Does exporting increase productivity?
Evidence from India

Megha Mukim*
London School of Economics
March 15, 2011

Abstract

Although there is much empirical evidence to show that more productive firms become exporters, the literature has been inconclusive regarding productivity benefits from exporting. This paper disentangles the direction of the causality to show that exporting improves firm performance. It uses Indian plant-level data (over 1989-2008) for over 15,000 firms across 2-digit product categories, to test for learning-by-exporting effects. In the paper, I generate measures of total factor productivity by estimating production functions using firm-level data. I follow Olley and Pakes (1996) and use investment to deal with the problem of simultaneity and endogenous exit. To deal with the self-selection bias, I use a two-fold identification procedure: propensity score matching and instrumental variables techniques. I find that exporting has a strong positive effect on the productivity of export-market entrants, but that this effect tapers over time. My results are robust to different identification techniques and model specifications. This paper contributes to the empirical literature by measuring the effects of learning-by-exporting.

JEL Classification: F14, D24
Keywords: Trade, Productivity, Learning-by-Exporting

* Department of International Development, London School of Economics, Email: m.mukim@lse.ac.uk; Website: http://personal.lse.ac.uk/mukim. I am grateful to Diana Weinhold and TN Srinivasan for helpful comments and suggestions, and to David Weinstein, Amit Khandelwal and other participants at the Columbia Trade Colloquium for useful discussions.
1 Introduction

An important empirical finding of the literature (Roberts and Tybout 1997, Clerides et al 1998, Bernard and Jensen 1999, Van Biesebroeck 2006, Alvarez and Lopez 2005) is that exporters are more productive than non-exporters. There are two mechanisms that can explain this difference - the first is related to self-selection and the second refers to learning-by-exporting. Exporters may be more productive than their counterparts, who only supply the domestic market, simply because more productive firms are able to engage in export activity and compete in international markets. The second, more important mechanism, and one which this paper will focus on, is learning-by-exporting by firms, in other words, post-entry productivity benefits. The idea is that when firms enter into export markets they gain new knowledge and expertise, which allows them to improve their level of efficiency. However, the two effects are not mutually exclusive – it is possible that high productivity firms that enter the export market continue to improve their productivity because of their exposure to exporting.

In this paper, I examine the link between exporting and productivity for manufacturing firms in India. I study firm-level data in India at the National Industrial Classification (NIC) 2-digit classification level for the period 1990-2008. With the exception of the last few years, this period saw significant entry into exports markets – see Figure 1. I find three main results in this paper. First, I demonstrate that firms’ productivity is in fact positively related to their participation in export markets. Second, I show that part of the increase in productivity is accounted for by learning-by-exporting, after having controlled for any self-selection into exporting. And finally, I find that the learning-by-exporting effect is the highest for the first few years after entry into exporting and then begins to level off.

The remainder of the paper is organised as follows: The next section provides a descriptive overview of the theoretical and empirical literature on exporting and firm productivity. Section III describes the data and shows that exporters are in fact different from non-exporters in a number of ways. Section IV outlines the empirical specification for unbiased production function estimates, and the identification of gains from exporting. Section V carries out robustness checks and Section VI concludes.

2 Theory and Evidence

The potential link between trade and economic growth has been fundamental to international and to development economics. This paper concerns itself with the question of whether firms achieve higher productivity growth by becoming exporters.

The empirical and theoretical literatures have moved forward in fits and starts. Earlier endogenous growth models (Grossman and Helpman 1991, Rivera-Batiz and Romer 1991) predicted that international technology diffusion through exposure to export
markets could boost within-plant productivity\(^1\). Traditional export-led growth hypothesis (Kaldor 1970, Dixon and Thirlwall 1975) posited that external demand would enable firms to exploit economies of scale\(^2\) leading to productivity growth. Firms could also invest in productivity-enhancing technology in anticipation of larger export markets (Goh 2000).

Another microeconomic channel is the reallocation of economic activity across firms within industries. Models by Helpman and Krugman (1985) and Krugman (1994) predicted that average productivity could rise if resources were shifted to industries with lower average costs. Heterogeneous firm models (Melitz 2003 and Bernard et al 2003) also argue that the existence of trade costs allows only the most productive firms to enter export markets. As low productivity firms exit, output and employment are reallocated towards higher productivity firms and average industry productivity increases. In other words, it is the reallocation of activity across firms, and not intra-firm productivity growth, that drives industry-level productivity.

The robust empirical findings of a number of papers in the literature provides backing for the predictions of these theoretical models and demonstrates that differences between exporters and non-exporters could arise whether or not exporting enhanced productivity.

The empirical literature on productivity and exporting over the years has grown quite rapidly. There are many studies that find little or no evidence of learning-by-exporting. See Table 1 for a summary of some key studies. For instance Bernard and Jensen (1999) find that the benefits from exporting for American firms are unclear. Although employment, growth and profitability are higher for exporters, productivity and wage growth is not superior. Kim (2000) finds only marginal increases in productivity following trade liberalisation in Korea. Delgado et al (2002) find evidence of higher productivity for exporters versus non-exporters, and evidence of self-selection of more productive firms into the export market. However, they do not find much evidence to support the learning-by-doing hypothesis, and if so, only for younger exporters. Castellani (2002) finds evidence of productivity gains associated with increases in export intensity. Other studies, such as Isgut (2001) and Clerides et al (1998), that use a variety of econometric methods and data from several countries, also conclude in favour of the self-selection and against the learning-by-exporting hypothesis. Only the most productive firms have a sufficient cost advantage to overcome transportation costs and compete internationally. Exporters are more productive than non-exporters, not because there are any benefits associated with export activities, but they are simply more productive from the outset. Some studies look at trade liberalisation in general and not just exporting. Hung et al (2004) find that exporting activity itself does not seem to promote productivity in the US, and that it is import competition that attributed for the largest part of labour productivity in manufacturing during 1996-2001. Fernandes (2007) finds that trade liberalisation has a strong positive impact on plant productivity in Columbia.

\(^1\) Industrial-level productivity could also rise if individual firms’ new technological learning’s spill over and positively affect the total stock of knowledge for all firms, thus raising aggregate productivity.

\(^2\) Firms move to a lower point on the average cost curve since a rise in output is accompanied by a less than proportionate rise in average costs.
However, some studies do find empirical support for post-entry productivity gains. For instance, Kraay (1999) for China, Bigsten et al (2004) for sub-Saharan Africa, and Aw et al (2000) for Taiwan, find evidence supporting the learning-by-exporting hypothesis. Loecker (2007) finds that Slovenian export entrants become more productive once they start exporting, and that the productivity gap between exporters and their domestic counterparts widens over time. He also finds that productivity gains are higher for firms exporting towards higher-income regions. Biesebroek (2005) finds evidence that sub-Saharan exporters are more productive than their counterparts who only serve the domestic market, and that the former enjoy increasing rates of productivity growth – in support of the learning-by-exporting hypothesis. Other examples of some studies are Castellani (2002), Baldwin and Gu (2003, 2004), Blalock and Gertler (2004), Girma et al. (2004) and Greenaway and Kneller (2008). However, not all studies are able to describe the source of these learning effects. A notable exception is Baldwin and Gu (2004). From their analysis of Canadian plants they conclude that exporters learn from participation in export markets through channels that include new innovations, as well as technology transfer from abroad and investments in absorptive capacity such as human capital.

