

# FIRM DYNAMICS IN SOUTH AFRICA<sup>1</sup>

By Mpho Tsebe<sup>2</sup>, Veron Vukeya<sup>2</sup>, Christine Lewis<sup>3</sup>, Flavio Calvino<sup>3</sup> and Chiara Criscuolo<sup>3</sup>

*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 of formal sector firms was probably 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 are consistent with the low rates of entrepreneurial activity and the presence of barriers to firm entry highlighted in the existing literature on the South African economy. As in other countries, young firms have disproportionately contributed to employment growth and remained net job creators even as GDP growth slowed. Nonetheless, large firms are particularly prominent in the South African economy, including as net job creators.*

## Introduction

1. The slowdown in the South African economy and the 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 drivers of job creation. To date, the study of firm dynamics in South Africa has been hindered by a lack of firm-level data. Some studies have used 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 does not capture young firms well (Kerr *et al.*, 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 Statistics survey of employers, Kerr *et al.* (2014) find that large firms are better net creators of jobs than small firms. This contrasts with findings for other (mostly OECD)

---

<sup>1</sup> 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 Lukas Klein-Rueschkamp, Lenka Wilderova and participants at a internal seminars for their feedback and suggestions. Further thanks go to Pedro Herrera Gimenez for statistical assistance and Raquel Paramo and Heloise Wickramanayake for editorial assistance.

<sup>2</sup> South African National Treasury at the time of work on this paper.

<sup>3</sup> Organisation for Economic Co-operation and Development (OECD)

countries that young firms have been the main creators of jobs (Criscuolo *et al.*, 2014; Calvino *et al.*, 2015). More recent work with the SARS-NT Panel has also highlighted worrying dynamics in the South African economy. Fedderke *et al.* (2018) focus on the manufacturing sector and find that the firm exit rate has exceeded the firm entry rate and that entry and exit are mostly by small firms. Overall, entry and exit contribute significantly to job creation and destruction but so does employment growth at existing firms. Using the SARS-NT Panel for all sectors, Kerr (2018) also finds that worker flows and job reallocation rates tend to decline with firm size.

3. Firm dynamics in South Africa have long been a concern of researchers and policy makers. Previous OECD *Economic Surveys* of South Africa have highlighted the role of product market regulation, licensing and access to finance in stifling competition and a dynamic business environment (OECD, 2008, 2013, 2015a, 2017a). The most recent *Survey* drew attention to large gaps in basic skills and entrepreneurial competencies, as well as access to finance, that are also likely to be weighing on entrepreneurship in South Africa. A recent series of studies commissioned by the South African National Treasury showed the pervasiveness of barriers to entry across sectors as well as heterogeneity in how these barriers originate and operate (Roberts, 2016).

4. 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, described more fully in the following section. These include that the tax dataset by its nature comprises firms that have chosen to register for company income tax, whereas business register data are generally more comprehensive. Dynamics in the informal sector cannot be captured until a firm formalises. As an emerging economy, the informal sector is likely to be larger in South Africa than advanced OECD economies but labour force data suggest that South Africa's informal sector is smaller than other emerging economies at similar income levels (OECD, 2017a). The South African data are for the post-crisis period when employment growth slowed, whereas the data for other countries span periods of healthier growth.

5. 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 the analysis suggests that firm entry rates have been below exit rates in recent years 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 coverage of the dataset as well as the presence of barriers to entry or to formalisation highlighted by other research for South Africa (OECD, 2017a). 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***

6. 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 *et al.*, 2018). The data are available from the 2007/08 tax year onwards.

7. 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.

8. 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.

9. 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 macro-sectors: manufacturing, non-financial business services (including trade, accommodation, transport, real estate, administration and professional services) and construction. This facilitates cross-country comparisons, as described below.

10. 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. 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. Only firms that have information on their economic sector are kept. Taking 2013/14 as an example, the number of firms is reduced from 781 085 to 247 803 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 and less complete later years. 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**

Tax-registered firms

	Full SARS-NT Corporate Income Tax panel		Active firms with sector classification		Active firms in manufacturing, construction and non-financial business services	
	Number of firms	Firms with positive employment	Number of firms	Firms with positive employment	Number of firms	Firms with positive employment
2007/08	608 271	201 212	233 937	131 065	162 305	92 918
2008/09	903 521	228 217	251 308	160 055	173 467	112 878
2009/10	940 321	268 032	223 041	209 811	150 814	141 980
2010/11	925 339	276 665	247 789	236 015	166 613	158 662
2011/12	898 335	276 421	250 944	236 762	167 709	158 144
2012/13	873 755	275 225	253 313	244 373	168 560	162 573
2013/14	781 085	267 407	253 973	247 803	168 189	164 097

*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***

11. 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, social security and tax 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 this data infrastructure to study employment and business dynamics across countries (see among others Criscuolo *et al.*, 2014; Calvino *et al.*, 2015; 2016; OECD, 2016, 2017b).

12. 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 country 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.

13. 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 *et al.* (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 pre-defined econometric models estimated on the micro-data, 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.

14. 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) or for three “macro sectors”, namely manufacturing, construction and non-financial business services.

15. The DynEmp approach allows a great degree of harmonisation along different dimensions and significantly limits sample selection effects 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***

16. 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 *et al.*, (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.

17. Relative to other countries in the DynEmp project, coverage of the SARS-NT Panel is constrained to only firms that are registered for company income tax that actually submit their tax returns. This implies that the Panel excludes firms in the informal sector, sole proprietorships and partnerships and micro-enterprises that are registered for the turnover tax (turnover below ZAR 1 million, or USD 75 000).<sup>4</sup> Consequently, the dataset is likely to disproportionately omit very young firms and very small firms (as discussed in Appendix B) and the findings should be interpreted in the context of the formal sector only. However the problem is mitigated by the fact that South Africa’s informal sector appears to be relatively small given its income level (OECD, 2017a). Informal employment is around 30% of total employment compared to two-thirds in sub-Saharan Africa and half in Latin America (Oosthuizen *et al.*, 2017) According to the labour force survey, self-employment represents just over 10% of total employment, which is low given South Africa’s level of income per capita (OECD, 2017a).

18. 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

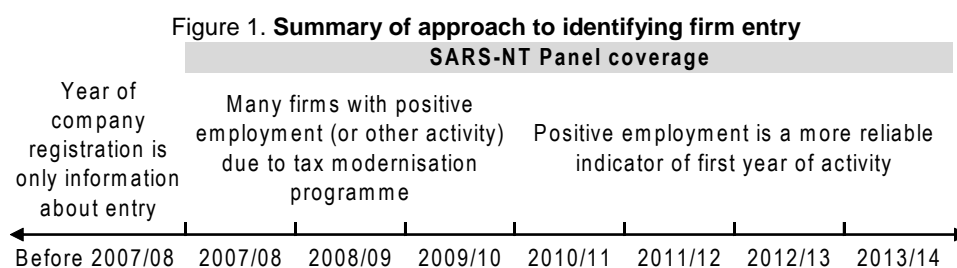
---

<sup>4</sup> 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.

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 income 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 the firm 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. Indeed, other research using the SARS-NT Panel uses the first year of positive employment for birth year, rather than year of registration (Fedderke *et al.*, 2018; Kerr, 2018).

19. To address these two issues, the definition of a firm's creation date is altered from the standard DynEmp definition. 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 and detailed in in Appendix A. This means that more use is made of entry based on positive employment than for other countries in DynEmp v.2.



## The distribution of firms in South Africa

20. 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 sectors, 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.

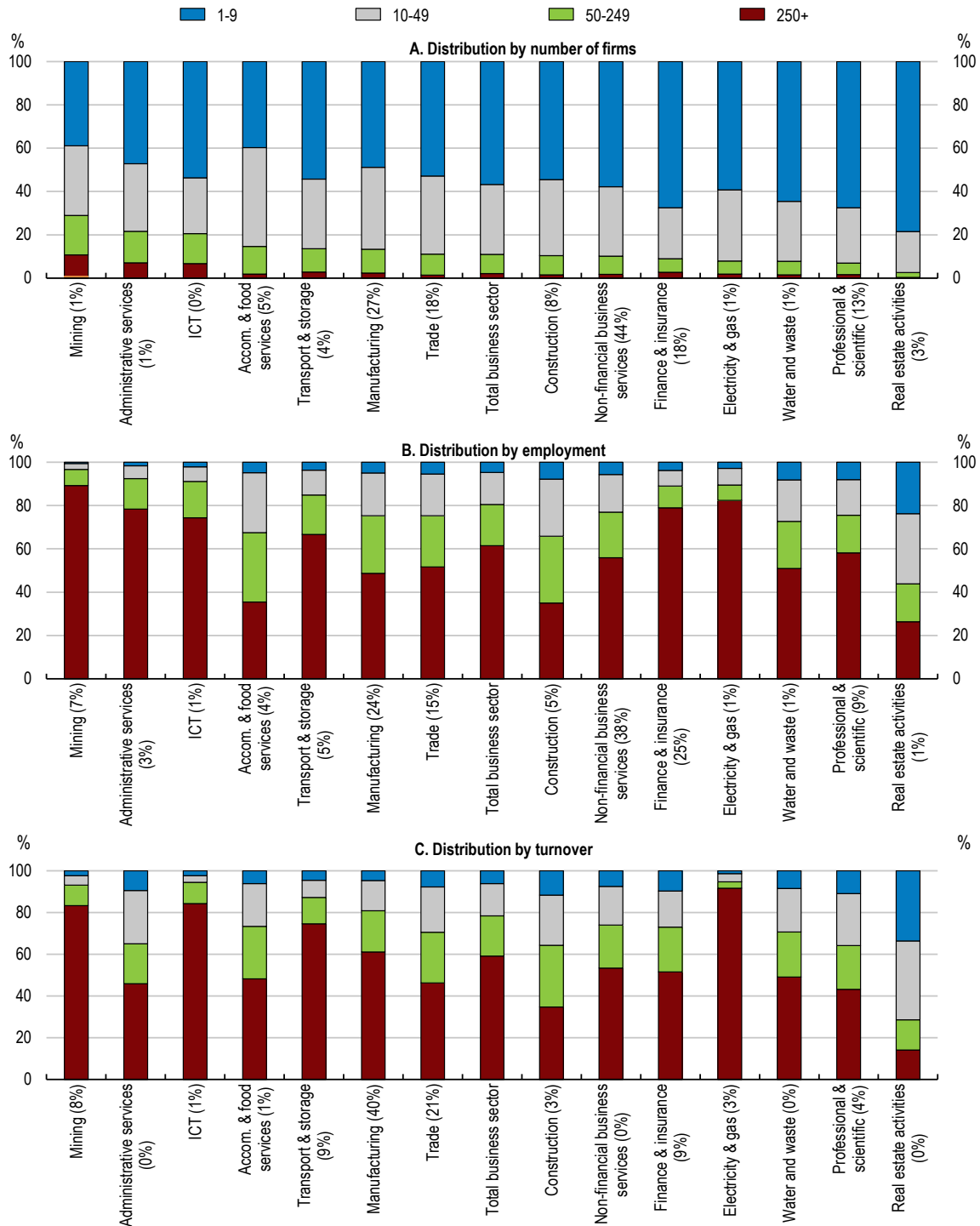
21. This section focuses on the sectors within the "business sector" which account for around three-quarters of total gross value-added; that is, agriculture, public administration, health, education, arts and recreation and other services are excluded. The manufacturing, construction and non-financial business services "macro sectors" are the focus of the most attention because these are the sectors the focus of cross-country comparisons using 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.*, 2018).

22. Micro- and small enterprises account for almost 90% of business sector firms in the dataset. They 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 in manufacturing is lower than in other sectors. Mining and administrative services have the lowest share of small firms.

23. Larger firms account for the bulk of employment, with three-fifths of business sector employees in large firms. This share is higher 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). Because of its combination of relatively few small firms and many very large firms, mining firms are, on average, the largest across sectors while the converse is true for real estate activities (Figure 3). Large electricity and gas firms typically have almost 4 000 employees. 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.

Figure 2. Firm size varies considerably across sectors

Distribution of firms within industries, average over 2011/12-2013/14

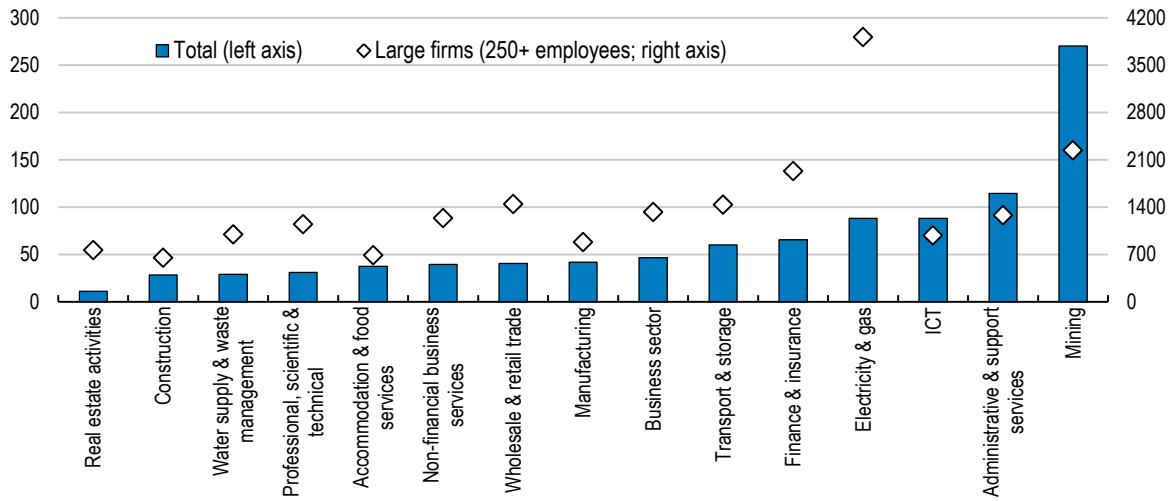


Note: Firm size is defined by number of employees, in size brackets shown. Sectors are ordered by the share of firms with less than 50 employees by number. Numbers in parentheses denote each sector's share of the total business sector. There are fewer firms underlying Panel C as not all firms with employment data have valid turnover data. Data are for employing firms registered for company income tax; see Pieterse et al. (2018) for details.

