



WIDER Working Paper 2017/107

## **Defining high-growth firms in South Africa**

Mulalo Mamburu\*

April 2017

**Abstract:** Traditionally, much of the research on economic growth drivers has been focused on small and medium enterprises. In recent years the academic focus on small businesses has shifted to a particular group of firms that are interesting from an economic growth and policy development perspective, namely high-growth firms. While the standard definition that is recommended by the Organisation for Economic Co-operation and Development is widely accepted, various scholars have used different definitions concentrating on different variables such as turnover or employment growth. Using new South African firm-level data, the study hypothesizes that the identification of high-growth firms is highly sensitive to the measure of firm growth, such that different firm growth measures will return samples of firms with significantly different demographic characteristics. These differences will then have an impact on the findings of analyses based on these growth measures. They will also have implications for public policy recommendations that seek to encourage the emergence of high-growth firms.

**Keywords:** firm growth, high-growth firms, South Africa, entrepreneurship

**JEL classification:** C33, C38, L25, L26

**Acknowledgements:** This work has been conducted towards a degree at the University of the Witwatersrand, South Africa. I would like to thank Gareth Roberts and Duncan Pieterse for their inputs and support. This paper was presented at the poster session in UNU-WIDER and the National Treasury conference on ‘Growth and Development Policy: New Data, New Approaches, and New Evidence’, in Pretoria conference 31 November – 1 December 2016.

---

\* National Treasury of South Africa, Pretoria, South Africa, email: [mulalo.mamburu@treasury.gov.za](mailto:mulalo.mamburu@treasury.gov.za)

This study has been prepared within the UNU-WIDER project on ‘[Regional growth and development in Southern Africa](#)’.

Copyright © UNU-WIDER 2017

Information and requests: [publications@wider.unu.edu](mailto:publications@wider.unu.edu)

ISSN 1798-7237 ISBN 978-92-9256-331-8

Typescript prepared by Lesley Ellen.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

UNU-WIDER acknowledges specific programme contribution from the National Treasury of South Africa to its project ‘[Regional growth and development in Southern Africa](#)’ and core financial support to its work programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

## **1 Introduction**

Traditionally, much of the research on economic growth drivers has been focused on small and medium enterprises (SMEs). In recent years the academic focus on small businesses has shifted to a particular group of firms that are interesting from an economic growth and policy development perspective, namely high-growth firms. Broadly defined as the fastest growing firms in an economy, these businesses have been shown to be responsible for a disproportionate level of net employment growth in developed countries, despite the high mortality and low growth rates that are observed among SMEs in general (Shane 2009). Moreover, they have been found to be among the most productive and innovative firms in their industries, making them significant contributors to economic growth (Li and Rama 2013).

As such, scholars such as Shane (2009) argue that blanket policies that seek to encourage all SME activities are inefficient, as not all SMEs contribute to economic development; they argue that policy should focus rather on identifying firms that have high growth potential and seek to create an environment which would support these businesses. With the general consensus in the literature shifting towards this line of thinking, the policy focus among many governments and institutions is gradually shifting towards entrepreneurship policy which supports high-growth firms, as opposed to traditional SME policy (e.g. Lilischkis 2011).

An appropriate point of departure in the discussion of high-growth firms is how they are defined. While the standard definition that is recommended by the OECD is widely accepted, various scholars have used different definitions concentrating on different variables such as turnover or employment growth. From the Birch Index to composite measures by other scholars such as Acs et al. (2008), the literature features a vast array of measures that define high growth in different ways, such that there is presently no consensus on how to identify high-growth firms. As noted by Delmar (1997), the use of different measures of growth impedes the extent to which empirical results in the literature may be compared. Furthermore, these methodological differences have implications for what policy makers can draw from the research on high-growth firms.

This study will make use of data from administrative tax records to analyse how the identification of high-growth firms in South Africa may be sensitive to the definition of growth used. It will draw on the literature and compile a literature survey on the definition and impact of high-growth firms. The paper will then discuss the data that will be used in the analysis, which was previously unavailable and is novel in the context of a middle-income economy. Section 4 will discuss the high-growth firm definitions to be used and will present the methodology that will be followed. Thereafter, the econometric findings of the study will be presented and discussed, followed by concluding remarks.

## **2 Literature review**

### **2.1 The rise of high-growth firms (HGFs) in research**

In his analysis of employment creation in the United States, Birch (1979) argued that the job-generating firm is not a large corporation but rather a small business. Despite some criticism that this finding evoked, e.g. Brown et al. (1990) and Davis et al. (1996), it became very influential among policy makers and inspired a theory in the literature that small businesses are vital to the job creation process (Acs 1992; Storey 1994). In a later breakthrough study, Birch and Medoff

(1994) refine the theory about who creates jobs and suggest that most creation of jobs by small firms occurs within a small group of firms, which they termed 'gazelles'. The size of the firm, they argued, is less significant and it is, rather, rapid growth that characterized these businesses.

The finding that a small group of HGFs contributes disproportionately to the creation of jobs in the economy has since inspired much similar research in other countries. Henrekson and Johansson (2010) conduct a survey of the literature and find 20 studies across different developed economies, all supporting the view that a small group of HGFs account for a disproportionately large share of employment growth. They have also been found to be among the most productive and innovative firms in their industries, making them significant contributors to economic growth (Li and Rama 2013). Moreover, these enterprises are also found to have positive impacts on their industries; for instance, Bos and Stam (2014) find that an increase in the number of gazelles in a given industry has a positive effect on the subsequent growth of the industry.

## **2.2 Key issues of interest in research**

Central to the issue of the prevalence of HGFs, another critical matter is what is meant by a high-growth firm. The OECD-Eurostat convention has been broadly suggested, where a HGF is defined as one that initially possesses ten or more employees, or that has at least four times the national per capita income in annual revenues, and that experiences average annualized growth greater than 20 per cent per annum over a three-year period. There exists heterogeneity in the choice of growth indicator in the literature. However, most studies use either turnover or employment growth. Gazelles are defined as young HGFs, aged up to five years old.

Another matter of importance in analysing the literature on HGFs is to distinguish between the impact of HGFs and that of SMEs in general on the economy. The World Bank's Enterprise Surveys distinguish between two types of entrepreneurs, the necessity entrepreneur who typically starts a micro enterprise due to a lack of employment opportunities and has no employees, and the opportunity entrepreneur who starts an enterprise due to a perceived opportunity in the market. It is generally accepted in the literature that HGFs emerge from firms founded by the opportunity entrepreneur (ANDE 2012) as these are the firms that seek to grow and employ more people.

Scholars such as Brown et al. (2014) suggest that future HGFs will emerge from today's stock of SMEs and, therefore, public policy should not only focus on fostering and supporting new business but also on helping SMEs that are potential HGFs to realize growth. Others such as Shane (2009) take a firmer stance and argue that blanket policies that seek to encourage all SME activities are inefficient as not all SMEs contribute to economic development; they argue that policy should focus rather on identifying firms that have high growth potential and seek to create an environment which would support these businesses. Coad et al. (2014) argue, however, that high growth among HGFs is found not to be persistent over time, and it would be ill-advised to build public policy aimed at targeting HGFs. What is common among the different perspectives, though, is the desire to cultivate conditions that are favourable to the emergence of businesses that will drive employment and productive growth. The policy focus among many governments and institutions, therefore, is gradually shifting towards entrepreneurship policy as opposed to traditional SME policy (e.g. Lilischkis 2011).

Due to this desire to create an environment that supports HGFs, much of the research is dedicated to identifying the characteristics of firms that become HGFs and their entrepreneurs. Research on European start-ups (Audretsch 2012) suggests that educational background, prior experience in the relevant industry, and prior experience as an entrepreneur, or working at a start-up are all significant in determining whether a firm will exhibit high growth. In the African context,

Fafchamps and Woodruff (2014) conduct an experiment to test whether business plan competitions are able to identify potential HGFs, and they find that ability and management practices are significant determinants of whether a firm will experience higher growth.

The literature also considers other firm-specific characteristics that might determine the growth of firms. Goedhuys and Sleuwaegen (2010) find that firms in Africa that engage in product innovation and have transport means and internet connectivity in the form of their own website also display higher growth rates. This suggests that infrastructure allows firms to create wider input and output markets, which allows them to grow faster. The broader literature such as research on European Union states (Mitusch and Schimke 2011), has also suggested that firms that are active in cooperations, are situated within a cluster, have access to project funding, and are part of growing industries are more likely to become HGFs.

