) revenue function $r_{jt} = p_{jt} + y_{jt} = [y_{jt}(\eta_{jt}+1) - \varphi_{jt}]/\eta_{jt}$. Note that the demand elasticity, η_{jt}, and the demand shifter, φ_{jt}, are flexible and can depend on the export decisions of the firms. With monopolistic competition, firm-level markup equals $\mu_{jt} \equiv \eta_{jt}/(\eta_{jt}+1)$. Hence, the revenue production function can be rewritten in the following form using (4):

$$r_{jt} = p_{jt} + y_{jt} = \frac{\alpha_L}{\mu_{jt}} l_{jt} + \frac{\alpha_M}{\mu_{jt}} m_{jt} + \frac{\alpha_K}{\mu_{jt}} k_{jt} + \frac{1}{\mu_{jt}}\omega_{jt} + (1 - \frac{1}{\mu_{jt}})\varphi_{jt} + \frac{1}{\mu_{jt}}\xi_{jt}, \tag{7}$$

where l_{jt}, m_{jt}, and k_{jt} are the logarithm of L_{jt}, M_{jt}, and K_{jt} respectively. This equation is an extension of Klette and Griliches (1996) to the more general case with firm heterogeneity in markup.

Three lessons are learned from this equation. First, if there is markup ($\mu_{jt} > 1$), the output elasticities and, as a result, returns to scale will be understated if we estimate the revenue production function as a proxy of the physical production function. Second, the revenue productivity, typically defined as $\frac{1}{\mu_{jt}}\omega_{jt} + (1 - \frac{1}{\mu_{jt}})\varphi_{jt}$ in the literature, can bias the estimate of the policy effect on productivity, as it combines the effect on productivity, markup, and demand factors. Third, firm heterogeneity in markup enters the error term and, if unobserved, will bias the estimates of (7). It is also difficult to find valid instrumental variables for markup, because markup correlates with inputs, outputs, and market conditions.[7]

[7]Relevant discussion can also be found in Bond et al. (2021) and Doraszelski and Jaumandreu (2021). Doraszelski and Jaumandreu (2019) also discusses the problem of omitting the demand heterogeneity when

3.2 Joint Estimation of Markup and Production Function

We develop a new approach to consistently estimate (7) for productivity, markup, and returns to scale, using the widely available revenue and input expenditure data.. Our strategy is to find a reliable proxy for markup and control it in the production estimation process. Specifically, we use the widely available data on the revenue-to-total variable cost ratios to control for firm heterogeneity in markup. This proxy is implied by firms' cost minimization assumptions in the spirit of De Loecker and Warzynski (2012).

Assume that firms choose labor and material to minimize production costs to produce output Y_{jt}, given productivity and quasi-fixed capital stock. Denote $\mathbf{X}_{jt}^{V} = (L_{jt}, M_{jt})$ as the vector of variable inputs and $\mathbf{P}_{jt}^{V} = (P_{jt}^{L}, P_{jt}^{M})$ the corresponding variable input prices. The associated Lagrangian function is:

$$\mathcal{L}(\mathbf{X}_{jt}^{V}) = \mathbf{P}_{jt}^{V} \cdot \mathbf{X}_{jt}^{V} + \lambda_{jt} \left(Y_{jt} - e^{(\omega_{jt} + \xi_{jt})} L_{jt}^{\alpha_L} M_{jt}^{\alpha_M} K_{jt}^{\alpha_K} \right). \tag{8}$$

The Lagrange multiplier λ_{jt} represents the marginal cost of production. Utilizing the first-order conditions for both labor and material, we can derive the relationship between markup, returns to scale, and revenue-variable cost ratio as follows:

$$\mu_{jt} = (\alpha_L + \alpha_M) \cdot \frac{R_{jt}}{P_{jt}^{L} L_{jt} + P_{jt}^{M} M_{jt}} = S^{V} \cdot \tilde{\mu}_{jt}, \tag{9}$$

where $S^{V} \equiv (\alpha_L + \alpha_M)$ is the sum of output elasticities of all Variable inputs, and it measures the returns to scale of variable inputs (RTSV). (9) shows that markup is known from the raw markup, up to RTSV.[8] Our strategy is to use this equation to control for unobserved markup when estimating the revenue function (7).

To proceed, we parameterize the demand shifter in the following form, $\varphi_{jt} = \bar{y}_{Jt} - \eta_{jt}\bar{p}_{Jt} + \beta_{exp}^{D} D_{jt}^{exp} + (1 + \eta_{jt})\, \epsilon_{jt}^{D}$. Here $\bar{y}_{Jt}$ and $\bar{p}_{Jt}$ are the (logarithm) average output volume and prices in the industry J at the prefecture level. ϵ_{jt}^{D} is an idiosyncratic demand shock that are i.i.d across firms and over time. This parameterization allows for two important features. First, in the domestic market, the demand shifter varies across firms and over time, as embodied by the demand elasticity and i.i.d demand shock. Second, export may increase the firms' market size, as captured by the term $\beta_{exp}^{D} D_{jt}^{exp}$, where D_{jt}^{exp} is an export dummy.

estimating the production function using the control function approach, even when quantity output data is available.

[8] In fact, the relationship among the three terms in (9) is general: it holds for any forms of production function and demand function, in which case the RTSV may be flexible and vary across firms and over time.

Replacing the markup μ_{jt} in (7) by (9) and rearranging yield the following markup-adjusted revenue function:

$$\begin{aligned}\tilde{\mu}_{jt} \cdot (r_{jt} - \bar{r}_{Jt}) =& \frac{\alpha_L}{(\alpha_L + \alpha_M)} l_{jt} + \frac{\alpha_M}{(\alpha_L + \alpha_M)} m_{jt} + \frac{\alpha_K}{(\alpha_L + \alpha_M)} k_{jt} + \frac{1}{(\alpha_L + \alpha_M)} \omega_{jt} \\ &- \frac{1}{(\alpha_L + \alpha_M)} \bar{y}_{Jt} + \beta^D_{exp} \tilde{\mu}_{jt} D^{exp}_{jt} - \frac{\beta^D_{exp}}{(\alpha_L + \alpha_M)} D^{exp}_{jt} + \frac{1}{(\alpha_L + \alpha_M)} (\xi_{jt} + \epsilon^D_{jt}),\end{aligned} \tag{10}$$

where $\bar{r}_{Jt}$ is the average revenue of industry J in logarithm form. Because the raw markup $\tilde{\mu}_{jt}$ is observed in the data, the left-hand side is just data. This equation forms our main estimation equation, and it can be estimated using standard approaches such as Levinsohn and Petrin (2003) and Ackerberg (2016). We use Levinsohn and Petrin (2003) in our main results, but the results are robust when using alternative methods, as shown in the robustness check. The only minor difference is that because α_l and α_m are not separately identified from variations in l_{jt} and m_{jt}, we will need variations in the industry average output quantity index, $\bar{y}_{jt}$, to identify the returns to scale, $\alpha_l + \alpha_m$ as in Klette and Griliches (1996). We use $m_{jt} = m(k_{jt}, \omega_{jt}, D^{exp}_{jt}, \tilde{\mu}_{jt}, \bar{y}_{Jt}, \bar{p}_{Jt})$ as a control function to recover $\omega_{jt} = m^{-1}(m_{jt}, k_{jt}, D^{exp}_{jt}, \tilde{\mu}_{jt}, \bar{y}_{Jt}, \bar{p}_{Jt})$. Therefore, in the first step, we estimate the following equation using the non-linear least square method:

$$\tilde{\mu}_{jt} \cdot (r_{jt} - \bar{r}_{Jt}) = \frac{\alpha_L}{(\alpha_L + \alpha_M)} l_{jt} + \phi_{jt}(m_{jt}, k_{jt}, \bar{y}_{Jt}, D^{exp}_{jt}, \tilde{\mu}_{jt} D^{exp}_{jt}) + \frac{1}{(\alpha_L + \alpha_M)} \tilde{\xi}_{jt}, \tag{11}$$

where $\tilde{\xi}_{jt}$ refers to $(\xi_{jt} + \epsilon^D_{jt})$, and $\phi_{jt}(m_{jt}, k_{jt}, \bar{y}_{Jt}, D^{exp}_{jt}, \tilde{\mu}_{jt} D^{exp}_{jt})$ is a non-parametric function of m_{jt}, k_{jt}, $\bar{y}_{Jt}$, D^{exp}_{jt}, and $\tilde{\mu}_{jt} D^{exp}_{jt}$ as implied by the model. We use a cubic polynomial as an approximation of $\phi_{jt}(m_{jt}, k_{jt}, \bar{y}_{Jt}, D^{exp}_{jt}, \tilde{\mu}_{jt} D^{exp}_{jt})$. By estimating (11), we can identify the coefficient before l_{jt}, $\alpha_L/(\alpha_L + \alpha_M)$, denoted as $\hat{\alpha}$, the error term $\tilde{\xi}_{jt}/(\alpha_L + \alpha_M)$, and the non-parametric function $\hat{\phi}_{jt}(m_{jt}, k_{jt}, \bar{y}_{Jt}, D^{exp}_{jt}, \tilde{\mu}_{jt} D^{exp}_{jt})$.

We solve the productivity, ω_{jt}, as a function of estimated $\hat{\phi}_{jt}$, observed variables, and parameters as follows:

$$\omega_{jt} = \frac{\alpha_m}{1 - \hat{\alpha}} \hat{\phi}_{jt} - \alpha_M m_{jt} - \alpha_K k_{jt} + \bar{y}_{Jt} - \frac{\alpha_M \beta^D_{exp}}{1 - \hat{\alpha}} \tilde{\mu}_{jt} D^{exp}_{jt} + \beta^D_{exp} D^{exp}_{jt} \tag{12}$$

Using the above equation to replace ω_{jt} and ω_{jt-1} in the productivity evolution process in (5) yields

$$
\begin{aligned}
\frac{\alpha_m}{1-\hat{\alpha}}\hat{\phi}_{jt} &= \alpha_M m_{jt} + \alpha_K k_{jt} - \bar{y}_{Jt} + \frac{\alpha_M \beta^D_{exp}}{1-\hat{\alpha}}\tilde{\mu}_{jt} D^{exp}_{jt} - \beta^D_{exp} D^{exp}_{jt} + \beta^\omega_0 \\
&+ \beta^\omega_1 \big(\frac{\alpha_m}{1-\hat{\alpha}}\hat{\phi}_{jt-1} - \alpha_M m_{jt-1} - \alpha_K k_{jt-1} + \bar{y}_{Jt-1} - \frac{\alpha_M \beta^D_{exp}}{1-\hat{\alpha}}\tilde{\mu}_{jt-1} D^{exp}_{jt-1} + \beta^D_{exp} D^{exp}_{jt-1}\big) \\
&+ \beta^\omega_{exp} D^{exp}_{jt} + \epsilon^\omega_{jt}. \qquad (13)
\end{aligned}
$$

In the second stage, we estimate (13) using the standard General Moment Method(GMM) approach, exploiting the moment conditions based on the irrelevance between the current productivity innovation (ϵ^ω_{jt}) and a set of instrumental variables including the current capital stocks, the past variable inputs, the past export status, and the past raw markup. The details can be found in appendix A.2. After the estimation, we can calculate the markup μ_{jt} by (9).

3.3 Extension: Translog-Form Production Function

One limitation of the Cobb-Douglas production function is that the output elasticities are constant. Hence, the variation of markup only comes from that of raw markup. This subsection shows that our empirical framework can naturally extend to more general production functions. We consider the translog production function as an example:

$$
\begin{aligned}
y_{jt} =& \alpha_L l_{jt} + \alpha_M m_{jt} + \alpha_K k_{jt} + \alpha_{LL} l_{jt} l_{jt} + \alpha_{MM} m_{jt} m_{jt} + \alpha_{KK} k_{jt} k_{jt} \\
&+ \alpha_{LM} l_{jt} m_{jt} + \alpha_{LK} l_{jt} k_{jt} + \alpha_{MK} m_{jt} k_{jt} + (\omega_{jt} + \xi_{jt}) \qquad (14)
\end{aligned}
$$

