

New Methods and Sources of Estimating and Analyzing Student Attainment

Sherman Dorn, University of South Florida

Growth-corrected graduation rates

From Preston and Coale (1982), the probability of leaving a population from cause i :

$$\frac{l_i}{l_0} = \frac{\int_0^{\infty} D_i(a)e^{r(x)kx} da}{B}$$

where l_i/l_0 is the proportion of the equivalent synthetic population that would leave through cause i (usually cause-specific mortality), $D_i(a)$ is the instantaneous number of decrements (deaths) at age a through cause i , B is the number of births, and $r(x)$ is the instantaneous proportional growth rate for the population.

Translating into a school population, starting at any age α or the time since enrollment:

$$l_g \equiv \frac{\sum_{t=(r-\alpha)_+}^{t=(r-\alpha)_+ + \frac{1}{2}(t-r)} D_{gt} e^{r(x)kx}}{B}$$

where D_{gt} is the total number of graduations in the t th interval, $(r-t)_+$ captures the average growth and net migration/mortality rates in the interval, t_+ is the width (in years) of the interval, and B is the total number of birthdays at age α (or entering matriculants, for higher education data). (Also see Preston 1987.)

Residual net flow

Consider a single age interval and its various decrements and additions in a period:

$$AN = \text{Birthdays (x)} - \text{Birthdays (x+1)} + \text{Graduations} + \text{Net migration} - \text{Dropouts} + \text{Returnees} - \text{Deaths.}$$

Grouping the last four terms together as an unadjusted net flow,

$$NF_1 = AN - B_1 + AO_1 + G_1$$

where NF_1 is the unadjusted net flow, B_1 is the number of students who age into the i th interval (or have matriculation anniversaries corresponding to the beginning of the interval), AO_1 is the number of students who age out of the interval (or whose matriculation anniversaries push them out of the interval), and G_1 is the number of students who graduate while in the interval. (An equivalent equation is for rates rather than numbers.)

NF_1 is a residual net flow. The greater difficulty in identifying net dropping out for most student populations is distinguishing it from migration. My tentative strategy is to use average 7-10 net flow as the population baseline.

School life spent in grade(s) x

Combining the variable-rate expressions for life expectancy and the expected years of life spent in state G (with incidence $g(a)$ at age a), one can estimate the expected proportion of school life spent in some state G starting at age α :

$$G'_\alpha = \frac{\int_\alpha^{\infty} g(a)c(a)e^{r(x)kx} da}{\int_\alpha^{\infty} c(a)e^{r(x)kx} da}$$

where $c(a)$ is the instantaneous distribution of the (student) population at age a . As with growth-corrected graduation rates, one uses a discrete version for estimation purposes (not shown here). G'_α is useful when operationalized as time spent in a range of grades (for a whole student population) or in a specific year (where concerns are raised about retention in 9th grade).

Progress through grades ~ parity progression?

Progress from one grade to another is an ordinal sequence, but since skipping a grade is relatively rare, progress (or retention) can be treated akin to parity progression, and Schertmann's (2002) suggestions for deriving age-specific rates from successive cross-sectional surveys are applicable:

$$gp(x, n; 0, t) \equiv \left(\frac{1}{2n} - \frac{1}{2t} \right) (GL(x+n, t) - GL(x, 0)) + \left(\frac{1}{2n} - \frac{1}{2t} \right) (GL(x+n, 0) - GL(x, t))$$

where $gp(x, n; 0, t)$ is the grade progress for the age interval $(x, x+t)$ between times 0 and t and $GL(x, t)$ is the mean grade level of students aged x at time t . (Where the age interval and period are identical, the second term on the right drops out.) **Critical assumption:** $GL(x, t)$ is not correlated with transitions in and out of student population. This estimate is only appropriate for ages below when students noticeably begin dropping out.

See side panels for specific examples, details, and references. Research support from the University of South Florida College of Education. Contact: dorn@mail.usf.edu.