A recent working paper by Tabrizy and Trofimenko (2010) study the learning-by-exporting effects for firms in India. They find no evidence of post-entry productivity gains and conclude that productivity differences between exporters and non-exporters are explained only by self-selection. There are some crucial differences between this paper and their study. Although the sample of firms comes from the same data source, their period of study covers only a ten-year period (1998-2008) whilst this paper studies firms over a twenty-year period (1989-2008). Additionally their measure of productivity, which has been calculated using Levinsohn-Petrin techniques, controls for simultaneity bias but does not control for endogenous exit. As capital-intensive firms are better able to weather a negative productivity shock and thus more likely to survive in the market, and since exporters also tend to be more capital-intensive than non-exporters, this implies that their productivity estimate does not account for the large downward bias on the capital coefficient. They deflate firm-level data using a national wholesale price index and not industry-specific input and output deflators. Most importantly, their paper describes pre-and post-entry productivity differentials between exporters and non-exporters mainly by using dummy variables for different types of exporting behaviour. Since they do not control for self-selection using robust identification techniques, I am unable to compare the reasons underlying our diverging conclusions regarding the effects of learning-by-exporting for Indian firms.

It is interesting that even for the same countries different studies (albeit, not always for overlapping periods in time) find confirming or conflicting evidence of learning-by-exporting effects. For instance, both papers that study the United States (Bernard and Jensen 1999, and Hung et al 2004) find no evidence of learning-by-exporting. There are three studies for Germany, of which the first finds a positive result (Bernard and Wagner 1997), while the later two (Wagner 2002, and Arnold and Hussinger 2004) find no effect. Both studies for the United Kingdom (Girma et al 1004 and Greenaway and Kneller 2008) find positive effects. For Columbia, the first two studies (Clerides et al 1998 and Isgut 2001) find no evidence, while a later study (Fernandes 2007) finds some evidence. A quick and dirty regression shows that if a study focuses on a developed country or if it
uses cross sectional analysis, it is less likely to find learning-by-exporting effects (the number of firms in the sample or panel analyses don’t seem to matter much).

It is possible that firms in developing countries that are further away from the technological frontier and that export to other, perhaps more developed, markets also make larger strides in productivity increases. According to the Global Economic Prospects report (World Bank 2008), progress in developing countries reflects the absorption of pre-existing technologies and not at-the-frontier inventions. The technology achievement index\(^3\) for developing countries clearly shows that India is a laggard – the highest value is 0.25 for the United States. Thus, one might expect Indian exporters to enjoy large productivity increases as compared their other domestic counterparts, whilst US exporters may not be much more productive than non-exporters.

3 Data and preliminary analysis

3.1 Data description

Firm-level data on output and inputs is drawn from the Prowess database. Prowess is a corporate database that contains normalised data built on a sound understanding of disclosures of over 20,000 companies in India. The database provides financial statements, ratio analysis, fund flows, product profiles, returns and risks on the stock market etc. The Centre for Monitoring the Indian Economy (CMIE), which collects data from 1989 onwards, assembles the Prowess database. The database contains information on 23,168 firms for the year 1989 – 2008, yielding a total of 437,283 observations. On average there are 8 years of data on each firm. However, data are either not available or are reported as missing values for a number of observations for different variables such as sales, capital stock and wages. After cleaning the data\(^4\), the final dataset contains 15,292 firms for the years 1989-2008, yielding a total of 106,712 observations.

There is a large degree of firm heterogeneity in terms of size and age. Firms in the sample also include both exporters and non-exporters – a total of 7,716 firms enter the export market at least once over the period of study. However, some caveats should be mentioned here. It is not mandatory for firms to supply data to the CMIE, and one cannot tell exactly how representative of the industry is the membership of the firms in the organisation. Prowess covers 60-70 percent of the organised sector in India, 75 percent of corporate taxes and 95 percent of excise duties collected by the Government of India (Goldberg et al 2010\(^5\)). Large firms, which account for a large percentage of industrial

---

\(^3\) The technology achievement index is published by the United Nations Development Programme and combines (a) the indicators of human skills (mean years of schooling in the population age 15 and older and enrollment ratio for tertiary-level science programs); (b) the diffusion of old innovations (electricity consumption per capita and telephones per capita) and of recent innovations (Internet hosts per capita and high- and medium-tech exports as a share of all exports); and (c) the creation of technology (patents granted to residents per capita and receipts of royalties and license fees from abroad). The index is constructed as simple averages of these indicators within subgroups and then across groups.

\(^4\) I exclude firms for which data on sales, gross assets and wages are missing.

\(^5\) Quoted in earlier version of NBER working paper.
production and foreign trade, are usually members of the CMIE and are more likely to be included in the database. And so, the analysis is based on a sample of firms that is, in all probability, taken disproportionately from the higher end of the size distribution. As Tybout and Westbrook (1994) point out, a lot of productivity growth comes from larger plants, so a more comprehensive study might have found smaller average residual effects.

3.2 Preliminary Analysis: Are exporters more productive?

At the outset I am interested in knowing whether the facts found in the literature – that exporters differ from non-exporters – also hold for firms in India. Regressing firm characteristics on a dummy for whether the firm exports (and controls), a number of studies have documented that exporters differ from non-exporters in important ways. Following Bernard and Jensen (1999) and others, I run the following OLS regression:

\[ x_{itkj} = \alpha + \beta \text{EXP}_{itkj} + \gamma \text{Controls}_{itkj} + \sum_t \delta_t \text{Time}_t + \sum_k \lambda_k \text{Ind}_k + \sum_j \xi_j \text{State}_j + \varepsilon_{ikt} \]  

(1)

where \( x_{itkj} \) refers to the characteristics of firm \( i \) at period \( t \) active in industry \( k \) in state \( j \), \( \text{EXP} \) is an export dummy equal to one when the firm is an exporter and zero otherwise. Firm-specific controls include the size (number of employees), the age and the type of firm (private domestic, private foreign, public or mixed). I also control for industry, year and location effects, where subscripts \( k \), \( t \) and \( j \) run through the number of industries (\( \text{Ind} \)), years (\( \text{Time} \)), and states (\( \text{State} \)) respectively. The interest lies in the coefficient \( \beta \) that tells us whether the relevant firm characteristic is different for exporting firms relative to non-exporting ones. Moreover it has a clear economic interpretation, i.e. it reveals the percentage differential between exporters and non-exporters.

Table 3 shows that exporters have a higher wage bill (143 per cent higher), operate on a larger scale, add higher value, sell more and invest more than non-exporters.6 These results are in line with a number of papers in the literature (see Table 1). The strong positive association between productivity and participation in export markets could reflect the self-selection of better firms into export markets and/or it could reflect the effect of exporting on productivity. The point of this paper is to disentangle the effect of exporting on productivity by controlling for the self-selection effect.

4 Empirical Specification

Following the influential papers of Bernard and Jensen (1999) and Clerides et al (1998), the literature has used two main methods to measure learning-by-exporting effects. The first method consists of separating the sample into mutually exclusive groups, such as exporters and non-exporters, to assess differences in plant performance between these groups (see Loecker 2007, Greenaway and Kneller 2008, Girma et al 2004). The second method of measurement of learning-by-exporting effects consists of one or more

---

6 Nominal values are deflated using NIC-level output and input price indices.

The measure of plant performance used in this paper to assess the presence of learning-by-exporting effects is total factor productivity (TFP)\(^7\). I use the two-step approach to estimate the causal effect of exporting on firm-level productivity, and later in the paper I also use the one-step/direct approach, among others, as a robustness check.

### 4.1 Estimating Productivity

The estimation of production functions can be affected by two different sources of bias. Since firms’ inputs and outputs are simultaneously chosen, inputs will be correlated with any shocks, say demand or productivity shocks, that would be captured in the error term and co-efficient estimates will be biased. Under fairly general assumptions\(^8\), Levinsohn and Petrin (2003) show that under simple OLS estimations the labour co-efficient will be upward biased and the capital co-efficient will be downward biased, implying that productivity estimates will be upward biased for more capital-intensive firms (such as exporters). On the other hand, the selection problem is generated by the relationship between the unobserved productivity variable and the shutdown decision. In this case, firms’ choices on whether to exit the export market depend on their productivity. Olley and Pakes (1996) obtain consistent production function estimates controlling for the fact that firms’ choices on whether to exit the market depends on their productivity\(^9\).