Through a similar procedure as that used to derive (7), we can derive the translog-form revenue production function. After substituting the markup and φ_{jt} into it, we can derive the estimation equation:

$$
\begin{aligned}
\tilde{\mu}_{jt} \cdot (r_{jt} - \bar{r}_{Jt}) =& \frac{\alpha_L}{(\alpha^*_L + \alpha^*_M)} l_{jt} + \frac{\alpha_M}{(\alpha^*_L + \alpha^*_M)} m_{jt} + \frac{\alpha_K}{(\alpha^*_L + \alpha^*_M)} k_{jt} + \frac{\alpha_{LL}}{(\alpha^*_L + \alpha^*_M)} l_{jt} l_{jt} \\
&+ \frac{\alpha_{MM}}{(\alpha^*_L + \alpha^*_M)} m_{jt} m_{jt} + \frac{\alpha_{KK}}{(\alpha^*_L + \alpha^*_M)} k_{jt} k_{jt} + \frac{\alpha_{LM}}{(\alpha^*_L + \alpha^*_M)} l_{jt} m_{jt} + \frac{\alpha_{LK}}{(\alpha^*_L + \alpha^*_M)} l_{jt} k_{jt} \\
&+ \frac{\alpha_{MK}}{(\alpha^*_L + \alpha^*_M)} m_{jt} k_{jt} + \frac{1}{(\alpha^*_L + \alpha^*_M)} \omega_{jt} - \frac{1}{(\alpha^*_L + \alpha^*_M)} \bar{y}_{Jt} + \tilde{\mu}_{jt} \beta^D_{exp} D^{Exp}_{jt} \\
&- \frac{1}{(\alpha^*_L + \alpha^*_M)} \beta^D_{exp} D^{Exp}_{jt} + \frac{1}{(\alpha^*_L + \alpha^*_M)} \tilde{\xi}_{jt}, \qquad (15)
\end{aligned}
$$

where the output elasticity α_X^*, $X = \{L,\ M,\ K\}$ are given by $\partial y_{jt}/\partial x_{jt}$ and they are functions of l_{jt}, m_{jt}, and k_{jt}.[9] Hence, the output elasticities are variable across firms and over time. Because we cannot separate the error term, $\tilde{\xi}_{jt}$ and the RTSV, $(\alpha_L^* + \alpha_M^*)$, which depends on the input usage l_{jt}, m_{jt}, and k_{jt}, the OLS method is no longer suitable to estimate (15). Hence, we use the GMM method in the first step to estimate (15). Specifically, from (15) we express the measurement error term $\tilde{\xi}_{jt}$ as a function of observed data and the unobserved productivity.[10] By proxying the unobserved productivity using a control function as above, we can rewrite the error term as $\tilde{\xi}_{jt} = \bar{y}_{Jt} - \Phi_{jt}\left(\tilde{\mu}_{jt}\left(r_{jt} - \bar{r}_{Jt}\right), l_{jt}, m_{jt}, k_{jt}, \tilde{\mu}_{jt} D_{jt}^{exp}\right)$, where Φ_{jt} a cubic polynomial of $\tilde{\mu}_{jt}\left(r_{jt} - \bar{r}_{Jt}\right), l_{jt}, m_{jt},\ k_{jt}$, and $\tilde{\mu}_{jt} D_{jt}^{exp}$. Given that the error term $\tilde{\xi}_{jt}$ is i.i.d. and uncorrelated with firm production decisions, we could estimate the model by using the following moment conditions:

$$E\left(\xi_{jt}\left(\beta^{\mathbf{\Phi}}\right)\begin{pmatrix}\left(\tilde{\mu}_{jt-1}\left(r_{jt-1} - \bar{r}_{Jt-1}\right)\right)^a l_{jt}^b m_{jt}^c k_{jt}^d \left(\tilde{\mu}_{jt} D_{jt}^{Exp}\right)^e \\ D_{jt}^{Exp}\end{pmatrix}\right) = 0 \tag{16}$$

$$\text{where} \quad a + b + c + d + e \in \{0, 1, 2\}.$$

$\beta^{\mathbf{\Phi}}$ is the parameters to be estimated in the function $\Phi(\cdot)$. Given the estimates of $\beta^{\mathbf{\Phi}}$ and the error term $\hat{\xi}_{jt}$, the productivity can be expressed as follows:

$$\begin{aligned}\omega_{jt} =& \left(\alpha_L^* + \alpha_M^*\right)\left[\tilde{\mu}_{jt} \cdot \left(r_{jt} - \bar{r}_{Jt}\right)\right] - (\alpha_L l_{jt} + \alpha_M m_{jt} + \alpha_K k_{jt} + \alpha_{LL} l_{jt} l_{jt} + \alpha_{MM} m_{jt} m_{jt} \\ &+ \alpha_{KK} k_{jt} k_{jt} + \alpha_{LM} l_{jt} m_{jt} + \alpha_{LK} l_{jt} k_{jt} + \alpha_{MK} m_{jt} k_{jt} - \bar{y}_{Jt}) - (\alpha_L^* + \alpha_M^*)\beta_{exp}^D \tilde{\mu}_{jt} D_{jt}^{Exp} \\ &+ \left(\alpha_L^* + \alpha_M^*\right)\beta_{exp}^D D_{jt}^{Exp} - \hat{\xi}_{jt}.\end{aligned} \tag{17}$$

Then, similar to the second stage estimation procedures for the Cobb-Douglas production function discussed in the above sub-section, we express the productivity innovation shock ϵ_{jt}^{ω} term as a function of parameters to be estimated $\epsilon_{jt}(\alpha_L, \alpha_M, \alpha_K, \alpha_{LL}, \alpha_{MM}, \alpha_{KK}, \alpha_{LM}, \alpha_{LK}, \alpha_{MK}, \beta_{exp}^D)$.

[9] Specifically, $\alpha_L^* \equiv \alpha_L + 2\alpha_{LL} l_{jt} + \alpha_{LM} m_{jt} + \alpha_{LK} k_{jt}$, $\alpha_M^* \equiv \alpha_M + 2\alpha_{MM} m_{jt} + \alpha_{LM} l_{jt} + \alpha_{MK} k_{jt}$, $\alpha_K^* \equiv \alpha_K + 2\alpha_{KK} k_{jt} + \alpha_{LK} l_{jt} + \alpha_{MK} m_{jt}$.

[10] $\tilde{\xi}_{jt} = (\alpha_L^* + \alpha_M^*) \cdot \tilde{\mu}_{jt}\left(r_{jt} - \bar{r}_{Jt}\right) - (\alpha_L l_{jt} + \alpha_M m_{jt} + \alpha_K k_{jt} + \alpha_{LL} l_{jt} l_{jt} + \alpha_{MM} m_{jt} m_{jt}$ $+ \alpha_{KK} k_{jt} k_{jt} + \alpha_{LM} l_{jt} m_{jt} + \alpha_{LK} l_{jt} k_{jt} + \alpha_{MK} m_{jt} k_{jt} + \omega_{jt}) + \bar{y}_{Jt} - (\alpha_L^* + \alpha_M^*)\beta_{exp}^D \tilde{\mu}_{jt} D_{jt}^{exp} + \beta_{exp}^D D_{jt}^{exp}$

Then, we can estimate these parameters using GMM with moment conditions as follows:

$$E\left(\epsilon_{jt}(\alpha_L,\alpha_M,\alpha_K,\alpha_{LL},\alpha_{MM},\alpha_{KK},\alpha_{LM},\alpha_{LK},\alpha_{MK},\beta^D_{exp})\begin{pmatrix} l_{jt-1}\\ m_{jt-1}\\ k_{jt}\\ l_{jt-1}l_{jt-1}\\ m_{jt-1}m_{jt-1}\\ k_{jt}k_{jt}\\ l_{jt-1}m_{jt-1}\\ l_{jt-1}k_{jt}\\ m_{jt-1}k_{jt}\\ \tilde{\mu}_{jt-1}D^{Exp}_{jt-1}\\ D^{Exp}_{jt-1}\end{pmatrix}\right)=0 \quad (18)$$

We use the estimation results of the translog-form production function as a robustness check and the estimation results of the Cobb-Douglas production function as the main discussion.

4 Estimation Results: Productivity, Markup, and Returns to Scale

We apply the model to the Chinese manufacturing industry from 2000 to 2006. This section first reports the estimation results. Then we calculate the production efficiency gains from export and how much the exporters and consumers can benefit from it.

4.1 Returns to Scale and Export Premium on Productivity and Demand

Column 1 of Table 3 reports the estimation results. There are three main findings. First, export has an economically and statistically significant effect on productivity, as captured by the parameter β^{ω}_{exp}. This echoes the findings in Li et al. (2017) in their researched industries (leather shoes, shirts, and suits). Exporting increases the exporters' productivity in the period by 1.5% on average. This result is at the higher end of the estimates, compared with the literature that uses revenue data without correcting for the unobserved output price bias (e.g. Lileeva and Trefler, 2010; Grieco et al., 2022). The improved productivity reduces the exporters' production costs, giving them competitive advantages in the market and potentially a higher profit.

Second, exporting increases the firm's market size. As shown by the estimate of β^D_{exp}, exporting increases the firm's demand shifter by 19.4% on average. The increased demand shifter gives the exporter a larger market size, potentially leading to higher profitability for the firm.

Third, firm-level production demonstrates substantial increasing returns to scale after correcting for firm heterogeneity in markup. In the table, the (total) returns to scale are captured by $S(RTS)$, which is the sum of output elasticities for capital, labor, and material. In our main estimation, returns to scale are equal to 1.093. This result is consistent with findings in De Loecker et al. (2016), who use the output quantity data to estimate the translog production function with the adjustment for the potential input price bias. They find that 68% of the sample exhibit increasing returns to scale. Moreover, they find that firms' marginal cost decreases in output quantity. As shown later, this result is robust using different specifications of production functions and other estimation methods. The estimate is higher than that reported in the literature using revenue data, in which the unobserved firm heterogeneity in markup biases down the estimate of returns to scale, as discussed in Section 3.

Table 3: Estimation Results

Parameter	(1) Ours	(2) K&G(1996)	(3) Original	(4) Deflated
β^ω_{exp} (Productivity Effect)	0.015	0.002	0.002	0.003
	(0.0005)	(0.0005)	(0.0002)	(0.0003)
β^D_{exp} (Demand Effect)	0.194	0.028	0.019	0.015
	(0.0064)	(0.0015)	(0.0009)	(0.0011)
α_L	0.081	0.057	0.041	0.042
	(0.0008)	(0.0007)	(0.0004)	(0.0004)
α_M	0.980	0.926	0.910	0.917
	(0.0051)	(0.0053)	(0.0042)	(0.0066)
α_K	0.032	0.024	0.015	0.018
	(0.0006)	(0.0032)	(0.0015)	(0.0020)
S (RTS)	1.093	1.007	0.965	0.976
	(0.0052)	(0.0017)	(0.0028)	(0.0045)
S^V (RTSV)	1.061	0.983	0.951	0.959
	(0.0056)	(0.0048)	(0.0041)	(0.0065)
Observations	1,234,292	1,234,292	1,234,292	1,234,292

Note: Standard errors (clustered at the firm level) in parentheses.

Increasing returns to scale provides a new source of gains from exports. Given that firms have a larger market and produce more after export, increasing returns to scale reduces the firm's marginal costs and improves its efficiency. The result echoes early studies Scherer

(1980) and Panzar (1989), who illustrated the existence of economies of scale in both case studies and theory. This result also provides empirical support to the new growth theory (Romer, 1986) and new trade theory (Krugman, 1980, 1995), whose results are mainly based on the assumption of increasing returns to scale at the aggregate economy. Our results show that increasing returns to scale exist even at the firm level.

As a comparison, columns (3)-(4) report the estimation results using the widely-used approaches in the literature based on revenue data to estimate the revenue production function (7) without correcting for firm heterogeneity in markup. Column (3) uses the original revenue as a proxy for output, and column (4) uses the deflated revenue instead. As shown in the table, without correcting for firm heterogeneity in prices (and markup), the traditional methods find a negligible effect of export on revenue productivity (0.2-0.3%), which is one order lower in magnitude compared with our results. This is in contrast to Garcia-Marin and Voigtländer (2019), which shows that TFPR would increase as long as markup increases (to be shown in Section 4.2) under constant returns to scale. Our results are intuitive. As discussed in Section 2.3, with increasing returns to scale, TFPR may not increase even if export increases markup because increasing returns scale with enlarged market size drives down TFPR.[11]

The revenue-based approaches in the literature also underestimate the effect of export on market size, as shown in the estimates of β^D_{exp} in columns (3) and (4). Notice that the markup-adjusted revenue function in (10) can be rewritten as: $r_{jt} - \bar{r}_{Jt} = \frac{1}{\mu_{jt}}(\alpha_L l_{jt} + \alpha_M m_{jt} + \alpha_K k_{jt} + \omega_{jt} - \bar{y}_{Jt} + \tilde{\xi}_{jt}) + \left(1 - \frac{1}{\mu_{jt}}\right) \beta^D_{exp} D^{exp}_{jt}$. The term $(1 - \frac{1}{\mu_{jt}})$ before the export variable D^{exp}_{jt} biases down the true estimates of export effects on productivity. Moreover, by ignoring firm heterogeneity in markup, the literature also underestimates the returns to scale as shown in columns (3) and (4). This is as expected following our discussion in Equation (7).

Interestingly, when controlling for common markup at the industry level when estimating the revenue production function (7) using the method proposed by Klette and Griliches (1996), the estimated returns to scale become larger (1.007) than those estimated using traditional approaches as reported in Column (2). This is consistent with the insights and results in Klette and Griliches (1996). However, it is still lower than our estimate because it ignores firm heterogeneity in markup. More importantly, the Klette-Griliches approach still substantially understates the productivity and demand gains from exports. This is because, by assuming constant demand elasticity, the approach fails to capture the changes in markup before and after the firm starts exporting, which has an impact on the estimation of the productivity

[11] We calculated the implied changes in TFPR based on (3) and the estimates of changes in markup, productivity, and sales. The implied changes in TFPR after export is -0.002, which is close to our estimate in Table 3. The decomposition results are summarized in the Appendix Table B2.

and demand premium of exports.

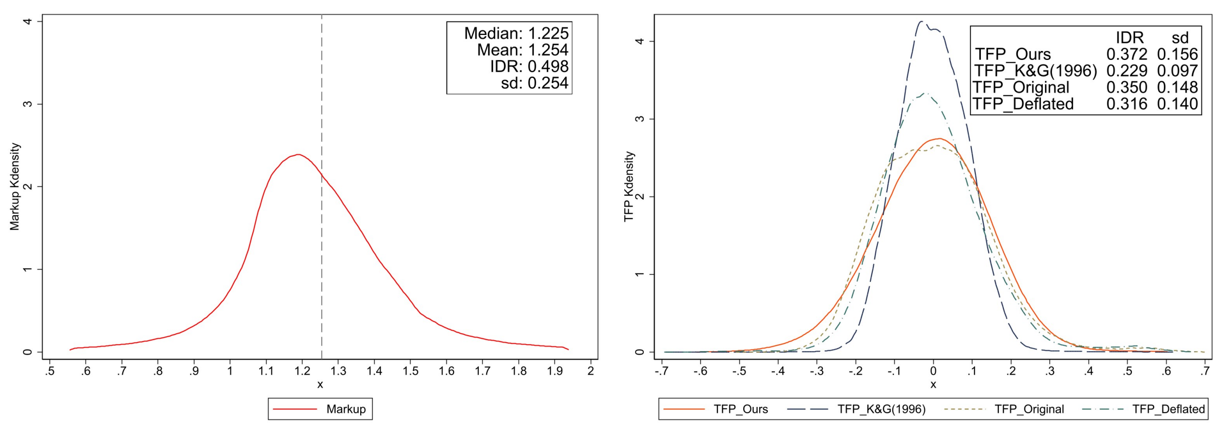
Firm Heterogeneity in Markup and Productivity

After estimating the output elasticities, we can calculate the firm-level markup using (9) and productivity using (12). We plot their kernel distributions in Figure 2.