Calculation Details

Migration and mortality

Because of conflation among residual decrements and increments, especially for teens, it is impossible to estimate age-specific migration. A proxy for the underlying general combined migration and mortality rate (**I**) is the average of the 7-10 year olds' age-specific net-flow rates.

One can, in addition, explicitly model the effects of changing constant or age-varying migration (i.e., having under- or overestimated migration). For Virginia, 1998-99, the estimated likelihood of ever leaving the system with a standard diploma for 14-year-olds is as follows for constant migration and mortality rates varying between -0.01 and +0.05:

Migration+mortality	Probability of graduating with standard diploma
-0.010	81.7%
-0.005	79.9%
0	78.1%
0.005	76.3%
0.007*	75.5%
0.010	74.6%
0.020	71.3%
0.050	62.3%

*—the mean net-flow rate for 7-10 year olds.

Given the low mortality of school-age children, modeling the effects of age-specific mortality is unlikely to greatly affect graduation rates.

Distributing graduates among ages

Graduates are distributed proportionate to 12th-grade enrollment, aged from the prior fall to the spring (when most students graduate).

Net flows and teens' age intervals

Because of the uneven distribution of school-leaving within single years of age for teens, the net-flow calculations use a multi-age interval, to reduce the potential distortions to estimates of birthdays into and out of the interval. Otherwise, estimates for individual ages are unstable.

References

- Preston, Samuel H. 1987. Relations among Standard Epidemiologic Measures in a Population. *American Journal of Epidemiology* 126: 336-45.
- Preston, Samuel H., and Ainsley J. Coale. 1982. Age Structure, Growth, Attrition, and Accession; A New Synthesis. *Population Index* 48: 217-59.
- Schmertmann, Carl P. 2002. A Simple Method for Estimating Age-Specific Rates from Sequential Cross Sections. *Demography* 39: 287-310.

Estimated Net Flows, Virginia Public Schools, 1996-97 and 1997-98 school years

(1) Age interval	(2a) Enrollment	(2b)	(3) Graduates	(4) Raw growth	(5) Average enrollment	(6)(±) Birthdays	(7) Raw net flow	(8) Flow rate	(9) Adjusted rate
j	$N_j(t)$	D_{gr}	ΔN_j	N_j	$B_j(6^+)$ $\frac{AO_{j-1}(6^+)}{2}$	$\frac{(2a)_{j-1} + (2b)_j}{2}$	NF_j	$\frac{(4)+(3)}{-(6)+(6^+)}$	$\frac{(7)/(5)}{(8)-1}$
	'96-'97	'97-'98							
5	64,242	62,305		-1,937	63,274	76,681	13,116	0.145	
6	91,532	89,120		-2,412	90,326				
7	91,590	92,886		1,296	92,238	92,209	1,132	0.012	0.005
8	89,365	92,500		3,135	90,933	92,045	756	0.008	0.001
9	88,534	89,966		1,432	89,250	89,666	481	0.005	-0.002
10	87,138	88,894		1,756	88,016	88,714	294	0.003	-0.004
11	87,392	87,365		-27	87,379	87,252	273	0.003	-0.004
12	85,169	87,710		2,541	86,440	87,551	126	0.001	-0.006
13	84,140	85,102		962	84,621	85,136	465	0.005	-0.002
14	84,952	85,136		184	85,044	84,638	888	0.010	0.003
15	84,081	85,731	6	1,650	84,906	85,342	227	0.003	-0.005
16	82,583	83,745	341	1,162	83,164	83,913			
17	75,890	77,842	14,319	1,952	76,866	80,213			
18	25,966	27,628	33,042	1,662	26,797	51,759	} -14,895	-0.080	-0.087
19	5,544	5,479	12,544	-65	5,512	15,723			
20+	2,674	2,578	2,006	-96	2,626	4,061	-2,151	-0.819	-0.826

I: Assumed general net migration and mortality rate (average of net flow rates for 7-10 year olds): 0.007.

See **Calculation Details** sheet for the older teen age interval.

