Inflow of new teachers into the public system

Martin Gustafsson
17 March 2015

1 Introduction

This short report explains the derivation of provincial and national values for one of the Action Plan indicators, namely indicator 14, which reads as follows: *The number of qualified teachers, aged 30 and below, entering the public service as teachers for the first time during the past year.* The current report thus provides a guide for the methodology of future derivations of values for this indicator, plus a source for national and provincial publications. What were considered the best possible indicator values for 2014 appear in Table 2 below.

In some ways, this report updates a part of an earlier report I produced, titled 'Teacher supply patterns in the payroll data', dated 3 March 2009. The current report looks very similar to a report with the same title as this one, but dated 20 February 2014. That report used data from the years 2004 to 2012 (to a limited extent to 2013), whilst this one uses data from 2004 to 2014, in other words for two additional years. (The previous report did not really make use of 2013 data as those data were from May 2013, and therefore out of line with the months of other years. For the current report, for 2013 the more appropriate data from November 2013 were obtained.)

2 The data

Persal data for all educators across all provinces for the years 2004 to 2014 were used. For each year, data for one month were used. The months were not completely consistent, but all in the range of September to November. The qualification variable (REQV) was not available in the downloads of all years. The data consistency situation is reflected in the following table.

Table 1: Data inconsistencies

Year	Month	REQV values available?
2004	Sep	Yes
2005	Oct	Yes
2006	Oct	Yes
2007	Oct	Yes
2008	Sep	Yes
2009	Sep	No
2010	Oct	No
2011	Oct	Yes
2012	Oct	No
2013	Nov	Yes
2014	Oct	Yes

3 Imputation of missing qualifications values

Note that even where REQV values were available for a particular year, they were often missing for individual educators. For the indicator in question, a key concern is whether the educator was qualified or not in a particular year. Anyone with REQV 13 or above is supposed to be considered qualified. The rules for imputing the qualified (value 1) or not qualified (value 0) values, where they were missing, were as follows:

a. If there was no value for any year for the educator, the value remained missing.

b. If there was a value 0 (unqualified) in an earlier year, and there was no value 1 (qualified) for an earlier year, the value would become 0 in the current year.

c. If there was a value 0 in a future year, and there was no value 1 for the educator for any future year, the value would become 0 in the current year.

d. If after following the above steps there were still missing values and the educator was marked as qualified in any year, then the missing value would become 1 (qualified).

4 Methodological and data issues

4.1 The extent of multiple entry

What prevents a straightforward use of just two years of Persal data at a time when calculating indicator values, is the fact that some educators enter more than once. Someone may appear as a joiner in 2005, then be absent during 2006 and 2007, and reappear as a joiner in 2008. Specifically, analysis of the data indicated that 5.9% of the 589,230 people who were educators in the 2004 to 2014 period, entered the system more than once in the years 2005 to 2014, where entering the system once means being absent in the previous year and present in the current year. The great majority of this multiple entry manifests itself in two entries. 5.3% of the 589,230 educators entered twice during the period in question. The phenomenon of multiple entry means that examining joining whilst using the data from just two consecutive years is likely to result in an over-estimation of the number of joiners. This strengthens the justification for using data from as many years as possible for any one educator, as is done in this report.

4.2 REQV data problems

The earlier 2009 report referred to problems with delayed updating of the REQV qualifications variable in Persal. Educators are appointed, but often it takes months or even years before the correct REQV value is entered on Persal. In the meanwhile, the value remains missing. There also seems to be an impression that non-qualified values (REQV 12 and below) are sometimes inserted on Persal pending an inquiry into the educator's qualification status. These issues obviously create problems for anyone trying to track the number of joiners to the public system who are qualified. The imputation of qualifications values described above would to some degree deal with the problems, but less so for more recent years. If an educator has recently joined, and data entry delays have resulted in no insertion of any qualification value in Persal, then there would be no imputation. However, as argued below, it seems as if the problem with delayed entry of REQV values is less serious in recent years than it was when the 2009 report was written.