The markup is greater than 1 on average, with the mean at 1.254 and median at 1.225. This explains why estimating the revenue production function without correcting for markup may understate the returns to scale. There is also large heterogeneity in markup across firms/years, with an interdecile range of 0.498. Comparing the 90 percentile with the 10th percentile, the markup difference is almost 50 percentage points.

The heterogeneity in markup is even higher than productivity. The interquartile range of markup (0.498) is significantly larger than that of productivity (0.372). This result highlights the importance of taking into account firm heterogeneity in markup in the analysis of firm heterogeneity and performance. This result also stresses why controlling common markup as in Klette and Griliches (1996) may still substantially bias the estimates of productivity and the export premium. Meanwhile, the different dispersions of productivity estimated by different methods also suggest different heterogeneity levels in firms. Our method suggests that firms are more heterogeneous in terms of productivity than the traditional estimates.

Figure 2: Markup and TFP Dispersion (LP)



4.2 Export Premium on Markup

The larger market size and improved efficiency may allow the exporters to charge a higher markup. In this section, we estimate the impact of export on markup based on the following equation:

$$\mu_{jt} = \beta_0^{\mu} + \beta_{exp}^{\mu} D_{jt}^{exp} + \beta^X \mathbf{X}_{jt} + \gamma_j + \gamma_t + \xi_{jt}^{\mu}, \tag{19}$$

where D_{jt}^{exp} is the export dummy as in equation (5) and our parameter of interest, β_{exp}^{μ}, captures the average exporting effect on markup. We control for a set of firm characteristics $\mathbf{X}_{jt}$, including firm size (measured by labor employed) and capital intensity. We also control for firm FE and time FE to capture time-invariant firm characteristics and common time shocks. ξ_{jt}^{μ} is an i.i.d shock to the firm's markup. The estimation results are reported in Table 4.

Table 4: Exporting Effect on Markup (LP)

	(1) Markup	(2) Markup
D_{jt}^{exp}	0.008***	0.010***
	(0.0015)	(0.0015)
Firm Size (L)		YES
Capital Intensity		YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	1,234,292	1,234,292
Adjusted R^2	0.326	0.326

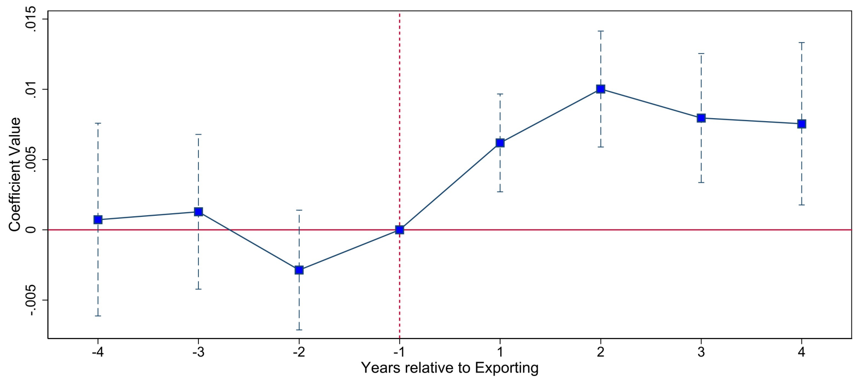
Note: Standard errors (clustered at the firm level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$

Column (1) reports the estimation results after controlling the two fixed effects, and column (2) further controls for the firm-specific variables, including firm size and capital intensity. We find that export increases a firm's markup by 0.8-1.0 percentage points. This result suggests that exporters are able to charge a higher markup after starting export, presumably because of the increased market size and improved production efficiency.

We also estimate the pre-trend and dynamic effects of export on markup, based on an extended model of (19). The results are reported in Figure 3.[12] On the horizontal axis of the figure, "-1" represents the year right before export, and "1" represents the first year of exporting. The other years are similarly defined. The result shows no obvious pre-trend. Before they

[12]We only kept four years before and four years after exporting in the dynamic graph because our database only covers the years from 2000 to 2006. Because we do not know when a firm started exporting if it was already exporting in 2000, the maximum number of years after starting to export that we can identify is 6 years (for those firms that started exporting in 2001). However, the number of observations that were 5 or 6 years before or after the firm started exporting is small, making the estimation of the dynamic effects at 5 and 6 years inaccurate. Therefore, we do not include them in the analysis of dynamic effects.

Figure 3: Dynamic Exporting Effect on Markup (LP)



Note: The range represents 95% confidence interval of the parameter estimates.

export, exporters' markups are not significantly different from those of non-exporters. It also confirms the large dynamic effect of export on markup. Compared with non-exporters, exporters' markup increases significantly from the first year of export, and the effects remain multiple years after exporting.

5 Validation with Output Quantity

When output quantity data are not available, the unobserved markup causes a bias in productivity and returns to scale when estimated using the revenue production function. When output quantity data are available, however, the quantity-based production function should generate consistent estimation results. We validate our new method by comparing our estimation results with that based on quantity-based output production function using a smaller sample from 2000 to 2006, which has output quantity at the firm level. The output quantity data come from the output quantity survey as discussed in Section 2.1 above. After merging it with our main data, there are about 444,475 observations with output quantity.

One problem is that we are unable to construct the firm-level price index accurately for multi-product firms because the data do not contain information to construct firm-product-level output prices. To construct accurate price information at the firm level, we kept only single-product firms. The output price of each single-product company is defined as the ratio of output revenue and reported output quantity. Finally, the sample contains 118,671 firms (321,280 observations) with output quantity data, of which 19,838 firms (16.72%) are exporters.

One caveat of estimating the physical output production function, as De Loecker et al. (2016)

and Li and Zhang (2022) point out, is that output quality and input quality (and output prices and input prices) may be positively correlated. This correlation may bias our estimates. This problem is especially serious when the physical output quantity is used as output, but the (deflated) input expenditure is used as a proxy for input quantity, as in most applications, including ours. In this case, the input quality and price are partially controlled by the input expenditure, but output quality enters the error term, which biases downward the estimates of output elasticities and thus returns to scale. In fact, even if we don't consider quality differences, input prices and output prices are still positively correlated due to the optimal pricing strategy of firms facing imperfect competition, leading to a similar bias in the production function estimation.

To address this potential problem, we follow the insight of De Loecker et al. (2016) to use the output price to control for input price differences. Specifically, we assume that the firm-level output prices are determined by variations in firm-level input prices and the market-level input and output price indices, so we can invert the output prices to construct an index of firm-level input prices using variations of firm-level output prices and market-level input and output price indices. Then, the "physical" input quantity is calculated by dividing the input expenditure by the firm-level input price index. We use it in the production function estimation. The data construction details, together with the detailed estimation process, are discussed in Appendix A.7.

Table 5 reports the estimation results based on the observed output quantity. As a comparison, we also report the estimation results using our method based on revenue data in this small sample. There are three major findings. First, the physical production function shows a significant increase in scale returns, with RTS = 1.079. The result is consistent with De Loecker et al. (2016), who find that 68% of their sample observations show increasing returns to scale by estimating the physical quantity production function after adjusting the potential biases caused by the input prices/quality. Second, exports significantly increase firm productivity by 0.5%. The effect is smaller than that of our main results, but the difference may be driven by the different samples. When estimating the model using our method (pretending no output quantity data are available) in the same smaller sample, we find similar results (0.6% for the productivity effect and 1.035 for the RTS) after correcting for the potential positive relationship between input and output prices (as proxies of input and output quality). The results further show that our results and the new estimation methods are robust. Third, consistent with our main results, using revenue deflated by the industry-year-level output deflator to proxy the output quantity underestimates the productivity effect of exports and the RTS. When using the constant markup as in Klette

Table 5: Results Using Quantity Data, Our Approach, and the Revenue-Based Methods

	(1) Quantity Output	(2)	(3) Revenue Output	(4)
Parameter		Ours	K&G(1996)	Deflated
β^{ω}_{exp} (Productivity Effect)	0.005	0.006	0.002	0.001
	(0.0006)	(0.0004)	(0.0008)	(0.0003)
β^{D}_{exp} (Demand Effect)		0.183	0.029	0.025
		(0.0097)	(0.0046)	(0.0017)
α_L	0.019	0.052	0.038	0.023
	(0.0006)	(0.0008)	(0.0010)	(0.0006)
α_M	1.003	0.963	0.984	0.974
	(0.0005)	(0.0011)	(0.0012)	(0.0004)
α_K	0.057	0.020	0.015	0.001
	(0.0188)	(0.0010)	(0.0043)	(0.0008)
S (RTS)	1.079	1.035	1.038	0.998
	(0.0190)	(0.0013)	(0.0039)	(0.0009)
S^V (RTSV)	1.021	1.015	1.023	0.997
	(0.0008)	(0.0012)	(0.0009)	(0.0007)
Observations	321,280	321,280	321,280	321,280

Note 1: Here we only keep the single-product firms (with quantity information) in the sample.

Note 2: Standard errors (clustered at the firm level) in parentheses.

and Griliches (1996), we can detect the IRS as expected, but the productivity effect is still negligible, like in our main results.

We also checked the robustness of the results by using data from both single- and multi-product firms. One problem with multi-product firms is that we are not able to precisely construct the firm-level output price index without knowledge of the values/prices of each output variety. Therefore, we used two approximate ways to construct the firm-level output. In the first, the firm-level output quantity is treated as the sum of the output quality of all output varieties, although a firm may produce multiple and different outputs. In the second, the output quantity is treated as the quantity of output variety with the maximum output quantity for this firm. Then, the firm-level output prices are defined as the ratio of the output value and the approximate output quantity. Neither of these two alternative approaches is accurate because the first approach may suffer from the problem of adding apples to pears, and the second approach may underestimate the output quality. However, they provide a validation that our results are not driven by the selection of single-product firms. The results are consistent and are reported in Appendix A.7.

6 Gains from Exporting

Improved productivity and increasing returns to scale imply that export can improve production efficiency. This efficiency, together with increased markup, guarantees that exporters can gain a greater profit. Moreover, can consumers gain (or lose) from exports by paying lower (or higher) output prices? This section answers these questions. First, we calculate the changes in production efficiency contributed by improved productivity and increasing returns to scale. Then, we calculate the impact on firms and consumers based on changes in their profitability and paid prices, respectively.

6.1 Export and Production Efficiency

To quantify the export premium in production efficiency and evaluate the contribution of improved productivity and increasing returns to scale, first note that the cost minimization problem as discussed in Section 3 implies the following marginal costs function[13],

$$\begin{aligned} mc_{jt} =& \frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M} y_{jt} - \frac{1}{\alpha_L+\alpha_M}\omega_{jt} + \frac{\alpha_M}{\alpha_L+\alpha_M} p_{jt}^M + \frac{\alpha_L}{\alpha_L+\alpha_M} p_{jt}^L \\ &- \frac{\alpha_M}{\alpha_L+\alpha_M}\ln(\alpha_M) - \frac{\alpha_L}{\alpha_L+\alpha_M}\ln(\alpha_L) - \frac{\alpha_K}{\alpha_L+\alpha_M} k_{jt} - \frac{1}{\alpha_L+\alpha_M}\xi_{jt}, \end{aligned} \tag{20}$$

where mc_{jt} is the log form of marginal cost, and other variables are similarly defined as before. We ignore the indirect effect of exports on input prices and capital stock.[14] Hence, the impact of export on marginal costs can be written in the following form,

$$\Delta mc^{exp} = \frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M}\Delta y^{exp} - \frac{1}{\alpha_L+\alpha_M}\beta_{exp}^{\omega}, \tag{21}$$

where Δy^{exp} denotes the increase in output caused by exporting. The first term on the right, $\frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M}\Delta y_j^{exp}$, captures the exporter's efficiency gains from increasing returns to scale. Given increasing returns to scale ($\alpha_L+\alpha_M>1$), producing more after export reduces the firm's marginal costs at a rate of $\frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M}$. The second term, $\frac{1}{\alpha_L+\alpha_M}\beta_{exp}^{\omega}$, captures the efficiency gains through improved productivity. Because $\beta_{exp}^{\omega}>0$, import reduces marginal costs.[15]

[13] Detailed derivation process is reported in Appendix A.5.

[14] By focusing on the direct gains from export via improved productivity and increasing returns to scale and ignoring the indirect gains from adjusting capital investment and input prices, our result can be thought of as the lower bound of the total gains from export.

[15] Of course, because improved productivity also induces more output, further allowing the exporter to gain more from increasing returns to scale (captured by the first term). Hence, more precisely, the second term captures the *direct* effect of export on production costs via improved productivity.

