Understanding the sharp primary level enrolment increases beginning in 2011

MARTIN GUSTAFSSON

Stellenbosch Economic Working Papers: WP08/2018

www.ekon.sun.ac.za/wpapers/2018/w082018

May 2018

KEYWORDS: Education demographics, South Africa, births
JEL: I21, J11, O15

ReSEP (Research on Socio-Economic Policy)
http://resep.sun.ac.za

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
SOUTH AFRICA

UNIVERSITEIT STELLENBOSCH UNIVERSITY

BER BUREAU FOR ECONOMIC RESEARCH

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

www.ekon.sun.ac.za/wpapers

Understanding the sharp primary level enrolment increases beginning in 2011

MARTIN GUSTAFSSON

November 2017

ABSTRACT

Enrolments at the primary level in South Africa increased sharply from around 2011. Over the six-year period 2009 to 2015, grade 1 enrolments increased by 13%. These increases were not expected, and came after many years of enrolment decline. The current paper concludes that the enrolment increases were due to population increases. They were not caused by fraudulent over-reporting or increases in grade repetition. They were clearly the outcome of a remarkable increase of around 13% in births, in particular during the years 2003 to 2005. This is confirmed by birth registrations data. After 2008, births declined somewhat and settled at a level which was around 6% lower than the 2005 to 2008 'plateau'. However, this decline was not large enough to take birth numbers back to their pre-2003 levels. A brief discussion of the aggregate statistics relating to the child support grant and anti-retroviral treatment, and of some available research on causation, leads to the conclusion that it is not easy to explain the increase in births, though the available evidence leans towards anti-retroviral treatment, rather than child support grants, as the most likely explanation. Further analysis of microdata may bring more certainty in future. The paper argues that better use could be made of the available data, all of which have problems, but which, when analysed together, can produce more reliable scenarios around future enrolments. Such scenarios are obviously vital for effective education planning.

Martin Gustafsson
Department of Economics
University of Stellenbosch
Private bag X1, 7602
Matieland, South Africa
E-mail: mgustafsson@sun.ac.za

The current report was produced as part of the author's work for the Department of Basic Education. A slightly amended version was included in the Department's first issue of its new journal *Basic Education Matters*, in 2018. Thanks for insights provided go to colleagues and peers at the Department of Basic Education (in particular Bheki Mpanza, Stephen Taylor, Johannes Malebye and John Ostrowick), ReSEP (Servaas van der Berg, Rulof Burger), and Statistics South Africa.

1 Introduction

This report is prompted by the need to understand large enrolment increases that have occurred in the ordinary schooling sector, starting in grade 1 in 2011, and continuing through to grade 7 in 2017. These increases need to be understood if effective responses are to be formulated. Section 2 outlines the grade-specific increases since 2011. Section 3 discusses briefly the possible factors that could lie behind the increases. Section 4 comprises the bulk of the report and uses a variety of data sources to investigate what appears to be easily the strongest contributing factor, namely increases in the sizes of birth cohorts, in particular since around 2003. Section 5 looks briefly at possible links between the expansion of the child support grant system, increased access to anti-retroviral treatment, and increases in births. Section 6 presents a couple of scenarios describing how changing birth cohort sizes are likely to impact on total school enrolments in the coming decades.

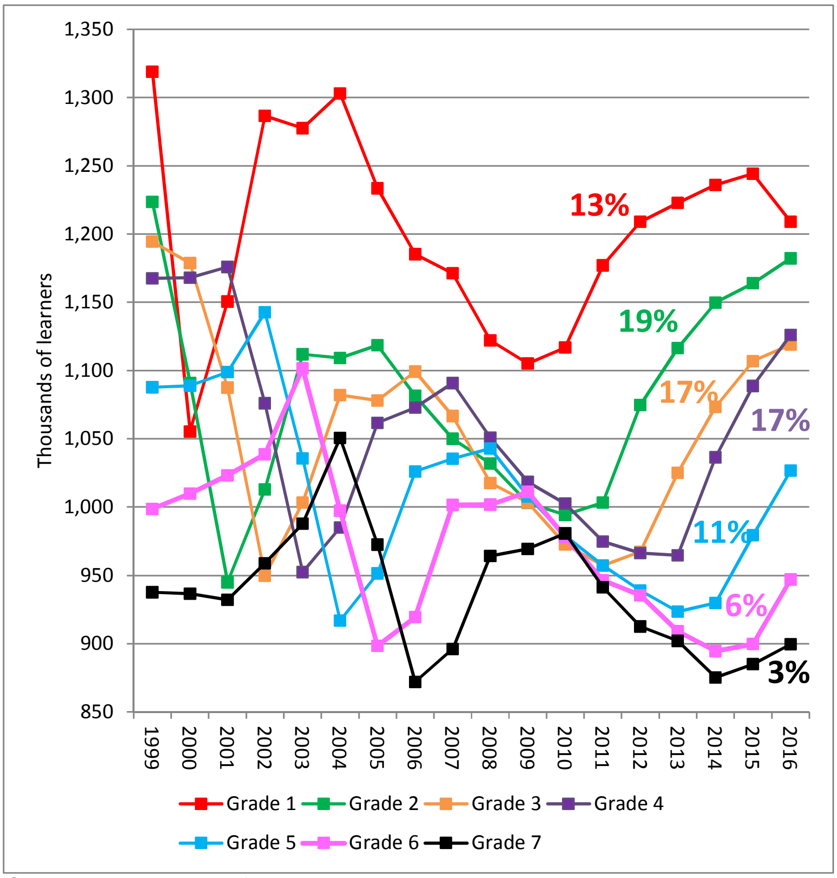
A key contribution of this report is to bring together several data sources, some of which have only become available recently, within one analysis. When viewed in isolation, each data source may not reveal much, but when viewed together with other sources, rather firm conclusions can be drawn. The report thus serves as a reminder of the importance of viewing a variety of data sources, wherever possible, when drawing conclusions around trends in the South African schooling system. The report moreover emphasises the importance of gauging the quality of the data, at the level of the microdata, using various methods.

2 Recent increases in primary level enrolments

Figure 1 below illustrates the grade-specific enrolment trends within the ordinary schooling system (public plus independent), for grades 1 to 7, as reflected in official Department of Basic Education (DBE) enrolment reports covering the period 1999 to 2016. One striking inconsistency is the low grade 1 enrolment level in 2000, but apart from this the trends appear consistent. In particular, increases and declines in one grade in one year are reflected in the next year and next grade. The focus of this report is the sharp increase seen in grade 1, in particular over the 2010 to 2012 period. Over the longer period 2009 to 2015, grade 1 enrolments increased by 13%. The subsequent decline between 2015 and 2016 is significant and is discussed below. If one considers only public schools, the 13% figure becomes 10.6%, pointing to even more rapid increases in the independent school sector than the public school sector. Over the period 2009 to 2015, the percentage of grade 1 learners in independent schools rose from 3.0% to 4.7%.

What is clear is that the post-2009 increases follow a trough in grade 1 enrolments in 2009 and that 2004 to 2009 saw very large declines in enrolment. These declines are understood to have been caused by less grade repetition, combined with an increase in the availability of grade R in schools, which reduced the 'baby-sitting' phenomenon in grade 1 whereby under-aged children were imposed on grade 1 classes by parents requiring child care. But the pre-2009 declines at the grade 1 level are not the subject of this report and as will be seen, the post-2009 increases relate to trends which are very different to the trends that are believed to underlie the earlier declines.

Figure 1: Official grade-specific enrolment totals 1999 to 2016



Source: Department of Basic Education, 2016a.
Note: Percentages refer to growth along the entire period of growth.

It is clear in Figure 1 that after 2009, the general pattern was that a large enrolment increase in one grade would be followed by a large increase in the next grade in the next year. For instance, the largest increase in grade 1 was in 2011, the largest increase in grade 2 was in 2012, and the largest increase in grade 3 was in 2013. Given this pattern, it should not come as a surprise that grade 6 saw a large increase in 2016. Enrolment figures for 2017 are not yet finalised, but one would expect a large increase in grade 7 in this year. Very importantly, the pattern suggests that large enrolment increases should be felt in secondary schools, specifically grade 8, starting in 2018.

As shown by the following table, which focuses on the grade 2 expansion years, there have been large increases in all provinces. Figures in Table 1 are from schools that could be found in all the three years in order to achieve greater certainty that trends were not being driven by schools missing from the data. Comparing the previous graph to the national figures in Table 1 suggests that missing schools is a small problem. The national figures in the graph are only around 1% higher than the national figures in the table. Whilst large increases are seen in all provinces, Limpopo (LP) has seen particularly large increases, whilst those in Eastern Cape (EC) and Northern Cape (NC) are well below the national average.

Table 1: Provincial increases in Grade 2 enrolment 2011-2013

	2011	2012	2013	% increase 2011-2013
EC	169,720	175,023	179,406	6
FS	55,279	59,476	62,396	13
GP	172,915	184,175	192,322	11
KN	217,270	235,377	239,513	10
LP	118,441	131,914	138,126	17
MP	81,059	88,594	92,047	14
NC	23,753	24,309	25,292	6
NW	68,246	73,315	74,061	9
WC	88,432	91,526	95,795	8
SA	995,115	1,063,709	1,098,958	10

Source: Department of Basic Education, 2017.

A similar comparison across years using a breakdown by the official poverty quintile of schools indicates that the smallest increases have been in the least poor (quintile 5) and the most poor (quintile 1) school communities. Here again, only schools with data in all three years were counted. The reason why national totals are a bit lower in Table 2 compared to Table 1 is that the data on quintile were in certain cases missing.