This paper follows Olley and Pakes (1996) – henceforth referred to as OP – to obtain consistent production function estimates. The OP approach uses investment to control for the simultaneity between inputs and outputs. Consider the following production function:

\[
Y_{it} = A_{it} L_{it}^{\beta_l} I_{it}^{\beta_i} K_{it}^{\beta_k}
\]

\(^7\) TFP measures the economic and technical efficiency with which resources are converted into products.

\(^8\) Levinsohn and Petrin (2003) consider the bias in three different cases: when only labour responds to the shock and capital is not correlated with labour (the labour co-efficient will be biased upwards, and the capital co-efficient will be unbiased); when only labour responds to the shock and capital and labour are positively correlated (the labour co-efficient will be biased upwards, and the capital co-efficient will be biased downwards); when labour and capital respond to the shock, the two are positively correlated and labour responds more strongly to the shock (the labour co-efficient will be biased upwards and the capital co-efficient will be biased downwards).

\(^9\) Olley and Pakes (1996) use investment as a proxy to control for the simultaneity problem, i.e. when inputs are endogenous to productivity.
where $Y_t$ is output, $A_{it}$ is total factor productivity, and $L_{it}, K_{it}, I_{it}$ represent labour, capital and investment, respectively. TFP is modelled as:

$$A_{it} = \exp(\omega_{it} + \epsilon_{it})$$

where $\omega_{it}$ is a firm-specific productivity shock known to the firm manager, but unknown to the econometrician, and $\epsilon_{it}$ is a zero-mean productivity shock realised after variable inputs have been chosen.

The production function is estimated as follows, where output is expressed as a function of the log of inputs and shocks:

$$y_{it} = \beta_0 + \beta_1 l_{it} + \beta_2 k_{it} + \beta_3 i_{it} + \omega_{it} + \epsilon_{it}$$

(4)

As mentioned above, there is a possibility that the coefficients on the variable input, i.e. labour, are upwardly biased and that there is a corresponding downward bias in the coefficient on the quasi-fixed input, i.e. capital. To obtain consistent production function estimates, I use the OP procedure. Estimation proceeds in two stages. In the first stage, I obtain the coefficients on labour by semi-parametric techniques. It is assumed that a firm’s demand for investment increases monotonically with productivity, conditional on capital. Then the inverse of the investment demand function depends only on observable inputs and capital, and its non-parametric estimate can be used to control for unobservable productivity, removing the simultaneity bias. Since productivity is assumed to affect capital with a lag, there is no simultaneity problem in estimation the coefficient on capital. Loecker (2010) argues that if firms that export are also firms that invest more, then using the OP procedure would overestimate the capital coefficient and underestimate the returns from exporting. I do not include past exporting experience when estimating productivity, although I do carry out a robustness check by including lagged exports as a state variable later in the paper.

The OP procedure also controls for the endogeneity of firm exit by computing survival probabilities for the firm. The probability that the firm survives in the market depends on lagged values of capital and the proxy for productivity. These probabilities control for the selection bias and are based on some threshold of productivity below which a firm exits the market. The survival probabilities are then introduced into the production function to generate the coefficient on capital. In Appendix A, I discuss the estimation algorithm for getting reliable estimates of the production function in more detail. Nominal values have been deflated using output (sales) and input (labour, capital, investment) deflators.

I re-run Equation (1) with the computed TFP levels and growth rates as the dependent variable, and find that exporters are 27.7 per cent more productive than non-exporters. I rescale time in such a way so that $t=1$ refers to the year when the firm

---

10 More specifically, it is assumed that productivity follows a Markov process: $\omega_{it} = E[\omega_{it} | \omega_{it-1}] + \xi_{it}$ where $\xi_{it}$ represents the unexpected part of current productivity to which capital does not adjust.

11 Where the input deflators are constructed by their weight in the Consumer Price Index basket, and the data is taken from the Central Statistical Organisation.
begins to export, $t = 2$ is when the firm has exported for a year and so on and so forth. The relationship between productivity and exporting is shown in Figure 3. The bold lines refer to the estimated coefficient for the dummy variable that equals 1 if the firm exports, after controlling for year, industry and firm characteristics (size and age). The dashed lines show the 99% confidence intervals around the parameter estimates. The levels of TFP jump at the time of entry into export markets and that they continue to grow after entry. The downward slope leading towards entry into the export market (i.e. towards $t=0$) could be explained by the much larger number of firms closer to $t = 0$, i.e. at $t = -1$, which lowers the average productivity per firm.\textsuperscript{12} In addition, the figure does not control for self-selection, and at this point the paper does not claim any causal effect.

\section*{4.2 Identification of productivity gains from exporting}

\textbf{Propensity Score Matching Techniques}

Following Girma et al (2004) and Loecker (2007), I control for the self-selection process, while testing for the learning-by-exporting hypothesis, by creating control groups using matching techniques based on average treatment models as suggested by Heckman et al (1997). The aim of this methodology is to evaluate the causal effect of exporting on productivity by matching export starters with non-exporters. The identifying assumption in estimating the treatment effect (i.e. exporting) comes from the introduction of the state variable – lagged productivity – in the matching procedure and this has a strong interpretation in the underlying structural framework. The method constructs a counterfactual that allows me to analyse how productivity of a firm would have evolved if it had not started exporting. The main problem in this type of analysis is that one does not observe the counterfactual and therefore it is necessary to match the exporting firm with a control group of similar firms that do not export.

The aim is to evaluate the causal effect of exporting on the performance indicator – here, TFP. Following Loecker (2007) I rescale the time periods in such a way that a firm starts exporting at $s = 0$. Let $\omega_{is}$ be the outcome at time $s$ - the productivity of firm $i$ at period $s$ - following entry in export markets at $s=0$ and the variable $START_i$ takes on the value one if the firm $i$ starts to export. The causal effect can be verified by looking at the difference: $(\omega_{is}^1 - \omega_{is}^0)$, where the superscript denotes the export behaviour. The crucial problem is that $\omega_{is}^0$ is not observable. I follow the micro-econometric evaluation literature (Heckman et al 1997) and I defined the average effect of export entry on productivity as:

$$E[\omega_{is}^1 - \omega_{is}^0 \mid START_i = 1] = E[\omega_{is}^1 \mid START_i = 1] - E[\omega_{is}^0 \mid START_i = 1]$$

\textsuperscript{12} At $t=0$ (firms: 4,507, average productivity: 0.18); $t=-1$ (firms: 3,463; average productivity: 0.25), at $t=-2$ (firms: 2,778; average productivity: 0.29), at $t=-3$ (firms: 2,250, average productivity: 0.36), at $t=-4$ (firms: 1,803; average productivity: 0.38), at $t=-5$ (firms: 1,477, average productivity: 0.44).
The key difficulty is to identify a counterfactual for the last term in Equation (6). This is the productivity effect that entrants in export markets would have experienced, on average, had they not exported. What is mainly of interest is the magnitude of the ‘impact’, labelled in red in Figure 4 and the main problem is the calculation of the counterfactual that is to be deducted from the total change.

This counterfactual is estimated by the corresponding average value of firms that remain non-exporters: $E[\omega_0^i | START_i = 0]$. An important feature of the construction of the counterfactual is the selection of a valid control group. In order to identify this group it is assumed that all the differences in productivity (except that caused by exporting) between exporters and the appropriately selected control group is captured by a vector of observables, including the pre-export productivity of a firm. The intuition behind selecting the appropriate control group is to find a group that is as close as possible to the exporting firm in terms of its predicted probability to start exporting. More formally, I apply the propensity score matching method as proposed by Rosenbaum and Rubin (1983). This boils down to estimating a probit model with a dependent variable equal to one if a firm starts exporting and zero elsewhere on lagged observables including productivity.