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



**Figure 3. Average firm size by sector**

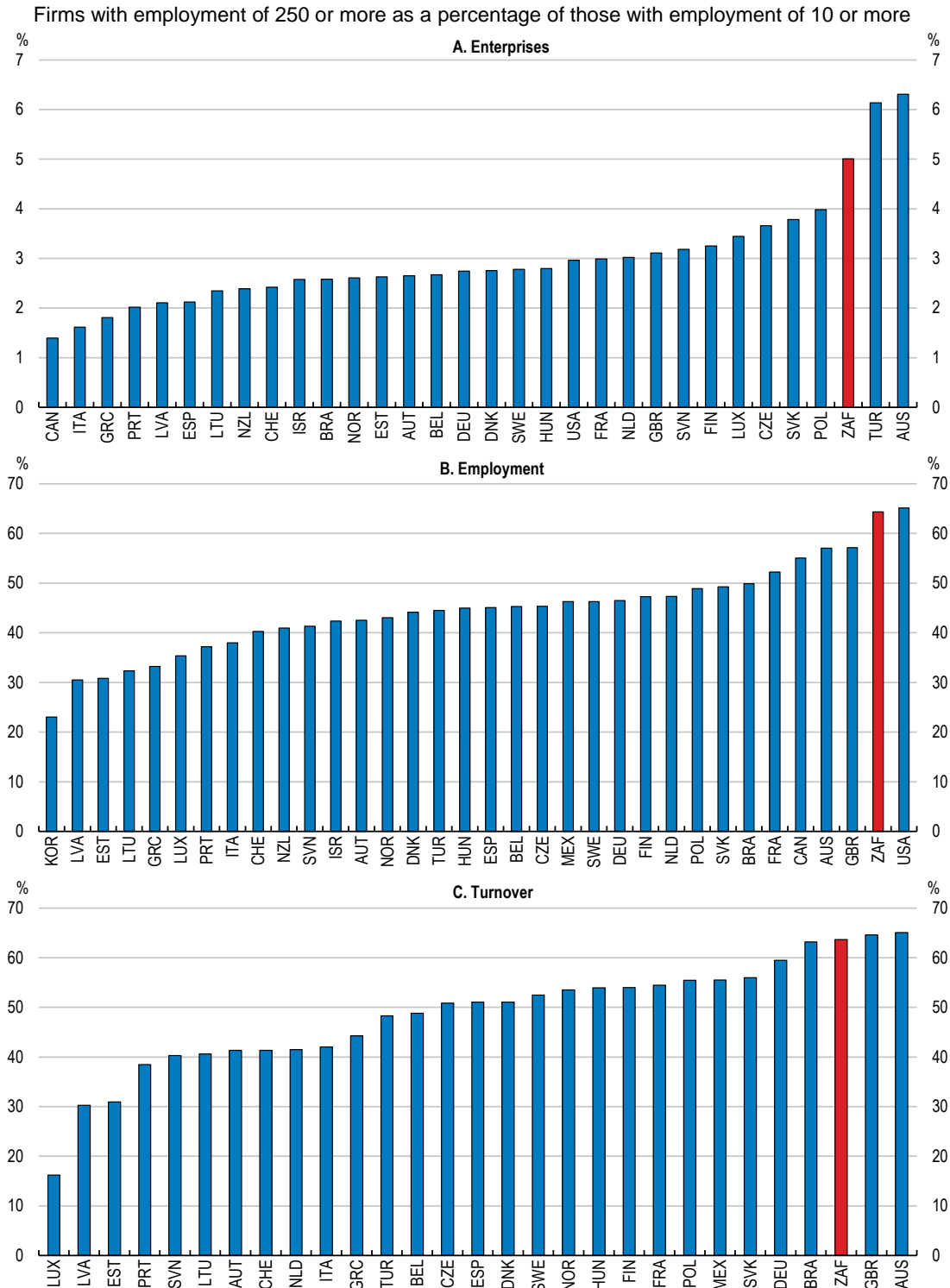


Note: Data are for employing firms registered for company income tax; see Pieterse, Gavin and Kreuser (2018) for details.

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 *et al.* (2018) and OECD (2008) (amongst others) have previously highlighted the concentration of the South African business sector.

**Figure 4. The share of large firms in the business sector is high**



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.

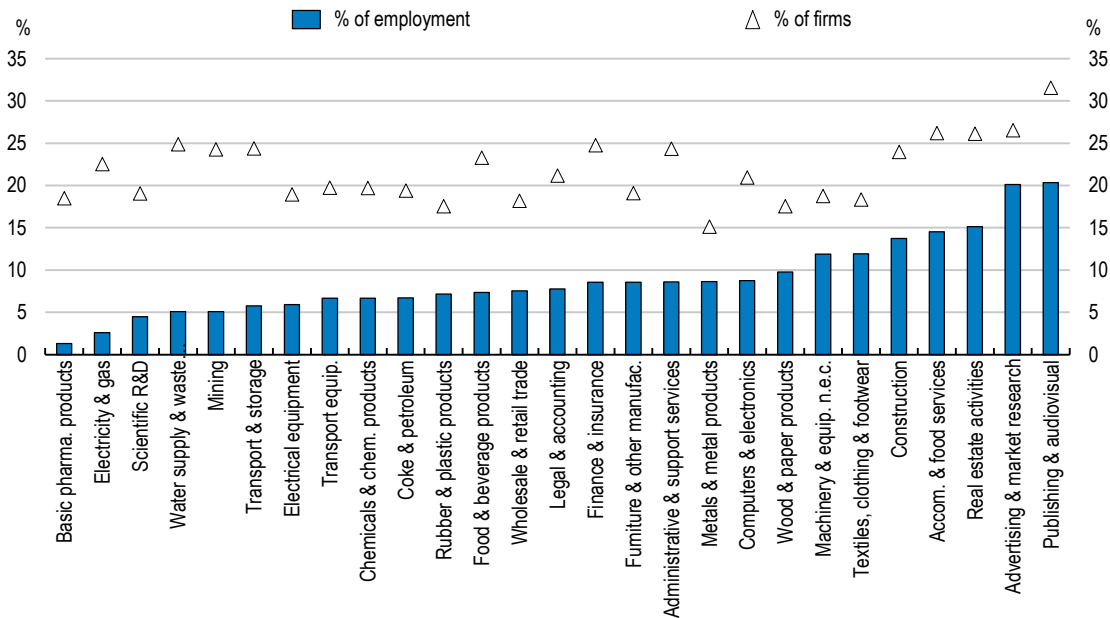
Source: OECD Structural Business Statistics database and Authors' calculations based on the SARS-NT Panel.

### Distribution of firms by age

25. A firm’s age can be calculated based on the 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 and so the analysis of firm age does not use these firms. 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 1% in pharmaceuticals manufacturing to 20% in advertising and market research and publishing, audiovisual and broadcasting (Figure 5). In sectors typically considered as having low barriers to entry, the employment share of start-ups is higher. Conversely the employment share of start-ups is lower in sectors that are more capital-intensive, like electricity and mining. This pattern may in turn reflect barriers to entry such as access to finance and market structure, as well as product market regulations. Indeed regulatory protection given to incumbents, particularly state-owned enterprises in network industries, may also contribute to concentration in those sectors (OECD, 2017a).

**Figure 5. Share of start-ups by sector**  
Firms aged 0-2 years by sector, 2013/14



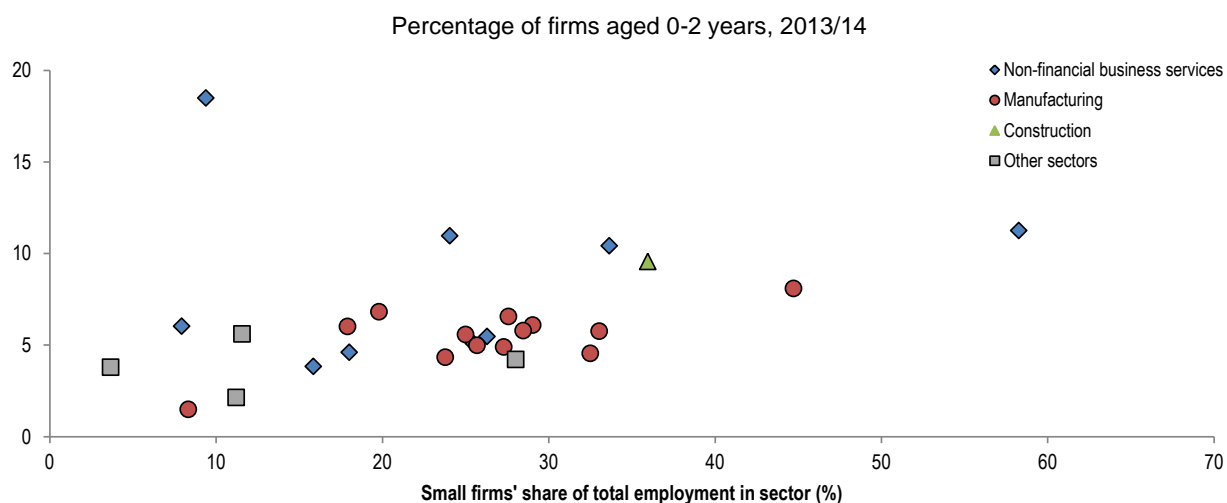
Note: Sectors are at the two-digit level.

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

27. The employment share of start-ups is partly dependent on the relative number of start-ups by sector. Consistent with the distribution of firms by firm size, there is a higher share of start-ups by number than by employment and it is more similar across sectors. This is partly determined by the fact that young firms tend to be small. Consequently, 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 6. **Young firms are more prominent in industries with more small firms**



Note: Sectors are at the two-digit level with colours indicating which macro-sector they belong to.

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

## Firm entry and exit

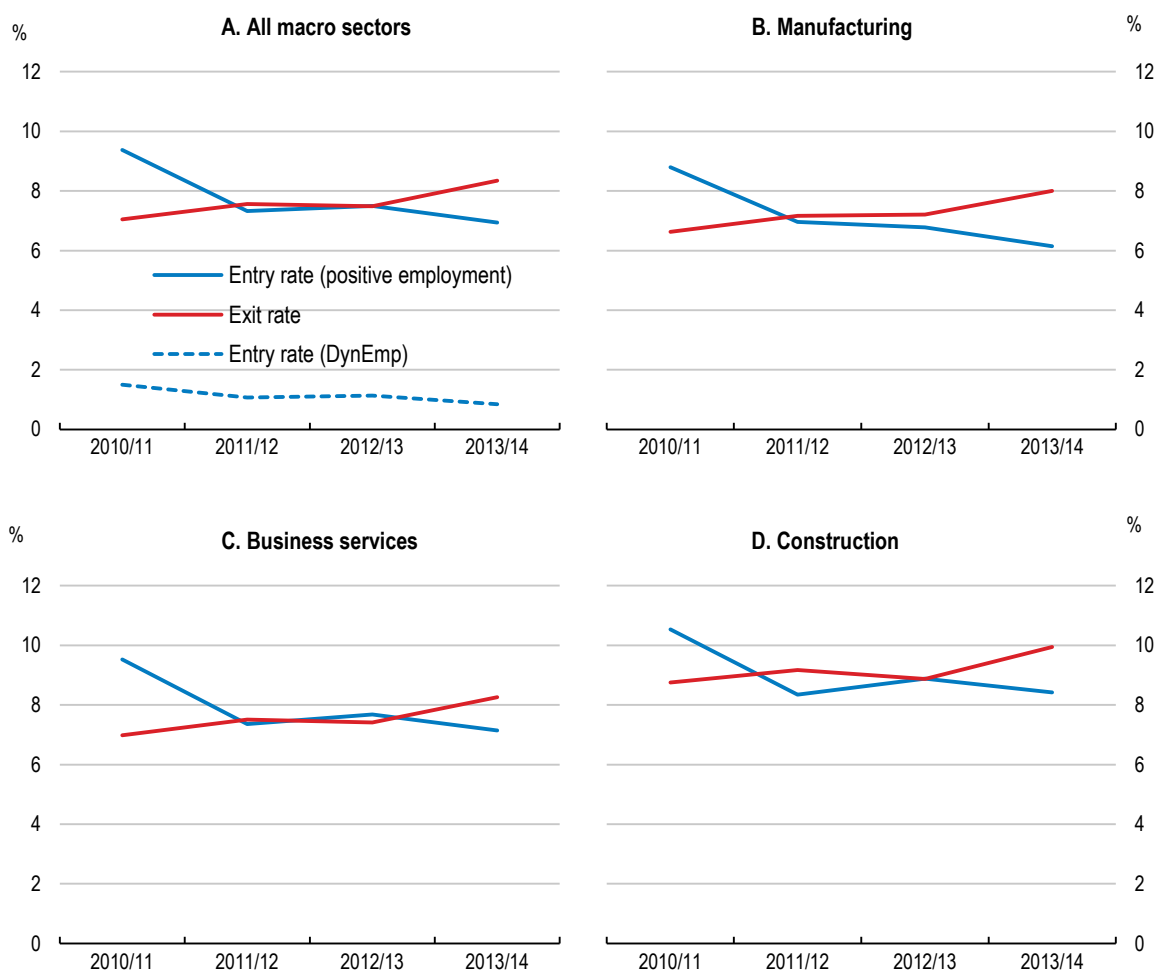
28. Firm entry and exit are key mechanisms for the introduction of new innovation and technologies. In a competitive economy, the process of creative destruction shifts labour and capital resources to the firms that are most productive and away from those that are least productive, boosting overall productivity 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. As mentioned above, the standard DynEmp v.2 approach is adjusted in several ways due to the characteristics of the SARS-NT database. For 2010/11 onwards, the first year of positive employment is used to indicate economic entry. Because entry due to mergers or non-economic reasons (such as improved compliance) cannot be reliably identified, entrants with first-year turnover in the top 5% of their sector are reclassified as existing firms with missing age. This screens out some implausibly large entrants. Without more information on the firms, more precise identification is difficult. However, the conclusions are little-changed when firms with over 100 employees in their first year are treated as existing firms rather than entrants (Figure C.3). 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 decline in 2013/14. These entry rates are lower than those in Brazil or Turkey (OECD, 2016) as well as other OECD countries (OECD, 2017c). Even though this includes some firms that are not genuinely new firms, the resulting over-estimation 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 were 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 exceeded 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.

