FIRM DYNAMICS IN SOUTH AFRICA

By Mpho Tsebe, Veron Vukeya, Christine Lewis, Flavio Calvino and Chiara Criscuolo

Until recently a lack of data meant that little was known about the distribution of firms and firm dynamics in South Africa. A new firm-level panel dataset based on tax data creates opportunities to better understand how firms enter, grow and exit. By using the OECD’s DynEmp framework, which was designed to create harmonised variables based on confidential firm-level data, this paper provides new insights about the dynamics of firms in South Africa and how these compare to other countries. One concerning finding is that the entry rate appears likely to have been below the exit rate in recent years, which means that firms are growing older. The relatively low start-up rate compared to other countries together with the higher average firm size of entrants may indicate the presence of barriers to firm entry. As in other countries, young firms have disproportionately contributed to employment growth and remained net job creators even as GDP growth slowed. Nonetheless, the data reinforce the prominence of large firms in the South African economy, including as net job creators.

Introduction

1. The slowdown in the South African economy and persistently low rate of job creation mean that it is more important than ever to understand the structure of the economy and the characteristics of firms that succeed and grow. At the same time the government is increasing the policy focus given to small firms as the driver of job creation. But the study of firm dynamics in South Africa has been hindered by a lack of firm-level data. Some studies have used the self-reported labour force survey data to infer the size distribution of firms (BER, 2016). Others have gained access to the quarterly employer survey, which covers large firms (Kerr, 2014). A new firm-level dataset created by the South African Revenue Service and National Treasury from tax data (“SARS-NT Panel”) contains a rich set of variables and follows firms over time, affording new insights into firm creation and growth in South Africa.

2. Reflecting the lack of data to date, research exploring the dynamics of South African firms is limited. Using the Quarterly Employment Survey of employers, Kerr et al. (2013) finds that job destruction rates are considerably higher than employment creation rates amongst smaller firms and that larger firms are better net creators of jobs. This contrasts with findings for other (mostly OECD) countries that young firms have been the main creators of jobs (Criscuolo, Gal and Menon, 2014; Calvino, Criscuolo and Menon, 2015).

1 The views expressed in this paper are those of the authors and do not necessarily reflect those of the OECD or of the governments of its member countries or the South African National Treasury. This paper is a draft and is not to be cited or circulated without the authors’ permission. Given that the database is new, the results herein should be treated as preliminary. The authors would like to thank Duncan Pieterse for assistance applying the DynEmp programme to the SARS-NT Panel and helpful conversations and comments. They would also like to thank Lenka Wilderova and participants at an internal seminar for their comments. Further thanks go to Pedro Herrera Gimenez for statistical assistance and Raquel Paramo for editorial assistance.

2 South African National Treasury

3 Organisation for Economic Co-operation and Development (OECD)
More recent work with the SARS-NT Panel has highlighted concerning dynamics in the South African economy. Fedderke, Obikili and Viegi (2016) find that the firm exit rate has exceeded the firm entry rate in the manufacturing sector. The bulk of entry and exit is by small firms (Kerr, 2016; Fedderke, Obikili and Viegi, 2016). This entry and exit contributes significantly to job creation and destruction, but so do existing firms. Job reallocation rates tend to decline with firm size (Kerr, 2016).

3. This paper uses the SARS-NT Panel to apply an OECD methodology for creating harmonised datasets based on confidential micro data, thereby enabling comparisons of firm dynamics across countries. There are naturally some caveats to this approach, as described in the following section, which include that the tax dataset by its nature comprises firms that have chosen to pay tax, whereas business register data are generally more comprehensive. The South African data are for the post-crisis period, whereas the data for other countries span periods of healthier growth. For an emerging economy the link to the informal sector is probably also important.

4. The next section explores the size composition of different economic sectors and shows that, notwithstanding the data differences, large firms are particularly prominent in South Africa compared with many other countries. Measuring entry and exit is difficult but firm entry rates are likely to have been below exit rates in recent years, at least in the formal sector. There are indications that the SARS modernisation and compliance programme in the late 2000s increased formalisation; from a research perspective, an unfortunate consequence is that this shortens the sample for the analysis. The analysis of job creation by new firms suggests that the start-up rate is low but firms that do enter are larger than in other countries, which may be due to the nature of the dataset but could also indicate the presence of barriers to entry or to formalisation. Analysis of firm age shows that while it is true that large firms dominate the economy, young firms have been important net job creators.

Data and methodology

The SARS-NT tax dataset

5. This paper takes advantage of a new firm-level longitudinal dataset that is available for South Africa, created as a joint research effort between the South African Revenue Service and the National Treasury: the SARS-NT Panel. The SARS-NT Panel is created from four data sources: (i) company income tax from registered firms who submit tax returns; (ii) employee data from employee income tax certificates submitted by employers; (iii) value-added tax data from registered firms; and (iv) customs records from traders (Pieterse, Gavin and Kreuser, 2016). The data are available from the 2007/08 tax year onwards.

6. The dataset is built up from the company income tax (CIT) data. Company income tax forms are completed by all companies that are residents for tax purposes. Each unique CIT reference number is associated with a single firm. A firm is defined in the Panel as an entity that is registered for company income tax and has completed an income tax form. The employee data are linked to the company income tax data by matching the identifiers of employers (the PAYE (Pay-As-You-Earn) reference number) to CIT reference numbers. Multiple branches and employees of the same firm are identified from the PAYE reference number.

7. Employment is measured in the SARS-NT Panel as the number of jobs, rather than individuals. If a person has more than one job (either simultaneously or during consecutive periods in the same tax year) they will be counted as two workers. Furthermore, the Panel does not contain employees that cannot be linked to a firm, including employees of government departments. Employers must register with the South African Revenue Service (and receive a PAYE reference number) if at least one employee is liable for personal income tax.
Firms can be classified into a sector based on several variables, including the profit code in the company income tax data or the main income source for individuals in the employee data, which is then mapped to sectoral classifications and the firm categorised based on the sector of the majority of employees. Following other research using the SARS-NT Panel, this paper uses the employee-based definition. For the most part, the analysis below focusses on three sectors: manufacturing, construction and non-financial business services.

The sample is further restricted to firms that are active in an economic sense by excluding dormant companies, share block companies and body corporates (the latter two are used for immovable property holdings) and bank nominee companies. To do this, only firms with positive employment in at least one year in the period 2007/08-2013/14 are included. Also, data from 2014/15 onwards are excluded because they are incomplete since the data used for this paper were extracted in mid 2016 and firms have 12 months to submit their returns. Finally, only firms that have non-missing information on their sectoral activity are kept. Taking 2013/14 as an example, the number of firms is reduced from 781,085 to 250,865 active firms with positive employment and non-missing sector classification (Table 1). The analysis focusses on the period 2010/11 to 2013/14, excluding the initial years of the database. Three-quarters of firms are in the business sector and two-thirds were in the manufacturing, construction or non-financial business services sectors; these three “macro sectors” are the main focus of the analysis.

