SAJE South African Journal of Economics
ESSA The Economic Society of South Africa (Est. 1925)
South African Journal of Economics Vol. 86:S1 January 2018

TOTAL FACTOR PRODUCTIVITY IN SOUTH AFRICAN MANUFACTURING FIRMS

CARL FRIEDRICH KREUSER† AND CAROL NEWMAN*

Abstract
We use firm-level data for the period 2010-2013 to estimate total factor productivity in the South African manufacturing sector. We examine differences in the level and growth of productivity across manufacturing subsectors and examine the heterogeneity in productivity levels within sectors. We find that productivity grew in most subsectors but that there is heterogeneity across subsectors in the pace of growth. We find that firm size is positively correlated with productivity and its growth rate. We also find that there is a productivity premium associated with engaging in R&D and international trade.
JEL Classification: D22, D24, O12
Keywords: total factor productivity, South Africa, heterogeneity, tax administration data

1. INTRODUCTION

The importance of manufacturing and industry in economic growth is well documented in the Kaldorian tradition.[1] Manufacturing firms are generally more productive than firms in the agricultural or services sectors and are an important source of job creation.[2] Several studies support the primacy of the manufacturing sector in determining output growth and employment creation (Wells and Thirlwall, 2003; Millin and Nichola, 2005; Li and Zhang, 2008; Mahmood *et al.*, 2014).[3]

The manufacturing sector in South Africa accounts for around 13% of gross domestic product (GDP). This is similar to Brazil (14%) but lower than Russia (16%) and India (17%), and significantly lower than China (30%).[4] As evident from Fig. 1, the contribution of the sector to GDP has been in decline over the last two decades. Although the contribution of the sector to output is also declining in the BRICS countries overall, the pace of decline has been faster in the case of South Africa. Growth in the sector has been

* Corresponding author: Professor, Department of Economics, Trinity College Dublin, College Green, Dublin 2, Ireland. E-mail: cnewman@tcd.ie
† Economics, Stellenbosch University
1 See Targetti (2005) for a discussion of Kaldor's (1967) contributions to development economics.
2 Recent work by Newman *et al.* (2016) highlights the fact that other sectors of the economy, such as traded services and agri-business, also hold similar potential for output, employment and productivity growth.
3 Specifically, Millin and Nichola (2005) examine the role of the manufacturing sector in South Africa using data from 1946 to 1998. Their findings, however, do not relate to the post-Apartheid period.
4 These figures are from the World Bank's World Development Indicators database (WDI, 2016).

doi: 10.1111/saje.12179

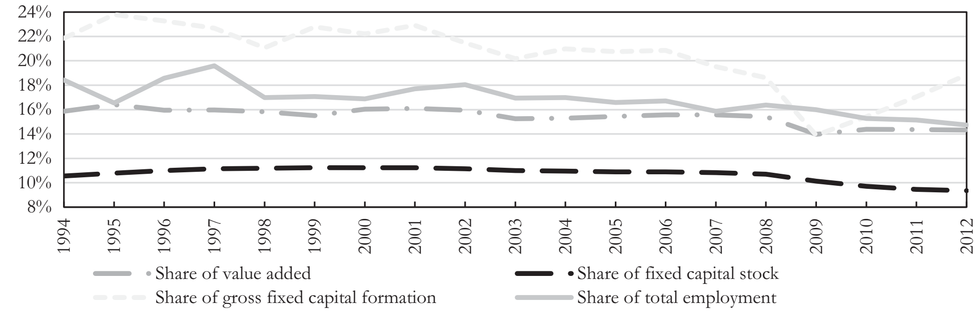

Figure 1. Manufacturing industry's share in gross value added, fixed capital stock, fixed capital stock formation and total employment
Note: All values are at constant 2010 prices. Note that the primary axis is on a logarithmic scale. Employment figures are obtained from the Post-Apartheid Labour Market Series (1994–2012) using cross-entropy weights.
Source: Authors' calculations based on DataFirst (1994−2012) and SARB (2014).

slow relative to the rest of the economy and the sector was also hit especially hard by the recent global financial crisis. This decline is reflected in the employment and investment numbers, with the sector representing a decreasing proportion of the total numbers employed and the proportion of total fixed capital formation. Moreover, Aghion *et al.* (2008) highlight the poor productivity performance of the sector when compared to manufacturing internationally and show that higher mark-ups in the sector, due to low product market competition, has had a negative impact on productivity growth in manufacturing.

These trends are worrying if the manufacturing sector is an important source of employment and productivity growth. Indeed, Tregenna (2008) shows the importance of the South African manufacturing sector as a source of demand for the service sector and argues that declines in manufacturing growth may have had a negative impact on economic growth. This conclusion is supported by Rodrik (2008), who attributes South Africa's slow growth and slow employment growth to weakness in manufacturing exports.[5] To address these concerns, recent policy efforts in South Africa have focused on employment-intensive economic growth with particular emphasis on the role of exporters and the manufacturing sector in creating jobs, specifically for low-skilled workers (DTI, 2010; DED, 2011; NPC, 2011).[6]

[5] There is some disagreement on the extent to which the decline in the manufacturing sector is of concern for economic growth. For example, Fedderke (2014) argues that the decline in employment in manufacturing in South Africa is partially due to high total factor productivity (TFP) growth in the sector and shows that TFP growth in manufacturing is high relative to the rest of the economy. Moreover, growth in South Africa has been more in the service and tertiary sectors, and so focusing on the manufacturing sector alone only gives a partial picture.

[6] This is evident in the Department of Trade and Industry's Industrial Policy Action Plan (DTI, 2010), the Department of Economic Development's New Growth Path (DED, 2011) and the National Planning Commission's National Development Plan (NPC, 2011; see also Black and Gerwel, 2014)

Given the importance of manufacturing in the South African context, understanding the productivity performance of the sector and its drivers is important in designing policies to promote the growth of the sector. The recent availability of firm-level tax administration data through collaboration between the South African National Treasury, United Nations University World Institute for Development Economics, and South African Revenue Services (SARS) has made it possible to accurately measure and analyse the productivity of manufacturing firms in South Africa for the first time. Indeed, the scarcity of South African firm-level studies has thus far been primarily driven by a lack of data (Bhorat and Lundall, 2004; Behar, 2010). Providing accurate and robust measures of total factor productivity (TFP) at the firm level allows for a comparison of productivity distributions and trajectories across manufacturing subsectors. It also allows for a better understanding with respect to the heterogeneity in productivity levels within sectors and its relationship with government policies, local labour markets and exposure to international markets, among others.

In this paper, we use the methodology of Ackerberg *et al.* (2006) on the South African administrative database to analyse the evolution of productivity in South Africa's manufacturing sector. We find that productivity grew on average between 2010 and 2013 but that there are some sectors that have seen productivity decline. We also find that productivity growth is largely driven by the firms that are most productive at the start of the period. We find significant heterogeneity in productivity within and between sectors. We find that TFP is increasing in the size category of the firm and that TFP growth is driven by the larger firms in terms of number of employees. We find that older firms are generally more productive as are firms that are engaged in international trade. We also consider the correlation between productivity and research and development (R&D) and find a positive relationship between R&D expenditure and TFP.

The rest of the paper is structured as follows. In Section 2, we describe in detail the data. In Section 3, we outline our approach to estimating productivity and present our core results. Section 4 presents a simple analysis of the factors related to productivity. Section 5 concludes.

2. DATA

We use tax administrative data obtained from SARS for the 2010–2013 period. The primary data source is the South African Corporate Income Tax (CIT) data. CIT data are collected by SARS annually with respect to the tax year that ends at the end of February each year. Firms are required to submit a corporate income tax return where they self-report items with respect to income, expenditures, equity and liabilities, capital items and tax credits. Almost all reporting items are compulsory, although firms are allowed to submit a "zero" where a specific field is not applicable to them. Firms are aware that they may be audited by SARS but do not know in any given year whether they will be selected for audit. More details on the steps involved in compiling the CIT database are provided in Supporting Information Appendix A.

The CIT database does not include information on the number of persons employed in the firm. We use employee income tax certificates (IRP5 forms) to construct a measure of labour employed by each firm (see Supporting Information Appendix A for more details). IRP5 data are aggregated for each Pay-As-You-Earn (PAYE) reference number. A table linking the PAYE reference numbers to the tax reference number of the firm in the

Table 1. Number of firms per year

	2010	2011	2012	2013
All firms	57,922	58,852	58,181	56,961
With sales	43,032	43,312	41,727	38,858
With value added	39,046	38,615	37,152	34,512
With capital	37,240	36,904	35,529	33,057
With labour	31,604	32,230	30,978	30,140
With lags	18,586	16,094	16,496	18,328
Sample	18,444	16,019	16,405	18,239

Source: Authors' calculations based on CIT-IRP5 data.

CIT dataset is used to match employees to firms. Companies, identified by a unique tax reference number, may have multiple PAYE numbers. We match all employees with a matching PAYE reference number to their corresponding tax reference number.

In Table 1, the number of firms in the CIT dataset belonging to the manufacturing sector, classified using either the firm's profit code[7] or the firm's industry code from the IRP5 data, is reported. A core set of variables on firms is required to estimate productivity. Table 1 documents the loss of observations due to missing data on these core variables. Several observations are lost when restricting the firms to those with positive and non-missing sales, value added and capital data. The availability of firms with labour data is indicative of an average matching rate of around 85% of viable firms. A constraint of the procedure we use to estimate productivity is that it requires the use of lags. This means that firms must be present in at least two consecutive periods for it to be included for analysis. Moreover, they must not be missing lagged values for the variables of interest. Restricting the sample to those that satisfy these criteria leads to just less than half of viable firms falling out of the sample each year. Further, we cut the top and bottom 1% of firms with respect to the value added to capital ratio in each year to eliminate outliers.

For our analysis, we classify firms according to the industry code reported by their employees in IRP5. We use the IRP5 classification because the two main industry classifiers of the CIT data, the industry code and profit code of the firm, are noisy and often contradictory. We use the industry code recorded for most employees of the firm in the IRP5 data. We convert these codes to the 2-digit level of the fourth revision of the International Standard Industrial Classification of All Economic Activities (ISIC4) (United Nations, 2008).

In calculating the total number of employees of the firm, each employee is weighted by the total number of periods they work at the firm. In Table 2, we compare the employment numbers based on our sample with those reported in the Quarterly Economic Survey (QES). As revealed, employment figures of manufacturing firms in our full sample is higher than the total employment figures reported in the QES. Once we restrict our sample to firms with sales and capital data, employment numbers in our sample drop to around 80% of the QES. In our restricted sample, where we only use firms that have the necessary data to compute TFP, employment numbers are between 40 and 50% of those reported in the QES data. It should be noted, however, that the QES data

[7] Note that the firm's profit code comes from the CIT data. This profit code is not used as a measure of industry in general as it is too inconsistent with other measures of the firm's industry. All results presented are robust to using the profit code as classifier, however.

Table 2. Total employment in sample

Data availability	2010	2011	2012	2013
All firms	1,424,450	1,491,902	1,529,811	1,591,282
Non-missing	909,066	1,013,459	1,031,075	1,040,918
Sample[1]	487,922	462,788	496,694	566,185
% of non-missing data	53.67	45.66	48.17	54.39
% of QES labour	40.69	39.78	43.09	49.31
QES	1,199,000	1,163,250	1,152,750	1,148,250

Notes: Quarterly Economic Survey (QES) employment figure is the average for the period ending in March of a given year starting from June in the previous year.
Source: Authors' calculations based on CIT-IRP5 and QES (StatsSA, 2010−2015b).
[1]The sample drops firms for which lags for any variable are not available at least one year prior to the year in question.

Table 3. Total fixed assets in millions of rands

	2010	2011	2012	2013
All firms	171,868	226,455	232,004	235,002
Non-missing	89,244	153,995	208,324	208,878
Sample[1]	73,387	70,124	77,523	89,236
% of non-missing data	82.23	45.54	37.21	42.72
% of QFS capital	23.42	19.86	20.97	23.85
QFS	313,372	353,134	369,648	374,212

Notes: The Quarterly Financial Statistics (QFS) book values are averages for the period ending in March in a given year and starting in June.
Source: Authors' calculations based on CIT-IRP5 and QFS (StatsSA, 2010−2015a).
[1]The CIT-IRP5 sample excludes firms that manufacture tobacco products.

are computed on the basis of payroll data of value-added tax (VAT) registered firms only, which are, in general, larger (StatsSA, 2015).

Fixed capital is taken from the CIT data (details are provided in Supporting Information Appendix A). We add the reported depreciation of the firm to the fixed asset value to get the value of fixed assets at the beginning of the year. We deflate this value using the manufacturing industry fixed capital investment deflators rebased to March 2012.[8] Lumpiness in fixed assets is controlled for using the two-year average of total assets.[9]

We compare our measures of fixed capital to the Quarterly Financial Statistics (QFS) collected by SARS and drawn from a sample of approximately 5,000 VAT-paying enterprises. The sample is drawn each year from a population of enterprises that account for around 95% of the total turnover per industry with adjustments made to account for the remainder (StatsSA, 2010−2015a). Comparing our total fixed assets measures with those in the QFS in Table 3 reveals that the sample covers between 19 and 24% of fixed assets reported for manufacturing firms in the economy as a whole. Although these figures are substantially smaller than those reported in the QFS, the fact that the QFS only surveys VAT registered entities suggests that they may be over estimating total capital stock as these entities are, in general, larger than non-VAT registered entities (Pieterse *et al.*, 2016).

[8] This is gross fixed capital formation in manufacturing (SARB KBP6082; see SARB, 2014). A caveat of using an industry-wide capital deflator is that it does not capture possible differences between sectors, or within sectors between firms, in capital goods inflation. In the absence of firms-specific or sector-specific capital goods prices this is unavoidable.
[9] This is the approach adopted by Hsieh and Klenow (2009) in estimating productivity for the United States, China and India.

Table 4. Value added in millions of rands

	2010	**2011**	**2012**	**2013**
All firms: value added	355,313	512,092	549,837	490,419
Sales	1,251,389	1,896,772	2,098,981	2,135,804
Cost of sales[1]	896,076	1,384,681	1,549,145	1,645,385
Non-missing: value added	269,279	378,021	387,067	387,553
Sales	1,019,126	1,549,420	1,674,806	1,837,568
Cost of sales	749,847	1,171,399	1,287,739	1,450,015
Sample[2] value added	134,385	140,612	159,150	181,827
Sales	505,448	519,309	608,508	716,484
Cost of sales	371,063	378,696	449,358	534,657
% of non-missing data	49.91	37.20	41.12	46.92
% of QFS value added	23.99	24.63	25.49	27.40
QFS[3]	560,092	570,961	624,390	663,487
Sales	1,534,438	1,683,978	1,944,606	1,989,060
Cost of sales	974,346	1,113,017	1,320,216	1,325,573

[1]All "Cost of sales" values are subtracted from "Sales."
[2]The sample drops firms for which lags for any variable are not available at least one year prior to the year in question.
[3]Turnover purchases.
Source: Authors' calculations based on CIT-IRP5 and QFS data (StatsSA, 2010−2015a).

Value added is computed as total sales minus the cost of sales (details are provided in Supporting Information Appendix A). Value added is deflated by the value added at basic prices deflator (SARB KBP6634; see SARB, 2014).[10] A comparison between the total value added of firms in our sample and those in the QFS is provided in Table 4. Similar to fixed assets, our sample covers around 24% of the total value added reported in the QFS estimates.

In Table 5, the total number of firms in our sample by industry is provided. Motor vehicles, trailers and semitrailers, fabricated metal products and other manufacturing are the largest manufacturing industries in terms of number of firms.[11] The smallest manufacturing industries, in terms of number of firms, include basic pharmaceutical products, beverages, leather and related products and computer, electronic and optical products.

We consider both weighted and unweighted estimates of productivity. In the weighted specification, for each industry j, we weight output (value added), capital, labour and intermediates (cost of sales) by the proportion of sales that firm i in time t contributes to total sales of all firms in industry j for the entire period in question, as in equation (1). This ensures that our TFP estimates give more weight to larger firms (in terms of sales) and so are a better representation of the manufacturing output of the sector.

$$\text{weight}_{ijt} = \frac{\text{sales}_{ijt}}{\sum_{i=1}^{n} \text{sales}_{ijt}}. \tag{1}$$

In Table 6, the mean log value added per firm in each industry and year is provided in its weighted and unweighted form. Firms producing pharmaceutical products, rubber

[10] Given that we only have an industry-level price deflator, price differences between firms within sectors and across sectors will not be controlled for.

[11] It should be noted that the motor vehicles sector also includes firms that manufacture parts and accessories for motor vehicles.

