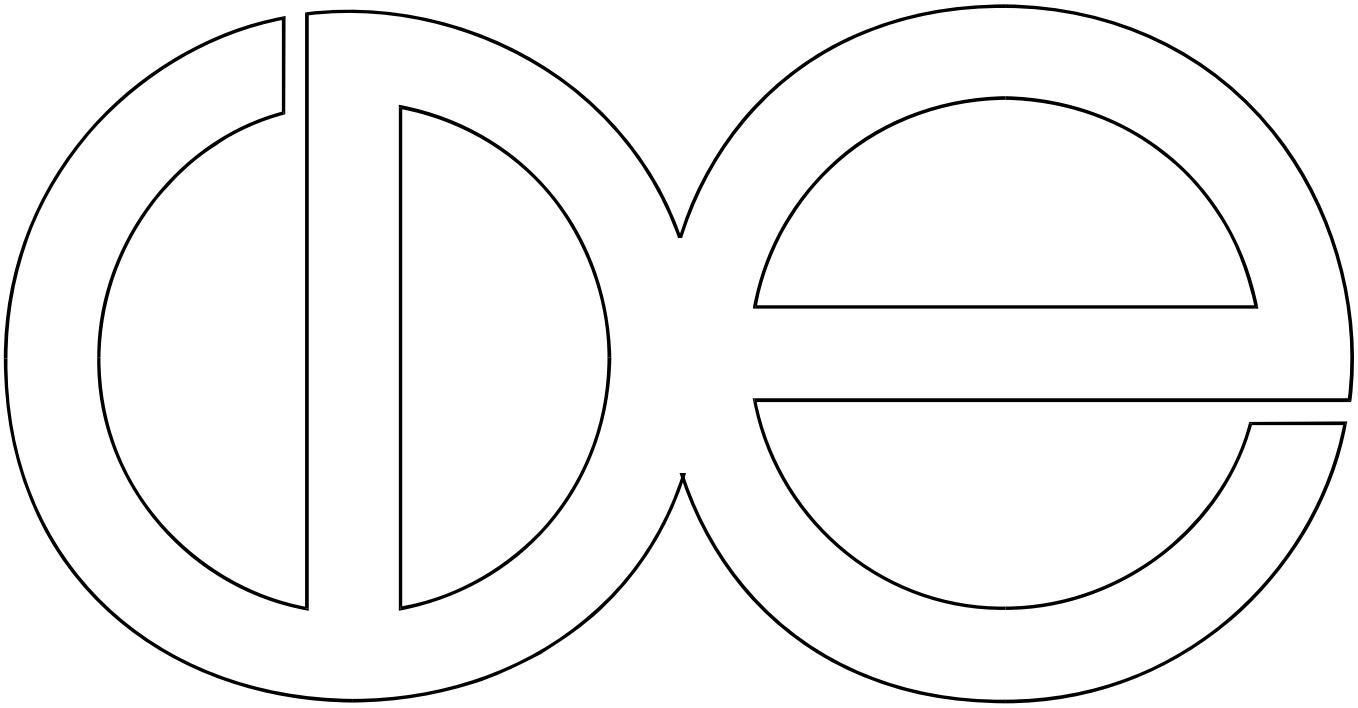
Center for Demography and Ecology

University of Wisconsin-Madison

Grade Retention in the Age of Standards-based Reform

Robert M. Hauser
Carl B. Frederick
Megan Andrew

CDE Working Paper No. 2007-04



Chapter 7

Grade Retention in the Age of Standards-based Reform

Robert M. Hauser
Carl B. Frederick
Megan Andrew

Center for Demography of Health and Aging
University of Wisconsin-Madison

Rev. February 15, 2007

This research has been supported in part by the Russell Sage Foundation, by the Vilas Estate Trust at the University of Wisconsin-Madison, and by a center grant for population research from the National Institute of Child Health and Human Development to the Center for Demography and Ecology at the UW-Madison.

Passed in January 2002, the No Child Left Behind Act (NCLB) codifies a long-standing shift away from an emphasis on school resources in education policy to an emphasis on standards-based reform. The shift stems in part from a large body of research suggesting that school resource inputs do not explain student achievement and, also, from sustained concern about the relative quality of education in the U.S. compared to other developed nations.[1] Several years before NCLB was instituted by the Bush administration, President Clinton declared his intention "to end social promotion."[2] This declaration led several state and local education agencies to institute new and, in many cases, test-based criteria for promotion from one elementary school grade to the next. Thus, most states began to move towards standards-based reform in the 1990s, well before the passage of NCLB.

State and federal movement towards standards-based reform and standardized testing is troubling. The past history of large-scale assessment provides many instances in which the use of tests to assess educational progress has encouraged their use to make decisions about students, whether or not the tests or their mode of administration are well-suited for high stakes decisions.[3] Because of the conjunction of standards-based reform and the use of test-based decisions about individual students as instruments of reform, we think it is reasonable to investigate whether retention in grade has increased in American schools from the late 1990s through the early years of the present decade. If this is the case, there are likely to be very serious, long-term effects on the educational

[1] Jencks et al. (1972); Hanushek (1986; 1996); U.S. Department of Education (1983).
[2] Clinton (1998); Steinberg (1999).
[3] National Research Council (1999).

attainment and subsequent life course of students who are held back.[4] That is, under past, current, and foreseeable educational regimes, students who are held back typically fail to catch up academically. Because they become over-age for grade, they are more likely to drop out and less likely to attend or complete college.[5] Moreover, it is well-established that minority and poor students are more likely to be held back in grade than majority and middle-class students, so students from deprived backgrounds are the most likely to suffer from an increase in grade retention.

In this paper, we assemble evidence about trends and differentials in grade-retention using data from the Census Bureau's October Current Population Survey and from state education agencies. Using descriptive statistics and multivariate models, we evaluate whether or not grade retention has increased under recent standards-based reforms, focusing in particular upon low-income students. Generally, we find a steady increase in grade retention, continuing an upward trend in grade retention rates since the 1970s. State and census data suggest that national increases in retention have generally been concentrated in the early grades and, in the case of the state data, occur in spurts. This increase has been steeper for children in the lowest income quartile from about 1998 until just before the NCLB Act was passed in 2002. However, retention rates temporarily spiked for children in the entire bottom half of the income distribution, implying fleeting, but deleterious effects of NCLB on retention rates for the poorest children Overall, in our multivariate models students in the lowest income quartile are 41 percent more likely to be retained than students in the highest quartile. In contrast, students in the second quartile are 14 percent more likely to be retained.

[4] Hauser, Pager, and Simmons (2004a); Hauser, Simmons and Pager (2004b).
[5] Andrew (2006).

It seems certain, then, that while the overall increases in grade retention have been modest since the 1990s, the increase is significant in real terms and has occurred to a large extent among low-income children, the very group NCLB is intended to help. Though standards-based reform policies at the federal and state level do not appear to have had extreme and lasting effects vis-à-vis grade retention, it remains possible that negative effects of using tests to make inappropriate decisions about individual student progress operate at the district or school level. In its first years of implementation, NCLB does not appear to have had direct and severe effects on national grade retention rates, but neither has it reduced the upward trend. We begin with a brief review of the literature.

Research on Grade Retention

Unfortunately, no national data-collection mechanism or repository for promotion or retention statistics exists, and most data on retention are based on indirect measures or limited samples. Extending from this general paucity of grade retention data, no national educational information system monitors the extent to which tests are used to make promotion or retention decisions. National trends in grade retention rates are mainly estimated using a proxy measure, the share of children who are below the modal grade for their age. There is a mix of uncertainty, approximation, and speculation about the prevalence of grade retention in American schools with often contradictory statistics.[6] Reliable and recent estimates based data from Current Population Surveys suggest, "At least 15% of pupils are retained between ages 6 to 8 and ages 15 to 17, and a large share of retention occurs either before or after those ages."[7] These estimates also suggest the prevalence of retention increases substantially over students' educational careers,

[6] For example, compare Karweit (1999) and Eide and Showalter (2001).
[7] Hauser, Pager, and Simmons (2004a:98).

growing by about 10 percentage points by ages 9 through 11 and by about another 5 percentage points by ages 12 through 14.

A number of correlates of grade retention have been identified. Research has established stark gender differences in grade retention; boys are more likely to be retained than girls at every level of the K-12 educational system.[8] Racial and ethnic differences in retention rates are also prominent. The National Research Council reports that shares of students below modal grade for age observed in Census data are relatively similar among racial and ethnic groups at ages 6 through 8, but they are 5 to 10 percent higher for blacks and Hispanics than for whites just three years later at ages 9 through 11.[9] By ages 15 through 17, the rate of being below modal grade for age ranges from 40 to 50 percent among blacks and Hispanics, but it is much lower among whites at 25 to 35 percent. Among the same age group, the rate among children in the lowest income quintile is 20 percentage points higher than those in the middle 60 percent of the income distribution and over 35 percentage points higher than those in the highest income quintile. However, the income differences may be decreasing after the mid 1990s.[10]

Beyond gender and race-ethnic differentials, a higher incidence of retention is associated with a disadvantaged socioeconomic background, living in a single-parent home, living in a central city or in the southern Census region, having been born to a teenage mother, having parents with low measured IQ and education, and having parents with a health or behavioral problem.[11]

---

[8] Byrd and Weitzman (1994); National Research Council (1999); Dawson (1998).
[9] National Research Council (1999).
[10] Hauser (2004:290).
[11] Corman (2003); Hauser et al. (2004a); National Research Council (1999); Hauser (2004).

Despite the intended benefits of grade retention, research suggesting it has positive effects on student outcomes is sparse and subject to methodological and analytical problems that render results suspect. For example, Alexander, Entwisle and Dauber argue that grade retention in the early primary grades halts the downward slide of low-achieving students and prepares them to succeed in later grades.[12] The positive effects appear for students who are retained once and only after the first grade in a cohort of 1000 Baltimore students. Over time, even among this select group of students, the apparent positive effects of grade retention diminish, so retained students reap no long-term positive benefits. Critics of Alexander et al.'s work have argued that the supposedly positive effects of grade retention actually reflect regression to the mean, and represent no real benefit in the first place.[13] Moreover, Alexander et al.'s data show significant detrimental effects of grade retention on a much larger number of students retained in the first grade. Finally, their analyses control for some student characteristics subsequent to the retention decision.[14]

In a second example, Jacob and Lefgren also find positive effects of grade retention using a regression discontinuity design with data from students in the Chicago Public Schools.[15] In their study, retention substantially increased the achievement of third graders in reading and math, but the effect on reading becomes insignificant two years after retention, and the effect on math achievement is barely significant at the $p = 0.05$ level and falls by half a year after retention. A number of aspects of the authors' research give reason for pause despite the sophisticated research design. While

[12] Alexander, Entwisle, and Dauber (2003).
[13] Shepard et al. (1996; 1998); Alexander (1998); Hauser (2005).
[14] Hauser (2005).
[15] Jacob and Lefgren (2004).

regression-discontinuity can help account for selection bias, only the effect of retention on math achievement for students originally retained in the third grade remains significant after two years time. Even at that, the effect is substantially smaller and barely significant despite the considerable sample size. This suggests that Jacob and Lefgren's findings may reflect regression to the mean, as in other research that has found a nominally positive effect of retention. Moreover, these results only speak to the effects of retention among a narrowly defined population of students: largely minority and low socioeconomic status; testing at about the 20th percentile of the national distribution in a single urban school district with strong incentives for test performance; and where a variety of programs are in place to enhance student achievement. For example, these included after-school tutoring for retained students beginning a third of the way through the study. Thus, results may not reflect the effects of retention so much as extra tutoring, which can be implemented without retaining a student, and, all else aside, may not generalize to the larger population.