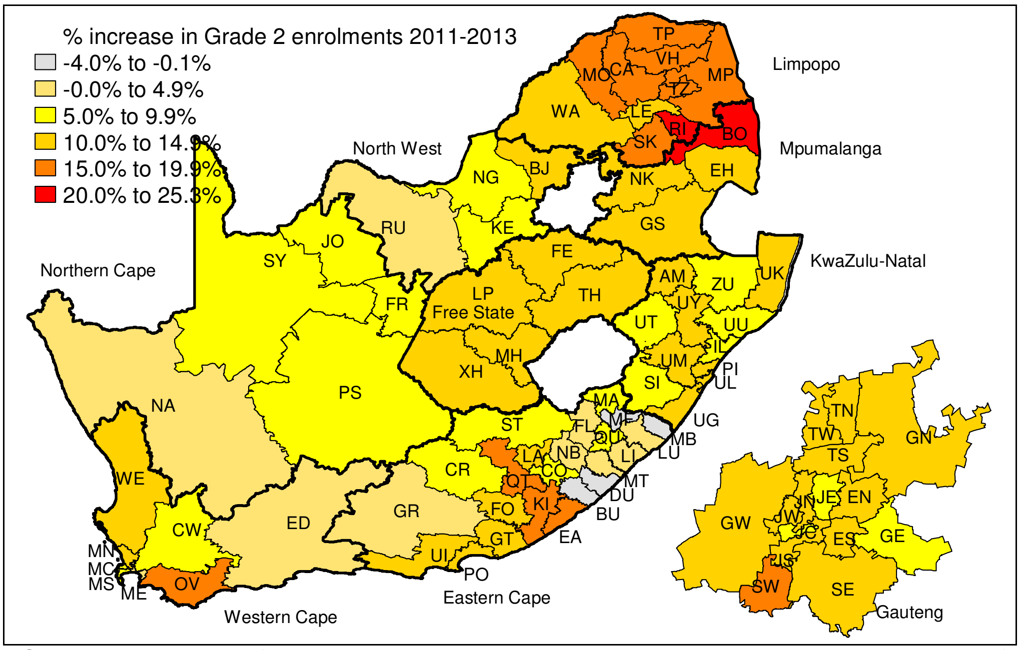
Table 2: Increases in Grade 2 enrolment 2011-2013 by quintile

Quintile	2011	2012	2013	% increase 2011-2013
1	288,655	307,960	314,961	9
2	201,607	217,594	225,075	12
3	255,368	274,672	285,739	12
4	109,790	116,712	120,418	10
5	104,022	108,362	111,912	8
Total	959,442	1,025,300	1,058,105	10

Source: Department of Basic Education, 2017.

The map in Figure 2 breaks the 2011 to 2013 grade 2 increases down by education district. What seems noteworthy is that in the provinces with the highest increases, Limpopo, Mpumalanga and Free State, all districts experienced increases of at least 10%.

Figure 2: Distribution of 2011-2013 Grade 2 enrolment increases



Source: Department of Basic Education, 2017. District codes are explained in Department of Basic Education, 2016b.

3 Possible explanations for the enrolment increases

What are the possible explanations for the post-2010 enrolment increases discussed in the previous section? (Though there was a relatively small increase between 2009 and 2010 in grade 1, the really large increases began between 2010 and 2011. Hence the title of the report refers to large increases beginning in 2011.)

The inflow into grade 1 could have been increasing. This could occur due to three factors. Birth cohorts entering grade 1 in a year could be larger than the preceding birth cohorts. The proportion of any birth cohort enrolled in school could be increasing. Or something could have prompted younger children to enter grade 1, so that more than a single age cohort were entering the grade at once.

The outflow from grade 1 could also be declining. The likely reason for this would be that grade repetition in grade 1 was increasing. This is sometimes referred to as 'churning'.

The possibility that the enrolment ratios of age-correct children were increasing to a large extent can be discounted because these ratios have been high, for instance around 98% for seven-year-olds since 2004[1]. The mix of ages within each grade and grade repetition as possible factors could have received more attention below, but that analysis was not done as it became clear that increases in the sizes of birth cohorts could easily account for the grade-specific enrolment increases after 2009 seen in Figure 1.

4 Clear evidence of increases in the birth cohorts

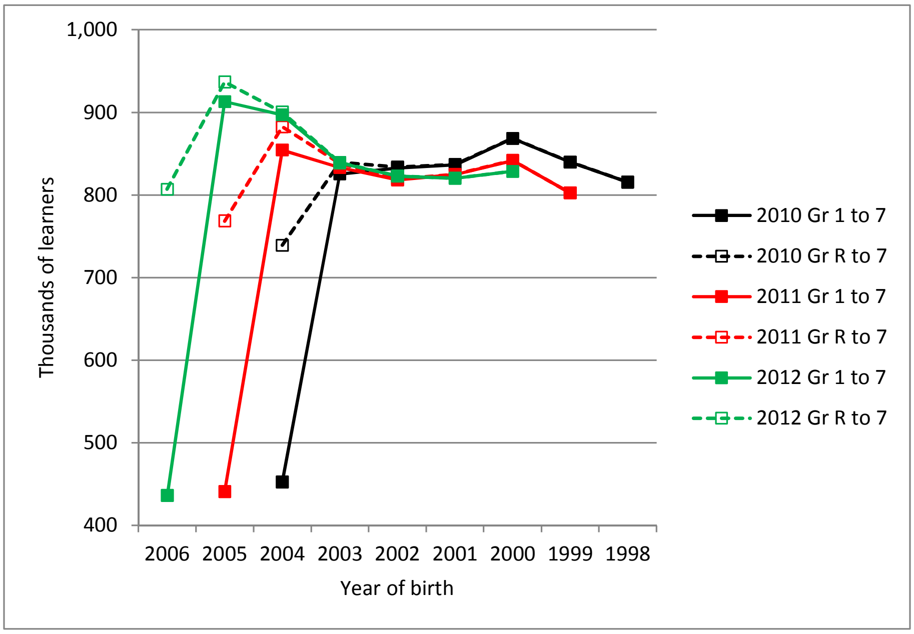
4.1 Evidence from the school data

Figure 3 below uses Annual Survey of Schools (ASS) data, which, since 2010, has required respondents, namely school principals, to break enrolment up year of birth. Prior to 2010,

[1] Department of Basic Education (2013b) and Department of Basic Education (2013c).

respondents broke enrolments up by age on the date of the survey, which was not optimal, as delays in the survey could then affect the statistics. The patterns seen in the 2012 ASS data differ considerably from those seen in the 2010 ASS. The major change is that the sizes of the largest birth cohorts have changed substantially. In particular, the number of children born in 2004 is 7% higher than the number of children born in 2003, and the difference between the 2005 and 2003 birth cohorts is 11.6%. In the analysis underlying Figure 3 only schools present in all three school years were used. If all schools were used, regardless of links across school years, the 11.6% figure would be 11.9%. These figures are both derived from the 2012 ASS curve including grade R. The fact that for many of the earlier birth cohorts, for instance children born in 2003, the levels against the vertical axis should be so similar for the three ASS surveys used, seems to confirm the reliability of the data. The drops at the left-hand end of all curves can be explained by the fact that many younger children have not entered the schooling system yet. Moreover, it should kept in mind that at the grade R level, many children are enrolled not in ordinary schools but in separate pre-schools and early childhood development (ECD) centres and would thus not be reflected in Figure 3.

Figure 3: Annual Survey of Schools age distributions



Source: Annual Survey of Schools dataset of the DBE.
Note: Each curve covers learners who would have turned between 6 and 12 during the year of the survey. For instance, those turning 6 years during the 2012 survey would have been born in 2006. In the Annual Survey of Schools forms for the years 2010 to 2012, schools were asked to provide numbers of learners according to year of birth. Before 2010, schools had been asked to provide number of learners by age on the day of the survey.

The relative stability in the sizes of age cohorts born between 1998 and 2003 seen in the above graph supports the argument made previously that the declines in, say, grades 1 to 4 enrolment levels in the years prior to 2009 were driven not so much by changes in the age structure of the population, but by factors internal to the schooling system, in particular declining grade repetition.

The 12% national increase between the 2003 and 2005 birth cohorts is broken down by province in the next table. Again, Limpopo emerges as a province which experienced exceptional enrolment growth. Clearly there is room to interrogate the data further than what is presented here to clarify apparent anomalies such as the increase the Eastern Cape birth cohorts being relatively high, at 14% (Table 3), whilst the grade 2 enrolment increase is just

6% (Table 1). There are several plausible explanations for patterns such as these, for instance declining grade repetition in Eastern Cape.

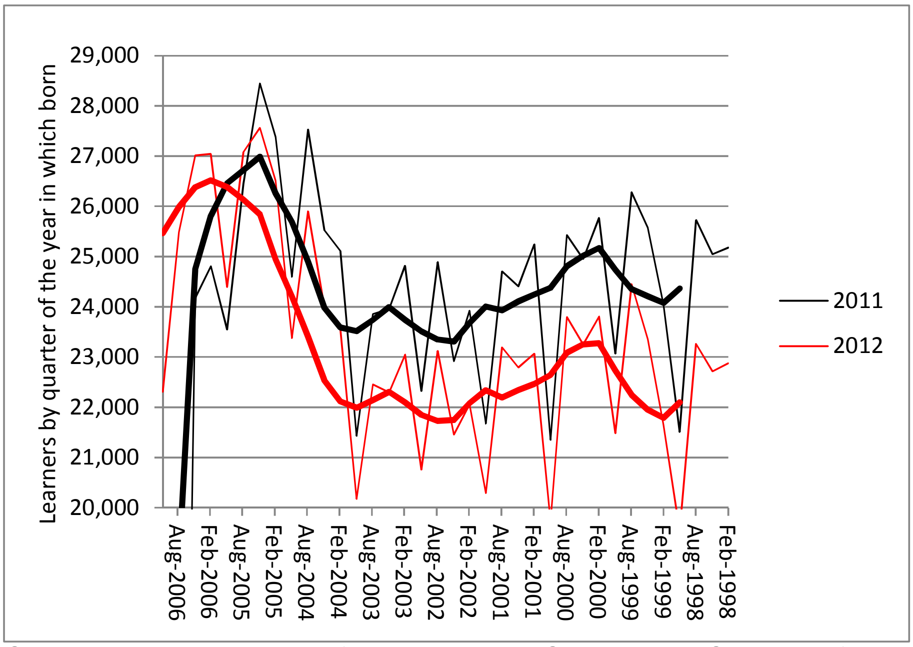
Table 3: Provincial details on 2003 to 2005 birth cohort increases in ASS

	% increase 2003 to 2005	Absolute increase 2003 to 2005	% of national increase
EC	14	18,421	19
FS	11	5,299	5
GP	9	13,347	14
KN	12	21,530	22
LP	19	19,921	20
MP	10	6,992	7
NC	3	553	1
NW	10	5,642	6
WC	7	5,829	6
SA	12	97,534	100

Source: Annual Survey of Schools dataset of the DBE.