Table 5. Number of firms by industry

Industry	**2010**	**2011**	**2012**	**2013**	**Total**
10: Food	1,651	1,340	1,392	1,598	5,981
11: Beverages	191	159	161	189	700
13: Textiles	538	469	481	532	2,020
14: Wearing apparel	449	352	354	394	1,549
15: Leather	231	215	212	216	874
16: Wood	423	366	372	402	1,563
17: Paper	317	295	299	331	1,242
18: Printing	1,037	935	967	1,057	3,996
19: Coke and refined petroleum	458	371	371	417	1,617
20: Chemicals	806	735	763	823	3,127
21: Pharmaceuticals	74	62	57	77	270
22: Rubber and plastics	482	463	494	535	1,974
23: Other minerals	667	570	581	642	2,460
24: Basic metals	805	705	741	798	3,049
25: Fabricated metals	1,947	1,876	1,935	2,153	7,911
26: Computer, electronic and optical products	257	233	246	288	1,024
27: Electrical equipment	398	376	370	435	1,579
28: Machinery n.e.c.	2,009	1,704	1,722	1,949	7,384
29: Motor vehicles	2,679	2,310	2,370	2,624	9,983
30: Transport equipment	318	241	253	304	1,116
31: Furniture	592	520	550	598	2,260
32: Other manufacturing	2,115	1,722	1,714	1,877	7,428
Total	18,444	16,019	16,405	18,239	69,107

Note: Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on CIT-IRP5 data.

and plastics products and coke and refined petroleum products report the highest unweighted average value added per firm, whereas firms producing furniture, wearing apparel or textiles report the lowest average value added per firm. Weighting the values as described change the rankings dramatically. The average value added per firm is still the highest among firms producing pharmaceutical products, whereas firms producing leather products and paper and paper products have the second and third highest weighted values, respectively. Firms producing food products, machinery and equipment not elsewhere classified, and other manufacturing products have the lowest weighted average value added per firm.

In Table 7, we report the sample statistics for our instrument, cost of sales. Firms in the production of coke and refined petroleum products, pharmaceuticals and beverages report the highest unweighted cost of sales per firm on average, whereas firms in the production of printing and recorded media, furniture and apparel report the lowest unweighted cost of sales per firm on average. We see dramatic shifts when weighting firms by sales contribution. Whereas firms in the production of coke and refined petroleum products and pharmaceuticals still have the highest average cost of sales per firm, those manufacturing leather products have higher average cost of sales than firms producing beverages. Interestingly, firms in the production of motor vehicles, other machinery and printing have the lowest average cost of sales per firm.

In terms of average capital stock (real values in logs), Table 8 reveals that firms in the production of beverages, pharmaceuticals or rubber and plastics products have the highest capital stock on average, whereas firms manufacturing apparel, motor vehicles and other transport equipment and leather products have the lowest values on average. In terms of weighted averages, firms producing pharmaceuticals, beverages and paper products have the highest average capital stock, whereas firms producing motor vehicles, other manufacturing products and other machinery have the lowest values.

Table 6. Sample statistics for value added

Industry	2010 Weighted	2010 Unweighted	2011 Weighted	2011 Unweighted	2012 Weighted	2012 Unweighted	2013 Weighted	2013 Unweighted
10: Food	4.11 (2.62)	14.78 (1.26)	4.41 (2.72)	14.94 (1.32)	4.5 (2.8)	14.97 (1.35)	4.48 (2.87)	14.96 (1.39)
11: Beverages	5.92 (3)	14.7 (1.52)	6.48 (3.1)	15.01 (1.6)	6.61 (3.23)	15.13 (1.65)	6.57 (3.21)	15.03 (1.65)
13: Textiles	5.31 (2.63)	14.47 (1.28)	5.66 (2.71)	14.65 (1.32)	5.75 (2.8)	14.69 (1.37)	5.74 (2.74)	14.67 (1.35)
14: Wearing apparel	5.34 (2.86)	14.43 (1.36)	5.58 (2.86)	14.56 (1.36)	5.61 (2.81)	14.58 (1.35)	5.57 (2.83)	14.56 (1.36)
15: Leather	6.51 (2.69)	14.73 (1.32)	6.82 (2.65)	14.9 (1.29)	6.95 (2.7)	14.94 (1.33)	6.89 (2.73)	14.92 (1.33)
16: Wood	5.75 (2.53)	14.56 (1.25)	5.9 (2.5)	14.64 (1.21)	5.97 (2.67)	14.67 (1.33)	5.96 (2.62)	14.68 (1.28)
17: Paper	6.16 (2.53)	14.9 (1.24)	6.54 (2.59)	15.09 (1.28)	6.7 (2.52)	15.16 (1.22)	6.68 (2.7)	15.13 (1.32)
18: Printing	4.59 (2.54)	14.52 (1.25)	4.87 (2.51)	14.67 (1.23)	4.97 (2.57)	14.71 (1.27)	4.9 (2.62)	14.68 (1.29)
19: Coke and refined petroleum	6.07 (2)	15.03 (0.94)	6.3 (2.06)	15.14 (0.96)	6.35 (2.14)	15.14 (0.98)	6.39 (2.21)	15.13 (1.03)
20: Chemicals	5.07 (2.76)	14.88 (1.33)	5.4 (2.7)	15.05 (1.32)	5.51 (2.76)	15.1 (1.33)	5.45 (2.83)	15.06 (1.38)
21: Pharmaceuticals	7.42 (3.21)	15.66 (1.63)	8.02 (3.38)	16 (1.75)	8.27 (3.32)	16.12 (1.72)	7.9 (3.43)	15.9 (1.76)
22: Rubber and plastics	6.07 (2.41)	15.03 (1.19)	6.32 (2.46)	15.17 (1.21)	6.47 (2.47)	15.24 (1.22)	6.41 (2.57)	15.21 (1.3)
23: Other minerals	5.34 (2.39)	14.66 (1.17)	5.52 (2.34)	14.75 (1.15)	5.59 (2.4)	14.77 (1.18)	5.7 (2.48)	14.84 (1.22)
24: Basic metals	4.82 (2.77)	14.94 (1.34)	5.01 (2.84)	15.03 (1.38)	5.28 (2.92)	15.16 (1.41)	5.4 (2.87)	15.22 (1.39)
25: Fabricated metals	4.46 (2.39)	14.84 (1.15)	4.72 (2.44)	14.99 (1.17)	4.86 (2.5)	15.05 (1.2)	4.91 (2.54)	15.07 (1.24)
26: Computer, electronic and optical products	6.28 (2.4)	14.85 (1.18)	6.48 (2.41)	14.97 (1.18)	6.63 (2.4)	15.02 (1.15)	6.63 (2.5)	15.03 (1.19)
27: Electrical equipment	5.62 (2.48)	14.73 (1.17)	5.94 (2.38)	14.88 (1.15)	6.11 (2.39)	14.97 (1.15)	6.04 (2.41)	14.93 (1.16)
28: Machinery n.e.c.	4.18 (2.43)	14.66 (1.18)	4.49 (2.44)	14.81 (1.18)	4.63 (2.51)	14.88 (1.22)	4.69 (2.6)	14.9 (1.27)
29: Motor vehicles	2.99 (2.69)	14.53 (1.22)	3.29 (2.74)	14.68 (1.23)	3.44 (2.77)	14.74 (1.24)	3.42 (2.82)	14.72 (1.27)
30: Transport equipment	5.97 (2.5)	14.69 (1.29)	6.31 (2.68)	14.89 (1.36)	6.63 (2.77)	15.01 (1.38)	6.84 (2.73)	15.15 (1.36)
31: Furniture	5.14 (2.41)	14.36 (1.19)	5.43 (2.47)	14.52 (1.24)	5.34 (2.55)	14.48 (1.26)	5.44 (2.54)	14.54 (1.26)
32: Other manufacturing	4.26 (2.64)	14.77 (1.29)	4.55 (2.6)	14.93 (1.27)	4.66 (2.71)	14.97 (1.32)	4.69 (2.8)	14.98 (1.36)
Total average	4.58 (2.73)	14.71 (1.25)	4.88 (2.75)	14.86 (1.26)	5 (2.81)	14.91 (1.29)	5.01 (2.86)	14.92 (1.32)

Notes: Standard errors in parentheses. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on CIT-IRP5 data.

Table 7. Sample statistics for cost of sales

	2010		2011		2012		2013	
Industry	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
10: Food	4.88 (3.028)	15.55 (1.621)	5.156 (3.101)	15.686 (1.651)	5.309 (3.194)	15.771 (1.696)	5.26 (3.267)	15.741 (1.733)
11: Beverages	6.855 (3.276)	15.642 (1.732)	7.354 (3.305)	15.886 (1.731)	7.288 (3.522)	15.807 (1.882)	7.459 (3.51)	15.921 (1.869)
13: Textiles	5.737 (2.978)	14.899 (1.615)	6.05 (3.1)	15.043 (1.693)	6.145 (3.176)	15.09 (1.726)	6.173 (3.12)	15.102 (1.709)
14: Wearing apparel	5.735 (3.353)	14.824 (1.834)	5.934 (3.351)	14.915 (1.853)	5.938 (3.266)	14.913 (1.798)	5.895 (3.345)	14.877 (1.862)
15: Leather	6.996 (2.969)	15.216 (1.579)	7.23 (2.989)	15.316 (1.607)	7.442 (3.021)	15.433 (1.616)	7.292 (3.161)	15.325 (1.757)
16: Wood	6.148 (2.869)	14.956 (1.565)	6.291 (2.865)	15.033 (1.553)	6.355 (3.001)	15.052 (1.62)	6.306 (3.032)	15.02 (1.687)
17: Paper	6.777 (2.843)	15.515 (1.524)	7.145 (2.896)	15.696 (1.555)	7.352 (2.881)	15.808 (1.555)	7.373 (3.066)	15.819 (1.656)
18: Printing	4.516 (2.798)	14.447 (1.522)	4.799 (2.802)	14.592 (1.523)	4.864 (2.864)	14.611 (1.572)	4.772 (2.94)	14.559 (1.613)
19: Coke and refined petroleum	8.117 (2.45)	17.07 (1.318)	8.334 (2.578)	17.173 (1.391)	8.43 (2.729)	17.217 (1.487)	8.507 (2.757)	17.256 (1.487)
20: Chemicals	5.456 (3.214)	15.264 (1.766)	5.747 (3.164)	15.398 (1.757)	5.855 (3.239)	15.448 (1.796)	5.844 (3.3)	15.451 (1.826)
21: Pharmaceuticals	7.751 (3.35)	15.999 (1.756)	8.198 (3.39)	16.178 (1.774)	8.538 (3.283)	16.385 (1.675)	8.215 (3.606)	16.216 (1.914)
22: Rubber and plastics	6.567 (2.755)	15.532 (1.509)	6.796 (2.748)	15.642 (1.476)	6.956 (2.785)	15.723 (1.499)	6.847 (2.847)	15.65 (1.555)
23: Other minerals	5.743 (2.811)	15.057 (1.55)	5.914 (2.74)	15.144 (1.506)	6.016 (2.77)	15.203 (1.525)	6.09 (2.875)	15.23 (1.578)
24: Basic metals	5.368 (3.191)	15.491 (1.721)	5.565 (3.281)	15.588 (1.769)	5.85 (3.371)	15.73 (1.816)	5.946 (3.329)	15.775 (1.792)
25: Fabricated metals	4.716 (2.83)	15.098 (1.566)	4.925 (2.893)	15.189 (1.608)	5.086 (2.943)	15.273 (1.63)	5.112 (2.97)	15.278 (1.663)
26: Computer, electronic and optical products	6.358 (2.832)	14.927 (1.595)	6.479 (2.827)	14.971 (1.595)	6.729 (2.907)	15.118 (1.645)	6.678 (3.01)	15.083 (1.688)
27: Electrical equipment	5.936 (2.959)	15.043 (1.634)	6.294 (2.753)	15.235 (1.505)	6.432 (2.765)	15.3 (1.504)	6.373 (2.819)	15.267 (1.544)
28: Machinery n.e.c.	4.369 (2.887)	14.849 (1.637)	4.7 (2.863)	15.026 (1.604)	4.87 (2.926)	15.115 (1.634)	4.925 (3.024)	15.14 (1.686)
29: Motor vehicles	3.69 (3.337)	15.23 (1.819)	3.996 (3.421)	15.382 (1.861)	4.162 (3.462)	15.467 (1.881)	4.161 (3.543)	15.464 (1.928)
30: Transport equipment	6.279 (2.883)	14.999 (1.658)	6.516 (3.131)	15.09 (1.804)	6.953 (3.224)	15.336 (1.835)	7.01 (3.241)	15.316 (1.919)
31: Furniture	5.542 (2.691)	14.769 (1.453)	5.789 (2.67)	14.883 (1.422)	5.696 (2.804)	14.836 (1.498)	5.765 (2.815)	14.862 (1.506)
32: Other manufacturing	4.519 (3.013)	15.033 (1.647)	4.777 (2.965)	15.155 (1.619)	4.918 (3.105)	15.229 (1.694)	4.941 (3.227)	15.235 (1.773)
Total average	5.029 (3.16)	15.156 (1.695)	5.299 (3.175)	15.282 (1.701)	5.441 (3.239)	15.354 (1.739)	5.45 (3.303)	15.354 (1.783)

Notes: Standard errors in parentheses. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on CIT-IRP5 data.

Table 8. Sample statistics for fixed capital

	2010		2011		2012		2013	
Industry	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**
10: Food	2.8 (2.91)	13.47 (1.72)	2.98 (3.04)	13.51 (1.8)	3.05 (3.15)	13.52 (1.87)	2.98 (3.29)	13.46 (1.98)
11: Beverages	5.24 (3.43)	14.02 (2.23)	5.71 (3.49)	14.24 (2.25)	5.83 (3.53)	14.35 (2.18)	5.75 (3.69)	14.21 (2.37)
13: Textiles	3.56 (3.06)	12.72 (1.94)	3.79 (3.21)	12.79 (2.05)	3.88 (3.36)	12.82 (2.16)	3.84 (3.34)	12.77 (2.18)
14: Wearing apparel	3.22 (3.29)	12.31 (2.1)	3.24 (3.3)	12.22 (2.15)	3.18 (3.39)	12.16 (2.27)	3.1 (3.6)	12.08 (2.46)
15: Leather	4.37 (3.26)	12.59 (2.13)	4.61 (3.06)	12.69 (1.91)	4.73 (3.22)	12.72 (2)	4.66 (3.14)	12.69 (1.91)
16: Wood	4.62 (2.69)	13.43 (1.57)	4.56 (2.7)	13.3 (1.59)	4.59 (2.83)	13.29 (1.67)	4.44 (3.01)	13.15 (1.87)
17: Paper	4.94 (3.01)	13.67 (1.89)	5.24 (3.07)	13.79 (1.91)	5.31 (3.11)	13.77 (1.97)	5.22 (3.3)	13.67 (2.1)
18: Printing	3.26 (2.9)	13.19 (1.8)	3.42 (2.96)	13.21 (1.87)	3.44 (3.08)	13.19 (1.96)	3.29 (3.19)	13.08 (2.07)
19: Coke and refined petroleum	3.99 (2.49)	12.94 (1.73)	4.16 (2.55)	13 (1.77)	4.2 (2.55)	12.99 (1.74)	4.24 (2.72)	12.98 (1.89)
20: Chemicals	3.49 (2.95)	13.3 (1.7)	3.7 (2.88)	13.35 (1.69)	3.77 (2.95)	13.36 (1.74)	3.7 (3.07)	13.31 (1.82)
21: Pharmaceuticals	5.43 (3.45)	13.68 (2.09)	5.92 (3.61)	13.9 (2.18)	6.18 (3.83)	14.03 (2.39)	6.14 (3.72)	14.14 (2.28)
22: Rubber and plastics	4.9 (2.65)	13.86 (1.58)	5.05 (2.76)	13.9 (1.65)	5.15 (2.86)	13.91 (1.73)	5.02 (2.94)	13.82 (1.81)
23: Other minerals	4.18 (2.8)	13.5 (1.78)	4.25 (2.8)	13.48 (1.77)	4.25 (2.89)	13.44 (1.85)	4.3 (2.99)	13.44 (1.93)
24: Basic metals	3.54 (2.97)	13.66 (1.71)	3.64 (3.02)	13.67 (1.71)	3.75 (3.09)	13.63 (1.76)	3.81 (3.15)	13.63 (1.83)
25: Fabricated metals	3.1 (2.58)	13.49 (1.54)	3.22 (2.67)	13.49 (1.6)	3.3 (2.8)	13.48 (1.7)	3.29 (2.86)	13.45 (1.77)
26: Computer, electronic and optical products	4.26 (2.57)	12.83 (1.59)	4.42 (2.68)	12.91 (1.73)	4.45 (2.88)	12.84 (1.92)	4.54 (2.91)	12.95 (1.87)
27: Electrical equipment	3.85 (2.55)	12.96 (1.45)	4.08 (2.57)	13.02 (1.54)	4.17 (2.63)	13.04 (1.6)	4.17 (2.65)	13.06 (1.61)
28: Machinery n.e.c.	2.67 (2.63)	13.15 (1.62)	2.83 (2.67)	13.15 (1.67)	2.93 (2.76)	13.18 (1.72)	2.93 (2.9)	13.14 (1.83)
29: Motor vehicles	1.1 (2.96)	12.64 (1.81)	1.26 (3.02)	12.64 (1.86)	1.3 (3.16)	12.6 (2)	1.24 (3.36)	12.55 (2.19)
30: Transport equipment	4.68 (2.85)	13.4 (1.86)	4.92 (2.97)	13.5 (1.89)	5.11 (3.12)	13.49 (1.97)	5.33 (3.22)	13.64 (2.07)
31: Furniture	3.63 (2.64)	12.86 (1.62)	3.76 (2.73)	12.86 (1.67)	3.58 (2.93)	12.72 (1.86)	3.55 (2.95)	12.64 (1.88)
32: Other manufacturing	2.7 (2.94)	13.22 (1.8)	2.8 (2.99)	13.18 (1.87)	2.87 (3.13)	13.18 (1.95)	2.87 (3.26)	13.16 (2.05)
Total average	3.05 (3.03)	13.18 (1.78)	3.21 (3.08)	13.2 (1.82)	3.27 (3.19)	13.19 (1.91)	3.24 (3.31)	13.15 (2.02)

Notes: Standard errors in parentheses. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on CIT-IRP5 data.

