

# Productivity, Size and Market Competition\*

Sudipta Ghosh

Amartya Lahiri

Swapnika Rachapalli

University of British Columbia   University of British Columbia   University of British Columbia

February 14, 2024

## Abstract

Standard theories of firm size predict a positive correlation between productivity and size. We examine plants in India and find support for this prediction only for exporting plants. For non-exporting plants we find no relationship between size and productivity. We show that this asymmetry between exporting and non-exporting plants is accounted for by product-level import tariffs. Higher tariffs reduce exposure to market competition which breaks the link between size and productivity for non-exporters. Our results illustrate the pitfalls of industrial policies that protect domestic firms from foreign competition.

**Keywords:** Productivity, size distribution, import tariffs, exports

**JEL Codes:** L11, O14, O47

## 1 Introduction

The relationship between firm size and productivity is an old topic of both academic and policy interest. Views regarding this relationship underpin public policy in disparate areas spanning industrial policy, trade policy, tax policy and credit policy. Standard models of establishment size such as Lucas (1978), Hopenhayn (1992) and Melitz (2003) predict that size should be positively correlated with productivity. Intuitively, if economic entities have scarce resources with which to organize production plans, optimal resource allocation would dictate that entities with greater productive efficiency should be allocated more productive inputs.

This insight and prediction of a positive correlation between size and productivity has been tested by many authors on diverse datasets with mixed results. Indeed, following the seminal contribution of Hsieh and Klenow (2009), deviations of empirical findings from the baseline prediction

---

\*We thank David Green and Devashish Mitra for helpful discussions. All errors are ours.

are typically viewed as evidence of resource misallocation. Work in Bartelsmann et al (2013), Garicano et al (2016) amongst others have linked such evidence of misallocations to distortions, labor policies and other frictions.

In this paper we examine the role of trade in affecting this predicted relationship. Specifically, how does openness to trade, along with the attendant competitive pressure that it brings, impact the conventionally hypothesized productivity-size relationship? This question is interesting because there is an old tradition of industrial policy directed at providing protected domestic markets for domestic producers in order allow scale based efficiency gains. This motivation for protectionism has found new proponents more recently with a rising wave of protectionist sentiment in many different countries and regions.

We examine data from the Annual Survey of Industries (ASI) for manufacturing plants in India and find that the predicted positive relationship between size and productivity only holds for plants that export their products. For non-exporting plants however we find no relationship between size and productivity. This difference in the behavior of exporting and non-exporting plants is statistically significant and robust to controlling for industry and year fixed effects.

In order to uncover the reasons for this different patterns we examine the role of market protection as measured by import tariffs. We find that controlling for product level import tariffs eliminates the difference in the size-productivity relationship between exporters and non-exporters. Importantly, higher import tariffs are associated with lower exports. We also show that exporters in our sample are exposed to greater market competition in foreign markets.

We interpret our results as providing a cautionary tale for the rising popularity of industrial policy in the form of restricting domestic market access to foreign competition. Raising tariffs tends to not only reduce exposure to foreign competition in domestic markets but also reduces exposure to international competition by also dampening the incentive to export. Both these effects weaken the relationship between productivity efficiency and firm size. A corollary of these results is that plants in protected sectors tend to grow for non-efficiency based reasons, a feature that is indicative of resource misallocations.

Our work is related to three distinct but connected bodies of work. The first is the role of resource misallocations in accounting for the size distribution of firms. This work has grown very sharply over the past decade since the contribution of Hsieh and Klenow (2009). The second literature connected to our paper is the work on import protection, scale economies and size. This literature owed its intellectual origins to the work of economists in the 1950s who provided an

intellectual foundation for import substitution policies on the grounds that such protection would allow young domestic firms to grow to some minimum efficiency scale. A description and overview of this literature can be found in Irwin (2020). A third strand of work related to this paper is the relationship between size distribution and productive efficiency of economic units of production. Examples of this work can be found in Lucas (1978) and its many followers. Our work is unique in bringing together these three strands of work in one conceptual framework.

We describe our data in the next section. Section 3 presents the main empirical results. Section 4 proposes an analytical structure to interpret the results along while the last section concludes.

## 2 Data

We use plant or “factory” level data for the period 2009-2015, which is collected by the Central Statistical Office of India in the Annual Survey of Industries (ASI). The primary unit of enumeration in the survey is a factory in the case of manufacturing industries. Data are based on returns provided by factories.

The ASI factory frame is classified into two sectors: the ‘census sector’ and the ‘sample sector’. The sample sector consists of small plants employing 20 to 199 workers if not using electricity and 10 to 199 workers if using electricity. The census sector comprises relatively large plants employing more than 200 workers. While the units in the census sector are approached for data collection on a complete enumeration basis every year, sample sector units are covered on the basis of a well designed sampling methodology. For all calculations we apply sampling weights so that our weighted sample reflects the population.

Our choice of years is based on the availability of information on exporter status and unique factory identifiers. The present study uses data on various plant level production parameters such as output, sales, labor, wage bill, capital, materials and energy. Our baseline analysis considers establishments as exporters if more than 20% of the establishments’ production share is exported<sup>1</sup>. All establishments are classified into industries based on NIC 2008 classification. We assign establishment into industries based on the main product produced by these establishments.