Source: Virginia Department of Education

Likelihood of Graduating, Virginia Public Schools, 1996-97 and 1997-98 School Years (for those having 14th birthdays as students)

(1)	(2a)	(2b)	(2c)	(3)	(4)	(5)	(6)	(7a)	(7b)	(7c)
Age int.	Diplomas (distributed as 12th graders)			Student growth rate	Migration-adjusted growth	Cumulative growth from age 14	Growth correction	Growth-corrected diplomas		
j	D_{gj}			r_j	$r_j - \mathbf{I}$	$\sum_{x=1}^j r_x (r_x - 1)^{-\frac{1}{2}} r_j (r_j - 1)$	$e^{\sum_{x=1}^j r_x (r_x - 1)^{-\frac{1}{2}} r_j (r_j - 1)}$	$D_{gj} e^{\sum_{x=1}^j r_x (r_x - 1)^{-\frac{1}{2}} r_j (r_j - 1)}$		
	Regular Special Certificate			(note a)	(note a)	(note b)	$e^{(5)}$	(2)·(6)		
	Regular Special Certificate							Regular Special Certificate		
14				0.002	-0.005	-0.0026	0.9974			
15	6			0.019	0.012	0.0009	1.0009	6		
16	332	5	4	0.014	0.007	0.0103	1.0103	335	5	4
17	13935	202	182	0.025	0.018	0.0226	1.0229	14,254	207	187
18	32156	466	421	0.062	0.055	0.0590	1.0608	34,111	494	447
19	12207	177	160	-0.012	-0.019	0.0768	1.0799	13,182	191	173
20+	1952	28	26	-0.037	-0.044	0.0453	1.0464	2,043	30	27
	60,587	878	793	Total			63,930	927	837	
	84,638	84,638	84,638	14th birthdays (see net-flow table)			84,638	84,638	84,638	
	Raw graduation proportions							Adjusted graduation probabilities		
	71.6%	1.0%	0.9%					75.5%	1.1%	1.0%
	Total	73.6%						Total	77.6%	

a: Growth rates are calculated from the net-flow table columns as $\ln[(2b)/(2a)]$, with \mathbf{I} from the same table. Here, $t=1$ for all j (including the open-ended interval).

b: Cumulative growth is calculated from age 14 through the middle of the age in question.

Source: Virginia Department of Education. Diplomas are from 1996-97 school year.

School Life Spent in 9th Grade, Virginia Public Schools, 14+ Year Olds, 1996-97 and 1997-98 School Years

(1)	(2a)	(2b)	(3)	(4)	(5)	(6)
Age int.	Proportion of students in 9th grade		Average proportion	Age distribution slice	Growth-adjusted age distribution	Proportion in 9th grade, weighted by adjusted age distribution
j			g_j	c_j	$c_j e^{\sum_{x=1}^{j-1} t_x (r_x - 1) + \frac{1}{2} t_j (r_j - 1)}$	$g_j c_j e^{\sum_{x=1}^{j-1} t_x (r_x - 1) + \frac{1}{2} t_j (r_j - 1)}$
			$\frac{(2a) + (2b)}{2}$	(note a)	(note b)	(3)*(5)
	1996 -1997	1997 -1998				
14	59%	57%	58%	0.233	0.232	0.135
15	35%	36%	36%	0.233	0.233	0.083
16	11%	11%	11%	0.228	0.230	0.025
17	4%	4%	4%	0.211	0.215	0.008
18	3%	2%	3%	0.073	0.078	0.002
19	3%	4%	3%	0.015	0.016	0.000
20+	20%	14%	17%	0.007	0.008	0.001
Total					1.013	0.255
Proportion of school life over 14 spent in 9th grade						
$\frac{\sum_j g_j c_j e^{\sum_{x=1}^{j-1} t_x (r_x - 1) + \frac{1}{2} t_j (r_j - 1)}}{\sum_j c_j e^{\sum_{x=1}^{j-1} t_x (r_x - 1) + \frac{1}{2} t_j (r_j - 1)}}$						
					0.252	

a: See net-flow table for student populations for each year and the average student population in each age interval.

b: See the net flow table for the growth-correction factor for each age interval.