Table 9. Sample statistics for employment

Industry	2010 Weighted	2010 Unweighted	2011 Weighted	2011 Unweighted	2012 Weighted	2012 Unweighted	2013 Weighted	2013 Unweighted
10: Food	−8.185 (2.491)	2.486 (1.326)	−7.808 (2.509)	2.722 (1.248)	−7.717 (2.582)	2.745 (1.266)	−7.738 (2.644)	2.743 (1.284)
11: Beverages	−6.604 (2.792)	2.182 (1.449)	−6.034 (2.761)	2.498 (1.378)	−5.965 (2.82)	2.554 (1.343)	−5.924 (2.858)	2.538 (1.397)
13: Textiles	−6.875 (2.5)	2.286 (1.333)	−6.513 (2.528)	2.48 (1.292)	−6.477 (2.605)	2.467 (1.318)	−6.438 (2.581)	2.491 (1.329)
14: Wearing apparel	−6.863 (2.882)	2.226 (1.557)	−6.648 (2.916)	2.332 (1.586)	−6.61 (2.855)	2.364 (1.569)	−6.589 (2.829)	2.393 (1.518)
15: Leather	−5.786 (2.619)	2.435 (1.434)	−5.499 (2.542)	2.587 (1.37)	−5.388 (2.597)	2.604 (1.397)	−5.392 (2.575)	2.642 (1.368)
16: Wood	−6.163 (2.461)	2.645 (1.317)	−5.999 (2.478)	2.742 (1.3)	−5.928 (2.575)	2.769 (1.32)	−5.95 (2.566)	2.764 (1.306)
17: Paper	−6.195 (2.479)	2.544 (1.313)	−5.802 (2.464)	2.748 (1.239)	−5.673 (2.412)	2.783 (1.2)	−5.72 (2.558)	2.726 (1.271)
18: Printing	−7.738 (2.42)	2.193 (1.216)	−7.503 (2.418)	2.29 (1.214)	−7.445 (2.481)	2.302 (1.248)	−7.508 (2.53)	2.278 (1.267)
19: Coke and refined petroleum	−6.574 (1.923)	2.379 (1.132)	−6.153 (1.955)	2.686 (1.018)	−6.113 (2.029)	2.675 (1.015)	−6.08 (2.067)	2.669 (1.04)
20: Chemicals	−7.58 (2.526)	2.227 (1.264)	−7.314 (2.478)	2.337 (1.233)	−7.234 (2.51)	2.359 (1.238)	−7.251 (2.592)	2.356 (1.291)
21: Pharmaceuticals	−5.391 (3.045)	2.858 (1.566)	−4.787 (3.136)	3.193 (1.608)	−4.542 (3.117)	3.305 (1.628)	−4.939 (3.214)	3.062 (1.637)
22: Rubber and plastics	−6.176 (2.259)	2.788 (1.159)	−5.962 (2.267)	2.885 (1.139)	−5.912 (2.302)	2.855 (1.168)	−5.998 (2.347)	2.805 (1.175)
23: Other minerals	−6.848 (2.296)	2.467 (1.256)	−6.638 (2.251)	2.592 (1.211)	−6.557 (2.254)	2.63 (1.182)	−6.506 (2.364)	2.634 (1.231)
24: Basic metals	−7.49 (2.595)	2.633 (1.301)	−7.337 (2.639)	2.686 (1.283)	−7.172 (2.669)	2.708 (1.282)	−7.046 (2.68)	2.783 (1.293)
25: Fabricated metals	−7.849 (2.284)	2.533 (1.149)	−7.648 (2.301)	2.616 (1.122)	−7.583 (2.357)	2.603 (1.15)	−7.556 (2.383)	2.61 (1.165)
26: Computer, electronic and optical products	−6.506 (2.196)	2.063 (1.078)	−6.337 (2.191)	2.155 (1.052)	−6.23 (2.238)	2.158 (1.068)	−6.222 (2.271)	2.183 (1.046)
27: Electrical equipment	−6.835 (2.37)	2.272 (1.171)	−6.58 (2.215)	2.361 (1.096)	−6.468 (2.231)	2.401 (1.117)	−6.514 (2.264)	2.38 (1.112)
28: Machinery n.e.c.	−8.416 (2.254)	2.064 (1.116)	−8.191 (2.248)	2.135 (1.105)	−8.072 (2.316)	2.174 (1.123)	−8.013 (2.411)	2.202 (1.17)
29: Motor vehicles	−9.341 (2.522)	2.2 (1.156)	−9.092 (2.544)	2.295 (1.127)	−8.989 (2.571)	2.316 (1.135)	−8.991 (2.625)	2.313 (1.141)
30: Transport equipment	−6.525 (2.36)	2.194 (1.245)	−6.288 (2.49)	2.286 (1.278)	−6.057 (2.524)	2.326 (1.286)	−5.852 (2.597)	2.453 (1.325)
31: Furniture	−6.909 (2.305)	2.318 (1.227)	−6.647 (2.341)	2.447 (1.227)	−6.704 (2.398)	2.437 (1.218)	−6.651 (2.444)	2.447 (1.244)
32: Other manufacturing	−8.136 (2.452)	2.378 (1.234)	−7.872 (2.448)	2.506 (1.222)	−7.783 (2.529)	2.528 (1.24)	−7.749 (2.6)	2.545 (1.262)
Total average	−7.784 (2.601)	2.344 (1.242)	−7.515 (2.608)	2.468 (1.216)	−7.428 (2.649)	2.485 (1.226)	−7.412 (2.697)	2.492 (1.244)

Notes: Standard errors in parentheses. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on CIT-IRP5 data.

In Table 9, the sample statistics of log labour per industry is given.[12] Firms in the production of pharmaceutical products, rubber and plastics products and wood and products of wood have the highest average number of workers per firm. Firms manufacturing computer, electronic and optical products, other machinery and printing have the lowest number of workers per firm on average.

In general, the figures suggest significant heterogeneity in industries even at the average level with the key variables roughly following the same relative rankings. That is industries with higher averages for value added are more likely to higher averages for cost of sales, employment and capital stock. A general trend emerges that industries with fewer firms share higher averages, a result especially clear in the pharmaceuticals, beverages and paper industries. This trend is intensified when weighting the firms by contribution to sales over the period. This may be indicative of high barriers to entry. It should be noted that a low number of firms in an industry does not appear to be the sole determinant of these factors with the rubber and plastics, basic metals and food industries having relatively high average costs, capital stock and employment levels despite having a large number of firms. There appears to be some positive relationship between cost of goods sold and chemical input intensity as firms in the paper, pharmaceuticals, coke and refined petroleum, rubber, basic metals and plastics sectors have high average values for both.[13] Again the strength of the relationship is not precise as the food and beverages industries report relatively low levels of chemical input intensity while the textile sector, a relatively low value added sector, has relatively high levels of chemical input intensity. The exceptionally low levels of average capital stock in the coke and refined petroleum sector is very surprising as the industry is marked by high average levels of all other variables of interest. While weighting firms in this sector brings these averages more in line with each other this low level may be indicative of measurement issues.[14] The means for the chemicals sector suggest that the industry is characterised by high value added, mid-level costs and capital stock and low employment input, suggesting that the industry may have high mark-ups, a topic analysed by Fedderke *et al.* (2016).

3. PRODUCTIVITY ESTIMATION

To measure productivity, we first estimate a production function for each 2-digit manufacturing sector and use the estimated parameters to back out a firm-specific measure of productivity. We also estimate the production function for each sector separately as the estimation assumes that all firms share a common technology; this is more realistic within 2-digit subsectors than for the manufacturing sector as a whole. Simultaneity between productivity shocks (observed by the firm but not the econometrician) and input choices leads to a bias in ordinary least square (OLS) estimates of the coefficients on these inputs in a standard production function. For example, firms that experience a

[12] Note that the negative weighted values are as expected because the number of employees is multiplied by the firm's weight before taking the natural logarithm.

[13] Using the StatsSA input output tables, chemical intensity here is defined as the share of nuclear fuel and other chemical inputs in inputs from all industries and compensation of employees.

[14] Here, it should be noted that weighting the variables still keeps the within firm relationship of capital, labour, value added and costs constant, it only changes the weight assigned to specific firms. That is, the weighted values are more representative of the industry as a whole.

negative productivity shock may decide to reduce their labour force or delay investment. This will lead to an upward bias in the coefficients on labour and capital. It is also possible that the employment decisions of firms are countercyclical, with higher productivity firms deciding to replace labour with more capital-intensive production processes. This would lead to a downward bias in OLS estimation of the coefficient on labour.

A common approach to estimating the production function parameters in the presence of such bias is to use a semiparametric estimator that applies some structure to the underlying decision-making process of firms (Olley and Pakes, 1996). In this paper, we use Ackerberg *et al.*'s (2006) modification of the Levinsohn and Petrin (2003) approach. Ackerberg *et al.*'s (2006) approach addresses multicollinearity issues that affect the identification of the parameters in the first stage of the estimation of the models of Olley and Pakes (1996) and Levinsohn and Petrin (2003). We estimate the model using a two-step generalised method of moments estimator (Wooldridge, 2009). A brief description of the approach is provided in Supporting Information Appendix B, with a more detailed exposition available in Newman *et al.* (2015). In each case, tests for under identification, weak identification and first-stage *F*-tests confirm the validity of the instruments.[15] We use higher-order terms of the instruments or additional lags to test for overidentification. Details of the instrument used to test for the overidentifying restrictions in each sector are also provided in Table 10. For all sectors, we find the lag of labour to be a suitable instrument for current period labour. We use the parameter estimates from the exactly identified system to back productivity to avoid loss of data due to the inclusion of additional lags and to reduce the potential for weak instruments. The results do not change much when we use different combinations of valid overidentifying restrictions.

Table 10 presents the OLS and instrumental variables estimates for the production function parameters for each 2-digit manufacturing subsector. We present the weighted estimates (preferred) as well as results for the unweighted and weighted balanced sample where we exclude all entrants and exits over the sample period for comparison.

The coefficient estimates on the labour and capital inputs are higher in the weighted sample than in the unweighted sample. This can be explained by the fact that in weighting the data before estimating the production function, we give more weight to larger firms (in terms of sales) who earn greater returns from their inputs than smaller firms. The capital coefficients in the balanced panel are larger than those in the restricted sample. This is in line with expectations given that the balanced panel captures survivors who, in general, are expected to have a higher capital stock.

Comparing our estimates to the OLS estimates, we find that in all cases the coefficient on labour is lower when the production function is estimated using the Ackerberg *et al.* (2006) approach. This makes sense if we believe that there is a positive correlation between labour and productivity shocks leading to an upward bias in the labour coefficient when using OLS. For the weighted sample OLS underestimates the capital coefficient, implying that the capital used by firms is negatively related to firm productivity leading to a downward bias.

[15] In Table 10, we also present p values for each test. The underidentification test is based on the Kleibergen-Park Lagrange multiplier statistic, the weak identification test is based on the Kleibergen-Park Wald *F* statistic, the *F*-test is based on the Angrist-Pischke multivariate *F*-test of excluded instruments in the first stage, and the test for the overidentifying restrictions is based on Hansen's *J*-test.

Table 10. Production function regression results

	Weighted		Unweighted		Balanced	
	IV	**OLS**	**IV**	**OLS**	**IV**	**OLS**
10: Food products						
Labour	0.353*** (0.0158)	0.621*** (0.00828)	0.319*** (0.0155)	0.463*** (0.00922)	0.317*** (0.0211)	0.444*** (0.0102)
Capital	0.440*** (0.0250)	0.366*** (0.00683)	0.129*** (0.0201)	0.353*** (0.00641)	0.507*** (0.0335)	0.548*** (0.00887)
N	5981	5981	5981	5981	3420	3420
N clusters	2226		2226		855	
R^2	0.986	0.929	0.999	0.726	0.993	0.953
R^2 centred	0.952		0.825		0.965	
OID variable	llab_lag2		k_lag4		llab_lag2	
Overidentification	0.178		0.752		0.542	
Underidentification	5.72e–103		0		0	
Weak identification	3352.8		3038.6		11,649.4	
Endogenous regressors	470.3		420.9		220.8	
Returns to scale	0.793	0.988	0.448	0.815	0.824	0.992
RT S.E.	(0.025)	(0.004)	(0.025)	(0.007)	(0.033)	(0.004)
RT=1, *P* value	0.0000	0.0012	0.0000	0.0000	0.0000	0.0584
11: Beverages						
Labour	0.421*** (0.0780)	0.801*** (0.0277)	0.375*** (0.0763)	0.612*** (0.0345)		
Capital	0.524*** (0.0796)	0.230*** (0.0220)	0.269*** (0.0696)	0.261*** (0.0213)		
	700	700	700	700		
N clusters	263		263			
R^2	0.989	0.921	0.998	0.707		
R^2 centred	0.946		0.811			
OID variable	llab_lag2		llab_lag2			
Overidentification	0.338		0.413			
Underidentification	0.0000103		0			
Weak identification	3785.4		456.7			
Endogenous regressors	28.34		24.03			
Returns to scale	0.946	1.031	0.643	0.873		
RT S.E.	(0.061)	(0.012)	(0.086)	(0.024)		
RT = 1, p value	0.3738	0.0121	0.0000	0.0000		
13: Textiles						
Labour	0.385*** (0.0277)	0.713*** (0.0147)	0.319*** (0.0269)	0.511*** (0.0167)	0.289*** (0.0371)	0.550*** (0.0191)
Capital	0.362*** (0.0375)	0.261*** (0.0116)	0.133*** (0.0317)	0.253*** (0.0106)	0.566*** (0.0550)	0.429*** (0.0162)
N	2020	2020	2020	2020	1088	1088
N clusters	750		750		272	
R^2	0.989	0.909	0.998	0.671	0.994	0.922
R^2 centred	0.941		0.797		0.948	
OID variable	k_2lag2		llab_lag2		llab_lag2	
Overidentification	0.102		0.119		0.171	
Underidentification	1.84e–47		0		0	