The variables in the ASI data that we use are the plant’s industry (two-digit and four-digit NIC), labor compensation, value-added, age (based on reported birth year), gross output, materials, energy and book value of the fixed capital stock. Specifically, the ASI reports the plant’s total wage

---

<sup>1</sup>For robustness we also consider firms as exporters if the share of exports in total sales is positive and/or  $> 10\%$ .

payments, bonus payments, and the imputed value of benefits. Our measure of labor compensation in the benchmark analysis is twages (we also use the sum of wages, bonuses, and benefits in other analyses). In addition, the ASI reports the book value of fixed capital at the beginning and end of the fiscal year net of depreciation. We take the average of the net book value of fixed capital at the beginning and end of the fiscal year as our measure of the plant’s capital. The table below provides a basic summary of coverage of the ASI data.

	Avg Gross Sales	Avg Employment	Number of Establishments
Exporters	8.55 Mn INR	391	17882
Non-Exporters	7.1 Mn INR	213	172825
All establishments	7.25 Mn INR	230	190807

Our baseline findings on the size-productivity correlation is based on TFP computed using Levinsohn-Petrin (henceforth LP) method using wage compensation as free variables and material and fuel as proxy variables. We construct seven labor size bins based on actual number of workers in each establishment.

We also use information on tariff rates for India from World Integrated Trade Solutions (WITS) from World Bank. The coverage of tariff rates at 4 digit HS classification is limited to years 2009-2013 for our sample period. So we also use information on MFN rates from the same source which has higher coverage. We match the 4-digit HS coded MFN and Tariff rates to 4-Digit CPC code derived from the product classification in ASI data (NPCMS classification). We then compute sales-share weighted average tariffs and MFN rates for each establishment which is our primary tariff measure.

After merging the ASI data with tariff measures, the summary stats of our final data is given below:

	Avg Gross Sales	Avg Employment	Number of Establishments
Exporters	4.467 Mn INR	359.5	13989
Non-Exporters	2.580 Mn INR	156	151053
All establishments	2.73 Mn INR	172	164451

