An Analysis of North Kingstown School Enrollment Projections

Tom Sgouros

4 January 2006

1 The data

Doing enrollment projections is a challenging task. The job involves predicting the future, which isn't a skill usually available to ordinary mortals. But it is possible to use trends to extrapolate past experience into the future, and by so doing, come up with projections that are defensible, and potentially more accurate than simple guesses. The trick is to identify the trends worth using.

Once identified, a projection technique involves mathematically projecting the trends out into the future. To get good projections, school demographers will adjust the coefficients slightly based on events that have already happened (the anomalous arrival of ten fourth graders in one year or a change in admission rules, for example) to arrive at a final result. This "tuning" of the mechanical projections is an important step in the process of making accurate projections, but it is beyond the scope of this critique, which is only a comparison of the mechanical starting point for projections.

The data we have to work with is the enrollment data shown in Table 1. This data represents the actual numbers of kids, counted on October 1 of each year, which is usually after the opening fluctuations in school population have settled down.

Everything to the right of the vertical line are the projections currently used by the superintendent's office in planning, and they are the subject of this analysis. There are a couple of notable features about these projections. For one thing, the number of kindergarten students is projected to decline quite a bit in the coming years. This appears to cause a decline of over 300 students in the elementary school population by 2011, a decrease of over 17% . Complaints about crowding the elementary grades will therefore be ameliorated by the declining enrollments.

Middle school enrollments are also slated to decline, almost as much as the elementary schools, and the high school enrollment is predicted to rise, though only by a few percent. Overall, school enrollment will decline by a few hundred children from the current levels.

Table 1: October 1 Enrollment Data, with administration projections

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
K	304	330	353	341	332	342	313	327	292	309	290	261	293	288	283	278	273	268	263
1	343	358	379	388	358	352	395	343	341	342	331	317	286	287	282	277	272	268	263
2	336	326	354	337	369	351	343	380	341	342	339	332	313	275	277	272	267	262	258
3	325	351	344	350	348	365	355	342	394	362	337	347	328	312	274	275	271	266	261
4	331	335	354	345	339	349	356	342	344	385	353	351	341	329	312	274	276	271	266
5	347	330	337	359	348	356	350	357	352	352	381	354	363	343	330	313	276	277	272
6	316	341	331	337	350	338	342	352	352	354	344	391	353	364	344	331	315	277	278
7	326	322	356	334	344	348	340	334	350	350	355	351	394	365	377	356	343	325	286
8	326	335	332	350	323	324	353	331	338	348	348	352	337	396	367	378	357	344	327
9	367	355	378	356	388	343	328	337	355	386	377	384	412	351	412	382	394	372	358
10	330	308	345	360	370	363	345	328	335	366	395	378	376	424	361	424	393	406	383
11	299	290	254	250	282	316	338	335	305	337	366	365	355	381	430	366	430	399	411
12	241	296	289	293	287	311	291	326	320	300	312	356	370	365	392	442	377	442	410
Total	4191	4277	4406	4400	4438	4458	4449	4434	4419	4533	4528	4539	4521	4479	4440	4370	4243	4176	4037
SPED	66	66	66	74	87	84	73	67	91	120	117	109	110	114	119	125	131	137	143
Grand	4257	4343	4472	4474	4525	4542	4522	4501	4510	4653	4645	4648	4631	4594	4560	4495	4373	4313	4180
HS	1237	1249	1266	1259	1327	1333	1302	1326	1315	1389	1450	1483	1513	1521	1595	1614	1594	1618	1562
Middle	968	998	1019	1021	1017	1010	1035	1017	1040	1052	1047	1094	1084	1125	1087	1065	1014	946	891
Elem	1682	1700	1768	1779	1762	1773	1799	1764	1772	1783	1741	1701	1631	1545	1475	1412	1361	1344	1320
K-5	1986	2030	2121	2120	2094	2115	2112	2091	2064	2092	2031	1962	1924	1833	1758	1690	1634	1612	1584