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	OLS	IV	OLS	IV	OLS
Weak identification	1521.4		1392.5		652.6	
Endogenous regressors	159.2		126.7		54.41	
Returns to scale	0.747	0.973	0.452	0.764	0.855	0.979
RT S.E.	(0.04)	(0.007)	(0.04)	(0.013)	(0.062)	(0.009)
RT = 1, p value	0.0000	0.0002	0.0000	0.0000	0.0192	0.0203
14: Wearing apparel						
Labour	0.274*** (0.0323)	0.685*** (0.0140)	0.235*** (0.0307)	0.499*** (0.0155)	0.267*** (0.0345)	0.490*** (0.0206)
Capital	0.270*** (0.0493)	0.245*** (0.0118)	0.0789*** (0.0282)	0.204*** (0.0107)	0.468*** (0.0494)	0.454*** (0.0186)
N	1549	1549	1549	1549	832	832
N clusters	602		602		208	
R^2	0.988	0.898	0.998	0.645	0.994	0.934
R^2 centred	0.941		0.807		0.964	
OID variable	k_lag4		llab_lag2		llab_lag2	
Overidentification	0.726		0.482		0.278	
Underidentification	1.24e–25		0		0	
Weak identification	1018.0		1275.0		793.7	
Endogenous regressors	64.58		55.60		55.45	
Returns to scale	0.545	0.931	0.314	0.703	0.735	0.944
RT S.E.	(0.056)	(0.008)	(0.04)	(0.013)	(0.052)	(0.009)
RT = 1, p value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
15: Leather and related products						
Labour	0.277*** (0.0415)	0.603*** (0.0213)	0.260*** (0.0412)	0.427*** (0.0229)	0.208*** (0.0593)	0.419*** (0.0281)
Capital	0.418*** (0.0647)	0.345*** (0.0173)	0.0916*** (0.0353)	0.304*** (0.0160)	0.475*** (0.0761)	0.556*** (0.0269)
N	874	874	874	874	516	516
N clusters	291		291		129	
R^2	0.992	0.901	0.998	0.655	0.995	0.922
R^2 centred	0.942		0.802		0.952	
OID variable	k_lag4		llab_lag2		k_lag4	
Overidentification	0.698		0.255		0.694	
Underidentification	5.78e–16		0		0	
Weak identification	1485.3		569.4		582.8	
Endogenous regressors	40.81		39.49		11.67	
Returns to scale	0.695	0.948	0.352	0.731	0.683	0.975
RT S.E.	(0.061)	(0.011)	(0.047)	(0.019)	(0.074)	(0.013)
RT = 1, p value	0.0000	0.0000	0.0000	0.0000	0	0.0538
16: Wood and products of wood and cork, except furniture						
Labour	0.371*** (0.0366)	0.662*** (0.0165)	0.301*** (0.0363)	0.522*** (0.0182)	0.352*** (0.0482)	0.526*** (0.0228)
Capital	0.557*** (0.0518)	0.308*** (0.0148)	0.199*** (0.0450)	0.273*** (0.0142)	0.641*** (0.0640)	0.422*** (0.0213)
N	1563	1563	1563	1563	832	832
N clusters	569		569		208	

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	OLS	IV	OLS	IV	OLS
R^2	0.990	0.913	0.998	0.676	0.995	0.936
R^2 centred	0.936		0.771		0.955	
OID variable	k_2lag3		llab_lag2		llab_lag4	
Overidentification	0.344		0.132		0.137	
Underidentification	5.17e–28		0		0	
Weak identification	550.1		503.9		230.4	
Endogenous regressors	86.58		60.96		40.56	
Returns to scale	0.928	0.97	0.5	0.795	0.993	0.948
RT S.E.	(0.049)	(0.008)	(0.059)	(0.014)	(0.059)	(0.009)
RT = 1, p value	0.1467	0.0001	0.0000	0.0000	0.8991	0.0000
17: Paper and paper products						
Labour	0.455*** (0.0403)	0.753*** (0.0183)	0.403*** (0.0371)	0.582*** (0.0208)	0.388*** (0.0430)	0.538*** (0.0218)
Capital	0.376*** (0.0617)	0.216*** (0.0145)	0.110*** (0.0387)	0.224*** (0.0132)	0.376*** (0.0481)	0.403*** (0.0194)
N	1242	1242	1242	1242	680	680
N clusters	438		438		170	
R^2	0.994	0.933	0.999	0.751	0.998	0.947
R^2 centred	0.953		0.830		0.966	
OID variable	k_lag4		llab_lag2		llab_lag2	
Overidentification	0.777		0.112		0.169	
Underidentification	0.000784		0		0	
Weak identification	4919.4		434.0		332.2	
Endogenous regressors	88.57		86.23		58.76	
Returns to scale	0.831	0.969	0.513	0.806	0.764	0.941
RT S.E.	(0.058)	(0.008)	(0.045)	(0.013)	(0.048)	(0.009)
RT = 1, p value	0.0038	0.0001	0.0000	0.0000	0.000	0.0000
18: Printing and reproduction of recorded media						
Labour	0.540*** (0.0239)	0.808*** (0.00907)	0.476*** (0.0239)	0.671*** (0.0109)	0.476*** (0.0296)	0.717*** (0.0124)
Capital	0.356*** (0.0278)	0.178*** (0.00736)	0.124*** (0.0221)	0.184*** (0.00697)	0.435*** (0.0332)	0.273*** (0.0109)
N	3996	3996	3996	3996	2464	2464
N clusters	1346		1346		616	
R^2	0.990	0.940	0.999	0.770	0.994	0.947
R^2 centred	0.952		0.829		0.960	
OID variable	llab_lag2		llab_lag2		llab_lag2	
Overidentification	0.432		0.556		0.386	
Underidentification	1.03e–64		0		0	
Weak identification	2403.4		2107.8		1280.0	
Endogenous regressors	419.9		348.8		206.6	
Returns to scale	0.896	0.986	0.6	0.855	0.911	0.989
RT S.E.	(0.028)	(0.004)	(0.031)	(0.008)	(0.032)	(0.005)
RT = 1, p value	0.0002	0.0005	0.0000	0.0000	0.0060	0.0258

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	OLS	IV	OLS	IV	OLS
19: Coke and refined petroleum products						
Labour	0.281*** (0.0270)	0.678*** (0.0136)	0.244*** (0.0259)	0.435*** (0.0167)	0.230*** (0.0396)	0.585*** (0.0189)
Capital	0.255*** (0.0354)	0.292*** (0.0105)	0.0738*** (0.0253)	0.234*** (0.00991)	0.287*** (0.0547)	0.394*** (0.0164)
N	1617	1617	1617	1617	896	896
N clusters	614		614		224	
R^2	0.993	0.871	0.999	0.516	0.996	0.889
R^2 centred	0.930		0.738		0.939	
OID variable	llab_lag2		k_lag4		llab_lag2	
Overidentification	0.490		0.420		0.475	
Underidentification	0.00000117		0		0	
Weak identification	8774.0		1121.5		359.2	
Endogenous regressors	100.7		85.88		34.71	
Returns to scale	0.536	0.97	0.318	0.669	0.518	0.98
RT S.E.	(0.041)	(0.009)	(0.038)	(0.017)	(0.061)	(0.012)
RT = 1, p value	0.0000	0.0014	0.0000	0.0000	0.0000	0.0831
20: Chemicals and chemical products						
Labour	0.407*** (0.0239)	0.699*** (0.0118)	0.362*** (0.0243)	0.548*** (0.0136)	0.366*** (0.0283)	0.512*** (0.0148)
Capital	0.466*** (0.0330)	0.323*** (0.0101)	0.142*** (0.0296)	0.314*** (0.00985)	0.538*** (0.0403)	0.515*** (0.0137)
N	3127	3127	3127	3127	1796	1796
N clusters	1086		1086		449	
R^2	0.989	0.926	0.999	0.702	0.995	0.953
R^2 centred	0.950		0.812		0.964	
OID variable	k_lag4		llab_lag2		llab_lag2	
Overidentification	0.935		0.290		0.160	
Underidentification	1.06e–62		0		0	
Weak identification	3099.8		2820.8		1306.0	
Endogenous regressors	249.4		204.4		142.2	
Returns to scale	0.873	1.023	0.504	0.861	0.903	1.028
RT S.E.	(0.032)	(0.005)	(0.035)	(0.01)	(0.038)	(0.006)
RT = 1, p value	0.0001	0.0000	0.0000	0.0000	0.0104	0.0000
Labour	0.254*** (0.0872)	0.760*** (0.0458)	0.235** (0.0915)	0.611*** (0.0524)		
Capital	0.407*** (0.0973)	0.244*** (0.0393)	0.197*** (0.0763)	0.263*** (0.0378)		
N	270	270	270	270		
N clusters	104		104			
R^2	0.994	0.934	0.999	0.762		
R^2 centred	0.964		0.875			
OID variable	llab_lag2		llab_lag2			
Overidentification	0.496		0.755			
Underidentification	0.0000203		0			
Weak identification	2041.9		57.81			

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	OLS	IV	OLS	IV	OLS
Endogenous regressors	7.451		5.910			
Returns to scale	0.661	1.005	0.432	0.874		
RT S.E.	(0.109)	(0.017)	(0.109)	(0.032)		
RT = 1, p value	0.0019	0.7890	0.0000	0.0001		
22: Rubber and plastics products						
Labour	0.352*** (0.0337)	0.621*** (0.0158)	0.305*** (0.0341)	0.458*** (0.0177)	0.297*** (0.0405)	0.496*** (0.0206)
Capital	0.585*** (0.0362)	0.360*** (0.0129)	0.286*** (0.0420)	0.349*** (0.0121)	0.682*** (0.0533)	0.501*** (0.0177)
N	1974	1974	1974	1974	1100	1100
N clusters	688		688		275	
R^2	0.992	0.917	0.999	0.693	0.995	0.935
R^2 centred	0.939		0.776		0.954	
OID variable	k_lag4		llab_lag2		llab_lag2	
Overidentification	0.141		0.183		0.120	
Underidentification	1.81e–27		0		0	
Weak identification	1505.0		1311.4		1740.7	
Endogenous regressors	109.2		79.82		57.17	
Returns to scale	0.938	0.981	0.591	0.807	0.979	0.997
RT S.E.	(0.037)	(0.007)	(0.054)	(0.013)	(0.056)	(0.008)
RT = 1, p value	0.0882	0.0077	0.0000	0.0000	0.7044	0.7165
23: Other non-metallic mineral products						
Labour	0.329*** (0.0282)	0.618*** (0.0139)	0.272*** (0.0285)	0.408*** (0.0150)	0.290*** (0.0339)	0.459*** (0.0179)
Capital	0.478*** (0.0398)	0.324*** (0.0111)	0.170*** (0.0357)	0.294*** (0.0100)	0.549*** (0.0411)	0.503*** (0.0152)
N	2460	2460	2460	2460	1432	1432
N clusters	885		885		358	
R^2	0.988	0.890	0.998	0.623	0.994	0.914
R^2 centred	0.925		0.736		0.942	
OID variable	llab_lag2		llab_lag2		llab_lag2	
Overidentification	0.216		0.388		0.973	
Underidentification	3.25e–49		0		0	
Weak identification	1535.4		1476.3		1052.7	
Endogenous regressors	124.4		86.70		67.48	
Returns to scale	0.807	0.943	0.442	0.702	0.839	0.962
RT S.E.	(0.039)	(0.007)	(0.043)	(0.012)	(0.042)	(0.008)
RT = 1, p value	0.000	0.0000	0.0000	0.0000	0.0001	0.0000
24: Basic metals						
Labour	0.444*** (0.0274)	0.679*** (0.0120)	0.398*** (0.0267)	0.549*** (0.0139)	0.428*** (0.0373)	0.585*** (0.0153)
Capital	0.547*** (0.0378)	0.332*** (0.0104)	0.195*** (0.0303)	0.322*** (0.0102)	0.618*** (0.0436)	0.441*** (0.0140)
N	3049	3049	3049	3049	1592	1592
N clusters	1121		1121		398	
R^2	0.987	0.930	0.998	0.713	0.994	0.952

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	OLS	IV	OLS	IV	OLS
R^2 centred	0.944		0.785		0.960	
OID variable	k_lag4		llab_lag2		llab_lag2	
Overidentification	0.283		0.148		0.108	
Underidentification	6.74e–48		0		0	
Weak identification	1714.6		1580.4		1048.0	
Endogenous regressors	236.0		220.1		118.7	
Returns to scale	0.991	1.011	0.593	0.871	1.046	1.026
RT S.E.	(0.037)	(0.005)	(0.038)	(0.01)	(0.035)	(0.006)
RT = 1, p value	0.8085	0.0345	0.0000	0.0000	0.1839	0.0000
Labour	0.443*** (0.0165)	0.731*** (0.00697)	0.377*** (0.0164)	0.587*** (0.00831)	0.395*** (0.0222)	0.596*** (0.00929)
Capital	0.518*** (0.0194)	0.268*** (0.00595)	0.190*** (0.0184)	0.254*** (0.00576)	0.578*** (0.0246)	0.408*** (0.00863)
N	7911	7911	7911	7911	4508	4508
N clusters	2731		2731		1127	
R^2	0.987	0.922	0.999	0.690	0.993	0.940
R^2 centred	0.938		0.766		0.950	
OID variable	k_lag4		llab_lag2		k_lag4	
Overidentification	0.610		0.530		0.643	
Underidentification	8.61e–121		0		0	
Weak identification	3299.1		3082.4		1248.1	
Endogenous regressors	587.4		478.7		264.7	
Returns to scale	0.961	0.998	0.566	0.841	0.973	1.003
RT S.E.	(0.019)	(0.003)	(0.024)	(0.007)	(0.022)	(0.004)
RT = 1, p value	0.0449	0.5838	0.0000	0.0000	0.2294	0.4170
26: Computer, electronic and optical products						
Labour	0.585*** (0.0497)	0.878*** (0.0185)	0.494*** (0.0486)	0.721*** (0.0247)	0.422*** (0.0707)	0.773*** (0.0270)
Capital	0.356*** (0.0487)	0.154*** (0.0149)	0.108** (0.0422)	0.156*** (0.0147)	0.471*** (0.0726)	0.264*** (0.0228)
N	1024	1024	1024	1024	544	544
N clusters	387		387		136	
R^2	0.991	0.919	0.998	0.660	0.995	0.929
R^2 centred	0.929		0.730		0.941	
OID variable	llab_lag2		llab_lag2		llab_lag2	
Overidentification	0.426		0.657		0.713	
Underidentification	3.92e–24		0		0	
Weak identification	1324.7		1320.6		1253.4	
Endogenous regressors	112.8		92.50		31.76	
Returns to scale	0.942	1.032	0.602	0.877	0.893	1.037
RT S.E.	(0.042)	(0.01)	(0.053)	(0.02)	(0.069)	(0.013)
RT = 1, p value	0.1623	0.001	0.0000	0.0000	0.1200	0.0035
27: Electrical equipment						
Labour	0.444*** (0.0348)	0.751*** (0.0149)	0.396*** (0.0347)	0.620*** (0.0178)	0.365*** (0.0431)	0.611*** (0.0209)

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	**OLS**	**IV**	**OLS**	**IV**	**OLS**
Capital	0.452*** (0.0415)	0.262*** (0.0130)	0.0960*** (0.0341)	0.241*** (0.0129)	0.636*** (0.0554)	0.423*** (0.0202)
N	1579	1579	1579	1579	896	896
N clusters	564		564		224	
R^2	0.992	0.926	0.999	0.696	0.996	0.941
R^2 centred	0.943		0.793		0.956	
OID variable	llab_lag2		llab_lag2		llab_lag2	
Overidentification	0.856		0.931		0.552	
Underidentification	4.18e–30		0		0	
Weak identification	586.9		548.3		216.0	
Endogenous regressors	124.5		109.4		53.73	
Returns to scale	0.896	1.012	0.492	0.862	1.001	1.034
RT S.E.	(0.042)	(0.007)	(0.046)	(0.015)	(0.049)	(0.009)
RT = 1, p value	0.0135	0.0903	0.0000	0.0000	0.9790	0.0001
28: Machinery and equipment n.e.c.						
Labour	0.474*** (0.0179)	0.817*** (0.00730)	0.401*** (0.0185)	0.661*** (0.00912)	0.410*** (0.0268)	0.670*** (0.0103)
Capital	0.464*** (0.0218)	0.205*** (0.00615)	0.152*** (0.0187)	0.201*** (0.00602)	0.597*** (0.0295)	0.359*** (0.00940)
N	7384	7384	7384	7384	4072	4072
N clusters	2752		2752		1018	
R^2	0.985	0.919	0.998	0.672	0.992	0.938
R^2 centred	0.937		0.769		0.953	
OID variable	k_lag4		llab_lag2		llab_lag4	
Overidentification	0.427		0.335		0.104	
Underidentification	2.90e–146		0		0	
Weak identification	5127.3		5071.9		1936.7	
Endogenous regressors	584.2		432.8		214.7	
Returns to scale	0.938	1.022	0.553	0.862	1.007	1.029
RT S.E.	(0.02)	(0.004)	(0.025)	(0.007)	(0.025)	(0.004)
RT = 1, p value	0.0021	0.0000	0.0000	0.0000	0.7682	0.0000
29: Motor vehicles, trailers and semi-trailers						
Labour	0.519*** (0.0143)	0.893*** (0.00486)	0.468*** (0.0139)	0.763*** (0.00685)	0.482*** (0.0178)	0.783*** (0.00702)
Capital	0.259*** (0.0162)	0.141*** (0.00399)	0.0935*** (0.0112)	0.144*** (0.00396)	0.366*** (0.0253)	0.255*** (0.00641)
N	9983	9983	9983	9983	6096	6096
N clusters	3482		3482		1524	

