Exporting/Importing and Firm Performance: Evidence from India

Arjun Gupta (2014592/823125)

January 4, 2020

Abstract

This research thesis provides empirical evidence whether we observe learning-by-doing and self-selection hypothesis amongst Indian manufacturing firms. The learning-by-doing hypothesis is shown by endogenously accommodating the the export/import decision in the Levinsohn-Petrin productivity estimation procedure. The self-selection phenomenon is checked with the help of a dynamic random effects probit model. The complementarity between exporting and importing is shown with the help of a dynamic bivariate probit model.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Econometrics and Mathematical Economics

Tilburg School of Economics and Management
Tilburg University

Contents

1 Introduction

There is vast empirical literature that has documented that exporters tend to outperform non-exporters in terms of wages, capital, productivity. Bernard and Jensen (1999) suggest that this can be due to the following phenomenon:

- Higher productivity leads to export (Self-selection)
- Export market participation increases productivity (Learning by doing)

Self-selection (SS) hypothesis states that more productive firms self-select into export as participation in the trade market is accompanied by additional costs such as transport costs, establishing a distribution channel, cost of traversing bureaucratic channels etc. This means that there are substantial sunk costs to participating in the trade market. Therefore, firms which are more productive enter the export market.

Learning-by-doing hypothesis for exporting (Haidar (2012)) states that exporting firms deal with tougher competition in the international market as compared to the domestic market, and therefore must improve their performance to remain active in the export market. Moreover, participating in the international market leads knowledge flows from international buyers to help post entry performance of export starters. This means that exporting should cause productivity spillovers as well.

However, most of the research in this field has been limited to exports and research on imports has been relatively low. The same hypothesis (self-selection and learning-by-doing) can apply to import behavior of firms. Since importing intermediaries also involves additional similar costs like taxes, transport costs, import duties etc., firms that are also more productive will *Self-Select* into entering the import market. Also, since a firm participating in the import market can have access to higher quality of intermediary goods, it might lead to lead to productivity benefits i.e the firms might exhibit *Learning-by-doing* phenomenon. Topalova and Khandelwal (2011) and Halpern, Szeidl, et al. (2011) find that improved access to foreign technology can boost productivity.

Since participating in one activity might increase the chances of participating in other as it might open up new information channels or because it fosters an increase in productivity and innovation. The former would decrease the cost of the the other activity. This means that there might be cost complementarity between exporting and importing.

India provides an interesting case as it liberalised its economy in 1992 which resulted in decrease import/export tariffs, deregulation of markets, reduction of taxes, and greater foreign investment. According to Topalova and Khandelwal (2011), *the government's trade policy under the Eighth Five-Year Plan (1992-97) ushered in radical changes to the trade regime by sharply reducing the role of the import and* *export control system. The share of products subject to quantitative restrictions decreased from 87 percent in 198788 to 45 percent in 1994-95, and the actual user condition on imports was discontinued. And since 1997, the decrease in output and input tariff has been very marginal.*. Therefore, it is important to find evidence of productivity gains from participation in the trade a market and how participation in one activity affects the participation in the other activity. It is also important to check whether participation in one trade activity complements participation in the other activity. This will help us in knowing the effect of one aspect of the trade market on the other.

So, my reasearch plan is to investigate:

- Self-selection hypothesis: Check whether more productive firms participate in the expprt/import market
- Learning-by-doing hypothesis: Check if there are productivity spillovers from participation in the export/import market
- Complementarity between exporting and importing: Check if there is complementarity between exporting and importing

2 Literature Review

Most papers on this can be differentiated into three different categories:

1. Importing and Productivity
2. Exporting and Producitivity
3. Complementarity between exporting and importing and its effect on productivity

In this section, I write down the major literature contributions towards my topic.

2.1 Importing and Productivity

If a firm resorts to importing inputs, then it can have access to higher quality of intermediate goods and might pave the way for future exporting by increasing the productivity or by providing cost-complementarity between the two activities.

Halpern, Szeidl, et al. (2011) studies effect of imports on productivity by estimating a structural model of importers in a panel of Hungarian firms. They find that imports have a significant and large effect on firm productivity, about one-half of which is due to imperfect substitution between foreign and domestic goods.

Topalova and Khandelwal (2011) find that the pro-competitive effects of the tariffs led firms to become more efficient, the larger impact appears to have come from increased access to foreign inputs.

2.2 Exporting and Productivity

Most studies find that there is self-selection of more productive firms into exporting, however there is mixed evidence for learning-by-exporting.

Aw, Mark J. Roberts, and Xu (2011) estimate a dynamic structural model that captures both the behavioral and technological linkages among R&D, exporting, and productivity. They find that Relative to a plant that does neither activity, export market participation raises future productivity by 1.96 percent, R&D investment raises it by 4.79 percent, and undertaking both activities raises it by 5.56 percent.

Bernard and Jensen (1999) test the self-selection and learning-by-doing hypothesis of exports on firms. They find that Good plants become exporters i.e. learn to export. and find that exporting increases the survival probability but it does not contribute towards productivity growth.

Mark J Roberts and Tybout (1997) quantify the effect of prior exporting decision on the current exporting decision and test the sunk cost hypothesis of these activities. They develop a dynamic discrete- choice model of exporting behavior that separates the roles of profit heterogeneity and sunk entry costs in explaining plants' exporting status and find that sunk costs are significant as prior export experience increases the probability of exporting by 60 percent.

In terms of work in this field with Indian firm level data, Haidar (2012) and Gupta, Patnaik, and Shah (2018) find evidence of self-selection of more productive firms into exporting, but they do not find evidence of learning-by-exporting.

2.3 Cost Complementarity between Exporting and Importing

As far as I am aware, there have been two major papers in this field i.e Aristei, Castellani, and Franco (2013) and Kasahara and Lapham (2013).

Aristei, Castellani, and Franco (2013) estimate a bivariate probit model to understand the two-way relationship between exporting and importing. Thy suse firm-level data for a group of 27 Eastern European and Central Asian countries from the World Bank Business Environment and Enterprise Performance Survey (BEEPS) over the period 20022008, and after controlling for size (and other firm-level characteristics), find that firms exporting activity does not increase the probability of importing, while the latter has a positive effect on foreign sales.

Kasahara and Lapham (2013) estimate a stochastic model of exporting and importing that incorporates heterogeneity in transport costs and estimate export and import complimentarities between the two trade activities. They find that policies which inhibit the importation of foreign intermediates can have a large adverse effect on the exportation of final goods.

Another research paper in this field, Muûls and Pisu (2009) find that productivity advantage of exports towards non-exporters may be overstated in the current literature, because imports were not taken into account as well as exports.

2.4 Productivity Estimation

Productivity estimation in all of the research papers mentioned is done using the methods highlighted below: Olley and Pakes (1992),Levinsohn and Petrin (2003), Ackerberg, Caves, and Frazer (2006) and Wooldridge (2009). These papers take inspiration from one another and the difference in estimation mentioned in these papers is very minimal. The method mentioned in Olley and Pakes (1992) have been explained in section 7.2

With regards to productivity estimation, De Loecker (2013) highlights the importance of endogenizing the export decision in the minimisation problem of the productivity estimation.

3 Data

I use annual firm level data from Centre for Monitoring Indian Economy (CMIE) which provides data from 1989 to 2017. Table 1 displays the variables I fetch from the database and their meanings.

I chose the variables which might be the most pertinent to my research question. The firm variables stated above are nominal values of Indian Rupees (Miilion). I fetch Wholseale Price Index (WPI), which provides the inflation rate of the wholesale prices and deflate the variables to give real values. Then, I clean the data to remove missing values and select firms with the broad industry classification code indicating that they are a manufacturing firm. Table 2 in the Appendix shows the composition of firms by year.

I restrict the time period of the study from 2001 to 2016 as those periods have relatively high number of firms. Since, firms are under no legal obligation to report their finances, this might mean that small firms are less likely to report their finances. Therefore, I do not analyse the effect of export/import on the survival

Table 1 Data Variables

Variable	Indicator
sa_company_name	Prowess company name
sa_finance1_year	Year
sa_sales	Sales
sa_rawmat_exp	Raw material expenses
sa_power_and_fuel_exp	Power & fuel
sa_salaries	Salaries & wages
sa_pat	Profit after tax
sa_gross_fixed_assets	Gross fixed assets
sa_export_goods	Export of goods(fob)
sa_export_serv	Export of services
sa_import_rawmat	Import of raw materials
sa_import_stores_spares	Import of stores and spares
sa_import_fg	Import of finished goods
nic.2digit	Broad industry classification code

probability of the firm and do not consider the entry/exit decision of a firm. However, this dataset includes all publicly listed firms as their firm financials are public information. This might affect the results as the dataset is biased towards bigger firms.

I create two additional variables *Export Value*, *Import Value*, *Domestic Sales* by adding the following variables from the Table 1.