### **2.3 The issue of methodological considerations**

The availability of more reliable data has encouraged more research that helps scholars and policy makers alike to better understand HGFs. However, a persistent issue that remains is the heterogeneity in the definitions used to select these firms. While the standard OECD-Eurostat definition is widely used, the literature still features various different definitions concentrating on different variables such as turnover or employment growth. Scholars (e.g. Delmar 1997; Delmar and Davidsson 1998; Almus 2002; Coad et al. 2014) have noted the lack of consensus on this matter and the significant variation in methodology across different studies in this field. For instance, in their survey of the literature, Henrekson and Johansson (2010) find four different measurements used along with different growth indicators.

Delmar and Davidsson (1998) suggest four issues that must be considered when measuring firm growth: the indicator of growth; measurement of growth; the period under consideration and the process of growth, that is, whether growth is acquired or organic. The source of much variation in the literature has been indicators and measurements of growth.

The indicator of growth refers to what variable we observe to identify growth. Sales and employment are most widely used in the literature as these are most widely accessible and present fewer shortcomings than other measures such as assets (Delmar et al. 2003). While sales has been noted as a preferred indicator due to its insensitivity to capital intensity and it applies to most firms, it is sensitive to inflation and for some start-ups it is possible for employment and assets to grow before there are any sales (Delmar et al. 2003). On the other hand, while employment may be the preferable choice when the rationale for analysis is a broader interest in employment creation (Schreyer 2000), it suffers the disadvantage that employment is dependent on labour productivity such that a firm can grow substantially with minimal employment growth (Delmar et al. 2003).

The choice of growth measure refers to whether we measure absolute or relative growth. It is widely acknowledged that relative measures such as percentage change thresholds favour smaller firms, while absolute measures such as raw changes in size over time are biased towards large firms (Almus 2002).

In order to reconcile both views, Birch (1987) developed an index which is a combined measure of growth as follows:

$$(x_{t1} - x_{t0}) \cdot \frac{x_{t1}}{x_{t0}}$$

where  $x_{t1}$  is the number of employees at time  $t_1$ . The use of other such indices is an approach that has been taken by other scholars (e.g. Brüderl and Preisendörfer 2000).

Delmar (1997) has argued that these differences in the methodology when studying growth have implications for comparing findings in the literature. For instance, different growth indicators are responses at different phases of firm growth, such that they may not be directly comparable. Furthermore, where a model is used to interpret the quantitative impact of independent variables on growth, as is the case in many studies on the determinants of firm growth, transformations such as logs which would provide for a better fit would also complicate the interpretation of the finding to the extent that they would not be directly comparable with other studies.

## **2.4 The implications of methodological differences**

Some studies have undertaken to analyse the significance of methodology on the results of firm growth studies. In their influential analysis of firm growth methodology, Delmar et al. (2003) argue that firms grow in different ways and the patterns of growth vary over time. They compare the populations of HGFs derived from six different growth measures and find little correlation among these definitions. To investigate the hypothesis regarding the patterns of firm growth, they further expand the set of variables used to define HGFs with further categories that measure the regularity of growth. Through cluster analysis they find groups of HGF definitions that exhibit similar growth patterns and demographic characteristics that in turn differ from other clusters. These findings strengthen their argument that a standard definition of HGFs will not necessarily be appropriate for all applications as it would ignore the different growth patterns of firms.

Daunfeldt et al. (2014) examine whether HGFs defined in different ways are the same firms across definitions and whether they are equally significant to economic outcomes. Using different combinations of various growth measures and indicators, they produce differences in mean age and size values between these definitions and show that the characteristics of HGFs differ depending on the choice of definition. Furthermore, they find that the relationship between HGF status and variables such as age and size is highly dependent on what definition is used. This is consistent with Delmar (1997) which finds that the significance of independent variables in firm growth analysis varies with the choice of definition.

The findings of these studies illustrate the significance of methodological considerations in firm growth research. An appropriate point of departure for such research in the South African context, then, would be to demonstrate how the results of firm growth analysis would be impacted by these issues, making use of firm-level data. Because the existing empirical evidence has overwhelmingly focused on developed economies, this research will also provide novel evidence regarding HGFs in the context of developing economies.

### 3 Data discussion

The analysis of growth-focused firms relies on the availability of reliable and detailed firm-level data. Such data has historically been difficult to obtain for empirical research in South Africa. However, the new South African Revenue Service and National Treasury Firm-Level Panel (SARS-NT panel) (SARS-NT 2016) now makes such analysis possible. South African administrative tax data received during 2015 from multiple sources were merged to create the unbalanced panel. These sources are company income tax (CIT) data from all CIT-registered entities who submit CIT forms; employee tax data from the IRP5 and IT3 forms; custom records from traders and value-added tax (VAT) data from VAT-registered firms. As of 2016, the panel contains administrative records for the 2008–14 period. The various fields in the related forms are used to construct variables in the panel. All variables that could be used to identify firms or individuals were removed to ensure anonymity. Pieterse et al. (2016) discuss the construction of the panel.

The panel provides a significant amount of data relating firms and employees. The CIT data contains comprehensive financial and tax information, firm demographic characteristics such as the sector in which it operates, employment information, and trade data for registered entities. At the employee level, the panel contains information regarding incomes, deductions, and payments made by firms related to their employees. This richness in the data makes it particularly useful in analysing firm performance over time.

The unit of analysis in the panel is the firm. Here a firm is a CIT-registered entity that has completed an IT14 and/or ITR14 form. A single firm can be associated with multiple branches, and some employee data is aggregated to the firm level, such as the number of jobs associated with the firm in a given period. As noted by Delmar et al. (2003), a key disadvantage of defining a firm as a given CIT number is that discontinuations in the data may occur where a firm changes ownership. To further distinguish firms of economic significance from other non-active CIT-registered entities such as dormant companies and body corporates, we consider only firms with non-missing and non-zero sales data as well as non-missing fixed capital stock and labour data.

Despite the obvious advantage of the richness offered by the data, it has potential sources of bias. Pieterse et al. (2016), in their detailed discussion of the construction and features of the panel, highlight that it only contains tax registered firms and, among those, only the firms that actually completed a tax return in the relevant period. This is particularly significant for unregistered small, young, or informal firms which may be of particular interest in the present analysis. On the other hand, the data features a larger sample size than survey data and its longitudinal nature allows us to track firm performance over time. These features of the panel allow for the analysis of many firm-level research questions which were previously difficult to tackle due to the limitation of data availability, such as the research questions which form the focus of this study.

We consider whether limiting the population to those firms in the sample that have non-missing sales, capital, and labour data will substantially skew the results of the study. Table 1 reports the distribution of the two populations across size, age, and sector in the year 2010. Firms that have missing sales, capital, and labour data also have missing characteristic information that is of particular interest in this analysis, as can be seen by the considerable variation in number of firms containing data across variables of interests in the entire population. On the other hand, firms that have this data tend to also have other information that is of interest. To draw a clearer comparison of the distributions, we ignore observations with missing data and report the distributions for the two populations in Table 1.

Table 1: The distribution of the populations across firm size, age, and sector in 2010

	Percentage of entire population (%)	Percentage of subpopulation with sales, labour, and fixed capital data (%)
<i>Firm size (number of employees)</i>		
1–4	39.55	34.12
5–9	21.28	21.14
10–19	16.39	18.02
20–49	13.11	15.29
50–199	7.44	8.83
200+	2.23	2.60
Number of firms	219,832	131,428
<i>Firm age (from birth year)</i>		
1	9.36	2.40
2–4	29.49	19.16
5–9	29.14	31.49
10–19	23.21	34.72
20+	8.81	12.23
Number of firms	785,370	131,903
<i>Sector classification</i>		
Manufacturing	25.39	29.52
Construction	8.22	8.49
Wholesale and retail trade	16.49	18.84
Finance, insurance, and real estate	20.24	17.17
Professional and technical services	11.58	10.06
Number of firms	244,840	121,511

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

The firm size distribution across the populations is broadly similar, with most firms having fewer than 20 employees; this is consistent with findings that SMEs constitute the bulk of firms in the economy (e.g. Kerr et al. 2013). The greatest difference between the two populations occurs among the smallest firms, which are under-represented in the subpopulation. The firm age distributions show greater variation as the subpopulation has a much smaller proportion of the youngest firms. It is also notable that the entire population is most concentrated in the ages 1–4 years, while the subpopulation features more firms aged 10–19 years. Finally, we present the distributions of the largest sectors in the panel, which show the largest variation between the two populations. Firstly, we note that while the financial services sector is larger than wholesale and retail trade in the population at large, the inverse is the case in the subpopulation. We also note the proportion of firms in the manufacturing sector is larger in the subpopulation than the population at large. However, it is the largest sector in both populations. No other significant differences are evident between the two distributions.