Table 1. Overview of the data

<table>
<thead>
<tr>
<th>Tax-registered firms</th>
<th>Full SARS-NT Corporate Income Tax panel</th>
<th>Active firms with sector classification</th>
<th>Active firms in manufacturing, construction and non-financial business services</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of firms</td>
<td>Firms with positive employment</td>
<td>Number of firms</td>
</tr>
<tr>
<td>2007/08</td>
<td>608,271</td>
<td>201,212</td>
<td>233,937</td>
</tr>
<tr>
<td>2008/09</td>
<td>903,521</td>
<td>228,217</td>
<td>251,308</td>
</tr>
<tr>
<td>2009/10</td>
<td>940,321</td>
<td>268,032</td>
<td>279,145</td>
</tr>
<tr>
<td>2010/11</td>
<td>925,339</td>
<td>276,665</td>
<td>289,188</td>
</tr>
<tr>
<td>2011/12</td>
<td>898,335</td>
<td>276,421</td>
<td>282,344</td>
</tr>
<tr>
<td>2012/13</td>
<td>873,755</td>
<td>275,225</td>
<td>274,410</td>
</tr>
<tr>
<td>2013/14</td>
<td>781,085</td>
<td>267,407</td>
<td>258,946</td>
</tr>
</tbody>
</table>

Note: The bold border shows the firms used in this paper. Active firms are defined as firms with positive employment in at least one year between 2007/08 and 2013/14. Some basic data cleaning is also done.

Source: Authors’ calculations based on the SARS-NT Panel

The DynEmp approach

The analysis of firm dynamics in South Africa relies on the methodological framework developed within the OECD DynEmp project. The OECD DynEmp project is based on a distributed data collection exercise that aims to create a harmonised micro-aggregated database to study employment and business dynamics across countries. The primary sources of analysis are confidential highly representative administrative micro-data, such as business registers or social security records. These data are micro-aggregated along several dimensions by national experts who run a common statistical code produced by the OECD DynEmp team. National experts also ensure that confidentiality is respected by implementing country-specific disclosure procedures. A number of policy papers and reports have already taken advantage of the richness of these data infrastructure to study employment and business dynamics across countries (see

11. Before including a country in the database, the output of the DynEmp program is carefully examined by the DynEmp team in the OECD Directorate for Science, Technology and Innovation, which interacts with the experts in each participating countries and provides country-specific checks aimed at identifying potential inconsistencies or irregular patterns in the underlying data. The South African data were accessed at the National Treasury and benefited from the expertise of other staff at the Treasury.

12. The version of the DynEmp routine that has been used for analysis is called “DynEmp v.2”. Details on the code’s functioning, its building blocks, main definitions used and output produced are extensively discussed by Criscuolo, Gal and Menon (2015). The main outputs produced by the DynEmp v.2 routine can be divided into: (i) flow data; (ii) transition matrices; and (iii) distributed regressions:

- The flow datasets contain annual statistics on gross job flows and on several statistical indicators of firm-level employment growth, such as mean, median, dispersion, and standard deviation.
- Transition matrices follow cohorts of firms at different points in time and summarize their employment growth trajectories after 3, 5 or 7 years. Given the available data for South Africa, the transition used in this analysis is three years.
- Distributed regressions consist instead of predefined econometric models estimated on the microdata, which are mainly aimed at collecting information on exit and employment growth dynamics. Only coefficients and few indicators on the quality of fit are then extracted.

13. The key variables that are used in this paper are age classifications, size classifications, and sector classifications. Sector classifications are either at the 2-digit sectoral level (based on the ISIC Rev. 4 sector classification system, although some very small sectors are aggregated) or for three “macro sectors”, namely manufacturing, construction and non-financial business services.

14. The DynEmp approach allows a great degree of harmonisation along different dimensions and significantly limits sample selection given the representativeness of the underlying data. This is particularly important when carrying out international comparisons. Accordingly, one of the advantages of adopting the DynEmp framework of analysis in this paper is that it allows comparisons of the patterns observed in South Africa with those observed in other participating countries. However, despite these efforts at harmonisation, some caveats associated with specificities of the data used need to be taken into account when comparing the outcomes to other countries. Some of these are discussed in the following sub-section.

**Key issues and caveats**

15. Although one advantage of using the DynEmp approach is that it facilitates international comparisons, differences in the underlying datasets should be kept in mind. General caveats that apply to the DynEmp v.2 data are: (i) that time coverage is country-specific; (ii) data for some countries are designed for research purposes and owing to methodological differences they may deviate from officially published national statistics; and (iii) there are some definitional differences in size and types of firms included in the underlying data source (see discussion below for the case of South Africa). For example, Brazil has one of the longest samples (1996-2012) while Costa Rica and Portugal have shorter samples (2010-12 and 2006-12) (Table 2 in Calvino, Criscuolo and Gal, (2016)). The analysis below focusses on 2010/11-2013/14 for South Africa. Because this period is after the financial crisis, and during a period of slowing GDP growth in South Africa, some comparisons may be affected.
16. Relative to other countries in the DynEmp project, only firms that are registered for company tax that actually submit their tax returns appear in the SARS-NT Panel. This implies that the Panel excludes firms in the informal sector, sole proprietors and partnerships, micro-enterprises that are registered for the turnover tax (turnover below ZAR 1 million, or USD 75 000). Consequently, the dataset is likely to disproportionately omit very young firms and very small firms (as discussed in Appendix B).

17. Other particularities of the SARS-NT Panel mean that defining the date of firm creation is more complicated than in other countries participating in DynEmp. The standard DynEmp approach is to define entry based on birth year recorded in the database and if this is missing, to supplement it with the first year of positive employment. However, there are some complications when applying this to the SARS-NT Panel:

- Firms self-select into the dataset by registering to pay company tax. This means that entry is overstated if there are changes in compliance behaviour that mean existing firms are choosing to pay tax. The early years of the tax dataset coincide with a period of modernisation at the South African Revenue Service that aimed to reduce costs of tax administration and compliance and increase SARS’ ability to enforce compliance (SARS, 2009). The combination of lower costs of compliance for individuals and firms, as well as increased probability of detection raised compliance during these years. Regulations were also changed so that all employees must be registered with the tax office even if they did not need to pay tax (unless no employees were liable for income tax). Consequently, in the early years of the dataset there are a large number of firms entering the database but it is not possible to identify if they are genuinely new firms.

- Birth year can be indicated by the year of registration with the Companies and Intellectual Property Office but there could be a long time before its starts operating. This means that the standard DynEmp approach which relies on company registration year would tend to characterise firms as older than they are in a practical sense and positive employment might be a better indicator of activity.

18. To address these two issues, the definition of a firm’s creation date is altered from the standard DynEmp definition, and is set out in Appendix A and summarised in Figure 1). The main effect is that more use is made of entry based on positive employment than for other countries in DynEmp v.2. Although it is not possible to calculate the true age of firms that have their first year of positive employment between 2007/08 and 2009/10, subsequent entrants can be identified with more confidence based on positive employment and older firms can be identified based on year of registration. This is summarised in Figure 1.

---

4 In principle, firms that have entirely very low paid employees are also excluded because if an employee earns less than ZAR 2 000 in a given tax year and no employee tax was deducted, the employee is not issued with an IRP5 or IT3a form (these are the two forms used for filing employee income). But this is unlikely to be material because the threshold is so low.
The distribution of firms in South Africa

19. Published data on the size distribution of South African firms are limited and not comparable with other countries. Existing data on the size distribution of firms are based on definitions of revenue (turnover) that differ across industries, as prescribed in the National Small Business Act. Financial data by firm size are published in Statistics South Africa’s Quarterly Financial Survey but the differences in the size definition depending on the sector limit the analysis, as does the focus on financial variables. The SARS-NT Panel allows a common size definition to be applied, based on employment as in the OECD Structural Business Statistics database, and the distribution to be compared for both employment and turnover. Following the OECD definitions, a micro enterprise has less than 10 employees, a small (including very small) firm has more than 9 and less than 50 employees, a medium-sized firm has more than 49 and less than 250 employees and a large firm has 250 employees or more.