In contrast to Alexander et al. and Jacob and Lefgren's work, most research suggests that grade retention is associated with negative student outcomes. Meta-analyses of retention effects have been particularly helpful in isolating the extent to which retention may help or harm students, and they generally show that grade retention is harmful to students. Prominent meta-analyses include those of Holmes and of Jimerson.[16]

Holmes assessed 63 studies spanning almost 90 years from 1900 through the 1980s. When promoted and retained students were compared one to three years later, the retained students' average levels of academic achievement were at least 0.4 standard

[16] Holmes (1989); Jimerson (2001).

deviations below those of promoted students. In these comparisons, promoted and retained students were the same age, but the promoted students had completed one more grade than the retained students. Promoted and retained students were also compared after completing one or more grades, that is, when the retained students were a year older than the promoted students but had completed equal numbers of additional grades. Here, the findings were less consistent, but still negative. When the data were weighted by the number of estimated effects, there was an initially positive effect of retention on academic achievement after one more grade in school, but it faded away completely after three or more grades. When the data were weighted by the number of independent studies, rather than by the estimated number of effects on achievement, the average effects were negligible in every year after retention. Of the sixty-three studies reviewed by Holmes, fifty-four yielded overall negative effects of retention, and only nine yielded overall positive effects. Holmes concluded, "On average, retained children are worse off than their promoted counterparts on both personal adjustment and academic outcomes."[17]

Jimerson updated Holmes' classic meta-analysis using 20 studies of the association between retention and academic achievement and socio-emotional adjustment spanning 1990 to 1999.[18] Jimerson combined studies comparing same-grade and same-age outcomes and does not compare differences in results by grade and age as Holmes does. However, Jimerson's results are consistent with Holmes, suggesting the associations between grade retention and student outcomes and characteristics have been quite stable over time. On average, retained students scored 0.39 standard deviations lower than students who were not retained on various academic achievement measures.

[17] Holmes (1989:27).
[18] Jimerson (2001).

Students who were retained scored 0.22 standard deviation units less than similar students who were not retained on socio-emotional outcomes.

Since the publication of Jimerson's meta-analysis, evidence against using retention as the primary method of remediation continues to emerge. Pagani et al. find negative effects of retention on academic achievement among children in Quebec using an autoregressive model of academic achievement.[19] Allensworth finds that low achieving eighth graders in Chicago who failed to pass through to the ninth grade were more likely to drop out of school.[20] In another regression discontinuity design using the same data as Jacob and Lefgren, Nagaoka and Roderick find no difference between third grade retainees' test scores and their regularly promoted peers two years after being retained.[21] Hong and Raudenbush use data from the Early Childhood Longitudinal Survey – Kindergarten Cohort (ECLS-K) with multilevel propensity score stratification and find that retaining children in kindergarten leads to lower academic achievement in the next school year.[22]

Evidently, grade retention is a common and prominent means to remedy academic failure in the U.S. educational system. Yet, the practice does not appear to bestow many of the intended benefits upon students who are retained in grade and incurs considerable costs. Specifically, one cost-benefit analysis suggests grade retention costs upwards of $3.5 billion a year, assuming a conservative 1 percent retention rate annually.[23] Finances aside, better solutions are available. The NRC report on high stakes testing concludes, "Neither social promotion nor retention alone is an effective treatment, and schools can

[19] Pagani et al. (2001).
[20] Allensworth (2004).
[21] Jacob and Lefgren (2004); Nagaoka and Roderick (2004).
[22] Hong and Raudenbush (2005).
[23] Eide and Goldhaber (2005:198).

use a number of possible strategies to reduce the need for these either-or choices—for example, by coupling early identification of such students with effective remedial education."[24]

Given the weight of the evidence of deleterious effects of grade retention on students in particular and society in general, it is logical to ask: In the current political climate, with its emphasis on standards-based reform, is grade retention increasing? The brief history of NCLB makes it difficult to discern its effects on trends in retention practices with the available data, either nationally or at the state level. However, because many states had already moved toward standards-based reform in the 1990s, we can get some traction on this question by looking at trends in retention during the years preceding the passage of NCLB as well as those immediately following.

Analytic Strategy

We begin with data from October Current Population Surveys, assessing trends in grade retention since 1996 for all students and by race, gender, and socioeconomic status. We then turn to data collected from the state educational agencies, comparing differences and similarities in the two sources of data. We ask three basic questions:

1) Given the diffusion of the standards-based reform within education and the subsequent use of tests in retention decisions, have grade retention rates increased over the past decade?
2) Are there different trends in grade retention in subpopulation groups, especially those defined by socioeoconomic status, or at specific grade levels?

[24] National Research Council (1999:278).

3) Are there regional and state variations in grade retention trends across time and, particularly, in the past decade? Do these data corroborate general trends observed in the census data?

Data and Methods

In our first analyses, we use data from October School Enrollment Supplements to the Current Population Survey. Since 1994, the survey has collected data on children's grade of enrollment in the previous year as well as the current year. This allows us to construct a direct measure of the probability that children are retained in grade. We limited our sample to people aged 5 to 20 in each survey from 1996 through 2004.[25] Additionally, we eliminated those who were enrolled below kindergarten in the year prior to the survey, those who were enrolled above twelfth grade in the current year of the survey, twelfth grade repeaters,[26] and observations that were missing one or both of the enrollment variables. This yielded an unweighted sample of 230,696 cases. Finally we eliminated 2,177 cases (1.0 percent) reporting a grade progression other than single retention, normal promotion, or double promotion (skipping one grade) for a final analytic sample of 228,519 cases.

For the logistic regressions, we have further restricted the age range to ages 5 through 17 in order to include data on parent's education and occupation.[27] Cases with missing data are dealt with in two different ways, depending on the reason data were missing. If an observation was missing income and head of household's education, it was

[25] We excluded data from the first two available years because of data quality issues.
[26] We must omit students who were enrolled in the twelfth grade in the previous year because our other exclusions eliminate all twelfth graders who were *not* retained. This would leave us with no variation with which to estimate our models.
[27] Above age 17, youth are less likely to live with their parents, thus breaking the link between school enrollment and social and economic characteristics of householders.

dropped from the sample. For observations missing data on occupation and spouse's education – because a household head was not in the labor force or was not married – we used a dummy variable adjustment procedure, which has been shown to be unbiased where data are missing because they could not exist.[28] Eliminating ineligible observations as well as those with missing data yielded a final analytic sample of 154,112 cases for the logistic regression analyses.

In addition to the year of the survey and the previous year's grade, we include both demographic and social background covariates. The demographic variables include gender, race, region, urban/rural residence, and the number of children living in the household. There are four categories of race-ethnicity: non-Hispanic black, non-Hispanic white, Hispanic (any race), and other. Region is also divided into four categories: Northeast, South, Midwest, and West. The urban residence dummy pertains to students who live in major central cities. It is important to note that children living in the same household are not necessarily siblings because the CPS is a household rather than a family survey.

The social background covariates include logged household income, the household head's education and occupational status, the spouse's education and occupational status, and a dummy variable indicating whether the household head (and spouse) own their home. Again, the household head and spouse are usually, but not necessarily, the child's parents. We divide the educational attainment measures into two variables in order to tease out piecewise linear effects of parental education before and after the high school to college transition.

[28] Allison (2002).

We also collected data on grade retention rates from state educational agencies. We build on earlier retention data through the mid-1990s provided by the states to the National Research Council's Committee on Appropriate Test Use.[29] We contacted states' educational offices and requested whatever grade retention trend data, particularly since the 1990s, were available, and we compiled these data together with the existing National Research Council data. Not all states collect these data or responded to our request. However, we are able to provide descriptive data on grade retention rates since the 1990s for 14 states across all regions of the U.S.

Incidence of Retention: CPS

We begin with a discussion of analyses employing CPS data. Overall, 2.73 percent of the CPS sample reported being retained in the year preceding the survey, and 0.34 percent reported experiencing a double promotion or 'skipping' a grade.[30] Figure 1 plots the trend in overall retention rates by year along with retention rates by race. Focusing on the total retention rate trend line, the absolute change is not very large; the largest difference between annual rates is 0.63 percent. However, there is a clear increasing trend from the beginning of this time period until the retention rate appears to level off near 3 percent per year after 2003.

FIGURE 1 ABOUT HERE.

How big is a difference of less than one percent at this level of retention? Assume that the difference between the maximum and minimum proportion of students retained is real and that the chance of being retained is independent and constant across grade levels. Over the twelve years from kindergarten through 11th grade, 75.5 percent of students are

[29] National Research Council (1999:138-146).
[30] For analytic purposes, the double-promotions are combined with normal progressions.

expected to reach their senior year of high school on time if subject to a 2.31 percent retention rate. This figure drops to 69.9 percent never retained if 2.94 percent of students are retained in each grade.

The assumptions made in these simulated examples are surely violated. For example, analyses of data from NELS88 show that children are rarely retained twice.[31] However, these illustrative estimates are not much higher than previous estimates of the proportion of children who have ever been retained.[32] Table 1 shows the number of children retained in each grade level in each year. Again, assuming that the chance of being retained is independent across grade levels, we can construct period retention rates for the synthetic cohort of students who progress through school at the observed rates of retention in each year. These estimates are displayed in the bottom row of Table 1. As in the calculation based on annual retention rates, the predicted proportion of students who enter 12th grade on time declines steadily from 75.8 percent in 1996 to 70.3 percent in 2005.

TABLE 1 ABOUT HERE.

The overall proportion of students retained varies across grade levels, as shown in Figure 2. In kindergarten 4.6 percent of students are retained in this time period, and in first grade more than seven percent of students are retained. This proportion decreases by two-thirds in second grade and hovers between one and a half and two percent until eighth grade, excepting a jump to 2.3 percent in seventh grade. Ninth graders report the highest probability of retention outside of the first two years of schooling at just over three percent, but this is still less than half the probability of first graders. In tenth and

[31] Andrew (2006); see also Shepard and Smith (1989:8).
[32] National Research Council (1999).

eleventh grades, the proportion retained falls back down to around two percent. First-grade retention stands out as considerably higher than in other grades. Though no research has assessed why this is the case, differences in academic preparation in the home prior to school entry are probably their source. These differences may not be as evident in kindergarten, which is not universal and has a considerably different, less academic curriculum, hence the jump in first grade retention.

FIGURE 2 ABOUT HERE.

Retention and Race

Figure 1 also displays the change in retention probabilities over time, disaggregated by race. Black students are at the highest risk of being retained, followed closely by Hispanics. White students and those of other races have the lowest percentages retained over time. The gap between black and white students is fairly constant until 2000 when black retentions begin to increase relative to those of whites. Likewise, the gap between Hispanics and whites is fairly constant, but with a slight increase from 1999 to 2001.

Blacks and Hispanics also experience more retention at specific grade levels than whites and students of other races, as shown in Figure 2. The gaps are most severe in kindergarten and first grade. After that the percentages seem to converge through the elementary and middle grades until they slightly diverge once more in high school. Overall, retention rates for minority children are more variable, particularly children in the "other" category. This variability is largely a function of the smaller sample of minority children in the CPS. Blacks and Hispanics each make up 13 percent of the

sample, while children of other races only make up 6.4 percent. The remaining two thirds of the sample are non-Hispanic Whites.

Retention and Gender

FIGURES 3 and 4 ABOUT HERE.

Figure 3 disaggregates the yearly trends in retention by gender. The aggregate trendline is included in Figure 3 (as it is in Figures 4 through 6) for reference. This confirms findings of previous studies that boys experience more retentions than girls. This finding appears in every year. The gap in the percentage of retentions between boys and girls is widest during the middle of this period, especially in 1999 and 2000; other than that the trends for boys and girls are remarkably parallel.