In (21), changes in output quantities Δy^{exp} are not observed in the data. It can be derived using (21), combined with the definition of (changes in) markup and revenue, as follows

$$\Delta \ln \mu^{exp} = \Delta p^{exp} - \Delta mc^{exp}, \tag{22}$$

$$\Delta r^{exp} = \Delta p^{exp} + \Delta y^{exp}. \tag{23}$$

where Δx represents changes in variable x as usual. The average changes in (logarithm) markup after export, $\Delta \ln \mu^{exp}$, has been estimated in Section 4.2 at 1%.[16] The average changes in revenue after export, Δr^{exp}, can also be estimated similarly, which is 20.974%. Combing (21)-(23) yields the solution for $\Delta p^{exp}, \Delta y^{exp}$ and Δmc^{exp} as the functions of observed estimates ($\Delta \ln \mu^{exp}$, Δr^{exp}, β^{ω}_{exp}) as follows:

$$\begin{aligned} \Delta y^{exp} &= (\alpha_L + \alpha_M)\left[\Delta r^{exp} - \Delta \ln(\mu)^{exp}\right] + \beta^{\omega}_{exp} \\ \Delta p^{exp} &= \left[1 - (\alpha_L + \alpha_M)\right]\Delta r^{exp} + (\alpha_L + \alpha_M)\Delta \ln(\mu)^{exp} - \beta^{\omega}_{exp} \\ \Delta mc^{exp} &= \left[1 - (\alpha_L + \alpha_M)\right]\left[\Delta r^{exp} - \Delta \ln(\mu)^{exp}\right] - \beta^{\omega}_{exp} \end{aligned} \tag{24}$$

Taking (24) into the data, we can calculate the average exporting changes on output quantities Δy^{exp}, output prices Δp^{exp}, and marginal costs Δmc^{exp}, which equal 22.657%, -1.683%, and -2.683%, respectively.

With changes in output quantity in hand, we can calculate the contribution of increasing returns to scale to production efficiency, as well as the impact of export on production efficiency based on (21). The results are summarized in Table 6. It is shown that export reduces the exporters' marginal cost of production by -2.683% on average, largely increasing the production efficiency. The improved productivity accounts for -1.381% of the cost reduction. The remaining part is contributed by increasing returns to scale by -1.302%. This result highlights the importance of returns to scale: it is as important as productivity gains and contributes to about half of the production efficiency gains from exports.

Table 6: Gains from Export

$\Delta mc_{jt} \downarrow$ 2.68%			Consumer Welfare	Exporter Profits
IRS	$\Delta \omega^{exp}$	$\Delta \mu^{exp}$	Price	Profits
$\downarrow$ 1.302%	$\downarrow$ 1.381%	$\uparrow$ 1%	$\downarrow$ 1.683%	$\uparrow$ 25.322%

[16] Export's effect on log markup almost equals the export's effect on markup numerically. See Table B11 for details.

6.2 Benefits to Firms and Consumers

Firm profitability. The increased production efficiency and markup naturally imply a gain in profitability for exporters. To quantify this effect, we calculate the changes in profits as implied by the model estimates. By the definition of profit $\Pi_{jt} = R_{jt} - (P^L_{jt} L_{jt} + P^M_{jt} M_{jt})$ and the definition of raw markup $\tilde{\mu}_{jt} \equiv R_{jt}/(P^L_{jt} L_{jt} + P^M_{jt} M_{jt})$, we can express the profit as $\Pi_{jt} = R_{jt}(\tilde{\mu}_{jt} - 1)/\tilde{\mu}_{jt}$. As shown in Appendix A.6, the first-order approximation of the impact of export on firms' profits can be written in the following form:

$$\Delta \pi^{exp} \simeq \Delta y^{exp} + \Delta p^{exp} + \frac{1}{\tilde{\mu}^{exp=0}\left(\tilde{\mu}^{exp=0} - 1\right)} \Delta \tilde{\mu}^{exp} \tag{25}$$

The changes in output quantity, prices, and markup after exporting have been calculated above. We further use the raw markup of all non-exporting observations to represent $\tilde{\mu}^{exp=0}$. Combining all information, we find that export raises firms' profit by 25.332% on average. A large part of this is due to the increased market size, but the improved efficiency and markup further improve the exporter's profits.

Consumer welfare. Given the increased markup after export, can consumers benefit from the improved efficiency? Since the markup in the log form can be expressed as $ln(\mu_{jt}) = p_{jt} - mc_{jt}$, we can derive the changes in consumer prices from information on changes in markup and marginal costs. The calculation above implies that the efficiency effect (-2.683%) dominates the increases in markup (1%). As a result, export reduces the average prices paid by consumers for products sold by firms that start exporting, by 1.683%. This means that although firms increase markup after exporting, they are able to pass through some of the export's efficiency premium to consumers by lowering the output prices.

Note that the output price of an exporter here is the firm-level average price, including both export and domestic sales. As a result, the reduction in exporters' output price may result from a lower export price or a reduced price for domestic consumers. In other words, it is possible that both foreign buyers and domestic consumers can benefit from improved efficiency, although more information is needed to answer this question precisely.

The back-of-the-envelope calculation in this subsection shows that both firms and consumers benefit from exports. After exporting, firms' profits increase by about a quarter due to increased productivity, markup, increasing returns to scale, and larger market size. Yet, due to improved production efficiency, firms are able to charge a lower price to achieve higher markup and the lower prices benefit consumer welfare.

7 Robustness Checks

This section shows that our results are robust to an alternative estimation approach and a more flexible production function that allows for variable output elasticities.

7.1 ACF Estimation

In our main results, we used Levinsohn and Petrin (2003). In this section, we use the alternative method provided by Ackerberg et al. (2015) to estimate all parameters jointly. The structural estimation results of all the specifications are displayed in Table 7.

Table 7: Estimation Results (ACF)

Parameter	(1) Ours	(2) K&G(1996)	(3) Original	(4) Deflated
β_{exp}^{ω} (Productivity Effect)	0.014	-0.004	0.002	0.003
	(0.0005)	(0.0020)	(0.0003)	(0.0009)
β_{exp}^{D} (Demand Effect)	0.212	0.035	0.022	0.020
	(0.0065)	(0.0060)	(0.0011)	(0.0027)
α_L	0.045	0.126	0.056	0.062
	(0.0039)	(0.0194)	(0.0026)	(0.0041)
α_M	1.014	0.918	0.884	0.873
	(0.0029)	(0.0114)	(0.0092)	(0.0314)
α_K	0.031	0.012	0.021	0.028
	(0.0005)	(0.0027)	(0.0025)	(0.0091)
S (RTS)	1.090	1.057	0.961	0.963
	(0.0054)	(0.0098)	(0.0050)	(0.0199)
S^V (RTSV)	1.059	1.045	0.940	0.935
	(0.0057)	(0.0111)	(0.0074)	(0.0290)
Observations	1,234,292	1,234,292	1,234,292	1,234,292

Note: Standard errors (clustered at the firm level) in parentheses.

The first two rows of Table 7 show that the exporting effects on productivity and demand shifter estimated using ACF are very similar to our main results as reported in Table 3. Besides, the exporting effects on productivity detected by the traditional methods are very small, as expected. The exporting effect on productivity using our method is still about five times as large as the three traditional methods, indicating that our discussion in the previous sections is robust. In terms of the output elasticity, notice that our method and Klette and Griliches (1996)'s method predicts increasing returns to scale, while the original revenue method and deflated revenue method show decreasing returns to scale. This, again, is consistent with the main results using Levinsohn and Petrin (2003).

After estimating the production function, we can calculate the new markup. We have also conducted a similar analysis of export's effect on markup as in Section 4.2. As shown in Table B3, the impact of export on markup based on ACF is almost the same as that in LP's case.

As in the main result, the increased productivity after export and increasing returns to scale (together with a larger market) imply an export premium on production efficiency. Given the new estimation results using ACF, we find that export reduced exporter's marginal costs by 2.561%, of which productivity channel and increasing returns to scale each contribute by about half. This is very close to our main result (see Table B4 for details). The improved production efficiency and markup[17] increases exporters' profits by 25.340%. Consumers also benefit from lower prices as well by 1.561%. This is because the improved efficiency allows the exporter to charge a higher markup at a lower price. All these results are very close to our main results, as summarized in Table B6.

7.2 Translog Production Function Estimation

In the Cobb-Douglas case in our main result, one limitation is that all variation in markup is driven by expenditure shares. This is because output elasticities are constant in the Cobb-Douglas case. We relax this assumption to consider the more flexible translog production function, in which case both expenditure shares and the flexible output elasticities contribute to the changes in markup. In this more flexible case, our key insight of controlling for firm heterogeneity in markup using the raw markup and estimate markup and production function jointly still holds. However, the estimation would be slightly adjusted to estimate all parameters in one step using GMM, as discussed in Section 3.3.

The estimation results are reported in Table B7. Our results are robust, except that we find an even larger export premium on productivity. In particular, the output elasticity is 0.074 for labor, 0.996 for material, and 0.039 for capital, indicating increasing returns to scale for total RTS at 1.109 and the variable RTSV at 1.070. The export effect is 0.030 for productivity and 0.220 for demand shifter. We also analyze the export effect on markup and find that export raises firms' markup by 1.3-1.7% on average (See Table B9), with a similar dynamic pattern as that in the Cobb-Douglas production function case (See Figure B2). From translog's case, we can see that when the output elasticities are variable, export shows even greater effects on productivity and markup, while the level of IRS and demand shifter

[17]see Table B5 for export's effect on markup and Table B12 for export's effect on log markup.

effect remain similar. With the flexible translog production function, the increasing returns to scale demonstrate substantial heterogeneity across firms, with a standard deviation of 0.043 and an inter-decile range of 0.110. All other results are qualitatively and quantitatively similar to our main results. This exercise suggests that our methodology can be applied to more general production functions and the main results are robust.

8 Conclusion

The productivity effect of exports has been the foundation for many trade policies. However, empirical studies typically detected limited/mixed productivity effects of exports using revenue data, even when exports increase markup. We show that this is because export may reduce output prices, due to efficiency gains from increased productivity and, more importantly, increasing returns to scale. Because output prices are typically unavailable in most micro datasets, we develop a new method to consistently estimate firm-level markup, productivity, and returns to scale jointly, using revenue and widely available variable inputs expenditure. The new method only requires cost minimization, but not profit maximization, as in De Loecker and Warzynski (2012).

We find three main results in the application to Chinese manufacturing industries. First, exports have large efficiency gains, by reducing marginal production costs by 2.683% on average. About half of the efficiency gains is contributed by increasing returns to scale; the other half is due to improved productivity after export. Second, improved production efficiency and a larger market allow exporters to charge a higher markup by 1% at a lower price. All together, firm profit increases by 25.332% after export in the Chinese manufacturing industry; consumers also benefit from lower prices by 1.683%. The estimation results are validated using a sub-sample with the output quantity information by estimating the physical output production function which is not affected by the unobserved markup heterogeneity. The results are also robust to alternative estimation methods, more flexible specification of the production functions, industry-separated estimation, and export instruments.

Our results suggest that the export productivity premium does exist when correctly estimated by dealing with the unobserved prices. After correcting for unobserved markup, production shows significant increasing returns to scale, which serves as an important source of efficiency gains from export, even at the firm level. We show that unlike Garcia-Marin and Voigtländer (2019), the TFPR effect of export may be limited even when markup increases when production shows increasing returns to scale. The results have implications in many other fields related to firm growth, industrial dynamics, macro growth, and international trade.

References

Ackerberg, D. A. (2016). Timing assumptions and efficiency: Empirical evidence in a production function context. *University of Michigan Mimeo*.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Akcigit, U. and Ates, S. T. (2023). What happened to us business dynamism? *Journal of Political Economy*, 131(8):2059–2124.

Amiti, M., Itskhoki, O., and Konings, J. (2019). International shocks, variable markups, and domestic prices. *The Review of Economic Studies*, 86(6):2356–2402.

Arkolakis, C. and Morlacco, M. (2017). Variable demand elasticity, markups, and pass-through. *Manuscript, Yale University*.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5):1998–2031.

Bai, X., Krishna, K., and Ma, H. (2017). How you export matters: Export mode, learning and productivity in China. *Journal of International Economics*, 104:122–137.

Basu, S. and Fernald, J. G. (1997). Returns to scale in us production: Estimates and implications. *Journal of Political Economy*, 105(2):249–283.

Bernard, A. B. and Jensen, J. B. (1999). Exceptional exporter performance: cause, effect, or both? *Journal of International Economics*, 47(1):1–25.

Blundell, R. and Bond, S. (2000). GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3):321–340.

Bond, S., Hashemi, A., Kaplan, G., and Zoch, P. (2021). Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics*, 121:1–14.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review*, 107(9):2784–2820.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of NAFTA. *The Review of Economic Studies*, 82(1):1–44.

Clerides, S. K., Lach, S., and Tybout, J. R. (1998). Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. *The Quarterly Journal of Economics*, 113(3):903–947.

Dai, M., Maitra, M., and Yu, M. (2016). Unexceptional exporter performance in China? The role of processing trade. *Journal of Development Economics*, 121:177–189.

De Loecker, J. (2011). Recovering markups from production data. *International Journal of Industrial Organization*, 29(3):350–355.

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

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

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

Doraszelski, U. and Jaumandreu, J. (2019). Using cost minimization to estimate markups.

Doraszelski, U. and Jaumandreu, J. (2021). Reexamining the De loecker & Warzynski (2012) method for estimating markups.

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, markups, and the gains from international trade. *American Economic Review*, 105(10):3183–3221.

Edmond, C., Midrigan, V., and Xu, D. Y. (2023). How costly are markups? *Journal of Political Economy*, 131(7):1619–1675.

Feenstra, R. C., Li, Z., and Yu, M. (2014). Exports and credit constraints under incomplete information: Theory and evidence from China. *Review of Economics and Statistics*, 96(4):729–744.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425.

Garcia-Marin, A. and Voigtländer, N. (2019). Exporting and plant-level efficiency gains: It's in the measure. *Journal of Political Economy*, 127(4):1777–1825.

Ge, Y., Lai, H., and Zhu, S. C. (2015). Multinational price premium. *Journal of Development Economics*, 115:181–199.

Greenaway, D., Gullstrand, J., and Kneller, R. (2005). Exporting may not always boost firm productivity. *Review of World Economics*, 141:561–582.

Greenaway, D. and Kneller, R. (2007). Industry differences in the effect of export market entry: Learning by exporting? *Review of World Economics*, 143:416–432.