The probability of starting to export is modelled as follows. $START$ is a dummy variable that is one at the time a firm starts exporting. The probability of starting to export, i.e. the propensity score, can be represented as follows:

$$\Pr(START_{i,0} = 1) = F(\omega_{i,-1}, size_{i,-1}, CONTROLS_{i,-1})$$ (7)

where $F(.)$ is the normal cumulative distribution function. The re-scaling of the time periods implies that the probability of starting to export is regressed on variables prior to this period $s = 0$ and I use the subscript ‘$-1$’ to denote this. The most important variable in estimating the propensity score estimation clearly is the productivity variable. Differences in productivity will be conditioned on pre-export levels of productivity and the size of the firm. I also include a full set of industry and year dummies to control for common aggregated demand and supply shocks. I use nearest-neighbour one-to-one matching, with replacement$^{13}$.

Let the predicted export probability for firm $i$ (which is an eventual exporter) be denoted by $p_i$. The matching is based on the method of the nearest neighbour, which selects a non-exporting firm $j$ that has a propensity score $p_j$ closest to that of the export entrant. This results in a group of matched exporting and non-exporting firms needed in order to evaluate the causal impact of exporting on productivity. Following both Girma et al (2004) and Loecker (2007) I match within each 2-digit NIC sector and therefore create control groups within narrowly defined sectors as opposed to matching across the entire manufacturing sector. This is likely to be important as the marginal effect of various variables on the probability of starting to export may differ substantially between different sectors due to different technological and market conditions that firms face in different industries. This implies that I estimate the probability to start exporting for each industry separately, allowing the coefficients to vary across the various industries.

$^{13}$ Matching with replacement tends to reduce bias, and can be performed in cases where the control group is smaller than the treatment group.
However, I am unable to control for other differences between firms that produce the same product, for instance, quality, mark-ups, employee skill sets etc.

Once I have this counterfactual in hand I use a difference-in-differences (DID) methodology to assess the impact of exporting on productivity. Following Loecker (2007) the estimator of the learning-by-exporting effect ($\beta_{LBE}$) is calculated in the following way. Assume $N$ firms that started exporting and a set $C$ of control firms and $\omega^{1}$ and $\omega^{c}$ are the estimated productivity of the treated and the controls respectively. Denote $C(i)$ as the set of control units matched to a firm $i$ with a propensity score of $p_i$. The number of control firms that are matched with an observation $i$ (starter) is denoted as $N_i^c$ and the weight $w_{ij} = \frac{1}{N_i^c}$ if $j \in C(i)$ and zero otherwise. In this way every firm $i$ that started exporting is matched with $N_i^c$ control firms. I stress that the matching is always performed at the time a firm starts exporting and $s = \{1,2,.....,S\}$ denotes the time periods after the decision to start exporting, i.e. at $s = 0$. I introduce two estimators getting at the productivity effect at every time period $s$ (Equation 8) and a cumulative productivity effect (Equation 9). The estimator $\beta_{LBE}^s$ at every period $s$ after the decision to start exporting is given by:

$$\beta_{LBE}^s = \frac{1}{N_s} \sum_i \left( \omega_{is}^1 - \sum_{j \in C(i)} w_{ij} \omega_{js}^c \right)$$  \hspace{1cm} (8)$$

In words, I estimate the productivity premium of firms that started exporting at each period $s$ compared with (a weighted average of) productivity of a control group based on nearest neighbour matching at every period $s$. However, since I am also interested in how starting to export impacts the productivity trajectory of a firm, I estimate the average cumulative treatment effect. This is the productivity gain gathered over a period $S$ after the decision to start exporting. The estimator $\beta_{LBE}^S$ is given by:

$$\beta_{LBE}^S = \frac{1}{N_S} \sum_i \left( \sum_{s=0}^{S} \omega_{is}^1 - \sum_{s=0}^{S} \sum_{j \in C(i)} w_{ij} \omega_{js}^c \right)$$  \hspace{1cm} (9)$$

This provides me with an average cumulative productivity gain at every time period and plotting these estimated coefficients over time gives me a relation between time ($s$) and the productivity gain. The estimate in Equation (9) gives us the productivity premium new exporters have gathered over time. This implies that the entire productivity path of export entrants is compared to that of the control group, whereas the estimate in Equation (8) estimates the productivity premium at the each time period $s$.

Also, the cumulative estimator $\beta_{LBE}^S$ is not equal to the sum of the pure time estimator $\beta_{LBE}^s$ due to the unbalanced data. Formally, $\sum_s \beta_{LBE}^s = \beta_{LBE}^S$ since $N$ varies with $s$.

There are two reasons why firms might start drop out of the sample as the total number of years of exporting (i.e. the value of $s$) increases – less productive firms that are unable to survive exit the export market, and/or firms that began to export in recent years
drop out because the sample period ends. If the former is true, the average treatment effect will be upwardly biased. However, if the latter is true then we should not expect to see any bias. In the first five years of exporting, I lose between 7-14 percent of firms because the sample period comes to an end. The remainder of the firms exit the export market, with the largest number of firms exiting the market in the first two years of exporting. And so, the coefficients could be biased upwards since I am unable to control for endogenous exit using matching.

The propensity matching score method also assumes that there exists a region of ‘common support’, where the treated and control propensity scores overlap, and over which a robust comparison can be made. A straightforward way to check overlap is a visual analysis of the density distribution of the propensity score in both groups. Firms that fall outside of the region of common support are disregarded and for these firms the treatment effect cannot be estimated. In this case the proportion of such firms is small, and thus the estimated effect on the remaining firms can be viewed as representative.

The problem with propensity score matching techniques is that it eliminates the selection problem based only on observables, and that it assumes away every possible problem with the error terms – this could include endogeneity and/or measurement error. To deal with these shortcomings I also use instrumental variables techniques to obtain identification.

**Instrumental Variables Technique**

I also obtain identification of productivity gains by using instrumental variable techniques. A good instrument would be a variable that affects the productivity of the firm only through its effect on the firm’s exporting behaviour or its decision to export. I use effectively applied tariffs faced by exporting firms. These are tariffs weighted by exports and destination country for each industrial sector from 1990 to 2008. Since these are tariffs imposed by destination countries, individual firms in India should not be able to affect the level of tariffs. And in theory, these tariffs should have an effect on the exporting behaviour, and thus on firm-level productivity.

One might argue that increases in firm-level productivity in India could have a direct impact on tariffs faced by exporting firms if destination markets raised their tariffs to protect their domestic markets. If this were the case, one would expect to see a positive relationship between exporting and tariffs. But in fact tariffs and exports are negatively related (see Figure 6), and so it would seem more likely that lower tariffs drive higher export volumes. Alternatively one may argue that an increase in firm-level productivity leads to an increase in exports, which in turn creates pressure from exporting firms from the originating country, i.e. India, on destination markets to lower their tariffs. In other words, that exports are driving tariffs. However, India’s share of the global export market is tiny, and since independence has varied between 0.5 and 2.5 percent and between 0.5 and just over 1 percent over the period of study. When disaggregated by sector the range varies between 0.32 and 0.55 percent. When further disaggregated by the top 20 country destinations (that make up almost 75 percent of the total market share of Indian exports), the range varies between 5.72 and 0.09 percent (see Figure 9). So, a more likely story would be that Indian exporters are in fact price takers in global markets. It has been
argued that India is able to punch above its weight at the WTO, for instance through anti-dumping investigations. In fact in 2002, India overtook the US to become the highest initiator of anti-dumping cases at the WTO. However, the point of anti-dumping cases is mainly to try and protect domestic markets from cheap imports. For instance, most anti-dumping cases were brought out against China, Brazil, and Taiwan etc. It could, however, be the case that countries may still use anti-dumping as a clever negotiating tool to increase market access.