## 4 Methodology

This study borrows much of its design from Delmar et al. (2003), who define six different growth definitions by which a firm may be considered a HGF and then compare the samples derived from these indicators. A survey of the literature informs the definitions that will be compared in the study. For each of the definitions, we observe firms in the sample for three-year periods ending 2012, 2013, and 2014. We then aggregate the firms fulfilling the criteria in each period to constitute the sample of HGFs for a given definition.

### 4.1 Eurostat-OECD relative growth definitions

In the Eurostat-OECD Manual on Business Demography Statistics (Eurostat and OECD 2007), it is recommended that in studying high-growth enterprises, both employment and turnover are observed. In order to remove the distortion caused by the growth of micro enterprises, they also recommend employment and turnover thresholds to ensure that only economically relevant growth is observed, while the threshold must be low enough to avoid excluding too many firms. They recommend the following definitions for HGFs:

1. A firm that initially possesses at least ten employees or has at least four times the national per capita income in annual revenues and that experiences *average annualized employment growth* of at least 20 per cent per annum, over a three-year period.
2. A firm that initially possesses at least ten employees or has at least four times the national per capita income in annual revenues and that experiences *average annualized turnover growth* of at least 20 per cent per annum, over a three-year period.

We apply this definition to obtain samples of firms that experienced rapid growth for the periods ending 2012, 2013, and 2014. For per capita income, we use the World Bank's GDP per capita in current prices at the onset of each period, that is, the GDP per capita in 2010, 2011, and 2012 respectively.

### 4.2 Absolute growth definitions

Another widely used growth measure in the literature is the absolute growth of employment or turnover over a three-year period. More specifically, these measures define HGFs as:

3. A firm with at least ten employees that is among the top 10 per cent of all firms in terms of annual average on *absolute employment growth*, over a three-year period.
4. A firm with at least ten employees that is among the top 10 per cent of all firms in terms of annual average on *absolute turnover growth*, over a three-year period.

We apply this definition by observing the absolute annual change in employment and turnover and defining those firms with the highest annual changes in all three years for the periods ending 2012, 2013, and 2014 respectively as HGFs. Like Daunfeldt et al. (2014) and Delmar et al. (2003), we impose a minimum threshold of ten employees to exclude growth of micro firms. We also use the 10 per cent cut-off point as recommended by Storey (1994).

### 4.3 The Birch Index

Other scholars have advocated for the use of composite measures to mitigate the limitations of absolute and relative measures. The most popular among these is arguably the Birch Index, which is commonly applied to employment and turnover growth:

5.

$$(x_{t1} - x_{t0}) \cdot \frac{x_{t1}}{x_{t0}}$$

where  $x$  is employment.

6.

$$(y_{t1} - y_{t0}) \cdot \frac{y_{t1}}{y_{t0}}$$

where  $y$  is turnover.

We define a HGF as one which is among the top 10 per cent of all firms in terms of the Birch Index in all three years over a three-year period.

### 4.4 Other composite measures

Davis et al. (1996) have also proposed a composite growth measure which is used by US statistical agencies:

7.

$$\frac{(x_{t1} - x_{t0})}{(x_{t1} + x_{t0})} / 2$$

where  $x$  is employment.

8.

$$\frac{(y_{t1} - y_{t0})}{(y_{t1} + y_{t0})} / 2$$

where  $y$  is turnover.

We observe this growth measure for all firms and define a HGF in the same way used for the Birch Index.

Acs et al. (2008) propose another growth indicator:

9. They define the employment growth quantifier (EGQ) as the product of the absolute and relative change in employment over a four-year period, and define a high impact firm as one whose turnover has at least doubled and with an EGQ of at least two over the four-year period.

A few adjustments are made to this definition to make the observed growth more comparable to those of the other definitions in the study. Firstly, we observe the change in employment over a three-year period rather than the four-year period, as all other definitions in the study use this period. Secondly, we then define a HGF as one whose turnover has increased by at least 75 per cent rather than the 100 per cent and with an EGQ of 1.5 rather than 2 as specified by Acs et al. (2008), to account for the shorter observed time frame. In keeping with the other samples, we observe firms for three-year periods ending 2012, 2013, and 2014 and aggregate these to obtain a sample for EGQ HGFs.

Lastly, in their sensitivity analysis of the HGF definitions proposed by the Eurostat-OECD Manual, Petersen and Ahmad (2007) suggest an alternative composite measure which considers both employment and turnover growth of a firm:

10.

$$\sqrt{x_{t1}/x_{t0} \cdot y_{t1}/y_{t0}}$$

where  $x$  is employment and  $y$  is turnover.

We apply this growth measure to all firms and define HGFs in the same manner used for the Birch Index. Table 2 summarizes the definitions. We denote the OECD definitions as relative growth definitions, and the Davis, Haltiwanger and Schuh (1996) definition as DHS growth definitions for brevity.

Table 2: Summary of HGF definitions

Definition	Criteria
Relative employment	Average annualized employment growth of at least 20% per annum over three years.
Relative turnover	Average annualized turnover growth of at least 20% per annum over three years.
Absolute employment	Top 10% of all firms in terms of annual average absolute employment growth over three years.
Absolute turnover	Top 10% of all firms in terms of annual average absolute turnover growth over three years.
Birch employment	Top 10% of all firms in terms of Birch employment index in three consecutive years.
Birch turnover	Top 10% of all firms in terms of Birch turnover index in three consecutive years.
DHS employment	Top 10% of all firms in terms of DHS employment index in three consecutive years.
DHS turnover	Top 10% of all firms in terms of DHS turnover index in three consecutive years.
EGQ	Firms with EGQ of at least 1.5 and turnover growth of at least 75% over three years.
OECD composite	Top 10% of all firms in terms of OECD composite in three consecutive years.

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

The literature makes a distinction between organic, acquisition, and total growth. Delmar et al. (2003) note that organic and acquisition growth have distinct societal and firm-level implications, while they also have different impacts on firm performance. The panel features a variable indicating whether a firm has entered into a sale or purchase agreement in a given year. However,

this variable is very sparsely populated such that it is an unreliable indicator of acquisitions. We therefore consider total growth in the study.

We explore the differences between the samples generated by the different HGF definitions in the correlation analysis shown in Table 3. Notably, there is little correlation between the relative growth samples and the absolute growth samples even within the same growth indicator. We also observe that there is weak correlation between the samples generated from the same measure of growth using different indicators; for instance, the correlation between absolute employment and absolute turnover is only 0.275. Some of the composite measures seem to be more correlated with other definitions in the study; in particular, we note that the Birch and DHS definitions are more correlated to the absolute growth measures based on the same indicator. The OECD definitions, however, which are growing increasingly popular in the literature, seem to be largely uncorrelated with other measures of high growth. This suggests that we are dealing with heterogeneous groups of firms, with each definition deriving a different group of firms. This is consistent with findings in Delmar et al. (2003) and Daunfeldt et al. (2014).

Table 3: Correlations between HGF samples of different definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Relative employment	1	0.383	0.202	0.140	0.181	0.135	0.167	0.135	0.139	0.074
(2) Relative turnover	0.383	1	0.140	0.179	0.126	0.189	0.117	0.189	0.192	0.074
(3) Absolute employment	0.202	0.140	1	0.275	0.829	0.261	0.741	0.261	0.138	0.134
(4) Absolute turnover	0.140	0.179	0.275	1	0.249	0.229	0.229	0.811	0.109	0.082
(5) Birch employment	0.181	0.126	0.829	0.249	1	0.242	0.882	0.243	0.129	0.150
(6) Birch turnover	0.135	0.189	0.261	0.229	0.242	1	0.226	0.865	0.115	0.098
(7) DHS employment	0.167	0.117	0.741	0.229	0.882	0.226	1	0.223	0.123	0.154
(8) DHS turnover	0.135	0.189	0.261	0.811	0.243	0.865	0.223	1	0.115	0.085
(9) EGQ	0.139	0.192	0.138	0.109	0.129	0.115	0.123	0.115	1	0.145
(10) OECD composite	0.074	0.074	0.134	0.082	0.150	0.098	0.154	0.085	0.155	1

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

To further illustrate the heterogeneity in the samples defined by the HGF definitions, we produce mean differences in age for the samples in Table 4. We observe that there is considerable variation between the age profiles of the samples generated by each definition. For instance, HGFs defined by absolute turnover growth tend to be older. Notably, we also observe large age differences between samples defined by the same indicator but with a different measure: there are considerable differences between relative employment and absolute employment growth as well as relative turnover and absolute turnover growth. However, firms that grew in relative terms, whether employment or turnover growth was observed, are of a more similar age profile.