The DBE’s Learner Unit Record Information Tracking System (LURITS) offers an opportunity to verify the ASS trends. LURITS was established largely to make fraudulent over-reporting by school principals very difficult, as LURITS requires the entry of an array of personal details on each learner. The format in which the LURITS data were made available would have made a national analysis cumbersome. A year’s data for each province is spread across around ten files. Moreover, the data are reportedly better for certain provinces than others. Limpopo’s data were explored as this was said to be one of the better provinces. The LURITS data do in fact confirm the patterns seen in the ASS. The curves in Figure 4 below are based on data where around 10% of learners appeared to be missing in 2011, and 16% in 2012. The LURITS data include independent schools. The more recent 2012 data point to the 2005 birth cohort being around 20% larger than the 2003 birth cohort. This is close to the 19% increase for LP seen in the first column of Table 3.

Figure 4: Limpopo 2011-2012 age distributions



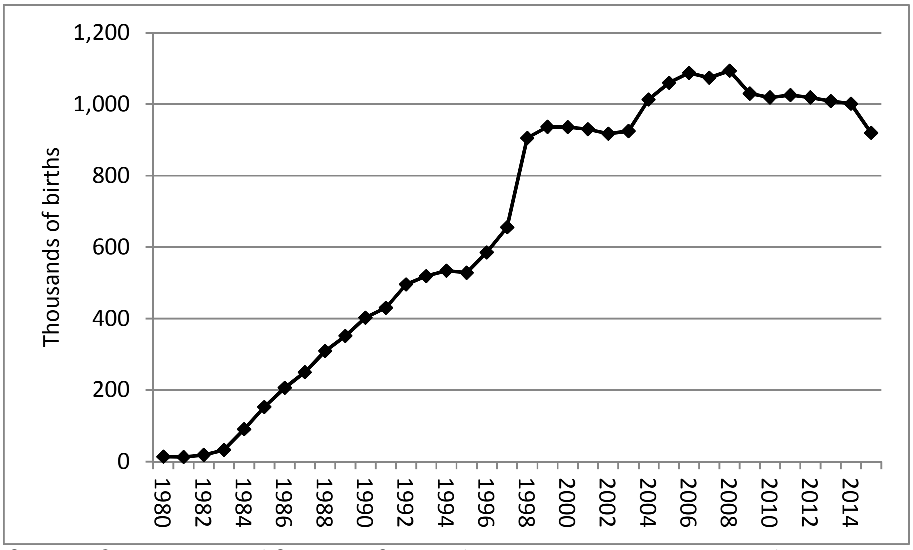
Source: Learner Unit Record Information Tracking System (LURITS) dataset of the DBE.
Note: Thin lines illustrate the trend for quarters, where for instance all dates of birth in the range 1 January 2004 to 31 March 2004 would be rounded to 15 February 2004. The thick curves are moving average curves using four values (i.e. four quarters) for each average. The fact that the 2012 levels should be so much lower than the corresponding 2011 levels is due to the fact that there are more missing data in 2012 than 2011.

Given the exceptionally large birth cohort increases in Limpopo, patterns with respect to surnames in LURITS were examined to see whether younger learners displayed different surnames to older learners, something which could confirm immigration from neighbouring Zimbabwe and Mozambique as contributing factors. However, this analysis revealed no interesting trends. For instance, if one ranked surnames by the number of learners in each birth cohort, no systematic change in the rankings over time emerged.

4.2 Evidence from Home Affairs birth registration data

The Home Affairs birth registrations data for 1998 to 2015, made available by Stats SA through the University of Cape Town's DataFirst portal, offers an opportunity for verifying the trends seen in the ASS and LURITS data of the DBE. However, the Home Affairs data must be examined with much caution. Though the title of the dataset indicates that data go back to 1998, in fact they go back to 1880, though for early years there is clearly an under-counting of births. This can be seen in Figure 5. The title of the dataset ('1998 to 2014') seems to suggest that from 1998 the data can be taken seriously. What is particularly telling is that the 2003 to 2005 birth cohort increases seen in the school data are mirrored in the Home Affairs data, at least if one considers the country as a whole. The curve in Figure 5 increases by 14.5% between 2003 and 2005 (this is close to the 12% seen for the country in Table 3).

Figure 5: Birth registrations data 1980-2015

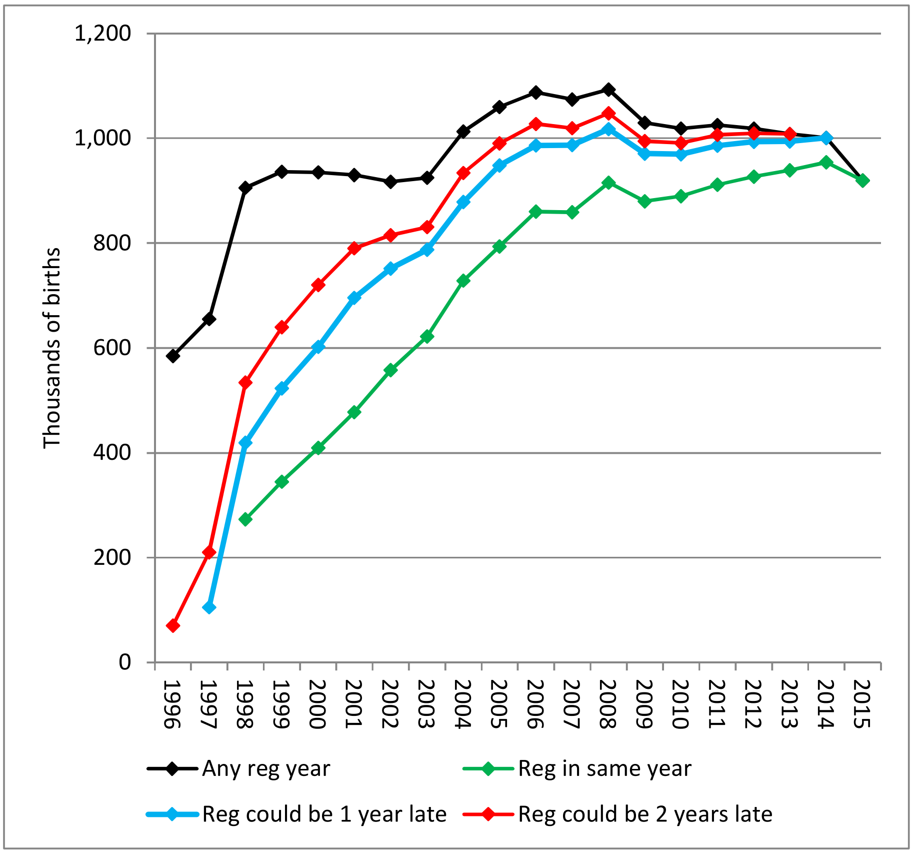


Source: Own analysis of Statistics South Africa, 2015. Despite the title of this dataset, it contains details of births up to 2015, not 2014.

A key question from an education planning perspective is what the trend is with respect to birth cohorts beyond 2005. Figure 5 suggests that after 2008, births declined fairly substantially, and stayed at a lower level, though a level that was still higher than the pre-2003 level. Specifically, annual births for the years 2009 to 2013 are 6% lower than for the years 2006 to 2008. This suggests that grade 1 in schools could begin to experience somewhat smaller incoming birth cohorts from 2016 (most learners turn seven in grade 1). In fact, a decline in grade 1 enrolments in 2016 is precisely what we see in Figure 1. Thus with respect not just to the earlier increases, but also with respect to subsequent decreases, the schools data mirror the birth registrations data. Unfortunately, ASS data, which would be age-specific, for years after 2013 is not ready for analysis. For the period of declines, it is thus necessary to compare birth registrations to enrolment by grade (and not age).