20. This section focuses on the sectors within the “business sector” which account for around three-quarters of total gross value added; that is, agriculture, public administration, health, education, arts and recreation and other services are excluded. The manufacturing, construction and non-financial business services “macro sectors” are the focus of the most attention because these are the sectors covered by DynEmp. These three sectors account for around 13%, 4% and 40% of South Africa’s gross value added, respectively. However, it should be noted that manufacturing and transport and storage, as well as financial services, electricity, gas and water, are over-represented in the SARS-NT database (Pieterse et al., 2016).

21. Micro- and small enterprises are most prevalent in business services such as real estate activities and professional services (Figure 2, Panel A). Indeed, small professional services firms account for around 12% of all firms in the dataset. Manufacturing micro- and small enterprises account for around one-quarter of all firms in the dataset, even though the share of micro and small firms is lower than other sectors, because of the sector’s size. Mining and administrative services have the lowest share of small firms.

22. Larger firms account for the bulk of employment in capital-intensive industries and industries with large parastatal companies (mining, ICT, electricity and gas, transport and storage) as well as financial and insurance services (Figure 2, Panel B). Indeed, because of their importance in the dataset, employment in large financial services and manufacturing firms accounts for one-third of total employment in the dataset. By contrast, the highest share of employment in micro-enterprises (less than 10 employees) is in service industries such as real estate and professional services as well as construction and water and waste management. Because of its combination of relatively few small firms and very large firms, mining firms are, on average, the largest across sectors while the converse is true for real estate activities (Figure 3).
Figure 2. Firm size varies considerably across sectors
Distribution of firms within industries, average over 2011/12-2013/14, per cent

Note: Firm size is defined by number of employees, in size brackets shown. Sectors are ordered by the share of small firms by number. Numbers in parentheses denote each sector’s share of the total. There are fewer firms underlying Panel C as not all firms with employment data have valid turnover data. Data are for employing firms registered for company income tax; see Pieterse et al. (2016) for details.

Source: Authors’ calculations based on the SARS-NT Panel.
23. The distribution of sectors based on turnover (i.e., revenue) is similar to that of employment (Figure 2, Panel C). Manufacturing turnover is more concentrated in large firms than its employment is, which could reflect the link between size, productivity and capital intensity in this sector. Small firms account for a higher share of turnover relative to their employment share in most services sectors, particularly administrative and support services and financial and insurance services. This may imply some small firms in these industries have higher productivity. Likewise, in construction, micro enterprises account for a disproportionate share of turnover relative to employment.

![Figure 3. Average firm size by sector](image)

Note: Data are for employing firms registered for company income tax; see Pieterse et al. (2016) for details.

Source: Authors’ calculations based on the SARS-NT Panel.

24. Large firms are more prominent in South Africa than elsewhere. Large firms account for 65% of all employment and turnover in firms larger than 10 employees (micro enterprises are excluded to adjust for differences database coverage) (Figure 4). This share is similar to that in the United States, United Kingdom and Australia but much higher than the 45-50% in the typical country in this sample. Likewise, large firms account for a comparatively high share of firms by number: 5% of business sector firms have more than 250 employees. Cross-checks using the share of employment in the manufacturing sector from the labour force survey as the denominator also lead to the conclusion that the South African business sector is characterised by the dominance of large firms (Figure C.1). This is not a new finding: Fedderke, Obikili and Viegi (2016) and OECD (2008) (amongst others) have previously highlighted the concentration of the South African business sector.
The share of large firms in the business sector is high

Firms with employment of 250 or more as a percentage of those with employment of 10 or more

---

**A. Enterprises**

---

**B. Employment**

---

**C. Turnover**

---

**Note:** Data for South Africa are for 2012/13 and 2013 for other countries. Panel B is based on employees for Canada, Korea and the United States. The business sector comprises: mining; manufacturing; electricity and gas; water supply and waste management; construction; wholesale & retail trade; transport and storage; accommodation and food services; information and telecommunication services; finance and insurance services; real estate and rental activities; professional, scientific and technical services and administrative and support services.

**Source:** OECD Structural Business Statistics database and Authors’ calculations based on the SARS-NT Panel.
**Distribution of firms by age**

25. A firm’s age can be calculated based on year of registration, or some measure of beginning activity. For reasons discussed above and detailed in Appendix A, from 2010/11 onwards the first year with positive employment is used to denote entry and calculate age. Prior to this, employment is not a reliable indicator of entry. In the period from 2007/08-2009/10, it is difficult to distinguish genuine economic entrants. In the years before 2007/08 birth year can be based on company registration as a proxy for entry. The median gap between entry based on measures of activity and based on company registration is around 3 years for recent entrants, suggesting that the proxy is reasonable though imperfect (Figure A.4). Using age classes rather than age in years overcomes some of the measurement difficulty.

26. Start-ups (aged 0-2 years) are the main focus as they are more clearly identified. There is considerable variation across industries (Figure 5). The employment share of start-ups ranges from 2% in pharmaceuticals to 18% in publishing, audiovisual and broadcasting (Figure 5). In sectors typically considered as having low barriers to entry, such as accommodation and food services, the employment share of start-ups is higher. Conversely the employment share is lower in sectors that are more capital-intensive, like electricity and mining. This may in turn reflect barriers to entry such as access to finance and market structure.

27. There is a higher share of start-ups by number than by employment and it is more similar across industries. This is partly determined by the fact that young firms tend to be small. The employment share of start-ups is higher in industries where small firms predominate (Figure 6). On the other hand, large firms are typically older: 94% of employment at firms in manufacturing and business services with over 500 employees is at firms that are at least 6 years old. Amongst firms that are older (at least six years old), the median age rises with firm size.

**Figure 5. Share of start-ups by sector**

Firms aged 0-2 years by sector, average over 2012/13-2013/14

Note: Sectors are at the two-digit level.

Source: Authors’ calculations based on the SARS-NT Panel.
Firm entry and exit

28. Firm entry and exit are key mechanisms for the introduction of new innovation and technologies. In a competitive economy, the process of creative destruction shifts labour and capital resources to the firms that are most productive and away from those that are least productive, boosting overall productivity and wages. This section looks at firms’ entry and exit rates over time, by size, and by firm age to shed light on how this mechanism works in South Africa.

29. Due to the characteristics of the SARS-NT database, the standard DynEmp v.2 approach is adjusted in several ways. For 2010/11 onwards, the first year of positive employment is used to indicate economic entry with adjustments to improve the measure of age (detailed in Appendix 1). Because mergers and non-economic entry cannot be reliably identified, entrants with turnover in that year in the top 5% of their sector are reclassified as existing firms with missing age as a way of screening implausibly large entrants. As in DynEmp, exit is defined on the basis of the last year with positive employment. The entry and exit rates are defined as the number of entrants or exiting firms as a percentage of the total number of firms with positive employment.

30. Based on this adjusted definition the entry rate averaged around 8% over 2010/11-2013/14 (Figure 7). There is a slight downward trend. These entry rates are lower than those in Brazil or Turkey (OECD, 2016). Even though this includes some firms that are not genuinely new firms, the resulting overestimation should be smaller than the under-estimation using the standard DynEmp definition of entry.