1. Export Value: Sum of the exports of goods and exports of services $sa\_export\_goods + sa\_export\_serv$
2. Import Value: Sum of imports of raw materials, stores and spares, finished goods and capital goods $sa\_import\_rawmat + sa\_import\_stores\_spares + sa\_import\_fg + sa\_import\_cap\_goods$
3. Domestic Sales: Total Sales - Export Sales $sa\_sales - Export\ Value$

Then, I create four dummy variables of trade market participation:

- Export Dummy: $d_{it}^{X} = 1$ if $Export\ Value > 0$
- Import Dummy: $d_{it}^{X} = 1$ if $Import\ Value > 0$

This gives four categories on the basis of which we can divide the firms in the data set:

- None: Firms that do not participate in the export/import market

Table 2 Compostion of firms by year

Year	Number of firms
1988	23
1989	60
1990	44
1991	56
1992	74
1993	148
1994	437
1995	562
1996	385
1997	388
1998	818
1999	1130
2000	3346
2001	3847
2002	3889
2003	4447
2004	4675
2005	5151
2006	5344
2007	5456
2008	5638
2009	5903
2010	5917
2011	5330
2012	4737
2013	4320
2014	4050
2015	3087
2016	2477
2017	3

Table 3 Composition of firms based on trade market participation

Year	None	Export only	Import only	Both	Total
2001	0.28	0.08	0.18	0.46	3847
2002	0.29	0.08	0.18	0.46	3889
2003	0.30	0.08	0.18	0.44	4447
2004	0.32	0.09	0.16	0.43	4675
2005	0.35	0.08	0.16	0.41	5151
2006	0.34	0.08	0.16	0.42	5344
2007	0.35	0.08	0.16	0.42	5456
2008	0.34	0.08	0.16	0.42	5638
2009	0.35	0.08	0.15	0.42	5903
2010	0.36	0.08	0.16	0.40	5917
2011	0.37	0.08	0.15	0.41	5330
2012	0.35	0.07	0.14	0.44	4737
2013	0.33	0.08	0.14	0.45	4320
2014	0.31	0.08	0.14	0.47	4050
2015	0.24	0.08	0.14	0.54	3087
2016	0.20	0.07	0.14	0.58	2477

- Export only: Firms that participate in the export market only
- Import only: Firms that participate in the import market only
- Both: Firms that participate in both export/import market

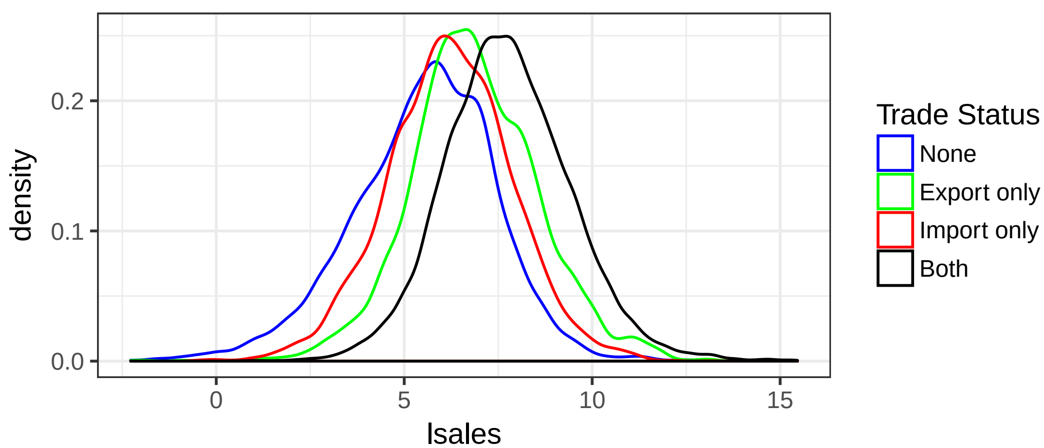
Table 3 displays the composition of firms according to their trade market participation status. It is seen that number of firms that do not participate in the trade market is low, around 20 to 35 %. Surprisingly, the number of firms that participate in the trade market is really high. Another interesting feature is that the number of firms that participate only in the import market is higher than the firms that participate only in the export market. This must mean that the demand for foreign intermediaries is really high. Also, almost 50 % of firms in each year participate in both export/import market. This might mean that there is a very high proportion of firms that participate in the both the trade market activities relative to the firms that participate only in one trade market activity.

4 Descriptive Statistics

4.1 Exporter/Importer Performance

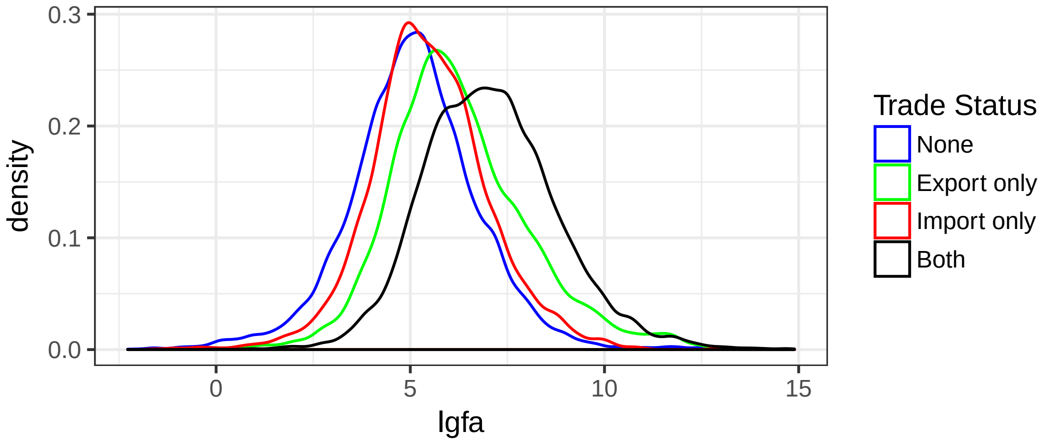
As a first step to see that firms that participate in the trade market are better performers than firms that do not participate in the trade market, I calculate the mean and standard deviation and create the density plots for log of sales, gross fixed assets, salaries and expenditure on power and fuel for firms with different trade activity status. Tables 4, 5, 6 and 22 display the results for the variables mentioned above. It can be seen that firms that participate in the trade market have a distribution that is more skewed towards the right for all the variables mentioned above. In the case of sales, firms that do not participate in the trade market have a mean of 5.52 whereas firms that participate in either export or import have a mean value of 6.86 and 6.82 respectively. Also, firms that participate in both export/import have higher mean of 7.74. It is also seen that firms that participate in the both export/import market have higher mean gross fixed assets, salaries, expenditure on power and fuel than firms that participate in only export and only import. On the other hand, the standard deviation in all the cases is very similar. This indicates that firms participating in the trade market has an positive effect on the observable characteristics of the firm.

Table 4 Summary statistics of Sales (log)



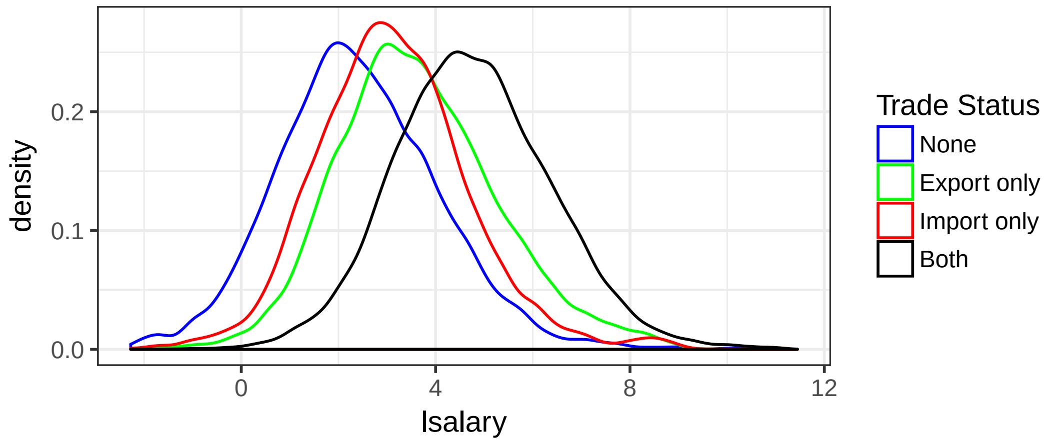
Status	mean	sd
None	5.52	1.91
Import only	6.86	1.72
Export Only	6.2	1.66
Both	7.74	1.69

Table 5 Summary statistics of Gross Fixed Assets (log)



Status	mean	sd
None	5.09	1.68
Import only	6.26	1.8
Export Only	5.49	1.53
Both	7.06	1.73

Table 6 Summary statistics of Salary (log)



Status	mean	sd
None	2.41	1.67
Import only	3.63	1.66
Export Only	3.11	1.56
Both	4.76	1.63

4.2 Trade Premia

Trade premia is defined as the difference in attributes of firms based on their trade participation status. I estimate the trade premia using the following Fixed Effect (FE) regression:

$$X_{it} = \alpha + \beta_1 d_{it}^X + \beta_2 d_{it}^M + \beta_3 d_{it}^X * d_{it}^M + \beta_4 age_{it} + \epsilon_{it}$$

where X_{it} is firm level characteristics such as Sales, Gross Fixed Assets, Expenditure on Raw Material and Salary, d_{it}^X is the export dummy,d_{it}^M is the import dummy, the interaction term between these two variables and age_{it} is the age of the firm. I estimate this equation using fixed-effects regression controlling time fixed effects. Table 7 displays the results for the above regression.

It is seen coefficients for both export and import dummy are positive and significant at 1% significance level. This means that firms that participate in the export/import market have more capital and assets and spend more on salary and

Table 7 Export and Import Premia

	Dependent variable:			
	Sales	Gross Fixed Assets	Raw Materials	Salary
	(1)	(2)	(3)	(4)
d_{it}^X	0.422***	0.209***	0.422***	0.311***
	(0.014)	(0.011)	(0.016)	(0.012)
d_{it}^M	0.430***	0.188***	0.451***	0.287***
	(0.012)	(0.009)	(0.013)	(0.010)
Age_{it}	0.039***	0.048***	0.041***	0.062***
	(0.001)	(0.001)	(0.001)	(0.001)
$d_{it}^X * d_{it}^M$	−0.060***	0.023*	−0.037**	−0.004
	(0.016)	(0.012)	(0.019)	(0.014)
Fixed Effects	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes
Observations	73,882	73,882	73,882	73,882
R^2	0.142	0.171	0.137	0.255
Adjusted R^2	0.026	0.059	0.021	0.154
F Statistic (df = 18; 65079)	599.067***	745.988***	575.117***	1,234.501***
Note:			*p<0.1; **p<0.05; ***p<0.01	

raw materials than firms that do not participate in the trade market. The interaction term between the export and import dummy is very low in two cases and not significant in the other two. This means that firms that participate in the both export and import have higher sales,assets etc. than firms that participate in one of these trade activities. The age variable also has a positive coefficient and is significant at 1% significance level. Therefore, the older a firm becomes the higher its capital, assets etc. become.