Table 5 shows the mean differences in firm size for the HGF samples. There are large variations between the firm sizes of the samples, even within a given growth measure. For instance, firms with the highest relative growth in employment have on average 413 fewer employees than those which experienced the highest growth in turnover. We again observe a considerable difference between firms that experienced absolute and relative growth, as firms that experienced the highest turnover growth in relative terms have on average 376 fewer employees than those which grew in absolute terms. As with the age profile, firms that experience relative growth either in employment or turnover are of a more similar size.

These differences in age and size profiles reinforce the suggestion that these are heterogeneous groups of firms. This has considerable implications for any empirical evidence that may be generated from a model using a given definition. For instance, econometric evidence regarding the impact of firm size on the probability of a given firm being a HGF would largely depend on how we define a HGF. A valuable extension on these findings, then, would be to compare the regression results of a given model specification applied to the different definitions, to assess whether these results will be consistent across the different HGF definitions.

Table 4: Differences in mean age between HGF samples of different definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Relative employment	-	1.833***	3.539***	-1.408***	2.944***	0.271***	2.204***	-0.463***	2.500***	2.964***
(2) Relative turnover	-1.833***	-	1.707***	-3.241***	1.111***	-1.561***	0.371***	-2.296***	0.668***	1.132***
(3) Absolute employment	-3.539***	-1.707***	-	-4.947***	-0.596	-3.268***	-1,335	-4.002***	-1.039***	-0.575***
(4) Absolute turnover	1.408***	3.241***	4.947***	-	4.352***	1.680***	3.612***	0.945	3.909***	4.373***
(5) Birch employment	-2.944***	-1.111***	0.596	-4.352***	-	-0.740***	-3.407	0.443***	0.021***	1.933***
(6) Birch turnover	-0.271***	1.561***	3.268***	-1.680***	0.740***	-	1.933***	-0.734	2.229***	2.693***
(7) DHS employment	-2.204***	-0.371***	1.335	-3.612***	3.407	-1.933***	-	-2.667***	0.296***	0.760***
(8) DHS turnover	0.463***	2.296***	4.002***	-0.945	-0.443***	0.734	2.667***	-	2.963***	3.427***
(9) EGQ	-2.500***	-0.668***	1.039***	-3.909***	-0.021***	-2.229***	-0.296***	-2.963***	-	0.464***
(10) OECD composite	-2.964***	-1.133***	0.575***	-4.373***	-1.933***	-2.693***	-0.760***	-3.427***	-0.464***	-

Note: \*p>0.10. \*\*p>0.05. \*\*\*p>0.01.

Source: Author's calculations based on SARS-NT panel.

Table 5: Differences in mean firm size between HGF samples of different definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Relative employment	-	9.824*	47.601***	-366.276***	-85.629***	-213.632***	-16.561***	-198.845***	-38.971**	33.169***
(2) Relative turnover	-9.824*	-	37.777***	-376.100***	-95.453***	-223.456***	-26.385***	-208.669***	-48.795**	23.345**
(3) Absolute employment	-47.601***	-37.777***	-	-413.876	-133.230***	-261.232	-64.162***	-246.446***	-86.572***	-14.432***
(4) Absolute turnover	366.276***	376.100***	413,876	-	280.647***	152.644***	349.714*	167.431	327.304***	399.445***
(5) Birch employment	85.629***	95.453***	133.230***	-280.647***	-	-128.003***	69.068***	-113.216***	46.658***	118.798***
(6) Birch turnover	213.632***	223.455***	261.232	-152.644***	128.003***	-	197.070***	14.787***	174.660***	246.801***
(7) DHS employment	16.561***	26.385***	64.162***	-349.714*	-69.068***	-197.070***	-	-182.284***	-22.410***	49.730***
(8) DHS turnover	198.845***	208.669***	246.446***	-167.431	113.216***	-14.787***	182.284***	-	159.874***	232.014***
(9) EGQ	38.971**	48.795**	86.572***	-327.304***	-46.658***	-174.660***	22.410***	-159.874***	-	72.140**
(10) OECD composite	-33.169***	-23.345**	14.432***	-399.445***	-118.798***	-246.801***	-49.730***	-232.014***	-72.140**	-

Note: \*p>0.10. \*\*p>0.05. \*\*\*p>0.01.

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

## 5 Specifying a model

The observed variation in sample characteristics from the different definitions suggests that there may be further implications for econometric results derived from a given model. That is, the significance of a given variable in the probability of a firm being a HGF may depend on what definition of HGF is being used. To test this, we define a linear probability model (LPM) and use each definition as the dependent variable to analyse whether the significance of independent variables in the model varies with HGF definition. We use a LPM for several reasons. Firstly, the coefficients resulting from such a regression are more readily interpretable than those of logistic models. Secondly, Angrist and Pischke (2008) suggest that using LPM is appropriate if we are only interested in the mean effect of the independent variable i.e.  $E(Y=1 | X)$  rather than the distribution of outcomes.

We use the probit model defined by Daunfeldt et al. (2014) as the basis for our model, with a few changes. Most significantly, we use pooled ordinary least squares (OLS) to estimate our model, rather than panel methods, as we are interested in the variation between firms at any point in time rather than variation in a given firm over time. The following is defined:

$$FIRMGROWTH_i = \alpha + \beta_1 age + \beta_2 age^2 + \beta_3 size + \beta_4 size^2 + \gamma sector + \sum year\ dummies + \varepsilon$$

where  $FIRMGROWTH_i$  is the growth measure  $i$  and takes the value 1 if a given firm fulfils that definition; age is the firm age; size is the number of the firm's employees; age and size have been squared to control for nonlinearity;  $\gamma sector$  is a categorical variable for firm sector and  $\varepsilon$  is an error term. We also include year dummies to control for period-specific effects. We note that firm sector is classified according to the Standard Industrial Classification (SIC) system.

Table 6 shows the preliminary results from the regressions of each HGF sample, with standard errors reported in parentheses. Firm size is found to have a positive, nonlinear relationship with all definitions of firm growth. Size has a more positive impact on the likelihood of being a HGF the larger a firm is, regardless of what growth measure or indicator is used. This is contrary to Daunfeldt et al. (2014) where the impact of firm size is found to be dependent on the definition of HGF used.

Firm age always has a negative impact on the likelihood of a firm being a HGF whenever growth is measured on turnover, regardless of what measure or index is used. This is in line with previous findings such as Daunfeldt et al. (2014) and Birch (1981) that age is a more important growth determinant than firm size. On the other hand, firm age has a positive impact on the likelihood of being a HGF whenever employment growth is measured, regardless of the measure used. This would suggest that firms are more capable of expanding their workforce the older they become. Notably, the age effect is particularly strong among the relative growth definitions.

Table 6: Results from regressions on HGF samples

	Independent variables			
	Age	Age <sup>2</sup>	Size	Size <sup>2</sup>
Relative employment	-0.004209*** (0.00012)	0.023026*** (0.00091)	-0.00003*** (0.00)	0.00666*** (0.00021)
Relative turnover	0.00022* (0.00011)	-0.03342*** (0.00088)	-0.00002*** (0.00)	0.00477*** (0.00015)
Absolute employment	-0.00041*** (0.00006)	0.00135*** (0.00042)	-0.00003*** (0.00)	0.00867*** (0.00017)
Absolute turnover	0.00079*** (0.00008)	-0.00437*** (0.00055)	-0.00004*** (0.00)	0.01114*** (0.00018)
Birch employment	-0.00032*** (0.00005)	0.00073*** (0.00037)	-0.00002** (0.00)	0.00691*** (0.00012)
Birch turnover	0.00090*** (0.00008)	-0.00584*** (0.00052)	-0.00003*** (0.00)	0.0091*** (0.00013)
DHS employment	-0.00030*** (0.00005)	0.00047 (0.00032)	-0.00002*** (0.00)	0.00591*** (0.00010)
DHS turnover	0.00068*** (0.00008)	-0.00371*** (0.00056)	-0.00004*** (0.00)	0.01066*** (0.00017)
EGQ	-0.00025*** (0.00002)	0.00025 (0.00015)	0.00*** (0.00)	0.00094*** (0.00004)
OECD composite	-0.00001 (0.00001)	-0.00041*** (0.00008)	0.00*** (0.00)	0.00022*** (0.00001)

Note: \*p>0.10. \*\*p>0.05. \*\*\*p>0.01.