Figure 7. **Entry and exit**



Note: The data shown cover the manufacturing, construction and non-financial business services sectors. See Appendix A for data definitions.

Source: Authors' calculations based on SARS-NT database

32. 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. A more structural explanation is proposed by Fedderke *et al.* (2018) 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. A third explanation is that it is linked to a broader cross-country phenomenon; Criscuolo *et al.* (2014) show that the activity of young firms declined over 2001-11.

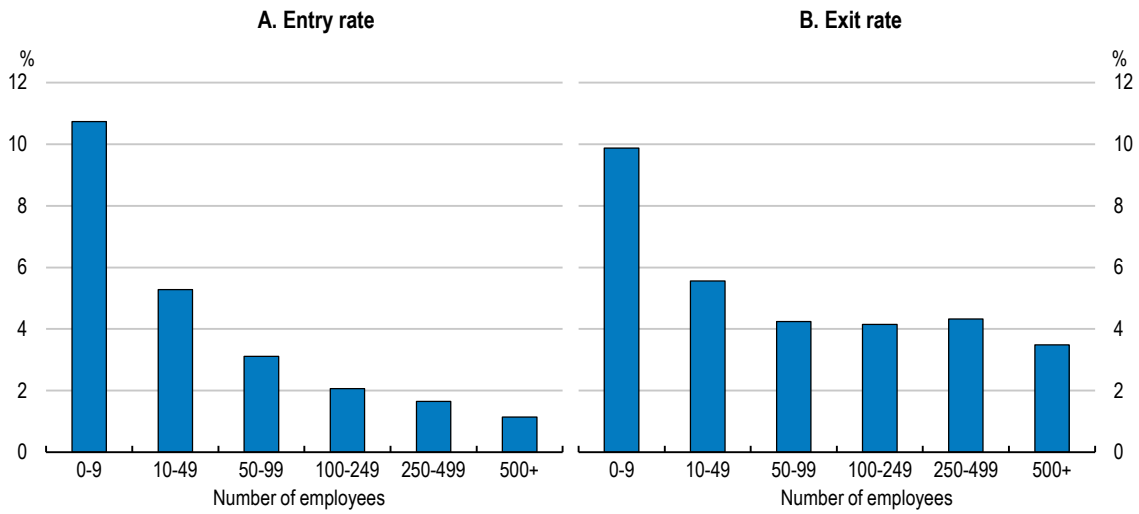
33. The data show that the definition of entry makes a large difference. The standard DynEmp definition that uses year of company registration where available (and first year of positive employment otherwise) results in a very low entry rate (at around 1% over the sample period). The DynEmp measure will understate 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 of firms with much longer gaps (Figure A.4). As mentioned above, 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.

34. These conclusions seem fairly robust to other checks. 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. As a cross-check, positive turnover and positive expenses were considered as indicators of activity-based entry or exit. These variables are available fewer firms but allow entry and exit to be independent of the process in which the company and employee databases are matched during the creation of the SARS-NT Panel. These measures also suggests that entry rates have been below exit rates (Figure A.3).

#### ***Characteristics of entrants and exiting firms***

35. As would be expected, entry and exit rates decline with firm size. Micro enterprises (less than 10 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. Exit rates are fairly stable over time, with only a slight increase for micro enterprises (Figure 9). Fedderke *et al.* (2018) 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, 2017a). On the other hand, the positive entry rate of large firms suggests that the cleaning above may be insufficiently strict to clean all implausibly large entrants (perhaps because turnover is missing for some firms); this reinforces the conclusion that the entry rate is likely to be somewhat overstated. But removing all large entrants (with over 100 employees) does not change the results materially.

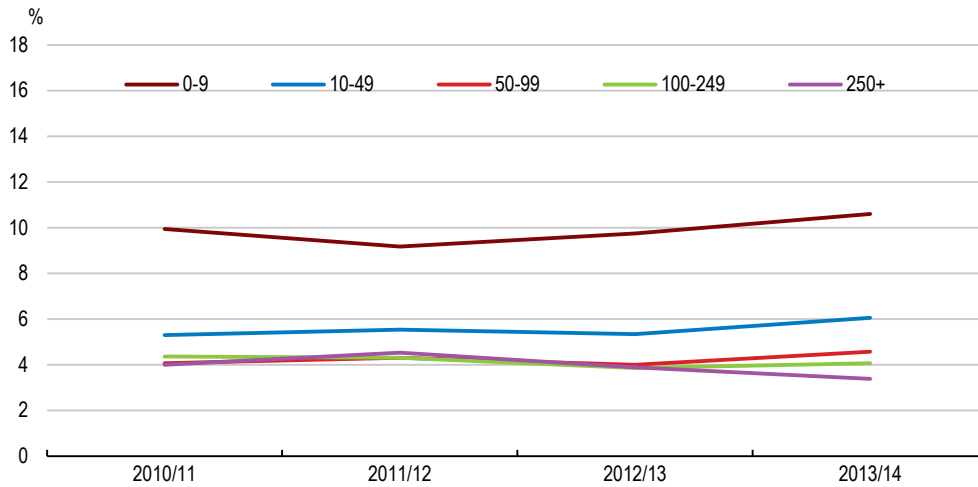
Figure 8. Average entry and exit rates by size class



Note: Size is defined by number of employees. Data shown are for the manufacturing, construction and non-financial business services sectors 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.

Figure 9. Exit rates by size over time



Note: Data shown are for the manufacturing, construction and non-financial business sectors. See Appendix A for details.

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

36. Looking at exit rates by firm age shows that young firms have a higher exit rate than old firms. Over 2012/13-2013/14 the exit rate averaged 8% for young firms (less than three years old) and 7% for those six or more years old. The finding for young firms fits with other work showing that young firms are typically more vulnerable and their employment growth more volatile (Calvino *et al.*, 2016). However, firms older firms 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.

37. It is possible that the exit rate is pushed up by micro enterprises shifting to the turnover tax regime described above. 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.

### **The dynamics of firm growth and job creation**

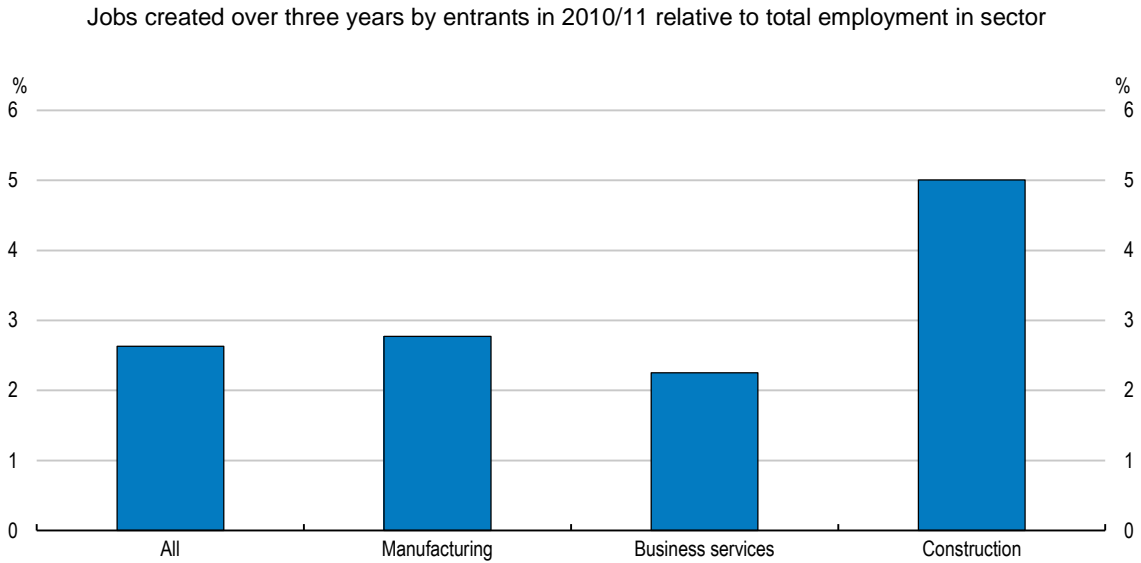
38. New firms contribute directly to employment dynamics through three channels: they create jobs by entering the market with positive employment; they destroy jobs if they close; and they create and destroy jobs by hiring and firing workers. Transition matrices in the DynEmp framework allow the employment 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.

39. A better understanding of the contribution that new firms make towards employment growth can be gained using the growth decomposition framework developed by Calvino *et al.* (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 (Figure 11). 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.



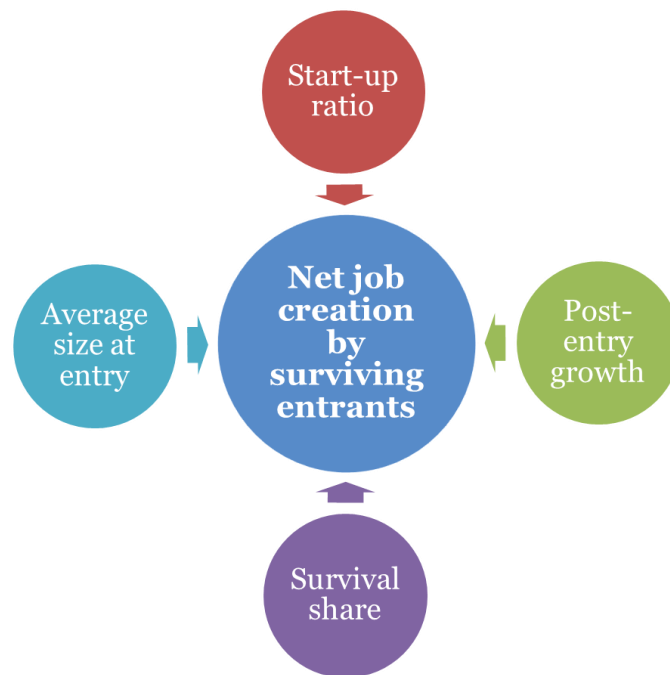
Figure 10. **Net job creation by surviving entrants**



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**

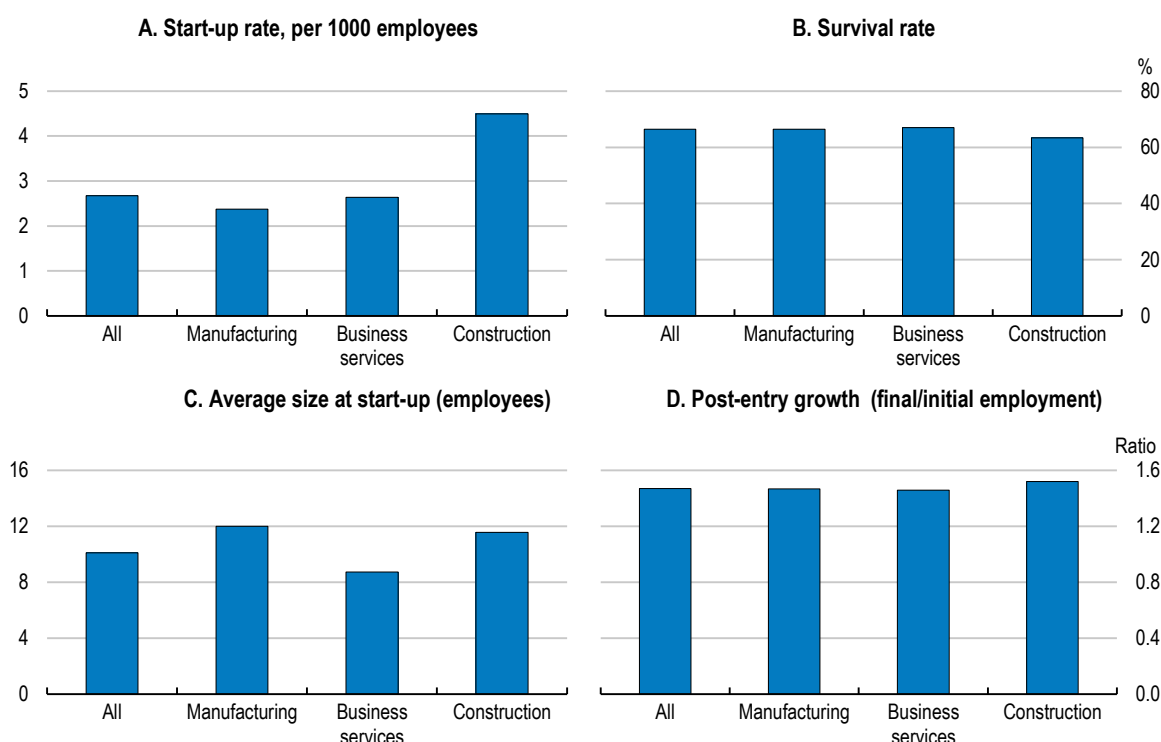


Source: Calvino, Criscuolo and Menon (2015).