Grieco, P. L., Li, S., and Zhang, H. (2016). Production function estimation with unobserved input price dispersion. *International Economic Review*, 57(2):665–690.

Grieco, P. L., Li, S., and Zhang, H. (2022). Input prices, productivity, and trade dynamics: long-run effects of liberalization on Chinese paint manufacturers. *The RAND Journal of Economics*, 53(3):516–560.

Hall, R. E. (1988). The relation between price and marginal cost in U.S. industry. *Journal of Political Economy*, 96(5):921–947.

Hall, R. E., Blanchard, O. J., and Hubbard, R. G. (1986). Market structure and macroeconomic fluctuations. *Brookings Papers on Economic Activity*, 1986(2):285–338.

Holmes, T. J. and Schmitz Jr, J. A. (2001). A gain from trade: From unproductive to productive entrepreneurship. *Journal of Monetary Economics*, 47(2):417–446.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

Keller, W. (2010). International trade, foreign direct investment, and technology spillovers. In *Handbook of the Economics of Innovation*, volume 2, pages 793–829. Elsevier.

Klette, T. J. and Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11(4):343–361.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959.

Krugman, P. (1995). Increasing returns, imperfect competition and the positive theory of international trade. *Handbook of International Economics*, 3:1243–1277.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Li, S. and Zhang, H. (2022). Does external monitoring from the government improve the performance of state-owned enterprises? *The Economic Journal*, 132(642):675–708.

Li, Y., Smeets, V., and Warzynski, F. (2017). Physical productivity and exceptional exporter performance: Evidence from a chinese production survey.

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity... for some plants. *The Quarterly Journal of Economics*, 125(3):1051–1099.

Marschak, J. and Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica*, pages 143–205.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Mo, J., Qiu, L. D., Zhang, H., and Dong, X. (2021). What you import matters for productivity growth: Experience from Chinese manufacturing firms. *Journal of Development Economics*, 152:102677.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Panzar, J. C. (1989). Technological determinants of firm and industry structure. *Handbook of Industrial Organization*, 1:3–59.

Raval, D. (2023). Testing the production approach to markup estimation. *The Review of Economic Studies*, 90(5):2592–2611.

Rhee, Y. W., Ross-Larson, B. C., and Pursell, G. (1984). Korea's competitive edge: managing the entry into world markets.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):1002–1037.

Scherer, F. (1980). *Industrial market structure and economic performance*. Chicago: Rand McNally., 2nd edition.

Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6):1181–1222.

Wagner, J. (2007). Exports and productivity: A survey of the evidence from firm-level data. *World Economy*, 30(1):60–82.

Yu, M. (2015). Processing trade, tariff reductions and firm productivity: Evidence from Chinese firms. *The Economic Journal*, 125(585):943–988.

A Appendix - Derivations of the Model

A.1 Markup Derivation for Cobb-Douglas Production Function

Firm j's cost minimization problem can be described as:

$$\begin{aligned}&\min_{\{L_{jt},M_{jt}\}} P^L_{jt}L_{jt} + P^M_{jt}M_{jt}\\ &s.t. \quad e^{(\omega_{jt}+\xi_{jt})}L^{\alpha_L}_{jt}M^{\alpha_M}_{jt}K^{\alpha_K}_{jt} \geq Y_{jt}.\end{aligned} \tag{A1}$$

The corresponding Lagrangian function is:

$$\mathcal{L} = P^L_{jt}L_{jt} + P^M_{jt}M_{jt} + \lambda_{jt}\left[Y_{jt} - e^{(\omega_{jt}+\xi_{jt})}L^{\alpha_L}_{jt}M^{\alpha_M}_{jt}K^{\alpha_K}_{jt}\right]$$

The First Order Condition w.r.t. L_{jt} and M_{jt} are shown as follows:

$$P^L_{jt} = \lambda_{jt}\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}}, \tag{A2}$$

$$P^M_{jt} = \lambda_{jt}\frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}}, \tag{A3}$$

where $Y_{jt}(\cdot)$ represents the production function. Multiplying both sides by L_{jt} and M_{jt} on the above two equations, respectively, and adding them together yield:

$$P^L_{jt}L_{jt} + P^M_{jt}M_{jt} = \lambda_{jt}\Big[\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}}L_{jt} + \frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}}M_{jt}\Big].$$

Dividing revenue R_{jt} ($R_{jt} = P_{jt}Y_{jt}$) by both sides yields:

$$\frac{R_{jt}}{P^L_{jt}L_{jt} + P^M_{jt}M_{jt}} = \frac{P_{jt}}{(\alpha_L + \alpha_M)\,\lambda_{jt}}.$$

λ_{jt} is the shadow price of producing one unit of output, representing the marginal cost $\lambda_{jt} \equiv MC_{jt}$. By definition, the markup $\mu_{jt} = P_{jt}/MC_{jt}$. We can get the expression for firm markup μ_{jt} as follows:

$$\begin{aligned}\mu_{jt} &= (\alpha_L + \alpha_M)\cdot\frac{R_{jt}}{P^L_{jt}L_{jt} + P^M_{jt}M_{jt}}\\ &= (\alpha_L + \alpha_M)\cdot\tilde{\mu}_{jt}\end{aligned} \tag{A4}$$

A.2 The Second Stage Estimation

Recall the AR(1) productivity evolution process:

$$\omega_{jt} = \beta_0^\omega + \beta_1^\omega \omega_{jt-1} + \beta_{exp}^\omega D_{jt}^{exp} + \epsilon_{jt}^\omega. \tag{A5}$$

where ϵ_{jt}^ω is uncorrelated with ω_{jt-1} and D_{jt}^{exp}. Thus by the formula of the OLS estimator, the parameters can be expressed as:

$$\begin{bmatrix}\beta_0^\omega & \beta_1^\omega & \beta_{exp}^\omega\end{bmatrix}' = (\mathrm{X}'_{\mathrm{jt-1}}\mathrm{X}_{\mathrm{jt-1}})^{-1}\mathrm{X}'_{\mathrm{jt-1}}\omega_{\mathrm{jt}}, \tag{A6}$$

where $\mathrm{X}_{\mathrm{jt-1}}$ is the vector given by $\mathrm{X}_{\mathrm{jt-1}} = [1\ \omega_{\mathrm{jt-1}}\ \mathrm{D}_{\mathrm{jt}}^{\mathrm{exp}}]$. Using equation (12) to substitute ω_{jt} and ω_{jt-1} in equation (A6), and combine the resulting equations (A6) and (13). Then, in the second step, we could express the ϵ_{jt}^ω term as a function of parameters to be estimated $\epsilon_{jt}(\alpha_M, \alpha_K, \beta_{exp}^D)$ and adopt the simple General Moment Method with moment conditions:

$$E\left(\epsilon_{jt}(\alpha_M, \alpha_K, \beta_{exp}^D)\begin{pmatrix} m_{jt-1} \\ k_{jt} \\ \tilde{\mu}_{jt-1} D_{jt-1}^{exp}\end{pmatrix}\right) = 0. \tag{A7}$$

With the estimates of $\hat{\alpha}$ estimated in the first step, we can get the full estimation of all the parameters.

A.3 Markup Derivation for Translog Production Function

Firm j's optimization problem can be described as:

$$\min_{\{L_{jt}, M_{jt}\}} P_{jt}^L L_{jt} + P_{jt}^M M_{jt}$$
$$s.t.\ e^{(\omega_{jt}+\xi_{jt})} L_{jt}^{\alpha_L} M_{jt}^{\alpha_M} K_{jt}^{\alpha_K} L_{jt}^{\alpha_{LL} l_{jt}} M_{jt}^{\alpha_{MM} m_{jt}} K_{jt}^{\alpha_{KK} k_{jt}} L_{jt}^{\alpha_{LM} m_{jt}} M_{jt}^{\alpha_{MK} k_{jt}} L_{jt}^{\alpha_{LK} k_{jt}} \geq Y_{jt}. \quad \text{(A8)}$$

The corresponding Lagrangian function is::

$$\begin{aligned}\mathcal{L} =& P_{jt}^L L_{jt} + P_{jt}^M M_{jt} + \lambda_{jt} \left[Y_{jt} \right. \\ & \left. - e^{(\omega_{jt}+\xi_{jt})} L_{jt}^{\alpha_L} M_{jt}^{\alpha_M} K_{jt}^{\alpha_K} L_{jt}^{\alpha_{LL} l_{jt}} M_{jt}^{\alpha_{MM} m_{jt}} K_{jt}^{\alpha_{KK} k_{jt}} L_{jt}^{\alpha_{LM} m_{jt}} M_{jt}^{\alpha_{MK} k_{jt}} L_{jt}^{\alpha_{LK} k_{jt}} \right]\end{aligned}$$

The First Order Conditions w.r.t. L_{jt} and M_{jt} are shown as follows:

$$\begin{aligned} P_{jt}^L &= \lambda_{jt} \frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}}, \\ P_{jt}^M &= \lambda_{jt} \frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}}, \end{aligned}$$

where $Y_{jt}(\cdot)$ represents the translog production function. Multiplying both sides by L_{jt} and M_{jt} on the above two equations, respectively, and adding them together yield:

$$P_{jt}^L L_{jt} + P_{jt}^M M_{jt} = \lambda_{jt} \Big[\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}} L_{jt} + \frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}} M_{jt} \Big].$$

Dividing revenue R_{jt} ($R_{jt} = P_{jt} Y_{jt}$) by both sides yields:

$$\frac{R_{jt}}{P_{jt}^L L_{jt} + P_{jt}^M M_{jt}} = \frac{P_{jt}}{(\alpha_L^* + \alpha_M^*)\,\lambda_{jt}},$$

where $\alpha_L^* \equiv \alpha_L + 2\alpha_{LL} l_{jt} + \alpha_{LM} m_{jt} + \alpha_{LK} k_{jt}$, $\alpha_M^* \equiv \alpha_M + 2\alpha_{MM} m_{jt} + \alpha_{LM} l_{jt} + \alpha_{MK} k_{jt}$. λ_{jt} is the shadow price of producing one unit of output, representing the marginal cost $\lambda_{jt} \equiv MC_{jt}$. By definition, the markup $\mu_{jt} = P_{jt}/MC_{jt}$. We can get the expression for firm markup μ_{jt} as follows:

$$\begin{aligned} \mu_{jt} &= (\alpha_L^* + \alpha_M^*) \cdot \frac{R_{jt}}{P_{jt}^L L_{jt} + P_{jt}^M M_{jt}} \\ &= (\alpha_L^* + \alpha_M^*) \cdot \tilde{\mu}_{jt} \end{aligned} \quad \text{(A9)}$$

A.4 Discussions on the Potential Underlining Market Structure for the VES Demand Function

The discussion here is to offer another possible market structure, except for the monopolistic competition with variable demand elasticity case in the main text, that could generate a VES demand function for our benchmark estimation. Similar exercises have also been done by, for example, Edmond et al. (2015) and Edmond et al. (2023), where they quantitatively studied the gains from trade and welfare costs by setting up the endogenously variable markup. Arkolakis and Morlacco (2017) also offers a very good review of different settings for variable markups.

Our setting includes three dimensions – industries, products, and firms. Within each industry J, there is a continuum of product varieties k, and there is a certain number of firms producing imperfectly substitute products under the variety k. Each firm will only produce under one product variety and one industry. But there are limit firms producing each product variety, so the market in each product variety is Oligopoly.

Eventually, similar to Atkeson and Burstein (2008), our model predicts that: the market share of firm j in product variety k is decreasing in firm j's price P_{jt}, and the markup of firm j is increasing in its market share of firm j in product variety k. The two mechanisms here may offer one explanation for the origin of the markup's gain – marginal cost decreases after exporting (by productivity gain and IRS); thus, exporting firms can charge a lower price to gain more market share in the Bertrand competition, hence the markup increases.

A.4.1 Households

As in Caliendo and Parro (2015), we assume there are $J = 1, ..., N$ industries and representative households maximize the utility by consuming final goods Y_{Jt}. The preferences of the households are given by:

$$U_t = \prod_{J=1}^{N} C_{Jt}^{\alpha_J}, \text{ where } \sum_{J=1}^{N} \alpha_J = 1, \tag{A10}$$

The market between final goods C_{Jt} is assumed to be perfectly competitive, so the final goods producers don't have the market power and cannot decide the price. Since the utility is in the Cobb-Douglas form, the market share of final goods J is fixed and equals to α_J.