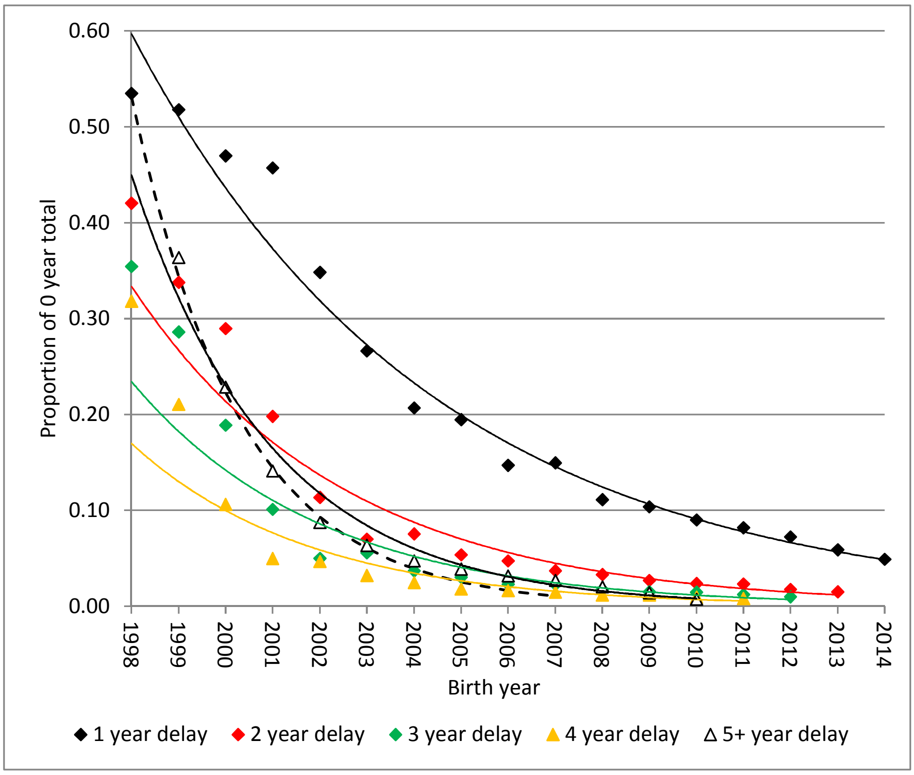
One aspect of the birth registrations data warrants further investigation. Many births are registered not in the year of birth, but in a later year. There is thus delayed registration by the parents or caregivers. This probably explains the 2014 to 2015 dip in Figure 5. This dip is more likely to be a reflection of the fact that some children born in 2015 have not been registered yet, as opposed to an actual decline in births in 2015. The next two graphs illustrate patterns in this regard. In Figure 6, 'Reg in the same year' reflects birth registered in the year that the birth occurred, 'Reg could be 1 year late' would be the previous curve plus births registered with a delay of just one year, and so on. The curve 'Any reg year' is the same as the curve in Figure 5 above. The data used for Figure 6 were used to generate Figure 7. To illustrate, the '1 year delay' points are calculated as follows. Each point is the number of births occurring in, say, year Y but registered in year $Y + 1$, divided by the number of births occurring in year Y and being registered in year Y. The lower this ratio, the faster the system would be at getting births registered. Similar series of points were plotted for a two-year delay, a three-year delay, and so on. The fact that all ratios have been declining points to faster registrations. The trendlines in Figure 6 are exponential regression lines for the five series of points. These regression lines are used to predict what the ratios are in all years up to 2015. This then allows one to estimate, say, how many births occurring in 2015 are likely to be registered one year late (so in 2016), two years late (so in 2017), and so on. This in turn allows one to estimate births to 2015, including births which are registered late[2].

Figure 6: Birth registrations delays (I)



[2] The formula for one year's births would be $B = B_0 + r_1 B_0 + r_2 B_0 + r_3 B_0 + r_4 B_0 + r_5 B_0$, where B is total births in a year, and B_0 is births where the year of birth and the year of registration is the same. The ratio r_3 would be births with a registration delay of 3 years, divided by B_0, and so on for the other subscripts to r. Only for the most recent years, where the relevant data do not exist, would r be estimated by a regression. For instance, this would occur for r_3 when dealing with births in 2013. Because the data only go as far as 2015, it would not be possible to determine from the data r where registration was delayed by three years, as registration would then occur in 2016 (three years after 2013). In such cases, the coefficients a and b in the following regression would be estimated for a particular series of r, for instance r_3: $\ln(r) = a + bY$. Here Y is year of birth. Thereafter, r_3 would estimated as $r_3 = e^{(a+bY)}$, where e is 2.718 (the base of a natural logarithm).

Figure 7: Birth registrations delays (II)



The outcome of the analysis described above is the 'Adjusted' curve in Figure 8. 'Original' in Figure 8 a repeat of the curve from Figure 5. The adjusted figures confirm that after 2008 there was a real decline in births, but that between 2009 and 2012 births per year remained roughly stable. The fact that the 2014 to 2015 dip remains could be a reflection of delays in the *capturing* of data, as opposed to delays in the registrations process. It is possible that some Home Affairs registrations which occurred in 2015, reflecting 2015 births, were not fed through to the database. But this is of course conjecture. Yet without more years of data, it seems safest to dismiss the 2014 to 2015 dip as a data issue.

Figure 8: Birth registrations with adjustments for delays

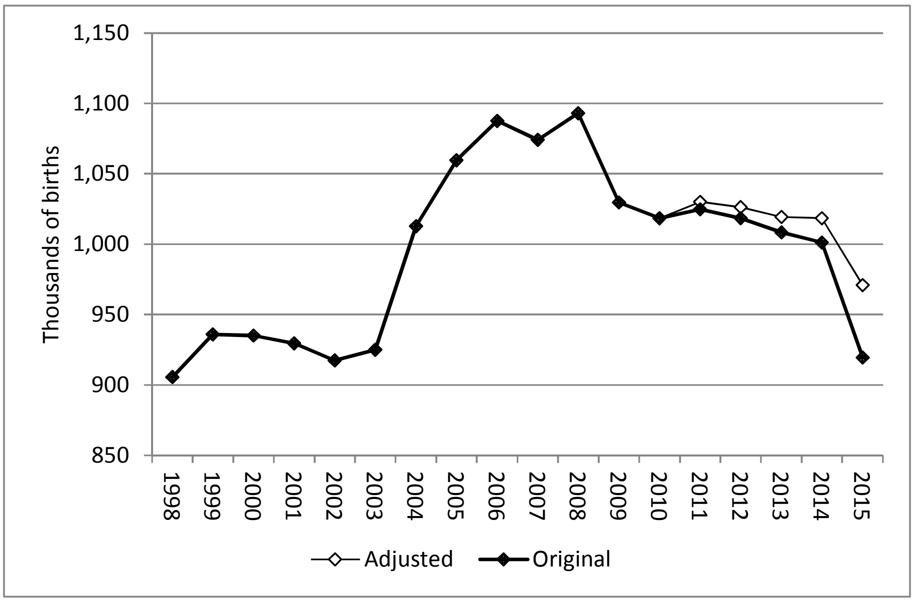


Table 4 below provides a breakdown by province of the 2003 to 2005 increase of 14.5% in the birth cohort sizes mentioned above. The metadata indicate that 2011 provincial boundaries were used for all years, so changes to these boundaries legislated in 2005 would not influence the provincial patterns. A crucial comparison is one between Table 4 and Table 3 above. One might expect a high degree of correspondence between the figures in the two tables. Limpopo still emerges as a province with strong growth. However, four provinces, Gauteng, Mpumalanga, North West and Western Cape, have considerably higher figures when the Home Affairs data are used compared to the ASS data. In the case of Mpumalanga and North West, this could be because children migrated from those provinces to, say, Gauteng between around 2004 and when these children were ready to enter school. However, possible explanations for the Gauteng and Western Cape discrepancies are not as obvious. Further analysis of both data sources could throw more light on this matter.

Table 4: Provincial details on 2003 to 2005 birth registrations increase

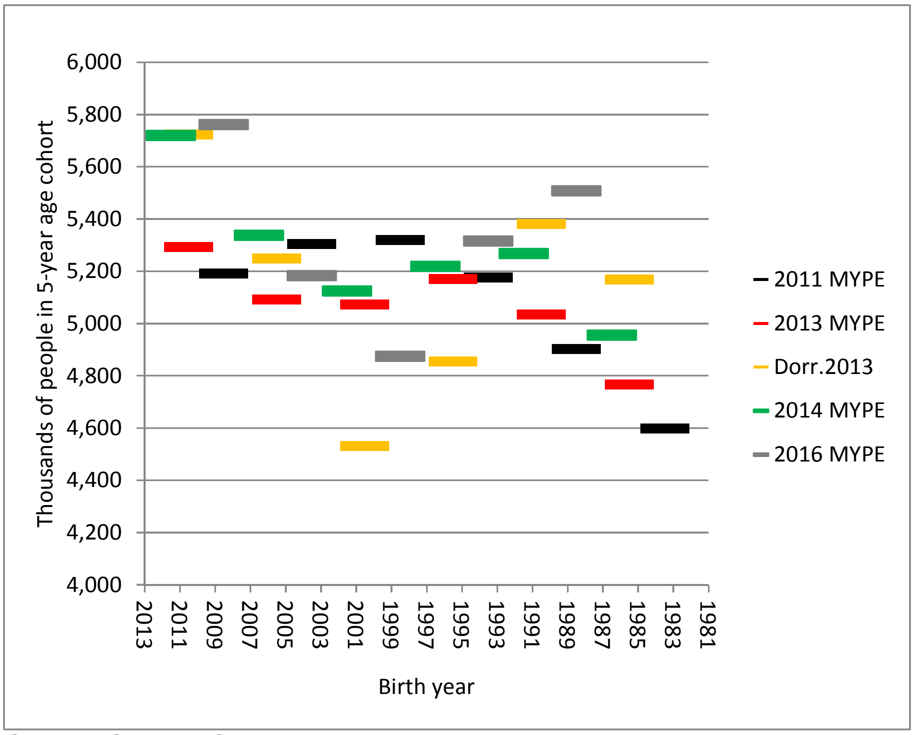
	% increase 2003 to 2005	Absolute increase 2003 to 2005	% of national increase
EC	13	17,167	13
FS	12	6,002	4
GP	15	25,491	19
KN	9	19,234	14
LP	23	21,989	16
MP	16	11,906	9
NC	9	2,047	2
NW	20	13,545	10
WC	19	16,783	13
SA	15	134,164	100

Source: Own analysis of Statistics South Africa, 2015.

4.3 Official population figures

The official mid-year population estimates (MYPE) of Stats SA should guide planning in the education sector. However, these estimates have not reflected the kinds of growth in the birth cohorts discussed above, at least not until the 2014 MYPE. Figure 9 below illustrates the sizes of the five-year age cohorts reported in the 2011, 2013, 2014 and 2016 mid-year population estimates. Alternative mid-year population estimates for 2013 produced by Dorrington (2013), who has disputed the estimates of Stats SA, are also reflected in the graph. The 2014 estimates indicate that there would be 4% more children alive in 2014 born in the 2004-2009 period than in the 1999-2004 period. The corresponding figure obtained if one uses the Home Affairs trends seen in Figure 8 is 12% (of course the figures are not fully comparable mainly because the Home Affairs figures are birth registrations, without any accounting for post-registration deaths, but roughly the figures are comparable). The 2014 MYPE thus gets closer to reflecting the increases seen the Home Affairs and enrolment data than previous MYPEs, yet the 2014 MYPE still seems to under-estimate the increase. On the other hand, the 2014 MYPE does point to a very large increase in post-2009 births. Specifically, children alive in 2014 born in the 2009-2014 period are 7% more numerous than those born in the preceding 2004-2009 period. This is very different to what is seen if one examines the Home Affairs birth registrations data (with adjustments for late registrations), which point to a *decline* of around 4% across these periods.