### 3 Empirical Results

#### 3.1 Preliminary Observations

##### 3.1.1 TFP computed using Levihnson-Petrin Method

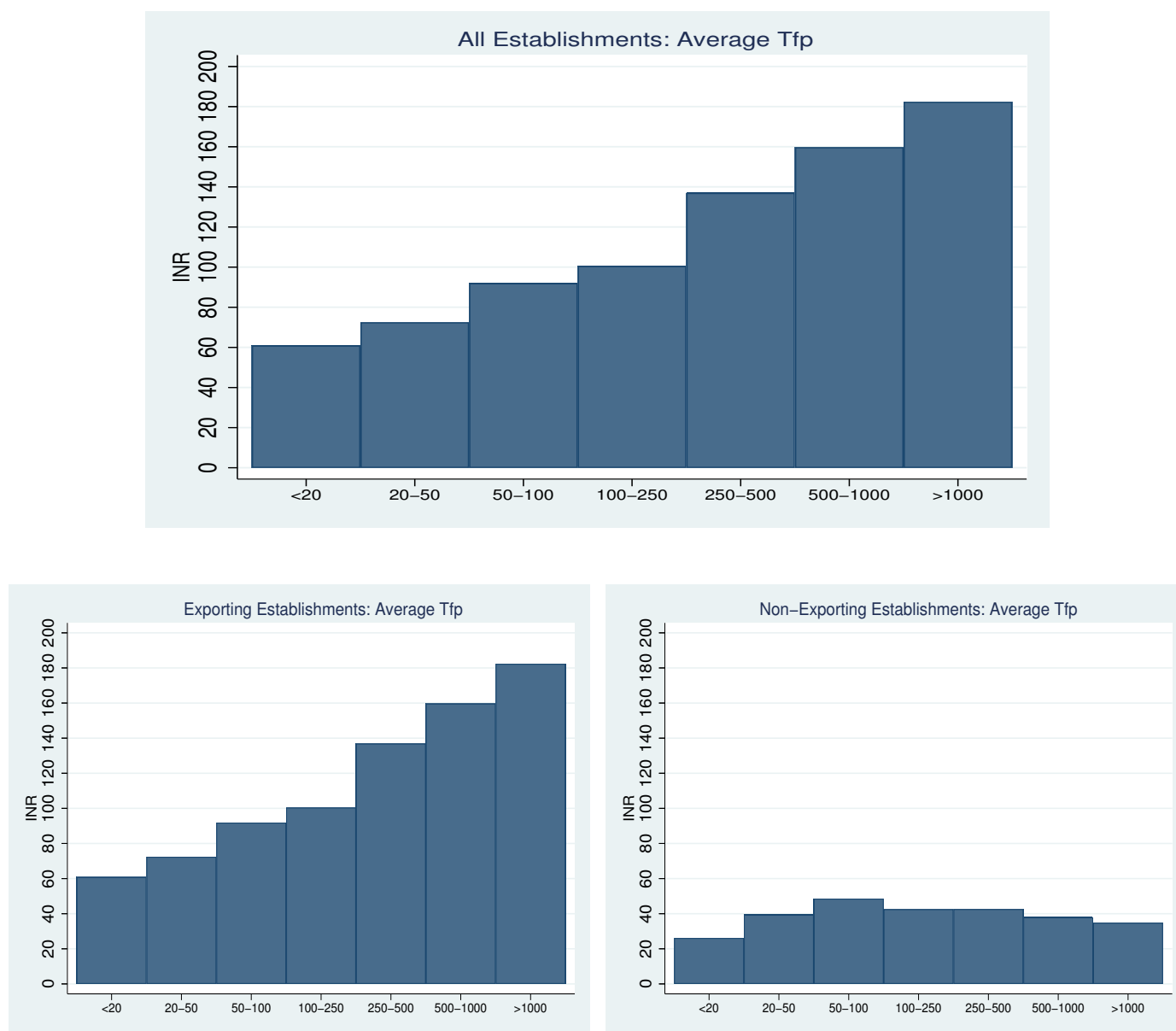


Figure 1

### 3.1.2 TFP computed using Olley-Pakes Method

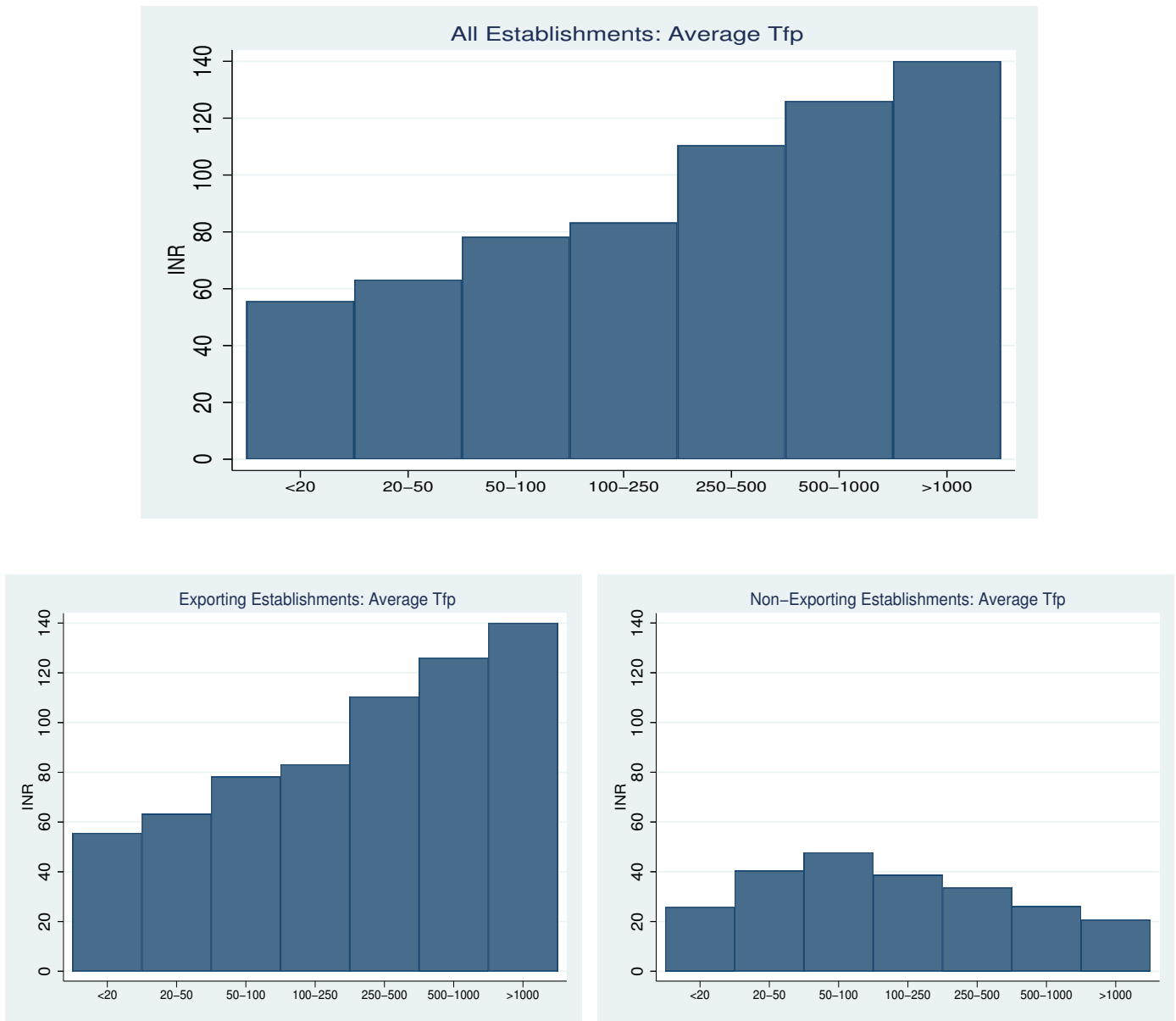


Figure 2

### 3.2 Regression: Establishment size vs Establishment productivity (LP)

$$tfp = \alpha + \sum_{j=1}^7 \beta_j labor\_category_j + \sum_{j=1}^7 \gamma_j i.export\_dummy \times i.labor\_category_j + Fixedeffects + \epsilon \quad (3.1)$$