A.4.2 Final Goods: Monopolistic Competition

Similar to Akcigit and Ates (2023)[18], producers of final goods in industry J supply Y_{Jt}, $Y_{Jt} = C_{Jt}$ by market clearing condition, according to the following production technology:

$$Y_{Jt} = \left[\int_0^1 {Y_{kt}}^{1-\frac{1}{\sigma}} dk\right]^{\frac{\sigma}{\sigma-1}}, \tag{A11}$$

[18]The difference is that we change the technology here to be CES.

where Y_{kt} denotes the amount of intermediate variety $k \in [0,1]$, and σ is the constant elasticity of substitution between intermediate varieties. The market for intermediate varieties is monopolistic competition. So as is standard, the theoretical price index $\bar{P}_{Jt}$ for the final goods is given by:

$$\bar{P}_{Jt} = \left[\int_0^1 (P_{kt})^{1-\sigma} \, dk \right]^{\frac{1}{1-\sigma}} . \tag{A12}$$

The product-industry relative demand functions for the output are given by:

$$Y_{kt} = \left(\frac{P_{kt}}{\bar{P}_{Jt}} \right)^{-\sigma} Y_{Jt} \tag{A13}$$

A.4.3 Intermediate Goods: Oligopoly

Within the intermediate product variety k in industry J is oligopoly (with Bertrand Competition). In each industry J and product variety k, there are N_{kt} firms selling the same kind of goods. Output in each intermediate product variety k within industry J is given by a CES production function:

$$Y_{kt} = \left[\sum_{j=1}^{N_{kt}} (Y_{jt})^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \tag{A14}$$

Where Y_{jt} denotes the output of firm j in product variety k in the industry J, and ρ is the constant elasticity of substitution between firms' production. As in Atkeson and Burstein (2008), we assume that firms' goods are imperfect substitutes, and goods within a variety are more substitutable than goods across varieties, i.e., $1 < \sigma < \rho < \infty$. As is standard, the theoretical price index $\bar{P}_{Jt}$ for the industry output is given by:

$$P_{kt} = \left[\sum_{j=1}^{N_{kt}} (P_{jt})^{1-\rho} \right]^{\frac{1}{1-\rho}} . \tag{A15}$$

The firm-product relative demand functions for the output are given by:

$$Y_{jt} = \left(\frac{P_{jt}}{P_{kt}} \right)^{-\rho} Y_{kt}. \tag{A16}$$

We also know that the market share of firm j in product k is given by:

$$S_{jt}^k = \frac{P_{jt}Y_{jt}}{P_{kt}Y_{kt}} = \left(\frac{Y_{jt}}{Y_{kt}}\right)^{-\frac{1}{\rho}} \frac{Y_{jt}}{Y_{kt}} \tag{A17}$$

$$= \left(\frac{Y_{jt}}{Y_{kt}}\right)^{\frac{\rho-1}{\rho}} \tag{A18}$$

$$= \left(\frac{P_{jt}}{P_{kt}}\right)^{1-\rho}. \tag{A19}$$

From the above equation, we can also see that:

$$S_{jt}^k = \left(\frac{P_{jt}}{\left[\sum_{i=1}^{N_{kt}} \left(P_{it}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}}\right)^{1-\rho} = \frac{P_{jt}^{1-\rho}}{\sum_{i=1}^{N_{kt}} \left(P_{it}\right)^{1-\rho}} \tag{A20}$$

$$\Rightarrow \frac{\partial S_{jt}^k}{\partial P_{jt}} = \frac{(1-\rho)\, P_{jt}^{-\rho}}{\sum_{i=1}^{N_{kt}} \left(P_{it}\right)^{1-\rho}} - \frac{(1-\rho)\, P_{jt}^{1-\rho} P_{jt}^{-\rho}}{\sum_{i=1}^{N_{kt}} \left(P_{it}\right)^{1-\rho}} = -\frac{\rho-1}{P_{jt}} S_{jt}^k \left(1 - S_{jt}^k\right) \tag{A21}$$

$$\Rightarrow \frac{\partial S_{jt}^k}{\partial P_{jt}} \leq 0, \tag{A22}$$

which implies that the market share of firm j in product k is decreasing in P_{jt}.

A.4.4 Market Structure: VES Demand Function

As in Atkeson and Burstein (2008), we assume that firms play a static game but with a price (Bertrand) competition.[19] Specifically, each firm j chooses its price P_{jt} for product variety k which belongs to industry J taking as given the prices chosen by the other firms in the economy, as well as the input prices and the industry price $\bar{P}_{Jt}$ and quantity Y_{Jt}. Note that under this assumption, each firm does recognize that product prices P_{kt} and quantities Y_{kt} vary when that firm changes its quantity Y_{jt}.

Then, firm j in producing product k which belongs industry J solve the profit maximization at time t:

$$\max_{P_{jt}, Y_{jt}} P_{jt}Y_{jt} - Cost(Y_{jt}) \tag{A23}$$

$$\text{s.t.} \quad Y_{jt} = \left(\frac{P_{jt}}{P_{kt}}\right)^{-\rho} \left(\frac{P_{kt}}{\bar{P}_{Jt}}\right)^{-\sigma} Y_{Jt} \tag{A24}$$

Equation (A24) is derived from combining equation (A13) and equation (A16), which implies how the firm-industry relative demand is influenced by the firm-product relative price and the product-industry relative price. However, neither the ρ nor σ represents the demand

[19] As discussed in Arkolakis and Morlacco (2017), the Cournot competition case is qualitatively the same as the Bertrand case. Another derivation method of Bertrand competition can also be found in Amiti et al. (2019).

elasticity. To derive the demand elasticity, we first rewrite the profit function in the form of:

$$\pi_{jt} = P_{jt}\left(\frac{P_{jt}}{P_{kt}}\right)^{-\rho}\left(\frac{P_{kt}}{\bar{P}_{Jt}}\right)^{-\sigma} Y_{Jt} - Cost(Y_{jt}) \tag{A25}$$

$$= P_{jt}^{1-\rho} P_{kt}^{\rho-\sigma} \bar{P}_{Jt}^{\sigma} Y_{Jt} - Cost(Y_{jt}) \tag{A26}$$

$$= P_{jt}^{1-\rho}\left\{\left[\sum_{j=1}^{N_{kt}} (P_{jt})^{1-\rho}\right]^{\frac{1}{1-\rho}}\right\}^{\rho-\sigma} \bar{P}_{Jt}^{\sigma} Y_{Jt} - Cost(Y_{jt}) \tag{A27}$$

And the FOC of P_{jt} gives:

$$\frac{\partial \pi_{jt}}{\partial Y_{jt}} = 0 \tag{A28}$$

$$\Rightarrow (1-\rho) P_{jt}^{-\rho}\left[\sum_{j=1}^{N_{kt}} (P_{jt})^{1-\rho}\right]^{\frac{\rho-\sigma}{1-\rho}} \bar{P}_{Jt}^{\sigma} Y_{Jt} + \frac{\rho-\sigma}{1-\rho}\left[\sum_{j=1}^{N_{kt}} (P_{jt})^{1-\rho}\right]^{\frac{\rho-\sigma}{1-\rho}-1} (1-\rho) P_{jt}^{-\rho} P_{jt}^{1-\rho} \bar{P}_{Jt}^{\sigma} Y_{Jt}$$

$$= MC_{jt} \frac{\partial \left(\frac{P_{jt}}{P_{kt}}\right)^{-\rho}\left(\frac{P_{kt}}{\bar{P}_{Jt}}\right)^{-\sigma} Y_{Jt}}{\partial P_{jt}} \tag{A29}$$

$$\Rightarrow (1-\rho) Y_{jt} + (\rho-\sigma)\left(\frac{P_{jt}}{P_{kt}}\right)^{1-\rho} Y_{jt} = MC_{jt}\left[-\rho\frac{Y_{jt}}{P_{jt}} + (\rho-\sigma)\left(\frac{P_{jt}}{P_{kt}}\right)^{1-\rho}\frac{Y_{jt}}{P_{jt}}\right] \tag{A30}$$

$$\Rightarrow \frac{P_{jt}}{MC_{jt}} = \frac{\rho - (\rho-\sigma)\left(\frac{P_{jt}}{P_{kt}}\right)^{1-\rho}}{(\rho-1) - (\rho-\sigma)\left(\frac{P_{jt}}{P_{kt}}\right)^{1-\rho}} \tag{A31}$$

Since we know that $(P_{jt}/P_{kt})^{1-\rho} = S_{jt}^{k}$, and markup is defined as $\eta_{jt}/(1+\eta_{jt})$.[20] So we can derive the variable demand elasticity:

$$\mu_{jt} = \frac{\eta_{jt}}{1+\eta_{jt}} = \frac{\rho - (\rho-\sigma) S_{jt}^{k}}{(\rho-1) - (\rho-\sigma) S_{jt}^{k}} \tag{A32}$$

$$\Rightarrow \eta_{jt} = -\left[S_{jt}^{k}\sigma + \left(1 - S_{jt}^{k}\right)\rho\right] \tag{A33}$$

As discussed in Arkolakis and Morlacco (2017), we can see that the markup is increasing in S_{jt}^{k}. From the previous section, we have already shown that the market share is decreasing in price. Thus, the two mechanisms here may offer an explanation of the markup's increase under this market structure – marginal cost decreases after exporting (by productivity gain and IRS), and exporting firms can charge a lower price to gain more market share; hence the markup increases.

[20] This definition of markup does not rely on monopolistic competition. Because the general demand elasticity is defined by $\eta \equiv (dY/Y)/(dP/P) = (dY/dP)\cdot(P/Y)$. From the F.O.C. of the profit maximization $\max_P\{PY - Cost(Y)\}$, s.t. $Y = D(P)$, we will always have $Y + P\cdot(dY/dP) = MC\cdot(dY/dP) \Rightarrow Y + \eta\cdot Y = MC\cdot\eta\cdot(Y/P) \Rightarrow P/MC = \eta/(1+\eta)$.

In order to construct the relative demand function with the demand elasticity, we substitute ρ with the relationship $\rho = -\left[\left(\eta_{jt} + S_{jt}^k \sigma\right) / \left(1 - S_{jt}^k\right)\right]$ in equation (A24), and thus we can derive the demand function[21] in the form of:

$$\frac{Y_{jt}}{Y_{Jt}} = \left(\frac{P_{jt}}{P_{kt}}\right)^{\frac{\eta_{jt}+S_{jt}^k\sigma}{1-S_{jt}^k}} \left(\frac{P_{kt}}{\bar{P}_{Jt}}\right)^{-\sigma} \tag{A34}$$

$$= \left(\frac{P_{jt}}{\bar{P}_{Jt}}\right)^{\eta_{jt}} \left[\left(\frac{Y_{jt}}{Y_{kt}}\right)^{\frac{(\rho-\sigma)}{\rho} S_{jt}^k} \left(\frac{Y_{kt}}{Y_{Jt}}\right)^{-\frac{(\rho-\sigma)}{\sigma}\left(1-S_{jt}^k\right)}\right] \tag{A35}$$

$$= \left(\frac{P_{jt}}{\bar{P}_{Jt}}\right)^{\eta_{jt}} e^{\bar{y}_{jt}} \tag{A36}$$

where $\bar{y}_{jt} \equiv \frac{(\rho-\sigma)}{\rho\sigma}\left[S_{jt}^k \sigma \ln\left(\frac{Y_{jt}}{Y_{kt}}\right) + \left(1 - S_{jt}^k\right)\rho \ln\left(\frac{Y_{Jt}}{Y_{kt}}\right)\right]$, which can be interpreted as a weighted average of the firm-product relative demand and the industry-product relative demand. As a practical solution, we approximate it with an exponential function $e^{\beta_{exp}^D D_{jt}^{exp}} e^{(1+\eta_{jt})\epsilon_{jt}^D}$. Taking the logs of this demand function will give us the log demand function used in our estimation:

$$y_{jt} = \eta_{jt} p_{jt} - \eta_{jt} \bar{p}_{Jt} + \bar{y}_{Jt} + \beta_{exp}^D D_{jt}^{exp} + \left(1 + \eta_{jt}\right)\epsilon_{jt}^D \tag{A37}$$

[21] We need to aggregate to the industry level, i.e., J-level, because we don't have the product k-level data in the revenue dataset.

A.5 Marginal Cost Derivation and Decomposition of TFPR

Combing (A2) and (A3) yields :

$$\frac{L_{jt}}{M_{jt}} = \frac{\alpha_L P^M_{jt}}{\alpha_M P^L_{jt}}. \tag{A38}$$

Substituting it into the Cobb-Douglas production function yields the expression of M_{jt}:

$$\begin{aligned} Y_{jt} &= e^{(\omega_{jt}+\xi_{jt})} \left(\frac{\alpha_L P^M_{jt}}{\alpha_M P^L_{jt}} M_{jt} \right)^{\alpha_L} M^{\alpha_M}_{jt} K^{\alpha_K}_{jt} \\ \Rightarrow M_{jt} &= Y_{jt}^{\frac{1}{\alpha_L+\alpha_M}} \left[e^{(\omega_{jt}+\xi_{jt})} K^{\alpha_K}_{jt} \left(\frac{\alpha_L P^M_{jt}}{\alpha_M P^L_{jt}} \right)^{\alpha_L} \right]^{\frac{-1}{\alpha_L+\alpha_M}}. \end{aligned} \tag{A39}$$

The total variable costs function can be expressed as:

$$\begin{aligned} Total\ Variable\ Costs_{jt} &= P^L_{jt} L_{jt} + P^M_{jt} M_{jt} \\ &= P^L_{jt} \left(\frac{\alpha_L P^M_{jt}}{\alpha_M P^L_{jt}} M_{jt} \right) + P^M_{jt} M_{jt} \\ &= \frac{\alpha_L + \alpha_M}{\alpha_M} P^M_{jt} M_{jt}. \end{aligned} \tag{A40}$$

Substituting equation (A39) in (A40), and taking the derivative of Y_{jt}, we get the expression of marginal costs in the natural logarithmic form as follows:

$$\begin{aligned} mc_{jt} =& \frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M} y_{jt} - \frac{1}{\alpha_L+\alpha_M} \omega_{jt} + \frac{\alpha_M}{\alpha_L+\alpha_M} p^M_{jt} + \frac{\alpha_L}{\alpha_L+\alpha_M} p^L_{jt} \\ & - \frac{\alpha_M}{\alpha_L+\alpha_M} \ln(\alpha_M) - \frac{\alpha_L}{\alpha_L+\alpha_M} \ln(\alpha_L) - \frac{\alpha_K}{\alpha_L+\alpha_M} k_{jt} - \frac{1}{\alpha_L+\alpha_M} \xi_{jt}, \end{aligned} \tag{A41}$$

where the lowercase variables represent the natural logarithmic form of the original variables. We fix and abstract away changes in input prices and capital stock. So the changes in marginal costs can be expressed in the following form:

$$\Delta mc = \frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M} \Delta y - \frac{1}{\alpha_L+\alpha_M} \Delta\omega. \tag{A42}$$