This section verifies the presence of trade premia in our dataset. This further substantiates the point that firms that participate in one trade market activity (export or import) have better sales, gross fixed assets etc than firms that do not participate in the trade market. Furthermore, firms that participate in the both of the trade market activities have a higher trade premia than the firms that participate in one of the trade activities. In the next section, I descriptively check whether exports have an effect on imports and vice-versa.

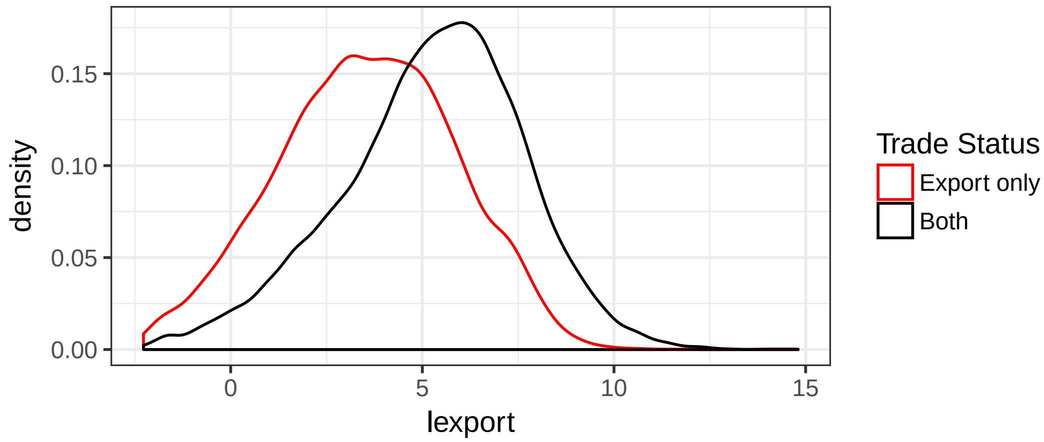
4.3 Complementarity between Exporting and Importing

The difference between density plots of exporting firms that only participate in the export market and firms that participate in import market as well will help give a better idea whether there is an effect of one activity on the other. Table 8 displays the the density plot of the log values of export for firms that participate only in the export market and for firms that participate in both export/import market.

It is seen in Table 8 that firms that participate in both the export/import market have a higher mean of exports (5.24) than firms that only participate in the export market (3.56). This suggests that importing has a positive effect on exporting such that the firms export more than they would if they did not participate in the import market.

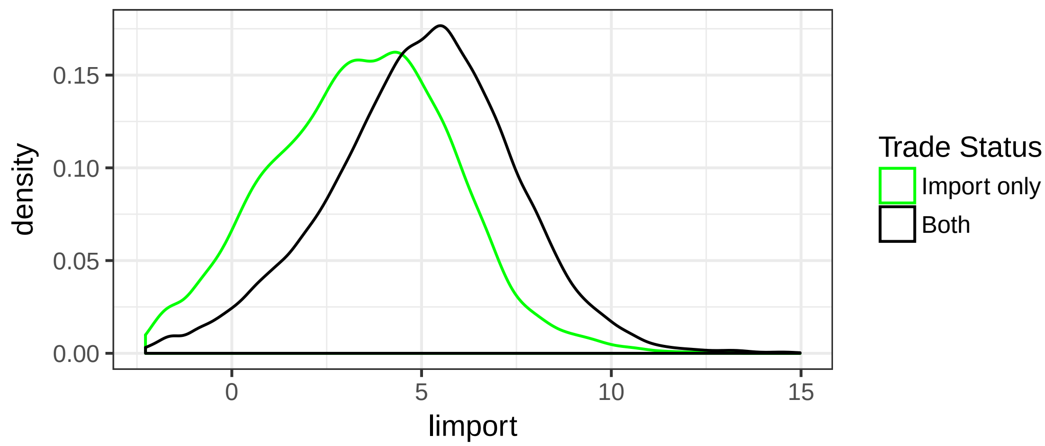
Table 9 displays the density plot of the log value of import for firms that participate only in the import market and for firms that participate in both export/import market. It is seen in Table 9 that firms that participate in both the export/import market have a higher imports (4.95) than firms that only participate in the import market (3.41). This suggests that exporting has a positive effect on importing firms.

Table 8 Summary statistics of Export (log)



Status	mean	sd
None	0	0
Import only	0	0
Export Only	3.56	2.3
Both	5.24	2.39

Table 9 Summary statistics of Import (log)



Status	mean	sd
None	0	0
Import only	3.44	2.38
Export Only	0	0
Both	4.99	2.45

Tables 8 and 9 suggest that both these activities have a positive effect on the other and this might be because importing complements exporting and vice-versa. Therefore, there might be complementarity between these activities that needs further research.

I estimate the trade premia similar to the regression in the previous section on exporting value of the discrete decision to import and vice-versa. This is done by estimating the following equation using a fixed-effects (FE) regression:

$$ln(Export)_{it} = \alpha + \beta_1 d_{it}^{M} + +\beta_2 age_{it} + \epsilon_{it}$$

$$ln(Import)_{it} = \alpha + \beta_1 d_{it}^{X} + \beta_2 age_{it} + \epsilon_{it}$$

Table 10 Export and Import Premia

	Dependent variable:			
	Export	Import	Capital Productivity	Profit to Sales
	(1)	(2)	(3)	(4)
d_{it}^{M}	0.506^{***}		0.159^{***}	0.061^{***}
	(0.026)		(0.009)	(0.004)
d_{it}^{X}		0.458^{***}	0.097^{***}	0.021^{***}
		(0.024)	(0.011)	(0.005)
Age_{it}	0.068^{***}	0.051^{***}	-0.005^{***}	-0.001^{**}
	(0.003)	(0.002)	(0.001)	(0.0004)
$d_{it}^{X} * d_{it}^{M}$			-0.027^{**}	-0.005
			(0.013)	(0.006)
Fixed Effects	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes
Observations	38,359	44,174	69,374	72,401
Note:				*p<0.1; **p<0.05; ***p<0.01

The first two columns of table 10 display these results. It is seen that the discrete decision to import has a positive and significant effect on the value of imports such that the discrete decision to participate in the import market increases the value of export by 0.506 . The discrete decision to export also has a positive and significant effect on the quantity of imports by increasing the value of imports by 0.458. This further suggests the presence of complementarity between exporting and importing.

In the next section, I investigate the effect exporting/importing on crude measures of productivity.

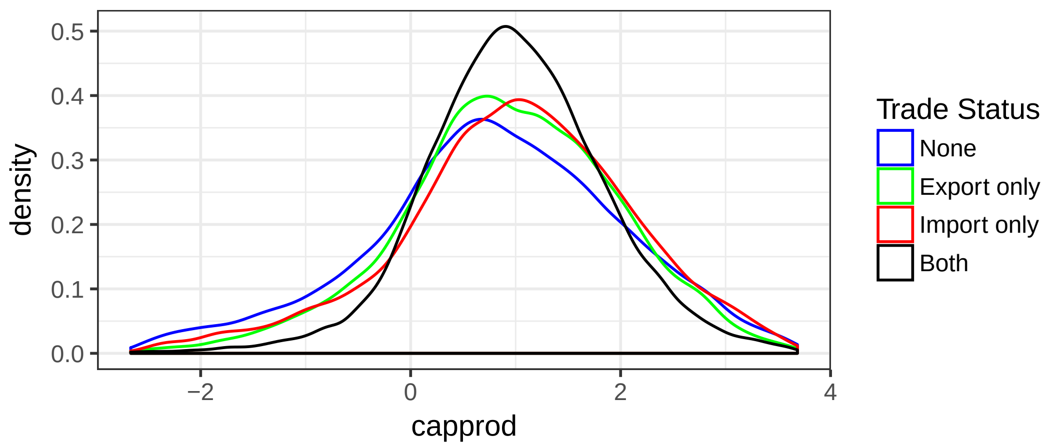
4.4 Productivity and Export/Import

Gupta, Patnaik, and Shah (2018) define a rough measure of productivity known as *capital productivity*. It is defined as the log of value added per unit of capital used by a firm:

$$log(VA_{it} - log(K_{it}))$$

where VA_{it} is firm-level value added, computed as total industrial sales plus change in stock minus power and fuel expenditures, and raw material expenses. Table 11 displays the summary statistics for this variable based on the trade activity status. It can be seen that mean of capital productivity increases as activity status moves from *None* to *Export only/Import only* to *Both*, whereas the standard deviation also decreases.

Table 11 Summary statistics of Capital Productivity (log)

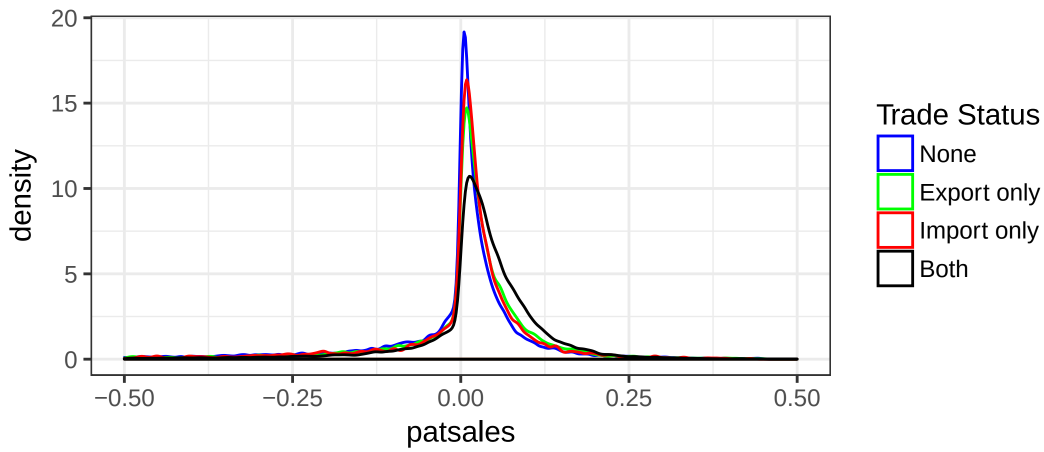


Status	mean	sd
None	0.8	1.2
Import only	0.94	1.03
Export Only	1	1.1
Both	0.99	0.86

It is seen in table 11 that firms participating in the trade market increases the capital productivity as firms participating in the trade market have a higher mean than those firms that do not participate in the trade. Firms that do not participate in the trade market have a very high standard deviation of profit to sales.