31. The exit rate averaged 8% over the period and has increased, especially in 2013/14. That said, the uptick in 2013/14 should be interpreted with some caution as there may be firms that have been especially slow in filing their tax returns. Given that these estimates of entry are likely to be overstated due to non-economic entry, it is likely that the exit rate exceed the entry rate in recent years. This is worrying given the link between start-ups and economic growth, employment growth and productivity growth, as well as the apparent domination of large firms in many industries.
The trends in entry and exit are similar for manufacturing, construction and non-financial business services, which points to economy-wide factors behind these trends. There are several potential explanations. One is economic activity: the South African economy grew at an average rate of 5% in the five years before the 2009 recession but afterwards GDP growth slowed from 3.3% in 2010/11 to 2.4% by 2013/14. This likely contributed to a tough business environment. It may also be linked to a broader cross-country phenomenon; Criscuolo, Gal and Menon (2014) show that the activity of young firms declined over 2001-11. A more structural explanation is proposed by Fedderke, Obikili and Viegi (2016) who show that the sub-sectors of the manufacturing sector with high concentration and high barriers to entry (proxied by asset requirements) tend to have low rates of entry and high mark-ups. Recent case studies of specific value chains commissioned by the National Treasury highlight a range of barriers to entry including vertical integration and market power (Roberts, 2016). The OECD’s product market regulation indicators point to relatively high regulatory barriers to entrepreneurship.

The data show that the definition of entry makes a large difference. The standard DynEmp definition that uses year of company registration where available results in a very low entry rate (at around 1% over the sample period). The DynEmp measure will underestimate entry in the period considered here (2010/11-2013/14) if firms have registered with the company office (CIPC) well before they started operating. In practice the median gap between year of registration and beginning to operate appears to be three years but there are a number firms with much longer gaps (Figure A.4). The measure of entry based on positive employment will overstate entry if compliance has improved or the matching of employees to firms has improved over time or mergers or other types of non-economic entry are not adequately controlled for. Thus the two definitions could be considered as lower and upper bound estimates of the true entry rate if there was full compliance and firms started operating when they registered their company name. Broadening the data to all industries (agriculture, mining, financial services and public-sector related industries) increases the entry rate slightly but does not change the conclusion. Using positive turnover or positive expenses as an indicator of activity-based entry or exit means that entry and exit are independent of the process in which the company and employee databases are matched. This measure also suggests that entry rates have been below exit rates (Figure A.2).
Characteristics of entrants and exiting firms

34. As would be expected, entry and exit rates decline with firm size. Micro enterprises (0-9 employees) have the highest entry and exit rates, while very large firms (500+ employees) have low entry and exit rates (Figure 8). Indeed, micro enterprises account for about three-quarters of firms entering and 70% of firms exiting over 2010/11-2013/14. Fedderke, Obikili and Viegi (2016) also show that small firms account for the majority of firms entering and exiting in South Africa. This share of entrants that are micro enterprises is somewhat lower than in other countries which is consistent with the challenges faced by entrepreneurs highlighted in recent OECD Economic Survey of South Africa (OECD, 2017b).

Figure 8. Average entry and exit rates by size class

Note: Size is defined by number of employees. Data shown are for manufacturing, construction and non-financial business services over the period 2010/11 to 2013/14. Entry (exit) rates are calculated as the number of entering (exiting) units with positive employment over total number of units with positive employment in each size category shown.

Source: Authors’ calculations based on the SARS-NT Panel.

35. A closer look at exit dynamics shows that in each year the share of firms exiting is higher for young firms (less than three years old) and micro enterprises (Figure 9). The finding for young firms fits with other work showing that young firms are typically more vulnerable and their employment growth more volatile. However, firms aged three or more years comprise a much larger share of exiting firms than young firms due to the age distribution of South African firms. Regressing exit (as a dummy variable) on interactions between age and size classes (allowing for time fixed effects) suggests that as would be expected, young firms and micro enterprises (0-2 years and up to 9 employees) have the highest exit probability, followed by other micro enterprises. Thereafter the probability declines and is close to zero for firms with 20 or more employees that are more than 2 years old.

36. It is possible that the exit rate is pushed up by micro enterprises shifting to the turnover tax regime. However, this is unlikely for two reasons. Firstly, the Davis Tax Committee (2014) noted that take-up of the scheme had been disappointing, with only 7 827 firms registered in July 2013. Secondly, inspection of exit rates by firm size for turnover ZAR 200 000 either side of the threshold (ZAR 1 million) does not reveal any stark differences.
Figure 9. Exit rates by age and size over time

A. By age (years)

B. By size (number of employees)

Note: Data shown are for manufacturing, construction and non-financial business sectors. There is no observation for 0-2 years in 2010/11 because firms that appear with positive employment in the preceding two years are not clearly entrants and their age is treated as missing. See Appendix A for details.

Source: Authors’ calculations based on the SARS-NT Panel

The dynamics of firm growth and job creation

37. New firms contribute to employment dynamics through three channels: they create jobs by entering the market with positive employment; they destroy jobs if they close; and they create and destroy jobs by hiring and firing workers. Transition matrices in the DynEmp framework allow the employment growth performance of firms to be tracked and compared. The adjusted DynEmp definition that relies on the first year of positive employment from 2010/11 onwards is used for this analysis. In the SARS-NT sample, the cohort of firms that were created in 2010/11 contributed 2.6% to net job creation over 2010/11-2013/14 (Figure 10). The rate of net job creation from entrants was highest in construction and lowest in non-financial business services.

38. A better understanding of the contribution that new firms make towards employment growth can be gained using the growth decomposition framework developed by Calvino, Criscuolo and Menon (2015). The framework decomposes the performance of new firms into four key indicators: the start-up ratio; average size at entry; survival rate; and post-entry growth. Together, these indicators describe how many firms enter, how big they are, how many survive, and how they grow. The measures are described as follows:

- The start-up ratio measures the relative weight of entrepreneurship in the economy and is calculated as the number of entrants relative to the country’s total employment.

- The survival share reflects the extent to which the selection process of entrants is strong in an economy and is measured as the number of units that survive until (or beyond) the third year of life relative to the total number of starting units.

- Average size at entry measures the average number of employees for entrants.

- Average post-entry growth measures the growth performance and the scale-up potential of surviving start-ups. It is measured as the final over initial employment ratio of surviving entrants.
Figure 10. **Net job creation by surviving entrants**

Jobs created over three years by entrants in 2010/11 relative to total employment in sector

Note: The graph illustrates the growth of surviving entrants between 2010/11 and 2013/14. Business services excludes financial and insurance services. See Calvino, Criscuolo and Menon (2015) for full details of the calculation.

Source: Authors’ calculations based on the SARS-NT Panel.

Figure 11. **Components of the contribution of entrants to net job creation**

Source: Calvino, Criscuolo and Menon (2015).

39. Figure 12 shows each of the four components of job growth from entrants in 2010/11 for the three main sectors. There were 2.7 start-ups per thousand employees in these sectors that were still operating three years later (Panel A). On average, firms had 10 employees when they began operating (Panel C). Two-thirds of the new firms that entered in 2010/11 survived three years (Panel B), typically growing by 50% (Panel D).
40. There are some differences in dynamics between sectors:

- The low start-up ratio and high average size at entry in the manufacturing sector could be an indicator of barriers to entry, perhaps related to start-up capital. However, employment growth is also higher in this sector.
- The construction sector has the highest number of start-ups per thousand employees. There has been substantial infrastructure investment that may have contributed to growth in the sector; over 2010/11-2013/14 public sector infrastructure spending increased by 8% per year on average.
- Firms entering the business services sector also have a lower average size at start-up than manufacturing and the highest survival rate among the three macro sectors.