40. 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). Sectoral specificities play an important role when examining the dynamics of

entry, survival, and growth of businesses. In manufacturing, the start-up ratio is low while the high average size at entry is high. This could be an indicator of barriers to entry, perhaps related to start-up capital. The construction sector has the highest number of start-ups per thousand employees. This points to relative ease of entry. It may also be due to increased demand as there has been substantial infrastructure investment; over 2010/11-2013/14 public sector infrastructure spending increased by 8% per year on average. Firms entering the business services sector have a lower average size at start-up than manufacturing and the highest survival rate among the three macro sectors.

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

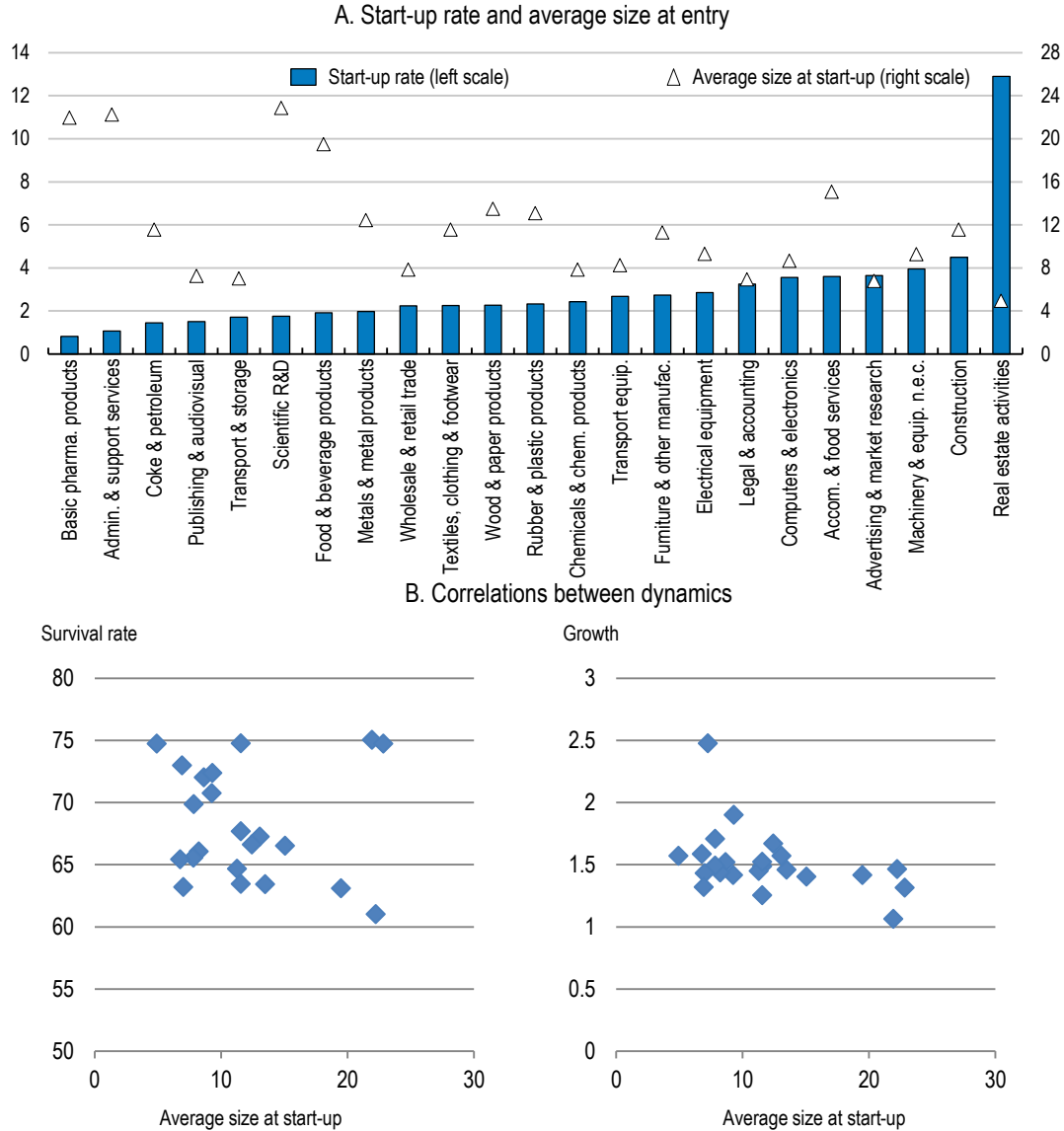


Note: The graph illustrates the four components of the growth decomposition of surviving entrants between 2010/11 and 2013/14. Business services excludes financial and insurance services. See Calvino, Criscuolo and Menon (2015) for full details of the decomposition.

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

41. A more detailed decomposition of the firm-level dynamics shows that there is considerable variation between sectors in all four components of net employment growth (Figure 12). In 2010/11, average employment at entry ranged from 5 employees in real estate and rental activities, which also had the highest start-up rate, to 22 employees in manufacturing of basic pharmaceutical products, which had the lowest start-up rate. There is a strong negative relationship between start-up rates and average size at entry, even excluding outliers. The detailed data also show that subsequent employment growth is higher in sectors with firms that are smaller at entry. The average survival rate by sector is negatively correlated with the average size at entry.

Figure 13. Sectoral dynamics



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.

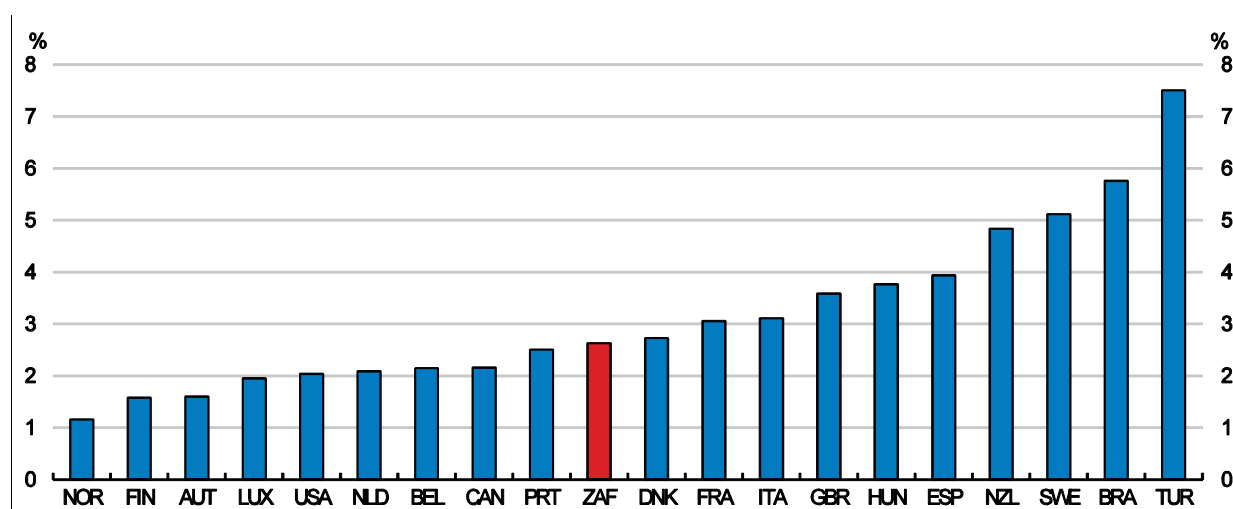
42. The empirical literature also points to considerable and persistent within-industry heterogeneity (Bartelsman and Doms, 2000; Syverson, 2011). Sector-specific dynamics can depend on factors such as capital intensity, the presence of significant economies of scale, sunk costs, or concentration (Siegfried and Evans, 1994; Audretsch, 1991), the stage of development of the industry (Agarwal and Gort, 1996; Klepper, 1996), or the nature and evolution of its technological base (Audretsch, 1995; see also Santarelli and Vivarelli, 2007). In their study of mark-ups in South Africa's manufacturing sector, Fedderke *et al.* (2018) show that barriers to entry affect the relationship between concentration, firm size and entry and exit dynamics differently across sub-sectors of manufacturing. While they find evidence that mark-ups are a barrier to entry in sectors such as food and food products, textiles, clothing and footwear, and chemicals manufacturing but not in beverages, coal or transport equipment, amongst others.

43. As discussed above, the entry rate is overstated because corporate activity is not identified. Figure 8 shows that there are some large entrants. Reclassifying firms with 100 or more employees from entrants to existing firms has the expected effect of lowering the average size at start-up, to 9 employees (Figure C.4). 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.

### *Firm dynamics compared to other countries*

44. Comparing the net job creation by entrants and the sources of this growth helps to 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 14). 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 *et al.*, 2015; Figure 15). 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 1 000 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 14. 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.

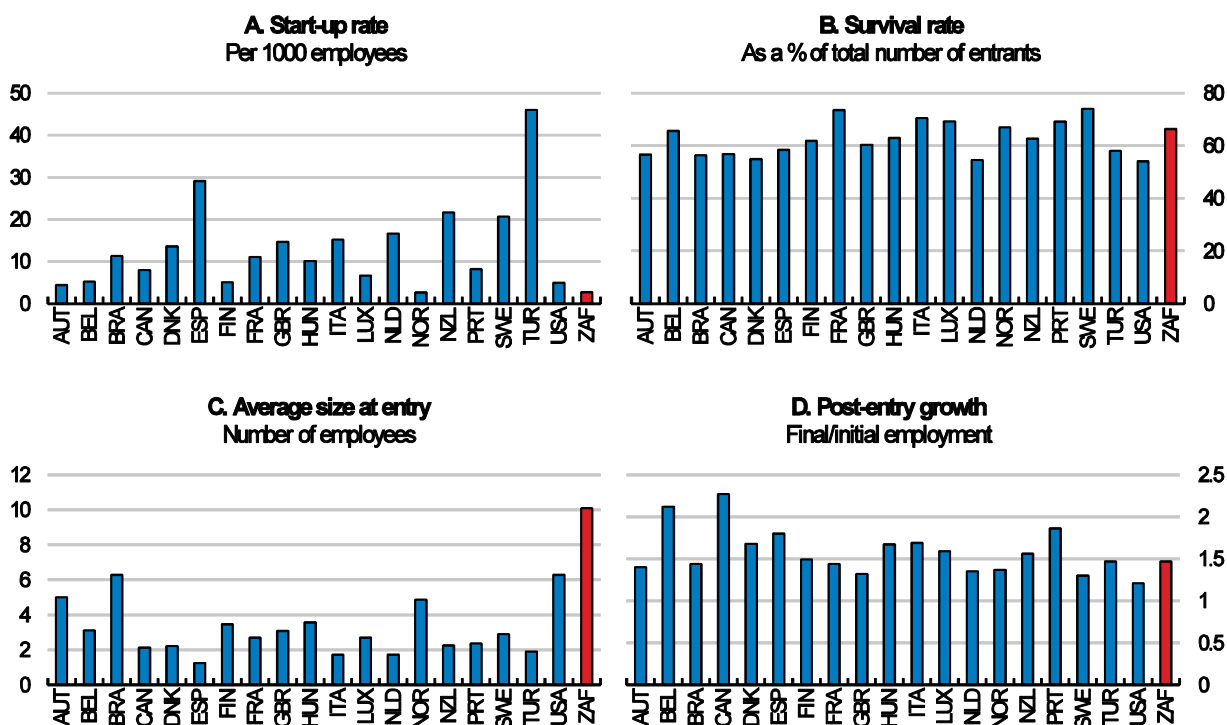
Source: SARS-NT Panel; Calvino, Criscuolo and Menon (2015); OECD (2016); DynEmp v.2 database (accessed in June 2017); Authors' calculations.

45. Significant cross-country differences in business dynamics have been highlighted by previous studies (see for instance Bartelsman *et al.*, 2005; 2009; Criscuolo *et al.*, 2014) and these appear related to a wide variety of factors, with an important role played by institutions. In this context, Calvino, Criscuolo and Menon (2016) highlight the role of bankruptcy regulations, contract enforcement and access to finance

as important factors for a dynamic business environment, Bravo-Biosca *et al.* (2016) focus on the role of labour market institutions and of a well working financial system, while Andrews *et al.* (2014) highlight how policy mediates the extent to which tangible resources flow to innovative firms. Relatedly, Klapper *et al.* (2000) examine the role of entry regulations as barriers to firm creation, and their effects on the size of entrants and on incumbents' growth, while additional contributions corroborate the fact that industries more financially dependent grow relatively faster where financial markets are more developed (see among others Beck *et al.*, 2008; Aghion *et al.*, 2007). In a South African context, Fedderke *et al.* (2018) find that mark-ups are higher in South Africa's manufacturing sector than in Finland's.

46. 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 such as product market regulation, bankruptcy regulations, licensing, access to finance and skills (discussed in OECD, 2018); (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 related to this, the informal sector and the potential for start-ups to begin as sole proprietorships, which would be outside of the SARS-NT 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. But the high survival rate could reflect structural factors such as high barriers to entry and weak market selection. The survival rate 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; the latter countries also has fairly large firms dominating the economy (Figure 4).

Figure 15. Cross-country comparison of growth decomposition



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.