	(1)	(2)	(3)	(4)	(5)
	tfp_all	tfp_all	tfp_all	tfp_all	tfp_all
20-50	9.426*** (21.63)	9.233*** (21.22)	8.675*** (20.17)	8.259*** (19.18)	-8.890** (-5.55)
50-100	16.50*** (29.05)	16.39*** (28.88)	16.32*** (29.06)	15.49*** (27.33)	-19.20*** (-6.20)
100-250	22.85*** (36.23)	22.35*** (35.35)	21.75*** (34.62)	20.44*** (32.35)	-31.38*** (-6.74)
250-500	29.44*** (29.93)	27.34*** (27.85)	26.10*** (26.81)	25.34*** (25.99)	-44.47*** (-7.18)
500-1000	28.35*** (20.82)	26.36*** (19.44)	25.29*** (18.85)	24.60*** (18.32)	-62.02*** (-7.99)
>1000	27.69*** (15.53)	26.16*** (14.75)	25.35*** (14.45)	23.21*** (13.31)	-72.32*** (-6.92)
Exporters	36.41*** (20.99)	37.45*** (21.68)	37.74*** (22.11)	35.79*** (21.15)	30.14*** (6.94)
Exporters X 20-50	21.75*** (9.40)	21.37*** (9.29)	22.53*** (9.93)	21.97*** (9.80)	4.465 (1.00)
Exporters X 50-100	32.65*** (13.17)	32.06*** (13.01)	32.44*** (13.35)	30.55*** (12.71)	38.54** (5.67)
Exporters X 100-250	47.74*** (19.95)	46.99*** (19.73)	48.55*** (20.68)	45.86*** (19.74)	63.01*** (13.25)
Exporters X 250-500	87.20*** (29.73)	88.10*** (30.21)	89.82*** (31.24)	87.62*** (30.82)	103.6*** (10.01)
Exporters X 500-1000	126.1*** (34.57)	127.2*** (35.08)	129.1*** (36.11)	127.7*** (36.14)	136.7*** (14.08)
Exporters X >1000	156.5*** (37.44)	157.5*** (37.89)	160.0*** (39.06)	161.7*** (39.94)	186.0*** (11.01)
Constant	21.01***	25.60***	27.62***	71.17***	47.97***

### 3.3 Regression: Establishment size vs Establishment productivity (LP) with MFN

$$\begin{aligned}
 tfp = & \alpha + \sum_{j=1}^7 \beta_j labor\_category_j + \sum_{j=1}^7 \gamma_j i.export\_dummy \times i.labor\_category_j + \\
 & \sum_{j=1}^7 \delta_j \times i.export\_dummy \times i.labor\_category_j \times c.mfn\_rates + FixedEffects + \epsilon
 \end{aligned} \tag{3.2}$$



	(1)	(2)	(3)	(4)
	tfp_all	tfp_all	tfp_all	tfp_all
20-50	5.848*** (6.48)	5.363*** (5.93)	5.338*** (5.90)	5.069*** (5.66)
50-100	10.78*** (10.45)	10.11*** (9.77)	10.07*** (9.74)	10.05*** (9.78)
100-250	14.62*** (13.83)	14.21*** (13.37)	14.19*** (13.36)	13.69*** (12.90)
250-500	17.62*** (12.30)	16.77*** (11.68)	16.72*** (11.65)	15.97*** (11.19)
500-1000	15.77*** (8.48)	15.07*** (8.11)	14.95*** (8.05)	14.66*** (7.95)
>1000	15.76*** (6.57)	15.00*** (6.26)	14.85*** (6.20)	14.73*** (6.20)
Exporters	-8.358* (-1.97)	-8.132 (-1.93)	-8.319* (-1.97)	-8.175 (-1.96)
Exporters X 20-50	45.67*** (8.92)	46.15*** (9.04)	46.43*** (9.10)	48.60*** (9.65)
Exporters X 50-100	43.57*** (8.10)	43.81*** (8.17)	44.21*** (8.25)	46.60*** (8.80)
Exporters X 100-250	52.44*** (10.04)	52.99*** (10.17)	53.37*** (10.25)	57.04*** (11.09)
Exporters X 250-500	109.5*** (18.16)	110.7*** (18.41)	111.0*** (18.46)	113.1*** (19.06)
Exporters X 500-1000	149.9*** (21.42)	151.5*** (21.72)	151.9*** (21.78)	153.3*** (22.27)
Exporters X >1000	181.8*** (24.47)	182.6*** (24.65)	182.7*** (24.67)	185.8*** (25.40)
m_mfnstariff	0.406*** (7.22)	0.409*** (7.27)	0.406*** (7.22)	0.600*** (10.30)
20-50 X m_mfnstariff	-0.00192 (-0.02)	0.0110 (0.14)	0.0139 (0.18)	-0.0284 (-0.37)
50-100 X m_mfnstariff	0.0616 (0.74)	0.0983 (1.18)	0.102 (1.23)	0.0687 (0.83)
100-250 X m_mfnstariff	-0.0364 (-0.49)	-0.00123 (-0.02)	-0.000739 (-0.01)	-0.0275 (-0.37)
250-500 X m_mfnstariff	0.0297 (0.32)	0.0430 (0.47)	0.0460 (0.50)	0.0368 (0.40)
500-1000 X m_mfnstariff	0.176 (1.33)	0.188 (1.42)	0.191 (1.45)	0.177 (1.35)
>1000 X m_mfnstariff	0.209 (1.58)	0.229 (1.72)	0.234 (1.78)	0.175 (1.35)