Figure 9: Mid-year population estimates (5-year age cohorts)

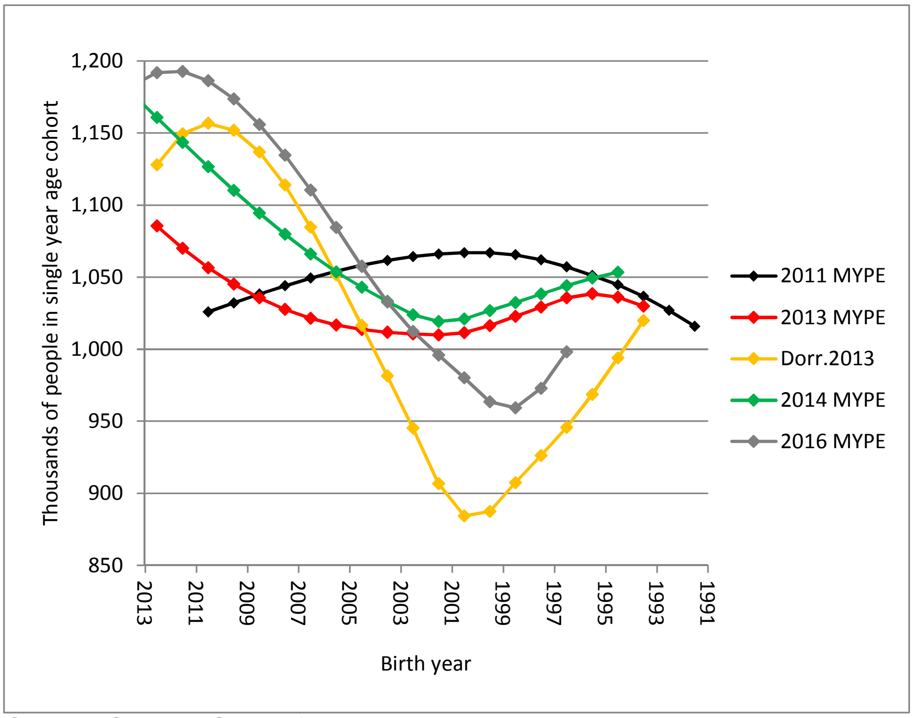


Sources: Statistics South Africa, 2013; Dorrington, 2013.
Note: Each horizontal bar in the graph is positioned above the point on the horizontal axis representing the mid-point of the five-year age group in question.

One can 'backward engineer' the process of arriving at five-year population groups in order to recreate the single-year birth cohorts, through the use of Sprague coefficients[3]. This was done to produce the estimates reflected in the following graph. None of the curves seen below are really consistent with the curves seen in the Home Affairs data (Figure 8) or the ASS data (Figure 3). In one respect, the 2016 MYPE curve is more consistent with the enrolment trends than the 2014 MYPE curve. The former curve is twice as steep as the latter curve, producing a 5% increase between 2003 and 2005, compared to the 2% increase one obtains from using the 2014 MYPE curve. Of course both of these figures are well below the 12% of the ASS data or the 15% of the Home Affairs data.

[3] Châu, 2003.

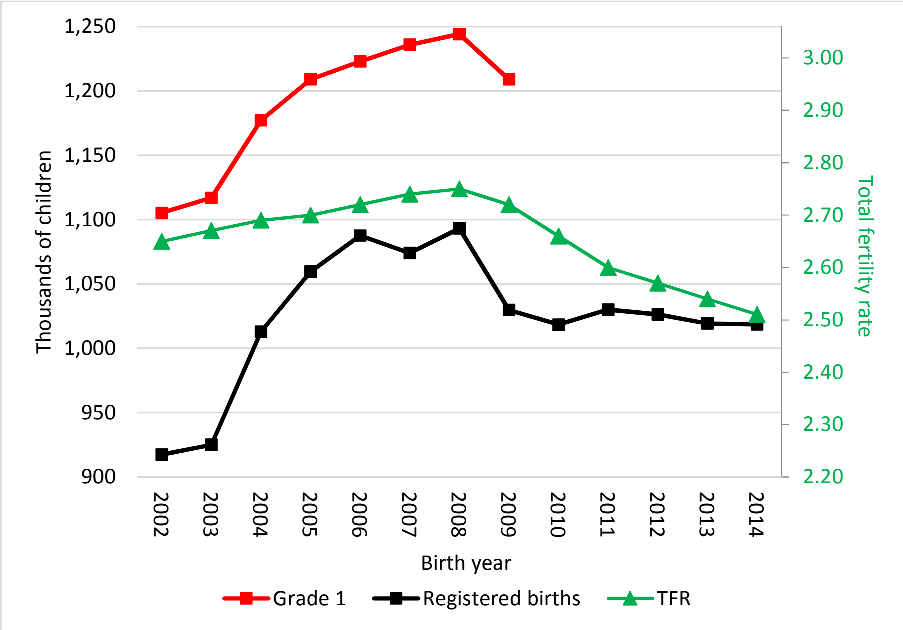
Figure 10: Mid-year population estimates with Sprague smoothing



Sources: Statistics South Africa, 2013; Dorrington, 2013.

If there is a trend in the recent MYPE publications which mirrors fairly closely the rise and decline in births seen in the Home Affairs data, and suggested by the enrolment data, it is the total fertility rate (TFR) values published in the 2016 MYPE. This is illustrated below. The TFR values do in fact trace the increase followed by a decline seen in the other two sources. In all three curves, the peak is in 2008. Here the 2008 enrolment value reflects grade 1 enrolment not in 2008, but seven years later, in 2015, using the assumption that children tend to turn seven when they are in grade 1. However, the TFR rise and fall is smaller than in the other two data sources. For instance, whilst the grade 1 curve in Figure 11 represents an increase of 13%, TFR rises by just 4% over this period.

Figure 11: Enrolments, births and fertility

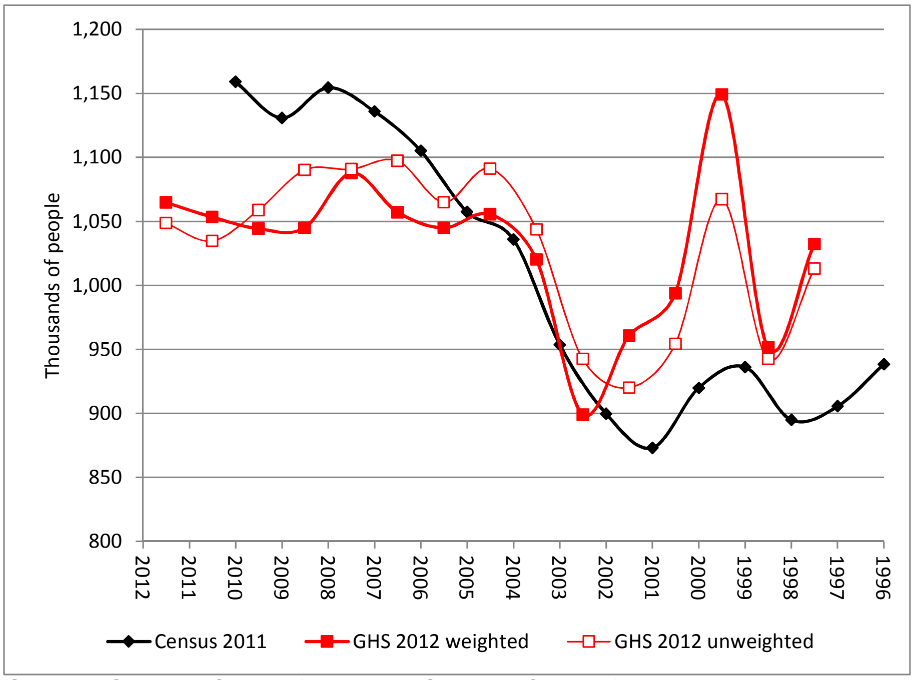


Sources: Enrolments as for Figure 1, registered births as for Figure 8, total fertility rate (TFR) from Statistics South Africa (2016: 4).
Note: The green curve should be read against the right-hand vertical axis.

4.4 Age patterns in the 2011 census and other household data

Figure 12 below reflects the population by age, converted to birth cohorts, according to the 2011 Census and the 2012 General Household Survey (GHS). Arguably these curves are less influenced by modelling assumptions than the mid-year population estimates discussed above. Both household collections confirm the existence of large increases in the size of birth cohorts beginning in around 2003. For instance, the GHS points to an increase between 2001 and 2005 of 9% using the weighted data (2001 is arguably a better base than 2002, which represents an exceptional dip). The distribution using unweighted GHS data is illustrated to confirm that the weighting does not alter the general age distribution patterns in a substantial way. The 2011 Census data point to an increase of 11% between the birth years 2003 and 2005 (close to the 12% seen in the ASS data) and an 18% difference between 2002 and 2005 (this is a larger difference than what one finds in the ASS or Home Affairs data). Importantly, the 2011 Census points to continued growth in the size of birth cohorts beyond 2005, where the Home Affairs data point to a levelling off followed by a decline. It is noteworthy that a bulge at approximately the birth year 1999 or 2000 is visible in both data sources reflected below, but also in the ASS data (see Figure 3).

Figure 12: Stats SA household survey birth cohort distributions



Sources: Statistics South Africa, 2014; Statistics South Africa, 2012.
Note: The horizontal axis refers to birth year. The unweighted GHS values have been recalibrated by inflating each birth year value by a common factor that made total respondents across all years equal the corresponding total from the weighted data. Markers on the curves were positioned against the horizontal axis in such a way that they reflected the survey's definition of age within the year, and years on the axis reflect the middle of the calendar year (30 June).