Figure 4 shows the retention rates by gender and grade. While boys are consistently retained with greater frequency than girls across all grades, the gap begins to widen in the sixth grade—perhaps a consequence of puberty—and is most severe in the ninth grade. This gap closes by the 11th grade.

Retention and Income

FIGURES 5 and 6 ABOUT HERE.

In figures 5 and 6 the trends are disaggregated by income quartile. There are no surprises here: Children who come from the most prosperous families experience the least retentions. The biggest gap between adjacent categories over time is that between the lowest and second quartile. The smallest such gap is between the third and fourth quartiles, possibly indicating a threshold effect of family income. Again figure 6 shows a now-familiar pattern – the biggest income differences in retention rates occur during

kindergarten and first grade; rates converge during the middle grades and then diverge again during high school.

Multivariate Analysis

Next we turn to our multivariate analysis. In this portion of this analysis, we employ two basic models: a standard logistic regression and a logistic regression model with proportionality constraints. These constraints can provide a clearer picture of how students' characteristics behave similarly in regards to predicting grade retention. In combination, these multivariate models test the descriptive patterns discussed above. We begin with the standard logistic regression. This model estimates retention in the year prior to the survey as a function of the grade in which a student was enrolled in the previous year, the year of the survey, and both demographic and social background variables:

$$\text{logit}(\text{P}(Y_i = 1)) = \alpha_j X_{ij} + \alpha_k W_{ik} + \sum_h \beta_h Z_{ih} \quad (1)$$

where Y_i indicates whether student *i* reports being retained, α_j is the intercept for year *j*, indexed by the dummy variable, X_{ij}, α_k is the intercept for grade *k*, indexed by the dummy variable, W_{ik}, and β_h is the coefficient of demographic and social background covariate Z_{ih}. Estimates from this model are listed in the first panel of Table 2. As written, we have explicitly divided the effects of the various groups of variables into two classes. The α s are additive intercepts that adjust for the mean levels of retention in each grade level and year, and the β s are the effects of individual characteristics, holding the grade and year specific retention levels constant. In both of the following models, the

social background and demographic characteristics are expressed as deviations from their means to facilitate interpretation.

The intercepts for year of survey follow the trend in Figure 1. The intercept for 1997 is not reliably different than the omitted category (1996) but the differences with 1996 are significant in each subsequent year. The intercept for each grade is significantly different from that of the omitted category, kindergarten, and the effects of previous year's grade also follow the trends shown in Figure 2. Average first graders are more likely to be retained than kindergartners. The rest of the grade levels have lower intercepts. The intercepts for students in ninth, tenth and seventh grades in order are the next higher intercepts relative to kindergarten with the other covariates held constant.

The effects of race-ethnicity are in the expected directions but only the black-white difference is significant at even the 5 percent level. The exponentiated coefficient indicates that black students are 31 percent more likely to be retained during this period than white students. The difference between boys and girls is also significant. Boys are 26 percent more likely to be retained than are girls. Additional children living in the household also increase one's chances of experiencing retention, all else being equal; each additional child increases the odds of retention by 7 percent.

The differences between places of residence are all significant. Relative to children in the Northeast, Southeasterners are 27 percent more likely to be retained. Both Midwesterners and Westerners are 24 percent less likely to be retained compared with their Northeastern counterparts. Residing in a major central city increases the odds of being retained by 37 percent relative to students in smaller cities, suburban, and rural areas.

Only five of the social background variables significantly affect the likelihood that a student was retained in the past year. Each year of father's postsecondary education reduces the likelihood of retention by 4.4 percent. Students whose parents own their home are fifteen percent less likely to be retained. An increase of ten points in the spouse's occupational status score decreases the odds of retention by almost five percent. Income is also negatively associated with retention, and, again, we see evidence of the threshold effect between students in the lowest and second quartile that we saw in the bivariate relationships. Students in the lowest income quartile are 41 percent more likely to be retained than students in the highest quartile. Students in the second quartile are 14 percent more likely to be retained. Students in the third highest quartile do not have significantly different odds of retention compared to students in the highest income quartile.

INSERT TABLE 2 ABOUT HERE.

We now turn to the second model, the logistic regression model with proportionality constraints on effects of the covariates (LRPC). We find this model useful because it allows us to estimate important differences and similarities in the effect of students' characteristics across grade and time in a single equation. Alternatively, one could run separate, fully interacted models by year and grade of retention that allow the slopes of all covariates to vary freely. However, this strategy is cumbersome, and it produces an overwhelming number of coefficients to interpret. Moreover, effects of the covariates may not vary substantially. If this is the case, the fully interactive models will not provide any more information than the LRPC model.[33] By estimating a single-

[33] In supplementary analyses, we compared the LRPC model presented here with a model including all possible interactions between student characteristics, grade of retention, and year. The LRPC model

equation model, we assume changes in the association between retention and demographic and social background characteristics behave similarly across time, excepting a proportional or scalar difference. In this way, we provide a more parsimonious characterization of differentials in grade retention. The LRPC model (equation 2) is similar to the model in equation 1 except it constrains grade level and year interactions for each demographic and social background covariate:[34]

$$\text{logit}(\text{P}(Y_i = 1)) = \alpha_j X_{ij} + \alpha_k W_{ik} + \sum_h \beta_h Z_{ih} + \lambda_j \left(\sum_h \beta_h Z_{ih} \right) + \lambda_k \left(\sum_h \beta_h Z_{ih} \right) \qquad (2)$$

In this model $\alpha_j, \alpha_k, \beta_h, X_{ij}, W_{ik},$ and Z_{ih} are defined as above.[35] The coefficients λ_j and λ_k scale the effects of the social background and demographic covariates by the same proportion for each grade level, *j*, and year, *k*. These estimates are listed in the second panel of Table 2. According to standard model fit statistics, the LRPC model better describes the data. Compared to the simple logistic regression, the LRPC model reduces the AIC by 13.15 and is the preferred model according to the likelihood ratio test.[36] There appear to be real, systematic differences in the effects of the covariates across grades and years, but the evidence in favor of the LRPC model is not unequivocal.

The interpretation of the LRPC coefficients may be counterintuitive. As mentioned above each λ is a scalar that proportionally increases or decreases the

---

presented here is preferred by the AIC (a difference of 278) and the BIC (a difference of 3848) but not by the LL ratio test (a difference of 440 with 359 *df*). The last finding is not surprising, given the exceptionally large size of the sample; see Raftery (1995) and Weakliem (1999).

[34] Hauser and Andrew (2006).

[35] As with all logistic regression models, this model is identified by fixing the variance of the error to a constant ($\pi\gamma/\sqrt{3}$). Thus, true differences across groups are empirically indistinguishable from heterogeneity in the conditional variance by groups. Scaling effects, as with all coefficients in this and similar models, should be interpreted with this caveat in mind; see Allison (1999) and Mare (2006).

[36] Due to the large sample size, the LRPC actually increases BIC by 176. The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are similar to a log-likelihood test but provide a more conservative estimate of model fit for nested and non-nested models; see Burnham and Anderson (1998) and Raftery (1995).

magnitude of the demographic and social background coefficients. Because we treat grade level and year as nominal categories, the β-coefficients are the effects for the omitted category – kindergarteners in 1996. The effects for other grades and other years are obtained by factoring the linear predictor out of equation 2. This yields a scalar multiplier of $1+\lambda_j+\lambda_k$ for each regression coefficient. Thus, negative signs do not imply a change in the direction of the effect unless $\lambda_j+\lambda_k<-1$, which is never the case in these estimates. An example may be helpful. For the sake of simplicity, consider a slightly revised model from the one estimated here, a model with only a proportionality constraint for each year. Assume we are specifically interested in the scalar for 1997, λ_{1997}. Recalling that the total scaling effect in a model must be equivalent to the sum of the specified scalars plus one, if λ_{1997} were equal to -0.34, the magnitude of the β-coefficient of each variable Z_{ih} in 1997 would be two thirds the size of the corresponding coefficient in 1996, i.e., $1+(-0.34)=0.66$. That is, the proportionality constraints imply that slopes of the demographic and social background characteristics change similarly across grade level and across time.

The composite scalars are listed in Table 3. Because they combine additively, the trends are easily summarized in Figures 7 (year) and 8 (grade). All else equal, the magnitudes of the coefficients on social background characteristics are generally not reliably different from kindergarten levels, but there are three statistically significant positive spikes – at grades three, seven, and nine. These grades are indicated with circles in Figure 8. The increased effects of social background at these key promotional gates, grades three and nine, indirectly suggest that retention decisions are being made with

high stakes tests in mind, either as a result of poor test scores or in anticipation thereof. The seventh grade scalar is only marginally different from that in kindergarten ($p = 0.049$).

TABLE 3 ABOUT HERE

FIGURES 7 and 8 ABOUT HERE

As above, the trend line in the scalar values over time must be interpreted with caution because only three years – 1997, 2003, and 2005 – are statistically distinguishable from the omitted category, 1996, at the five percent level. These years are indicated with circles in Figure 7. Holding grade-level constant, there is a large drop in the scalar between 1996 and 1997. Further analysis with additional years of data will be necessary to learn whether the 2004 peak is an aberration or may signal a longer upward turn in the effects of demographic and social background characteristics on retention rates. Excepting the spike in 2004, the LRPC model indicates that a reduction in ascriptive barriers to promotion – in which the influence of social background is generally decreasing after 2001 – is concurrent with, but not necessarily a direct result of NCLB.

We use the predicted differences between students from households in the lowest income quartile and those in the highest to illustrate the substantive interpretation of the composite scalars in Table 3. The baseline coefficient for children in the lowest income quartile from Table 2 is .34. This means that low income kindergartners in 1996 were 41 percent more likely than high income kindergartners in 1996 to be retained the following year. The implied coefficient for kindergartners in 1997 is 0.34*0.673=0.23. Therefore, low income kindergartners the next year were only 26 percent more likely to be retained.

Similarly, the implied coefficient for third graders in 2001 is 0.34*1.46 = 0.50, which indicates an increase in relative retention rates of 65 percent between low and high income students (exp(0.5) = 1.65).

It is important to emphasize that increased effects of social background variables do not imply that retention decisions are made directly on the basis of race, gender, or socioeconomic standing. Instead, they probably reflect standards-based reform, both before and after the passage of NCLB, that is characterized by an increase in the amount of testing, combined with well-documented racial and socioeconomic test-score gaps. Other evidence suggests that new as well as older tests are used to make retention decisions. Since the CPS data do not include any measure of academic ability or performance, these findings are consistent with a retention system based increasingly on test scores, but we lack direct evidence of that. However, large-scale assessments are not widely used in kindergarten, first, or second grades, so other mechanisms must account for trends at those grade levels. One possibility is that the anticipation of later, test-based retention accounts for increased retention in the primary grades.[37]

Incidence of Retention: State Reports

Generally, grade retention rates in the CPS data have increased very modestly since the mid-1990s amid the shift to a more rigorous standards-based regime. Do state-level reports of retention data corroborate this picture of grade retention trends? To answer this question, we collected data from state educational agencies and combined them with existing data on grade retention rates by state.[38] These data are available upon request for some states by year and grade. Perhaps not surprisingly, state retention rate

[37] Jacob (2005).
[38] We have not attempted to reconcile state-to-state differences in the definition of retention.

data confirm many of the broader trends observed in the CPS data. Overall growth in state retention rates is marginal though some slight upward growth is observed over time in some grades in some states. Sudden spikes are observed in retention rates for some states, but retention rates generally return to previous levels in the next school year. Absolute levels of grade retention by region reflect noted disparities in the literature: Southern states have considerably higher rates of retention than other states. Data on the state of New York are limited, but it appears that, at least in the case of 9th grade retention, rates for New York are comparable to those in the South, that is, higher than in neighboring states in the Northeast and in the Midwest.