In the regression above, productivity is regressed on labor size categories interacted with exporter dummy and MFN rates. The simple average MFN tariff rates at 4-digit HS product classification, aggregated at establishment level. This data is not available for the years 2013 and 2014. These years are omitted from the regression.

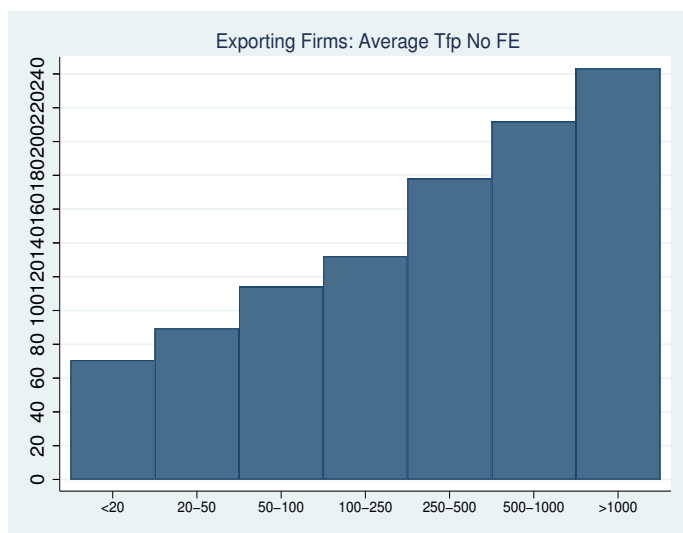
The systematically increasing size-productivity relationship holds strongly for all specifications in the size class greater than 50-100 among exporters. But the triple interaction terms imply that for exporters as MFN rates increase, this systematic increasing relationship is broken. In most size classes for exporters the productivity is falling in imposed MFN rates. For all size class among exporters, productivity is falling MFN rates. This however is not the case for non-exporters.

Here the number of observations is different from the previous table since MFN rates are not available for the years 2013 and 2014.

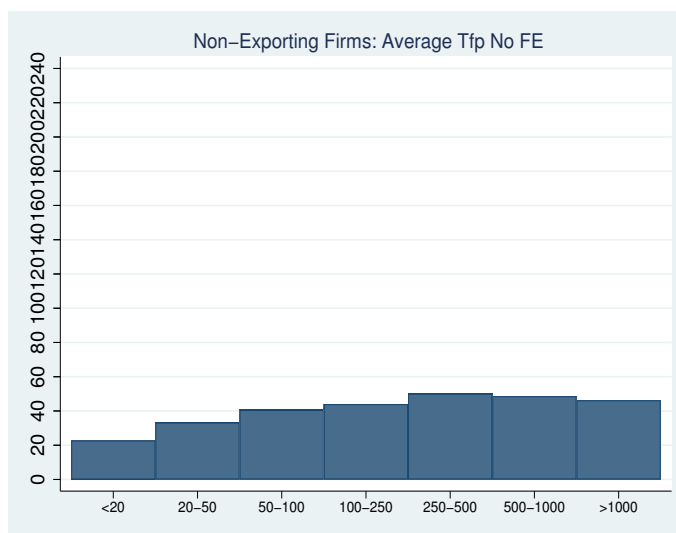
The next section has the predicted productivities from the above two regressions for both exporters and non-exporters along with the actual data plots. Section 4.4 shows the effect of MFN on productivity for both exporters and non-exporters. We see that For exporters with MFN imposed, the size productivity relationship is broken, however the size-productivity relationship in total for exporters is too strong to be countered by MFNs. Section 6.5 shows the predicted average TFP subtracted the effect of MFN. For smaller size class, the negative effect of MFN is stronger while for larger firms, this effect diminishes.