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	**OLS**	**IV**	**OLS**	**IV**	**OLS**
R^2	0.984	0.946	0.999	0.735	0.991	0.961
R^2 centred	0.961		0.831		0.971	
OID variable	llab_lag2		llab_lag3		llab_lag2	
Overidentification	0.189		0.162		0.199	
Underidentification	3.85e–159		0		0	
Weak identification	4880.8		4714.7		2418.1	
Endogenous regressors	976.9		945.1		601.8	
Returns to scale	0.778	1.034	0.562	0.908	0.847	1.038
RT S.E.	(0.018)	(0.003)	(0.016)	(0.006)	(0.025)	(0.003)
RT = 1, p value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
30: Other transport equipment						
Labour	0.504*** (0.0348)	0.709*** (0.0172)	0.442*** (0.0338)	0.572*** (0.0209)	0.366*** (0.0450)	0.606*** (0.0255)
Capital	0.533*** (0.0548)	0.288*** (0.0141)	0.295*** (0.0838)	0.274*** (0.0137)	0.646*** (0.0748)	0.376*** (0.0226)
N	1116	1116	1116	1116	520	520
N clusters	476		476		130	
R^2	0.990	0.919	0.998	0.705	0.996	0.927
R^2 centred	0.933		0.759		0.948	
OID variable	k_lag4		llab_lag2		llab_lag2	
Overidentification	0.316		0.457		0.447	
Underidentification	4.02e–31		0		0	
Weak identification	2371.7		2477.5		1547.4	
Endogenous regressors	182.8		158.4		60.89	
Returns to scale	1.037	0.997	0.738	0.847	1.011	0.982
RT S.E.	(0.058)	(0.009)	(0.098)	(0.017)	(0.079)	(0.012)
RT = 1, p value	0.5206	0.7719	0.0073	0.0000	0.8870	0.1448
31: Furniture						
Labour	0.390*** (0.0267)	0.715*** (0.0128)	0.354*** (0.0259)	0.566*** (0.0148)	0.339*** (0.0367)	0.546*** (0.0192)
Capital	0.482*** (0.0302)	0.271*** (0.0108)	0.153*** (0.0350)	0.256*** (0.0103)	0.594*** (0.0397)	0.448*** (0.0177)
N	2260	2260	2260	2260	1288	1288
N clusters	804		804		322	
R^2	0.990	0.921	0.999	0.711	0.994	0.939
R^2 centred	0.944		0.809		0.956	
OID variable	k_lag4		llab_lag2		llab_lag2	

Table 10. Continued

	Weighted		Unweighted		Balanced	
	IV	OLS	IV	OLS	IV	OLS
Overidentification	0.500		0.270		0.181	
Underidentification	8.94e–44		0		0	
Weak identification	1516.7		1672.3		5125.5	
Endogenous regressors	179.2		177.7		84.00	
Returns to scale	0.872	0.986	0.507	0.821	0.934	0.993
RT S.E.	(0.032)	(0.006)	(0.043)	(0.011)	(0.035)	(0.007)
RT = 1, p value	0.0001	0.0212	0.0000	0.0000	0.0610	0.3566
32: Other manufacturing						
Labour	0.408*** (0.0168)	0.751*** (0.00727)	0.340*** (0.0165)	0.589*** (0.00875)	0.371*** (0.0219)	0.598*** (0.0104)
Capital	0.439*** (0.0233)	0.248*** (0.00592)	0.123*** (0.0175)	0.242*** (0.00566)	0.561*** (0.0262)	0.412*** (0.00925)
N	7428	7428	7428	7428	3912	3912
N clusters	2872		2872		978	
R^2	0.984	0.919	0.998	0.685	0.991	0.937
R^2 centred	0.938		0.777		0.950	
OID variable	k_lag4		llab_lag2		k_lag4	
Overidentification	0.262		0.116		0.606	
Underidentification	2.51e–148		0		0	
Weak identification	3991.4		3710.5		2123.2	
Endogenous regressors	525.3		408.3		252.7	
Returns to scale	0.847	0.999	0.463	0.831	0.932	1.01
RT S.E.	(0.025)	(0.004)	(0.023)	(0.007)	(0.027)	(0.004)
RT = 1, p value	0.0000	0.8015	0.0000	0.0000	0.0112	0.0258

Note: OID, overidentification; RT, returns to scale. Production functions for the balanced panel for sector 11 (beverages) and sector 21 (pharmaceuticals) could not be estimated because of too few observations. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on CIT-IRP5 data.

We observe capital elasticities (weighted estimates) ranging from 0.255 to 0.585. Firms manufacturing rubber and plastics products, wood products, basic metal products, transport equipment, beverages and fabricated metal products have the highest capital coefficients, above 0.5. These sectors are among the smaller sectors in the sample when measured by total value added but yield the highest value-added returns to capital. Firms manufacturing coke, motor vehicles and apparel have the lowest capital coefficient, below 0.3. Firms in the production of computer, electronic and optical products, printing, motor vehicles and transport equipment have labour elasticities above 0.5. Firms on the low end of the elasticity distribution include coke and refined petroleum products, pharmaceuticals, apparel and leather products. We find evidence of constant returns to scale only for firms producing beverages, wood products, computer, electronic, and optical products, basic metals and transport equipment, whereas all other industries are characterised by decreasing returns to scale.

4. TFP IN SOUTH AFRICAN MANUFACTURING

We use the elasticity estimates presented in Table 10 to estimate productivity for each firm in each sector using equation (2).

$$\hat{\omega}_{ijt} = y_{itj} - \hat{\beta}_{lj} l_{ijt} - \hat{\beta}_{kj} k_{ijt}, \tag{2}$$

where $\hat{\omega}_{ijt}$ is the estimated (log) of TFP for firm i in sector j in time t, y is log value added, l is the log of labour, k is the log of capital, $\hat{\beta}_{lj}$ is the estimated labour elasticity for sector j and $\hat{\beta}_{lj}$ is the estimated capital elasticity for sector j. Given that firms in different sectors (by assumption) use different technologies, we cannot compare the level of productivity across sectors. We can, however, compare the growth trajectory. In Table 11, we present the growth in average TFP between 2010 and 2013 for each sector, where the level of TFP in each sector is indexed to the average value of TFP in 2010. The best performing sectors in terms of productivity growth are the chemicals, coke and refined petroleum and non-metallic mineral products sectors.[16] The worst-performing sectors were firms in the production of leather products, pharmaceutical products and wood products.

In Fig. 2a and Fig. 2b, we provide scatter plots illustrating the differences across sectors in the TFP growth rate of firms in the top 25% and bottom 25% of the TFP distribution, respectively. The growth rate of the top and bottom 25% of firms in each sector (relative to all other firms in the sector) is estimated using equation (3).

$$\text{TFP\_Growth}_i = \alpha_0 + \alpha_1 \text{Position\_2010}_i, \tag{3}$$

where TFP_Growth_i is the growth rate in TFP of firm i over the entire time period and Position_2010_i is an indicator for whether the firm was in the top or bottom 25% of the TFP distribution in 2010. This estimate allows us to analyse the relative growth rates of the top and bottom performers in each industry. The size of the circles indicates the size of the industry in terms of contribution of that industry to total value added to the manufacturing sector.

[16] It should be noted that changes in productivity over time may be due to either real productivity changes or the entry and exit of firms given that we are working with an unbalanced panel. However, this effect is unlikely to be high, as our estimates are robust to including controls for exit.

Table 11. TFP index per year

Industry	2010	2011	2012	2013
10: Food	100	104.380	106.471	103.739
11: Beverages	100	101.619	102.112	102.295
13: Textiles	100	101.274	104.762	102.443
14: Wearing apparel	100	102.849	102.279	100.406
15: Leather	100	96.550	97.585	97.614
16: Wood	100	99.825	103.487	99.093
17: Paper	100	99.895	100.141	104.643
18: Printing	100	97.110	101.910	102.193
19: Coke and refined petroleum	100	110.258	111.425	109.374
20: Chemicals	100	99.815	111.569	108.080
21: Pharmaceuticals	100	103.170	98.251	98.322
22: Rubber and plastics	100	101.128	101.142	101.684
23: Other minerals	100	104.123	105.372	109.326
24: Basic metals	100	101.518	102.244	101.451
25: Fabricated metals	100	100.878	101.626	102.496
26: Computer, electronic and optical products	100	100.430	100.686	100.032
27: Electrical equipment	100	100.774	101.859	100.121
28: Machinery n.e.c.	100	102.378	103.352	104.457
29: Motor vehicles	100	102.828	103.340	104.460
30: Transport equipment	100	102.646	103.795	102.425
31: Furniture	100	103.505	103.219	103.890
32: Other manufacturing	100	100.969	102.784	103.059

Note: Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on results from TFP regressions on CIT-IRP5 data.

In Fig. 2a, we find a weak positive relationship between the growth rate of the top 25% of firms in terms of productivity in each industry and the average growth rate for the industry as a whole. In Fig. 2b, we see a very weak negative relationship between the growth rate of the bottom performing 25% of firms and the growth of the industry as a whole. This suggests, perhaps unsurprisingly, that the productivity growth of the industry is driven by the most productive firms in the industry. This is particularly the case for sectors that appear in the upper right quadrant of Fig. 2a, namely, those producing chemicals and coke and refined petroleum products. The size of the industry does not appear to be related to the growth of the industry as a whole or the growth of the top and bottom performing 25% of firms, although it does appear that a number of the smaller sectors have slower average growth in TFP.

A well-documented stylised fact relating to the manufacturing sector in both developed and developing country economies is that there is considerable heterogeneity in firm-level productivity, even within narrowly defined sectors (Tybout, 2000; Syverson, 2011). To examine the extent of heterogeneity in the South African context, we plot the distribution of productivity for each subsector and across different firm characteristics. To make meaningful comparisons across subgroups of firms, we de-mean by the industry-year average to control for industry- and year-specific shifts.

Figure 3 presents the distribution of productivity across different subsectors of manufacturing, and the distribution of productivity in each subsector for each year of the sample is presented in Supporting Information Appendix C (Fig. C1). A wider dispersion indicates a greater level of heterogeneity across firms within sectors. Also of interest is the extent to which there is a greater concentration of firms in the upper part, or the lower part, of the distribution. In Fig. 3a, the productivity distributions for food and beverages sectors are observed to have relatively high dispersions. In both cases there is a high concentration of firms in the right tail, indicating that these sectors, particularly the beverages sector, have a disproportionate number of high productivity firms relative to the mean.

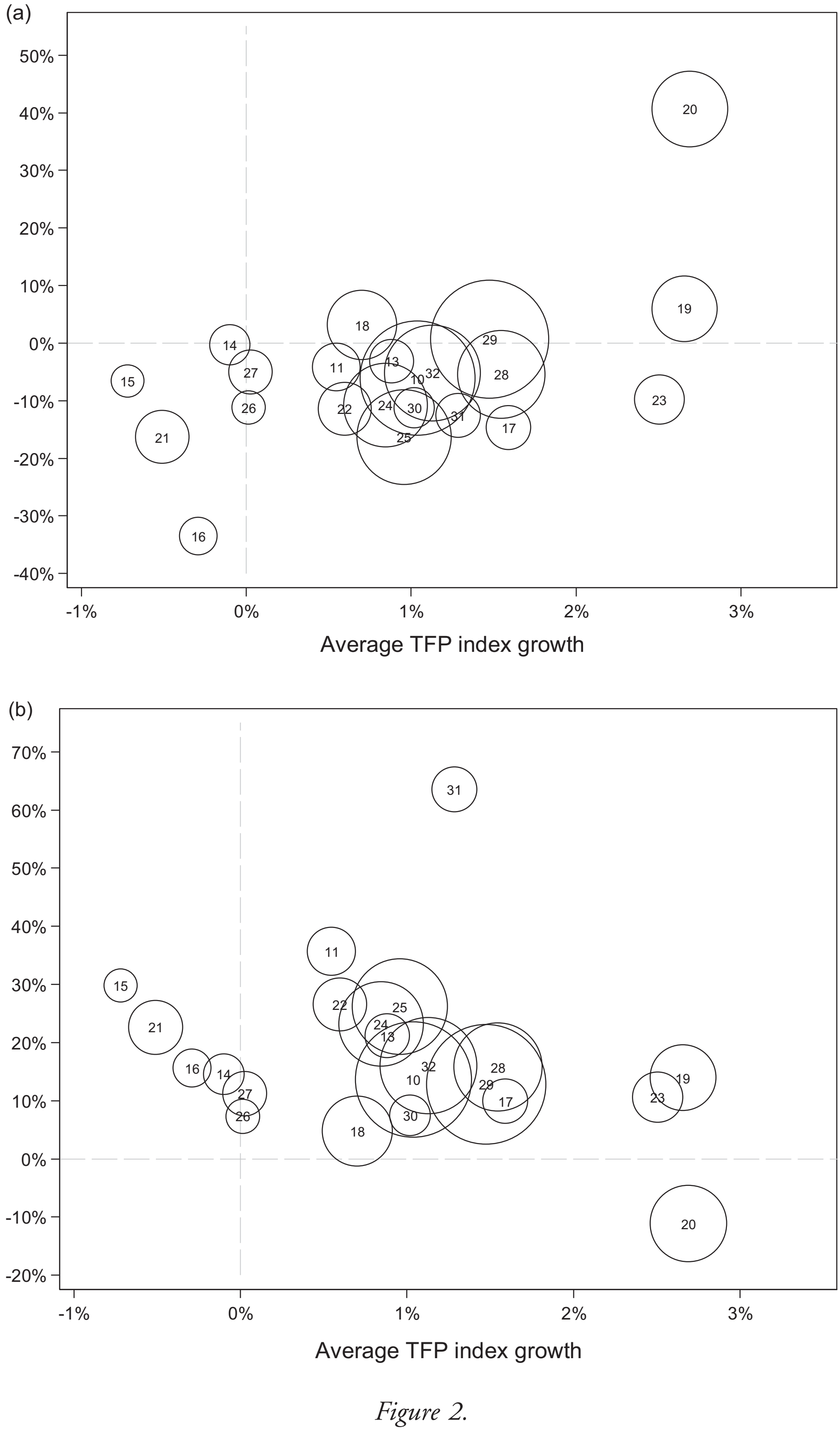
(a)
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
-1%
0%
1%
2%
3%
Average TFP index growth
20
19
18
29
14
13
32
10
28
11
27
15
23
22
24
30
26
31
17
25
21
16
(b)
70%
60%
50%
40%
30%
20%
10%
0%
-10%
-20%
-1%
0%
1%
2%
3%
Average TFP index growth
31
11
15
22
25
21
24
13
32
28
19
16
14
10
29
27
23
17
26
30
18
20

Figure 2.

Figure 2. TFP growth and growth of top and bottom performing firms by industry share in total value added
Note: Industry codes are as follows: 10 Food; 11 Beverages; 13 Textiles; 14 Apparel; 15 Leather; 16 Wood; 17 Paper; 18 Printing; 19 Coke and refined petroleum; 20 Chemicals; 21 Pharmaceuticals; 22 Rubber and plastics; 23 Non-metallic mineral products; 24 Basic metals; 25 Fabricated metals; 26 Computer, electronic and optical products; 27 Electrical equipment; 28 Machinery not elsewhere classified (n.e.c.); 29 Motor vehicles; 30 Other transport equipment; 31 Furniture; 32 Other manufacturing. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on results from TFP regressions on CIT-IRP5 data.

In Fig. 3b, the TFP distributions for textiles, apparel and leather manufacturing firms are shown to be very wide relative to that of other industries. This illustrates that in these more traditional sectors a large number of high productivity and low productivity firms co-exist. This is suggestive of rigidities or other distortions that allow low productivity firms to continue to operate (Hsieh and Klenow, 2009). This is particularly the case for the apparel sector, although the distribution does seem to have tightened somewhat over the timeframe of our analysis.