Some notable patterns are apparent in the state data, and particular states present interesting case studies in that regard. For example, when growth in retention did occur from the early 1990s through the present, this growth was often concentrated in kindergarten and the early primary grades. Tennessee and Texas both evince this pattern. Based on trend data for Tennessee for the 1980s through the 2003-04 school year, we see that kindergarten retention steadily grows over time with a 1 percentage point jump between the 2000-01 and 2001-02 school years. The state maintained a similar rate of retention in the next year. Texas shows a similar pattern to that of Tennessee with about a 1 percent increase in kindergarten retention between the 1998-99 and 2003-04 school years. Texas also saw nearly a 1 percent increase in the retention rate in second grade in the same time span, 40 percent of which occurred between the 2000-01 and 2001-02 school years. In approximately the same time span, Texas also saw about a 1.5 percent increase in grade retention in the third grade as well. Connecticut similarly exhibits a 1 percent increase in kindergarten grade retention in the same time period. In the available

data, grade retention rates at higher grades show overall decline or stability in these states.

In contrast, retention rates in Kentucky in the primary grades remain relatively stable across time for the data available, but grade retention increases in the secondary grades. Growth in retention rates for the secondary grades is characterized by unusual spikes in this state. For example, grade retention in the 9th grade alone jumps a comparatively large 2 percent in the 1998-1999 school year. In the following school year (1999-2000), the same cohort of students was subject to a 1.2 percent increase in retention. Alabama follows a similar trend line. In the available data, grade retention decreases in the primary grades but increases after about the 2000-01 school year for all secondary grades. In the 9th, 10th, and 12th grades, this increase was steady and retention rates did not return to the previous levels in subsequent years. In the 4-year span from the close of the 2000-01 school year to the 2004-05 school year, retention increased 1.5 percent in the 9th grade alone in Alabama.

Another common pattern in the state data was a sharp spike in retention rates in one year, followed by a return the previous level in the next school year. This pattern was observed in Wisconsin. Overall, Wisconsin exhibited the low and relatively stable retention rates that are characteristic of other Midwestern states, but in stark contrast to many Southern states. Despite the relative stability of retention rates in the state of Wisconsin, there is an unusual spike in 8th grade retention in the 2000-01 school year, when the retention rate unexpectedly doubled, jumping 1 percent. Yet, the 8th grade retention rate returned to normal within the next year or so. The same pattern occurs in Maine in multiple grades. Between the 2000-01 and 2001-02 school years, grade

retention jumped about 2.25 percent in kindergarten before returning to the previous level of retention in the following year. Grade retention jumps 2 percent in the 3rd grade in the 1999-2000 school and similarly returns to the previous level in the next year. At the same time, overall retention rates in Maine show little movement. North Carolina also shows similar spikes in grade retention rates in the 3rd, 5th, and 8th grades—key gateway grades—in the 2001-02 school year.

Ohio represents a striking example of the spike pattern observed in other states. Trend data on grade retention rates for the state of Ohio show that in the 1999-2000 school year, *all* primary grades (K-8) exhibit increases in grade retention rates on the order of 2.5 to 3.7 percent, but retention rates returned to previous levels in the next school year. Yet, at least for kindergarteners, retention rates jump right back up again in the 2001-02 school year. Twelfth grade retention jumped a seemingly improbable 8 percent between 2003-04 and 2004-05 – from 3 to 11 percent – and this is likely attributable to the introduction of a more demanding high school exit examination in 2004-05 in that state.

Despite several temporary spikes, retention rates have not been highly responsive to the introduction and expansion of standards-based reforms since the 1990s associated with high stakes testing and concerns over social promotion. Some states, such as New Mexico, show growth in overall retention rates since the introduction of NCLB in January 2002 or since the 1990s under states' individual accountability initiatives, but this trend is not uniform. Instead, if retention rates do increase over time in states, this increase is concentrated in kindergarten and the early primary grades or key transition grades such as the 9th grade. Alternatively, if increases in grade retention rates are

apparent and appear to be concurrent with the accountability and standards movement, these increases are usually reversed in the next year or soon thereafter. Thus, the story in grade retention rates over time, if any, appears to be that states do exhibit short-term increases in grade retention rates associated with standards-based reform but reverse the increases in the next year. Displaying retention trend data by grade does illustrate an important point. Though states' overall retention rates often showed little to no growth – even decreases in grade retention from the early 1990s on – this often masks other important trends, such as increased kindergarten retention.

Conclusion

Comparing data from the October Current Population survey and the state education agencies yields two consistent findings. First, there has not been an abrupt and sustained increase in the levels of retention in the decade of accountability just before and after passage of NCLB. The CPS data indicate increasing levels of retention since 1996, but this follows a longer trend beginning in the 1970s.[39] It may be the case that the move to standards-based reform since the 1990s is the source of this upswing in retention since 1996, for a large number of states implemented such reforms well-before the 2002 passage of the NCLB. However, without more detailed data regarding when individual states implemented a standards-based agenda, we are unable to definitively respond to this possibility. We also observed a decline in differential retention rates by social background and demographic characteristics through 2005, excepting an abrupt increase in 2004. It may be the case that state-implemented accountability regimes, which were not uniformly tied to federal funding and strict student achievement guidelines, may not have had as severe implications for students' progress through school as NCLB. Thus, it

[39] Frederick and Hauser (2005).

is possible a significant portion of states' accountability regimes, implemented well-before NCLB, weakened the link between sociodemographic characteristics and retention. Obviously, this remains conjectural given the available data.

Some significant spikes in retention are observed in the data, particularly the state data and near the time NCLB was passed in the CPS data. Again, it is difficult without more information to explain these sudden increases and subsequent corresponding declines in grade retention rates. Immediately available information suggests that, at least in the case of Ohio, these spikes may be tied to the use of tests to make promotion decisions. This spike pattern also mimics other trends observed in the accountability literature in the case of student test scores. For example, Koretz observes a sharp decline in student test scores with the introduction of new accountability tests, which are typically followed by a subsequent sharp increase in test scores within a few years.[40] It may be the case that tests are in fact being used to make promotion decisions but that as test scores return to previous levels, due to coaching (as Koretz posits) or some other mechanism, the retention rate returns to previous levels. The trends observed by Koretz and seen in the states' data jointly suggest this may be the case.[41] However, this is by no means a definitive conclusion, and it remains to be tested in future research. Some spikes in state retention rates also occurred just before the passage of the NCLB Act, offering slight evidence in support of our conjecture that many states' accountability regimes may not have been as stringent in their consequences for students before it was certain that federal legislation would mandate strict achievement guidelines and sanctions.

[40] Koretz (2002).
[41] Ibid.

Anticipating the passage of NCLB, some states may have begun assessing stronger penalties at the student level.

Second, both sources of data suggest retention very early in a student's career – in kindergarten and first grade – may have been somewhat more responsive to the increasing popularity of accountability regimes. This trend is troubling. Even Alexander, Entwisle and Dauber, who offer a qualified endorsement of grade retention, conclude that students who are retained in the first grade experience the greatest negative consequences, both academically and emotionally.[42] As state testing regimes mature under NCLB, and more data become available, it will be possible to see whether these trends continue. However, we note that testing in the lower grades is not mandated by NCLB, so any link between NCLB and increased retention in kindergarten or first grade must be an anticipatory effort to stave off possible academic problems, later in a student's educational career, rather than a direct consequence of the federal law. Indeed, evidence from the Chicago public schools suggests this is an underlying source of the observed increase in grade retention in the early primary grades. Jacob finds that grade retention increased 3.6 percent in the first and second grades, 64 percent of which was attributable to the introduction of high stakes tests in the third grade.[43]

Recommendations

1. Local, state, and national data systems for grade retention information.

Accumulated observational evidence strongly supports the finding that retention in grade has no lasting, positive effects on students' academic achievement or socio-

[42] Alexander, Entwisle, and Dauber (2003).
[43] Jacob (2005).

emotional development and that, on the contrary, it has negative long-term effects, most notably increasing the cost of K-12 schooling and the likelihood of high school dropout. In this paper, as in other efforts to monitor trends and differentials in retention, we have had to rely on a short-time series of demographic data and on incomplete administrative reports from individual states, which are of doubtful comparability. Neither of these data sources includes any measure of academic achievement or any direct measure of the use of large scale assessments or other tests in decisions about retention. Given the importance—and likely growing importance—of retention as a characteristic of schooling in the age of accountability, along with clear differentials in the likelihood of retention by race-ethnicity and socioeconomic status, local, state, and national data systems should be created to monitor the extent, correlates, and consequences of grade retention. For example, districts and states should be required to track and monitor retention rates for major population groups, just as they now monitor test score performance under NCLB. Indeed, one could argue that valid interpretation of test-score differences across time and across population groups cannot be made without information on retention. Just as school absences and accommodations affect aggregate academic achievement, so also do absences of age appropriate students from tested populations. Information about retention should be included in the Common Core of Data and in the National Assessments of Educational Progress.

2. Mandatory supplementary services for students at-risk of grade retention.

A more immediate need is to make NCLB supplementary services mandatory, either as a concomitant or substitute for retention, and to provide them earlier. Extant NCLB provisions include supplemental services for students attending schools that fail to

make proficiency targets in a consecutive number of years, but these services are voluntary and receive secondary emphasis compared to other facets of NCLB such as testing and measuring student achievement. Research on the supplementary service provision of NCLB is scarce but generally suggests considerable problems with delivery, arguably due to the voluntary nature of these services and their secondary importance in the act. The take-up rate of these services among eligible students is low—just 11 percent for tutoring services.[44] Moreover, students receiving tutoring services face high student to teacher ratios and do not necessarily regularly attend. However, when services are mandatory and delivery problems are addressed, students can and do benefit. Jacob and Lefgren's work on grade retention in the Chicago public schools suggests the positive effects associated with grade retention in their work is largely the result of mandatory tutoring services for at-risk students.[45] Increased supplementary services are especially important because of the likelihood that they will be received by poor students, who, as we have shown, are disproportionately affected by grade retention.

3. Large scale, national randomized experiment on grade retention.

If there is real doubt about the effects of grade retention—a position that can be argued persuasively because so much of the available evidence is purely observational—then there should be a large scale, national experiment on grade retention with random assignment of poorly performing individual students to retention, promotion, or promotion with remediation.[46] Education researchers have traditionally rejected the idea of such an experiment on the grounds that it would be harmful to the retained students, but there is no ethical bar to that if we really do not know whether or not retention is

[44] Farkas and Durham (2007).
[45] Jacob and Lefgren (2004).
[46] Krueger (1999); Burtless (2002).

harmful. Moreover, we would argue that it is surely more harmful in the long run to subject poorly performing students to a harmful educational practice—whether it be promotion or retention—than to find out which policy is more beneficial to students.

4. Resources and incentives to implement and follow guidelines for appropriate use of high stakes tests.

Finally, the history of testing practice confirms that when tests are given, educators will use them to make decisions about individual students, whether or not they are well-designed for that purpose. The massive growth of testing under accountability-based reforms and, especially, NCLB, implies an obligation to use tests wisely and fairly for the purposes for which they were designed. Otherwise, as our findings suggest, the equity goals of NCLB may well be undermined by negative, unintended consequences of inappropriate test use. There is no lack of guidelines for the appropriate use of high stakes tests,[47] but there are few resources or incentives to educate and guide test users or to enforce appropriate use. Given the role of the federal government in mandating increased testing under NCLB, we think that appropriate regulations and enforcement mechanisms for test use should be included in the forthcoming reauthorization of NCLB.[48]

Given the tentative conclusions that can be made with the present data, more research is clearly needed. We will continue to exploit the CPS data to monitor grade retention trends. We also hope to obtain more detailed data on state accountability regimes. We want to combine this information with CPS data in order to more closely track the relationship between grade retention and testing. With these more detailed data, we will be able to address fully the questions of whether and how much accountability

[47] American Educational Research Association (1998); National Research Council (1999).
[48] See, for example, U.S. Department of Education, Office for Civil Rights (2000).

regimes, particularly after the passage of NCLB, have affected student promotion decisions. For the time being, we conclude that grade retention rates have not been dramatically responsive to the increasing popularity of accountability regimes, both prior to and immediately after the passage of NCLB. However, for particular groups, such as first-graders, accountability regimes characterized by high-stakes testing appear to have marked effects. Time will tell if this remains the case.