### 3.4 Implied Productivities

#### 3.4.1 No Fixed Effects

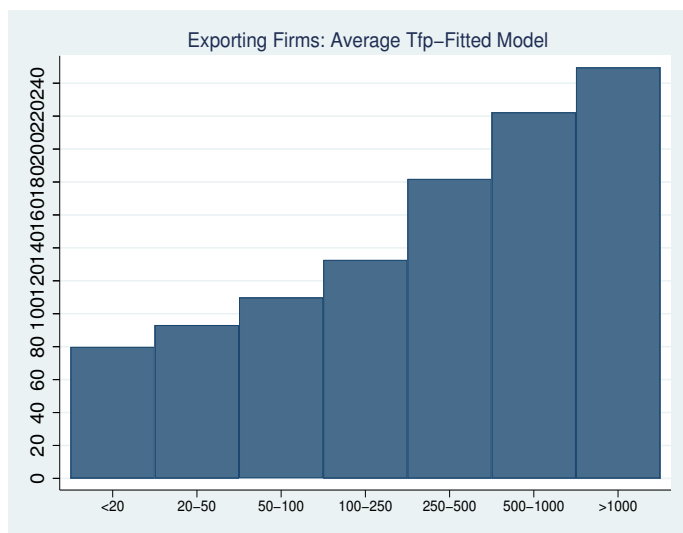


(a)

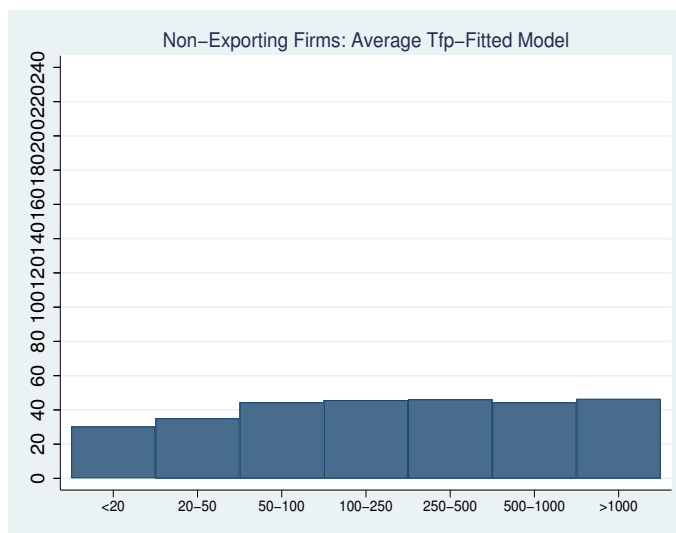


(b)

### 3.5 Without Tariffs/MFNs

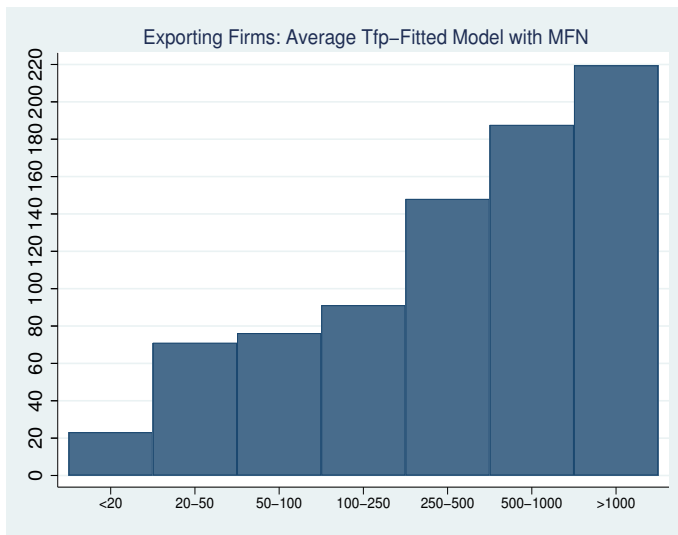


(a)

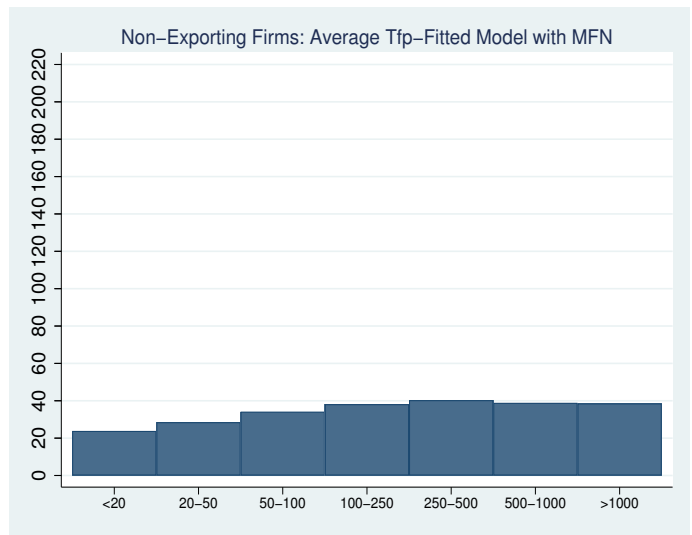


(b)