I use effectively applied tariff rates taken from the Trade Analysis and Information System (TRAINS)\textsuperscript{14}. The classification system used is SITC, as this corresponds well to the Indian NIC system. Since WITS provides tariff data for only agricultural and manufacturing industries, I am forced to drop services while carrying out the OLS and IV estimation. I re-run Equation (1) with productivity as the dependent variable and I instrument the dummy variable for exporting and also for starting to export with tariffs, both simple and weighted. I take firm fixed-effects to control for the differential effects that changes in industry-level tariffs would have on individual firms depending on their characteristics. Thus, the instrument is time varying and applies within firms.

Recall, that productivity here has been calculated after having controlled for the potential simultaneity of input choices and unobserved productivity, and so I am mainly trying to control for any reverse causality between productivity and the export (or starting to export) status of the firm. I run the regression using both simple averages and weighted averages, using two-stage least squares techniques. I check the exogeneity of the export status using the Durbin-Wu Hausman specification test, and find that the results of the IV estimates are preferable. The instrumented coefficient remains positive and significant. The F-statistic is above the rule-of-thumb value of 10 in the case of weighted tariffs, but well below for simple tariffs in the case of the export dummy. This could be because a number of countries which serve as important export destinations for Indian producers apply tariffs over and above a particular value of exports – this effect is only captured by import-weighted tariffs, and not by simple tariffs. In effect weighted tariffs are the true indicator of the barriers faced by exporters in India.

I also find that instrumenting for exports raises the coefficient marginally. My endogenous variable, exports, is weakly correlated with my instruments, whether simple or weighted tariffs (see Table 6 for first-stage results). In general, the weaker the correlation between the instrument and the variable being instrumented, the greater is the population variance of the coefficient. The increase in the instrumented coefficient could also be owing to heterogeneity in export effects, implying that the marginal return to exporting is higher for lower productivity firms, who are also more affected by changes in effectively applied tariffs. Biesebroeck (2005) finds a similar result using ethnicity of the owner as an instrument, and Card (2001) elaborates on such econometric results when using instrumental variables.

\textsuperscript{14} TRAINS data is made available through the World Integrated Trade Solution (WITS) database.


5 Robustness checks

5.1 One-Step Procedure

As a check I use the one-step procedure and include lagged exporting behaviour directly in the production function. As in the two-step procedure, I use the OP method and use investment to correct for the endogeneity of input choices with respect to productivity. It is assumed that the firm manager observes current productivity $\omega_{it}$ before making profit-maximising choices of labour and investment to be combined with capital, to produce output. To obtain coefficients on production and export experience variables, I modify the OP estimation procedure. Following Biesebroeck (2005) I introduce a dummy for lagged exports directly into the estimation of production functions. The main identifying assumption is that the firm manager takes production and export experience as state variables like capital; hence their coefficients are obtained in the same stage as that of capital. The estimation method is outlined in Appendix A.

The coefficients generated by the OP method are provided below. Column (1) provides simple OLS estimates and column (2) provides estimates using investment as a proxy to control for the simultaneity bias. OLS estimates have been generated controlling for industry and year and are expected to be biased for two main reasons – they do not control for entry and exit of firms (and thus over-estimate industry level productivity) and they do not control for the simultaneity problem, i.e. when firm’s inputs and outputs are chosen simultaneously. Thus, the labour coefficient is meant to be biased upwards and the capital coefficient biased downwards using simple OLS estimations. This result is borne out in the production function coefficients compared across OLS and OP methods.

What is of main interest here is the coefficient on the export dummy, which is a lagged dummy variable for whether the firm exported in the last period. The coefficient provides further evidence for the learning-by-exporting theory. In both cases, whether productivity is generated by simple OLS or by a modified OP procedure, exporting has a positive and significant effect on firm productivity.

5.2 Exiting export markets

I am interested in knowing how entry, continued stay and also exit impact firm-level productivity. Thus, I follow Bernard and Jensen (1999) and run the following modified regression to estimate the causal impact of export entry on productivity:

$$\frac{1}{T}(\omega_{iT} - \omega_{i0}) = \beta_0 + \beta_1 start_{iT} + \beta_2 stop_{iT} + \beta_3 continue1_{iT} + \beta_4 continue2_{iT} + controls + \epsilon_{iT}$$

(10)

where $start_{iT} = 1$ if a firm did not export at 0 but exports at $T$, $stop_{iT} = 1$ if a firm exported at 0 but no longer at $T$ and $continue1_{iT} = 1$ if a firm exports at both $T$ and 0, interacted with the number of years of exporting up to 5 years, and $continue2_{iT} = 1$ is interacted with the number of years of exporting up to 10 years. Firm-specific controls include the size (number of employees), the age of the firm and the type of firm. I also control for industry, year and location effects.
Bernard and Jensen’s estimate for productivity, however, is a residual from an OLS production function without controlling for the simultaneity bias. It is also clear however, from Equation (10) that self-selection into export markets is not controlled for, and that new exporters are not compared to similar non-exporting firms. Thus, I modify the regression and run it for my full and for my matched sample. In other words, I control for the self-selection of firms into the export market, and for the simultaneity bias for my production function estimates.

According to the results, the year of entry into the export market is associated with an increase in productivity for the firm. However, the effect disappears if the firm ‘continues’ to export. The variable ‘continue’ is interacted with the number of years of exporting in the short (5 years) and in the medium term (10 years)\textsuperscript{15}. What is interesting is that exiting the export market has a negative effect on firm productivity. It should be kept in mind that although the OP method controls for endogenous exit from the sample, it does not control for endogenous exit from export markets. And so the negative coefficient on ‘stop’ could be an overestimate since I am unable to observe firms that exit both, the export market and the sample, at the same instant. I also provide the results from the regression of current productivity on when the firm starts, continues and stops exporting. The results are broadly similar, except that continuing to export is positive and significant for the short and medium run, although the value of the coefficient is drastically reduced and is no longer significant for 15 years and over.

5.3 Export intensity

Another interesting sub-question is whether productivity is determined not just by participation in export markets, but also by the intensity of that participation. Following Castellani (2002) and Girma et al (2004) I regress firms’ TFP on previous period’s TFP, exports as a proportion of sales, age and type of firm, industry, year and location, for my complete and my matched sample of firms. I find that the value of total exports has a positive and significant effect on firms’ productivity – implying that the more that a firm exports, the higher the productivity of the firm. In this case, a percentage point increase in export volume leads to a 1.09\% rise in productivity.

5.4 Business Groups

The sample also provides me with information on whether firms belong to particular ownership groups. These are large business groups that own a number of firms, usually operating in similar industries, along the vertical or horizontal scale of production. There are a total of 583 business groups where the mean number of firms within each group is 6.5, the minimum 2 and the maximum 127\textsuperscript{16}. As an example, take the ‘Rane Group’, which has 12 firms that produce steering systems, engine valves, brake linings etc. In a given year, between 1 to 5 firms within the Rane Group export.

One might expect knowledge spillovers to be high for firms within the same business group since there would be fewer or no restrictions on technology sharing. Technological spillovers are one of the channels through which firms with access to

\textsuperscript{15} The possible range is (1,20), with firms exporting for an average of 4-5 years.

\textsuperscript{16} The business group with the most firms is the Tata Group.
foreign markets became more productive. It could be theorised that non-exporting firms might be able to access better technologies or production processes or designs if other firms within the same business group had access to foreign markets through exporting.

I regress productivity of non-exporters in a given business group on the total number of exporting firms within the same business group in that year, controlling for the location, age and size (number of employees) of firms. I also control for the NIC industry group to ensure that I measure spillovers between firms with higher technological relatedness. I find that the more the exporters within the same business group, the higher the productivity of non-exporters, although the effect is quite marginal. The effect on productivity is stronger if I use lagged values.