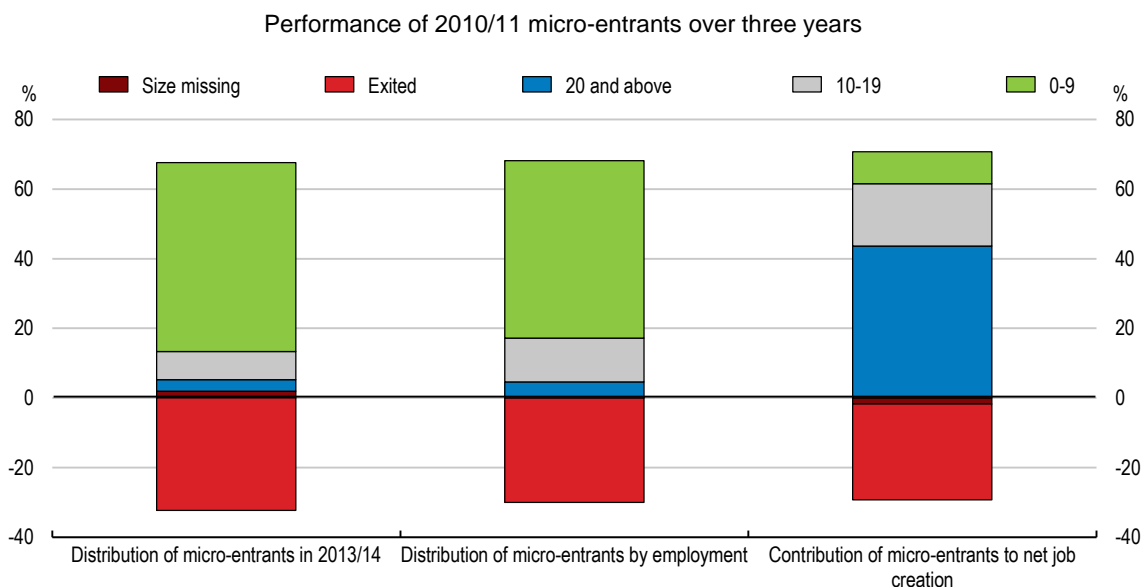
Source: SARS-NT Panel; Calvino, Criscuolo and Menon (2015); OECD (2016); DynEmp v.2 database (accessed in June 2017); Authors' calculations.

### Growth of micro-entrants

47. 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 *et al.*, 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.

48. In South Africa, there are more surviving start-ups than exiting start-ups, which is in line with other Dynemp countries (Figure 16, see also Figure 11 in Calvino *et al.*, 2015, and Figures 10 and 11 in Criscuolo *et al.*, 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 *et al.*, 2016). There is some evidence of stronger dynamics in manufacturing and construction than business services, perhaps linked to market selection (Figure 17). 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.

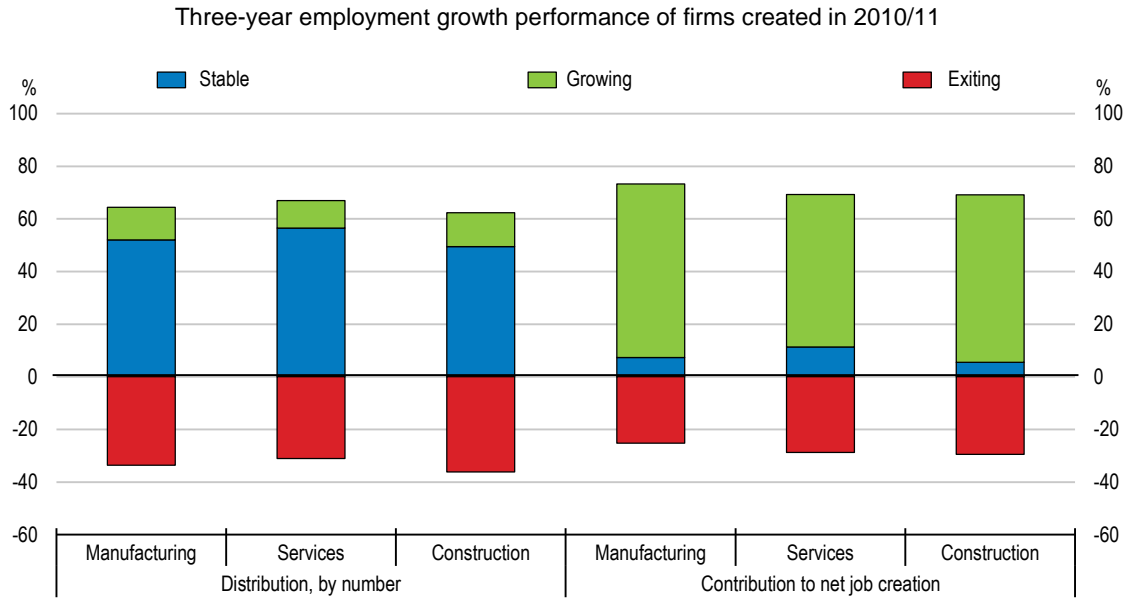
Figure 16. Few micro-entrants grow but they are important job-creators



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

**Figure 17. Performance of micro-entrants across key sectors**

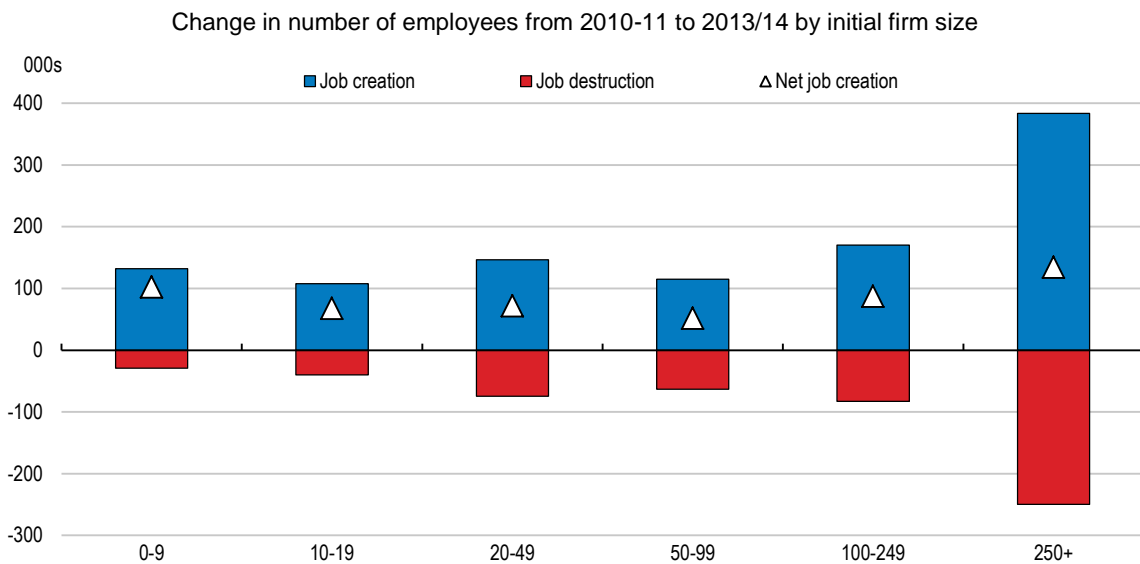


Note: A small share of micro-entrants that are missing size information are excluded.

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

49. 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 18). 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.

**Figure 18. Job creation and destruction by existing surviving firms**



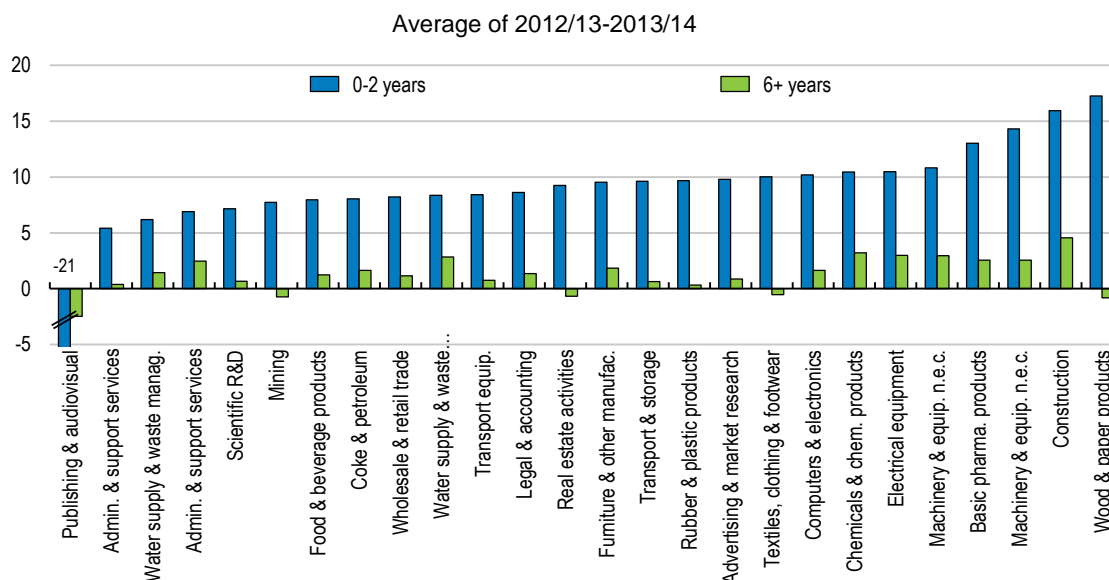
Note: Industries shown are in the manufacturing, construction and non-financial business services sectors. Jobs created by entrants and exiting firms are not included.

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

## The role of age in firm growth

50. As for other countries, and in other studies, young firms grow faster than old firms. In 2012/13-2013/14 the difference between average employment growth of young (0-2 years old) and older (6 or more years old) firms was more than 4 percentage points in all industries considered here except publishing, 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 and petroleum manufacturing and accommodation and food services made the largest net contributions to sectoral job creation in these years. The data also reveal that old firms were net job destroyers in 2012/13 and 2013/14. Taken together, these findings point 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 *et al.*, 2014; Haltiwanger *et al.*, 2013).

Figure 19. Average employment growth by sector and age group



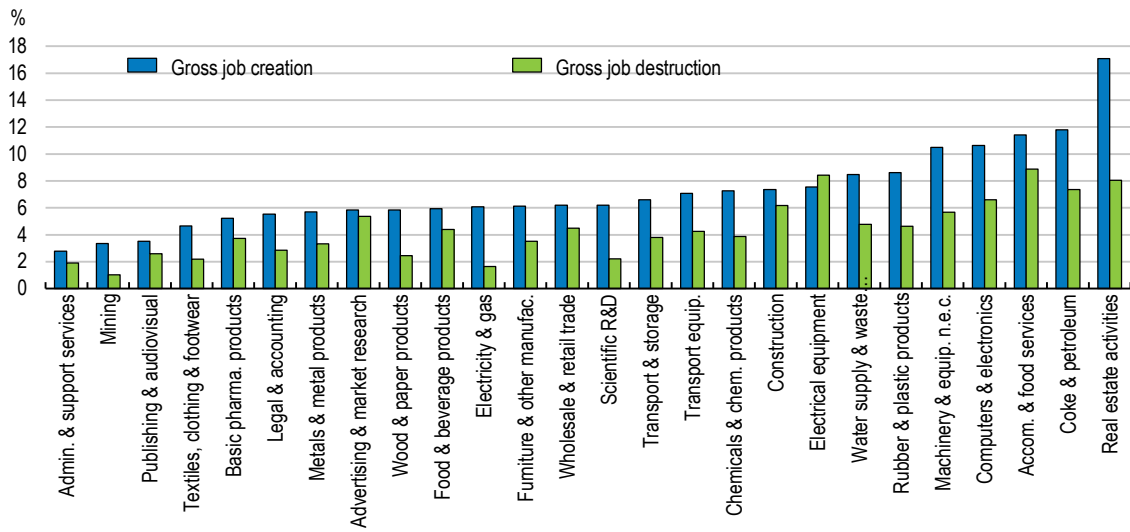
Note: Firms aged 3-5 years are not well identified and are omitted here (see Appendix A for details). Sectors shown are in the manufacturing, construction and non-financial business services sectors.

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



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

Small firms' share of sector's gross job creation and destruction, 2012/13-2013/14 average



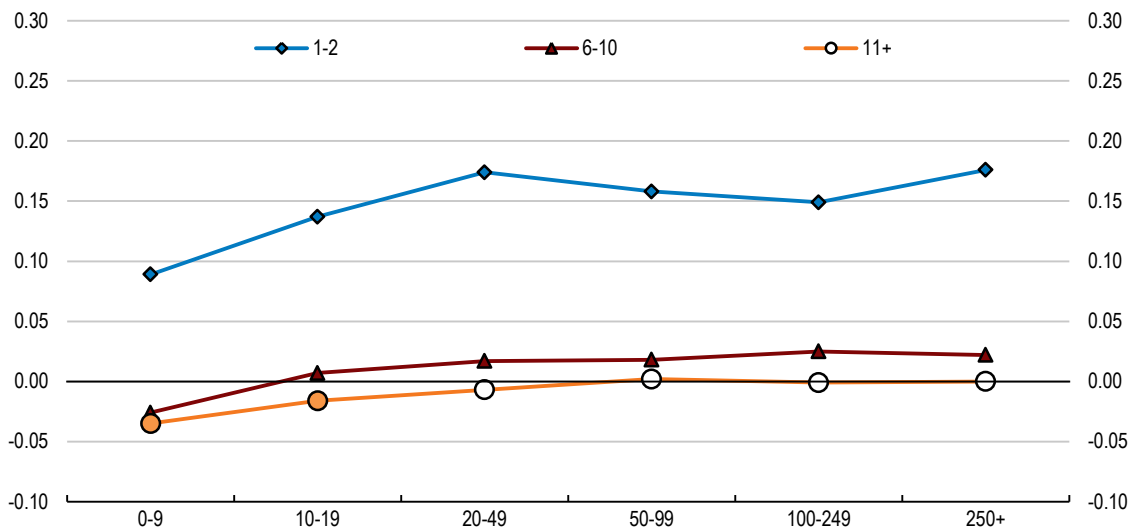
Note: Small young firms are firms with less than 50 employees and are 0-2 years old. Industries shown are in the manufacturing, construction and non-financial business services sectors. Industries are ordered by their net contribution to job creation (i.e. gross job creation less gross job destruction).