2 The method

To arrive at these results, for the past several years, the NK administration has been using a simple linear extrapolation to do its school enrollment projections. The technique involves using an average of the differences between one year and the next to develop a linear relationship, and then extrapolating out from there. For example, the last five entries in the sixth grade are:

2000	2001	2002	2003	2004	2005
352	352	354	344	391	353

The differences between the years are:

00 - 01	01 - 02	02 - 03	03 - 04	04 - 05
0	+2	-10	+47	-38
0%	.57%	-2.8%	+14%	-9.7%

Averaging over the five years, this gives an average rate of change of .34% per year (the numbers may not appear to add because of rounding). So the sixth grade years from 2006 on out were therefore estimated at each 1.034 times the fifth grade population of the year before:

	2005	2006	2007	2008	2009	2010
5th	363	343	330	313	276	277
6th	353	364	344	331	315	277

The 2006 number for the sixth grade is 1.034 times the fifth grade number for 2005, and so on. The administration did not do any tuning of their predictions (though see section 3.1).

The kindergarten populations for each year are estimated by applying the average annual change to the kindergarten population of the year before. The special ed populations are done the same way.

3 The problem

The difficulty with this technique is twofold. First, it is applying an average of trends in one set of numbers (population changes within a grade level) to predict a trend in an entirely different thing (population changes from one grade to the next). Though it is possible that regularities in the data might justify such an approach, there is no obvious justification for predicting enrollment data this way. (Though see section 4.)

The other important problem is that this method takes no notice of any extrinsic trends that affect enrollment. According to the State planning department's analyses of population trends, the number of very young children in North Kingstown has been declining for a number of years, but their projections are that the decline has essentially bottomed out, and that populations are leveling off and will begin to rise again in a few years. But the school department's projections—which do not take into account *any* extrinsic information but only the previous year's kindergarten enrollment—show only a continuous decline in kindergarten population.

3.1 Mechanical projections

The consensus among school demographers is that no numerical technique, applied blindly to a school district's population patterns, will produce predictions of any value. Along with the random fluctuations of population, any district will experience changes in education law or changes in their program that have ramifications to the school population. For example, NK changed the kindergarten eligibility age a couple of years ago. Applying a formula in a way that didn't take account of this change would produce a poor prediction. (And in fact the administration projections correctly omit this year from the data used to make their forecast.)

Other important changes from recent years are the opening of the new high school, reconfiguration of the grade schools, economic trends that may have had an impact on families' choice of private schools, and demographic changes in the town.

4 An alternate

One possible improvement to the current method would be to use an average of the change between grades, rather than the average of the change in one grade from year to year. This would make sense, since you'd essentially be tracking a cohort of children from one year to the next. This is called the "cohort survival ratio" method (CSR), and is a popular technique, in use in many districts nationwide. The average can be calculated in several different ways, ranging from a straight mean of the previous several years, to an exponential average, which gives slightly better results due to its emphasis of recent years.[1]

We can check the accuracy of the two techniques by applying them to data we already have. That is, we can use our two techniques to project enrollments for past years, based on the data in earlier years, and see which one performs better when compared to the actual data.

What we find is that the cohort survival ratio method yields estimates substantially better than the current method, before taking into account extrinsic changes. The systems were tested against each other, and against the real data, by generating (purely mechanical) predictions for years where we have real data, too. For predictions in the years 2003-2005, the CSR method produced errors more than 40% smaller than the current method, even three years out (that's the 2005 prediction using 2002 data).

The following table shows the errors of the two

[1] An exponential mean weights the most recent measurements more heavily than the older contributions to the mean. The exact formula depends on how heavily you choose to emphasize the difference. In this case, we've used $t_n = t_{n-1}/2 + t_{n-2}/4 + t_{n-3}/8 + t_{n-4}/8$.

methods, compared to the actual data. The row labeled "2003" contains an accounting of the errors (the standard deviations, actually) for the projections for 2003, 2004, and 2005 compared to the actual data from those years, using data from the preceding years to formulate the projection coefficients.