6 Concluding remarks

This paper analyses the effect of exporting on firm-level productivity over a period that saw a large increase in the number of exporting firms. Whilst descriptive statistics find that exporters are larger, more capital-intensive and add greater value than non-exporters, what is really interesting is the causal effect of exporting on firm performance. After calculating unbiased production function estimates, I then go on to identify the effect of entering export markets after controlling for the self-selection of more productive firms into such markets.

I use propensity-matching techniques and instrumental variable techniques to disentangle the effect of exporting on firm-level productivity. Using the first technique, I construct a set of ‘control firms’ and then evaluate the effect of the ‘treatment’, i.e. exporting. I find that exporting does indeed lead to a positive and significant effect on the productivity of firms that begin to export, and that this effect remains persistent and even grows as the firm continues to export for the first few years. When I include the decision to export within the production function of the firm, it continues to positively effect productivity. I then use effectively applied tariffs faced by Indian exporters in world markets to control for the reverse causality between exporting and productivity. I find, again, that the productivity advantage for exporters continues to persist after controlling for the self-selection bias. There is also evidence that gains from exporting are the highest in the first few years of exporting and then begin to taper off. These results are robust to different identification procedures and I also find evidence that productivity gains are reversed when the firm decides to exit the export market.

My results are very much in line with the previous empirical literature. Loecker (2007) uses propensity-score matching techniques to identify the effect of exporting on firms in Slovenia and finds that the annual productivity premium from exporting varies between 8 to 13 per cent. Using similar techniques, this paper finds that average productivity premium for firms in India varies between 9 to 13 per cent. Biesenbroeck (2005) incorporates lagged exports within the production function and finds robust evidence for a positive effect of exporting on productivity. He also uses the ethnicity of the firm owner as an instrument for the decision of firms to export, and again finds evidence for a causal and positive impact. Using different techniques, his estimate of the effect of exporting on productivity varies between 25 to 28 per cent for firms in Africa.
Using effectively applied tariffs as an instrument, my estimate of the effect of exporting on productivity is 29 per cent, whilst the effect of starting to export lies between 11 to 12 per cent.

This paper is an attempt to understand the productivity premium from exporting for a poor, rapidly developing country. There has been a dearth of studies for developing countries (with the exception of the papers mentioned above), where one would expect the gains to be the highest. Evidence on the determinants and computation of firm-level productivity in low-income countries is also rare. I use a double-fold identification test, and I have used an instrument, i.e. effectively applied tariffs, that has previously not been used in the literature to control for the self-selection effect. The use of tariffs nicely isolates the effect of exporting on productivity, and since these tariffs are hardly unique to the case of India, they could easily be applied in other settings.

And finally, why should such empirical work matter in a broader policy setting? A number of developing countries have spent time and resources in devising ways to encourage their firms to enter export markets, for instance through tax incentives, establishing export processing zones etc. While the underlying idea is that firms learn through producing larger and larger volumes, it is important to disentangle the effect of learning-by-exporting. Empirical evidence for a causal effect of exporting on productivity indicates that it may make sense to provide incentives to firms to enter export markets. However, once these firms begin to export the real productivity premium will depend on whether they are able to survive in these markets on their own. The most important takeaway from this paper is that participation in export markets, global export markets, makes firms more productive. However, there is research to show that it matters whom one exports to and that exporting to larger and more developed markets is associated with greater productivity gains. It also seems intuitive that if exports increase primarily because of preferential access to markets that are insulated from global competition, one might expect the productivity premiums from exporting to be lower.
**Table 1: Summary of empirical findings on exports and productivity**

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Sample</th>
<th>Methodology</th>
<th>Evidence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aw and Hwang (1995)</td>
<td>Taiwan</td>
<td>2,832 firms; 1986</td>
<td>Translog production function, cross-section</td>
<td>√</td>
</tr>
<tr>
<td>Bernard and Wagner (1997)</td>
<td>Germany</td>
<td>7,624 firms; 1978-92</td>
<td>Panel Data</td>
<td>√</td>
</tr>
<tr>
<td>Bernard and Jensen (1999)</td>
<td>US</td>
<td>60,000 plants; 1984-92</td>
<td>Linear probability with fixed effects</td>
<td>√</td>
</tr>
<tr>
<td>Wagner (2002)</td>
<td>Germany</td>
<td>353 firms; 1978-89</td>
<td>Panel data; Matching</td>
<td>√</td>
</tr>
<tr>
<td>Alvarez and Lopez (2005)</td>
<td>Chile</td>
<td>5,000 plants; 1990-96</td>
<td>Ordered probit; pooled data</td>
<td>√</td>
</tr>
</tbody>
</table>

Van Bisebroeck (2005) Sub-Saharan Africa 1916 firms (9 countries); 1992-1996 GMM, Maximum likelihood, Olley and Pakes production function; Cross-section √ √


*\( \omega^e > \omega^ne \): Exporters more productive than non-exporters; LBE: Learning-by-exporting effects. 1: Some learning from exporting in the case of Morocco. 2: Learning associated with export intensity. Source: Girma et al (2004), modified and updated

### Table 2: Data Summary

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firms (#)</td>
<td>15,292</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations (#)</td>
<td>106,712</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>1,092</td>
<td>0.05</td>
<td>842,770</td>
</tr>
<tr>
<td>Gross Assets</td>
<td>1,827</td>
<td>0.08</td>
<td>1,352,683</td>
</tr>
<tr>
<td>Investment</td>
<td>22.3</td>
<td>-154.69</td>
<td>14,239</td>
</tr>
<tr>
<td>Wage Bill</td>
<td>89.3</td>
<td>0</td>
<td>56,567</td>
</tr>
<tr>
<td>Raw Materials Bill</td>
<td>514.2</td>
<td>0</td>
<td>269,567</td>
</tr>
<tr>
<td>Electricity Bills</td>
<td>48.5</td>
<td>0.06</td>
<td>14,988</td>
</tr>
<tr>
<td>Age</td>
<td>17.4</td>
<td>0.04</td>
<td>174</td>
</tr>
</tbody>
</table>

Note: Sales, assets, wages, raw materials and electricity are reported in INR '000s.

### Table 3: Exporters differ from non-exporters

<table>
<thead>
<tr>
<th>Firm Characteristic (x)</th>
<th>( \beta )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average wage bill</td>
<td>1.43***</td>
<td>.54</td>
</tr>
<tr>
<td>Gross Assets</td>
<td>1.19***</td>
<td>.43</td>
</tr>
<tr>
<td>Gross value added</td>
<td>1.43***</td>
<td>.45</td>
</tr>
<tr>
<td>Sales</td>
<td>1.67***</td>
<td>.45</td>
</tr>
<tr>
<td>Investment</td>
<td>1.07***</td>
<td>.25</td>
</tr>
<tr>
<td>Observations (min/max)</td>
<td>46,426/57,609</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Estimated learning by exporting effects

<table>
<thead>
<tr>
<th>s</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome: productivity</td>
<td>( \beta_{LBE} )</td>
<td>.090***</td>
<td>.138***</td>
<td>.138***</td>
<td>.112***</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>(.019)</td>
<td>(.021)</td>
<td>(.023)</td>
<td>(.024)</td>
</tr>
<tr>
<td># treated</td>
<td>6,131</td>
<td>4,956</td>
<td>4,355</td>
<td>3,757</td>
<td>3,267</td>
</tr>
<tr>
<td># controls</td>
<td>5,641</td>
<td>4,690</td>
<td>4,171</td>
<td>3,625</td>
<td>3,156</td>
</tr>
<tr>
<td>Outcome: cumulative productivity</td>
<td>( \beta_{LBE} )</td>
<td>.090***</td>
<td>.121***</td>
<td>.142***</td>
<td>.149***</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>(.019)</td>
<td>(.014)</td>
<td>(.013)</td>
<td>(.023)</td>
</tr>
<tr>
<td># treated</td>
<td>6,131</td>
<td>11,087</td>
<td>15,442</td>
<td>19,199</td>
<td>22,466</td>
</tr>
<tr>
<td># controls</td>
<td>5,641</td>
<td>10,331</td>
<td>14,502</td>
<td>18,127</td>
<td>21,283</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001
Standard errors in parenthesis