Then, the changes of TFPR can be decomposed as:

$$\begin{aligned} \Delta \text{TFPR} &= \Delta p + \Delta\omega = \Delta \ln\mu + \Delta mc + \Delta\omega \\ &= \Delta \ln\mu + \frac{1-(\alpha_L+\alpha_M)}{\alpha_L+\alpha_M} \Delta y - \frac{1}{\alpha_L+\alpha_M} \Delta\omega + \Delta\omega \\ &= \Delta \ln\mu + \big(1 - \frac{1}{\alpha_L+\alpha_M}\big)(\Delta\omega - \Delta y). \end{aligned} \tag{A43}$$

A.6 Approximation of Firm Profit Changes after Exporting

The firm j's profit function is:

$$\Pi_{jt} = R_{jt} - (P_{jt}^{L} L_{jt} + P_{jt}^{M} M_{jt}). \tag{A44}$$

Based on the definition of raw markup $\tilde{\mu}_{jt} \equiv R_{jt}/(P_{jt}^{L} L_{jt} + P_{jt}^{M} M_{jt})$, the profit function can be expressed as:

$$\begin{aligned} \Pi_{jt} &= \frac{R_{jt}(\tilde{\mu}_{jt} - 1)}{\tilde{\mu}_{jt}} \\ \pi_{jt} &= r_{jt} + \ln\left(\frac{\tilde{\mu}_{jt} - 1}{\tilde{\mu}_{jt}}\right). \end{aligned} \tag{A45}$$

The approximation of the firm profit changes after exporting is shown as follows:

$$\begin{aligned} \Delta\pi^{exp} &= \pi^{exp=1} - \pi^{exp=0} \\ &= \left(r^{exp=1} - r^{exp=0}\right) + \left\{\left[\ln\left(\frac{\tilde{\mu} - 1}{\tilde{\mu}}\right)\right]^{exp=1} - \left[\ln\left(\frac{\tilde{\mu} - 1}{\tilde{\mu}}\right)\right]^{exp=0}\right\} \\ &= \left(r^{exp=1} - r^{exp=0}\right) + \ln\left(\frac{\tilde{\mu}^{exp=1} - 1}{\tilde{\mu}^{exp=0} - 1}\right) - \ln\left(\frac{\tilde{\mu}^{exp=1}}{\tilde{\mu}^{exp=0}}\right) \\ &= \Delta r^{exp} + \ln\left(1 + \frac{\tilde{\mu}^{exp=1} - \tilde{\mu}^{exp=0}}{\tilde{\mu}^{exp=0} - 1}\right) - \ln\left(1 + \frac{\tilde{\mu}^{exp=1} - \tilde{\mu}^{exp=0}}{\tilde{\mu}^{exp=0}}\right) \\ &\simeq \Delta r^{exp} + \frac{\tilde{\mu}^{exp=1} - \tilde{\mu}^{exp=0}}{\tilde{\mu}^{exp=0} - 1} - \frac{\tilde{\mu}^{exp=1} - \tilde{\mu}^{exp=0}}{\tilde{\mu}^{exp=0}} \\ &= \Delta y^{exp} + \Delta p^{exp} + \frac{1}{\tilde{\mu}^{exp=0}\left(\tilde{\mu}^{exp=0} - 1\right)}\Delta\tilde{\mu}^{exp} \\ &= \Delta y^{exp} + \Delta p^{exp} + \frac{1}{(\alpha_L + \alpha_M)\tilde{\mu}^{exp=0}\left(\tilde{\mu}^{exp=0} - 1\right)}\Delta\mu^{exp} \end{aligned} \tag{A46}$$

A.7 Quantity Survey Data

The output quantity data is from the Annual Product Quantity Survey of the Industrial Enterprises (APQSIE). In this dataset which spans from 2000-2008, the firms have reported their product type (by both product name and 5-digit product code) and the output quantity for each product type with the quantity unit in each year.[22] Besides, since this dataset is also collected by the NBS of China, the firms share the same ID and name that can be easily merged to our main ASIE dataset to acquire the information shown in the Table 1 (e.g. total sales, labor employed, material expenditure, capital stock, etc.), and the final merged sub-sample with the quantity information is also ranged from 2000 to 2006.

A.7.1 Quantity Adjustment w.r.t. Units

Before merging the APQSIE data with the ASIE data, we need to process the quantity output data to make sure the output quantity of different firms is as comparable as possible. To generate the true price data, we only keep those firms with a single product. Then we adjust the quantity w.r.t. the quantity units.

Specifically, we only keep the observations with the quantity unit information. First, for each unit in the list (individual, piece, kilowatt, pair, unit, ton, block, set, square meter, handle, branch, volume, strip, slice, cubic meter, meter, door), we multiply the quantity by 10,000 if the quantity unit is "ten thousand units" and then replace the quantity unit with the corresponding unit without "ten thousand". Next, we multiply the quantity by 1,000 if the quantity unit is "kilometer" and then replace the quantity unit with "meter". If the quantity unit is "core kilometer", we multiply the quantity by 1,000 and then replace the quantity unit with "meter". Similarly, if the quantity unit is "pair kilometer", we multiply the quantity by 2,000 and then replace the quantity unit with "meter". We also multiply the quantity by 1,000 if the quantity unit is "ton", "total ton", "comprehensive ton", or "evaporation ton", and then replace the quantity unit with "kilogram". If the quantity unit is "pair", we multiply the quantity by 2 and then replace the quantity unit with "unit". Finally, we drop observations where the quantity unit is "kilowatt-hour", which is only used for measure the electricity power generation in the sample. Besides, since we use the sales revenue as the output in the main sample, after merging the APQSIE with the main sample, we first generate the output price by output value divided by the output quantity. Then, we use the sales revenue divided by the output price to derive the quantity.

Moreover, we trim the quantity and price data by the largest and smallest 1% to avoid the influence of some extreme values. Eventually, the dataset includes 118,671 firms with 321,280 observations, among which there are 19,838 firms (16.72%) are exporters.

A.7.2 Input Deflators

Another important adjustment we have made is the deflators. Since we only have the material expenditure data, we borrow the industry-year level input deflator P_{Jt}^{M} from Brandt et al.

[22] The product code is missed in the years 2006–2008, and the quantity unit is missed in the years 2000–2003, which requires us to make use of the information in other years to complete the missing information in the years above.

(2017) to get the deflated (log) material m_{jt}. However, we notice the fact that the (log) quantity output shares a very low correlation with the deflated (log) material, compared to the (log) sales revenue: The difference between y_{jt} and the r_{jt} is the removal of the price

Table A1: Correlations between Outputs & Material

	(1)	(2)
	y_{jt}	r_{jt}
m_{jt}	0.305	0.977
	(0.0000)	(0.0000)

Note: P-values in parentheses.

information, which is positively correlated with the material as shown from the Table A1. Inspired by Li and Zhang (2022), we consider the difference is caused by the correlation between input quality and output quality. y_{jt} only contains the product quantity information, where the product quality information is removed, and we are using the industry-year-level material deflator to proxy m_{jt}, where the input quality information is still contained, so r_{jt} shows a higher correlation with m_{jt}. If we directly use the industry-year-level-deflated material in the estimation, the results would be biased. To illustrate this, we write down the (log) production function with both the output quality and input quality[23]:

$$y_{jt} + \chi_{jt}^{Y} = \alpha_L l_{jt} + \alpha_M \left(m_{jt} + \chi_{jt}^{M}\right) + \alpha_K k_{jt} + \omega_{jt} + \xi_{jt}^{\chi}, \tag{A47}$$

where y_{jt} is the quantity output, with the output quality χ_{jt}^{Y}, and m_{jt} is the quantity input, with the input quality χ_{jt}^{M}. If we estimate the production function using the quantity output information from the APQSIE and material expenditure deflated by the industry-year-level material deflator $\bar{P}_{Jt}^{M}$, we are actually estimating the function:

$$y_{jt} = \alpha_L l_{jt} + \alpha_M \ln\left(\frac{E_{jt}^{M}}{\bar{P}_{Jt}^{M}}\right) + \alpha_K k_{jt} + \omega_{jt} + \left(\xi_{jt}^{\chi} - \chi_{jt}^{Y}\right), \tag{A48}$$

$$= \alpha_L l_{jt} + \alpha_M \left(m_{jt} + \chi_{jt}^{M} - \bar{p}_{Jt}^{M}\right) + \alpha_K k_{jt} + \omega_{jt} + \left(\xi_{jt}^{\chi} - \chi_{jt}^{Y}\right), \tag{A49}$$

$$= \alpha_L l_{jt} + \alpha_M \tilde{m}_{jt} + \alpha_K k_{jt} + \omega_{jt} + \tilde{\xi}_{jt}^{\chi}, \tag{A50}$$

where $\tilde{m}_{jt}$ is the quality-inclusive material. However, as illustrated in De Loecker et al. (2016) and Li and Zhang (2022), the input quality is (positively) correlated with the output quality, so the error term $\tilde{\xi}_{jt}^{\chi}$ is correlated with the $\tilde{m}_{jt}$, biasing the estimation results. Besides, even though we take the quality problem into account, the firm-level input price, which is partially contained in the $E_{jt}^{M}/\bar{P}_{Jt}^{M}$, is still positively correlated with the output price, so the endogeneity problem is still severe.

To avoid this endogeneity problem, we make use of the output price to construct a firm-year-level input price index based on the assumption of the linear relationship between input

[23]Detailed discussions can also be found in De Loecker et al. (2016).

quality and output quality:

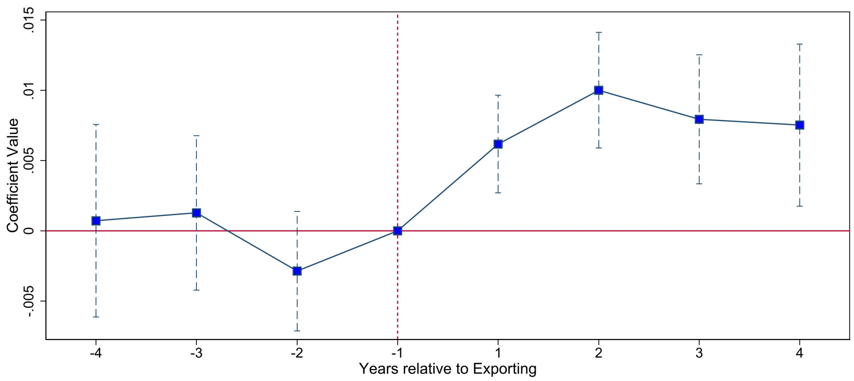
$$\tilde{P}_{jt}^{M} = \bar{P}_{Jt}^{M} \times \frac{P_{jt}}{\bar{P}_{Jt}}, \tag{A51}$$

where $\bar{P}_{Jt}^{M}$ and $\bar{P}_{Jt}$ are industry-year-level deflators w.r.t. material and output as in Brandt et al. (2017), and P_{jt} is the output price calculated by output value divided by the output quantity described before. After using $\ln\left(E_{jt}^{M}/\tilde{P}_{jt}^{M}\right)$ to proxy the material, the correlation between the quantity output and material becomes 0.997.

During the estimation, except for controlling for the city, ownership, and industry fixed effect as in the traditional methods, we also control for the product code $\times$ year fixed effect in the first stage to control for the remaining inconsistency between the quantity output of different firms. Furthermore, in order to be consistent with the estimation in the main sample, we construct the new industry-year-level material deflator $\tilde{P}_{Jt}^{M}$ by calculating the corresponding mean of $\tilde{P}_{jt}^{M}$ and the new industry-year-level output deflator $\tilde{P}_{Jt}$ by calculating $\bar{P}_{Jt} \times \tilde{P}_{Jt}^{M}/\bar{P}_{Jt}$ for both our methods and the traditional methods using the revenue information.

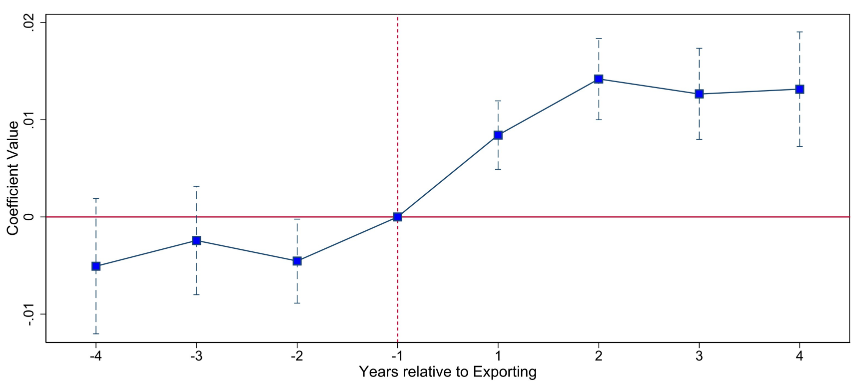
B Appendix - Figures and Tables

Figure B1: Dynamic Exporting Effect on Markup (ACF)



Note: The range represents 95% confidence interval of the parameter estimates.

Figure B2: Dynamic Exporting Effect on Markup for Translog Production Function



Note: The range represents 95% confidence interval of the parameter estimates. Firm size and capital intensity are controlled in the estimation.

Table B1: Single Product Firms' Price Change

Parameter	(1) P_{jt}	(2) $\ln(P_{jt})$	(3) P_{jt}	(4) $\ln(P_{jt})$
D_{jt}^{exp}	-8.335	0.015	-12.449*	-0.037***
	(6.9451)	(0.0104)	(6.9448)	(0.0103)
Firm Size (Sales)			YES	YES
Capital Intensity			YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Adjusted R^2	0.715	0.935	0.715	0.937
Observations	69,367	69,367	69,367	69,367

Note 1: Here we only look at the single-product firms' price change.