	K Projections		Actual K Data	
	Admin	CSR	Admin	CSR
2000	23.0	24.9	22.8	23.7
2001	27.9	33.1	30.0	31.3
2002	28.6	24.4	25.2	22.2
2003	22.4	19.3	22.4	16.0
2004	22.4	18.1	22.3	16.5
2005	23.4	17.5	20.5	12.1

The first pair of columns uses the administration's projection method for the K numbers, and the second pair uses the actual kindergarten enrollments. An error of 12.1 means that you can expect most grades will be within 12 students of the projection. This is a 3.4% error (without tuning), quite a bit better than the 6.7% error of the administration method.

The administration's current method actually produced better estimates than CSR for tests over 2000-2001, but these years are problematic for testing purposes, since they saw a sudden and dramatic shift in the cohort survival ratios, largely due to extrinsic factors. For one thing, the new high school opened in 2001, and the attrition rates in 11th and 12th grades dropped significantly, while the jump up from 8th to 9th grades increased significantly. At the same time, attrition rates in the elementary grades increased. This, of course, has nothing to do with the high school, but is possibly explained by the stock market decline, which may have made private school a less appealing option for many families. Since these changes were sudden and due to extrinsic factors, and were not addressed specifically by the test, it's hard to regard the success of the current method as more than accidental. Certainly it would be unwise to mechanically extrapolate these trends into the future: North Kingstown isn't going to open a new high school every year.

5 Predicting the future

The CSR method is used to advance a cohort through the grades. But it does not help come up with the original kindergarten population to advance. The administration's current method for predicting kindergarten classes (and special ed) is to extrapolate from the trend of the previous four or five years. But this kind of extrapolation from intrinsic variables only works so long as no extrinsic changes happen beyond the schools' walls.

The raw census data, as reported by the US Census Bureau, seems to support the projection that K populations will decline over time. But when that data is combined with records of births, building permits, and other measures of immigration and emigration, one might see a different trend. In fact, demographers in the RI statewide planning department do see a countervailing trend taking over in the next few years. Each year, they publish population projections at five year intervals 25-30 years out from the present, for five-year age cohorts. That is, they predict the number of children aged from 0-4 years, the number from 5-9, and so on.[2] Here are the data for children in North Kingstown:

Age range	2000	2005	2010	2015
0-4	1804	1623	1583	1770
5-9	1994	1619	1638	1613
10-14	1917	1819	1628	1659
15-19	1770	1613	1659	1931

The projections are that recent declines in the numbers of small children in town will stop quite soon. The planning department numbers imply that the number of five-year-olds in town will be close to the same in 2010 as it is today. This makes the administration projections that K classes will decline by around 10% in the next five years seem improbable.

One could (and probably should) also use this data to modify the middle and high school cohort survival rates in the far-future predictions, but that hasn't been done in the current analysis.

The statewide planning numbers are hardly the last word in predicting the future. It's not clear, for example, how well they accommodate changing details of the housing or job markets (though they do try). For example, the effects of the runaway real estate market and the lack of affordable housing in suburban towns like North Kingstown is sure to have an important impact on whether young families can find their way to live in our town. Still, these projections are prepared through real survey data and analysis by demographic professionals, and many people and institutions in Rhode Island use them in their own projections. It seems curious to omit notice of this work in the administration enrollment projections.

Apart from the population trends, no other extrinsic factors were considered in the present analysis. The economic outloook for the next five years isn't great, but no one is predicting apocalypse on even the scale of the 2000-2001 stock market meltdown. The legal changes, which constitute the other big category of extrinsic factors, are also not accounted for, largely out of ignorance. Adapting them would be relatively straightforward, at least outside of the special education realm.