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

The initial findings regarding the relationship between firm age and size and the likelihood of a firm becoming a HGF therefore supports the evidence that samples generated by different HGF definitions have significantly different demographic characteristics. If these differences do not extend to other factors of economic interest, such as profitability, then the variation between HGF samples may not have a significant impact on policy making and the analysis of these firms. If, however, the demographic variation between HGF samples extends to such factors then the implication for economic analysis and policy making will be more serious. It is therefore useful to compare the relationship between the different types of HGFs and some variables of economic interest, to understand how these correlates are associated with the different types of firms identified by these definitions.

The richness of the SARS-NT panel allows us to consider a wide array of factors such as sector data, financial performance, and trade activity. Firstly, we are interested in the firm sector and whether a given firm is a subsidiary of a group or a larger firm, as evidence from the literature suggests that firms which are subsidiaries or part of an affiliation are more likely to experience rapid growth (e.g. Delmar et al. 2003; Bjuggren et al. 2010). For sectors we use the SIC one-digit

classification, and for whether a firm is a subsidiary we consider a variable in the data where firms indicate whether or not they are a subsidiary of another firm.

For financial performance, we consider the real value added, capital–labour ratio, and net profit. The real value added will capture a firm’s contribution to total output, while the capital–labour ratio will capture the capital intensity of the firm and the net profit will indicate profitability. The real value added is defined as the difference between real sales and real cost of sales in Rands; the capital–labour ratio is the ratio between fixed capital stock and number of employees; net profit is a firm’s reported net profit in Rands.

The customs data in the panel allows us to analyse the import/export activity of firms with custom records. We consider whether a given firm imports or exports any goods, the number of countries it imports to and exports from, and the total value of imports and exports. The total values of imports and exports are reported in Rand values.

We also consider the research and development (R&D) expenditure of each firm, as this will give an indication of the focus on innovation and research of the enterprise. The amount spent on R&D is reported in Rand values.

Finally, we are also interested in whether a given firm accessed some sector-specific tax incentives; the tax incentives we consider are a deduction for commercial buildings and a deduction for buildings used in manufacturing; an exemption for films; a deduction for manufacturers and hotelkeepers; a deduction for learnership agreements registered in a given year and a deduction for R&D. For this variable we construct an indicator variable which is 1 if the firm claims a deduction for any of these incentives, and 0 otherwise.

Table 7: Descriptive characteristics of HGF samples and other firms, 2010–14

	Relative employment	Relative turnover	Absolute employment	Absolute turnover	Birch employment	Birch turnover
<i>Firm size</i>						
2010	7	7	84	68	93	67
2014	16	12	198	108	246	107
% change	129%	71%	135%	59%	165%	60%
<i>Real value added</i>						
2010	13,700,000.00	7,610,477.00	132,000,000.00	115,000,000.00	169,000,000.00	134,000,000.00
2014	18,900,000.00	15,100,000.00	203,000,000.00	190,000,000.00	250,000,000.00	220,000,000.00
% change	38%	98%	54%	65%	48%	64%
<i>Capital/labour ratio</i>						
2010	2,882,201.00	1,068,760.00	1,073,845.00	986,519.70	1,378,733.00	1,309,579.00
2014	215,434.00	236,696.70	148,743.40	1,281,882.00	152,725.90	2,159,057.00
% change	-93%	-78%	-86%	30%	-89%	65%
<i>Net profit</i>						
2010	3,755,017.00	2,970,566.00	28,600,000.00	31,900,000.00	34,500,000.00	37,600,000.00
2014	6,243,148.00	5,428,683.00	66,900,000.00	57,000,000.00	83,400,000.00	66,300,000.00
% change	66%	83%	134%	79%	142%	76%
<i>R&amp;D expenditure</i>						
2010	90,948.01	142,394.10	631,200.70	1,036,553.00	791,315.10	1,262,165.00
2014	52,770.23	41,072.15	127,041.70	108,284.00	166,771.50	133,734.50
% change	-42%	-71%	-80%	-90%	-79%	-89%
<i>Value of imports</i>						
2010	18,100,000.00	28,500,000.00	119,000,000.00	116,000,000.00	154,000,000.00	130,000,000.00
2014	28,500,000.00	52,900,000.00	205,000,000.00	228,000,000.00	261,000,000.00	254,000,000.00
% change	57%	86%	72%	97%	69%	95%
<i>Value of exports</i>						
2010	14,600,000.00	14,900,000.00	142,000,000.00	57,400,000.00	197,000,000.00	64,900,000.00
2014	19,900,000.00	24,800,000.00	136,000,000.00	104,000,000.00	183,000,000.00	117,000,000.00
% change	36%	66%	-4%	81%	-7%	80%
<i>Median firm age</i>						
	9	8	11	12	9	11

Table 7 (continued)

	DHS employment	DHS turnover	EGQ	OECD composite	All HGFs	All firms
<i>Firm size</i>						
2010	106	58	12	10	8	8
2014	288	94	47	90	14	9
% change	172%	62%	292%	800%	75%	13%
<i>Real value added</i>						
2010	198,000,000.00	109,000,000.00	3,694,794.00	3,430,641.00	19,900,000.00	12,800,000.00
2014	288,000,000.00	180,000,000.00	16,500,000.00	24,700,000.00	32,500,000.00	16,300,000.00
% change	45%	65%	347%	620%	63%	27%
<i>Capital/labour ratio</i>						
2010	1,639,775.00	1,147,639.00	238,551.40	460,773.40	2,502,564.00	1,359,324.00
2014	148,408.20	1,753,888.00	124,692.20	252,432.10	421,284.30	314,730.20
% change	-91%	53%	-48%	-45%	-83%	-77%
<i>Net profit</i>						
2010	40,200,000.00	30,200,000.00	1,638,297.00	2,448,151.00	6,088,322.00	3,844,678.00
2014	94,100,000.00	54,200,000.00	6,714,237.00	7,488,449.00	11,600,000.00	5,433,675.00
% change	134%	79%	310%	206%	91%	41%
<i>R&amp;D expenditure</i>						
2010	919,896.70	1,022,576.00	19,913.66	36,164.61	169,088.30	70,995.52
2014	199,621.00	106,536.10	27,946.00	6,206.11	49,955.42	43,803.82
% change	-78%	-90%	40%	-83%	-70%	-38%
<i>Value of imports</i>						
2010	141,000,000.00	110,000,000.00	7,749,824.00	8,462,649.00	37,000,000.00	14,300,000.00
2014	220,000,000.00	217,000,000.00	26,200,000.00	22,000,000.00	64,400,000.00	39,200,000.00
% change	56%	97%	238%	160%	74%	174%
<i>Value of exports</i>						
2010	221,000,000.00	55,600,000.00	2,724,705.00	2,933,610.00	36,400,000.00	13,500,000.00
2014	183,000,000.00	101,000,000.00	12,100,000.00	13,800,000.00	41,700,000.00	31,200,000.00
% change	-17%	82%	344%	370%	15%	131%
<i>Median firm age</i>						
	11	12	8	7	9	10

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

Table 7 reports descriptive characteristics for each definition across some variables of interest. We show the median firm age in 2010 and the median firm size as these variables are particularly sensitive to outliers in the data; we show average figures for all other variables. On aggregate, HGFs are younger than other firms. While the relative-growth, OECD composite, and EGQ definitions generate samples of small firms that are SME-sized, the other samples feature large firms that grow significantly in the period under review. These firms add the most to increases in employment. As expected, real value added for all HGF samples is considerably greater than that of other firms; on average, HGFs add around twice as much real value to output than other firms by 2014, with even more real value added by HGF samples featuring larger firms.