Source: Authors' calculations based on SARS-NT panel

51. 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 sectoral level. Firms with age missing and aged 3-5 years were included in the regression.

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

## Conclusions and next steps

52. 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 than manufacturing in South Africa. A key finding has been that the entry rate appears likely to have been below the exit rate in recent years, which means that South African firms are growing older overall. This is particularly true in manufacturing. These conclusions will be able to be made more firm as the time dimension of the SARS-NT Panel grows.

The DynEmp framework facilitates cross-country comparisons of dynamics of net employment creation by entrants and reinforces the impression that dynamics are weak in South Africa, notwithstanding the caveats relating to the data, as discussed above. But there are important differences within South African 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 such as regulation, capital intensity and, relatedly, access to finance.

53. Young firms have disproportionately contributed to employment growth in recent years. Even with slower GDP and employment growth, young firms remained net job creators overall. 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. However, a notable feature is the importance of large incumbent firms for job creation and destruction.

54. This paper contributes to a better understanding of firm dynamics in South Africa by establishing measures of dynamics across sectors and comparisons with other countries. These measures 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 or regulatory burdens. Exposure to the latter could be analysed using the OECD's measures of sectoral exposure to upstream network and services sector regulation, for instance. Sectoral concentration or market power of incumbents may also affect start-ups' performance (as in Fedderke *et al.*, 2016).

## APPENDIX A: DEFINING ENTRY AND EXIT

55. 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:

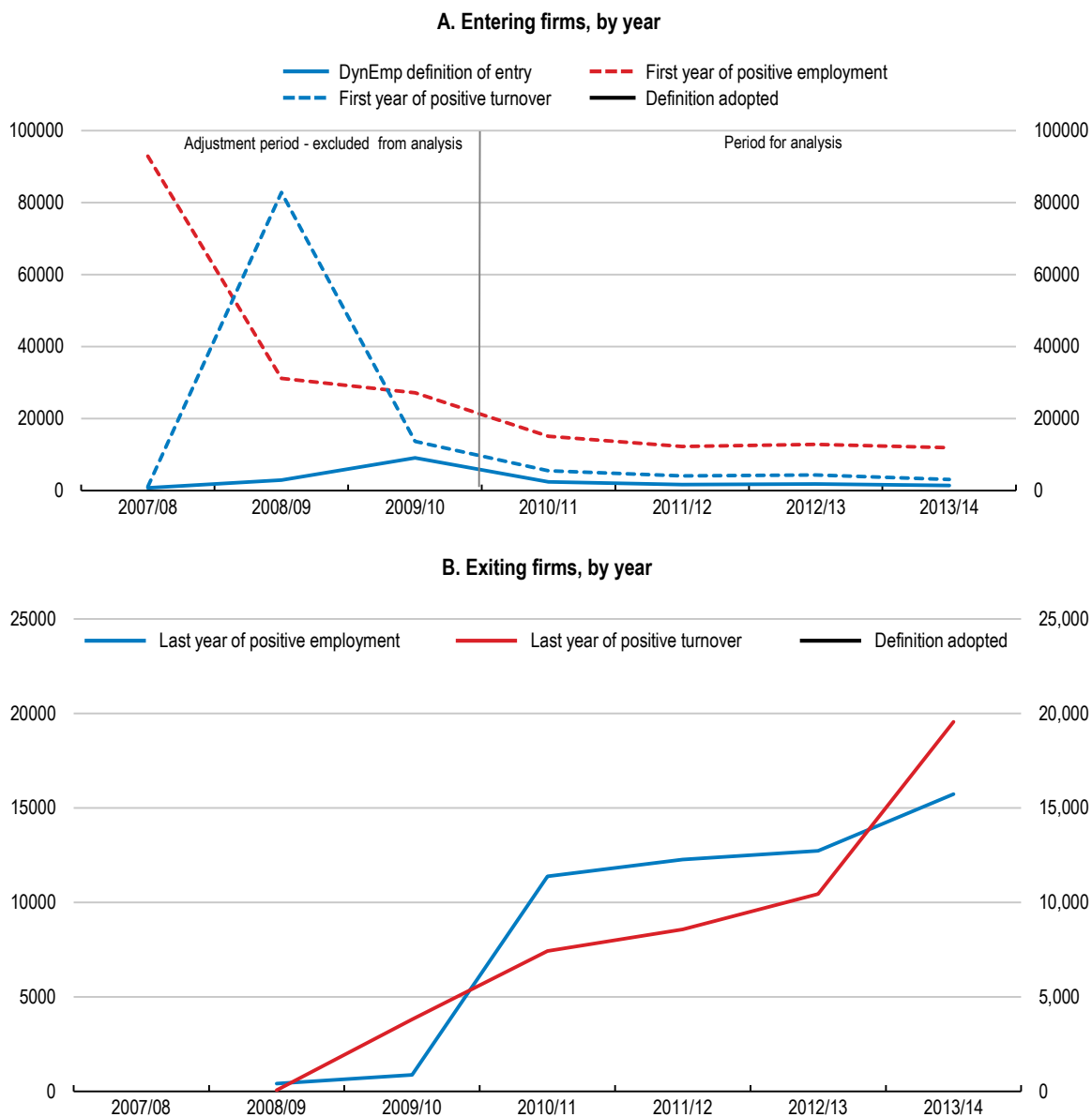
- The standard DynEmp v.2 definition of entry uses a firm's registration date if it exists, supplemented by first year of appearance with positive employment if the former is missing. But firms may register a company name and then wait years before deciding to build a company. The firm's true age would 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 that 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, 2018), which implies greater caution in using information from 2007/08.

56. 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 2009/10 (when the major compliance-related adjustment has occurred) that registered before 2007/08 are treated as having been active since their registration year.

57. 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 programm 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.) The justification of treating data 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).

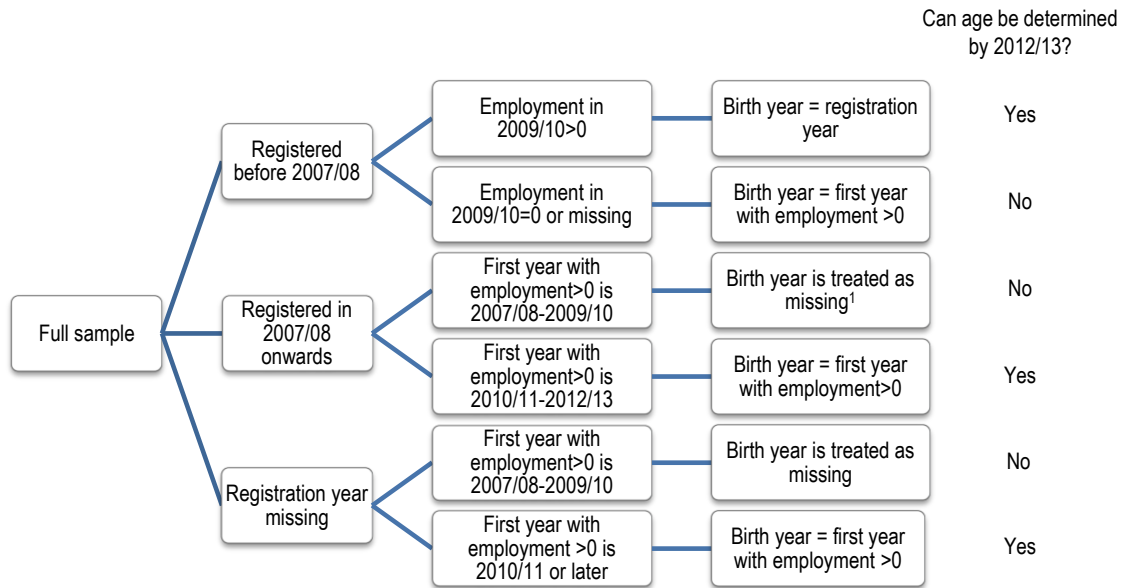
Figure A.1. Entry and exit based on different definitions



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

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

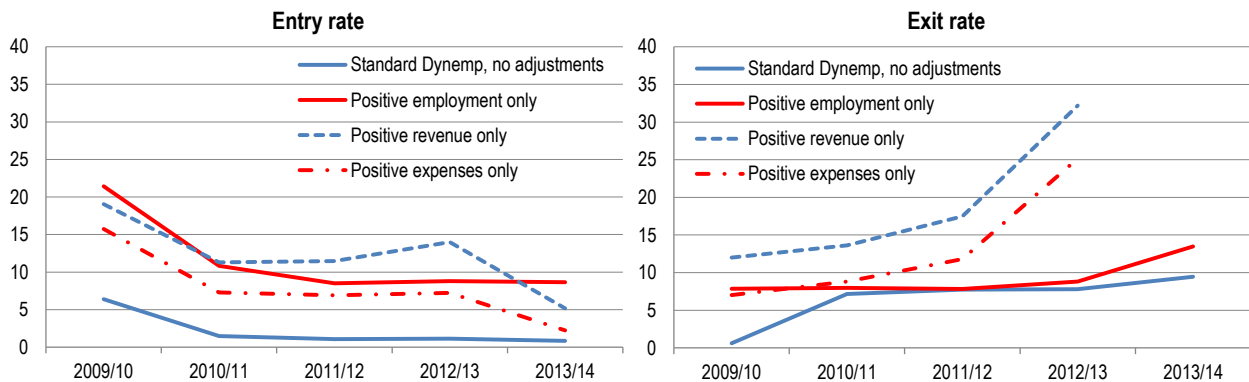
Figure A.2. Dating firm birth and classifying age



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.

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

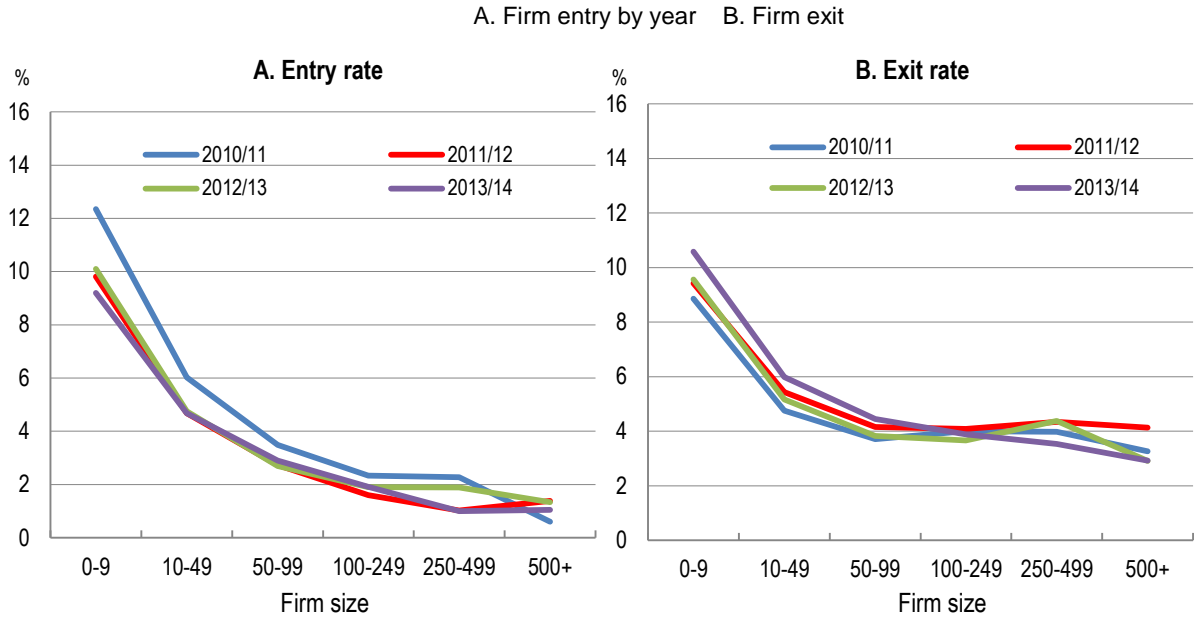


Note: Sectors covered are manufacturing, construction and non-financial business services for DynEmp and all industries (including where not specified) for employment, revenue and expenses. The exit rate in 2013/14 is implausibly high and omitted.

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

58. As mentioned in the main text, 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 key 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

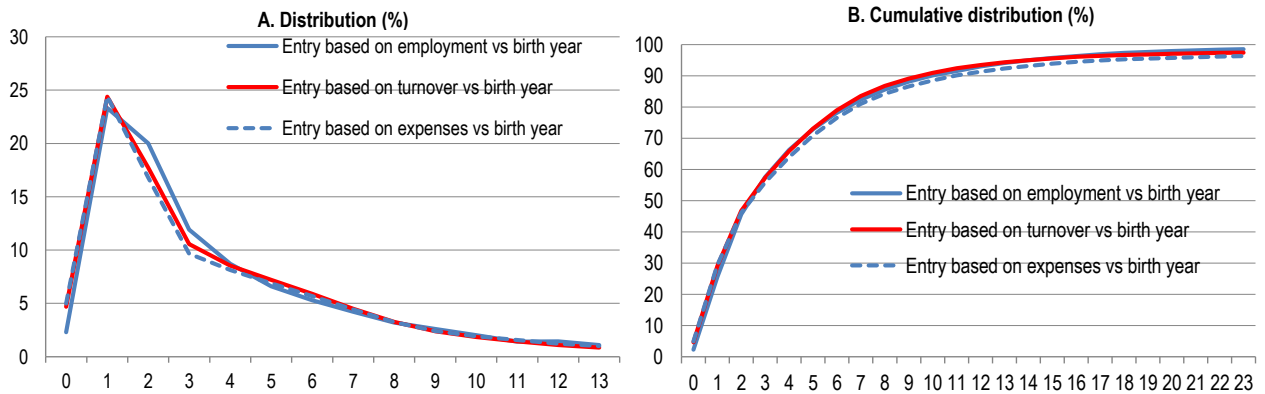


Note: Sectors covered are manufacturing, construction and non-financial business services.