5.1 Special Education

Projections of special education demographics are notoriously difficult. Legal trends matter at least as much or more than the population trends. Still, it

[2] This document can be seen online at http://www.planning.ri.gov/census/tp154.pdf. North Kingstown is on page 63.

seems important to point out that because the special ed population jumped from 91 to 120 between 2001 and 2002, the administration projections are that the special education population will grow at an average pace of almost 6% per year for the next five years. This is in spite of the fact that the special ed population has actually *declined* every year since 2002 except 2005, when it increased by a single student.

Weighting the average to favor the most recent year gives a growth rate of about 2.5% for the special ed population. Whether this is any more plausible a number than the administration's is difficult to say, but it is more in line with recent experience than the blind extrapolation.

6 Discussion

School enrollment projections using the methods outlined here are presented below, in Table 2. (These are purely mechanical, and untuned, for comparison with the administration projections.) You can see by comparison with Table 1, that there are significant differences in the projected populations, even a small number of years in the future. The CSR method predicts almost 300 more students by 2010 than the administration currently expects. The vast bulk of the difference is from the difference in the elementary school population. The CSR method predicts an elementary (K-5) school population of 1877 students in 2010, compared to 1612 according to administration projections.

An important note here is that the bulk of the difference is from the change in extrapolation method, not the change in K populations using the Census data. The CSR method, applied to the administration's own K projections still shows around 200 more students by 2010.

Obviously, this projection is based on current policy. For example, establishing district-wide full day kindergarten could change these numbers substantially.

At the high school, the CSR method shows a few percent fewer students than the administration projects in the near term, and shows the population making up the difference and even growing slightly faster than the administration estimates by 2009.

Table 2: October 1 Enrollment Data, with CSR projections (one cohort is highlighted for illustration)

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
K	304	330	353	341	332	342	313	327	292	309	290	261	293	291	289	287	285	284	282
1	343	358	379	388	358	352	395	343	341	342	331	317	286	323	321	318	316	314	313
2	336	326	354	337	369	351	343	380	341	342	339	332	313	284	321	319	316	314	312
3	325	351	344	350	348	365	355	342	394	362	337	347	328	315	286	323	321	318	316
4	331	335	354	345	339	349	356	342	344	385	353	351	341	327	314	285	322	320	317
5	347	330	337	359	348	356	350	357	352	352	381	354	363	348	333	320	291	328	326
6	316	341	331	337	350	338	342	352	352	354	344	391	353	364	349	334	321	292	329
7	326	322	356	334	344	348	340	334	350	350	355	351	394	356	367	352	337	324	295
8	326	335	332	350	323	324	353	331	338	348	348	352	337	385	348	358	344	329	316
9	367	355	378	356	388	343	328	337	355	386	377	384	412	384	439	396	408	392	375
10	330	308	345	360	370	363	345	328	335	366	395	378	376	411	383	438	395	407	391
11	299	290	254	250	282	316	338	335	305	337	366	365	355	358	391	364	417	376	387
12	241	296	289	293	287	311	291	326	320	300	312	356	370	351	354	387	360	412	372
Total	4191	4277	4406	4400	4438	4458	4449	4434	4419	4533	4528	4539	4521	4497	4495	4481	4433	4410	4331
SPED	66	66	66	74	87	84	73	67	91	120	117	109	110	112	113	116	119	122	125
Grand	4257	4343	4472	4474	4525	4542	4522	4501	4510	4653	4645	4648	4631	4609	4608	4597	4552	4532	4456
HS	1237	1249	1266	1259	1327	1333	1302	1326	1315	1389	1450	1483	1513	1504	1567	1585	1580	1587	1525
Middle	968	998	1019	1021	1017	1010	1035	1017	1040	1052	1047	1094	1084	1105	1064	1044	1002	945	940
Elem	1682	1700	1768	1779	1762	1773	1799	1764	1772	1783	1741	1701	1631	1597	1575	1565	1566	1594	1584
K-5	1986	2030	2121	2120	2094	2115	2112	2091	2064	2092	2031	1962	1924	1888	1864	1852	1851	1878	1866