41. As discussed above, the definition of entry is overstated because corporate activity is not identified. Figure 8 shows that there are some large entrants. Reclassifying firms with 250 or more employees from entrants to existing firms has the expected effect of lowering the average size at start up to 9 employees. The biggest effect is in construction. When birth year is used as the primary measure of entry when available (following the standard Dynemp v.2 approach), the start-up ratio is very low (0.4 per 1000 employees) and the average size at entry is around 7 employees. The actual start-up rate is likely to be closer to the higher estimates. These results are broadly unchanged when the performance of entrants in 2011/12 is considered instead (see Figure C.2). As the database grows over time, these issues will be able to be investigated further.

Figure 12. Decomposition of job creation by surviving entrants in 2010/11 over three years

A. Start-up rate, per 1000 employees
B. Survival rate

C. Average size at start-up (employees)
D. Post-entry growth (final/initial employment)

Note: The graph illustrates the four components of the growth decomposition of surviving entrants between 2010/11 and 2013/14. Business services excludes financial and insurance services. See Calvino, Criscuolo and Menon (2015) for full details of the decomposition.

Source: Authors’ calculations based on the SARS-NT Panel
Firm dynamics compared to other countries

42. Comparing the net job creation by entrants and the sources of this growth help put the South African figures in context, bearing in mind the caveats mentioned above. New entrants contributed fewer jobs relative to existing employment than in Brazil or Turkey (about 6% and 7.5%, respectively) (Figure 13). The high degree of cross-country variation in net job creation by new firms is driven by variation in the start-up ratio, average size at entry and post-entry growth (Calvino, Criscuolo and Menon, 2015; Figure 14). By contrast, the survival rate is fairly similar across countries, at around 60%. The main points of difference between the South African dataset and the other DynEmp countries are the especially low start-up rate per 1000 employees and the high average size of new firms; in Brazil and Turkey the start-up rate is much higher and the average size at start-up is lower.

Figure 13. Net job creation by surviving entrants relative to total employment

Notes: The graph illustrates the ratio between employment at time t +3 of surviving entrants and overall country employment at time t. Figures report the average for different time periods. For South Africa, t=2010/11. For other countries t =2001, 2004 and 2007, conditional on their availability. Sectors covered are: manufacturing, construction, and non-financial business services. Owing to methodological differences, figures may deviate from officially published national statistics. Data for some countries are still preliminary.

Source: SARS-NT Panel; Calvino, Criscuolo and Menon (2015); OECD (2016); DynEmp v.2 database (accessed in June 2017); Authors’ calculations.

43. The difference between the South African data and the other countries likely reflects a combination of four factors: (i) structural factors related to the South African economy; (ii) the slowing of the economy after the financial crisis (which may have also reduced access to finance for start-ups); (iii) the self-selection aspect of the SARS-NT tax database and (iv) some corporate activity being falsely identified as entry in the SARS-NT database. For instance, the relatively high survival rate despite difficult economic conditions could reflect selection factors if only firms that expect to survive bother to register for tax, or structural factors such as high barriers to entry and weak market selection. It is also linked to the very low exit rates for firms with 20 or more employees that have survived their first two years highlighted above as well as the low entry rate. Interestingly, the sources of firm growth in South Africa look most similar to those in Norway and the United States – countries which also have fairly large firms dominating the economy, as shown earlier.
Figure 14. Cross-country comparison of growth decomposition

A. Start-up rate
Per 1000 employees

B. Survival rate
As a % of total number of entrants

C. Average size at entry
Number of employees

D. Post-entry growth
Final/initial employment

Note: The graph illustrates the four components of the growth decomposition of surviving entrants between period t and t+3. Figures report the average for different time periods. For South Africa, t=2010/11. For other countries t = 2001, 2004 and 2007, conditional on their availability. Sectors covered are: manufacturing, construction, and non-financial business services. Owing to methodological differences, figures may deviate from officially published national statistics. Data for some countries are still preliminary.

Source: SARS-NT Panel; Calvino, Criscuolo and Menon (2015); OECD (2016); DynEmp v.2 database (accessed in June 2017); Authors’ calculations.

Looking at average size at entry in more detail sheds some light on the drivers of South Africa’s high average size. However, one difference from the decomposition above is that these data include all entrants, both those that survive at least three years (as in the previous figures) and those that exit within three years from entry. Overall, the entrants tended to have around 12 employees at entry but there is considerable variation across industries (Figure 15). Average employment at entry ranges between 5 employees in real estate and rental activities, and more than 60 in publishing and broadcasting services. That said, average size at entry might be driven by the episodic entry of very large firms, especially in sectors such as pharmaceuticals, R&D or administrative services and publishing. This could include greenfield investment by foreign firms, or by mergers and acquisitions that have not been purged from the data, as suggested by the large difference in average size at entry between the two years shown. Across the sub-sectors, manufacturers tend to be larger at entry in both 2011/12 and 2012/13, suggesting that the higher average size is a characteristic of South Africa’s manufacturing sector, as expected. A high size at entry could be a sign of barriers to entry including access to finance and capital intensity, or (in this dataset) preferences for sole proprietorships/partnerships, or a large number of mergers relative to truly new firms.
Figure 15. **Average size at entry by sector**

Number of employees, all entrants

Note: Industries shown are in the manufacturing, construction and non-financial business services macro sectors. Industries are at the two-digit sectoral level with colours indicating which of the three macro sectors they belong to. Differences with earlier figures are due to the fact that these calculations include firms that did not survive three years.

Source: Authors’ calculations based on the SARS-NT Panel.

45. The DynEmp approach also shows the importance of existing large firms in driving employment flows. The transition matrices contain the gross job creation and destruction of firms from 2010/11 to 2013/14. Large firms (with 250 or more employees) made the largest contribution to gross job creation and destruction as well as net job creation (Figure 16). But the faster growth of small firms, particularly micro enterprises, meant that they created many more jobs than they destroyed so that net job creation was not dissimilar to large firms. The dynamics of micro-entrants are explored in more detail in the next section.

Figure 16. **Job creation and destruction by existing surviving firms**

Change in number of employees from 2010-11 to 2013/14 by initial firm size

Note: Industries shown are in the manufacturing, construction and non-financial business services sectors. Jobs created by entrants and exiting firms are not included.

Source: Authors’ calculations based on the SARS-NT Panel.
**Growth of micro-entrants**

46. The cross-country work related to the DynEmp project has shown that start-ups and young firms are drivers of job creation. But only a small number of firms account for this growth in jobs. These are the innovative, transformational firms that scale up quickly. They are especially interesting because their growth is typically what offsets the job losses that come from the high exit rate of young firms (Calvino, Criscuolo and Menon, 2016). The SARS-NT Panel allows the growth of micro-entrants to be tracked over time, specifically, the status of the 2010/11 cohort of entrants in 2013/14.

47. In South Africa, there are more surviving start-ups than exiting start-ups, which is in line with other DynEmp countries (Figure 17, see also Figure 11 in Calvino, Criscuolo and Menon, 2015, and Figures 10 and 11 in Criscuolo, Gal and Menon, 2014). After three years, 12% of micro start-ups had grown into a higher size class, compared to an average of 3% and a maximum of 8% in other DynEmp countries. As in most of the DynEmp countries, micro-entrants are net job creators (Calvino, Criscuolo and Menon, 2016). There is some evidence of stronger dynamics in manufacturing and construction than business services, perhaps linked to market selection (Figure 18). In manufacturing and construction, a higher share of firms exited, and a higher share also grew beyond micro enterprises; this manifests itself in a higher share of job creation coming from growing firms. Nonetheless, the manufacturing jobs destroyed by micro enterprises closing exceeded those created by the survivors.