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**

	2007/08	2008/09	2009/10	2010/11	2011/12	2012/13	2013/14
0-9 employees	445	2,181	6,777	1,973	1,436	1,544	1,145
10-49 employees	244	631	2,068	372	241	284	236
50-99 employees	49	73	198	28	12	18	12
100-249 employees	18	37	63	6	7	9	5
250-499 employees	5	4	12	2		2	
500+ employees	7	1	5			1	1
<b>Total</b>	<b>768</b>	<b>2,927</b>	<b>9,123</b>	<b>2,381</b>	<b>1,696</b>	<b>1,858</b>	<b>1,399</b>
Memo: micro-enterprises % of total	58	75	74	83	85	83	82

**B. Number of firms based on positive employment without screening out largest firms**

	2007/08	2008/09	2009/10	2010/11	2011/12	2012/13	2013/14
0-9 employees		4,256	10,560	11,155	8,987	9,418	8,562
10-49 employees		1,174	3,191	3,275	2,727	2,841	2,751
50-99 employees		128	344	401	334	337	322
100-249 employees		73	151	193	148	177	169
250-499 employees		12	29	54	37	50	34
500+ employees		7	13	22	26	29	39
<b>Total</b>		<b>5,650</b>	<b>14,288</b>	<b>15,100</b>	<b>12,259</b>	<b>12,852</b>	<b>11,877</b>
Memo: micro-enterprises % of total		75%	74%	74%	73%	73%	72%

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

	2007/08	2008/09	2009/10	2010/11	2011/12	2012/13	2013/14
0-9 employees				11,377	8,808	9,251	8,438
10-49 employees				3,065	2,426	2,550	2,545
50-99 employees				287	233	241	267
100-249 employees				112	81	102	107
250-499 employees				33	16	31	17
500+ employees				7	17	17	14
<b>Total</b>				<b>14,881</b>	<b>11,581</b>	<b>12,192</b>	<b>11,388</b>
Memo: micro-enterprises % of total				76%	76%	76%	74%

*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

59. 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**

	SARS-NT Panel	Labour force survey Survey of Employers and self-employed
<b>Definitions</b>		
Firm	Registered for tax Positive employment in at least one year during the sample period	Employer or own-account worker
Employment by firm size	Number of employees (jobs) in firm	Number of employees at "place of work"
<b>Coverage by type of firm</b>		
Informal firms	Not included	Included (but firms only include the self-employed)
Formal firms:		
Own-account workers / sole proprietorships	Not included	Included (firms and employment)
Partnerships	Not included	Included (firms and employment)
Micro-enterprises	Included if paying company income tax and PAYE for at least one employee but excluded if registered for turnover tax	Firms included if one person owns it. employment less well identified
Small to large firms (10 or more employees)	Well identified in this dataset	Larger firms are less well identified because employment is inferred from respondants' estimates.

60. 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**

	By number of firms ('000)				By employment ('000)			
	1-9	10-19	20-49	50+	1-9	10-19	20-49	50+
<b>Manufacturing</b>								
SARS-NT dataset	24.6	10.2	9.1	15.5	104.4	140.2	280.2	1659.7
	(41.5)	(17.1)	(15.3)	(26.1)	(4.8)	(6.4)	(12.8)	(76)
LFS – all	183.4	13	3.6	4.1	345.5	164.5	258.9	876
	(89.9)	(6.4)	(1.8)	(2)	(21)	(10)	(15.7)	(53.3)
LFS – formal	46.8	11	3.5	3.6	161.9	162.6	258.8	875.5
	(72.1)	(17)	(5.4)	(5.5)	(11.1)	(11.1)	(17.7)	(60)
<b>Construction</b>								
SARS-NT dataset	8.2	2.9	2.5	5	33.8	39.5	74.9	316.6
	(44.4)	(15.6)	(13.2)	(26.7)	(7.3)	(8.5)	(16.1)	(68.1)
LFS – all	278.3	17.4	11.2	4.5	597.3	219.4	215.9	291.8
	(89.4)	(5.6)	(3.6)	(1.4)	(45.1)	(16.6)	(16.3)	(22)
LFS – formal	58.7	16.2	10	4.2	218.9	218.2	214.4	291.5
	(65.9)	(18.2)	(11.2)	(4.7)	(23.2)	(23.1)	(22.7)	(30.9)
<b>Business services</b>								
SARS-NT dataset	48.8	15	12.4	23.5	191.4	205.1	379.1	2820.6
	(48.9)	(15)	(12.4)	(23.6)	(5.3)	(5.7)	(10.5)	(78.4)
LFS – all	1236.7	41.3	25.3	13.8	2348.1	792.1	906.9	1470.8
	(93.9)	(3.1)	(1.9)	(1.1)	(42.6)	(14.4)	(16.4)	(26.7)
LFS – formal	314.6	38.7	24	12.8	938.8	789.5	905.4	1469.8
	(80.6)	(9.9)	(6.2)	(3.3)	(22.9)	(19.2)	(22.1)	(35.8)

*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.

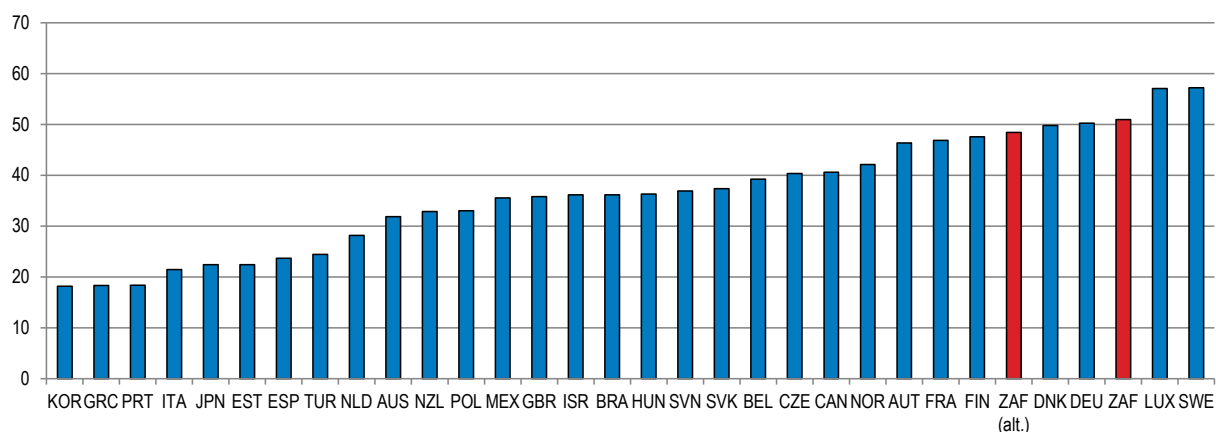
*Source:* Authors' calculations based on the SARS-NT Panel and Labour Market Dynamics database

## APPENDIX C. ROBUSTNESS CHECKS

### The importance of large firms

61. 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. **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.

Source: OECD Structural Business Statistics database; OECD Labour Force database; SARS-NT Panel.

### Definition of employment

62. As discussed in the text and Pieterse, Gavin and Kreuser (2018), 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.

63. 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

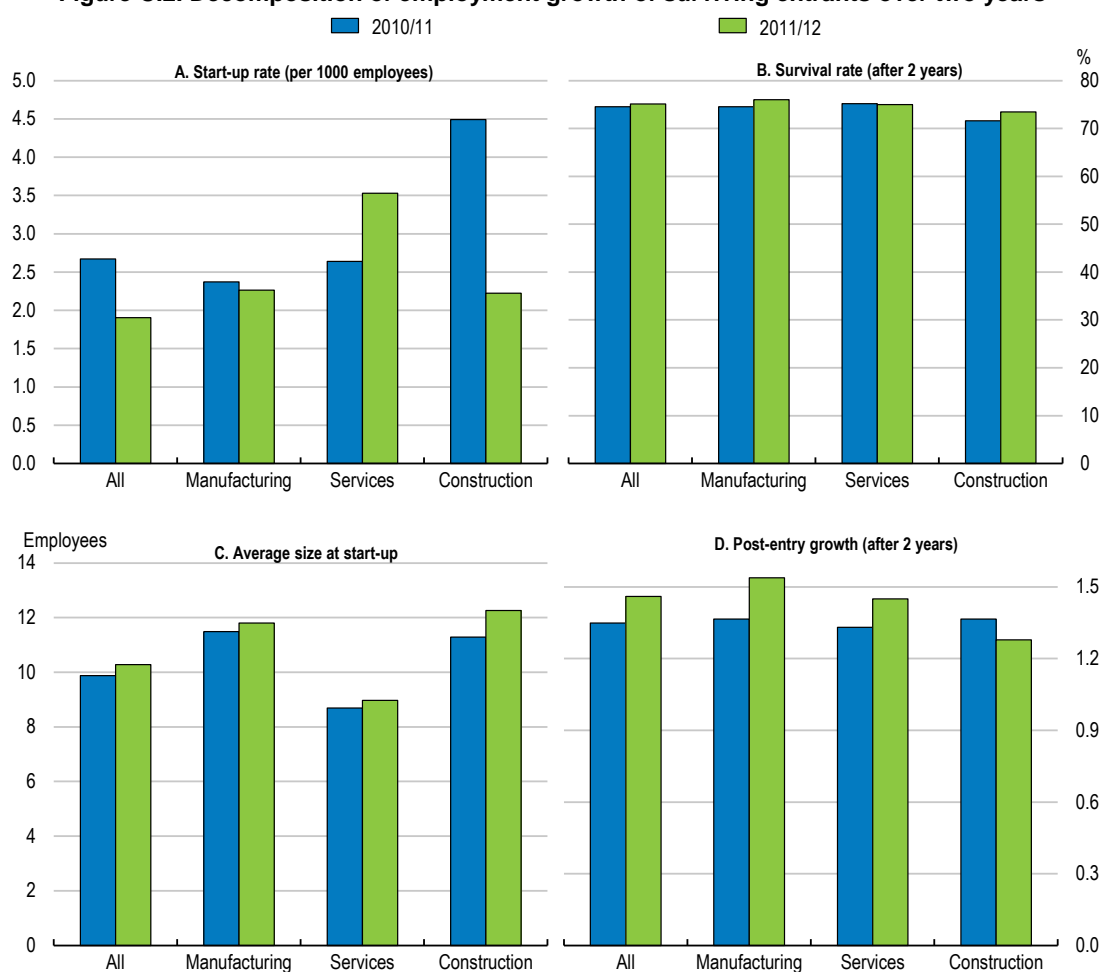
proportion of days worked raises the start-up rate and lowers the average size at entry of survivors somewhat, but the start-up rate would still be comparatively low and the average size comparatively high.

### Sensitivity of the year chosen for the growth decomposition

64. 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).

65. 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**

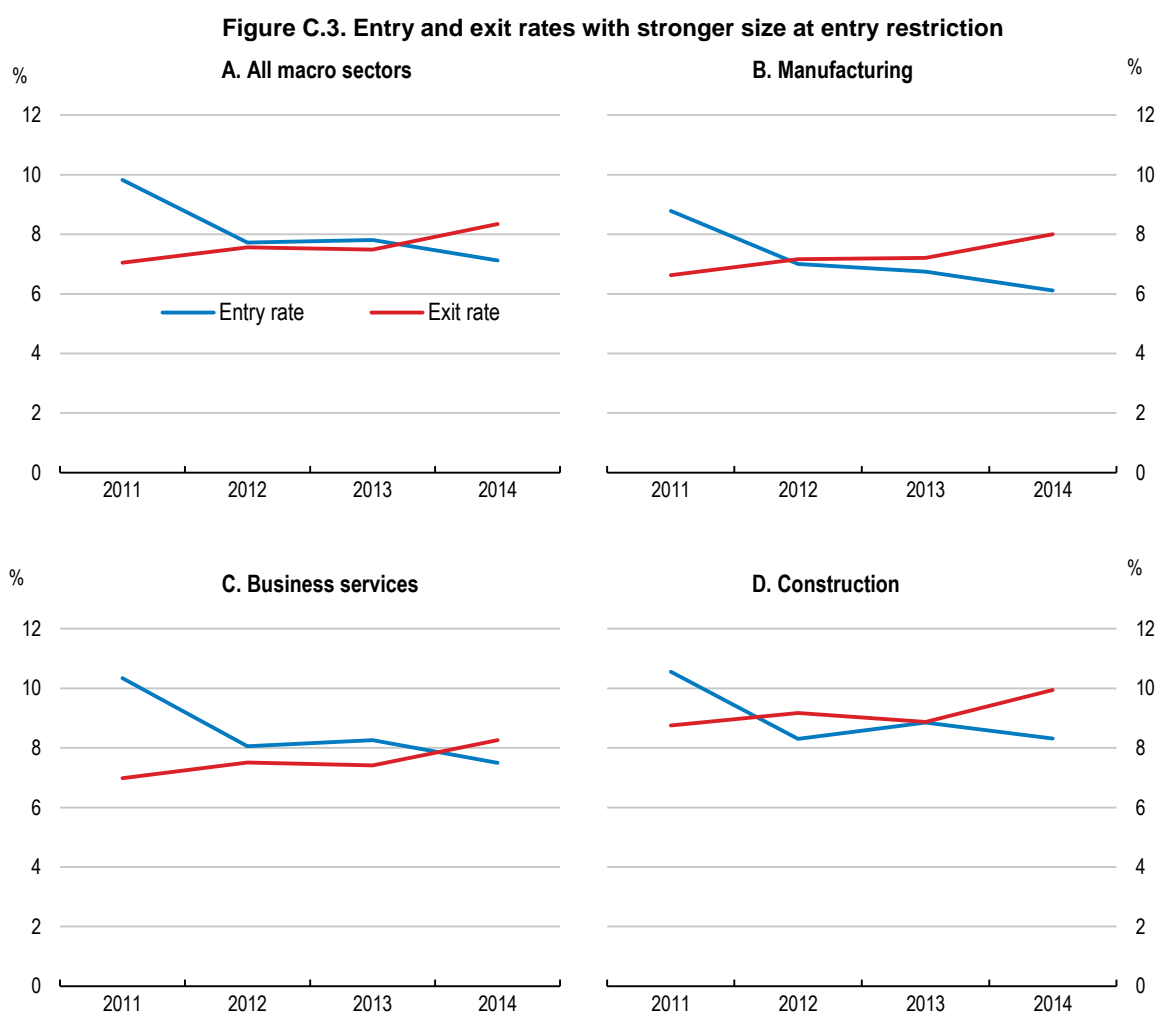


Note: Sectors covered are manufacturing, construction and non-financial business services. Post entry growth is the ratio of final to initial employment.