In Fig. 3c, firms in the production of wood products are shown to have a tight dispersion but this appears to be widening over time (Supporting Information Appendix C), with more firms forming part of the left tail of the distribution. This

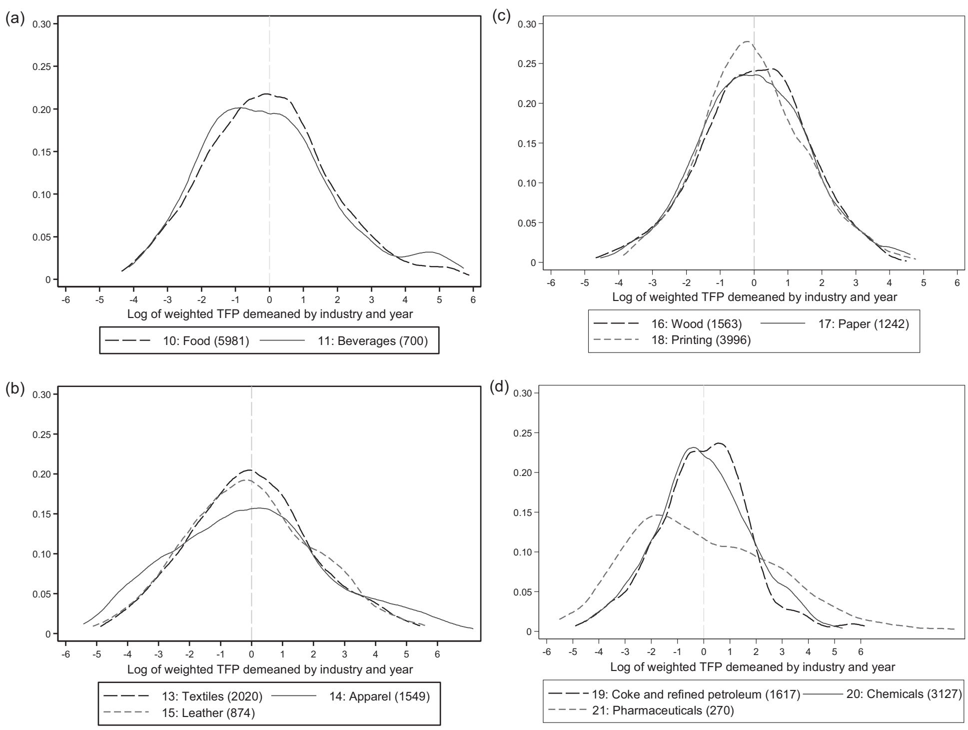

Figure 3. TFP distribution by manufacturing subsector
Note: Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on TFP results from regressions on the CIT-IRP5 data.

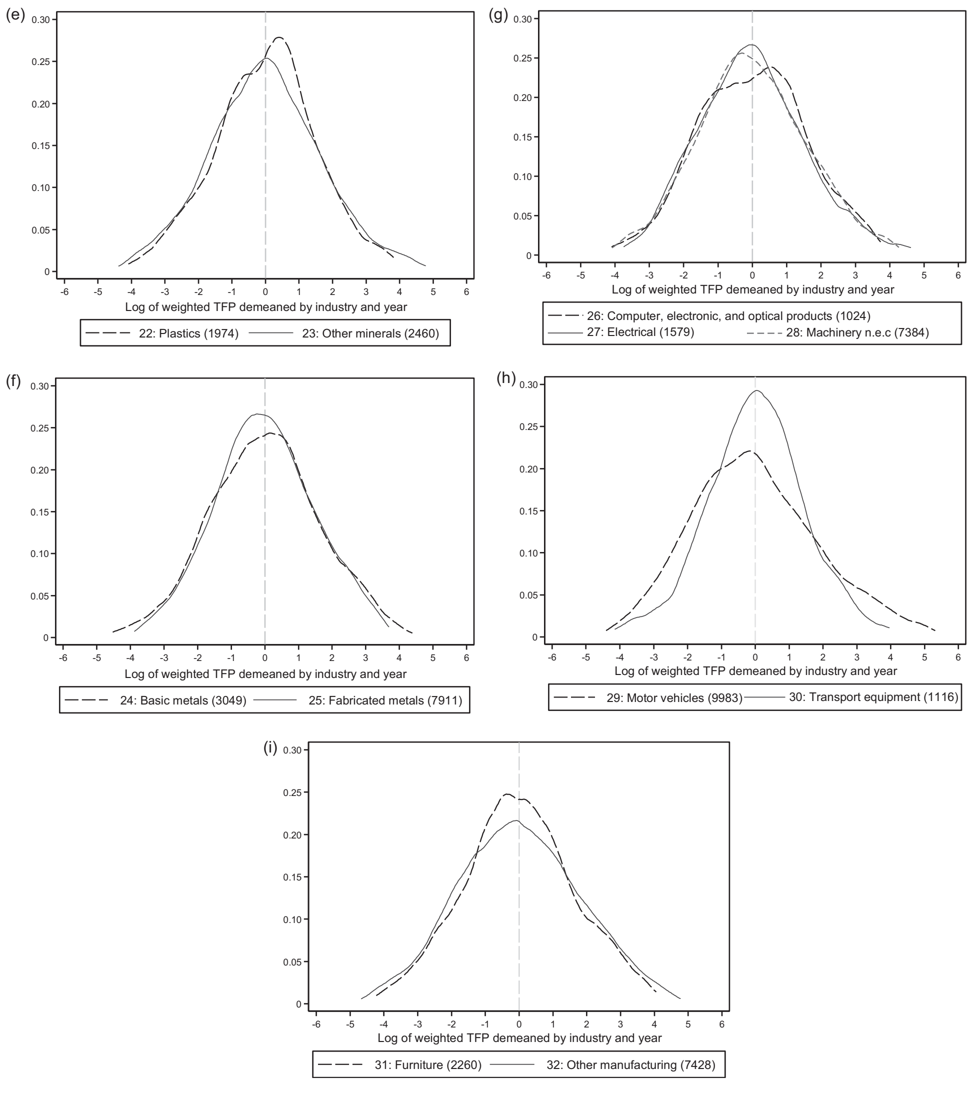

Figure 3. (Continued)

is suggestive of either more low productivity firms entering into the sector or some firms becoming relatively less productive over time. Firms in the production of paper products, on the other hand, are evenly dispersed compared to other industries, with the distribution becoming bimodal over time. Printing products have a tight dispersion with an increasing density at the left tail, or an increase in the number of low-productivity firms.

In Fig. 3d, firms in the production of coke and refined petroleum products are shown to be relatively widely dispersed with a mode above the mean suggesting a greater number of low-productivity firms. In contrast, firms in the production of chemical products are dispersed with a mode below the mean and a relatively large density above the mean, pointing to more high-productivity firms. The pharmaceutical products sector is the most widely dispersed. Interestingly, this sector is also the one with the fewest firms. In

Supporting Information Appendix C, we observe that an increasing density of firms in the right tail, or an increase in the number of firms with relatively higher productivity, drives the widening in the dispersion of the distribution over time.

In Fig. 3e, firms in the production of rubber and plastics products are shown to be very tightly distributed, but we observe a dramatic widening of the distribution of the firms over time with the sector appearing to be tending towards bimodality, suggesting growth in the number of both high and low productivity firms within the sector. The productivity of firms producing other nonmetallic mineral products has a medium dispersion that appears to be tightening over time.

In Fig. 3f, firms in the production of fabricated metals and basic metals are shown to be very tightly distributed as are firms in the production of computer, electronic and optical products and the production of electrical equipment as revealed in Fig. 3g.

In Fig. 3h, firms in the production of transport equipment are shown to be very tightly dispersed and the distribution appears to be tightening over time. Firms in the production of motor vehicles, on the other hand, are relatively widely dispersed with increasing density in the left tail over time. Finally, in Fig. 3i, furniture firms are shown to have a tighter than average dispersion in productivity with very short tails. Other manufacturing firms have an average dispersion that appears to be slowly widening.

In sum, we observe considerable heterogeneity in productivity across firms within sectors but this also differs to a great extent across sectors. It is not the purpose of this study to explain why such heterogeneity exists but it clearly points to rigidities, distortions or regulatory restrictions that prevent resources from being allocated efficiently within sectors. This is particularly the case for the sectors with the widest productivity dispersion including in particular wearing apparel and pharmaceuticals.[17]

Figure 4 illustrates the productivity distribution for firms in different size categories. We consider eight size categories in total: firms with 1–4 employees, 5–9 employees, 10–19 employees, 20–49 employees, 50–99 employees, 100–249 employees, 250–999 employees and 1,000+ employees. We find that average productivity increases with firm size and that the distribution is narrowest for firms with between 10 and 19 employees. The level and dispersion in productivity is very different at the two extremes of the size distribution. For micro firms, those with fewer than five employees, we find that average productivity is much lower than for firms with more than 1,000 employees. Productivity is also widely dispersed among micro firms, suggesting that there is a lot of heterogeneity in productivity levels within this size category. The largest firms also appear to have a wide and bimodal distribution, suggesting that there are distortions at the top end of the size distribution that allow for large amounts of heterogeneity in the productivity of large firms.

Figure 5 illustrates the productivity distribution for firms in different age categories. We consider five age categories in total: firms in existence for less than five years, firms aged 5–10 years, 10–20 years, 20–40 and 40+ years. The youngest firms have the lowest productivity level and also exhibit a wide dispersion in the productivity distribution. The

[17] See Boonzaaier *et al.* (2017) for a detailed exploration of misallocation in the South African manufacturing sector as measured by the extent of dispersion in the productivity distribution of firms within sectors.

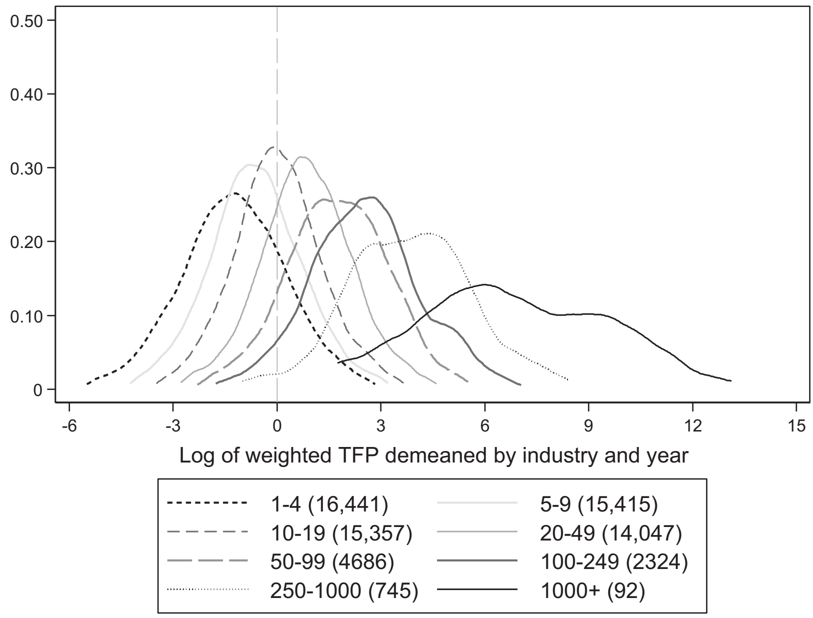

Figure 4. TFP demeaned by industry and year by employment category
Source: Authors' calculations based on TFP results from regressions on the CIT-IRP5 data.

average productivity level appears to be increasing with firm age and the distribution of productivity narrowing as firms approach the 20–40-year category. Firms older than 40 years appear to be substantially more productive than younger firms on average but with

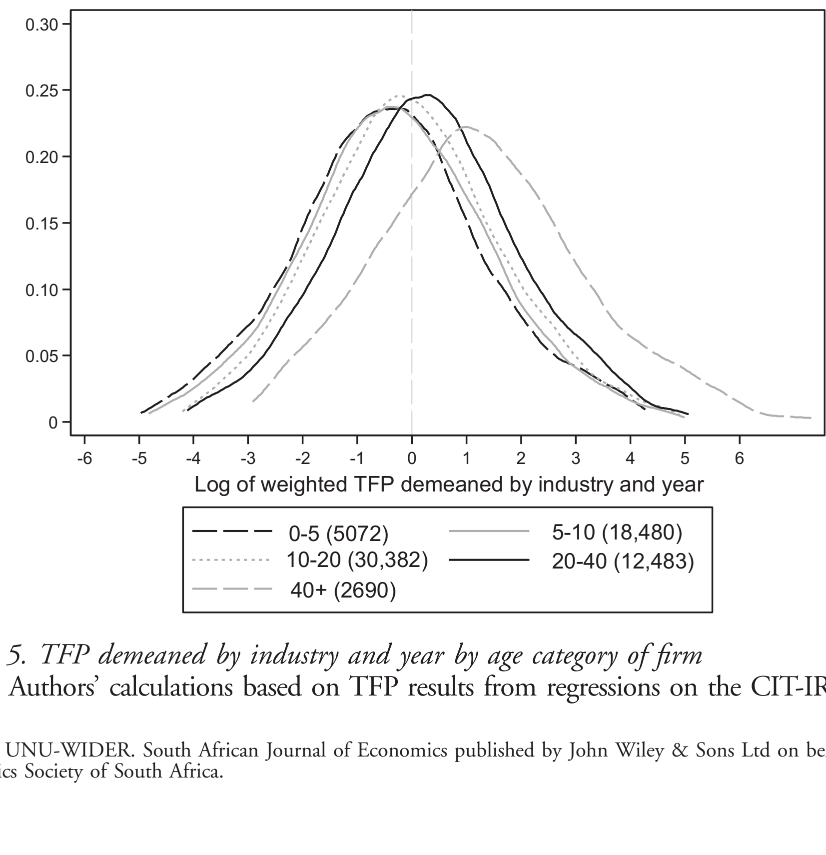

Figure 5. TFP demeaned by industry and year by age category of firm
Source: Authors' calculations based on TFP results from regressions on the CIT-IRP5 data.

a relatively wide distribution suggesting that low productivity firms manage to survive. This provides further suggestive evidence of distortions in the South African economy that prevent the efficient allocation of resources. (See Boonzaaier *et al*., 2017 for a detailed analysis of this issue in the South African context.)

5. PRODUCTIVITY AND ITS CORRELATES

The construction of unbiased firm-specific productivity measures provides an important basis for analysing the determinants of manufacturing productivity in South Africa and paves the way for future research aimed at determining the causal drivers at work. To motivate future research in this area, we explore the relationship between productivity and firm-specific characteristics including the size and age of the firm, R&D, uptake of government incentives for R&D and worker training (learnership agreements),[18] and engagement with trade. Before examining how these variables are correlated with TFP and TFP growth, we examine some of the trends within sectors in access to R&D and learnership incentives and engagement in trade.

Table 12 shows that R&D expenditure is highest in the pharmaceuticals, chemicals and other manufacturing sectors while R&D is lowest in the furniture, apparel, wood, motor vehicles and coke and refined petroleum sectors. The aggregate average R&D over all sectors was between 16 and 18% between 2010 and 2012. While average R&D over all firms appears to have increased between 2010 and 2012, a dramatic drop is observed in 2013. R&D expenditures grew fastest in the pharmaceutical sector between 2010 and 2012, with a particularly sharp drop in 2013. Sectors unaffected by this decline include the textiles, apparel, paper, computer and optical products, electrical machinery and motor vehicles sectors.

In terms of tax deductions for R&D, Table 13 shows that computer, beverage, chemical, other manufacturing and pharmaceutical sectors are the most significant benefactors. Wood, printing, paper, coke and refined petroleum and wearing apparel industries have no R&D tax deductions in our sample. R&D tax deductions appear to follow general R&D expenditure as they also fall rather dramatically in 2013. This may be due to changes in the tax incentive documentation required, although again the shift may be caused by reporting differences over the years in question.

Table 14 shows that average absolute learnership tax deductions are highest in pharmaceuticals, motor vehicles, textiles and electrical machinery industries, whereas the deductions are lowest in manufacturing of wood, furniture, beverage and computer products. It should be noted that there are no learnerships in these industries in certain years. Learnership tax deductions appear to be increasing over time in the majority of industries, with growth in deductions being highest in pharmaceuticals, wearing apparel and textile industries and uptake not slowing down significantly in any sector. See Rankin *et al* (2014) for a detailed discussion on South Africa's learnership programme.

Table 15 provides information on the proportion of firms involved in specific kinds of international trade within industries. Pharmaceuticals, leather and computer manufacturing

[18] The R&D tax incentive takes the form of a tax deduction amounting to 150% of expenditure incurred directly for R&D and an accelerated depreciation deduction for capital expenditure incurred on machinery or plants used for R&D. The learnership incentive provides deductions to employers for engaging their employees in specific types of worker training.