4.3 Temporary versus permanent

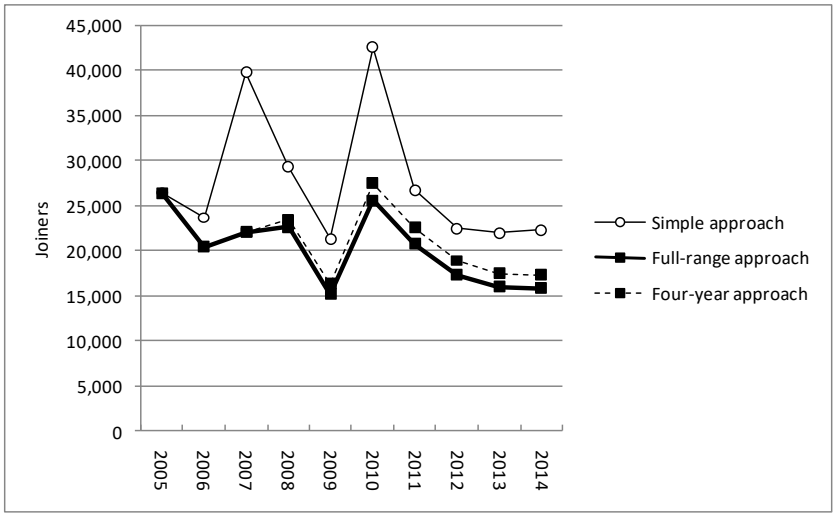
The fairly large extent to which educators are employed on a temporary basis suggests that indicator 14 values must take into account temporary appointments, and not just permanent ones. To illustrate, in 2014 10% of educators were employed on a temporary basis.

5 Results

Figure 1 below provides three different measures of joiners. In the 'simple approach', anyone who was not in the system in the previous year but was present in the current year is counted as a joiner. Given multiple entries, this measure would clearly over-estimate joining as many people would be counted more than once. The curve 'full-range approach' illustrates joiners

who were not present in any earlier year (of the years for which data were available). Using this measure could also result in an over-estimate of joining, in particular for earlier years. For instance, a joiner in 2005 would be classified a joiner if she was not present in the data in 2004. However, we do not know whether this person was present in 2003, in which case she should ideally not be counted as a 2005 joiner. For later years, however, one is less likely to encounter this kind of problem. The 'four-year approach' involves looking at just the previous three years, and if the person was not present during those three years, but was present during the current year, the person would be a joiner for the current year. The measure starts in 2007, because the earliest data are from 2004. Though this third measure can also result in an over-estimation of joining, its advantage is that the method is the same for each year, so it is possible to compare the trend over some years (specifically the 2007 to 2014 trend).

Figure 1: Effect of multiple entry on joiner statistics



So what relevant trends are seen in the previous graph? Clearly we should use the second or third measure, and given they are not too different from each other, it is not terribly important which one we use. The number of 'first-time' joiners per year has been uneven, within the range 15,000 to 25,000 roughly. The earlier 2009 report had argued that the public system needs to attract at least 10,100 new educators each year. This is simply to maintain the existing stock of educators, without any expansion. Clearly the actual inflow is not too bad in the sense that it has not dipped below 10,100 in any recent year. However, what proportion of the inflow in recent years consists of unqualified teachers? Following an approach used below for young teachers, it was calculated that 27% of the joiners illustrated in the 'four-year approach' curve of Figure 1 for the years 2010 to 2014 were unqualified. This still leaves the average number of qualified joiners per year slightly above the 10,100 threshold, specifically at around 11,700 for the years 2010 to 2014.

Ideally, virtually all the new inflow should consist of young teachers. As will be seen, the number of qualified young joiners does not reach 10,100, in fact it has remained at around a half to two-thirds of this level for many years. This should be a cause for concern. Figure 2 below repeats two of the curves from Figure 1 for educators aged 30 or below in the year of entry. There appears to have been a large increase in 2010, to around 10,000 young joiners per year, using the more reliable 'four-year approach' measure.