It is also worth noting that absolute values reflected in the census data tend to be higher than values reflected in the Home Affairs birth registrations data. For example, the 2008 birth year peak in the 2011 Census data seen in Figure 12 is 7% above the corresponding value in Figure 8. Could this point to an under-registration problem in the Home Affairs data, something which would have large implications for the analysis presented here? Earlier analysis by Gustafsson (2012) and other analysts within the Department of Basic Education suggests that a more likely explanation is that there is an over-weighting problem in the census data.

4.5 How do the South African changes compare internationally?

One indicator that is relevant for understanding the magnitude of the challenge of school expansion, and which lends itself to a cross-country comparison, is the size of the population born in the last five years, divided by the size of the population born during the preceding five years. Using the adjusted births seen in Figure 8, and ignoring the impact of factors such as post-registration deaths, one arrives at a value for this indicator in 2010 of 1.09. This would represent births during 2006-2010 over births during 2001-2005. In 2008, the indicator value would have been as high as 1.14. What have other comparable countries seen in recent years? In 2010, 15 middle income countries (16% of all middle income countries) faced a ratio of at least 1.09. These countries included Nigeria (a ratio of 1.14), Pakistan (1.13), Zambia (1.10) and Angola (1.24), countries commonly cited as high fertility countries[4]. South Africa is thus not alone in facing the current kinds of challenges to accommodate a growing number of learners in its schooling system. However, the numbers suggest that South Africa will for some years be within a group of countries which must display particularly creative approaches to ensure that all children have access not just to schooling, but also schooling which gets

[4] Own analysis of UN: UN Statistics Division (2008).

better, not worse, over time. Enrolment increases tend to stretch resources, and increase class sizes, trends which can easily undermine the quality of schooling.

5 Increases in births, the child support grant and ARV access

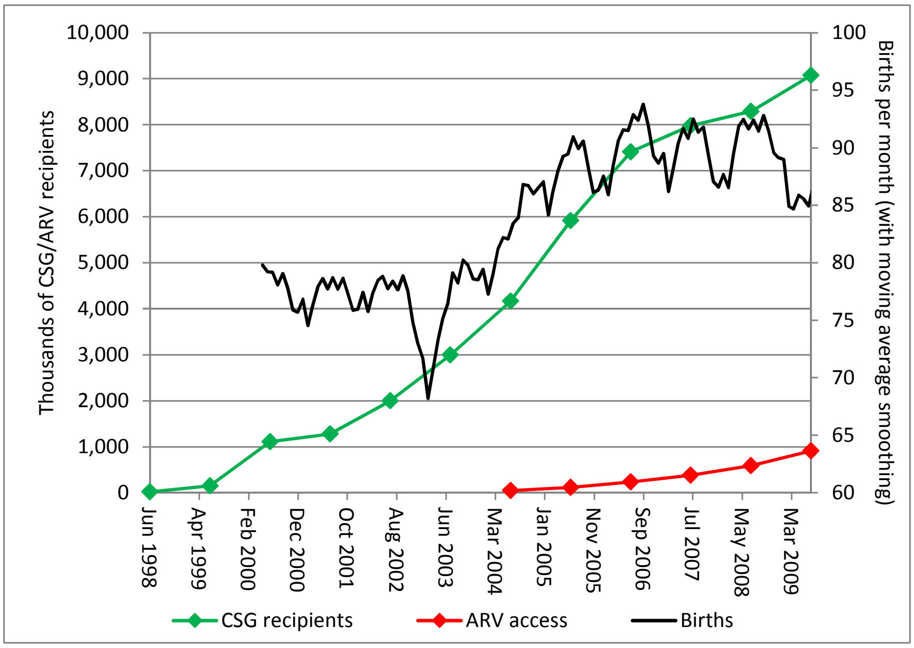
It is not the intention of this report to produce an analysis of why the increases in the birth cohorts described above have occurred. However, a few patterns in the data which could be relevant to answer this question are presented here. It should be emphasised that a proper explanation would require considerable analysis of correlations over time and at as low a geographical unit as possible, and even then there is no guarantee that the question will be answered conclusively.

Figure 13 displays the trends for two variables which could have influenced births. The number of child support grant (CSG) recipients increased from zero to nine million in the decade 1998 to 2009. However, this trend has been fairly smooth, with no striking change around the years 2003 to 2005, when the large increases in births occurred.

The Home Affairs birth trend displayed in Figure 13 uses monthly, not annual, totals in order to provide a more precise picture of when the increases in births began. There are clearly important within-year fluctuations. These reflect, for instance, fewer births each November and a particularly large number of births each September. If one calculates the ratio of births in a month, divided by births in the same month a year ago, the highest ratio in the post-2000 period is that of January 2004, which comes to 1.21 (births in January 2004 were 21% higher than births in January 2003).

One change which did occur in 2004, was the beginning of a rapid rise in access to antri-retroviral (ARV) treatment. However, if this influenced births, the question is why an increase in births began as early as January 2004. If access to ARVs, and a perceived reduction in the risk of mother-to-child HIV transmission, influenced people's decisions to have children, one would expect the increase in births to begin a year later, around January 2005. As with the CSG trend, there is no immediately striking pattern which suggests strongly a clear link. However, ARV access is still the explanatory factor for which we have the strongest evidence in the available published research, as will be shown below.

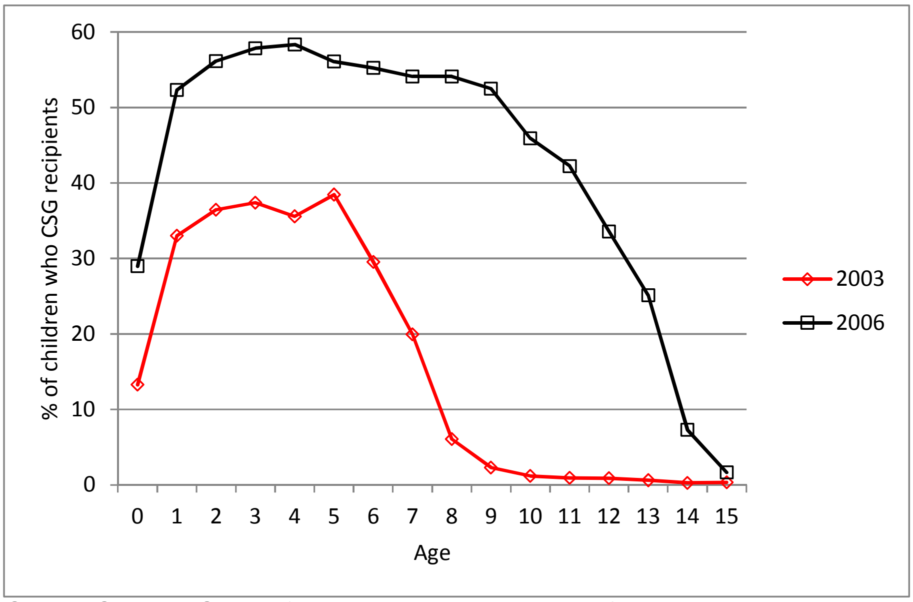
Figure 13: Child support grant and ARV recipients 1998-2009



Source: Child support grant (CSG) recipients from Department of Social Development (2012: 117). Anti-retroviral (ARV) access from Johnson (2012: 24). Both CSG and ARV values are positioned in the middle of the year (30 June) for which the data corresponds. The births curve uses monthly values, with a moving average trendline covering five points (source is the microdata referred to under Figure 5.

The next graph, which draws from the General Household Survey data, points to important aspects of the CSG trend. Most of the increase in the number of CSG recipients in the period 2003 to 2006, when births rose, can be attributed to older children, specifically children older than around seven. It seems likely that increases in the number of younger recipients, not older recipients, would have a stronger impact on decisions to have children. Figure 14 thus reduces further the apparent likelihood of the CSG as a major contributing factor to more births. However, it should be re-emphasised that any firm conclusions would need to be based on analysis of microdata, and patterns over time at a local level.

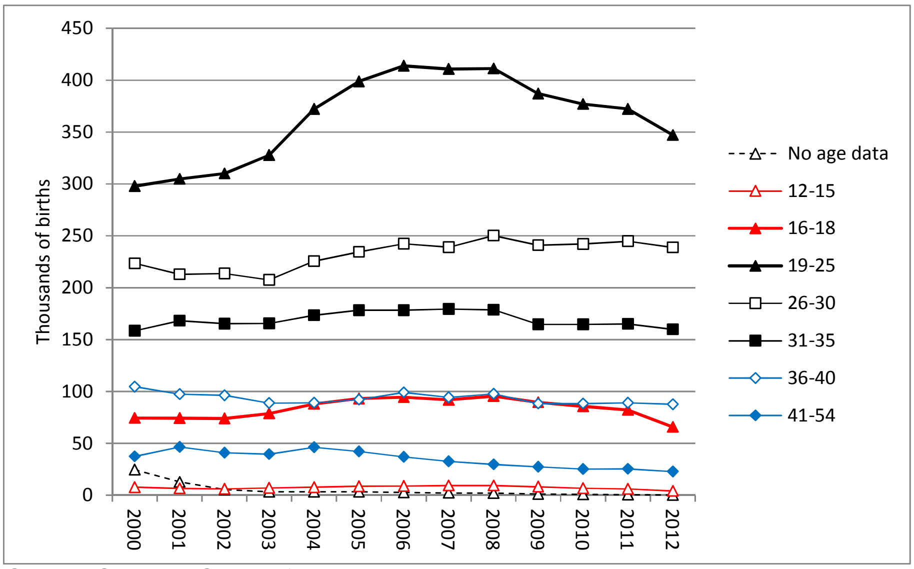
Figure 14: Child support grant recipients 2003-2006 by age



Source: Statistics South Africa (2012) – survey datasets for 2003 and 2006 analysed.