We also note that HGFs are generally more capital-intensive than other firms, and their capital–labour ratios increased more in the period under review than the average firm in the sample. It is notable, however, that the change in capital intensity depends on the indicator used in the definition; samples of turnover-based definitions increased their capital intensity while employment-based definitions became more labour-intensive. This would be consistent with employment-growth HGFs switching activities from capital to labour or expanding through an increased reliance on labour rather than capital in response to increased demand and turnover-growth HGFs become increasingly mechanized and reducing their reliance on labour, or meeting expanding demand through a greater reliance on capital.

The R&D expenditure for all firms in the sample is highly volatile throughout the period under consideration, with large variations in reported expenditure from year to year; on average, however, HGFs spend more on innovation activities than other firms. We also consider the potential impact of the changes in legislature regarding the R&D tax allowance. In October 2012, the R&D allowance was amended from an automatic tax deduction to one that required approval by an adjudication panel administered by the Department of Science Technology. Indeed, among the HGFs, the definitions which identify larger firms all experienced drastic reductions in reported innovation expenditure in the 2014 financial year; however, the lack of subsequent years after the change inhibits the possibility of identifying a trend after 2014.

HGFs are found to be far more profitable than other firms, and their growth in net profit was more rapid than other firms in the period considered. Interestingly, among the HGF samples featuring larger firms, the employment-growth HGFs grew more in terms of profitability than turnover-growth HGFs. HGFs also have larger links to international markets: they import a higher value of goods, and they also export more.

Table 8: Sector distributions of HGF samples and other firms in 2010

	Relative employment	Relative turnover	Absolute employment	Absolute turnover	Birch employment	Birch turnover
Manufacturing	28.94	27.04	29.01	36.17	26.57	35.65
Construction	8.57	8.51	8.6	3.74	9.93	4.26
Wholesale and retail trade	18.97	16.38	15.05	30.19	13.62	28.08
Finance, insurance, and real estate	16.4	18.85	16.87	9.11	18.88	9.98
Professional, technical, and scientific activities	9.88	11.55	8.75	5.26	8.94	5.64

  

	DHS employment	DHS turnover	EGQ	OECD composite	All HGFs	All firms
Manufacturing	25.33	35.75	25.43	22.89	28.88	26.14
Construction	9.92	3.98	10.76	10.24	8.32	7.76
Wholesale and retail trade	13.45	29.96	15.54	8.43	18.78	17.11
Finance, insurance, and real estate	20.37	9.33	18.26	24.1	16.76	14.87
Professional, technical, and scientific activities	9.2	5.15	10.98	13.86	10.08	8.53

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

Table 8 reports the sector distribution for each HGF definition. We focus on manufacturing and the four larger services sectors, namely: construction; wholesale and retail trade; financing, insurance, and real estate; and professional, technical, and scientific activities. HGFs are most concentrated in the manufacturing sector among all the definitions but most notably so for the definitions that yield larger firms. For the definitions that yield larger, turnover-growth firms, HGFs are also heavily concentrated in wholesale and retail trade, with many fewer firms in the other sectors. On the other hand, more HGFs emerge from the finance sector when the definitions yielding larger, employment-growth firms are considered. For the smaller, younger HGFs, the finance sector is also significant, and they are less concentrated in a given sector than among the large, turnover-growth HGFs. We also note that those definitions which are most concentrated in the manufacturing and wholesale sector are also those which yield samples of firms that became more capital-intensive in the 2010–14 period. Those firms which became more labour-intensive in this period are more evenly distributed across the sectors.

There are therefore several notable characteristics of HGFs that make them economically interesting. Consistent with the literature, there is a general finding that although HGFs are a small proportion of firms in the economy, they are economically impactful. The nature and extent of the impact, however, varies by the definition used to identify HGFs, as they evidently grow very differently and display different characteristics. In light of this, it is worthwhile to proceed with a regression analysis to determine what is associated with the likelihood of becoming a given type of HGF and how the correlates of the different HGFs samples vary.

## 6 Regression results

We can incorporate each of these variables in the above-specified LPM to analyse the association between the variable of interest and the type of firm growth. For ease of analysis, we sort the ten definitions into groups of criteria that derive similar samples and compare these groups of definitions with each other. To group similar definitions with one another, a simple cluster analysis based on that used by Delmar et al. (2003) is employed. This form of analysis divides data into groups or clusters which are similar between them and dissimilar to objects in other groups (Madhulatha 2012).

We use hierarchical clustering with Ward’s method and Euclidean distances to determine the number of clusters for the definitions. An examination of the cluster tree resulting from this clustering suggests that the definitions may be divided into three groups. The first cluster is comprised of Birch employment, DHS employment, and absolute employment growth HGFs; the second cluster consists of the OECD composite, EGQ, relative employment, and relative turnover growth HGFs; the third cluster is comprised of Birch turnover, DHS turnover, and absolute turnover growth HGFs. These clusters have the characteristics shown in Table 9.

Table 9: Cluster characteristics in 2010

	Name	N	Average firm size	Average firm age
Cluster 1	Large employment growers	3,069	372	12.40
Cluster 2	Small growers	42,806	44	8.92
Cluster 3	Large turnover growers	5,417	379	13.25
All HGFs		51,292	79	9.34

Source: Author’s calculations based on SARS-NT panel (SARS-NT 2016).

Cluster 1, the large employment growers, features those definitions which yield larger, older firms that add the largest number of jobs over the period. Cluster 2, the small growers, is comprised of those definitions which yield smaller, younger HGFs which experience rapid growth relative to their base. Cluster 3, the large turnover growers, features definitions which yield the large, mature firms in the sample that grow in terms of turnover. We note that there are far more HGFs in cluster 2 than the other two clusters, such that HGF definitions which lean towards small firms will also yield a large number of small firms, while definitions which yield large firms will also yield a small number of large firms.

We then use these clusters as the samples on which the LPM is applied for a correlate analysis, to see what the firms in each cluster look like. Table 10 reports the regression results for the LPM incorporating different variables of interest. The dependent variable in each specified equation is an indicator variable for whether a firm is in a given HGF cluster. As an additional dependent variable, we also define a discrete variable for the number of HGF criteria a given firm satisfies, where the variable takes on the value 10 if a firm is identified as a HGF using all ten definitions. This variable is used to measure the association of a given correlate with satisfying more HGF definitions.

Table 10a: Sectors of HGF clusters

	Cluster 1	Cluster 2	Cluster 3	All HGFs
Agriculture	0.0025*** (0.0007)	0.0052** (0.0022)	-0.0121*** (0.0009)	-0.0162*** (0.0048)
Mining	0.0137*** (0.0019)	0.0163*** (0.0038)	-0.0158*** (0.0020)	0.0185 (0.0113)
Electricity, gas, and water	0.0062*** (0.0009)	0.0188*** (0.0033)	-0.0027** (0.0012)	0.0625*** (0.0073)
Construction	0.0034*** (0.0005)	-0.0027* (0.0016)	-0.0191*** (0.0005)	-0.0376*** (0.0032)
Wholesale and retail trade	0.0001 (0.0003)	0.0026** (0.0012)	0.0165*** (0.0006)	0.0419*** (0.0027)
Transport and storage	0.0068*** (0.0007)	0.0060*** (0.0021)	0.0097*** (0.0010)	0.0603*** (0.0052)
Catering and accommodation	-0.0063*** (0.0006)	-0.0218*** (0.0020)	-0.0316*** (0.0005)	-0.1384*** (0.0037)
Information and communication	-0.0017 (0.0021)	-0.0382*** (0.0057)	-0.0182*** (0.0023)	-0.0784*** (0.0140)
Finance, insurance, and real estate	0.002*** (0.0003)	-0.0428*** (0.0011)	-0.0173*** (0.0004)	-0.0793*** (0.0025)
Professional, technical, and scientific activities	0.0027*** (0.0004)	-0.0213*** (0.0014)	-0.0150*** (0.0005)	-0.0445*** (0.0029)
Other services	-0.0057*** (0.0004)	-0.1435*** (0.0016)	-0.0253*** (0.0004)	-0.2580*** (0.0030)

Note: \*p>0.10. \*\*p>0.05. \*\*\*p>0.01.