I also use another rough measure of productivity i.e *Profit to Sales* which is calculated by dividing the profit of a firm with its sales. Profit to sales is calculated by dividing the profit after tax with sales. This measure can be interpreted as a profitability measure. Table 12 shows the summary statistics for this variable.

Table 12 Summary statistics of Profit to Sales



Status	mean	sd
None	-0.1	0.43
Import only	-0.03	0.29
Export Only	-0.06	0.34
Both	0.01	0.21

The same pattern is observed for these variable as well. The mean of these values increases from -0.1 when a firm is not participating in the trade market to -0.03 and -0.06 when a firm is participating in import and export respectively to 0.01 when a firm is participating in both of the trade market activities.

The last two columns of table 10 estimate the trade premia for the crude measures of the productivity defined above i.e Capital Productivity and Profit-to-sales ratio. It is seen that both of the crude measures of productivity react positively to the discrete decisions to import and export. The discrete decision to participate in the export market increases the capital productivity by 0.159 and profit to sales ratio by 0.061. The discrete decision to import increases the capital productivity by 0.097 and profit to sales ratio by 0.021. Moreover, since the interaction term is really low in one case and is not significant in the other. This suggests that participation in both activities leads to higher productivity than participation in one activity.

4.5 Sunk Cost (Transition Probability)

Table 13 displays the transition probabilities observed in the data and the values in the brackets represent the number of firms. It is seen that there are very high levels of persistence from *None* in t-1 to *None* in t. This means that there must be high sunk costs to enter in the export/import market since only 12% of the firms that do not participate in the trade market in t-1 start participating in the trade market in t. Moreover, high levels of persistence are also observed in *Both* (91.5%). The levels of persistence in *Import Only* and *Export Only* is not as high. A large portion of firms switch to participating in both the trade market activities. This might mean that participating in the export in time period in t-1 complements participating in the import market in time period t and vice-versa. Also, the number of firms that flip-flop (i.e switch trade market status) is low, which provides further evidence that there are fixed costs to participating in the stock market as well.

Table 13 Transition probability

T-1/ T	None	Import Only	Export Only	Both
None	0.877	0.068	0.033	0.022
	(16431)	(1275)	(623)	(416)
Import Only	0.131	0.732	0.011	0.127
	(1304)	(7278)	(107)	(1259)
Export Only	0.134	0.023	0.664	0.179
	(650)	(113)	(3215)	(865)
Both	0.016	0.036	0.033	0.915
	(467)	(1030)	(963)	(26491)

This section provided descriptive evidence that firms that participate in the trade market are bigger (Trade premia), have higher productivity than firms that do not participate in the trade market. It provide evidence that exporting firms export more if they participate in the import market as well and vice-versa. It also provided evidence that there is persistence in trade status (especially when the firms are not participating in the trade market and when they are participating in the export and import market).

In the next section, I proceed to calculate productivity using more sophisticated techniques used in the literature and check the endogenous effect of the decision to export/import on productivity. Then, check if the sunk cost hypothesis and self-selection hypothesis is observed using a dynamic random effects probit model. I end the next section by estimating a dynamic bivariate probit model to examine the complementary nature of these two activities.

5 Analysis

I check if the phenomenon seen in the descriptive statistics are also observed using more sophisticated techniques.

I divide this into three parts and check whether:

- Learning-by-Doing: How does lagged choice of export/import impact productivity?
- Self-Selection and Sunk Cost Hypothesis : Selection of more productive firms into exporting/importing and persistence of these activities
- Complementarity between exporting and importing (Lagged and Contemporaneous): Does engaging in one activity complement participation in the other?

5.1 Learning-by-doing

I assume that the firms have a Cobb-Douglus production function:

$$y_{it} = \beta_l l_{it} + \beta_k K_{it} + \omega_{it} + \eta_{it} \tag{1}$$

where l_{it} is the labor, K_{it} is capital, ω_{it} is the total factor productivity or unobserved productivity and η_{it} is a unknown shock affecting the firms output. However, the OLS estimation provides biased results as it does not account for simultaneity problem i.e the productivity of a firm should be correlated with the inputs of production.

There is a vast literature on the estimation of productivity starting with the seminal paper in productivity estimation: Olley and Pakes (1992) and subsequent modifications of their method: Levinsohn and Petrin (2003), Ackerberg, Caves, and Frazer (2006) and Wooldridge (2009). The estimation strategy of Olley and Pakes (1992) is highlighted in section 7.2.

I use the methods shown in Levinsohn and Petrin (2003) and Ackerberg, Caves, and Frazer (2006) to estimate Cobb-Douglus parameters.

Levinsohn and Petrin (2003) **LP** uses a strategy similar to Olley and Pakes (1992) but use intermediate input demand as the function to invert out ω_{it}. Here, the intermediated material demand function is given by:

$$m_{it} = m_{it}(\omega_{it}, k_{it})$$

This function is assumed to be montonically increasing and therefore productivity can be found by inverting the function above. Therefore, we can write the Cobb-Douglus equation as:

$$y_{it} = \beta_l l_{it} + \phi_{it}(m_{it}, K_{it})$$

where $\phi_{it}(m_{it}, K_{it}) = \beta_k K_{it} + \omega_{it}(m_{it}, K_{it})$ They suggest the use of a third degree polynomial to approximate ϕ_{it} and substitute it into the equation above to give the following result:

$$y_{it} = \beta_l l_{it} + \sum_{i=0}^{3}\sum_{j=0}^{3-i}\delta_{ij}k_{it}^{i}m_{it}^{j}$$

The coefficient is β_l is estimtated by OLS using the equation above as they assume that the labor is a flexible input i.e there are no labor adjustment costs and $\hat{\phi_{it}}$ is estimtated by subtracting the effect of labor from the fitted value of the above equation:

$$\hat{\phi_{it}} = \hat{y_{it}} - \hat{\beta_l} l_{it} = \sum_{i=0}^{3}\sum_{j=0}^{3-i}\hat{\delta_{ij}}k_{it}^{i}m_{it}^{j}$$

Therefore, So, for any value of β_k^*:

$$\hat{\omega_{it}} = \hat{\phi_{it}} - \beta_k^* k_{it}$$

$$y_{it}^* = y_{it} - \beta_l l_{it} = \beta_k K_{it} + \omega_{it}(m_{it}, K_{it})$$

Since it is also assumed that ω_{it} follows a first order markov process :

$$\omega_{it} = E[\omega_{it}|\omega_{it-1}] + \epsilon_{it}$$

They also assume a polynomial expansion of the expectation above to give:

$$\omega_{it} = \gamma_o + \gamma_1\omega_{it-1} + \gamma_2\omega_{it-1}^2 + \gamma_3\omega_{it-1}^3 + \epsilon_{it}$$

Therefore, the value of β_K, for which the expression below is minimized is chosen to be the coefficient of capital.

$$\min_{\beta_k^*} \sum (y_{it} - \beta_l \hat{l}_{it} - \beta_k^* K_{it} - E[\omega_{it}\hat{|}\omega_{it-1}])^2 \tag{2}$$

On the other hand, Ackerberg, Caves, and Frazer (2006)(ACF) use a similar strategy but, suggest that labour might be correlated with productivity. Therefore, they write the firms input material demand as a function of productivity, capital as well as labor :

$$m_{it} = f_{it}(\omega_{it}, k_{it}, l_{it})$$

Inverting this function for ω_{it} and substituting into the production function results in the following equation of the form:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + f_{it}^{-1}(m_{it}, k_{it}, l_{it}) + \epsilon_{it} \tag{3}$$

They suggest that the labor coefficient along with capital since it is correlated with productivity. They suggest the following steps:

1. Obtain $\phi_{it}(m_{it}, k_{it}, l_{it}) = \beta_l l_{it} + \beta_k k_{it} + f_{it}^{-1}(m_{it}, k_{it}, l_{it})$ by regressing y_{it} on polynomial approximation of $\phi_{it}(m_{it}, k_{it}, l_{it})$
2. Use the markovian nature of $\omega_{it} = E(\omega_{it}|\omega_{it-1}) + e_{it}$ and use the following moment equations to estimate β_K and β_l:

$$E[e_{it}|(k_{it}, l_{it-1})] = 0 \tag{4}$$

However, exogenously regressing lagged export/import variables on productivity (obtained by getting the residuals) suggests that past export/import performance does not impact the evolution of capital/labor. This has been highlighted by De Loecker (2013), and they suggest that the trade activities should be accommodated endogenously in the productivity evolution process. This is done by accommodating the the lagged export/import variable into the minimisation procedure of productivity:

$\omega_{it} = E(\omega_{it}|\omega_{it-1}, d^X_{it-1}, d^M_{it-1}) + e_{it}$

I use the log values of the following variables for the estimation procedure: gross fixed assets as a measure of capital, salary as a measure of labor and a expenditure on raw materials as a measure of intermediated input. Since productivity evolution is assumed to have a markovian nature , I assume the following form of productivity evolution:

$$\omega_{it} = \alpha_o + \alpha_1\omega_{it-1} + \alpha_2\omega_{it-1} + \alpha_3\omega^2_{it-1} + \alpha_4 d^X_{it-1} + \alpha_5 d^M_{it-1} + \alpha_6 d^X_{it-1} d^M_{it-1} + \nu_{it}$$

where d^X_{it-1} and d^M_{it-1} is the discrete decision to export and import respectively.

The estimates of the productivity evolution and the Cobb-Douglus coefficients using the Levinsohn-Petrin method are shown in Table 15 and Table 14 respectively.