Concerns that the child support grant was encouraging teenage pregnancies led to a number of studies into this matter. Makiwane, Desmond, Richter *et al* (2006) examine fertility trends over time, just amongst teenagers, using data from the period 1995 to 2001. They conclude that the CSG was not causing teenage pregnancies. Makiwane (2010) reaches a similar conclusion, and presents statistics showing that the fertility rates of teenagers declined steadily between 1980 and 2009. The Home Affairs data used for the current report shows an increase in the number of births registered by mothers aged 16 to 18, with the 2003 to 2006 increase being 20%. This is not the same as the fertility rate for teenagers, yet it appears that different data sources do reveal different trends. Figure 15 below describes the birth number trends for different age categories of mothers. The category displaying the largest 2003 to 2006 increase was in fact mothers aged 19 to 25, a trend which would not be compatible with the hypothesis that a surge in teenage pregnancies was the primary cause of the births surge. Finally, Mokoma (2008), using household data from the 1995 to 2004 period, concludes that the CSG did not raise teenage fertility rates.

Figure 15: Birth registrations by age of mother



Source: Statistics South Africa, 2015.
Note: The horizontal axis refers to birth year.

Not much research has been conducted into the link between ARV treatment and fertility, but the studies that exist do suggest that increases in the former are likely to lead to increases in the latter. Myer, Morroni and Rebe (2007) interviewed ARV patients in South Africa in 2005 and found that ARV access increased the interest amongst women in having a child. Myer, Carter, Katyal, *et al* (2010) used a panel of data covering the years 2003 to 2007 and drawing from seven African countries and found that greater ARV access did lead to more pregnancies, with South Africa displaying a particularly strong relationship. Specifically, this study finds that in South Africa women were 2.23 times more likely to become pregnant if they were accessing ARV treatment. If one does a rough calculation using the assumption (based on the ARV access curve of Figure 13) that around 13% of women giving birth were receiving ARV treatment, and using the 2.23 inflation value, on arrives at an increase in births of 16%, which is remarkably close to the 17.6% increase in registered births between 2003 and 2006 represented in Figure 8.

6 Implications for basic education budgets

There is no simple relationship between enrolment totals and the cost of delivering education in schools. Smaller schools can often absorb enrolment increases without hiring more teachers because classes are relatively small. To illustrate, 8% of grades 1 to 7 learners were enrolled in grades during 2013 where there were 25 or fewer learners in the grade. The very smallest schools would be forced to practice multi-grade teaching, but less small schools would not. Ongoing urbanisation, which shifts enrolments from smaller schools to larger schools, where ‘economies of scale’ allow for higher learner/educator ratios, further improves the capacity of the system to absorb overall enrolment increases.

What is presented below is a very basic, yet useful, set of projections indicating how the population changes outlined in this report are likely to impact on total enrolment and, by implication, on total basic education spending. Again, one should bear in mind the provisos relating to the link between total enrolment and total cost. Moreover, it should be emphasised that the data allow for much more robust work to be done than what is presented here. The intention with what follows is to provide just a rough ‘back of the envelope’ estimate of future trends.

First a set of annual birth number changes for 1985 to 2035 were generated. This was done three times, to produce three scenarios. The three sets of ratios reflecting birth number changes are as follows:

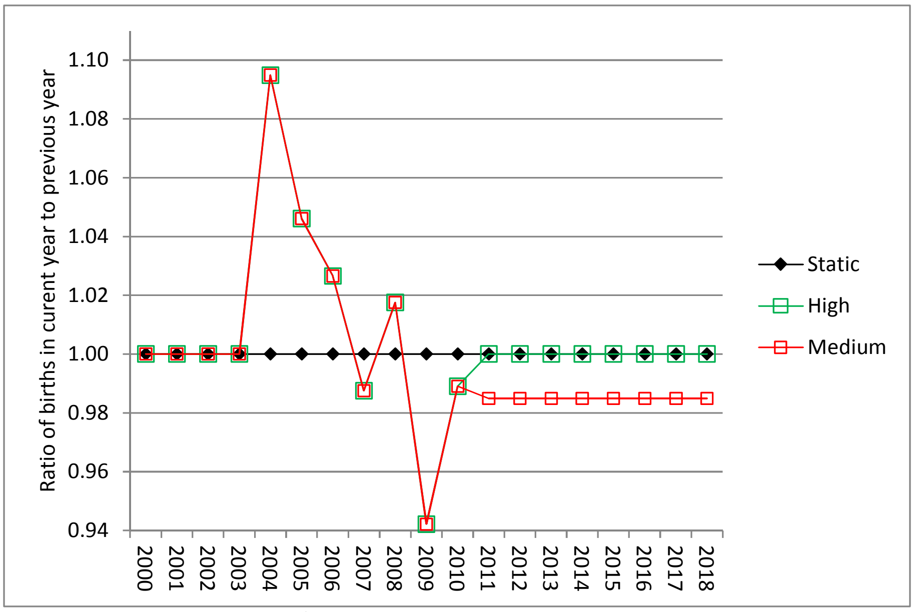
- **Static scenario.** For all years 1985 to 2035 the ratio of births in the current year to births in the previous year was assumed to be 1.00, meaning it was assumed that the number of learners entering Grade 1 was the same for all years. This is not a business-as-usual (BAU) scenario, but an unrealistically static scenario which simply helps to interpret the graphs below.

- **High impact scenario.** Here it was assumed that births remained constant up to 2003, and then starting rising substantially, as seen in the births curve in Figure 11 above. The 2003 to 2008 rise in births follows the curve of Figure 11. Thereafter, a decline in births from 2008 to 2010 along the lines of Figure 11 is assumed. From 2010 onwards, the number of births is assumed to remain static at the 2010 level.

- **Medium impact scenario.** This is like the previous scenario up to 2010. After 2010, it is assumed that births each year would be 0.9849 of births in the previous year, up to 2029. This downward slope is based on post-2008 fertility declines seen in the fertility (TFR) curve in Figure 11. By 2029, the fertility rate would reach the replacement rate of 2.0, and as it is assumed fertility would not dip below this level, after 2029 it is assumed that the number of births would not change.

The annual change ratios for the three scenarios and for the years 2000 to 2018 are illustrated in Figure 16 below.

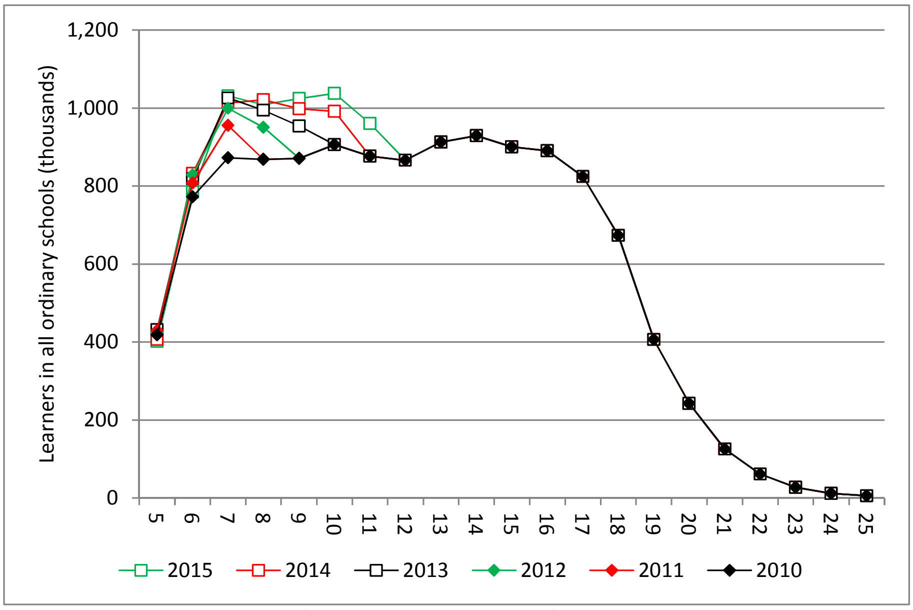
Figure 16: Annual change ratios for three scenarios



Note: The horizontal axis refers to birth year.

Next, enrolment by birth year in 2010 according to the Annual Survey of Schools, for the ages 5 to 25, was obtained. This is shown in the 2010 curve of the next graph.

Figure 17: Medium impact scenario enrolments by age 2010-2015



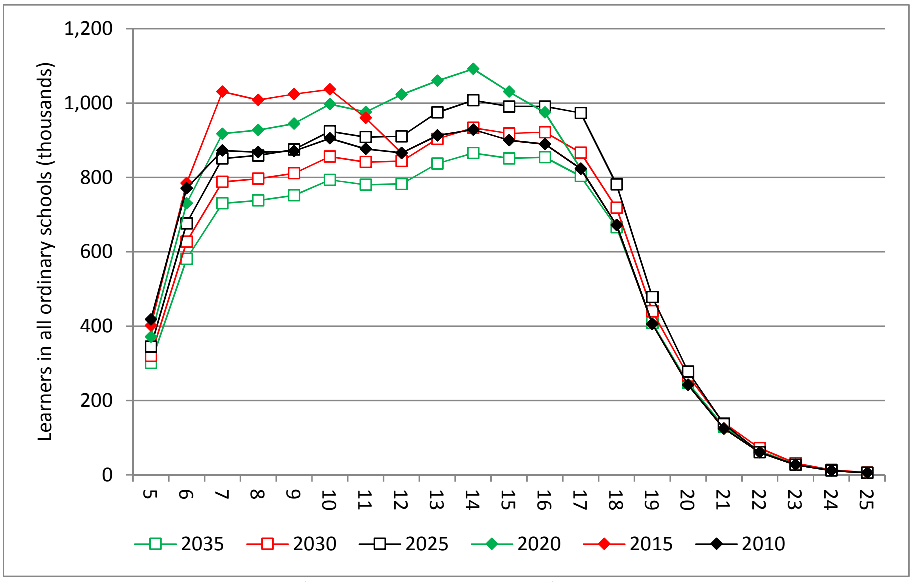
Note: The horizontal axis refers to age at the end of the calendar year.