Figure 17. **Few micro-entrants grow but they are important job-creators**

Performance of 2010/11 micro-entrants over three years

![Graph showing distribution of micro-entrants in 2013/14, distribution of micro-entrants by employment, and contribution of micro-entrants to net job creation.](image)

Note: Micro-entrants are firms with less than 10 employees. The sample is restricted to manufacturing, construction and non-financial business services.

Source: Authors’ calculations based on the SARS-NT Panel
The role of age in firm growth

As for other countries, and in other studies, young firms grow faster than old firms. In 2012/13 the difference between average employment growth of young (0-2 years old) and older (6 or more years old) firms was more than five percentage points in all industries considered here, and highest in the manufacturing sector (Figure 19). Data for 2012/13 and 2013/14 suggest that in all industries, small young firms (0-2 years and less than 50 employees) account for a large share of gross job creation but a smaller share of gross job destruction; that is, they are important net job creators (Figure 20). Small young firms in real estate activities, coke manufacturing and accommodation and food services made the largest net contributions to sectoral employment growth in these years. The data also reveal that old firms were net job destroyers in 2012/13 and 2013/14. Taken together, this points to the importance of understanding the role of firm age in employment dynamics, rather than firm size alone, a finding highlighted in studies of other countries (Criscuolo, Gal and Menon, 2014; Haltiwanger, Jarmin and Miranda, 2013).
Figure 19. **Average employment growth by sector and age group**

Average of 2012/13-2013/14, per cent

Note: Firms with 3-5 years are not well identified and are omitted here (see Appendix A for details). Sectors shown are in the manufacturing, construction and non-financial business services sectors. The publishing, audiovisual and broadcasting sector is not shown due to its extreme value: average employment growth in 0-2 year old firms was -22% and in 6 or more year-old firms 1%.

Source: Authors’ calculations based on the SARS-NT Panel.

Figure 20. **Small young firms are net job creators**

Share of gross job creation and gross job destruction in each sector, 2012/13-2013/14 average

Note: Small young firms are firms with less than 50 employees and are 0-2 years old. Industries shown are in the manufacturing, construction and non-financial business services sectors. Industries are ordered by their net contribution to job creation (i.e. gross job creation less gross job destruction).

Source: Authors’ calculations based on SARS-NT panel

49. The effects of size and age on employment growth can be considered simultaneously and controlling for economy-wide conditions using the distributed regressions in the DynEmp code (with some minor adjustments). Employment growth is regressed on interactions of size and age, with year fixed effects and robust standard errors at the industry level. In most other countries, small firms tend to grow faster than large firms — a violation of “Gibrat’s law” that posits that growth is independent of size — and this is only restored when age is controlled for. However, in South Africa, the effect of age — and being young more specifically - is more important than size across all size categories (Figure 21). Although the overall
explanatory power of the regressions is low, the slow, or negative, employment growth at older micro enterprises warrants more attention.

Figure 21. Net employment growth by age and size
Regression coefficients on employment growth from year earlier

Note: Employment growth is calculated as \((E_t - E_{t-1})/[0.5*(E_t + E_{t-1})]\). Filled markers denote that the coefficient is statistically significant at the 10% level. The regressions include year fixed effects and robust standard errors clustered at the 2-digit industry level.

Source: Authors’ calculations based on SARS-NT database.

Conclusions and next steps

50. The creation of the SARS-NT Panel has spurred a range of new research into South African firms. Combining this new dataset with the OECD’s DynEmp framework has generated new insights about the dynamics of firms in South Africa, particularly in manufacturing, construction and non-financial business services. The latter two sectors have been studied less often than manufacturing in South Africa.

51. Key findings relating to the firm creation, growth and destruction are:

- The entry rate appears likely to have been below the exit rate in recent years, which means that firms are growing older. The reasons behind this are not clear but seem to be economy-wide.
- Dynamics are somewhat different across sectors. Construction has a higher start-up rate and lower survival rate than manufacturing or non-financial business services. In the business services sector entrants are typically smaller than in other sectors, which may reflect lower barriers to entry.
- Two thirds of the micro entrants (firms with less than 10 employees at entry) in 2010/11 survived to 2013/14 but of these, most remained small. A small fraction (11%) grew beyond 10 employees but these disproportionately contributed to jobs growth.
- Young firms have disproportionately contributed to employment growth in recent years. Even with slower GDP and employment growth, young firms remained net job creators. However, a notable feature is the importance of large incumbent firms for job creation and destruction.

52. This paper contributes to a better understanding of firm dynamics in South Africa. The definitions and measures created can facilitate further empirical work to better understand the relative importance of
firm size and age and barriers to firm growth, and ultimately policy-making. For instance, one possibility is to use the sectoral variation to try to understand why the typical firm size at start-up is so high in some sectors. Potential candidate explanations are constraints in accessing finance or skilled workers. Regulatory burdens are disproportionately higher for small firms so the OECD’s measures of sectoral exposure to upstream network and services sector regulation can be used. Sectoral concentration or market power of incumbents may also affect start-ups’ performance (as in Fedderke, Obikili and Viegi, 2016).
APPENDIX A: DEFINING ENTRY AND EXIT

53. The special characteristics of the SARS-NT dataset means the standard DynEmp v.2 definition of a firm’s creation date has been adapted and the dataset used in the analysis is shortened to 2010/11-2013/14. There are several issues that needed to be addressed:

- The standard DynEmp v.2 definition of entry uses a firm’s registration date if it exists, supplemented by first year of appearance with positive employment if the former is missing. But firms may register a company name and then wait years before deciding to build a company. The firm’s true age would appear older than its actual age based on years of operation.

- Firms essentially select into the SARS-NT dataset because they choose to pay company income tax and to register their employees. Thus, in the first three years of the dataset (2007/08-2009/10), a large number of firms enter the database with positive employment or positive turnover (Figure A.1, Panel A). Moreover, the firms being measured as entrants in these years using the standard DynEmp definition are considerably larger than in later years (Table A.1, Panel A). Using the definition of entry based on positive employment leads to more large firms appearing in the data in each year, some of which are unlikely to be genuinely new firms (Table A.1, Panel B.) This indicates the upward bias in the measure of entry and average size at entry when using positive employment. The additional screening for non-economic entrants based on having turnover in the top 5% of the sector reduces the bias somewhat (Panel C.) Although this issue complicates analysis of firm dynamics, it is a sign the massive modernisation programme and compliance drive that SARS undertook led to a correspondingly huge increase in formalisation.

- To some extent entry is driven by the researchers’ ability to match employers PAYE records to firms’ CIT records. The working assumption is that this ability may have improved in the early years of the sample as data quality improved but that from 2010/11 onwards it is stable so there is no bias in the period of analysis.

- The quality of the data in 2007/08 is also lower than in later years (Pieterse, Gavin and Kreuser, 2016), which implies greater caution in using information from 2007/08.

54. Because of the importance of firm age, and particularly young firms, in empirical work, the classification system used aims to maximise the information that is reliable and take advantage of the full 2007/08-2013/14 dataset, as set out in Figure A.2. It identifies active firms using their first year of positive employment from 2010/11 and supplements this with information on their year of registration in the case of firms that are clearly older. The key features are:

- Firms with registration year from 2007/08 onwards are treated as being created in their first year of positive employment if the first year of positive employment is 2010/11 or later.