### 3.5.1 With MFN

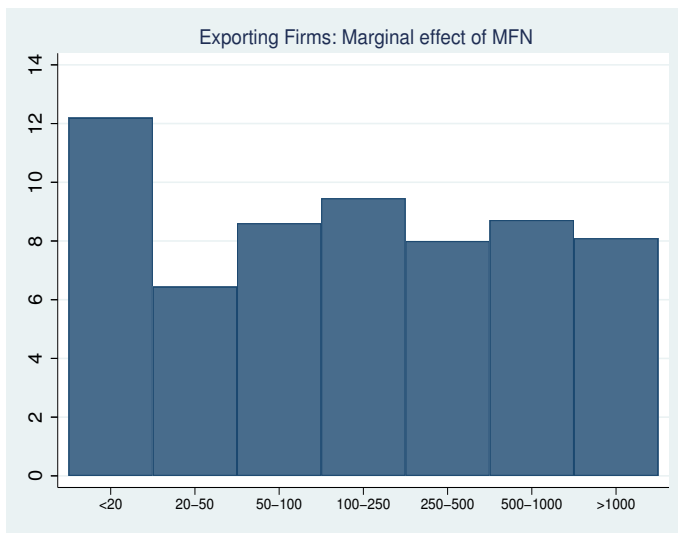


(a)

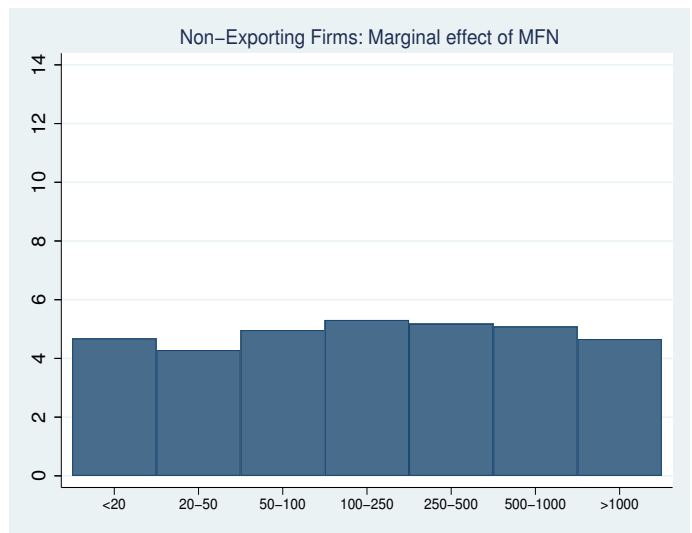


(b)

### 3.5.2 Effect of MFN



(a)



(b)

### 3.5.3 One-sided test for differences in coefficients

With state, year, 4-digit MFN rates and 2-digit NIC fixed effects :year  $\notin$  2013,2014

The one-sided test is based on the regression below:

$$\begin{aligned}
tfp = & \alpha + \sum_{j=1}^7 \beta_j labor\_category_j + \sum_{j=1}^7 \gamma_j i.export\_dummy \times i.labor\_category_j + \\
& \sum_{j=1}^7 \delta_j \times i.export\_dummy \times i.labor\_category_j \times c.mfn\_rates + IndustryFE + stateFE + yearFE + \epsilon
\end{aligned}
\tag{3.3}$$

The one sided test is testing if  $\delta_j < \delta_{j+1}$  (the triple interaction term) separately for exporting and non exporting firms. The interaction of non-exporting firms with MFN rates implies non-exporting firms which produce products which face MFN rates on imports of those products.

Hypothesis	F-value Non-Exporting Firms	F-value Exporting Firms
Labor: 20-50 -Labor:<20 = 0	-0.28	-5.53***
Labor: 50-100 -Labor:20-50 = 0	0.09	0.48***
Labor: 100-250 -Labor:50-100 = 0	-0.096	1.36***
Labor: 250-500 -Labor:100-250 = 0	0.064	-1.2**
Labor: 500-1000 -Labor:250-500 = 0	0.14	-0.1
Labor: >1000 -Labor:500-1000 = 0	-0.002	-0.08

Thus we see that for exporting firms facing positive MFN rates, the systematically increasing size-productivity relationship is lost.

### 3.6 Regression of exporting status on productivity

### 3.6.1 MFN Rates: Labor > 1000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy
exp_dummy									
ln_tfp_all	1.27*** (0.06)	1.39*** (0.08)	1.50*** (0.09)	1.39*** (0.08)	1.39*** (0.08)	1.50*** (0.09)	1.50*** (0.09)	1.39*** (0.08)	1.50*** (0.09)
age	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.01* (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.01* (0.00)	0.00 (0.00)
m_mfnstariff		-0.04*** (0.01)	-0.09*** (0.02)	-0.05*** (0.01)	-0.04*** (0.01)	-0.09*** (0.03)	-0.09*** (0.02)	-0.05*** (0.01)	-0.09*** (0.03)
_cons	-6.64*** (0.30)	-6.81*** (0.37)	-7.77*** (0.53)	-6.05*** (0.39)	-6.61*** (0.39)	-7.50*** (0.54)	-7.70*** (0.54)	-5.86*** (0.41)	-7.46*** (0.55)
Year fixed effects	NO	NO	NO	NO	YES	NO	YES	YES	YES
State fixed effects	NO	NO	NO	YES	NO	YES	NO	YES	YES
Industry fixed effects	NO	NO	YES	NO	NO	YES	YES	NO	YES
N	6437.00	4492.00	4448.00	4315.00	4492.00	4273.00	4448.00	4315.00	4273.00
r2_p	0.27	0.30	0.39	0.33	0.30	0.41	0.39	0.34	0.41