Table 14 Cobb-Douglus coefficients

	Value	Bootstrap Standard Errors
β_L	0.299	0.006
β_K	0.437	0.021

It is seen in Table 15 that productivity evolutions depends strongly on the past productivity and the coefficients suggest that there is a non linear relationship with past productivity. However, the interesting result is that lagged decision to export does not a have a significant effect on productivity. However, it is seen that the discrete decision to import does have a significant effect on productivity i.e the lagged decision to import causes the productivity to increase by 3 per cent. The interaction term between exporting and importing also does not have a significant effect on productivity. Table 25 and 24 display the Cobb-Douglus coefficients when the productivity evolution is dependent on the lagged continuous value of export and import rather than the lagged decision to export/import. and they show results similar to the ones shown in tables 14 and 15 for the discrete decision. They show that lagged continuous value of export does not have a significant effect on productivity, increase in lagged continuous value of import by 1 unit increases the productivity by 2.6 %.

Table 15 Productivity Evolution

	Dependent variable:
	ω_{it}
α_1	0.701***
	(0.005)
α_2	0.082***
	(0.002)
α_3	−0.008***
	(0.0003)
α_4	0.002
	(0.007)
α_5	0.035***
	(0.005)
α_6	−0.006
	(0.009)
α_0	0.341***
	(0.007)
Observations	62,487
Note:	*p<0.1; **p<0.05; ***p<0.01

As written before, De Loecker (2013) say that exogenously accommodating the decision to export/import implies that past export/import experience has no impact on direct technological improvements. Therefore, the coefficient of capital will be biased upwards if the decision to participate in the trade market correlated with capital. Table 16 displays the coefficients when the export/import decision is not endogenously accommodated in productivity evolution. In this case, the coefficient of capital is biases upwards (0.452 compared to 0.437 in the endogenous case) since the variation in output is attributed more towards capital rather than productivity.

Table 16 Cobb-Douglus coefficients

	Value	Bootstrap Standard Errors
β_L	0.295	0.005
β_K	0.452	0.021

The results from the estimation method mentioned in Ackerberg, Caves, and Frazer (2006) are shown in 17 and 18. The coefficient of labor is much lower in this estimation method which suggests that labor is not a flexible input which suggests that is also correlated to productivity. However, in terms of the endogenous effect of export/import on productivity is roughly the same. The export decision does not have a significant effect on productivity and the import decision has a 2.6% increase to the productivity. These results are similar to the results from the Levinsohn-Petrin estimation.

The procedure above does not account for immediate impact of the importing behavior to the output of a firm as it is assumed that participating in the import market only helps in increasing the productivity of the firm in the next period. This is a drawback of the estimation procedure.

Based on these results above, it is seen that manufacturing firms experience learning-by-importing and do not experience learning-by-exporting. In the next section, I check if more productive self-select into exporting/importing and if the sunk cost hypothesis is observed for the discrete decision to export and import with the help of a dynamic probit random effects model.

5.2 Self Selection and Sunk Cost hypothesis

The self-selection hypothesis states that entry into the trade market involves fixed and sunk costs and only the most productive firms can overcome these trade costs. Therefore, to participate in the export/import market a firm must pay a certain costs and only the most productive are able to pay the costs. To check this hypothesis, I estimate a dynamic random effects probit model similar to the model used in Mark J Roberts and Tybout (1997). I define d_{it}^X as the discrete decision to export, where Bellman equation for a profit maximising firm deciding to enter the

Table 17 Productivity Evolution ACF (Discrete)

	Dependent variable:
	ω_{it}
α_1	0.755^{***}
	(0.004)
α_2	0.073^{***}
	(0.002)
α_3	-0.008^{***}
	(0.0003)
α_4	0.00004
	(0.007)
α_5	0.026^{***}
	(0.005)
α_6	-0.005
	(0.009)
α_0	0.250^{***}
	(0.006)
Observations	62,487
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 18 Cobb-Douglus coefficients

	Value	Bootstrap Standard Errors
L_{it}	0.374	0.120
K_{it}	0.474	0.069

export market is:

$$V_{it}(S_{it}) = max_{d_{it}^X} E_t(\sum_{i=0}^{\infty} \delta^{t+i} R_{i,t+i}|S_{it}) \tag{5}$$

where δ is the one-period discount factor, S_{it} is the relevant state variables affecting the firms decision, R_{ij} is the revenue. The equation above can also be written as:

$$V_{it}(S_{it}) = max_{D_{it}^X}(\pi^D + d_{it}^X(\pi^X - f^X - c^X(1 - d_{it-1}^X)) + \sum_{j=t+1}^{\infty} \delta^{t-j} R_{ij}|S_{it}) \tag{6}$$

where π^D is the domestic profit, π^X is the export profit, f^X is the fixed cost of exporting and c^X is the sunk cost of exporting.

$$V_{it}(S_{it}) = max_{D_{it}^X} E(\pi^D + d_{it}^X(\pi^X - f^X - c^X(1 - d_{it-1}^X) + \delta E(V_{it}(S_{it+1})|S_{it}, d_{it}^X))) \tag{7}$$

Thus, a firm will participate in the export market if:

$$\pi_{it}^* = \pi^D + \pi^X + \delta E_t(V_{i,t+1}(S_{it+1}|S_{it}, D_{it}^X = 1) - V_{i,t+1}(S_{it+1}|S_{it}, D_{it}^X = 1)) - (f^X + c^X(1 - D_{it-1}^X)) \tag{8}$$

It is assumed that the state variables entering the Bellman equation are the following: $S_{it} = (K_{it}, l_{it}, \omega_{it}, d_{it-1}^X)$ i.e capital, labor, productivity and lagged decision to participate in the export market. In the model above, it is assumed that a firm pays a sunk cost if $d_{it-1}^X = 0$. Therefore, if there are no sunk costs, then according to the Bellman equation above, a firm will export as long as the the current profits plus the discounted expected value is greater than the fixed costs of exporting. The reduced form expression of the equation above is:

$$d_{it}^X = \begin{cases} 1, & \text{if } \pi_{it}^* \geq 0. \\ 0, & \text{if } \pi_{it}^* < 0. \end{cases} \tag{9}$$

To enable to reduced form estimation of the model, I write the equation above as a linear function of the relevant state variables along with dummy variables to adjust for industry and time effects, to give the equation below:

$$\pi_{it}^* = \gamma_1^X d_{it-1}^X + \gamma_3^X \hat{\omega}_{it} + \beta_1^X K_{it} + \beta_2^X L_{it} IndustrialDummy_i^X + TimeDummy_t^X + \alpha_i + \epsilon_{it}^X \tag{10}$$

So if the coefficient of d_{it-1}^X is significantly positive, it provides evidence of sunk cost to participate in the export market as this means that there is a persistence in the exporting behavior. If the sunk cost hypothesis did not hold then the probability of exporting should not depend on the lagged decision to export. And if the coefficient of ω_{t-1} is positive, then this proves the self-selection hypothesis as firms with high productivity have a higher probability of participating in the export market.

A similar model can be estimated with the discrete decision to import, since importing also involves additional fixed and sunk costs, a firm would be only participate in the import market if gets productivity benefits to overcome the costs.

Learning-from-importing has been demonstrated in the previous section. The Bellman equation of a firm deciding to enter the import market is the following:

$$V_{it}(S_{it}) = max_{D_{it}^M}(\pi + d_{it}^M(-f^M - c^M(1 - d_{it-1}^M) + \delta E(V_{it}(S_{it+1})|S_{it}, d_{it}^M))) \quad (11)$$

where π is total profit of a firm, f^M is the fixed cost of importing and c^M is the sunk cost of importing and $S_{it} = (K_{it}, l_{it}, \omega_{it}, d_{it-1}^M)$.

Since S_{it} contains productivity, this Bellman equation assumes that importing in time t provides benefits in t+1 as the decision to import improves the productivity in the next period. This has been shown in the learning-by-doing section. Here, a firm will participate in the import market if:

$$\pi^* = \pi + \delta E_t(V_{i,t+1}(S_{it+1}|S_{it}, d_{it}^M = 1) - V_{i,t+1}(S_{it+1}|S_{it}, d_{it}^X = 0)) - (f^M + c^M(1 - d_{it-1}^M)) > 0 \quad (12)$$

Therefore, a firm will import if profit plus discounted productivity benefits outweigh the costs to participate in the import market. This can be tested with a reduced form dynamic probit model similar to the discrete decision to export:

$$d_{it}^M = \begin{cases} 1, & \text{if } \pi^*. \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

Therefore, the reduced form equations for the **base model** of both the discrete choices can be written as:

$$d_{it}^{X*} = \gamma_1^X d_{it-1}^X + \gamma_3^X \hat{\omega}_{it} + \beta_1^X K_{it} + \beta_2^L L_{it-1} s i_i^X + \mu_t^X + \alpha_i + \epsilon_{it}^X \quad (14)$$

$$d_{it}^{M*} = \gamma_1^M d_{it-1}^M + \gamma_3^M \hat{\omega}_{it} + \beta_1^M K_{it} + \beta_2^L L_{it} s_i^M + \mu_t^M + \alpha_i + \epsilon_{it}^M \quad (15)$$

Here s_i is a vector of industry dummies and μ_t is a vector of time dummies.

To check the presence of lagged complementarity between exporting and importing, I add another variableto state space: lagged export choice into the import decision equation and lagged import choice to export decision equation. Therefore, the Bellman equations that accounts for lagged complementarity for the export decision is:

$$V_{it}(S_{it}) = max_{D_{it}^X} E(\pi^D + (\lambda * d_{it-1}^M + 1 - d_{it-1}^M) * d_{it}^X(\pi^X - f^X - c^X(1 - d_{it-1}^X) + \delta E(V_{it}(S_{it+1})|S_{it}, d_{it}^X))) \quad (16)$$

Here, λ has a value between 0 and 1 if a firm participated in the import market in the previous period.