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

We first consider the base model to analyse how firm sector affects the likelihood of being in a given cluster, where we use manufacturing as the base sector for comparison. Notably, the wholesale and retail trade sector is consistently more likely to produce HGFs than manufacturing, while the other service sectors are far less likely to do so in general. Among the large employment growers and the small growers, the HGFs are more likely to emerge from the primary sectors than from manufacturing. For the large employment growers, the services sectors are more likely to yield HGFs, especially in construction. On the other hand, firms focused on turnover growth, are more likely to emerge from manufacturing.

## 6.1 Financial characteristics

Next we consider some financial characteristics of HGFs. We add the logged capital–labour ratio of the firm to the base model and find that a significant and positive relationship exists between capital intensity and being in all three clusters, and the relationship is largest among large turnover growers. Capital intensity is also associated with the number of HGF definitions a firm fulfils: firms with a higher capital intensity also fulfil a larger number of HGF criteria. The logged net profit of a firm is also strongly correlated with the likelihood of becoming all types of HGFs, particularly large employment growers. We also test the association between the clusters and the log of real value added and find a strong, positive relationship between all of the clusters and this

variable, particularly among the small growers. Moreover, increasing real value added is highly associated with fulfilling more types of growth as defined by the different HGF criteria.

Table 10b: Firm characteristics of HGF clusters

	Cluster 1	Cluster 2	Cluster 3	All HGFs
Log of capital–labour ratio	0.0014*** (0.00)	0.0086*** (0.0002)	0.0050*** (0.00)	0.0292*** (0.0004)
Log of real value added	0.0036*** (0.0002)	0.0443*** (0.0004)	0.0256*** (0.0003)	0.1341*** (0.0015)
Log of net profit	0.0031*** (0.0002)	0.0178*** (0.0003)	0.0170*** (0.002)	0.0763*** (0.0012)

Note: \* $p > 0.10$ . \*\* $p > 0.05$ . \*\*\* $p > 0.01$ .

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

## 6.2 Trade activity

We also consider the impact of trade activity on the likelihood of being a HGF. We note that whether a firm imports any intermediary goods or exports final goods is strongly associated with the likelihood of HGF status. Involvement in international trade has a highly significant, positive relationship with the likelihood of falling within all of the clusters, which suggests that all forms of growth are complemented by the openness of a firm. The number of countries that a firm exports to is positively correlated with the likelihood of becoming a HGF of all types but, most especially, a large turnover grower. The logged total value of exports is also positively associated with HGF status, such that firms which export more are more likely to be all types of HGFs. In a specification where we include the number of countries a firm exports to, as well as the total value of exports, the value of exports is positive and significant for all clusters, but the number of countries exported to becomes insignificant.

Similarly, the logged value of imports has a strong, positive relationship with the likelihood of being a HGF. We note that the effect of imports on the likelihood of growth is greater for large employment growers and small growers than the effect of exports, such that access to imports is a more important determinant of growth than the value of exports. For all clusters, the number of countries from which imports are sourced is associated with an increasing likelihood of achieving high growth. We also test a specification where we include the number of countries a firm imports from as well as the total value of its imports, and while the coefficients for both the number of countries imported to and the logged value of imports is smaller than when regressed separately, both variables still have a positive, significant relationship with all HGF clusters and the number of criteria fulfilled.

Table 10c: Trader status of HGF clusters

	Cluster 1	Cluster 2	Cluster 3	All HGFs
Importer/exporter indicator	0.0032*** (0.0006)	0.0167*** (0.0014)	0.022*** (0.0008)	0.08689*** (0.0037)
Number of countries exported to	0.0007*** (0.0002)	0.0026*** (0.0003)	0.0063*** (0.0003)	0.020*** (0.0012)
Log of total value of exports	0.0028*** (0.0003)	0.0090*** (0.0006)	0.0173*** (0.0004)	0.0623*** (0.0019)
Number of countries imported from	0.0017*** (0.0002)	0.0043*** (0.0003)	0.0089*** (0.0003)	0.0337*** (0.0014)
Log of total value of imports	0.0031*** (0.0002)	0.0150*** (0.0005)	0.0160*** (0.0004)	0.0715*** (0.0016)

Note: \*p>0.10. \*\*p>0.05. \*\*\*p>0.01.

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

### 6.3 Other firm characteristics

Being a subsidiary of a group or parent company is found to have no significant impact on the likelihood of being in clusters 1 or 2, while it has a significant, positive impact on being a large turnover grower. Being a subsidiary is therefore found to be associated with revenue growth more so than employment growth. For all HGFs in general, being a subsidiary has a positive impact on the number of HGF definitions that a given firm fulfils.

Next we consider the relationship between logged spending on R&D activities and each cluster, and we find that this has a strong, positive association with all clusters, and particularly so with large turnover growers. We note, however, that the lack of additional data after the change in the R&D allowance casts some doubt on any identified relationship between R&D expenditure and a given type of firm growth.

Table 10d: Other firm characteristics of HGF clusters

	Cluster 1	Cluster 2	Cluster 3	All HGFs
Subsidiary	0.0015 (0.0031)	-0.0003 (0.0049)	0.0494*** (0.0039)	0.1091*** (0.0188)
Log of R&D expenditure	0.0063*** (0.0015)	0.0081*** (0.0024)	0.0185*** (0.0018)	0.0833*** (0.0089)
Tax incentive	-0.0774*** (0.0276)	-0.1803** (0.0738)	-0.013 (0.0477)	-0.4961*** (0.1730)

Note: \*p>0.10. \*\*p>0.05. \*\*\*p>0.01.

Source: Author's calculations based on SARS-NT panel (SARS-NT 2016).

Whether a firm accesses a tax incentive is found to have a significant impact on the likelihood of experiencing some forms of growth. While the relationship is insignificant for the large turnover growers, there is a sizable, negative association between accessing tax incentives and being a large employment grower or a small grower. This effect is large for smaller HGFs: accessing an incentive reduces the likelihood of being a HGF as defined by these definitions by around 18 per cent. It also has a significant negative effect on the number of HGF definitions a given firm fulfils. An alternative explanation for this, which may be more plausible, is that firms which are less likely to grow for other reasons are also those which are more likely to seek assistance in the form of a tax

incentive, such that accessing a tax incentive is negatively correlated with becoming any type of HGF.

From these regression results we may begin to understand what the growth of the HGFs in each of these clusters is associated with. Large employment growers, which experienced lower real value added in the period under review but generated more jobs than turnover-growth HGFs, are more likely to emerge from the services sectors than from manufacturing, and while they are labour intensive, capital intensity does have a positive impact on the likelihood of being such a HGF. While more small growers are found in the services sectors than larger firms, they are more likely to emerge from manufacturing. A firm is more likely to become a small grower if it has higher profit margins and higher real value added, but it is less likely to do so if it makes use of tax incentives. The large turnover growers, which generate more real value added but less employment, are much more concentrated in wholesale and retail trade as well as in manufacturing. They are more capital-intensive than other firms and contribute more to total output. Being a subsidiary helps the chances of being a large turnover grower by around 5 per cent, and so does investment in R&D. And across the board, international trade makes a difference: firms that import and export more goods from more countries are more likely to be HGFs.

## 7 Concluding remarks

This study has found that the selection of HGFs is highly sensitive to the definition of HGFs used. The low correlation between the different HGF samples supports the evidence that different growth measures and indicators will yield different groups of firms. In turn, this results in considerable variation in the demographic characteristics of the samples derived from each definition.

While size is found to have a positive impact on the likelihood of becoming all types of HGFs, the impact of firm age is dependent on the growth indicator considered. The sector concentration of HGFs also depends on the definitions used: HGFs yielded from definitions which favour large, turnover-focused firms are heavily concentrated in manufacturing, while the HGFs identified by other definitions have a stronger primary and services sector representation. Other characteristics such as capital intensity also vary by definition. Across the board, however, trade activity matters: firms linked to international markets are more likely to be HGFs.

These findings have implications for public policy. Firstly, policy that seeks to foster the emergence of more HGFs must be cognisant of how it identifies such enterprises. As suggested by scholars such as Delmar et al. (2003), this may advocate for the use of different definitions of HGFs depending on what the objectives of a given policy are. It also cautions against the adoption of a single identifier of HGFs such as the OECD definitions, as they generate samples of firms that are not necessarily appropriate for all forms of policy intervention. More broadly, the effect of firm sector and trade activity on the likelihood of becoming a HGF suggests that policy makers may have scope to consider whether policy is conducive to the emergence of HGFs. If trade policy impedes the procurement of intermediate inputs by businesses, it may be stifling the ambitions of growth-focused firms. If industrial policy champions industries that are less likely to produce HGFs, or more likely to produce those which do not further objectives such as employment creation, then such policy is misaligned with supporting HGFs.