Table 5: IV Estimation

<table>
<thead>
<tr>
<th>Model Predictors</th>
<th>OLS</th>
<th>Simple</th>
<th>IV</th>
<th>Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Export dummy</td>
<td>.291***</td>
<td>.292***</td>
<td>.292***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.017)</td>
<td>(.017)</td>
<td></td>
</tr>
<tr>
<td>Start dummy</td>
<td>.123***</td>
<td>.112***</td>
<td>.122***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.012)</td>
<td>(.013)</td>
<td></td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>5.28</td>
<td>247</td>
<td>10.50</td>
<td>25</td>
</tr>
<tr>
<td>#</td>
<td>61,191</td>
<td>61,191</td>
<td>61,191</td>
<td>61,191</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001
Standard errors in parenthesis (clustered at the firm level

The residuals could be correlated across firms or time. For instance, economy-level shocks would cause correlation between firms at a moment in time, and persistent firm-specific shocks could cause correlation across time. Thus, I also cluster standard errors by year, but this does not affect the results.
Table 6: First-Stage Results

<table>
<thead>
<tr>
<th></th>
<th>Export Dummy</th>
<th></th>
<th>Start Dummy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Simple tariffs</td>
<td>-.0007***</td>
<td></td>
<td>.003*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0000)</td>
<td></td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>Weighted tariffs</td>
<td>-.001***</td>
<td></td>
<td>.001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.562***</td>
<td>.563***</td>
<td>-.003</td>
<td>.028***</td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.009)</td>
<td>(.006)</td>
<td>(.005)</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>#</td>
<td>61,191</td>
<td>61,191</td>
<td>61,191</td>
<td>61,191</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001
Standard errors in parenthesis

Table 7: Production Function Coefficients

<table>
<thead>
<tr>
<th>Variables</th>
<th>OLS</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Capital</td>
<td>.506***</td>
<td>1.143***</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.457)</td>
</tr>
<tr>
<td>Labour</td>
<td>1.927***</td>
<td>.955</td>
</tr>
<tr>
<td></td>
<td>(.050)</td>
<td>(1.679)</td>
</tr>
<tr>
<td>Export dummy</td>
<td>660.59***</td>
<td>189.13*</td>
</tr>
<tr>
<td></td>
<td>(90.46)</td>
<td>(120.33)</td>
</tr>
</tbody>
</table>

# 102,209         101,702

* p<0.05, ** p<0.01, *** p<0.001
Bootstrap Standard errors in parenthesis
### Table 8: Exiting export markets

<table>
<thead>
<tr>
<th>Variables</th>
<th>Δ Productivity</th>
<th>Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All firms</td>
<td>Matched firms</td>
</tr>
<tr>
<td>Start</td>
<td>.106***</td>
<td>.095***</td>
</tr>
<tr>
<td></td>
<td>(.012)</td>
<td>(.011)</td>
</tr>
<tr>
<td>Continue1</td>
<td>.003</td>
<td>.003</td>
</tr>
<tr>
<td>(5 years)</td>
<td>(.002)</td>
<td>(.001)</td>
</tr>
<tr>
<td>Continue2</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>(10 years)</td>
<td>(.001)</td>
<td>(.001)</td>
</tr>
<tr>
<td>Stop</td>
<td>-.215***</td>
<td>-.153***</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.013)</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

R² | 0.01 | 0.01 | 0.25 | 0.25

# | 52,229 | 49,556 | 57,609 | 49,556

* p<0.05, ** p<0.01, *** p<0.001

Standard errors in parenthesis

### Table 9: Exporting, export intensity and TFP

<table>
<thead>
<tr>
<th>Variables</th>
<th>Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
</tr>
<tr>
<td>Log (ωi−1)</td>
<td>.819***</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
</tr>
<tr>
<td>Log (Exports)</td>
<td>.0107***</td>
</tr>
<tr>
<td></td>
<td>(.0009)</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
</tr>
<tr>
<td>State FE</td>
<td>Yes</td>
</tr>
</tbody>
</table>

R² | 0.78 | 0.78

# | 28,981 | 28,914

p<0.05, ** p<0.01, *** p<0.001

Standard errors in parenthesis
### Table 10: Business Groups

<table>
<thead>
<tr>
<th>Variables</th>
<th>Productivity (1)</th>
<th>Productivity (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Exporters (sum)</td>
<td>.0001***</td>
<td>(.0000)</td>
</tr>
<tr>
<td>Exporters (sum)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lagged</td>
<td>.0002***</td>
<td>(.0000)</td>
</tr>
<tr>
<td>Age</td>
<td>-.0007***</td>
<td>-.0008***</td>
</tr>
<tr>
<td></td>
<td>(.0004)</td>
<td>(.0005)</td>
</tr>
<tr>
<td>Size</td>
<td>.0000***</td>
<td>.0000***</td>
</tr>
<tr>
<td></td>
<td>(.0000)</td>
<td>(.0000)</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>#</td>
<td>25,606</td>
<td>22,200</td>
</tr>
</tbody>
</table>

p<0.05, ** p<0.01, *** p<0.001
Standard errors in parenthesis

### Figures

**Figure 1: Number of exporters**

Source: Based on Prowess, CMIE (Centre for Monitoring of the Indian Economy)
Figure 2: Technology Achievement Index

Source: Based on data from Global Economic Prospects (World Bank 2008)

Figure 3: TFP and entry into export markets
Figure 4: Identification of the true impact

![Graph showing identification of true impact with baseline and evaluation periods, and exports vs. doesn't export categories.]

Figure 5: Region of Common Support

![Graph showing density distributions for control and treated groups, indicating common support region.]

Baseline | Evaluation
--- | ---

Impact | Impact
--- | ---

Counterfactual | Change
--- | ---

Exports | Doesn't export
--- | ---

Density

Control | Treated
--- | ---
Figure 6: Tariffs and Exports

Source: World Integrated Trade Solution (WITS) and Trains

Figure 7: Indian Export share

Source: World Integrated Trade Solution (WITS) and Comtrade
Figure 8: Range of export shares by industries

Source: World Integrated Trade Solution (WITS) and Comtrade

Figure 9: Range of export shares by top 20 export markets and industries

Source: World Integrated Trade Solution (WITS) and Comtrade
Appendix A

A.1 Estimating Productivity

Following Olley and Pakes (1996) it is assumed that in year $t$ the manager observes the firm’s current productivity $\omega_t$ before choosing labour $l_t$ and investment $i_t$ to combine with the quasi-fixed input, capital $k_t$ for the production of output $y_t$. Output is expressed as follows:

\[ y_t = \beta_0 + \beta_1 l_t + \beta_2 i_t + \beta_3 k_t + \omega_t + \epsilon_t \quad (A1) \]

Inputs are divided into a freely variable ones ($l_t$, $i_t$) and the state variable capital ($k_t$). The error term is assumed to be additively separable in a transmitted component ($\omega_t$) and an i.i.d. component ($\epsilon_t$). The key difference between the former and the latter is that the former is a state variable and hence impacts the firm’s decision rules, while the latter has no impact on the firm’s decisions.