Table 12. Log R&D expenditure

	2010	2011	2012	2013	Total
10: Food	0.1926	0.2557	0.1909	0.1237	0.1877
	(1.397)	(1.652)	(1.439)	(1.161)	(1.411)
11: Beverages	0.1148	0.4375	0.2216	0.1832	0.2314
	(1.087)	(2.166)	(1.568)	(1.411)	(1.58)
13: Textiles	0.2241	0.2645	0.1579	0.1605	0.2011
	(1.588)	(1.669)	(1.297)	(1.372)	(1.488)
14: Apparel	0.0201	0.054	0.0000	0.0248	0.0244
	(0.414)	(0.691)	(0.0000)	(0.479)	(0.465)
15: Leather	0.1613	0.1069	0.1616	0.0565	0.1222
	(1.378)	(1.096)	(1.336)	(0.812)	(1.179)
16: Wood	0.0466	0.0238	0.062	0.0000	0.0329
	(0.68)	(0.448)	(0.822)	(0.0000)	(0.576)
17: Paper	0.0384	0.0439	0.039	0.1269	0.0635
	(0.673)	(0.741)	(0.663)	(1.153)	(0.841)
18: Printing	0.1221	0.0592	0.0733	0.0467	0.0757
	(1.09)	(0.751)	(0.861)	(0.751)	(0.877)
19: Coke and Ref. Pet	0.0427	0.1106	0.0622	0.0055	0.0534
	(0.655)	(1.054)	(0.832)	(0.11)	(0.736)
20: Chemicals	0.5065	0.5144	0.485	0.2855	0.4453
	(2.312)	(2.28)	(2.259)	(1.685)	(2.145)
21: Pharmaceuticals	0.4392	0.6962	0.9455	0.7125	0.6834
	(2.133)	(2.964)	(3.307)	(2.937)	(2.823)
22: Plastics	0.3345	0.1347	0.3309	0.1461	0.2359
	(1.85)	(1.18)	(1.876)	(1.214)	(1.566)
23: Other Minerals	0.1235	0.1004	0.0656	0.0000	0.0722
	(1.089)	(0.967)	(0.786)	(0.0000)	(0.828)
24: Basic Metals	0.1366	0.1609	0.2179	0.1217	0.1579
	(1.223)	(1.4)	(1.561)	(1.202)	(1.347)
25: Fabr. Metals	0.1043	0.1261	0.1183	0.0631	0.1016
	(1.017)	(1.165)	(1.131)	(0.821)	(1.035)
26: Comp. Elec. Optic. Prod	0.2637	0.1627	0.2124	0.2561	0.2265
	(1.736)	(1.416)	(1.659)	(1.754)	(1.654)
27: Electrical	0.2806	0.3464	0.2204	0.2495	0.2735
	(1.692)	(2.015)	(1.615)	(1.674)	(1.751)
28: Machinery n.e.c	0.2036	0.2219	0.2627	0.144	0.2059
	(1.524)	(1.569)	(1.718)	(1.285)	(1.525)
29: Motor Vehicles	0.0647	0.0564	0.0477	0.0438	0.0533
	(0.877)	(0.759)	(0.694)	(0.707)	(0.765)
30: Transport Equipment	0.0357	0.0386	0.209	0.0536	0.0805
	(0.628)	(0.586)	(1.502)	(0.915)	(0.963)
31: Furniture	0.0244	0.0288	0.0434	0.0000	0.0235
	(0.588)	(0.645)	(0.727)	(0.0000)	(0.56)
32: Other Manufacturing	0.2921	0.3587	0.3861	0.1855	0.3021
	(1.745)	(1.989)	(2.053)	(1.473)	(1.819)
Total Average	0.1674	0.1809	0.1821	0.1082	0.1584
	(1.338)	(1.404)	(1.415)	(1.103)	(1.316)

Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

firms appear to be most involved in trade of any kind while coke and refined petroleum, food and motor vehicles industries appearing to be less connected with global markets. Among trading firms, in the majority of sectors most firms engage in both import and export activities. The exceptions are the wood and furniture industries which a have a slightly larger number of firms exporting only, while the wearing apparel industry have slightly more importing firms than other trading firms. Overall, the total number of trading firms is increasing over time.

In Table 16, the proportion of firms in different industries within different age groups is presented. Firms within the 10–20-year age group are most common in all industries, followed by firms in the 5–10-year category. The food, beverage, other minerals, transport equipment and furniture industries have relatively more firms in the 0–5-year old category than firms in the 20–40-year old category, indicating a large number of incumbents in this sector within the sample in question.

Table 13. Log R&D tax incentive

	2010	2011	2012	2013	Total
10: Food	0.0073	0.0143	0.019	0.0141	0.0134
	(0.294)	(0.382)	(0.491)	(0.402)	(0.394)
11: Beverages	0.1851	0.264	0.1244	0.2067	0.1951
	(1.432)	(1.63)	(1.099)	(1.587)	(1.454)
13: Textiles	0.0504	0.086	0.0607	0.051	0.0613
	(0.817)	(1.059)	(0.923)	(0.821)	(0.904)
14: Apparel	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
15: Leather	0.0553	0.0000	0.0000	0.0000	0.0145
	(0.8200)	(0.0000)	(0.0000)	(0.0000)	(0.419)
16: Wood	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
17: Paper	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
18: Printing	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
19: Coke and Ref. Pet	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
20: Chemicals	0.2012	0.1987	0.1462	0.1142	0.0000
	(1.625)	(1.6)	(1.402)	(1.22)	(1.468)
21: Pharmaceuticals	0.0000	0.2601	0.2968	0.0000	0.1222
	(0.0000)	(1.946)	(2.119)	(0.0000)	(1.345)
22: Plastics	0.0292	0.0232	0.0224	0.0764	0.0388
	(0.635)	(0.494)	(0.493)	(1.02)	(0.706)
23: Other Minerals	0.0294	0.0215	0.0000	0.0000	0.013
	(0.561)	(0.506)	(0.0000)	(0.0000)	(0.38)
24: Basic Metals	0.0434	0.0777	0.0584	0.0183	0.0483
	(0.703)	(1.021)	(0.895)	(0.507)	(0.795)
25: Fabr. Metals	0.0229	0.0504	0.0348	0.0134	0.0297
	(0.506)	(0.768)	(0.623)	(0.365)	(0.577)
26: Comp. Elec. Optic. Prod	0.2619	0.276	0.2184	0.1497	0.2231
	(1.853)	(1.853)	(1.701)	(1.448)	(1.709)
27: Electrical	0.0655	0.1149	0.0733	0.0656	0.0791
	(0.917)	(1.277)	(1.009)	(0.964)	(1.045)
28: Machinery n.e.c	0.1059	0.1222	0.1438	0.0811	0.1119
	(1.187)	(1.289)	(1.396)	(1.047)	(1.229)
29: Motor Vehicles	0.0266	0.0252	0.0178	0.0185	0.0221
	(0.611)	(0.604)	(0.502)	(0.539)	(0.566)
30: Transport Equipment	0.1343	0.0000	0.0000	0.0000	0.0387
	(1.364)	(0.0000)	(0.0000)	(0.0000)	(0.733)
31: Furniture	0.0232	0.0274	0.0257	0.0000	0.0186
	(0.559)	(0.614)	(0.592)	(0.0000)	(0.503)
32: Other Manufacturing	0.1728	0.1425	0.1619	0.0912	0.1428
	(1.508)	(1.365)	(1.432)	(1.12)	(1.368)
Total Average	0.063	0.0667	0.0601	0.0392	0.0569
	(0.907)	(0.931)	(0.888)	(0.727)	(0.865)

Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

In Table 17, we report regression results on a weighted TFP measure. Consistent with Fig. 4 we find that larger firms, in terms of numbers employed, are more productive than smaller firms. The relationship between firm size and productivity has been explored extensively in the empirical literature, but the evidence is mixed. Using a similar approach to the one we use in this paper, Fernandes (2008) finds that smaller firms in Bangladesh manufacturing industries have higher TFP on average. Similarly, Söderbom and Teal (2004) find evidence of substantial allocative inefficiency in large manufacturing firms in Ghana due to higher labour costs than in smaller firms and more costly capital-intensive technology. In contrast, Van Biesebroeck (2005a) finds that larger firms are more productive in general (for Burundi, Ethiopia, Tanzania, Zambia, Kenya, Cote d'Ivoire, Ghana, Zimbabwe and Cameroon). He also finds that larger firms grow larger and become more productive

Table 14. Log learnerships

	2010	2011	2012	2013	Total
10: Food	0.0300	0.0656	0.0349	0.0438	0.0427
	(0.6010)	(0.8900)	(0.6400)	(0.7700)	(0.728)
11: Beverages	0.0000	0.0000	0.0924	0.0000	0.0211
	(0.0000)	(0.0000)	(1.131)	(0.0000)	(0.541)
13: Textiles	0.0666	0.2162	0.2845	0.2703	0.2065
	(0.883)	(1.622)	(1.831)	(1.77)	(1.562)
14: Apparel	0.0221	0.0319	0.1371	0.2522	0.1095
	(0.454)	(0.579)	(1.251)	(1.711)	(1.118)
15: Leather	0.0620	0.1787	0.1052	0.1285	0.1179
	(0.9190)	(1.494)	(1.068)	(1.301)	(1.211)
16: Wood	0.0000	0.034	0.0000	0.0366	0.0174
	(0.0000)	(0.64)	(0.0000)	(0.72)	(0.479)
17: Paper	0.1077	0.1099	0.1122	0.0719	0.0997
	(1.095)	(1.074)	(1.098)	(0.911)	(1.043)
18: Printing	0.0596	0.0625	0.0911	0.1258	0.0853
	(0.851)	(0.85)	(1.056)	(1.211)	(1.007)
19: Coke and Ref. Pet	0.0287	0.0658	0.0687	0.061	0.0546
	(0.608)	(0.886)	(0.92)	(0.858)	(0.817)
20: Chemicals	0.0295	0.1125	0.1079	0.0577	0.0755
	(0.588)	(1.131)	(1.111)	(0.811)	(0.929)
21: Pharmaceuticals	0.3124	0.6226	1.1615	0.7434	0.6862
	(1.781)	(2.676)	(3.594)	(3.05)	(2.809)
22: Plastics	0.0275	0.0765	0.0969	0.0406	0.0599
	(0.599)	(0.941)	(1.07)	(0.654)	(0.835)
23: Other Minerals	0.0172	0.0186	0.0182	0.0000	0.0133
	(0.44)	(0.438)	(0.434)	(0.0000)	(0.375)
24: Basic Metals	0.1525	0.1588	0.1968	0.124	0.1571
	(1.342)	(1.371)	(1.505)	(1.212)	(1.358)
25: Fabr. Metals	0.0956	0.1339	0.2014	0.2097	0.1614
	(1.049)	(1.247)	(1.516)	(1.551)	(1.362)
26: Comp. Elec. Optic. Prod	0.0000	0.059	0.0546	0.0000	0.0264
	(0.0000)	(0.884)	(0.851)	(0.0000)	(0.592)
27: Electrical	0.1483	0.1487	0.2107	0.2311	0.186
	(1.313)	(1.271)	(1.515)	(1.587)	(1.432)
28: Machinery n.e.c	0.0802	0.1118	0.1253	0.0812	0.0982
	(0.956)	(1.141)	(1.213)	(0.988)	(1.072)
29: Motor Vehicles	0.1897	0.2561	0.3596	0.3433	0.2856
	(1.458)	(1.692)	(2.012)	(1.953)	(1.788)
30: Transport Equipment	0.1059	0.0926	0.1327	0.082	0.1026
	(1.073)	(0.992)	(1.192)	(0.996)	(1.063)
31: Furniture	0.0561	0.0235	0.0000	0.0000	0.0202
	(0.783)	(0.527)	(0.0000)	(0.0000)	(0.476)
32: Other Manufacturing	0.044	0.1029	0.156	0.1485	0.1095
	(0.708)	(1.081)	(1.343)	(1.341)	(1.132)
Total Average	0.0807	0.1226	0.1598	0.1469	0.1265
	(0.962)	(1.188)	(1.357)	(1.308)	(1.209)

Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

faster. This is further supported by Arnold *et al*. (2008) who find a productivity premium for larger manufacturing firms in Africa. The relationship between firm size and productivity in the South African context appears to be consistent with the latter findings on the basis of our estimates. We also find that older firms are generally more productive than younger firms.

We find a positive and significant correlation between R&D expenditure and productivity. Similarly, R&D tax allowances are also shown to be positively correlated with TFP even after controlling for actual R&D expenditure. Firms with higher capital-labour ratios are found to be more productive on aggregate. We find significant productivity premiums for firms involved in international trade. Similar studies to ours also find a positive relationship between exporting and productivity (see Alvarez and Lopez, 2005: for Chile manufacturing firms; Cruz *et al*., 2016: for Mozambique; Fernandes, 2008: for

Table 15. Trade status

Industry	No trade	Exports	Imports	Imports and exports
10: Food	83.47%	4.89%	3.92%	7.72%
11: Beverages	60.82%	11.89%	4.42%	22.87%
13: Textiles	50.99%	6.37%	15.09%	27.54%
14: Apparel	59.95%	3.70%	18.59%	17.76%
15: Leather	44.30%	9.14%	13.54%	33.02%
16: Wood	72.86%	12.54%	5.71%	8.89%
17: Paper	56.86%	8.89%	10.81%	23.44%
18: Printing	75.63%	8.31%	5.16%	10.90%
19: Coke and Ref. Pet	91.00%	2.55%	1.15%	5.30%
20: Chemicals	53.16%	10.98%	8.30%	27.57%
21: Pharmaceuticals	41.56%	9.47%	15.23%	33.74%
22: Plastics	50.98%	11.28%	11.75%	25.98%
23: Other Minerals	79.58%	5.88%	6.13%	8.42%
24: Basic Metals	67.74%	11.20%	6.59%	14.48%
25: Fabr. Metals	65.05%	9.45%	8.31%	17.19%
26: Comp. Elec. Optic. Prod	46.86%	2.99%	14.16%	35.99%
27: Electrical	69.17%	8.03%	4.95%	17.85%
28: Machinery n.e.c	57.35%	5.74%	10.55%	26.36%
29: Motor Vehicles	83.03%	4.99%	3.80%	8.18%
30: Transport Equipment	63.47%	8.01%	9.97%	18.55%
31: Furniture	71.05%	12.24%	6.28%	10.42%
32: Other Manufacturing	48.05%	10.96%	11.04%	29.96%

Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

Bangladesh; Newman *et al.*, 2015: for Vietnam; and Van Biesebroeck, 2005b: for Sub-Saharan Africa). Finally, we find that TFP increases relative to 2010 levels in 2011 and 2012 but experiences a statistically significant decline in 2013.

In Table 18, we show the relationship between firm characteristics and future TFP growth of the firm. The level of real value added of the firm is shown to be negatively correlated with TFP growth whereas the capital-labour ratio of the firm is insignificantly positively correlated. Productivity growth is negative for firms in the bottom 25% of the value-added distribution and positive for firms in the top 25% of the distribution.

Table 16. Age categories

Industry	0–5	5–10	10–20	20–40	40+
10: Food	16.37%	32.51%	40.87%	8.59%	1.67%
11: Beverages	16.31%	31.71%	40.55%	8.23%	3.20%
13: Textiles	11.12%	25.14%	43.75%	15.81%	4.18%
14: Apparel	11.18%	27.50%	41.56%	17.01%	2.74%
15: Leather	11.16%	26.25%	38.36%	20.43%	3.80%
16: Wood	13.74%	30.99%	38.95%	13.21%	3.12%
17: Paper	10.97%	25.94%	47.38%	13.38%	2.33%
18: Printing	13.24%	27.73%	38.61%	15.67%	4.75%
19: Coke and Ref. Pet	11.36%	31.84%	43.08%	11.42%	2.30%
20: Chemicals	10.21%	25.39%	43.14%	17.95%	3.31%
21: Pharmaceuticals	8.64%	26.34%	43.62%	15.64%	5.76%
22: Plastics	9.99%	29.55%	40.42%	16.20%	3.83%
23: Other Minerals	16.25%	31.63%	37.00%	13.04%	2.08%
24: Basic Metals	12.84%	26.12%	38.85%	17.99%	4.20%
25: Fabr. Metals	10.13%	23.69%	41.63%	20.08%	4.47%
26: Comp. Elec. Optic. Prod	8.87%	30.91%	40.68%	17.55%	1.99%
27: Electrical	10.02%	28.26%	41.10%	17.66%	2.95%
28: Machinery n.e.c	11.17%	26.16%	39.95%	19.01%	3.71%
29: Motor Vehicles	13.15%	26.48%	41.06%	16.36%	2.95%
30: Transport Equipment	14.73%	31.13%	40.63%	11.65%	1.86%
31: Furniture	12.24%	30.00%	44.01%	11.97%	1.78%
32: Other Manufacturing	10.22%	25.61%	40.37%	19.46%	4.34%

Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

Table 17. OLS regressions of TFP on firm characteristics

	Log weighted TFP	
Variables	**Estimated coefficient**	**Standard error**
Firm size (base category: 1–4)		
5–9	0.475***	(0.016)
10–19	1.045***	(0.016)
20–49	1.786***	(0.017)
50–99	2.483***	(0.024)
100–249	3.02***	(0.033)
250-1,000	3.456***	(0.079)
Firm age (base category: 0–4)		
5–10	0.081***	(0.018)
10–20	0.123***	(0.018)
20–40	0.103***	(0.021)
40+	0.18***	(0.034)
Policy variables		
Log R&D expenditure	0.032***	(0.004)
Log R&D tax incentive	0.047***	(0.007)
Log amount deducted through learnership agreements	0.033***	(0.004)
Log capital-labour ratio	0.145***	(0.003)
Trade variables (base category: no trade)		
Firm exports	0.447***	(0.02)
Firm imports	0.501***	(0.02)
Firm imports and exports	0.979***	(0.016)
Year (base year: 2010)		
2011	0.138***	(0.015)
2012	0.116***	(0.015)
2013	−0.003	(0.015)
Model statistics		
N	67 091	
R^2	0.5219	
Adjusted R^2	0.5217	

Note: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.
Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

The growth rates of firms of different sizes appear to correlate to their position in the TFP distribution. Productivity growth of firms in the top 25% of the TFP distribution within their respective industries and year are generally lower than those of firms elsewhere in the distribution. Firms with smaller labour forces within the top 25% of the TFP distribution appear to be hardest hit, with only firms with between 100 and 249 employees having no penalty from being in the top 25%.