- Firms with positive employment in 2010/11 (when the major compliance-related adjustment has occurred) that registered before 2007/08 are treated as having been active since their registration year.

55. In 2012/13 and 2013/14 it is therefore possible to reliably identify firms that are 0-2 years old. Firms that are more than 6 years old can also be identified. In the implementation of DynEmp programme the data are left-censored at 2008/09 so firms born in 2008/09 will have a missing age class. (Firms born in 2008/09 would be included in the “six years or more” age class in 2014/15.)
from 2010/11 onwards as reliable is seen by the high and stable share of micro enterprises in all entrants from 2010/11 onwards (over 80%, Table A.1, Panel A). There is some difference in entry rates for 2010/11 but it is not large (Figure A.4).

Figure A.1. Entry and exit based on different definitions

A. Entering firms, by year
- DynEmp definition of entry
- First year of positive turnover
- Definition adopted

B. Exiting firms, by year
- Last year of positive employment
- Last year of positive turnover
- Definition adopted

Note: Sectors covered are manufacturing, construction and non-financial business services.

Source: Authors’ calculations based on the SARS-NT Panel.
Figure A.2. Dating firm birth and classifying age

Figure A.3. Entry and exit based on different definitions

Note: “Registered” refers to registration with the Company and Intellectual Property Commission.

1. To be more precise, birth year is treated as missing for firms with positive employment in 2007/08 but firms with positive employment in 2008/09 and 2009/10 are left-censored and would appear as being 6 or more years old in 2014/15.

56. As mentioned above, there is a gap between the year of registration and the year of becoming economically active. To investigate this further, the firm’s age based on the year of company registration is compared to the age based on measures of activity (Figure A.4). The focus is on the period from 2010/11 onwards when compliance-based entry should have stabilised. The gap between the measures is similar for the three measures of activity. The median is three years but there is a longer right tail. It is possible that the gap is caused by company names being registered some time in advance of trading beginning. Alternatively, the effect of increasing compliance may be very strong even after the initial few years. Because of the newness of the database there is some uncertainty around which explanation is most important. As the dataset grows, this will be easier to discern.
Figure A.4. **Entry and exit rates by year and firm size**

A. Firm entry by year  
B. Firm exit

Note: Sectors covered are manufacturing, construction and non-financial business services.

Source: Authors’ calculations based on the SARS-NT Panel.

Figure A.5. **Distribution of gap between different measures of age**

Age based on year of registration less age based on activity-based measure

Note: Sample covers macro sectors over 2010/11 to 2013/14

Source: Authors’ calculations based on the SARS-NT Panel.
### Table A.1. Size at entry over time

**A. Number of firms based on standard DynEmp definition of entry**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9 employees</td>
<td>445</td>
<td>2,181</td>
<td>6,777</td>
<td>1,973</td>
<td>1,436</td>
<td>1,544</td>
<td>1,145</td>
</tr>
<tr>
<td>10-49 employees</td>
<td>244</td>
<td>631</td>
<td>2,068</td>
<td>372</td>
<td>241</td>
<td>284</td>
<td>236</td>
</tr>
<tr>
<td>50-99 employees</td>
<td>49</td>
<td>73</td>
<td>198</td>
<td>28</td>
<td>12</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>100-249 employees</td>
<td>18</td>
<td>37</td>
<td>63</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>250-499 employees</td>
<td>5</td>
<td>4</td>
<td>12</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500+ employees</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>768</td>
<td>2,927</td>
<td>9,123</td>
<td>2,381</td>
<td>1,696</td>
<td>1,858</td>
<td>1,399</td>
</tr>
<tr>
<td>Memo: micro-enterprises % of total</td>
<td>58</td>
<td>75</td>
<td>74</td>
<td>83</td>
<td>85</td>
<td>83</td>
<td>82</td>
</tr>
</tbody>
</table>

**B. Number of firms based on adjusted definition of entry without cleaning largest firms**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9 employees</td>
<td>4,256</td>
<td>10,560</td>
<td></td>
<td>11,155</td>
<td>8,987</td>
<td>9,418</td>
<td>8,562</td>
</tr>
<tr>
<td>10-49 employees</td>
<td>1,174</td>
<td>3,191</td>
<td></td>
<td>3,275</td>
<td>2,727</td>
<td>2,841</td>
<td>2,751</td>
</tr>
<tr>
<td>50-99 employees</td>
<td>128</td>
<td>344</td>
<td></td>
<td>401</td>
<td>334</td>
<td>337</td>
<td>322</td>
</tr>
<tr>
<td>100-249 employees</td>
<td>73</td>
<td>151</td>
<td></td>
<td>193</td>
<td>148</td>
<td>177</td>
<td>169</td>
</tr>
<tr>
<td>250-499 employees</td>
<td>12</td>
<td>29</td>
<td></td>
<td>54</td>
<td>37</td>
<td>50</td>
<td>34</td>
</tr>
<tr>
<td>500+ employees</td>
<td>7</td>
<td>13</td>
<td></td>
<td>22</td>
<td>26</td>
<td>29</td>
<td>39</td>
</tr>
<tr>
<td>Total</td>
<td>5,650</td>
<td>14,288</td>
<td></td>
<td>15,100</td>
<td>12,259</td>
<td>12,852</td>
<td>11,877</td>
</tr>
<tr>
<td>Memo: micro-enterprises % of total</td>
<td>75%</td>
<td>74%</td>
<td></td>
<td>74%</td>
<td>73%</td>
<td>73%</td>
<td>72%</td>
</tr>
</tbody>
</table>

**C. Number of firms based on adjusted definition of entry**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9 employees</td>
<td>11,377</td>
<td>8,808</td>
<td>9,251</td>
<td>8,438</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-49 employees</td>
<td>3,065</td>
<td>2,426</td>
<td>2,550</td>
<td>2,545</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-99 employees</td>
<td>287</td>
<td>233</td>
<td>241</td>
<td>267</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100-249 employees</td>
<td>112</td>
<td>81</td>
<td>102</td>
<td>107</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250-499 employees</td>
<td>33</td>
<td>16</td>
<td>31</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500+ employees</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14,881</td>
<td>11,581</td>
<td>12,192</td>
<td>11,388</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memo: micro-enterprises % of total</td>
<td>76%</td>
<td>76%</td>
<td></td>
<td>76%</td>
<td>76%</td>
<td>76%</td>
<td>74%</td>
</tr>
</tbody>
</table>

**Note:** Sectors covered are manufacturing, construction and non-financial business services. The standard DynEmp definition of entry uses a firm’s birth year (registration year) supplemented by the first year the firm appears in the dataset with positive employment where birth year is missing. The adjusted definition of entry uses the first year the firm appears in the dataset with positive employment as set out in Appendix Table A.2.

**Source:** Authors’ calculations based on the SARS-NT Panel.
APPENDIX B. COMPARISON WITH OTHER DATA SOURCES

A comparison with other data sources

Because the tax dataset does not include very small businesses, the distribution of employment and turnover accounted for by small firms will be understated, even in the formal sector. For example, the self-reported labour force data imply that 11% of manufacturing sector employment was in micro-enterprises (less than 10 employees), compared to 5% here. (Including informal employment would take the total to 21%.) Likewise, the World Bank (2013) estimates that micro-enterprises employed around one-quarter of manufacturing sector employees in 2005-07 and around 45% of services sector employees. One other source of difference is that the analysis here is based on the definition of a firm, whereas employees may be considering the plant level data, which they know more about. In sum, no dataset is perfect and the findings that follow should be interpreted in the context of the formal sector only, knowing that many self-employed formal and informal workers fall outside the sample.