References

Alexander, Karl L. 1998. "Letter to the Editor." *Psychology in the Schools* 35:402-04.

Alexander, Karl L., Doris R. Entwisle, and Susan L. Dauber. 2003. *On the Success of Failure: A Reassessment of the Effects of Retention in the Primary Grades*. Cambridge: Cambridge University Press.

Allensworth, Elaine. 2004. "Ending Social Promotion: Dropout Rates in Chicago after Implementation of the Eighth-Grade Promotion Gate." Consortium for Chicago School Research, Chicago, Illinois.

Allison, Paul D. 1999. "Comparing Logit and Probit Coefficients across Groups." *Sociological Methods Research* 28:186-208.

Allison, Paul David. 2002. *Missing Data*. Thousand Oaks, Calif: Sage Publications.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. 1998. *Standards for Educational and Psychological Testing (Draft Standards)*. Washington, D.C.: American Psychological Association.

Andrew, Megan. 2006. "Retained and Re-Tracked? Preliminary Evidence on the Effects and Implications of Primary Grade Retention in the Timing of the Transition to Adulthood." Department of Sociology, University of Wisconsin, Madison, Wisconsin.

Burnham, K. P. and D. R. Anderson. 1998. *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer.

Burtless, Gary. 2002. "Randomized Field Trials for Policy Evaluation: Why Not in Education?" Pp. 179-98 in *Evidence Matters: Randomized Trials in Education Research*, edited by F. Mosteller and R. F. Boruch. Washington, DC: Brookings.

Byrd, R. S. and M. L. Weitzman. 1994. "Predictors of Early Grade Retention among Children in the United States." *Pediatrics* 93:481-87.

Clinton, William J. 1998. *Memorandum to the Secretary of Education*. Washington, DC: White House Press Release.

Corman, H. 2003. "The Effects of State Policies, Individual Characteristics, Family Characteristics, and Neighbourhood Characteristics on Grade Repetition in the United States." *Economics of Education Review* 22:409-20.

Dawson, P. 1998. "A Primer on Student Grade Retention: What the Research Says." *Communiqué* 26:28-30.

Eide, E. R. and M. H. Showalter. 2001. "The Effect of Grade Retention on Educational and Labor Market Outcomes." *Economics of Education Review* 20:563-76.

Eide, Eric R. and Dan D. Goldhaber. 2005. "Grade Retention: What Are the Costs and Benefits?" 31:20, 195-214.

Farkas, George and R. Durham. 2007. "The Role of Tutoring in Standards Based Reform." in *No Child Left Behind and Poverty*, edited by A. Gamoran. Washington, DC: Brookings Institution.

Frederick, Carl B. and Robert M. Hauser. 2005. "Have We Put an End to Social Promotion? Changes in Grade Retention Rates among Children Aged 6 to 17 from 1972 to 2003." in *American Sociological Association*. Philadelphia, Pennsylvania.

Hanushek, E. A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24:1141-77.

—. 1996. "A More Complete Picture of School Resource Policies." *Review of Educational Research* 66:397-409.

Hauser, Robert M. 2004. "Progress in Schooling." Pp. 271-318 in *Social Inequality*, edited by K. M. Neckerman. New York: Russell Sage Foundation.

—. 2005. "K.L. Alexander, D.R. Entwisle, S.L. Dauber, on the Success of Failure, a Reassessment of the Effects of Retention in the Primary School Grades, 2nd Edition." *Journal of School Psychology* 43:87-94.

Hauser, Robert M. and Megan Andrew. 2006. "Another Look at the Stratification of Educational Transitions: The Logistic Response Model with Partial Proportionality Constraints." Pp. 1-26 in *Sociological Methodology 2006*, edited by R. M. Stolzenberg. Washington, D.C.: American Sociological Association and Blackwell Publishers.

Hauser, Robert M., Devah I. Pager, and Solon J. Simmons. 2004a. "Race-Ethnicity, Social Background, and Grade Retention." Pp. 97-114 in *Can Unlike Students Learn Together? Grade Retention, Tracking, and Grouping*, edited by H. J. Walberg, A. J. Reynolds, and M. C. Wang. Greenwich, Connecticut: Information Age Publishing.

Hauser, Robert M., Solon J. Simmons, and Devah I. Pager. 2004b. "High School Dropout, Race-Ethnicity, and Social Background from the 1970s to the 1990s." Pp. 85-106 in *Dropouts in America: Confronting the Graduation Rate Crisis*,

edited by G. Orfield. Cambridge, Massachusetts: Harvard Educational Publishing Group.

Holmes, C. Thomas. 1989. "Grade Level Retention Effects: A Meta-Analysis of Research Studies." Pp. 16-33 in *Flunking Grades: Research and Policies on Retention*, edited by L. A. Shepard and M. L. Smith. London: The Falmer Press.

Hong, G. and S. W. Raudenbush. 2005. "Effects of Kindergarten Retention Policy on Children's Cognitive Growth in Reading and Mathematics." 27:205-24.

Jacob, B. 2005. "Accountability, Incentives and Behavior: Evidence from School Reform in Chicago." *Journal of Public Economics* 89:761-96.

Jacob, B. A. and L. Lefgren. 2004. "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." 86:226-44.

Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.

Jimerson, Shane R. 2001. "Meta-Analysis of Grade Retention Research: Implications for Practice in the 21st Century." *School Psychology Review* 30:420-37.

Karweit, Nancy L. 1999. "Crespar Report." Center for the Education of Students Placed at Risk, Center for Social Organization of Schools, Baltimore, Maryland.

Koretz, Daniel M. 2002. "Limitations in the Use of Achievement Tests as Measures of Educators' Productivity." *Journal of Human Resources* 37:753-77.

Krueger, Alan B. 1999. "But Does It Work?" Pp. 46.

Mare, Robert D. 2006. "Statistical Models of Educational Stratification: Hauser and Andrew's Models for School Transitions." Pp. 27-37 in *Sociological Methodology 2006*, edited by R. M. Stolzenberg: American Sociological Association and Blackwell Publishers.

Nagaoka, Jenny and Melissa Roderick. 2004. *Ending Social Promotion: The Effects of Retention*. Chicago, Illinois: Consortium for Chicago School Research.

National Research Council, Committee on Appropriate Test Use. 1999. *High Stakes: Testing for Tracking, Promotion, and Graduation*, Edited by J. P. Heubert and R. M. Hauser. Washington, DC: National Academy Press.

Pagani, L., R. E. Tremblay, F. Vitaro, B. Boulerice, and P. McDuff. 2001. "Effects of Grade Retention on Academic Performance and Behavioral Development." 13:297-315.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." Pp. 111-63 in *Sociological Methodology 1995*, edited by P. V. Marsden. Cambridge: Basil Blackwell.

Shepard, Lorrie A. and Mary Lee Smith. 1989. *Flunking Grades: Research and Policies on Retention*. London: The Falmer Press.

Shepard, Lorrie A., Mary Lee Smith, and Scott F. Marion. 1996. "Failed Evidence on Grade Retention." *Psychology in the Schools* 33:251-61.

—. 1998. "On the Success of Failure: A Rejoinder to Alexander." *Psychology in the Schools* 35:404-07.

Steinberg, Jacques. 1999. "Clinton Urges Tough Love for Students Who Are Failing." Pp. 18 in *New York Times*. New York, October 1.

U.S. Department of Education, National Commission on Excelence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform.* Washington, D.C.: National Commission on Excellence in Education.

U.S. Department of Education, Office for Civil Rights. 2000. *The Use of Tests When Making High-Stakes Decisions for Students: A Resource Guide for Educators and Policymakers*. Washington, DC: U.S. Department of Education.

Weakliem, David. 1999. "A Critique of the Bayesian Information Criterion in Model Selection." *Sociological Methods and Research* 27:359-97.

Table 1: Percentage of Students Retained by Previous Year Grade and Year of Survey										
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Kindergarten	4.91	3.63	3.72	3.69	4.84	5.12	4.29	4.68	5.77	5.45
First	5.62	6.63	7.13	4.52	6.99	6.71	6.87	9.12	9.09	9.76
Second	1.67	2.17	2.75	2.22	2.79	2.45	2.32	1.99	2.33	2.26
Third	1.57	1.78	2.24	3.38	2.11	1.71	1.67	2.49	2.05	1.81
Fourth	1.61	1.72	1.16	1.75	1.64	2.21	1.93	1.86	1.37	1.14
Fifth	1.38	1.04	1.16	1.31	2.11	1.49	1.69	1.55	1.33	1.60
Sixth	1.37	1.36	1.95	1.93	2.55	1.88	1.85	1.25	2.09	1.51
Seventh	1.76	1.73	2.18	3.05	3.07	2.28	1.90	2.94	1.44	2.98
Eighth	1.02	1.51	1.79	2.75	1.94	1.86	1.73	2.01	3.03	1.63
Ninth	3.13	2.93	2.21	4.04	2.67	4.11	2.29	3.45	2.66	2.69
Tenth	1.95	2.20	2.39	2.23	2.09	2.29	2.73	2.01	1.79	1.80
Eleventh	1.30	1.79	2.55	2.46	1.97	3.32	1.70	1.54	1.97	1.71
Proportion of the Synthetic Cohort that Enters 12th Grade "On Time"										
Percent	75.77	74.85	72.75	71.29	70.17	69.69	72.97	69.99	69.93	70.33

Table 2: Results from Logits Predicting Retention in the Year Prior to the Survey

Variable	Simple Logit Coef	SE	LRPC Coef	SE
Previous Year's Grade				
First Grade	0.4038***	0.0576	0.4122***	0.0599
Second Grade	-0.8351***	0.0794	-0.8201***	0.0824
Third Grade	-0.9403***	0.0820	-1.0759***	0.0933
Fourth Grade	-1.0689***	0.0856	-1.0706***	0.0893
Fifth Grade	-1.1415***	0.0878	-1.1409***	0.0910
Sixth Grade	-1.2232***	0.0911	-1.2784***	0.0980
Seventh Grade	-0.7735***	0.0780	-0.8332***	0.0837
Eight Grade	-1.0566***	0.0865	-1.0837***	0.0903
Ninth Grade	-0.5613***	0.0742	-0.6352***	0.0797
Tenth Grade	-0.7239***	0.0809	-0.7681***	0.0842
Eleventh Grade	-0.8451***	0.0882	-0.8689***	0.0898
Year of Survey				
1997	0.0563	0.0825	0.1245	0.0901
1998	0.2481**	0.0799	0.3008***	0.0873
1999	0.2364**	0.0807	0.2531**	0.0888
2000	0.2793***	0.0808	0.3319***	0.0879
2001	0.2768***	0.0780	0.3169***	0.0851
2002	0.1808*	0.0797	0.2271**	0.0866
2003	0.2374**	0.0802	0.2887***	0.0868
2004	0.3366***	0.0795	0.3613***	0.0863
2005	0.1924*	0.0819	0.2496**	0.0882
Demographic Characteristics				
Black	0.2688***	0.0595	0.2569***	0.0698
Hispanic	0.0328	0.0581	0.0623	0.0602
Other Race	-0.0107	0.0762	-0.0063	0.0788
Male	0.2318***	0.0350	0.2400***	0.0496
Midwest	-0.2707***	0.0551	-0.3097***	0.0712
South	0.2381***	0.0496	0.2245***	0.0610
West	-0.2750***	0.0549	-0.3018***	0.0708
Major Central City	0.3183***	0.0658	0.3329***	0.0827
Number of Children in Household	0.0714***	0.0137	0.0704***	0.0171
Social Background Characteristics				
Head's Occupation	-0.0009	0.0012	-0.0015	0.0013
Head's K-12 Education	-0.0082	0.0109	-0.0123	0.0108
Head's Postsecondary Education	-0.0455***	0.0123	-0.0480***	0.0142
Spouse's Occupation	-0.0046***	0.0014	-0.0043**	0.0016
Spouse's K-12 Education	-0.0023	0.0122	0.0051	0.0122
Spouse's Postsecondary Education	0.0089	0.0049	0.0087	0.0052
Lowest Income Quartile	0.3427***	0.0666	0.3609***	0.0865
Second Income Quartile	0.1307	0.0581	0.1553*	0.0645
Third Income Quartile	-0.0720	0.0546	-0.0749	0.0581
Home Ownership	-0.1600***	0.0439	-0.1613***	0.0501
Lambda j				
First Grade			-0.0644	0.1191