To calculate enrolment by age curves for years beyond 2010, the following formula was used. Here *E* is enrolment, *y* is the year beyond 2010, *a* is age (in the range 5 to 25), *r* is the annual change ratio and *b* is birth year. Essentially enrolment for a current and specific age *a* is the previous year's value for the same age adjusted up or down according to earlier changes in the number of births. Mortality rates are thus assumed to be constant across birth cohorts. The subscript to *r* is essentially saying that the birth year to use is determined by taking 2005 (the birth year for 5 year olds in 2010), moving backwards a number of years depending on the difference between the current age *a* and the minimum age 5, and then adding the difference between the current year *y* and the starting year 2010.

$$E_{y,a} = E_{y-1,a}\left(r_{b=2005-(a-5)+(y-2010)}\right)$$

The resultant curves for years up to 2015, using the medium impact scenario, are included in Figure 17 above. Curves for later years in the medium impact scenario are shown in the next graph.

Figure 18: Medium impact scenario enrolments by age 2010-2035

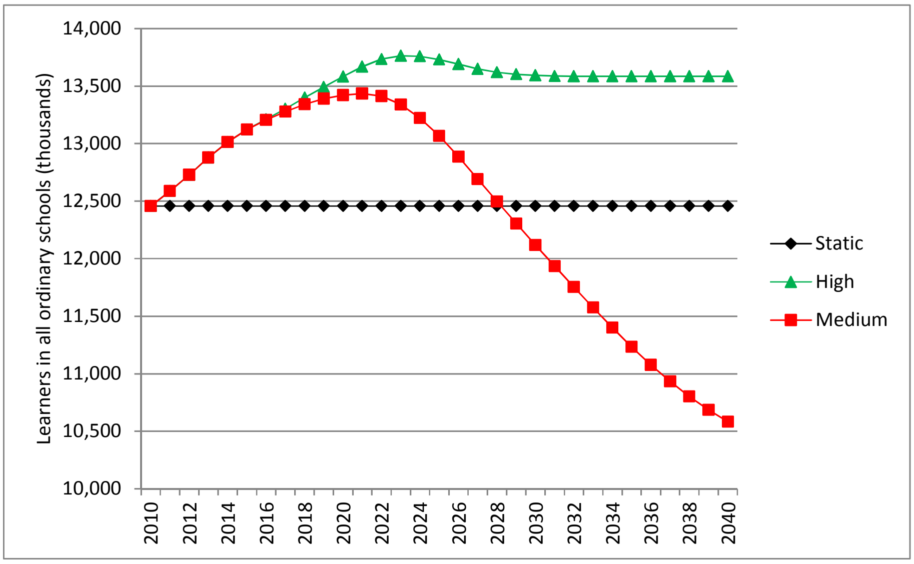


Note: The horizontal axis refers to age at the end of the calendar year.

In Figure 18 it is clear that the general trend is for enrolment by age to increase and then decrease, with the lowest values being reached by 2035. This is because the medium scenario envisages first the impact of the higher numbers of births after 2003 moving through the system, and thereafter the declining post-2008 birth numbers bringing about enrolment declines.

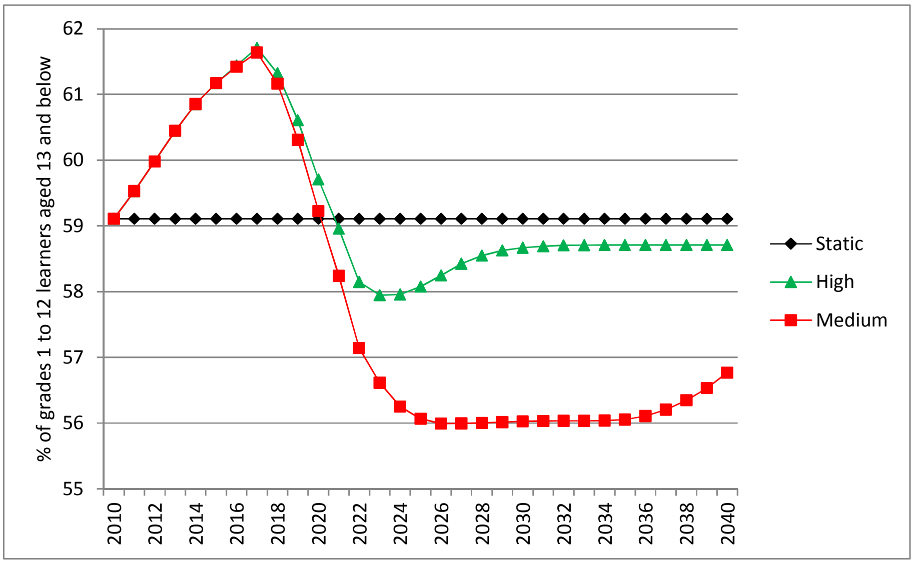
Figure 19 below, illustrates the projected enrolments for all three scenarios. By design, the static scenario results in no enrolment change. The high impact scenario leads to a peak in 2023 which is 10% above the 2010 baseline, at almost 13.8 million learners. In the medium scenario, the peak of 2021 is 8% above the baseline.

Figure 19: Enrolment changes to 2040 for all scenarios



Of course the challenge for the schooling system is not just to deal with more learners, but also to deal with enrolments which will move across ages and grades, something which requires sufficient flexibility, so that resources can be moved from one level of schooling to another as demand shifts. Figure 20 below describes how the post-2003 increase in births brings about a larger percentage of enrolments at the primary level, with a peak in 2017, after which enrolment demand begins to shift in favour of the secondary level (as discussed above, the post-2003 rise in births will reach grade 8, the start of the secondary level, in 2018). The medium scenario results in a particularly large shift towards a higher secondary-level share of enrolments, something one would expect with large fertility declines after 2008.

Figure 20: Percentage of enrolments at the primary level



7 Conclusion

This report has presented important information on the impact of large changes in birth patterns occurring in the years following 2003. It would have been good to have this information earlier as planning and adjusting the schooling system takes time. In many ways, what confirms the births trend is that different data sources display similar patterns. These data sources have their limitations and may not have been taken seriously if viewed separately and on their own. Developing countries such as South Africa experience serious data quality problems. These are problems which should be resolved, but cannot be resolved overnight. This underlines the importance of using whatever data are available, and employing triangulation techniques, in order to ensure that policies are informed by as much evidence as possible.

References

Châu, T.-N. (2003). *Demographic aspects of educational planning*. Paris: UNESCO: IIEP.
Department of Basic Education (2013). *2013 school realities*. Pretoria. [Earlier issues also consulted.]

Department of Basic Education (2013b). *Report on progress in the schooling sector against key indicators*. Pretoria.

Department of Basic Education (2013c). *Macro indicator report*. Pretoria.

Department of Basic Education (2017). Snap Survey of ordinary schools 1997-2016 (dataset). Pretoria. Available from: <http://www.datafirst.uct.ac.za> [Accessed January 2017].

Department of Basic Education (2016a). *2016 school realities*. Pretoria. Available from: <http://www.education.gov.za/> [Accessed January 2017]. Earlier publications in this series also used.

Department of Basic Education (2016b). *Report on progress in the schooling sector against key learner performance and attainment indicators*. Pretoria. Available from: <http://www.education.gov.za> [Accessed September 2016].

Department of Social Development (2012). *The South African child support grant impact assessment*. Pretoria. Available from: <http://www.unicef.org/southafrica/SAF_resources_csg2012s.pdf> [Accessed July 2014].

Dorrington, R. (2013). *Alternative South Africa mid-year estimates, 2013*. Cape Town: Centre for Actuarial Research.

Gustafsson, M. (2012). *The gap between school enrolments and population in South Africa: Analysis of the possible explanations*. Stellenbosch: University of Stellenbosch.

Johnson, L.F. (2012). Access to antiretroviral treatment in South Africa, 2004-2011. *The Southern African Journal of HIV Medicine,* 43: 22-19.

Makiwane, M. (2010). The Child Support Grant and teenage childbearing in South Africa. *Development Southern Africa*, 27(2): 193-204.

Makiwane, M., Desmond, C., Richter, L. & Udjo, E. (2006). *Is the Child Support Grant associated with an increase in teenage fertility in South Africa?* Pretoria: HSRC.

Myer, L., Carter, R.J., Katyal, M. & Toro, P. (2010). Impact of antiretroviral therapy on incidence of pregnancy among HIV-infected women in Sub-Saharan Africa: A Cohort Study. *PLoS Medicine*, 7(2).

Myer, L., Morroni, C. & Rebe, K. (2007). Prevalence and determinants of fertility intentions of HIV-infected women and men receiving antiretroviral therapy in South Africa. *AIDS patient care and STDs,* 21(4): 278-285.

Mokoma, T.P. (2008). *An analysis of the impact of child support grant on teenage fertility rate in South Africa*. Stellenbosch: Stellenbosch University.

Statistics South Africa (2012). General Household Survey 2012 (dataset). Pretoria. [Datasets for earlier years also analysed.]

Statistics South Africa (2013). *Mid-year population estimates, South Africa 2013*. Pretoria. [Earlier and later issues also consulted.]

Statistics South Africa (2014). Census 2011 10% sample (internet-based Nesstar data querying facility). Pretoria. Available from: <http://www.statssa.gov.za> [Accessed July 2014].

Statistics South Africa (2015). *South Africa – recorded live births 1998-2014 (dataset)*. Pretoria. Available from: <https://www.datafirst.uct.ac.za> [Accessed April 2017].

Statistics South Africa (2016). *Mid-year population estimates, 2016*. Pretoria. Available from: <http://www.statssa.gov.za> [Accessed January 2017].

UN: UN Statistics Division (2008). World Population Prospects 2008 (dataset). New York. Available from: <http://data.un.org> [Accessed November 2009]