Similarly, the Bellman equation for the import decision is:

$$V_{it}(S_{it}) = max_{D_{it}^M}(\pi + (\lambda * d_{it-1}^M + 1 - d_{it-1}^M) * d_{it}^M(-f^M - c^M(1 - d_{it-1}^M) + \delta E(V_{it}(S_{it+1})|S_{it}, d_{it}^M))) \quad (17)$$

Here, λ has a value between 0 and 1 if a firm participated in the export market in the previous period. And the reduced form equation of the two Bellman equations that account for lagged complementarity are the following:

$$d_{it}^{X*} = \gamma_1^X d_{it-1}^X + \gamma_2^M d_{it-1}^M + \gamma_3^X \hat{\omega}_{it} + \beta_1^X K_{it} + \beta_2^X L_{it} + s_i^X + \mu_t^X + \alpha_i + \epsilon_{it}^X \quad (18)$$

Here, it is important to notice that γ_2^M measures the effect of lagged importing on exporting after accounting for the productivity benefits of importing, since the equation already uses the productivity estimates. Therefore, if $\gamma_2^X > 0$, then this means that importing in time t leads to decrease in cost of exporting at time t+1.

$$d_{it}^{M*} = \gamma_1^M d_{it-1}^M + \gamma_2^M d_{it-1}^M + \gamma_3^M \hat{\omega}_{it} + \beta_1^M K_{it} + \beta_2^M L_{it} + s_i^M + \mu_t^M + \alpha_i + \epsilon_{it}^M \quad (19)$$

In the equation above, if $\gamma_2^X > 0$, then this means that exporting in time t leads to decrease in cost of importing at time t+1.

I use the dynamic random effects probit specification with Wooldridge method (Wooldridge (2005)) which treats the initial conditions problem by accounting for the correlation of the initial value with α:

$$\alpha_i = \gamma d_{i1} + \tilde{\alpha_I}$$

where $\tilde{\alpha_I} \sim N(0, \sigma_\alpha^2)$

I use the log value of gross fixed assets as a measure of capital, log value of wages as a measure of labor and use productivity estimates of the Levinsohn-Petrin method in the previous section.

Table 19 Dynamic Random Effects Probit (Estimates)

	Export Decision		Import Decision	
	(1)	(2)	(3)	(4)
	Base	Lagged Complementarity	Base	Lagged Complementarity
d^X_{it-1}	1.834***	1.786***		0.380***
	(71.18)	(68.65)		(16.31)
d^M_{it-1}		0.370***	1.600***	1.554***
		(14.67)	(66.32)	(63.85)
$\hat{\omega}_{it}$	0.216***	0.198***	0.277***	0.266***
	(15.83)	(14.70)	(21.51)	(21.16)
K_{it}	0.0669***	0.0467***	0.109***	0.100***
	(5.17)	(3.65)	(8.99)	(8.51)
L_{it}	0.210***	0.186***	0.213***	0.186***
	(15.09)	(13.53)	(17.06)	(15.20)
d^X_{i1}	1.333***	1.264***		
	(29.35)	(28.62)		
d^M_{i1}			1.081***	0.986***
			(28.27)	(26.62)
_cons	-3.264***	-3.122***	-3.655***	-3.617***
	(-25.05)	(-24.47)	(-28.64)	(-29.17)
rho	0.392	0.373	0.348	0.321
σ^2_α	0.804***	0.771***	0.731***	0.687***
	(0.0245)	(0.023)	(0.021)	(0.021)
Log-Likelihood	-14513.0	-14406.6	-15879.7	-15749.5
Industry Dummies	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes

t statistics in parentheses

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

Table 20 Dynamic Random Effects Probit (Average Marginal Effects)

	Export Decision		Import Decision	
	(1)	(2)	(3)	(4)
	Base	Lagged Complementarity	Base	Lagged Complementarity
d^X_{it-1}	0.218***	0.212***		0.0495***
	(53.32)	(53.84)		(16.34)
d^M_{it-1}		0.0439***	0.209***	0.202***
		(14.65)	(53.75)	(53.94)
$\hat{\omega}_{it}$	0.0257***	0.0235***	0.0362***	0.0347***
	(15.99)	(14.82)	(21.92)	(21.55)
K_{it}	0.00797***	0.00553***	0.0142***	0.0130***
	(5.19)	(3.66)	(9.03)	(8.54)
L_{it}	0.0250***	0.0220***	0.0279***	0.0243***
	(15.20)	(13.62)	(17.27)	(15.37)
N	62485	62485	62485	62485

t statistics in parentheses

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

Table 19 and 20 displays the estimates and average marginal effects of the dynamic random effects probit model respectively. Columns 1 and 3 display the result for the base reduced model of importing and exporting, whereas columns 2 and 4 display the estimates when accounting for the lagged complementarity between exporting and importing.

The likelihood ratio test to see if the base model is rejected by the model that accounts for lagged complementarity gives the following results:

1. Export:
   H_o: Base Model, H_1:Model with lagged discrete import variable. $LR = 2[ln(model2) - ln(model1)] = 2[-14406.6 + +14513.0] = 106.4$ The critical value of $\chi^2_{1;0.95} = 3.84$. Therefore H_o is rejected.

2. Import:
   H_o: Base Model, H_1:Model with lagged discrete export variable. $LR = 2[ln(model2) - ln(model1)] = 2[-15749.5 + 15879.7] = 130.19$ The critical value of $\chi^2_{1;0.95} = 3.84$. Therefore H_o is rejected.

In both cases, the model that accounts for the lagged complementarity performs better than the base model. Thereofore, I interpret the model that acoounts for the lagged complementarity.

Export: The state dependence parameter is positive significant at 1% level with the highest magnitude amonst all of the coefficients. This means that there is persistence in exporting behavior which confirms that sunk-cost hypothesis. The average marginal effect of the lagged decision to export is 0.218. Productivity also has a significant and positive effect on exporting which provides evidence of self-selection of high productivity into exporting. Capital and Labor also have a positive and significant effect, which tells us that a bigger firm has a higher chance of participating in the export market. The lagged import is significant and positive and this provides evidence that importing at time t increases the probability of exporting at time $t+1$ and cofirms the presence of cost complementarity. The value of $sigma_{\alpha}^2$ in the table does not account for the contribution of the intial condition. Therefore, the actual unobserved heterogeneity $\sigma_{\alpha}^2 = \lambda^2 * V\hat{ar}(d_{i1}^X + sig\hat{ma}_{\alpha}^2 = 1.16$. This means that individual effects capture about 54% of the unexplained variance. This suggests there are variables other than the ones used in the estimation that contribute towards the export market participation decision.

Import: The results for the import decision are quite similar to the results for the export decision. The state dependence parameter is the most highest amongst all variables and significant at 1% level. This means that there is persistence in importing behavior and confirms the sunk cost hypothesis for import decision. The average marginal effect of the lagged decision to export is 0.202. Productivity is also significant at 1% and has a positive effect on importing. This provides evidence of self-selection of high productivity into importing. Capital and Labor also have a positive and significant effect, which confirms that bigger firms have a higher probability of import participation. The lagged export coefficient is significant and positive and this provides evidence that exporting at time t increases the probability of importing at time $t + 1$. The lagged cost complementarity hypothesis is also confirmed for the import decision. The unobserved heterogeneity is high as it explains about 48% of the unexplained variation in the data and the variance of the unobserved heterogeneity is significant. This value is lower than the value when the lagged export decision is not included in the dynamic specification (Model 1)

These results provide evidence of the sunk-cost hypothesis, self-selection of higher productivity and bigger firms into exporting/importing and lagged complementarity. However, this estimation does not account for the contemporaneous complementarity between exporting and importing. The next section shows the results of a dynamic bivariate probit model which accounts for simultaneous complementarity as well.

5.3 Complementarity between exporting and importing(Lagged and Contemporaneous)

The bellman equation of a profit maximising firm deciding to export and import is the following:

$$V_{it}(S_{it}) = max_d(\pi_{it}^d + d_{it}^X \pi_{it}^X + F(d_{it}, d_{it-1}) + \beta E(V_{it}(s_{it+1}|s_{it}, d_{it}))) \tag{20}$$

where $d_{it} = (d_{it}^X, d_{it}^M)$ and $S_{it} = (K_{it}, L_{it}, \omega_{it}, d_{it-1}^X, d_{it-1}^M)$. $F(d_{it}, d_{it-1})$ is the fixed/sunk costs the firms must pay to export/import and it takes the following form:

$F(d_{it}, d_{it-1}) =$

1. $f^x + c^X(1 - d_{it-1}^X)$ for $(d_{it}^X, d_{it}^M) = (1, 0)$
2. $f^M + c^M(1 - d_{it-1}^M)$ for $(d_{it}^X, d_{it}^M) = (0, 1)$
3. $\lambda[f^x + f^M + c^X(1 - d_{it-1}^X) + c^M(1 - d_{it-1}^M)]$ for $(d_{it}^X, d_{it}^M) = (1, 1)$
4. 0 for $(d_{it}^X, d_{it}^M) = (0, 0)$

Here f^X is the fixed cost of exporting,C^X is the sunk cost of exporting, f^M is the fixed cost of importing, f^M is the fixed cost of importing and λ captures the degrees of complementarity between exporting and importing.

The reduced form model of the bellman equation is a dynamic bivariate probit model. Then, the bivariate dynamic probit model in this takes the following form:

$$d_{it}^X = \begin{cases} 1, & \text{if } d_{it}^{X*} > 0. \\ 0, & \text{if } d_{it}^{X*} < 0. \end{cases} \tag{21}$$

$$d_{it}^M = \begin{cases} 1, & \text{if } d_{it}^{M*} > 0. \\ 0, & \text{if } d_{it}^{M*} < 0. \end{cases} \tag{22}$$

The discrete decision of exporting and importing is modelled as a function of previous import and export status controlling for lagged firm characteristics and industry and time fixed effects.

$$d_{it}^{X*} = \gamma_1^X d_{it-1}^X + \gamma_2^X d_{it-1}^M + \gamma_3^X \hat{\omega}_{it} + \beta_1^X K_{it} + s_i^X + \mu_t^X + \epsilon_{it}^X \tag{23}$$

Here γ_1 identifies the state dependence coefficient, γ_2 accounts for the fact that participating in one activity in time t-1 improves the odds of participating in the other at time t (lagged complimentarity),γ_3 accounts for the self-selection mechanism, β_1 accounts for capital at time t-1 and s_i^M μ_t^M are industrial and time dummies respectively.