Second Grade			-0.0873	0.1613
Third Grade			0.6338**	0.2212
Fourth Grade			0.0265	0.1795
Fifth Grade			-0.0045	0.1846
Sixth Grade			0.3379	0.2095
Seventh Grade			0.3716*	0.1891
Eight Grade			0.1921	0.1928
Ninth Grade			0.4832*	0.1946
Tenth Grade			0.3597	0.1948
Eleventh Grade			0.1870	0.1973
Lambda k				
1997			-0.3270*	0.1439
1998			-0.2556	0.1425
1999			-0.0484	0.1577
2000			-0.2572	0.1457
2001			-0.1738	0.1449
2002			-0.2255	0.1470
2003			-0.2845*	0.1446
2004			-0.0527	0.1608
2005			-0.3557*	0.1479
Missing Head's Occupation	0.0810	0.0659	0.0690	0.0662
Missing Spouse's Occupation	0.1357***	0.0401	0.1463**	0.0465
Constant	-3.4203***	0.0756	-3.4418***	0.0818
Observations	154112		154112	
Log Likelihood	-15682.027		-15656.454	

Asterisks indicate significance levels: * = p < 0.05, ** = p < 0.01, *** = p < 0.001

Table 3: Combined Scalar for Each Grade Level in Each Year

Grade	Year									
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Kindergarten	1.000	0.673	0.744	0.952	0.743	0.826	0.774	0.716	0.947	0.644
First	0.936	0.609	0.680	0.887	0.678	0.762	0.710	0.651	0.883	0.580
Second	0.913	0.586	0.657	0.864	0.655	0.739	0.687	0.628	0.860	0.557
Third	1.634	1.307	1.378	1.585	1.377	1.460	1.408	1.349	1.581	1.278
Fourth	1.026	0.700	0.771	0.978	0.769	0.853	0.801	0.742	0.974	0.671
Fifth	0.996	0.669	0.740	0.947	0.738	0.822	0.770	0.711	0.943	0.640
Sixth	1.338	1.011	1.082	1.289	1.081	1.164	1.112	1.053	1.285	0.982
Seventh	1.372	1.045	1.116	1.323	1.114	1.198	1.146	1.087	1.319	1.016
Eighth	1.192	0.865	0.936	1.144	0.935	1.018	0.967	0.908	1.139	0.836
Ninth	1.483	1.156	1.228	1.435	1.226	1.309	1.258	1.199	1.431	1.127
Tenth	1.360	1.033	1.104	1.311	1.102	1.186	1.134	1.075	1.307	1.004
Eleventh	1.187	0.860	0.931	1.139	0.930	1.013	0.961	0.903	1.134	0.831