Figure 2: Effect of multiple entry on young joiner statistics

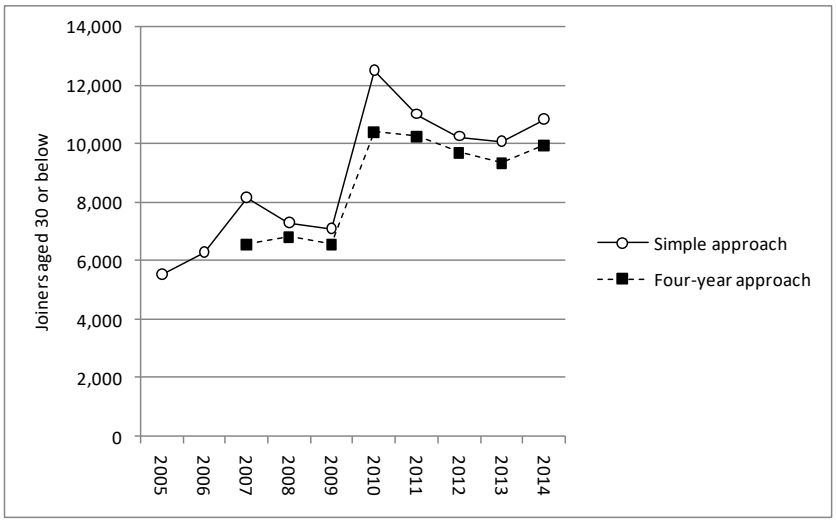
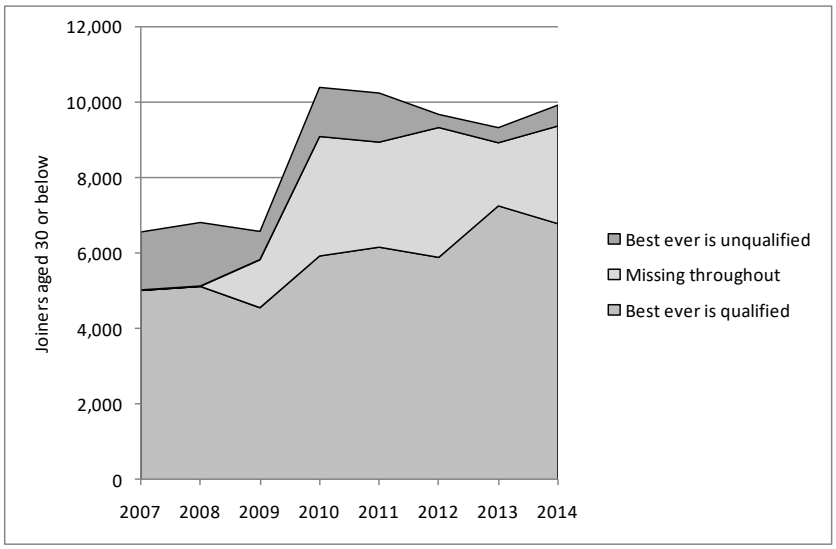


Figure 3 illustrates the breakdown of the 'four-year approach' curve from the previous graph, by qualification status. Here 'best ever is qualified' means the educator has REQV 13 or above in at least one year in the range 2004 to 2014, 'best ever is unqualified' means that the educator had REQV values, but all below the REQV 13 threshold, and 'missing throughout' means there was no qualification data for the educator from any year. The problem of missing values is clearly seen. The 7,929 joiners with missing qualifications values for the years 2011 to 2013 were examined further to see to what extent these were likely to be qualified educators. There is in fact little to suggest that they were. Of the educators in question, 62% were not present in the October 2014 Persal data, suggesting they were fairly transient employees. The 38% who were present in the 2014 data had in 90% of cases the nature of appointment category 'abnormal appointment' or 'periodical remuneration'. Such educators are less likely to be qualified educators. 52% of the joiners are in KwaZulu-Natal, and a further 13% in Gauteng. It is difficult to draw firm conclusions from the available data, but it seems as if few, if any, of the educators within the wedge 'missing throughout' in the graph below were qualified educators. Thus it seems that few of these educators were qualified educators for whom the REQV status had yet to be filled in on Persal. The problem of delayed entry of REQV values on Persal appears not be as serious in recent years as it was in the years studied within the earlier 2009 report. Specifically, since 2010 the 'missing throughout' wedge in the graph has become a bit narrower.

Figure 3: Qualifications status of young joiners



It seems as if the wedge 'best ever is qualified' in the previous graph represents the best figures obtainable with the available data for indicator 14. Figures could be slight under-estimates insofar as the wedge 'missing throughout' is likely to include a few educators who are qualified but whose REQV values were not entered on Persal. But the figures would also be slight *over*-estimates because only four years of data were used. Had more years of data been used, some joiners would have to be discarded, specifically those who had entered in some earlier year. But as Figure 1 shows, this over-estimation is unlikely to exceed around 10% (see the gap between the 'full-range approach' and 'four-year approach' curves). The next table provides actual figures, including provincial figures, for what I would argue are the best set of indicator values (namely 'best ever is qualified' from Figure 3).

Table 2: Young qualified joiners 2007-2014

	2007	2008	2009	2010	2011	2012	2013	2014
EC	706	818	960	1,146	213	281	189	237
FS	296	369	221	396	504	540	590	424
GP	947	872	876	1,099	1,154	1,345	1,607	1,580
KN	1,796	1,707	1,316	1,649	2,196	1,535	2,116	1,711
LP	165	226	234	310	486	395	597	593
MP	372	379	270	435	433	579	521	679
NC	132	143	129	109	148	154	162	197
NW	157	93	158	207	301	343	579	544
WC	459	528	407	592	742	734	912	835
SA	5,030	5,135	4,571	5,943	6,177	5,906	7,273	6,800

Note: Only people who were aged 30 or below in the year in which they were counted appear. A four-year approach was followed, meaning for instance that for someone to be a joiner in 2007, they should not have been present in Persal in 2004, 2005 or 2006 (hence four years of data were considered). Only educators who at any point (in any of the eight years) were qualified according to the Persal data were counted.