Since $\omega_t$ is known to the manager but unknown to the econometrician and may be positively correlated with $l_t$ and $i_t$, it generates a potential simultaneity bias that is addressed by the following estimation procedure. The firm’s variable input demands, derived from profit maximisation, depend on privately known productivity and capital. Investment’s demand function is given by:

\[ i_t = i(\omega_t, k_t) \]

and it must be monotonic in all $\omega_t$ for all relevant $k_t$ to qualify as a valid proxy – implying that conditional on capital the demand for investment increases with productivity. Assuming that monotonicity holds, the input demand function can be inverted to obtain $\omega_t$ as a function of investment and capital, as below. Note that this function depends on observables only.

\[ \omega_t = \omega(i_t, k_t) \]

Consider the problem of self-selection. Firms with larger capital stocks can expect higher returns on capital even in the face of lower levels of productivity, and will choose to stay longer in the market. Thus the self-selection generated by the exit behaviour implies that the expectation of productivity will be decreasing in capital, leading to a negative bias in the capital coefficient.

The first stage of the estimation proceeds by rewriting Equation (A1) in a partially linear form:

\[ y_t = \beta_1 l_t + \phi(i_t, k_t) + \epsilon_t \quad (A2) \]
where,
\[ \phi(i, k, t) = \beta_0 + \beta_i i_t + \beta_k k_t + \omega(i, k, t) \]  \hspace{1cm} (A3)

Since \( E[\varepsilon_i | i, k, t] = 0 \), taking the difference between Equation (A2) and its expectation conditional on investment and capital generates the following expression:
\[ y_t - E[y_t | i, k, t] = \beta_t (I_t - E[I_t | i, k, t]) + \varepsilon_t \]  \hspace{1cm} (A4)

Equation (A4) is estimated by OLS (no constant) to obtain consistent parameter estimates for labour. The conditional expectations in Equation (A4) are the intercepts of locally weighted least squares (LWLS) regressions of output and labour on \((i, k, t)\). After obtaining estimates for \( \beta_t \), we estimate the function \( \phi(.) \) as a LWLS regression of \( y_t - \hat{\beta}_t I_t \) on \((i, k, t)\). If one were only concerned with the marginal productivities of the variable inputs (but not the co-efficient on the proxy variable) one could stop here. To obtain a capital co-efficient, a plant-level measure of productivity a more complete model for \( \phi(.) \) will be required since capital enters it twice. To estimate \( \beta_k \), in addition to the estimates of \( \beta_t \) obtained from the partially linear model, estimates of the survival probabilities are also used. These probabilities are given by:
\[ \Pr\{\chi_{t+1} = 1 | \omega_{t+1}(k, t), J_t\} = P_t \]  \hspace{1cm} (A5)

where \( \chi \) is defined as the indicator function and is equal to zero if the firm exits the market, and \( J_t \) refers to the information available at time \( t \). In the implementation the probability of survival is estimated by fitting a probit model of \( \chi_{t+1} \) on the state and proxy variables, as well as their squares and cross products.

In the next stage the expectation of \( y_{t+1} - \beta_I I_{t+1} \) conditional on information at time \( t \) and survival is given as:
\[ y_{t+1} - \beta_I I_{t+1} = \beta_k k_{t+1} + g(P_t, \phi_I - \beta_k k_t) + \xi_{t+1} + \varepsilon_{t+1} \]  \hspace{1cm} (A6)

where, \( \xi_{t+1} = \omega_{t+1} - E[\omega_{t+1} | \omega_t, \chi_{t+1} = 1] \), and is the unexpected productivity shock and is independent and identically distributed (i.i.d.). The unknown function \( g(.) \) is approximated by a second-order polynomial in \( \phi_I - \beta_k k_t \) and \( P_t \). The estimates of \( \beta_t \), \( \phi_t \) and \( P_t \) are substituted in (A6) for the true values of \( \beta_t, \phi_t \) and \( P_t \), to then obtain estimates of \( \beta_k \) by minimising the sum of squared residuals in equation (A6). Since the estimation routine involves three steps, the stat command implemented uses the clustered bootstrap errors, treating all observations for a single firm as one cluster.

Following Loecker (2007) with the coefficients of the production function in hand, I then recover a productivity measure for the firm \( i \) in industry \( j \) at time \( t \):

---

18 Readers interested in the details of the estimation system can refer to Yasar et al (2008).
19 It should be noted here that this measure of productivity if not the true unobserved productivity shock. It also includes the i.i.d. component which is assumed to be zero on average.
\[ \omega_{ijt} = y_{ijt} - \beta_{ijl_{ijt}} - \beta_{ijk_{ijt}} \]

## A.2 Introducing Export

In the one-step approach, lagged export status (or export experience, i.e. cumulative years of exporting) is directly introduced in the production function:

\[ y_t = \beta_0 + \beta_{lj} l_t + \beta_{ji} i_t + \beta_{ki} k_t + \beta_{EX} EX_{t-1} + \omega_t + \epsilon_t \]  

(A7)

The first stage of the estimation is very close to that described above for Equation (A1). The main difference is that the productivity function resulting from the inversion of the investment demand function depends on additional state variables, the export dummy variable: \( \omega_t = \omega(i_{it}, k_{it}, EX_{it-1}) \). The first stage of the estimation algorithm is in fact almost identical to introducing lagged export status as an input, however it is now also interacted with all terms of the polynomial in capital and investment as given in the equation below:

\[ y_t = \beta_{ji} i_t + \phi_t(i_t, k_t, EX_{t-1}) + \epsilon_t \]  

(A8)

where

\[ \phi_t(i_t, k_t, EX_{t-1}) = \beta_0 + \beta_{k} k_t + \beta_{i} i_t + \omega_t(i_t, k_t, EX_{t-1}) \]

In the second estimation step, the probability that a firm exits the sample is captured by the probability that the end-of-period productivity falls below the exit threshold: \( \text{Prob (exit after period } t) = \text{Prob } [\omega_{t+1} \leq \omega_{t+1}(K_{t+1}, EX_{t+1})]_{\omega_{t+1}, \omega_t} \). This can be written as an unknown function of current observables by substituting the transition equations for the state variables and the previously used expression for productivity:

\[ \text{Prob (exit)} = P_t(K_{t+1}, EX_t, \omega_t) = P'_t(K_{t+1}, i_t, EX_{t-1}, \Delta EX_t), \text{ or equivalently, } \\
P'_t(K_{t+1}, i_t, EX_{t-1}, EX_t) \]

Both the lagged and the current export status are needed because next period’s productivity depends on current productivity and lagged export status belongs in the equation that predicts the unobservable \( \omega_t \). Current export status belongs in the exit-threshold, because it moves the production function, hence also the profit function, for the next period. The production function coefficients of both state variables, capital and lagged export status, are recovered in the last estimation step.

The equation for the last stage is:

\[ y_{t-1} - \beta_{lk_{t+1}} = \beta_{k} k_{t+1} + \beta_{EX} EX_t + g(P_t, \phi_t - \beta_{EX} EX_{t-1} - \beta_{k} k_t) + \xi_{t+1} + \epsilon_{t+1} \]  

(A9)

The polynomial in the three variables will improve the estimation in the last stage when identifying the capital coefficient. When introducing the lagged export status dummy as an input in the production process, one has to identify the coefficient on the lagged export status in this stage as well. This implies that one has to assume that export
status only affects the average future of the productivity distribution and hence leaves no scope for learning by exporting to be a heterogeneous process across firms. In addition, it also implies that the effect is time-invariant or that every year exporting raises output (conditioned on capital and labour) by the coefficient estimated on the export dummy.

It is clear that the estimation algorithm that controls for export status has an impact on the estimated production function coefficients. Compared to the standard Olley and Pakes (1996) approach, it is expected that the labour coefficient will be lower since export status is strongly correlated with the productivity shock. In addition to investment and capital, export proxies for productivity shocks that are unobserved. The identifying assumption to estimate the capital co-efficient in the standard OP method is that any shock in productivity between period $t$ and $t+1$ is uncorrelated with the capital stock at $t+1$. If export status is not controlled for, part of the unobserved productivity shock (at time $t$) correlated with the export status end up in the error term.

References