Figure 1: Percentage of Students Retained by Race and Year

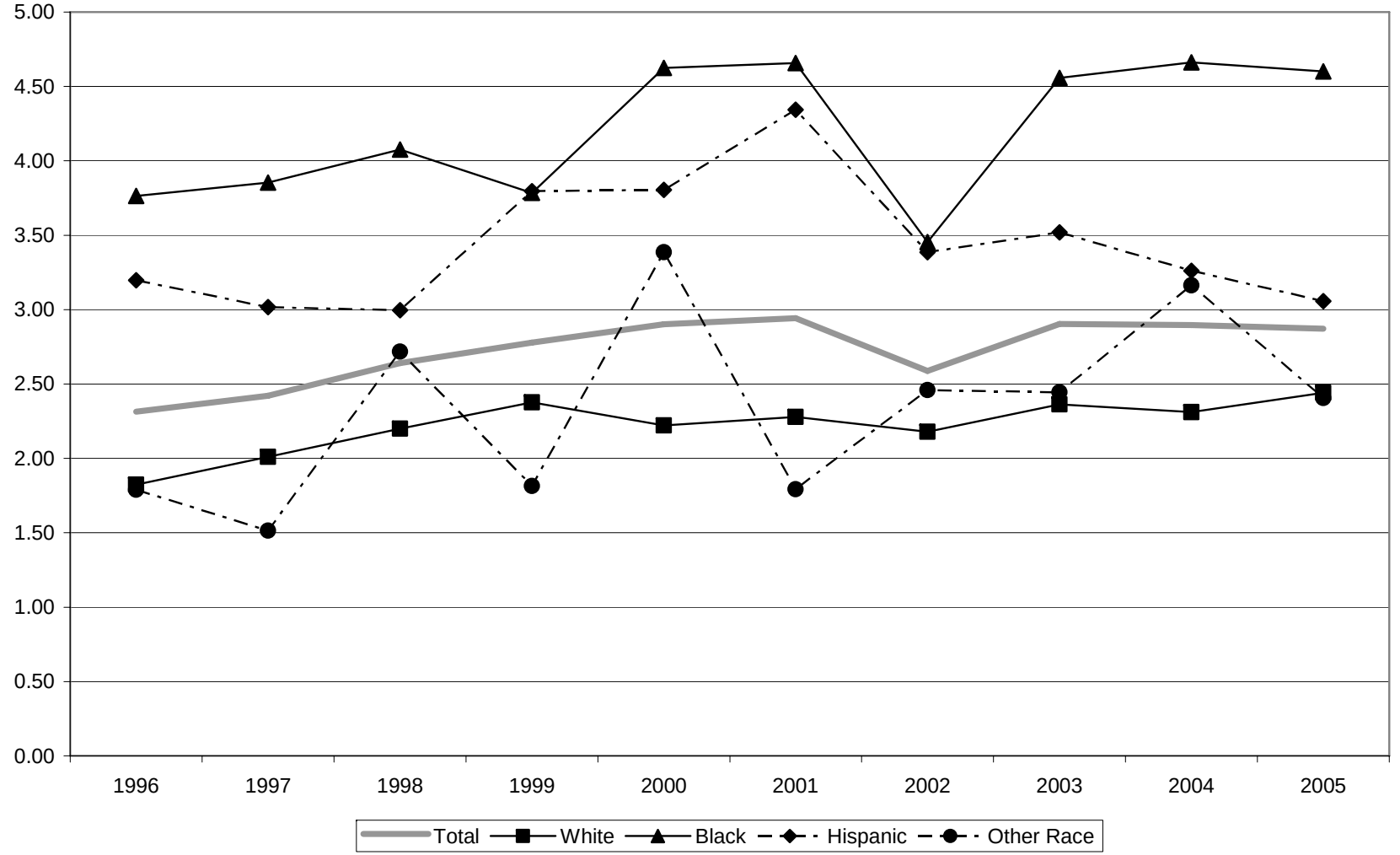


Figure 2: Percentage of Students Retained by Race and Grade

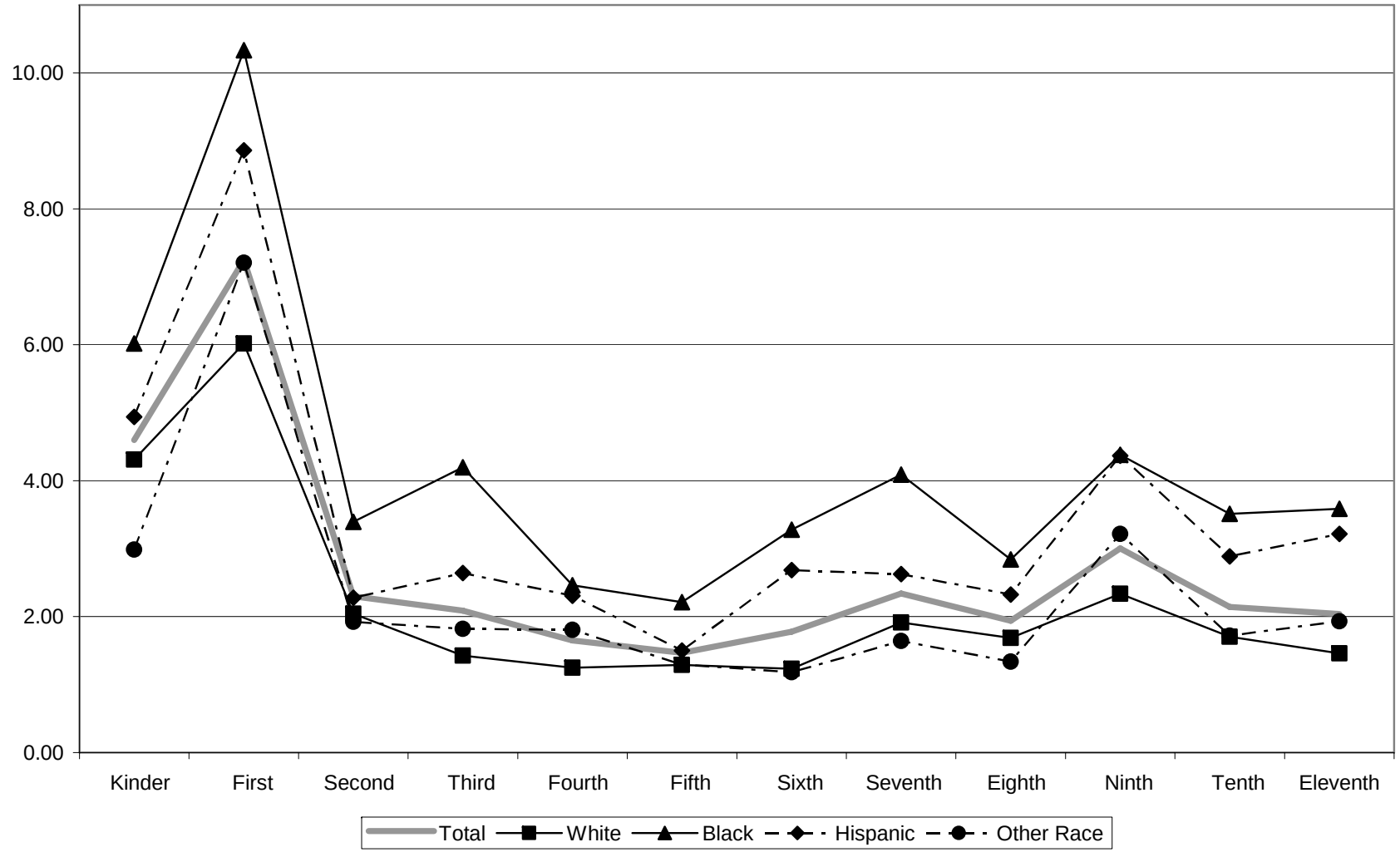


Figure 3: Percentage of Students Retained by Gender and Year

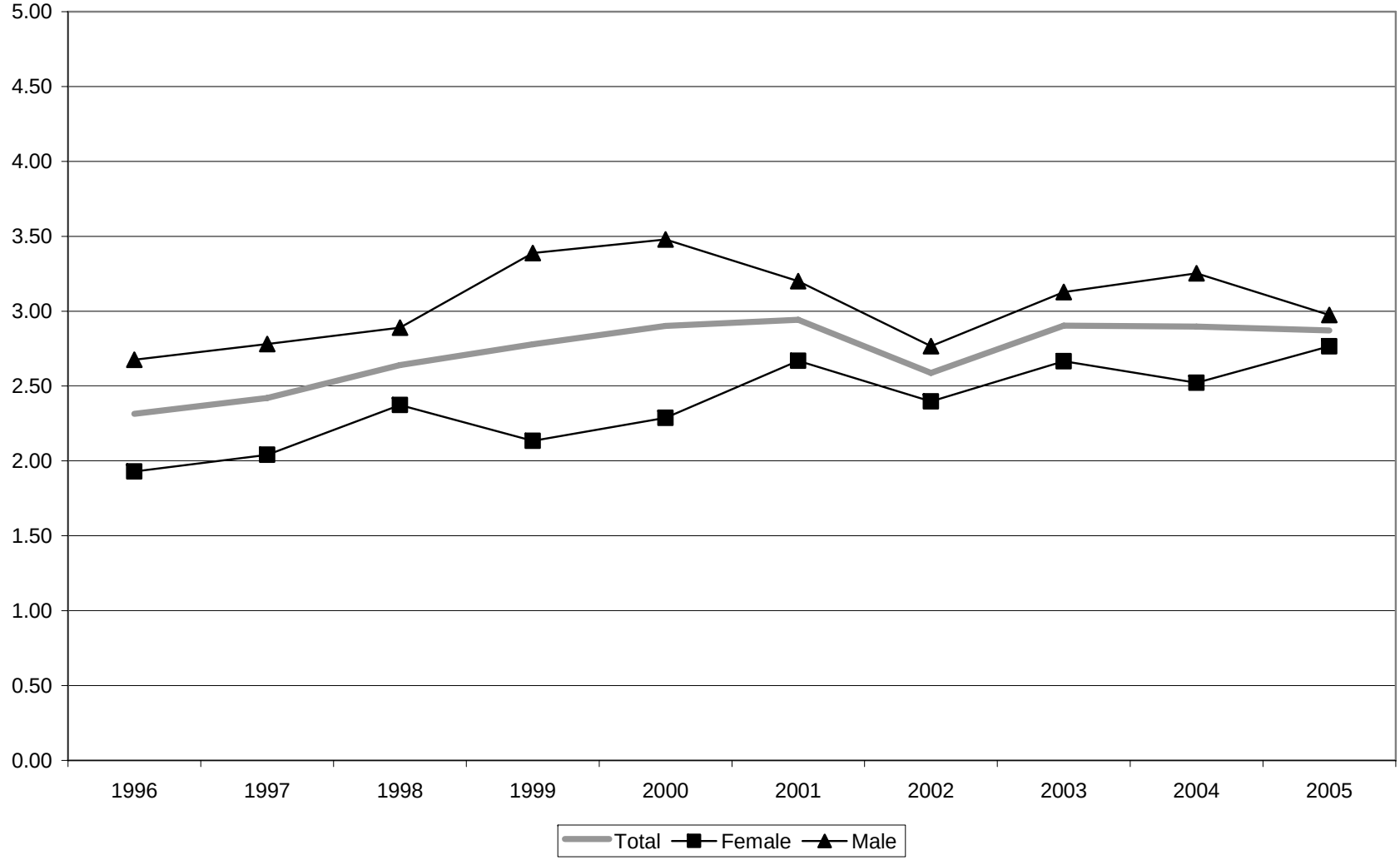


Figure 4: Percentage of Students Retained by Gender and Grade

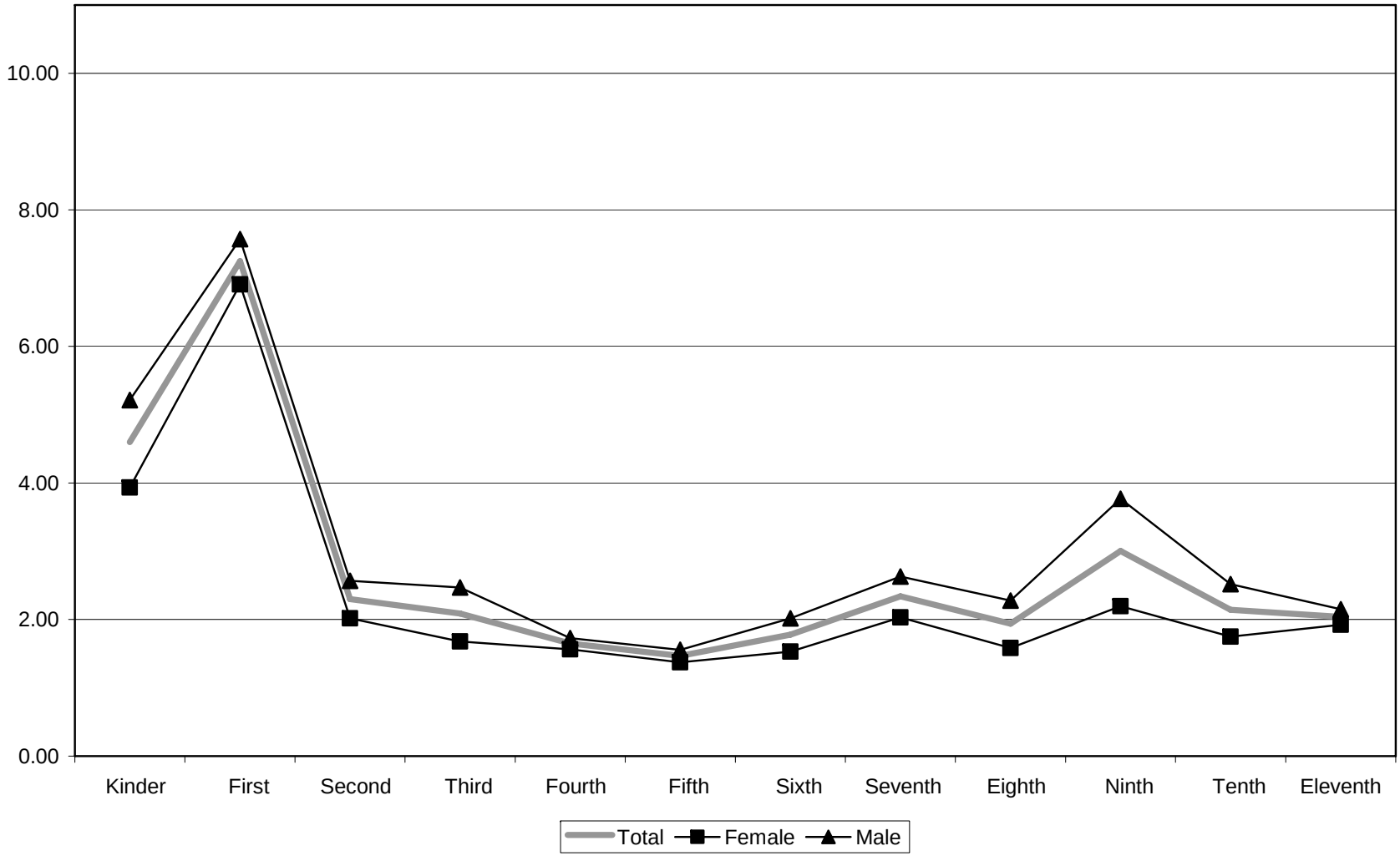


Figure 5: Percentage of Students Retained by Income Quartile and Year

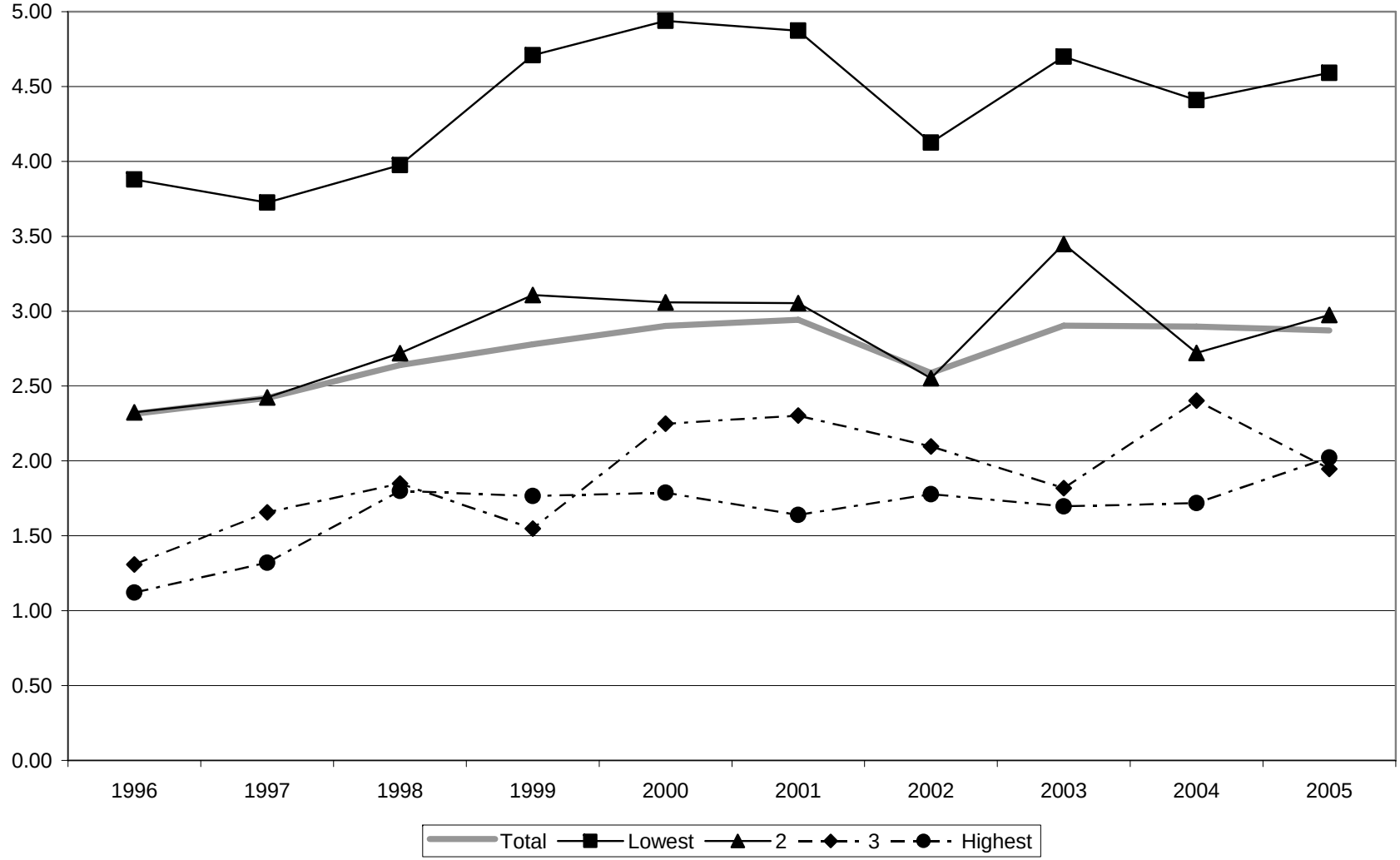