Table B.1. Comparison of coverage of two data sources used to assess the distribution of firms

<table>
<thead>
<tr>
<th>Definitions</th>
<th>SARS-NT Panel</th>
<th>Labour force survey</th>
<th>Survey of Employers and self-employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm</td>
<td>Registered for tax</td>
<td>Employer or own-account worker</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive employment in at least one year during the sample period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment by firm size</td>
<td>Number of employees (jobs) in firm</td>
<td>Number of employees at “place of work”</td>
<td></td>
</tr>
<tr>
<td>Coverage by type of firm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informal firms</td>
<td>Not included</td>
<td>Included (but firms only include the self-employed)</td>
<td></td>
</tr>
<tr>
<td>Formal firms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own-account workers / sole proprietorships</td>
<td>Not included</td>
<td>Included (firms and employment)</td>
<td></td>
</tr>
<tr>
<td>Partnerships</td>
<td>Not included</td>
<td>Included (firms and employment)</td>
<td></td>
</tr>
<tr>
<td>Micro-enterprises</td>
<td>Included if paying company tax and PAYE for at least one employees but excluded if registered for turnover tax</td>
<td>Firms included if one person owns it. employment less well identified</td>
<td></td>
</tr>
<tr>
<td>Small to large firms</td>
<td>Well identified</td>
<td>Less well identified. The labour force survey categories</td>
<td></td>
</tr>
</tbody>
</table>

The definitional differences between the two data sources are in the expected directions. The SARS-NT dataset contains fewer small firms and more large firms in absolute and relative terms, even when informality is excluded. The differences between informal and formal sector estimates are largest for small firms; that is, there is little informality in large firms.
### Table B.2. Comparison of data sources

<table>
<thead>
<tr>
<th></th>
<th>By number of firms ('000)</th>
<th>By employment ('000)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-9</td>
<td>10-19</td>
</tr>
<tr>
<td><strong>Manufacturing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SARS-NT dataset</td>
<td>24.6</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>(41.5)</td>
<td>(17.1)</td>
</tr>
<tr>
<td>LFS – all</td>
<td>183.4</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>(89.9)</td>
<td>(6.4)</td>
</tr>
<tr>
<td>LFS – formal</td>
<td>46.8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>(72.1)</td>
<td>(17)</td>
</tr>
<tr>
<td><strong>Construction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SARS-NT dataset</td>
<td>8.2</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>(44.4)</td>
<td>(15.6)</td>
</tr>
<tr>
<td>LFS – all</td>
<td>278.3</td>
<td>17.4</td>
</tr>
<tr>
<td></td>
<td>(89.4)</td>
<td>(5.6)</td>
</tr>
<tr>
<td>LFS – formal</td>
<td>58.7</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>(65.9)</td>
<td>(18.2)</td>
</tr>
<tr>
<td><strong>Business services</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SARS-NT dataset</td>
<td>48.8</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>(46.9)</td>
<td>(15)</td>
</tr>
<tr>
<td>LFS – all</td>
<td>1236.7</td>
<td>41.3</td>
</tr>
<tr>
<td></td>
<td>(93.9)</td>
<td>(3.1)</td>
</tr>
<tr>
<td>LFS – formal</td>
<td>314.6</td>
<td>38.7</td>
</tr>
<tr>
<td></td>
<td>(80.6)</td>
<td>(9.9)</td>
</tr>
</tbody>
</table>

**Note:** The labour force data are weighted but there is no imputation for the firms in the SARS-NT dataset that did not report the size of their workplace. Business services exclude financial services. The number of firms is calculated from the labour force survey based on the number of respondents reporting being self-employed and the size of the company that they worked in. The number of employees in each size firm is based on all responses to the latter question. This is the methodology adopted in BER (2016) to estimate the number of small firms in South Africa.

**Source:** Authors’ calculations based on the SARS-NT Panel and Labour Market Dynamics database.
APPENDIX C. ROBUSTNESS CHECKS

The importance of large firms

59. One way of abstracting from the differences in the coverage of the smallest firms that hinder cross-country comparability is to focus on the role of the largest firms and compare these to overall employment. That is, if employment in the large firms is well measured and the labour force survey is representative, then the ratio should be a good indicator of the importance of large firms in a sector. This is shown for manufacturing employment in Figure C.1. There are two values for South Africa because the labour force estimate is below the estimate from the SARS-NT Panel, which is also the case in Denmark, Finland, France, Luxembourg, Norway and Sweden. In these cases, the estimates of large firms’ share of employment are still above 40%. This comparison is further evidence that South African manufacturing employment is relatively concentrated in large firms.

![Figure C.1](image)

**Figure C.1. Large firms’ share of manufacturing employment**
Share of manufacturing employment according to labour force survey, per cent

*Note:* Large firms are those with 250 or more employees. ZAF (alt) uses manufacturing employment in the SARS-NT Panel as the denominator.


Definition of employment

60. As discussed in the text and Pieterse, Gavin and Kreuser (2016), there are 12 different measures of employment all of which are based on the number of employees. The first set of employment measures is the *total periods set*. This set has three unweighted and weighted variables respectively which make use of the number of periods worked and total number of periods in a tax year. The weighted variables account for instances where individuals were not employed for the full duration of a particular year. The second set of the six employment measures uses the periods employed start and end dates (*the dates set*) to weight workers and assign them to a firm in its financial year.

61. Re-running the decomposition of job creation by new entrants using these other definitions of employment does not generally change the conclusions although there are some differences that could be investigated further. For instance, using the “periods” definition but weighting employment by the proportion of days worked raises the start-up rate and lowers the average size at entry of survivors somewhat, but the start-up rate would still be comparatively low and the average size comparatively high.
Sensitivity of the year chosen for the growth decomposition

To ensure that the growth decomposition is not sensitive to the choice of year of entry, the calculation was replicated for firms that entered in 2011/12. Because the last year of the database is 2013/14, the decomposition was run for performance over two years, rather than three years and the outcome compared to the results for two-year performance of the 2010/11 cohort (Figure C.2).

Moving to two-year performance raises the survival rate to 75% and lowers post-entry growth slightly, as expected. The broad inter-sectoral patterns seen in the three-year performance remain, although there are some differences in the levels, especially for construction. Construction still has the highest start-up rate and manufacturing has the lowest. The business services sector has the lowest average size at start-up. However, there seem to be some cohort effects between surviving entrants from 2010/11 and 2011/12, particularly comparing the start-up rate in construction, which falls dramatically in 2011/12. The relative dynamics of manufacturing firms compared to the non-financial business sector is similar for 2010/11 and 2011/12.

**Figure C.2. Decomposition of employment growth of surviving entrants over two years**

- **A. Start-up rate (per 1000 employees)**
- **B. Survival rate (after 2 years)**
- **C. Average size at start-up**
- **D. Post-entry growth (after 2 years)**

*Note: Sectors covered are manufacturing, construction and non-financial business services. Post entry growth is the ratio of final to initial employment.*

*Source: Authors’ calculations based on the SARS-NT Panel.*
REFERENCES


http://dx.doi.org/10.1787/eco_surveys-zaf-2017-en