Note 2: Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table B2: Implied changes in TFPR

(1) Markup	(2) Productivity	(3) IRS	(4) TFPR
1%	0.084%	-1.302%	-0.218%

Table B5: Markup Decomposition (ACF)

(1) Marginal Cost	(2) Price	(3) Markup
-2.561%	-1.561%	1%

Table B6: Gains from Exporting (ACF)

	(1) Exporter Profit	(2) Consumer Welfare
Gains from Exporting	25.340%	1.561%

Table B3: Exporting Effect on Markup (ACF)

	(1) Markup	(2) Markup
D_{jt}^{exp}	0.008***	0.010***
	(0.0015)	(0.0015)
Firm Size (L)		YES
Capital Intensity		YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	1,234,292	1,234,292
Adjusted R^2	0.326	0.326

Note: Standard errors (clustered at the firm level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table B4: Marginal Cost Decomposition (ACF)

(1) Productivity	(2) IRS	(3) Marginal Cost
-1.305%	-1.256%	-2.561%

Table B7: Estimation Results for Translog Production Function

Parameter	(1) Ours
β_{exp}^{ω} (Productivity Effect)	0.030
	(0.0006)
β_{exp}^{ω} (Demand Effect)	0.220
	(0.0044)
S (RTS)	1.109
	(0.0028)
S^V (RTSV)	1.070
	(0.0025)
α_L^*	0.074
	(0.0014)
α_M^*	0.996
	(0.0014)
α_K^*	0.039
	(0.0003)
α_L	0.066
	(0.0030)
α_M	0.800
	(0.0019)
α_K	-0.023
	(0.0001)
α_{LL}	0.034
	(0.0001)
α_{MM}	0.017
	(0.0000)
α_{KK}	0.008
	(0.0000)
α_{LM}	-0.023
	(0.0002)
α_{LK}	-0.012
	(0.0001)
α_{MK}	-0.002
	(0.0001)
Observations	1,234,292

Note: Standard errors (clustered at the firm level) in parentheses.

Table B8: Industry Separated Estimation Results (Ours)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
								Marginal Cost Decomposition		
Industry Code (name)	β^{ω}_{exp}	β^{D}_{exp}	α_L	α_M	α_K	S **(RTS)**	S^V **(RTSV)**	**Marginal Cost**	**IRS**	**Productivity**
13 (Processing of Food from Agricultural Products)	0.001	0.018	0.044	0.983	0.044	1.072	1.027	-0.006	-0.005	-0.001
	(0.0025)	(0.0206)	(0.0022)	(0.0224)	(0.0020)	(0.0223)	(0.0234)	(0.0050)	(0.0042)	(0.0024)
14 (Manufacture of Foods)	0.013	0.166	0.048	0.969	0.022	1.039	1.017	-0.016	-0.003	-0.013
	(0.0030)	(0.0448)	(0.0043)	(0.0323)	(0.0061)	(0.0288)	(0.0344)	(0.0075)	(0.0069)	(0.0029)
15 (Manufacture of Beverages)	0.023	0.159	0.064	0.927	0.040	1.031	0.991	-0.022	0.002	-0.024
	(0.0142)	(0.1435)	(0.0071)	(0.0740)	(0.0126)	(0.0671)	(0.0791)	(0.0173)	(0.0187)	(0.0163)
16 (Tobacco)	0.346	-4.517	0.063	0.738	0.076	0.877	0.801	-0.255	0.177	-0.432
	(0.3996)	(1.9147)	(0.0318)	(0.1432)	(0.0729)	(0.0959)	(0.1530)	(0.3984)	(0.1295)	(0.4499)
17 (Textiles)	0.008	0.169	0.069	1.080	0.030	1.179	1.149	-0.034	-0.027	-0.007
	(0.0020)	(0.0178)	(0.0029)	(0.0328)	(0.0017)	(0.0339)	(0.0351)	(0.0061)	(0.0068)	(0.0019)
18 (Garment, Foot Ware, and Caps)	0.022	0.179	0.085	0.770	0.039	0.894	0.855	-0.003	0.024	-0.026
	(0.0088)	(0.0718)	(0.0314)	(0.2848)	(0.0040)	(0.3193)	(0.3161)	(0.0345)	(0.0442)	(0.0106)
19 (Leather, Fur, Feathers, and Related Products)	0.011	0.110	0.106	1.143	0.010	1.259	1.249	-0.053	-0.044	-0.009
	(0.0039)	(0.0224)	(0.0083)	(0.0731)	(0.0071)	(0.0740)	(0.0803)	(0.0147)	(0.0155)	(0.0032)
20 (Timber, Manufacture of Wood, Bamboo, Rattan, Palm, and Straw Products)	-0.002	0.160	0.073	1.139	-0.029	1.182	1.211	-0.021	-0.023	0.002
	(0.0065)	(0.0542)	(0.0045)	(0.0297)	(0.0226)	(0.0155)	(0.0317)	(0.0116)	(0.0089)	(0.0054)
21 (Furniture)	0.010	0.269	0.054	1.036	0.007	1.098	1.090	-0.028	-0.019	-0.009
	(0.0163)	(0.3675)	(0.0083)	(0.1217)	(0.0097)	(0.1206)	(0.1288)	(0.0137)	(0.0324)	(0.0216)
22 (Paper and Paper Products)	0.009	0.377	0.049	0.949	0.030	1.028	0.998	-0.008	0.000	-0.009
	(0.0165)	(0.2316)	(0.0056)	(0.0957)	(0.0049)	(0.0966)	(0.1004)	(0.0141)	(0.0125)	(0.0158)
23 (Printing, Reproduction of Recording Media)	0.007	0.220	0.074	0.948	0.056	1.078	1.022	-0.011	-0.005	-0.007
	(0.0080)	(0.1481)	(0.0058)	(0.0359)	(0.0047)	(0.0397)	(0.0396)	(0.0089)	(0.0090)	(0.0086)
24 (Articles for Culture, Education, and Sport Activities)	0.007	0.155	0.092	0.973	0.018	1.084	1.065	-0.013	-0.006	-0.006
	(0.0060)	(0.0860)	(0.0097)	(0.0895)	(0.0076)	(0.0915)	(0.0982)	(0.0047)	(0.0101)	(0.0070)
25 (Petroleum, Coking, and Processing of Nuclear Fuel)	0.099	2.204	0.026	0.780	0.052	0.858	0.806	-0.087	0.036	-0.123
	(0.0464)	(0.8794)	(0.0047)	(0.0467)	(0.0091)	(0.0402)	(0.0475)	(0.0510)	(0.0277)	(0.0610)
26 (Raw Chemical Materials and Chemical Products)	0.011	0.172	0.056	1.021	0.026	1.104	1.078	-0.025	-0.015	-0.010
	(0.0026)	(0.0176)	(0.0023)	(0.0102)	(0.0022)	(0.0097)	(0.0111)	(0.0029)	(0.0024)	(0.0025)
27 (Manufacture of Medicines)	0.010	0.230	0.050	1.006	0.052	1.108	1.057	-0.016	-0.007	-0.009
	(0.0133)	(0.1779)	(0.0076)	(0.1167)	(0.0116)	(0.1132)	(0.1228)	(0.0118)	(0.0187)	(0.0162)
28 (Chemical Fibers)	0.001	0.565	0.043	0.985	0.018	1.047	1.028	-0.005	-0.004	-0.001
	(0.0180)	(0.1708)	(0.0066)	(0.0491)	(0.0072)	(0.0465)	(0.0520)	(0.0146)	(0.0103)	(0.0184)
29 (Rubber)	-0.003	0.219	0.059	1.064	0.007	1.129	1.123	-0.021	-0.023	0.002
	(0.0329)	(0.1510)	(0.0074)	(0.0825)	(0.0135)	(0.0743)	(0.0872)	(0.0314)	(0.0183)	(0.0292)
30 (Plastics)	0.004	0.355	0.058	0.968	0.030	1.056	1.026	-0.008	-0.004	-0.004
	(0.0152)	(0.2898)	(0.0069)	(0.1043)	(0.0079)	(0.1031)	(0.1107)	(0.0052)	(0.0199)	(0.0194)
31 (Non-Metallic Mineral Products)	0.010	0.293	0.058	0.996	0.019	1.073	1.054	-0.019	-0.001	-0.009
	(0.0110)	(0.3818)	(0.0070)	(0.1118)	(0.0124)	(0.1064)	(0.1186)	(0.0171)	(0.0256)	(0.0145)
32 (Smelting and Pressing of Ferrous Metals)	0.018	0.279	0.043	1.059	0.040	1.142	1.102	-0.014	0.002	-0.016
	(0.0079)	(0.0579)	(0.0038)	(0.0182)	(0.0029)	(0.0180)	(0.0192)	(0.0099)	(0.0058)	(0.0073)
33 (Smelting and Pressing of Non-Ferrous Metals)	0.002	1.665	0.047	0.798	0.059	0.904	0.845	0.025	0.027	-0.002
	(0.0141)	(0.3265)	(0.0033)	(0.0149)	(0.0037)	(0.0157)	(0.0155)	(0.0148)	(0.0077)	(0.0165)
34 (Metal Products)	0.001	0.173	0.059	1.086	0.033	1.178	1.145	-0.019	-0.018	-0.001
	(0.0035)	(0.0227)	(0.0031)	(0.0134)	(0.0024)	(0.0141)	(0.0145)	(0.0049)	(0.0039)	(0.0031)
35 (General Purpose Machinery)	0.005	0.222	0.057	1.048	0.030	1.134	1.105	-0.023	-0.018	-0.005
	(0.0023)	(0.0204)	(0.0023)	(0.0102)	(0.0016)	(0.0103)	(0.0110)	(0.0031)	(0.0025)	(0.0021)
36 (Special Purpose Machinery)	0.020	0.278	0.061	0.991	0.016	1.068	1.052	-0.027	-0.008	-0.019
	(0.0033)	(0.0816)	(0.0047)	(0.0526)	(0.0052)	(0.0515)	(0.0561)	(0.0077)	(0.0092)	(0.0038)
37 (Transport Equipment)	0.009	0.153	0.071	0.991	0.031	1.093	1.062	-0.022	-0.014	-0.008
	(0.0031)	(0.0260)	(0.0037)	(0.0156)	(0.0026)	(0.0153)	(0.0170)	(0.0037)	(0.0042)	(0.0029)
39 (Electrical Machinery and Equipment)	0.006	0.127	0.065	1.036	0.040	1.141	1.101	-0.021	-0.015	-0.006
	(0.0024)	(0.0220)	(0.0031)	(0.0239)	(0.0024)	(0.0242)	(0.0257)	(0.0041)	(0.0039)	(0.0022)
40 (Communication Equipment, Computers, and Other Electronic Equipment)	0.008	0.138	0.112	1.060	0.010	1.182	1.172	-0.028	-0.021	-0.007
	(0.0057)	(0.0341)	(0.0082)	(0.0628)	(0.0055)	(0.0664)	(0.0696)	(0.0076)	(0.0091)	(0.0051)
41 (Measuring Instruments and Machinery for Cultural Activities and Office Work)	0.008	0.241	0.076	1.013	0.017	1.107	1.089	-0.019	-0.012	-0.007
	(0.0080)	(0.0614)	(0.0068)	(0.0385)	(0.0046)	(0.0409)	(0.0419)	(0.0073)	(0.0058)	(0.0075)
42 (Artwork and Other Manufacturing)	0.010	0.143	0.084	0.942	0.023	1.050	1.027	-0.012	-0.003	-0.009
	(0.0018)	(0.0249)	(0.0045)	(0.0319)	(0.0031)	(0.0328)	(0.0352)	(0.0042)	(0.0041)	(0.0018)
Mean	0.010	0.239	0.063	1.003	0.029	1.095	1.066	-0.028	-0.001	-0.027
Observations	1,234,292									

Note: Standard errors (clustered at the firm level) in parentheses.

Table B9: Exporting Effect on Markup for Translog-form Production Function

	(1) Markup	(2) Markup
D^{exp}_{jt}	0.017***	0.013***
	(0.0016)	(0.0016)
Firm Size (L)		YES
Capital Intensity		YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	1,234,292	1,234,292
Adjusted R^2	0.329	0.330

Note: Standard errors (clustered at the firm level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table B10: Estimation Results of Export Instrument (Lagged Export Dummy)

Parameter	(1) Ours
β^{ω}_{exp} (Productivity Effect)	0.013
	(0.0037)
β^{ω}_{exp} (Demand Effect)	0.250
	(0.0868)
α_L	0.081
	(0.0007)
α_M	0.980
	(0.0020)
α_K	0.031
	(0.0008)
S (RTS)	1.092
	(0.0016)
S^V (RTSV)	1.061
	(0.0023)
Observations	1,234,292

Note: Standard errors (clustered at the firm level) in parentheses.

Table B11: Exporting Effect on Log-Markup (LP)

	(1) Log-Markup	(2) Log-Markup
D_{jt}^{exp}	0.009^{***}	0.010^{***}
	(0.0012)	(0.0012)
Firm Size (L)		YES
Capital Intensity		YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	1,234,292	1,234,292
Adjusted R^2	0.345	0.346

Note 1: Standard errors (clustered at the firm level) in parentheses.

Note 2: Control variables are in log-form.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table B12: Exporting Effect on Log-Markup (ACF)

	(1) Log-Markup	(2) Log-Markup
D_{jt}^{exp}	0.009^{***}	0.010^{***}
	(0.0012)	(0.0012)
Firm Size (L)		YES
Capital Intensity		YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	1,234,292	1,234,292
Adjusted R^2	0.345	0.346

Note 1: Standard errors (clustered at the firm level) in parentheses.

Note 2: Control variables are in log-form.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table B13: Exporting Effect on Log-Revenue

	(1) Log-Revenue	(2) Log-Revenue
D_{jt}^{exp}	0.211***	0.210***
	(0.0045)	(0.0045)
Capital Intensity		YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	1,234,292	1,234,292
Adjusted R^2	0.823	0.823

Note 1: Standard errors (clustered at the firm level) in parentheses.

Note 2: Control variables are in log-form.

* $p < .10$, ** $p < .05$, *** $p < .01$