We note, however, that there are limitations to the study. The lack of geographical data means that we cannot control for a possible location effect or observe the effect of location on the emergence of HGFs. The data does not feature information regarding the acquisition of other firms, such

that total firm growth cannot be disaggregated into acquired and organic growth. Furthermore, the short time period for which data is available limits the study in two ways. First, the change in South Africa's economic climate after 2008 may have had an impact on this study's findings regarding the characteristics of firms that grow. Second, the study does not consider how HGFs perform beyond their growth period; for instance, it does not consider whether HGFs continue to grow or whether they are more likely to survive beyond their fast-growth period. Access to the administrative data for the 2003–08 growth years would allow for comparison with the findings of this study and mitigate the first problem, while also allowing to consider the performance of these HGFs after their fast-growth period and thereby mitigating the second limitation.

Despite these limitations, however, the study highlights that when attempting to define HGFs for the South African context, it is important to recognize that there is no single definition suitable for all discussions regarding firm growth. Rather, a more nuanced view of the type of firm growth under consideration is warranted. In this way, research and policy will be more focused in the discussion concerning the contribution of growth-focused firms to the economy.

## References

- Acs, Z. (1992). 'Small Business Economics: a Global Perspective.' *Challenge*, 35(2): 38–44.
- Acs, Z., W. Parsons, and S. Tracy (2008). 'High Impact Firms: Gazelles Revisited'. Office of Advocacy Working Paper. US Small Business Administration. Washington, DC: Corporate Research Board LLC.
- Almus, M. (2002). 'What Characterizes a Fast-growing Firm?'. *Applied Economics*, 34(12): 1497–508.
- Angrist, J.D., and J.S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- ANDE (Aspen Network of Development Enterprises) (2012). *Small and Growing Businesses: Investing in the Missing Middle for Poverty Alleviation*. Washington, DC: The Aspen Institute.
- Audrestch, D. (2012). 'Determinants of High-growth Entrepreneurship'. Report prepared for the OECD/DBA International Workshop on High-growth Firms: Local Policies and Local Determinants, Copenhagen, 28 March. Available at: [https://www.oecd.org/cfe/leed/Audrestsch\\_determinants%20of%20high-growth%20firms.pdf](https://www.oecd.org/cfe/leed/Audrestsch_determinants%20of%20high-growth%20firms.pdf) (accessed on 19 April 2017).
- Birch, D. (1979). 'The Job Generation Process'. *MIT Program on Neighbourhood and Regional Change*, 302: 1979.
- Birch, D. (1981). 'Who Creates Jobs?'. *Public Interest*, 65: 3–14.
- Birch, D. (1987). *Job Creation in America: How Our Smallest Companies Put the Most People to Work*. New York, NY: Free Press.
- Birch, D., and J. Medoff (1994). 'Gazelles'. In L.C. Solmon and A.R. Levenson, (eds), *Labour Markets, Employment Policy and Job Creation*. Boulder, CO: Westview Press.
- Bjuggren, C.M., S.O. Daunfeldt, and D. Johansson (2010). 'Ownership and High-growth Firms'. Ratio Working Paper, 147. Stockholm: The Ratio Institute.
- Bos J., and E. Stam (2014). 'Gazelles, Industry Growth and Structural Change'. Discussion Paper Series Nr.11-02. Utrecht: Tjalling C. Koopmans Research Institute.

- Brown, C., J. Hamilton, and J. Medoff (1990). *Employers Large and Small*. Cambridge, MA: Harvard University Press.
- Brown, R., C. Mason, and S. Mawson (2014). 'Increasing the Vital 6%: Designing Effective Public Policy to Support High-growth Firms'. Working Paper No. 14/01. London: NESTA.
- Brüderl, J., and P. Preisendörfer (2000). 'Fast-growing Businesses: Empirical Evidence from a German Study'. *International Journal of Sociology*, 30(3): 45–70.
- Coad, A., S.O. Daunfeldt, W. Hözl, D. Johansson, and P. Nightingale (2014). 'Introduction to the Special Section'. *Industrial and Corporate Change*, 23(1): 91–112.
- Daunfeldt, S.O., N. Elert, and D. Johansson (2014). 'The Economic Contribution of High-growth Firms: Do Policy Implications Depend on the Choice of Growth Indicator?'. *Journal of Industry, Competition and Trade*, 14(3): 337–65.
- Davis, S., J. Haltiwanger, and S. Schuh (1996). *Job Creation and Destruction*. Cambridge, MA: MIT Press.
- Delmar, F. (1997). 'Measuring Growth: Methodological Considerations and Empirical Results'. In R. Donckels and A. Miettinen (eds), *Entrepreneurship and SME Research: On its Way to the Next Millennium*. Aldershot: Ashgate.
- Delmar, F., and P. Davidsson (1998). 'A Taxonomy of High-growth Firms'. In P.D. Reynolds, W.D. Bygrave, M. Carter, S. Manigart, C.M. Mason, G.D. Mayer, and K.G. Shaver, (eds), *Frontiers of Entrepreneurship Research*. Wellesley, MA: Babson College.
- Delmar, F., P. Davidsson, and W. Gartner (2003). 'Arriving at the High-growth Firm'. *Journal of Business Venturing*, 18(2003):189–216.
- Eurostat and OECD (2007). *Eurostat-OECD Manual on Business Demography Statistics*. Paris: OECD Publishing.
- Fafchamps, M., and C. Woodruff (2014). 'Identifying Gazelles: Expert Panel vs. Surveys as a Means to Identify Firms with Rapid Growth Potential'. Policy Research Working Paper; No. WPS 7647. Washington, DC: World Bank Group.
- Goedhuys, M., and L. Sleuwaegen (2010). 'High-growth Entrepreneurial Firms in Africa: a Quantile Regression Approach'. *Small Business Economics*, 34: 31–51.
- Henrekson, M., and D. Johansson (2010). 'Gazelles as Job Creators: A Survey and Interpretation of the Evidence'. *Small Business Economics*, 35(2): 227–44.
- Kerr, A., M. Wittenberg, and J. Arrow (2013). 'Job Creation and Destruction in South Africa'. Southern Africa Labour and Development Research Unit Working Paper Number 92. Cape Town: Southern Africa Labour and Development Research Unit.
- Li, Y., and M. Rama (2013). 'Firm Dynamics, Productivity Growth and Job Creation in Developing Countries: The Role of Micro- and Small Enterprises'. Background Paper for the World Development Report. Washington, DC: World Bank.
- Lilischkis S. (2011). 'Policies in Support of High-growth Innovative Smes'. INNO-Grips Policy Brief No.2. Bonn: INNO-Grips.
- Madhulatha, T. (2012). 'An Overview of Clustering Methods'. *IOSR Journal of Engineering*, 2(4): 719–25.
- Mitusch, K., and A. Schimke (2011). 'Gazelles: High-Growth Companies'. Task 4 Horizontal Report 5 (Final Report). Europe INNOVA Sectoral Innovation Watch for DG Enterprise and Industry. Brussels: European Commission.

- Petersen, D., and N. Ahmad (2007). *High Growth Enterprises and Gazelles—Preliminary and Summary Sensitivity Analysis*. Paris: OECD-FOR A.
- Pieterse, D., F. Kreuser, and E. Gavin (2016). 'Introduction to the South African Revenue Service and National Treasury Firm-level Panel'. WIDER Working Paper 2016/42. Helsinki: UNU-WIDER.
- SARS-NT (South African Revenue Service and National Treasury) (2016). 'Tax Statistics, July 2016'. Available at: <http://www.sars.gov.za/AllDocs/Documents/Tax%20Stats/Tax%20Stats%202016/Tax%20Stats%202016%20Full%20document%20web.pdf> (accessed on 15 July 2016).
- Schreyer, P. (2000). 'High-growth Firms and Employment'. OECD Science, Technology and Industry Working Papers, 2000/03. Paris: OECD Publishing.
- Shane, S. (2009). 'Why Encouraging More People to Become Entrepreneurs is Bad Public Policy'. *Small Business Economics*, 33: 141–49.
- Storey, D. (1994). *Understanding the Small Business Sector*. Abingdon: Routledge.