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

## Restricting size at entry

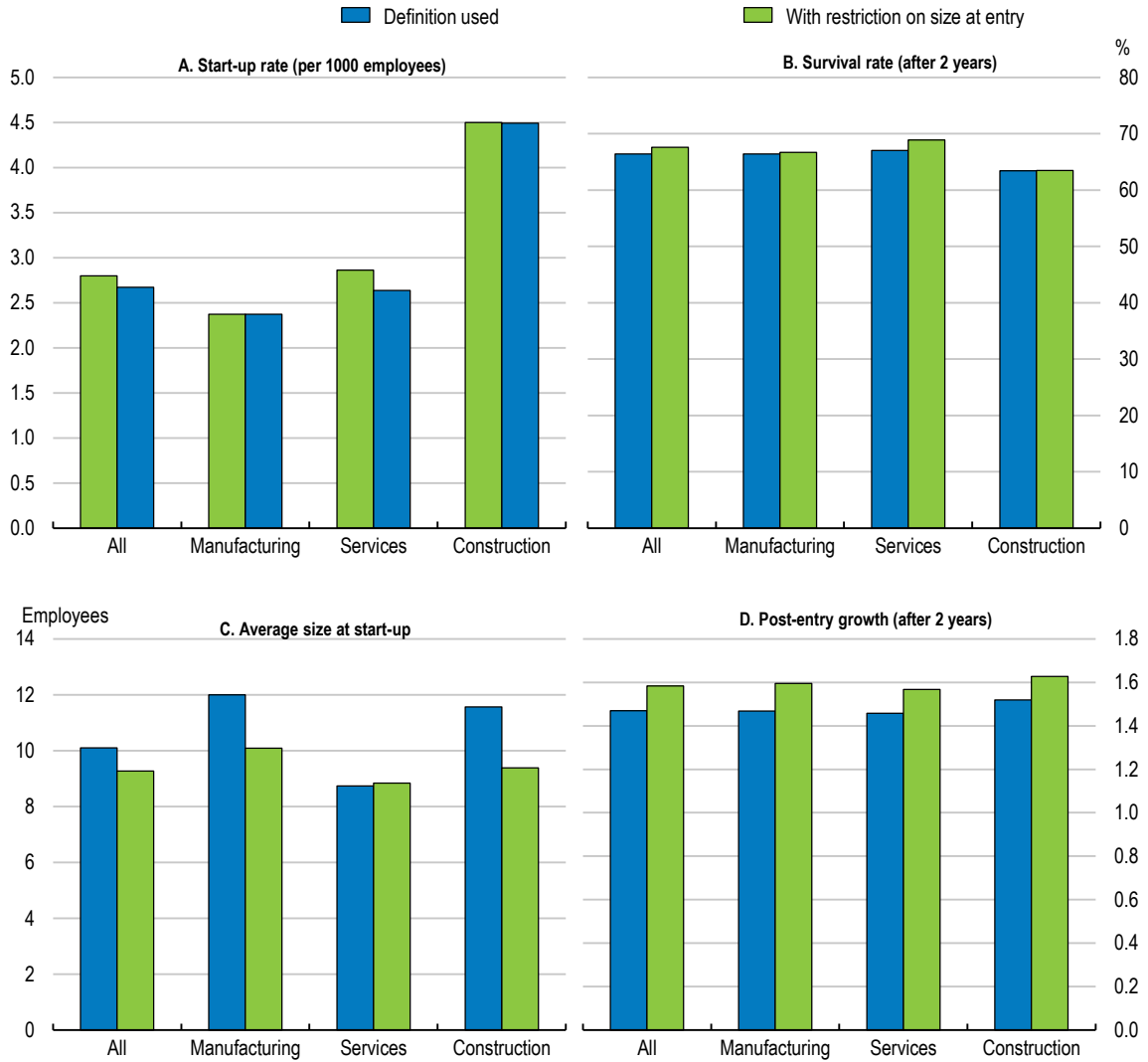
66. To assess the impact of “false entrants” on the dynamics of firm growth, the results were compared under the more restrictive assumption that any entrant of more than 100 employees is a non-economic entrant and should be excluded from the analysis. As shown in Figure C.3, the conclusions for entry and exit rates are not greatly affected although entry is above exit in the non-financial business services sector. The main effect on decomposition of net employment creation by surviving entrants is that the average size at entry falls in manufacturing and construction (Figure C4).



Note: Firms with more than 100 employees in their first year of positive employment are reclassified as existing firms of unknown age.

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

**Figure C.4. Decomposition of employment growth of surviving entrants over two years**



*Note:* Sectors covered are manufacturing, construction and non-financial business services. Post entry growth is the ratio of final to initial employment. Firms with more than 100 employees in their first year of positive employment are reclassified as existing firms of unknown age.

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

## REFERENCES

- Aghion, P., T. Fally and S. Scarpetta (2007), "Credit constraints as a barrier to the entry and post-entry growth of firms", *Economic Policy*, Vol. 22, No. 52, pp. 732-779.
- Andrews, D., C. Criscuolo and C. Menon (2014), "Do Resources Flow to Patenting Firms?: Cross-Country Evidence from Firm Level Data", *OECD Economics Department Working Papers*, No. 1127, OECD Publishing, Paris. <http://dx.doi.org/10.1787/5jz2lpmk0gs6-en>
- Audretsch, D. B. (1991), "New-firm survival and the technological regime". *The Review of Economics and Statistics*, Vol. 73, No. 3, pp. 441-450.
- Audretsch, D. B. (1995), "Innovation, growth and survival". *International Journal of Industrial Organization*, Vol. 13, No. 4, pp. 441-457.
- Bartelsman, E. J. and M. Doms (2000), "Understanding productivity: Lessons from longitudinal microdata". *Journal of Economic Literature*, Vol. 38, No. 3, pp. 569-594.
- Bartelsman, E., J. Haltiwanger and S. Scarpetta (2009), "Measuring and analyzing cross-country differences in firm dynamics". In Dunne, T., J. Bradford Jensen and M. J. Roberts (Eds.), *Producer dynamics: New evidence from micro data* (pp. 15-76), University of Chicago Press.
- Bartelsman, E., S. Scarpetta and F. Schivardi (2005), "Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries". *Industrial and Corporate Change*, Vol. 14, No. 3, pp. 365-391.
- Beck, T., A. Demirguc-Kunt, L. Laeven and R. Levine (2008), "Finance, firm size, and growth", *Journal of Money, Credit and Banking*, Vol. 40, No. 7, pp. 1379-1405.
- BER (2016), "The Small, Medium and Micro Enterprise Sector of South Africa", Research Note No. 1, University of Stellenbosch, Stellenbosch.
- Bravo-Biosca, A., C. Criscuolo and C. Menon (2016), "What drives the dynamics of business growth?". *Economic Policy*, Vol. 31, No. 88, pp. 703-742.
- Calvino, F., C. Criscuolo and C. Menon (2015), "Cross-country evidence on start-up dynamics", *OECD Science, Technology and Industry Working Papers*, 2015/06, OECD Publishing, Paris. doi: <http://dx.doi.org/10.1787/5jrxtkb9mxtb-en>
- Calvino, F., C. Criscuolo and C. Menon (2016), "No Country for Young Firms?: Start-up Dynamics and National Policies", *OECD Science, Technology and Industry Policy Papers*, No. 29, OECD Publishing, Paris. doi: <http://dx.doi.org/10.1787/5jm22p40c8mw-en>
- Coelho, D., C. H. Corseuil and M. N. Foguel (2017), "Employment growth of establishments in the Brazilian economy: Results by age and size groups", in *Business Dynamics and Productivity*, OECD Publishing, Paris. doi: <http://dx.doi.org/10.1787/9789264269231-6-en>

- Criscuolo, C., P. Gal and C. Menon (2014), "The Dynamics of Employment Growth: New Evidence from 18 Countries", *OECD Science, Technology and Industry Policy Papers*, No. 14, OECD Publishing, Paris. doi: <http://dx.doi.org/10.1787/5jz417hj6hg6-en>
- Davis Tax Committee (2014), "Small and Medium Enterprises: Tax Considerations", *Interim Report*, July 2014, Pretoria.
- Fedderke, J., N. Obikili and N. Viegi (2018), "Mark-ups and concentration in South African manufacturing sectors: an analysis with administrative data", *South African Journal of Economics*, Vol. 86, No. S1, pp. 120-140.
- Haltiwanger, J., R.S. Jarmin and J. Miranda (2013), "Who Creates Jobs? Small Versus Large Versus young", *The Review of Economics and Statistics*, Vol. 95, No. 2, pp. 347-361.
- Kerr, A. (2018), "Job flows, worker flows, and churning in South Africa", *WIDER Working Papers South African Journal of Economics*, No. Vol. 86, No. S1, pp. 141-66 2016/37, UNU-WIDER.
- Kerr, A., M. Wittenberg, and J. Arrow (2014), "Job Creation and Destruction in South Africa", *South African Journal of Economics*, Vol. 82, No. 1, pp. 1-18.
- Klepper, S. (1996), "Entry, exit, growth, and innovation over the product life cycle". *The American Economic Review*, Vol. 86, No. 3, pp. 562-583.
- Malerba, F. (2002), "Sectoral systems of innovation and production". *Research policy*, Vol. 31, No. 2, pp. 247-264.
- OECD (2008), *OECD Economic Surveys: South Africa 2008*, OECD Publishing, Paris. [http://dx.doi.org/10.1787/eco\\_surveys-zaf-2008-en](http://dx.doi.org/10.1787/eco_surveys-zaf-2008-en)
- OECD (2013), *OECD Economic Surveys: South Africa 2013*, OECD Publishing, Paris. [http://dx.doi.org/10.1787/eco\\_surveys-zaf-2013-en](http://dx.doi.org/10.1787/eco_surveys-zaf-2013-en)
- OECD (2015a), *OECD Economic Surveys: South Africa 2015*, OECD Publishing, Paris. [http://dx.doi.org/10.1787/eco\\_surveys-zaf-2015-en](http://dx.doi.org/10.1787/eco_surveys-zaf-2015-en)
- OECD (2015b), *The Future of Productivity*, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264248533-en>
- OECD (2016), "No Country for Young Firms?", Policy Note, Directorate for Science, Technology and Innovation Policy Note, June 2016.
- OECD (2017a), *OECD Economic Surveys: South Africa 2017*, OECD Publishing, Paris. [http://dx.doi.org/10.1787/eco\\_surveys-zaf-2017-en](http://dx.doi.org/10.1787/eco_surveys-zaf-2017-en)
- OECD (2017b), *OECD Economic Surveys: South Africa 2017*, OECD Publishing, Paris. [http://dx.doi.org/10.1787/eco\\_surveys-zaf-2017-en](http://dx.doi.org/10.1787/eco_surveys-zaf-2017-en)
- OECD (2017ba), *Business Dynamics and Productivity*, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264269231-6-en>
- OECD (2017c), *The Geography of Firm Dynamics: Measuring Business Demography for Regional Development*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264286764-en>.

- Oosthuizen, M., K. Lilenstein, C. Fernandez, J.C. Medellin and F. Fernandez (2017), “Rethinking the effect of informality on inclusive growth: Lessons from Colombia and South Africa for their regions”, *ELLA Research Paper Series: Economic Development: Informality*, Development Policy Research Unit University of Cape Town and Fedesarrollo
- Pieterse, D., C. Kreuser and E. Gavin (2016), “Introduction to the South African Revenue Service and National Treasury firm-level panel”, *WIDER Working Papers*, No. 42/2016, UNU-WIDER/South African Journal of Economics, Vol. 86, No. S1, pp. 6-39.
- Roberts, S. (2016), “An Agenda for Opening up the South African Economy: Lessons from Studies of Barriers to Entry”, CCRED Policy Brief, University of Johannesburg, Johannesburg.
- Santarelli, E. and M. Vivarelli (2007), “Entrepreneurship and the process of firms’ entry, survival and growth”. *Industrial and Corporate Change*, Vol. 16, No. 3, pp. 455-488.
- SARS (2009), *Annual Report: South African Revenue Service 2008/09*.
- Siegfried, J. J. and L. B. Evans (1994), “Empirical studies of entry and exit: a survey of the evidence”. *Review of Industrial Organization*, Vol. 9, No. 2, pp. 121-155.
- Statistics South Africa (2014), *Construction industry, 2014*, Report No. 20-20-01, Statistics South Africa, Pretoria.
- Syverson, C. (2011), “What determines productivity?”. *Journal of Economic Literature*, Vol. 49, No. 2, pp. 326-65.
- World Bank (2013), “Chapter 3: Jobs and Productivity”, in *World Development Report 2013*, World Bank, Washington DC.