The reduced form specification for import is similar to the export equation:

$$d_{it}^{M*} = \gamma_1^M d_{it-1}^M + \gamma_2^M d_{it-1}^X + \gamma_3^M \hat{\omega}_{it} + \beta_1^M K_{it} + s_i^M + \mu_t^M + \epsilon_{it}^M \tag{24}$$

The bivariate specification also allows for the contemporaneous complementarity between the two choices as ϵ_{it}^X and ϵ_{it}^M are allowed to be correlated. This gives gives the following form to the error structure:

$$\begin{pmatrix} \epsilon_{it}^X \\ \epsilon_{it}^M \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

The estimated ρ measures the correlation between the unobserved errors between the two activities. Therefore, this provides evidence of contemporaneous complementarity if it significantly positive after controlling for different effects. This model specification has been used to test the contemporaneous relationship by Aristei, Castellani, and Franco (2013) and Aw, Mark J. Roberts, and Xu (2011).

This model has a few drawbacks: it does not account for individual heterogeneity (α_i) and it does not account for the initial conditions problems. Therefore, the model assumes that d_{i1} is endogenously given and that the firm characteristics and industry and time dummy variables account for the individual heterogeneity between firms.

Table 21 Dynamic Bivariate Probit (Estimates)

	(1) Export Decsion	(2) Import Decision
d^X_{it-1}	2.539***	0.397***
	(152.29)	(23.28)
d^X_{it-1}	0.360***	2.175***
	(19.05)	(129.84)
$\hat{\omega}_{it}$	0.103***	0.165***
	(11.84)	(19.66)
K_{it}	0.00873	0.0788***
	(1.08)	(10.27)
L_{it}	0.135***	0.134***
	(16.29)	(16.52)
Constant	-2.212***	-2.768***
	(-28.56)	(-36.07)
ρ	0.439***	
	(29.48)	
LR test	$\rho = 0,\ \chi^2(1) = 868.863$	
	p=0.000	
Industry Dummies	Yes	
Time Dummies	Yes	

t statistics in parentheses

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

Table 21 displays the results for dynamic probit specification. All the coefficients are significant at 1% level and are quite similar to the coefficients in the previous section. The state-dependence coefficient has the strongest effect amongst all the variables, suggesting that there is persistence in both the activities and high sunk cost. There is a positive effect of productivity on both activities, providing further evidence of self-selection of firms into exporting and importing. The coefficients of capital, labor and productivity are positive and quite similar to each other. Firms which were importing in the previous year are more likely to be exporters and firms which were exporting in the previous year are more likely to be importing this year.

The results suggest simultaneous complementarity between exporters and importers as the likelihood-ratio test with the null hypothesis that the correlation between the unobserved errors is 0 is rejected at 1% significance level. The estimated value of ρ is 0.438 and is significantly different than zero. This suggests that there is simultaneous complementarity between exporting and importing.

6 Conclusion

The results from this section and descriptive statistics suggest a few overall themes of the data:

- Learning-by-doing: I estimate productivity using Levinsohn and Petrin (2003) such that they are dependent on the lagged export and import choices and I get the following results for the two trading activities:
    1. Export: Firms do not display learning-by-exporting as the coefficient of discrete choice of lagged export decision does not have a significant effect on productivity.
    2. Import: Firms display learning-by-exporting a the coefficient of discrete choice of lagged import decision does have a significant effect on productivity
- Self-Selection: I regress the lagged productivity values of the estimated productivity for $t = t-1, t-2$ and $t-3$ on the discrete choice of exporting and importing after controlling for firm characteristics and industry and time fixed effects and get the following result:
    1. Export: The coefficient for the discrete export choice has a positive effect significant at 1% level on the lagged values of productivity. This suggests that firms learn to export.
    2. Import: The coefficient for the discrete import choice also has a positive effect significant at 1% level on the lagged values of productivity. This suggests that firms learn to import as well.
- Complementarity between exporting and importing: I run a bivariate dynamic probit regression of discrete choice of exporting/importing on their lagged values, firm characteristics and industry and time dummies to get the following results:
    1. Export: There is strong persistence in exporting behavior, lagged import decision has a positive effect on current exporting behavior.
    2. Import: There is strong persistence in importing behavior, lagged export decision has a positive effect on current importing behavior.
    3. : Contemporaneous Complementarity: There is a strong presence of contemporaneous complementarity since the errors for the equations are highly correlated are significant at 1% level.

The bivaraite dynamic provide results suggest that there is strong cost complementarity between exporting at time t with importing at time t and t-1 and vice-versa. But, the learning by export result show that there is no learning by exporting. This must mean that importing must help in reducing the cost to export since they do not enter the productivity mechanism.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer (2006). "Structural identification of production functions". In:

Aristei, David, Davide Castellani, and Chiara Franco (2013). "Firms exporting and importing activities: is there a two-way relationship?" In: *Review of World Economics* 149.1, pp. 55–84.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu (2011). "R&D Investment, Exporting, and Productivity Dynamics". In: *American Economic Review* 101.4, pp. 1312–44. DOI: 10.1257/aer.101.4.1312. URL: http://www.aeaweb.org/articles?id=10.1257/aer.101.4.1312.

Bernard, Andrew B and J Bradford Jensen (1999). "Exceptional exporter performance: cause, effect, or both?" In: *Journal of international economics* 47.1, pp. 1–25.

De Loecker, Jan (2013). "Detecting learning by exporting". In: *American Economic Journal: Microeconomics* 5.3, pp. 1–21.

Gupta, Apoorva, Ila Patnaik, and Ajay Shah (2018). "Exporting and firm performance: Evidence from India". In: *Indian Growth and Development Review* 12.1, pp. 83–104.

Haidar, Jamal Ibrahim (2012). "Trade and productivity: Self-selection or learning-by-exporting in India". In: *Economic Modelling* 29.5, pp. 1766–1773.

Halpern, Adam Szeidl, et al. (2011). "Imported inputs and productivity". In: *Center for Firms in the Global Economy (CeFiG) Working Papers* 8, p. 28.

Kasahara, Hiroyuki and Beverly Lapham (2013). "Productivity and the decision to import and export: Theory and evidence". In: *Journal of international Economics* 89.2, pp. 297–316.

Levinsohn, James and Amil Petrin (2003). "Estimating production functions using inputs to control for unobservables". In: *The review of economic studies* 70.2, pp. 317–341.

Muûls, Mirabelle and Mauro Pisu (2009). "Imports and Exports at the Level of the Firm: Evidence from Belgium". In: *World Economy* 32.5, pp. 692–734.

Olley, G Steven and Ariel Pakes (1992). *The dynamics of productivity in the telecommunications equipment industry*. Tech. rep. National Bureau of Economic Research.

Roberts, Mark J and James R Tybout (1997). "The decision to export in Colombia: An empirical model of entry with sunk costs". In: *The American Economic Review*, pp. 545–564.

Topalova, Petia and Amit Khandelwal (2011). "Trade liberalization and firm productivity: The case of India". In: *Review of economics and statistics* 93.3, pp. 995–1009.

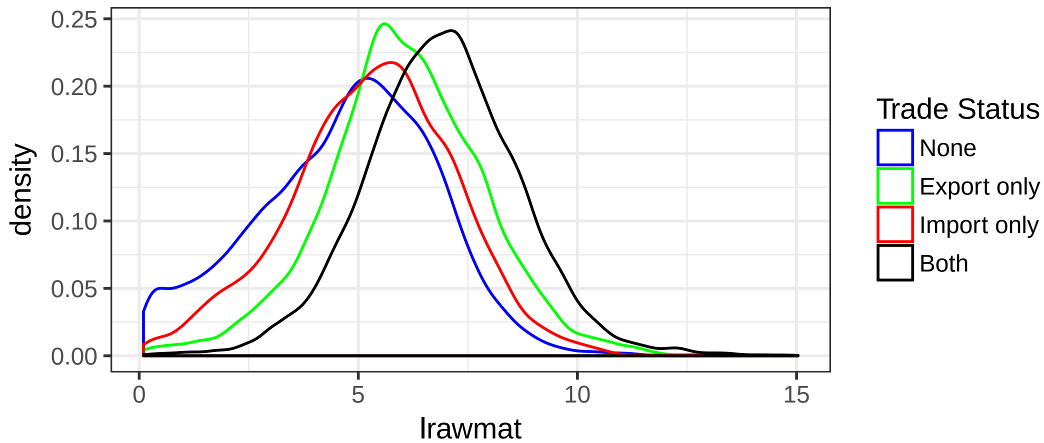
Wooldridge, Jeffrey M (2005). "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity". In: *Journal of applied econometrics* 20.1, pp. 39–54.

– (2009). "On estimating firm-level production functions using proxy variables to control for unobservables". In: *Economics Letters* 104.3, pp. 112–114.