Figure 6: Percentage of Students Retained by Income Quartile and Grade

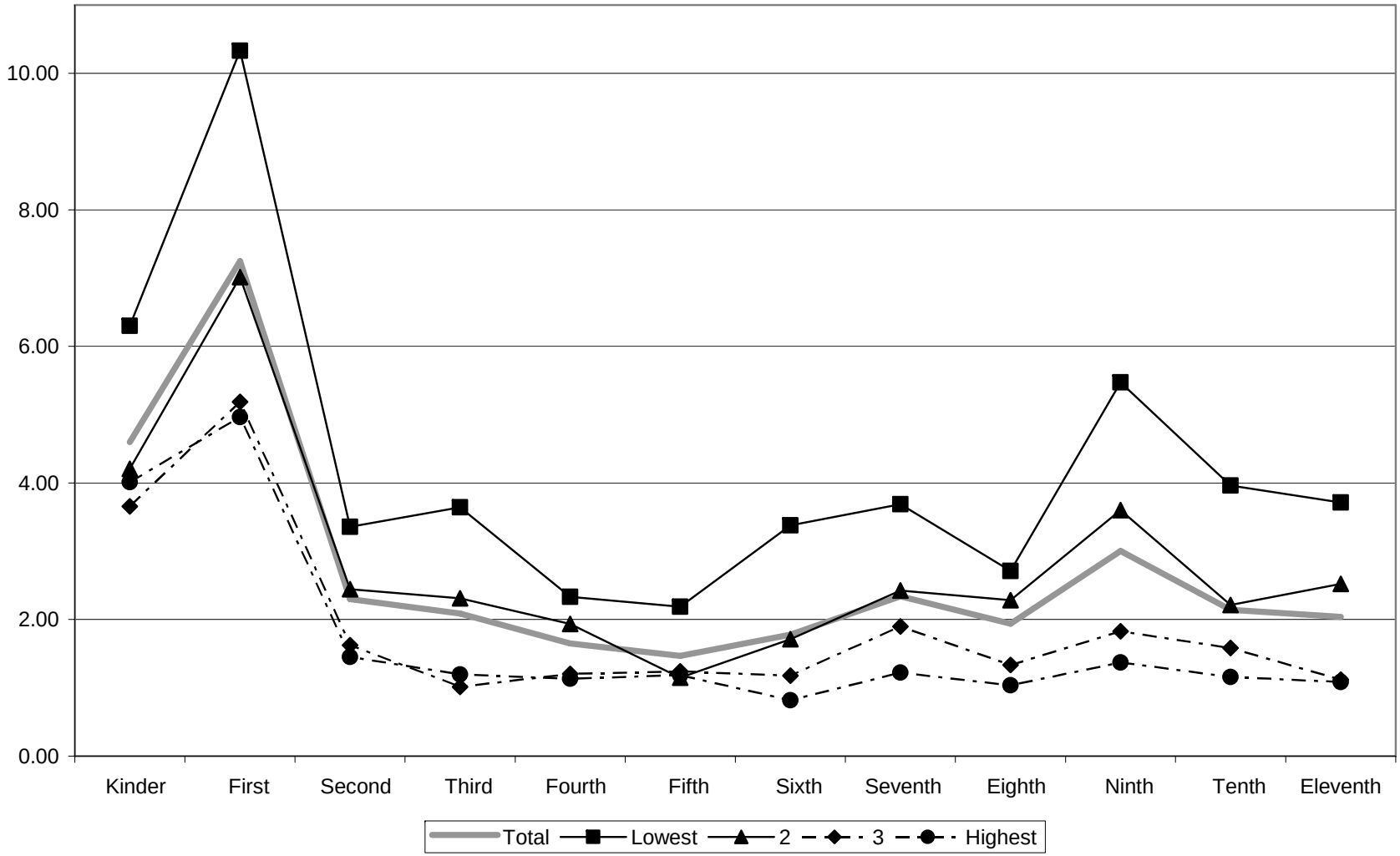


Figure 7: Changes in Scalar Value by Year

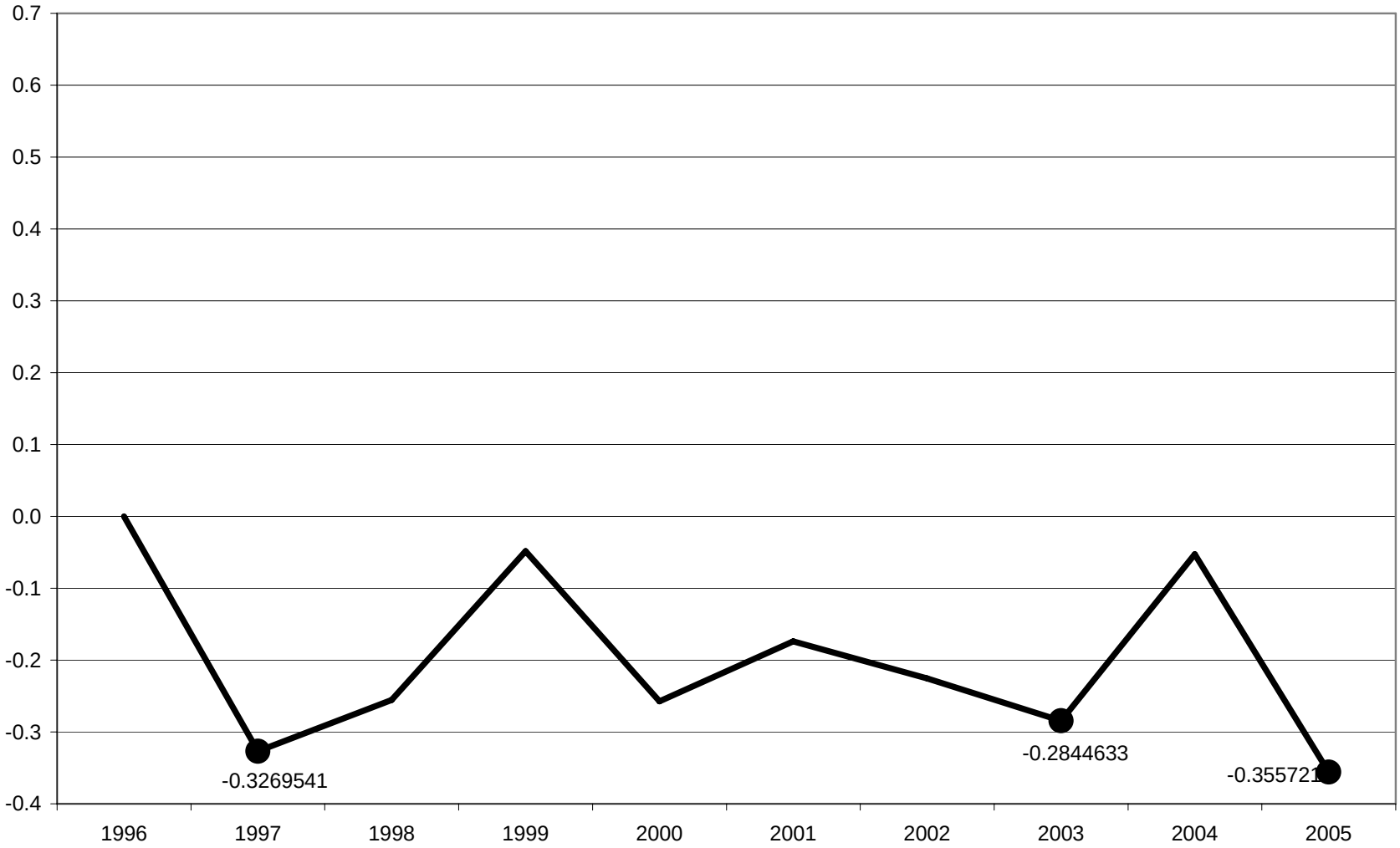
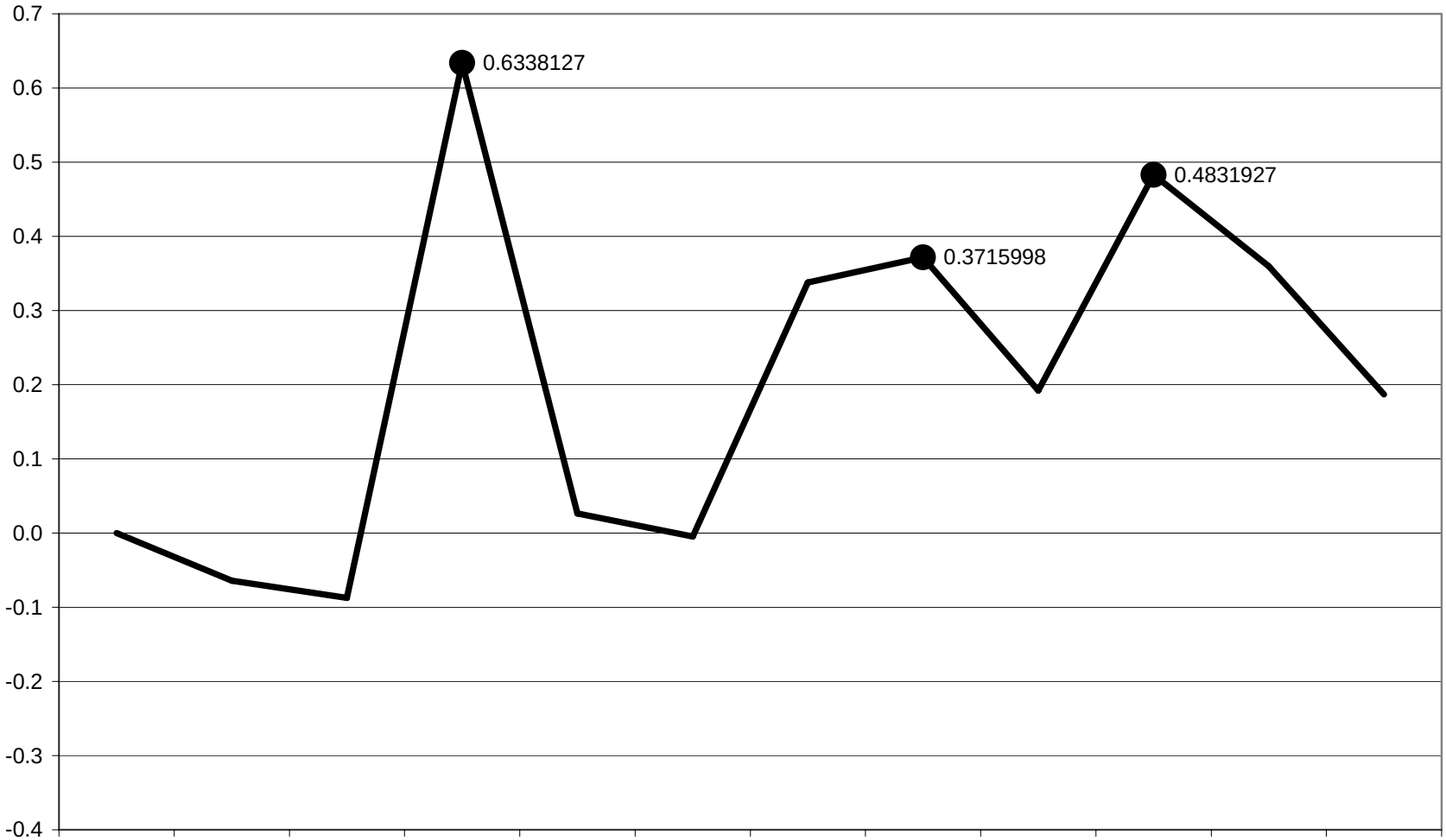


Figure 8: Changes in Scalar Value by Grade Level



Center for Demography and Ecology
University of Wisconsin
1180 Observatory Drive Rm. 4412
Madison, WI 53706-1393
U.S.A.
608/262-2182
FAX 608/262-8400
comments to: hauser@ssc.wisc.edu
requests to: cdepubs@ssc.wisc.edu