Productivity growth appears to be decreasing in firm size for firms in the bottom 25% of the TFP distribution. This trend suggests rather rapid TFP growth for new and unproductive firm, a conclusion strengthened by the decreasing returns to firm age. Although firms in the top 25% and bottom 25% of the TFP distribution grow at rates different from those of their comparator firms, the high coefficients on firm size show that it remains the case that larger firms grow faster than smaller firms on average. That being said, a firm with 1–4 employees at the bottom of the TFP distribution is expected have productivity growth higher than a firm with between 5 and 24 employees anywhere else in the distribution. It appears as though firm size in terms of number of employees plays a more important role in increasing productivity than value added.

Table 12 shows that importing firms have higher TFP levels than exporting firms in general; however, exporting firms become more productive at a faster rate than nonexporting firms and importing firms. Although tax allowances for learnerships are positively correlated to TFP growth, the coefficient is relatively small in magnitude.

Table 18. OLS regressions of TFP growth rate on firm characteristics

	TFP growth	
Variables	**Estimated coefficient**	**Standard error**
Log real value added	−0.025***	(0.005)
Log capital-labour ratio	0.008***	(0.002)
Labour category (base category: 1–4 employees)		
5–9	0.139***	(0.018)
10–19	0.186***	(0.018)
20–49	0.264***	(0.021)
50–99	0.335***	(0.035)
100–249	0.376***	(0.061)
250-1,000	−0.021	(0.216)
Top 25% of value added distribution in industry and year	0.032*	(0.018)
Bottom 25% of value added distribution in industry and year	−0.042**	(0.018)
Position in TFP distribution in industry and year with firm size interactions		
Top 25% of TFP distribution (Base Category: 1–4 employees)	−0.275***	(0.034)
Top 25% and 5–9 employees	0.116***	(0.043)
Top 25% and 10–19 employees	0.134***	(0.04)
Top 25% and 20–49 employees	0.125***	(0.039)
Top 25% and 50–99 employees	0.106**	(0.048)
Top 25% and 100–249 employees	0.095	(0.07)
Top 25% and 250-1,000 employees	0.434*	(0.226)
Bottom 25% of TFP distribution (Base Category: 1–4 employees)	0.188***	(0.021)
Bottom 25% and 5–9 employees	−0.00003	(0.026)
Bottom 25% and 10–19 employees	0.078***	(0.03)
Bottom 25% and 20–49 employees	0.106***	(0.041)
Bottom 25% and 50–99 employees	0.455***	(0.092)
Bottom 25% and 100–249 employees	0.284	(0.178)
Bottom 25% and 249-1,000 employees	2.713***	(0.479)
Age of firm (base category: 1–4)		
5–10	−0.04***	(0.014)
10–20	−0.045***	(0.013)
20–40	−0.048***	(0.015)
40+	−0.054**	(0.025)
Trade variables (base category: no trade)		
Firm exports	0.078***	(0.016)
Firm imports	0.061***	(0.016)
Firm imports and exports	0.081***	(0.012)
Policy variables		
Log R&D expenditure	−0.002	(0.003)
Log R&D tax incentive	0.006	(0.005)
Log tax incentive amount for learnership agreements	0.005	(0.003)
Industry (base category: food)		
11: Beverages	−0.104**	(0.046)
13: Textiles	−0.09***	(0.029)
14: Wearing apparel	0.046	(0.033)
15: Leather	−0.032	(0.041)
16: Wood	0.046	(0.032)
17: Paper	−0.051	(0.035)
18: Printing	0.008	(0.022)
19: Coke and refined petroleum	0.141***	(0.032)
20: Chemicals	0.02	(0.025)
21: Pharmaceuticals	−0.114	(0.073)
22: Rubber and plastics	0.001	(0.029)
23: Other minerals	0.033	(0.027)
24: Basic metals	−0.074***	(0.024)
25: Fabricated metals	0.008	(0.019)
26: Computer, electronic and optical products	−0.163***	(0.039)
27: Electrical equipment	−0.034	(0.032)
28: Machinery n.e.c.	0.02	(0.019)
29: Motor vehicles	0.018	(0.018)
30: Transport equipment	−0.076**	(0.038)
31: Furniture	−0.002	(0.027)
32: Other manufacturing	0.065***	(0.019)
Current year of firm (base category: 2010)		
2011	−0.154***	(0.01)
2012	−0.284***	(0.01)
Constant	0.078***	(0.027)
Model statistics		

Table 18. Continued

	TFP growth	
Variables	**Estimated coefficient**	**Standard error**
Observations	42,966	
R^2	0.0932	
Adjusted R^2	0.0920	

Notes: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Refer to United Nations (2008:63–67) for the exact contents of each industry.
Source: Authors' calculations based on TFP results from regressions on CIT-IRP5 data.

We find that after controlling for size, TFP distribution, export status, age and other variables, large numbers of industries are growing at around the same rate as the food sector. Firms that manufacture textile and computer, electronic and optical products are growing at a statistically significantly slower rate than the food sector. The general decline of productivity in the leather, pharmaceuticals and apparel industries are shown to be insignificantly different to the growth of the food sector at the firm level after conditioning on size, age, export status and other characteristics.

6. CONCLUSION

The recent availability of tax administration data for South Africa provides researchers and policymakers with a unique and invaluable opportunity to truly understand the dynamics of the private sector. In this paper, we present for the first time disaggregated TFP estimates across sectors, years and firm characteristics, which provides some new insights into the nature and performance of the manufacturing sector in South Africa.

We find that productivity grew in most sectors between 2010 and 2013, and there is heterogeneity across sectors in the pace of growth. We also find significant heterogeneity in productivity within and between sectors. We find that firm size (in terms of number of employees) is positively correlated with TFP and its growth rate. We also consider the correlation between productivity and R&D and find a positive relationship between R&D expenditure and TFP. Moreover, similar to other studies, we find that there is a productivity premium associated with engaging in international trade.

Understanding the drivers of firm performance is crucial in designing policies aimed at promoting and expanding the private sector, arguably the key driver of productivity, job creation and exports in the economy. Our analysis paves the way for future research into the factors driving productivity growth of manufacturing firms that can provide causal explanations for the significant heterogeneity in measured firm performance, even within narrowly defined sectors and size groups. This research will play an important role in shaping future industrial policy for the South African economy.

ACKNOWLEDGMENTS

We thank Channing Arndt, Rob Davies, Lawrence Edwards, Neil Rankin, John Rand, Asha Sunderam, Nonso Obikili, Finn Tarp and Johannes Fedderke for their inputs during this project. We are grateful to the following members of the South African National Treasury for allowing us to work on their premises, providing technical assistance and the institutional structure under which this research took place: Monale Ratsoma, Landon McMillan, Konstantin Makrelov, Catherine McLeod, Duncan Pieterse, Marle van Niekerk and Wian Boonzaaier. We thank Tasha Naughtin, Wayde Flowerday and Andre Steenkamp for their assistance in creating the database. We also thank the following members of the South African Revenue Service: Elizabeth

Gavin for promptly answering our questions relating to the data; Jaco de Beer for assisting us with the extraction of the data; and Theo Winkelmann for assisting us with extractions of the 2009 data. This research was carried out under UNU-WIDER's project on Regional Growth and Development in Southern Africa. The usual caveats apply.

REFERENCES

ACKERBERG, D., CAVES, K. and FRAZER, G. (2006). *Structural Identification of Production Functions*. MPRA Paper 38349. Munich: Munich Personal RePEc Archive (MPRA). Available at: https://mpra.ub.uni-muenchen.de/38349/1/MPRA_paper_38349.pdf [Accessed 12 May 2016].

AGHION, P., BRAUN, M. and FEDDERKE, J. (2008). Competition and productivity growth in South Africa. *Economics of Transition*, 16: 741-761.

ALVAREZ, R. and LOPEZ, R. A. (2005). Exporting and performance: Evidence from Chilean plants. *Canadian Journal of Economics*, 38: 1384-1400.

ARNOLD, J. M., MATTOO, A. and NARCISO, G. (2008). Services inputs and firm productivity in sub-Saharan Africa: Evidence from firm-level data. *Journal of African Economies*, 17: 578-599.

BEHAR, A. (2010). Would cheaper capital replace labour? *South African Journal of Economics*, 78: 131-151.

BHORAT, H. and LUNDALL, P. (2004). Employment, wages and skills development: Firm-specific effects—Evidence from a firm survey in South Africa. *South African Journal of Economics*, 72: 1023-1056.

BLACK, A. and GERWEL, H. (2014). Shifting the growth path to achieve employment intensive growth in South Africa'. *Development Southern Africa*, 31: 241-256.

BOONZAAIER, W., NEWMAN, C., RAND, J. and TSEBE, M. (2017). Resource misallocation and total factor productivity: Manufacturing firms in South Africa. UNU-WIDER Working Paper (forthcoming).

CRUZ, A., NEWMAN, C., RAND, J. and TARP, F. (2016). Learning by exporting: The case of Mozambican manufacturing. *Journal of African Economies*, 26: 93-118.

DataFirst. (1994-2012). *South Africa: Post-Apartheid Labour Market Series 1994–2012 (PALMS)*. Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/434 [Accessed 6 December 2014].

DED. (2011). *The New Growth Path: Framework*. Pretoria: Department of Economic Development (DED). Available at: http://www.gov.za/sites/www.gov.za/files/NGP%20Framework%20for%20public%20release%20FINAL_1.pdf [Accessed 29 December 2014].

DTI. (2010). *2010/11–2012/13 Industrial Policy Action Plan*. Pretoria: Department of Trade and Industry (DTI). Available at: http://www.pmg.org.za/questions/Q1569Annexure.pdf [Accessed 29 December 2014].

FEDDERKE, J. W. (2014). *Exploring Unbalanced Growth in South Africa: Understanding the Sectoral Structure of the South African Economy*. ERSA Working Paper 468. South Africa: Economic Research Southern Africa (ERSA). Available at: http://www.econrsa.org/system/files/publications/working_papers/working_paper_468.pdf [Accessed 6 December 2014].

FEDDERKE, J., OBIKILI, N. and VIEGI, N. (2016). Markups and concentration in South African manufacturing sectors. UNU-WIDER Working Paper Number 40/2016

FERNANDES, A. M. (2008). Firm productivity in Bangladesh manufacturing industries. *World Development*, 36: 1725-1744.

HSIEH, C. T. and KLENOW, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124: 1403-1448.

KALDOR, N. (1967). *Strategic Factors in Economic Development*. New York: Ithaca.

LEVINSOHN, J. and PETRIN, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70: 317-341.

LI, Y. and ZHANG, B. (2008). Development path of China and India and the challenges for their sustainable growth. *The World Economy*, 31: 1277-1291.

MAHMOOD, M., LEE, W., MAMERTINO, M., SAGET, C., MALGOUYRES, C. and GIOVANZANA, M. (2014). Growth patterns in developing countries. *World Employment and Social Outlook*, 2014: 17–32.

MILLIN, M. and NICHOLA, T. (2005). Explaining economic growth in South Africa: A Kaldorian approach. *International Journal of Technology Management and Sustainable Development*, 4: 47-62.

NEWMAN, C., RAND, J., TALBOT, T. and TARP, F. (2015). Technology transfers, foreign investment and productivity spillovers. *European Economic Review*, 76: 168-187.

———, PAGE, J., RAND, J., SHIMELES, A., SÖDERBOM, M. and TARP, F. (2016). *Made in Africa: Learning to compete in industry*. Washington, DC: Brookings Institution Press.

NPC. (2011). *National Development Plan: Vision for 2030*. Pretoria: National Planning Commission (NPC), Department of the Presidency. Available at: http://www.npconline.co.za/medialib/downloads/home/NPC%20National%20Development%20Plan%20Vision%202030%20-lo-res.pdf [Accessed 29 December 2014].

OLLEY, G. and PAKES, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64: 1263-1297.

PIETERSE, D., KREUSER, C. F. and GAVIN, E. (2016). Introduction to the South African Revenue Service and National Treasury firm-level panel. UNU-WIDER Working Paper Number 42/2016.

RODRIK, D. (2008). Understanding South Africa's economic puzzles. *Economics of Transition*, 16: 769-797.

SARB. (2014). *Quarterly Bulletin*, 274(December). Available at: https://www.resbank.co.za/Publications/Detail-Item-View/Pages/Publications.aspx?sarbweb=3b6aa07d-92ab-441f-b7bf-bb7dfb1bedb4&sarblist=21b5222e-7125-4e55-bb65-56fd3333371e&sarbitem=6530 [Accessed 6 December 2014].
StatsSA. (2010-2015a). *P0044—Quarterly Financial Statistics (QFS)*, March, June, September, December. Statistics South Africa (StatsSA) Publication. Available at: http://www.statssa.gov.za/?page_id=1866&PPN=P0044&SCH=6092 [Accessed 15 December 2015].
———. (2010-2015b). *P0277—Quarterly Employment Statistics (QES)*, March, June, September, December. Statistics South Africa (StatsSA) Publication. Available at: http://www.statssa.gov.za/?page_id=1866&PPN=P0277&SCH=6268 [Accessed 15 December 2015].
———. (2015). *P0044—Quarterly Financial Statistics*. Available at: http://www.statssa.gov.za/?page_id=1854&PPN=P0044&SCH=5844 [accessed 15 December 2015].
SÖDERBOM, M. and TEAL, F. (2004). Size and efficiency in African manufacturing firms: Evidence from firm-level panel data. *Journal of Development Economics*, 73: 369-394.
SYVERSON, C. (2011). What determines productivity? *Journal of Economic Literature*, 49: 326-365.
TARGETTI, F. (2005). Nicholas Kaldor: Key contributions to development economics. *Development and Change*, 36: 1185-1199.
TREGENNA, F. (2008). The contributions of manufacturing and services to employment creation and growth in South Africa'. *South African Journal of Economics*, 76: S175-S204.
TYBOUT, J. R. (2000). Manufacturing firms in developing countries: How well do they do, and why? *Journal of Economic Literature*, 38: 11-44.
UNITED NATIONS. (2008). *International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4*. New York: United Nations. Available at: http://unstats.un.org/unsd/cr/registry/isic-4.asp [Accessed 15 December 2015].
Van BIESEBROECK, J. (2005a). Firm size matters: Growth and productivity growth in African manufacturing. *Economic Development and Cultural Change*, 53: 545-583.
———. (2005b). Exporting raises productivity in Sub-Saharan African manufacturing firms. *Journal of International Economics*, 67: 373-391.
WELLS, H. and THIRLWALL, A. P. (2003). Testing Kaldor's growth laws across the countries of Africa. *African Development Review*, 15: 89-105.
WDI. (2016). *World Development Indicators 1960–2015*. Washington, DC: World Bank. Available at: http://data.worldbank.org/data-catalog/world-development-indicators [Accessed 12 May 2016].
WOOLDRIDGE, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104: 112-114.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Appendix A Database construction

Appendix B Productivity estimation

Appendix C Distribution of total factor productivity (TFP)