7 Appendix

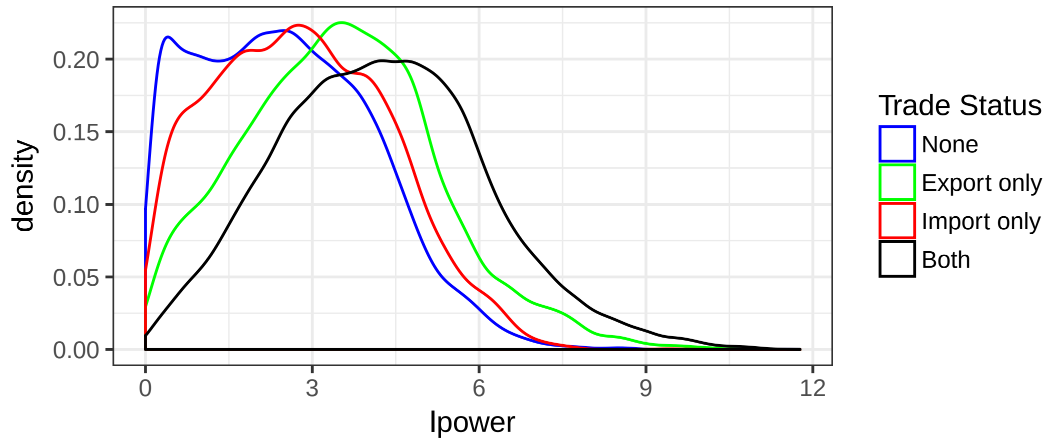
7.1 Appendix (Descriptive Statistics)

Table 22 Summary statistics of Expenditure on raw material (log)



Status	mean	sd
None	4.73	2.03
Import only	6.04	1.82
Export Only	5.32	1.88
Both	6.9	1.77

Table 23 Summary statistics of expenditure on power and fuel (log)



Status	mean	sd
None	2.51	1.56
Import only	3.5	1.75
Export Only	2.81	1.57
Both	4.24	1.88

7.2 Appendix (Productivity Estimation OP)

Olley and Pakes (1992) (OP) use the following strategy to estimate the Cobb-Douglus function:

The log transformation of the production function is written as :

$$y_{it} = \beta_o + \beta_l l_{it} + \beta_k K_{it} + \omega_{it} + \epsilon_{it} \tag{25}$$

where l_{it} is labour, K_{it} is capital, m_{it} is the demand for materials, ω_{it} is firm-level-productivity observable to the firm and ϵ_{it} are shocks to production.
The optimal investment decision of the firm is characterised by the following function:

$$i_{it} = f_K(l_{it-1}, K_{it}, \omega_{it}) \tag{26}$$

The optimal labor decision is characterised by the following function:

$$l_{it} = f_K(l_{it-1}, K_{it}, \omega_{it}) \tag{27}$$

These are the assumptions made by **OP**:

1. $f_K(l_{it-1}, K_{it}, \omega_{it})$ is invertible in ω_{it}
2. ω_{it} follows a first order markov process
3. Investment at period i is not active until period $t+1$: $K_{it+1} = (1-\delta)k_{it} + i_{it}$
4. Labor is a perfectly flexible input i.e there are no labor adjustment cost

The estimation procedure proceeds in two step:

1. Estimation of β_k: Since $\omega_{it} = f_K^{-1}(l_{it-1}, k_{it}, i_{it})$ and inserting this in the Cobb-Douglus equation to get:

$$y_{it} = \beta_l l_{it} + \phi_{it}(l_{it-1}, i_{it}, K_{it})$$

   where $\phi_{it}(l_{it-1}, i_{it}, K_{it}) = \beta_k K_{it} + f_K^{-1}(l_{it-1}, k_{it}, i_{it})$ $\phi_{it}(l_{it-1}, i_{it}, K_{it})$ is approximated using a polynomial expression and $beta_l$ is estimated by using OLS on the above equation

2. Estimation of β_k: Since ω_{it} follows a first order markov process, it can be written as:

$$\omega_{it} = h(\omega_{it-1}) + e_{it}$$

   Then ϕ_{it} can be written as:

$$\phi_{it} = \beta_k K_{it} + h(\omega_{it-1}) + e_{it}$$

$$\phi_{it} = \beta_k K_{it} + h(\phi_{it-1} - \beta_k k_{it-1}) + e_{it}$$

   The unknown form of function h is approximated by quadratic function and for any value of β_K to get:

$$\hat{\phi_{it}} = beta_k K_{it} + \gamma_0 \gamma_1 \omega_{it-1}^{\hat{\beta_k}} + \gamma_2 (\omega_{it-1}^{\hat{\beta_k}})^2 + \gamma_3 (\omega_{it-1}^{\hat{\beta_k}})^3$$

   This expression is minimised to get the estimate of β_k.

7.3 Appendix (Learning-by-doing)

7.4 Appendix (Dynamic Random Effects Probit Model)

Let i be the unit and t be the time. A dynamic random effects probit is written as:

$$y_{it}^* = \gamma y_{i,t-1} + x_{it}'\beta + \alpha_i + \epsilon_{it}; y_{it} = 1[y_{it}^* > 0]$$

where γ is the state dependence parameter.

There are three ways to estimate the above equation:

1. Treat y_{i1} as exogenously given and do not explain it

Table 24 Productivity Evolution

	Dependent variable:
	ω_{it}
α_1	0.693***
	(0.005)
α_2	0.083***
	(0.002)
α_3	−0.008***
	(0.0003)
α_4	0.003
	(0.007)
α_5	0.036***
	(0.005)
α_6	−0.006
	(0.009)
α_0	0.354***
	(0.007)
Observations	62,487
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 25 Cobb-Douglus coefficients

	Value	Bootstrap Standard Errors
β_L	0.299	0.006
β_K	0.429	0.020

Table 26 Productivity Evolution ACF (Continuous)

	Dependent variable:
	ω_{it}
α_1	0.781***
	(0.004)
α_2	0.062***
	(0.001)
α_3	−0.006***
	(0.0002)
α_4	0.004
	(0.008)
α_5	0.043***
	(0.006)
α_6	−0.006
	(0.009)
α_0	0.237***
	(0.007)
Observations	62,487
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 27 Cobb-Douglus coefficients ACF (Continuous)

	Value	Bootstrap Standard Errors
L_{it}	0.061	0.110
K_{it}	0.603	0.060

2. Heckman Method
3. Wooldridge Method

I use the Wooldridge method which assumes that α_i:

$$\alpha_i = \lambda y_{i1} + \tilde{\alpha_i}$$

Assumptions:

1. The i-units are a random sample
2. $\epsilon_{it} \sim N(0,1)$ is independent of x_i
3. $\tilde{\alpha_i} \sim N(0, \sigma^2_{alpha})$ is independent of x_i and ϵ_{it}

The likelihood function is given by:

$$L_i(\beta, \gamma, \sigma_\alpha, \lambda) = \int_{-\infty}^{\infty} \prod_{t=2}^{T} \Phi(x_{it}^{'}\beta + \gamma y_{it-1} + \lambda y_{i1} + \tilde{\alpha}) \frac{\phi(\tilde{\alpha})}{\sigma_{\tilde{\alpha}}} d\tilde{\alpha}$$

$$L(\beta, \gamma, \sigma_\alpha, \lambda) = \prod_{i=1}^{N} L_i(\beta, \gamma, \sigma_\alpha, \lambda : y_i, x_i$$

The marginal effects are given by:

$$\frac{\delta E[y_{it}|x_{it}, \alpha_i]}{\delta x_{it,j}} = \beta_j \phi(x_{it}^{'}\beta + \alpha_i) \tag{28}$$

I report the average marginal effects which are given by:

$$\frac{1}{N} \sum_i \phi(x_{it}^{'}\beta + \alpha_i)\hat{\beta}_j$$

7.5 Appendix (Dynamic Bivariate Probit Model)

Let the latent model be:

$$y^*_{i1,t} = x^T_{i1,t}\beta_1 + y_{i1,t-1}\gamma + \varepsilon_{i1}$$

$$y^*_{2ti} = x^T_{i2t}\beta_2 y_{i1,t-1}\gamma + \varepsilon_{i2}$$

we have the observed responses as,

$$y_{i1,t} = I(y^*_{i1,t} > 0)$$

$$y_{2i,t} = I(y^*_{2i,t} > 0)$$

The error distribution is given as follows,

$$(\varepsilon_{i1,t}, \varepsilon_{i2,t}) \sim N\left(0, \quad \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

The multivariate normal density of $f(y*_{i1,t}, y*_{i2,t})$ given the assumption above is given by:

$$\begin{aligned} f(y_{i1,t}, y_{i2}) = \frac{1}{2\pi\sigma_{y_{i1,t}}\sigma_{y_{i1,t}}\sqrt{1-\rho^2}} exp(\frac{-1}{2(1-\rho^2)}[\frac{(y_{i1,t}-\mu_{y_{i1,t}})^2}{\sigma^2_{y_{i1,t}}} + \\ \frac{(y_{i2}-\mu_{y_{i2}})^2}{\sigma^2_{y_{i2}}} - 2\rho\frac{(y_{i1,t}-\mu_{y_{i1,t}})(y_{i2}-\mu_{y_{i2}})}{\sigma_{y_{i1,t}}\sigma_{y_{i2}}}] \end{aligned} \tag{29}$$

Therefore the log-likelihood function is given by

$$\begin{aligned} L(\beta_1, \beta_2, \gamma_1, \gamma_2) = \Big(\prod P(Y_{1i,t} = 1, Y_{2i,t} = 1 \mid \beta_1, \beta_2, \gamma_1, \gamma_2)^{Y_{1i,t}Y_{2i,t}} \\ P(Y_{1i,t} = 0, Y_{2i,t} = 1 \mid \beta_1, \beta_2, \gamma_1, \gamma_2)^{(1-Y_{1i,t})Y_{2i,t}} \\ P(Y_{1i,t} = 1, Y_{2i,t} = 0 \mid \beta_1, \beta_2, \gamma_1, \gamma_2)^{Y_{1i,t}(1-Y_{2i,t})} \\ P(Y_{1i,t} = 0, Y_{2i,t} = 0 \mid \beta_1, \beta_2, \gamma_1, \gamma_2)^{(1-Y_{1i,t})(1-Y_{2i,t})}\Big) \end{aligned} \tag{30}$$

$$\begin{aligned} L(\beta_1, \beta_2) = \sum(\Phi(Y_{1i,t}Y_{2i,t} \ln\Phi(X_1\beta_1, X_2\beta_2, \rho) \\ +(1-Y_{1i,t})Y_{2i,t} \ln\Phi(-X_1\beta_1, X_2\beta_2, -\rho) \\ +Y_{1i,t}(1-Y_{2i,t}) \ln\Phi(X_1\beta_1, -X_2\beta_2, -\rho) \\ +(1-Y_{1i,t})(1-Y_{2i,t}) \ln\Phi(-X_1\beta_1, -X_2\beta_2, \rho)\Big)) \end{aligned} \tag{31}$$