Source: Virginia Department of Education.

Grade Progress, African-American Students Aged 6-13, Atlanta-Fulton Public Schools 1938-39—1940-41

(1)	(2a)	(2b)	(3a)	(3b)	(4)	(5)	(6)
Age	Average grade of enrollment	Estimated average grade of enrollment, at birthday	Longitudinal component	Correction component	Grade progress		
x			GL(x)	GL(x+n,t)-GL(x,0)	GL(x+n,0)-GL(x,t)	gp(x,n,t)	
			(averaging successive grade levels except 6yos)	(note a)	(note b)	.75 · (4) + .25 · (5)	(note c)
	1938	1940	1938	1940			
	-1939	-1941	-1939	-1941			
6	1.07	1.11	1.00	1.00	0.45	0.36	0.43
7	1.64	1.79	1.36	1.45	0.80	0.51	0.72
8	2.28	2.51	1.96	2.15	0.95	0.52	0.84
9	3.06	3.31	2.67	2.91	1.01	0.50	0.88
10	3.77	4.04	3.41	3.67	1.04	0.43	0.88
11	4.44	4.86	4.10	4.45	1.11	0.35	0.92
12	5.16	5.57	4.80	5.22	1.20	0.36	0.99
13	6.00	6.42	5.58	6.00			

a: Based on (3a) and (3b); see Schertmann (2002) for details. The calculation for 6-year-olds: $GL(7,1940-41) - GL(6,1938-39) = 1.45 - 1 = 0.45$

b: Based on (3a) and (3b). The calculation for 6-year-olds: $GL(8,1938-39) - GL(6,1940-41) = 1.36 - 1 = 0.36$

c: $n=1, t=2$. See Schmermann (and main poster) for the formula.

Source: Atlanta City and Fulton County Annual Reports of the Superintendent to the Georgia Department of Education, 1938-39 and 1940-41, in Georgia State Archives.

Substantive Questions

Historical Data (1930s-1960s)

Georgia county and city school-system annual-reports

- Age-grade tables by race late 1930s-mid 1960s
- Age-grade tables by sex from late 1950s-mid 1960s
- Graduates from 1940s
- 23 counties in four clusters: urban, northeastern hills, coastal plain, and Black Belt counties near Alabama

Goals of project

- Pilot methods
- Describe subregional variations in 20th c. closure of Black-White secondary enrollment gap

Grade Progress of 8-year-old Students by Race,

Five Georgia Counties, '38-39—'40-41 and '55-56—'56-57

County	'38-39—'40-41		'55-56—'56-57	
	White	Black	White	Black
Atlanta-Fulton	0.95	0.84	0.96	0.95
Bibb (Macon)	0.89	0.65	0.92	0.98
Glynn (coastal)	0.94	0.66	1.03	1.03
Liberty (coastal)	0.98	0.68	0.95	1.00
Randolph (Belt)	0.91	0.55	0.94	0.94

Contemporary Data

States with individual student databases/publish by age

- Florida
- Texas
- Virginia

Goals of project

- Pilot methods
- Compare to other estimates (esp. of graduation rates)
- Examine recent trends (high-stakes environment)

Virginia Comparisons, '96-97—'97-98 and '02-'03—'03-04

Statistic	'96-'97 —'97-98	'02-'03 —'03-'04
Average 7-10 net flow	0.007	0.020
Total PK-12 enrollment growth	0.011	0.015
Adjusted net flow, 16-19	-0.087	-0.069
Proportion of School Life 14+ spent in 9th grade	0.252	0.261
Likelihood of graduating with a standard diploma, 14+	75.5%	76.3%
Likelihood of graduating with other diploma, 14+	2.1%	5.4%

Extension of project to higher education...

If one looks at time since matriculation instead of ages, then these methods are equally applicable to higher education. Preston (1997) applied a similar translation of age for growth-corrected cancer mortality.