Standard errors in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 3.6.2 MFN Rates: Labor > 500

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy
exp_dummy									
ln_tfp_all	1.21*** (0.04)	1.27*** (0.05)	1.41*** (0.05)	1.30*** (0.05)	1.28*** (0.05)	1.41*** (0.06)	1.41*** (0.05)	1.30*** (0.05)	1.41*** (0.06)
age	-0.00 (0.00)	-0.01*** (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.00 (0.00)
m_mfnstariff		-0.03*** (0.00)	-0.07*** (0.01)	-0.04*** (0.01)	-0.03*** (0.00)	-0.07*** (0.01)	-0.07*** (0.01)	-0.04*** (0.01)	-0.07*** (0.01)
_cons	-6.34*** (0.19)	-6.32*** (0.22)	-7.55*** (0.28)	-6.84*** (0.60)	-6.07*** (0.22)	-7.78*** (0.66)	-7.41*** (0.29)	-6.59*** (0.61)	-7.66*** (0.67)
Year fixed effects	NO	NO	NO	NO	YES	NO	YES	YES	YES
State fixed effects	NO	NO	NO	YES	NO	YES	NO	YES	YES
Industry fixed effects	NO	NO	YES	NO	NO	YES	YES	NO	YES
N	16338.00	11276.00	11235.00	11170.00	11276.00	11129.00	11235.00	11170.00	11129.00
r2_p	0.25	0.27	0.37	0.32	0.27	0.40	0.37	0.32	0.40

Standard errors in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 3.6.3 MFN Rates: Labor > 250

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy
exp_dummy									
ln_tfp_all	1.05*** (0.03)	1.10*** (0.03)	1.23*** (0.04)	1.14*** (0.04)	1.11*** (0.03)	1.23*** (0.04)	1.23*** (0.04)	1.14*** (0.04)	1.23*** (0.04)
age	-0.00** (0.00)	-0.01*** (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.00* (0.00)
m_mfnstariff		-0.03*** (0.00)	-0.06*** (0.01)	-0.04*** (0.00)	-0.03*** (0.00)	-0.06*** (0.01)	-0.06*** (0.01)	-0.04*** (0.00)	-0.06*** (0.01)
_cons	-5.71*** (0.14)	-5.54*** (0.16)	-6.80*** (0.21)	-7.57*** (0.71)	-5.32*** (0.17)	-8.70*** (0.69)	-6.66*** (0.22)	-7.39*** (0.72)	-8.57*** (0.68)
Year fixed effects	NO	NO	NO	NO	YES	NO	YES	YES	YES
State fixed effects	NO	NO	NO	YES	NO	YES	NO	YES	YES
Industry fixed effects	NO	NO	YES	NO	NO	YES	YES	NO	YES
N	33800.00	22635.00	22619.00	22506.00	22635.00	22490.00	22619.00	22506.00	22490.00
r2_p	0.20	0.22	0.33	0.28	0.23	0.36	0.33	0.28	0.36

Standard errors in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



### 3.6.4 MFN Rates: Labor > 100

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy	exp_dummy
exp_dummy									
ln_tfp_all	0.86*** (0.02)	0.97*** (0.03)	1.08*** (0.03)	1.00*** (0.03)	0.97*** (0.03)	1.09*** (0.03)	1.09*** (0.03)	1.01*** (0.03)	1.09*** (0.03)
age	-0.00*** (0.00)	-0.01*** (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.00* (0.00)
m_mfnstariff		-0.03*** (0.00)	-0.05*** (0.01)	-0.03*** (0.00)	-0.03*** (0.00)	-0.05*** (0.01)	-0.04*** (0.01)	-0.03*** (0.00)	-0.05*** (0.01)
_cons	-4.99*** (0.10)	-4.96*** (0.13)	-6.04*** (0.18)	-7.76*** (0.59)	-4.72*** (0.14)	-8.63*** (0.58)	-5.84*** (0.19)	-7.56*** (0.59)	-8.47*** (0.58)
Year fixed effects	NO	NO	NO	NO	YES	NO	YES	YES	YES
State fixed effects	NO	NO	NO	YES	NO	YES	NO	YES	YES
Industry fixed effects	NO	NO	YES	NO	NO	YES	YES	NO	YES
N	62895.00	40162.00	40128.00	39685.00	40162.00	39651.00	40128.00	39685.00	39651.00
r2_p	0.15	0.18	0.29	0.24	0.18	0.32	0.30	0.24	0.33

Standard errors in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 3.7 Size productivity correlation

Exporter status	Labor size class	State, Year and NIC 2 digit FE
Exporter	0.259	Yes
Non-Exporter	0.158	Yes

Results are invariant to 4 digit Industry fixed effects and to firm fixed effects.