The above figures are not very different to figures calculated by Mfela Mahlangu in the DBE. In those calculations, data on appointments in Persal are used. That approach would, like the approach presented in the current report, not fully resolve issues related to multiple joining and missing qualifications data. Mahlangu's total young joiner figure for 2013 is around 5,300, so lower than the 7,273 seen above in Table 2. This is probably because the Table 2 figure includes qualified teachers who entered the system for the first time (in four years) on a *temporary* basis. Clearly, the number of young joiners has for many years been a little over half of what it should be, assuming that the 10,100 per year threshold represents a level of joining that allows for a sustainable teaching force.

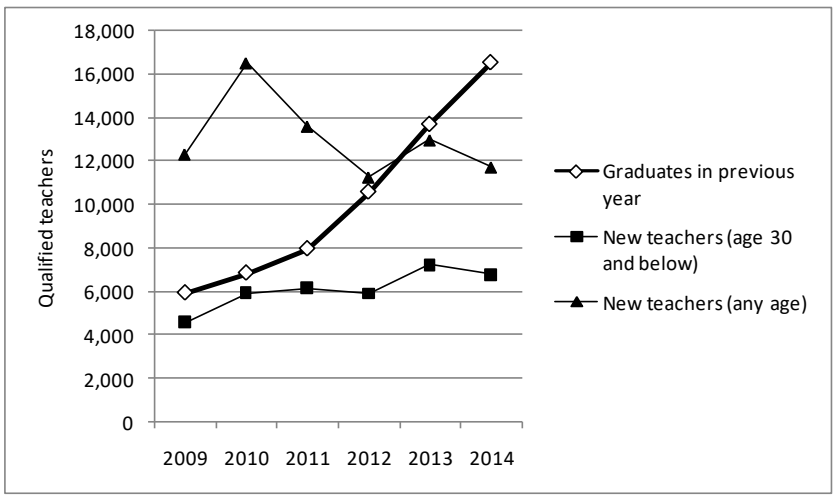
To complete the picture, in a sense, the following table provides the same statistics as Table 2, except here people of any age were considered.

Table 3: All qualified joiners 2007-2014

	2007	2008	2009	2010	2011	2012	2013	2014
EC	2,603	2,805	2,888	3,389	601	480	356	434
FS	1005	1,087	463	908	1,148	852	905	564
GP	3,264	3,093	2,608	3,418	2,669	2,849	3,073	2883
KN	3,506	3,261	2,189	3,161	3,836	2,487	3310	2718
LP	1,720	2,164	1,394	1,611	1,539	882	1217	1229
MP	1,630	1,760	884	1,328	914	1,172	1003	1170
NC	394	475	263	276	304	326	380	345
NW	941	342	675	1,033	1209	864	1217	1007
WC	1,458	1,835	968	1,398	1,392	1,363	1,522	1388
SA	16,521	16,822	12,332	16,522	13,612	11,275	12,983	11,738

It is important to see the trends described above in the light of increases in the number of university students graduating as teachers each year. Whilst the number of graduates increased by around 180% (in other words almost tripled) between 2008 and 2013, the number of young teachers, aged 30 or below, entering the public schooling system increased by only around 35% between 2009 and 2014. This confirms that there are serious problems attracting young graduates into working in public schools. The problem goes well beyond the issue of whether Funza Lushaka bursary recipients are placed in schools according to plan. These bursary recipients account for only around 2,500 graduates a year in recent years, meaning that the bulk of the gap between graduates and joiners to the public system is accounted for by newly qualified teachers who are not Funza Lushaka bursary recipients.

Figure 4: Graduation and joiner trends compared



Source: The figures for graduates are obtained from a pre-final version of the 2015 Funza Lushaka evaluation report of Presidency.
Note: The curve for new teachers aged 30 and below uses the same figures as 'Best ever is qualified' from Figure 3. The curve for new teachers of any age uses the same method as that for teachers aged 30 and below. As one might expect graduates from one year to become teachers in the next year, the graduates curve has been shifted to the right, so for instance the 2009 value is actually the number of graduates from 2008.