Economics Bulletin

Volume 44, Issue 1

Older and wiser? Relative age and college course failure

P. Wesley Routon
Georgia Gwinnett College

Jay K. Walker
Old Dominion University

Abstract

A student's relative age in their schooling cohort has been shown related to several measures of academic and labor market success. Here, we focus on a singular outcome: the probability of college course failure. Even within a sample constrained to students with traditional academic progression and who completed their college degree program, we find evidence relatively younger students were more likely to fail courses. The estimated impact is larger for males, minorities, and those with less academic success before college. Statistical significance remains constant across the parental income distribution. Students within the sample represent over 600 colleges and universities across the United States. Three different empirical methodologies are used: fixed effects regression, two-stage least squares, and regression discontinuity design. Results have implications for educational policy and provide motivation for further study of the relationship between relative age and collegiate success.

Citation: P. Wesley Routon and Jay K. Walker, (2024) "Older and wiser? Relative age and college course failure", *Economics Bulletin*, Volume 44, Issue 1, pages 1-10
Contact: P. Wesley Routon - prouton@ggc.edu, Jay K. Walker - j1walker@odu.edu.
Submitted: August 02, 2023. **Published:** March 30, 2024.

Section 1. Introduction

Though rules and norms differ internationally, and across U.S. states, where public education is required there near always exist laws regarding when a child should initiate schooling. As an example, in the U.S. state of Georgia a child must be five years old on or before September 1st of that year to enter kindergarten, the first year of schooling in the U.S. School starting age (SSA) policies most often result in academic cohorts where the oldest students are approximately a full year older than their youngest peers. As with everyone but especially young children, with age comes maturity and experience. Older students begin schooling with these advantages, which have the possibility to set them on a better academic trajectory. Thus, SSA policies have been both argued and found to result in disparities within schooling cohorts. These inequalities are sometimes called "relative age effects."

Relatively younger students have been shown to score substantially lower in elementary and middle school (McEwan and Shapiro 2008; Ponzo and Scoppa 2014) and are more likely to repeat an early grade (McEwan and Shapiro 2008). Bedard and Dhuey (2006) find that relative age has long-lasting effects on student performance across OECD countries. Relatively younger students have also been shown to have increased probabilities of receiving special education services (Dhuey and Lipscomb 2010). Some studies find relative age effects are more prominent among male students (e.g., McEwan and Shapiro 2008).

There is mixed evidence relative age continues to matter into adult labor markets. Fredriksson and Ockert (2014) find that SSA positively affects one's labor supply and increases earnings among those with low-educated parents, while Du et al. (2012) show the numbers of U.S. CEOs born in June or July (those formerly the youngest in their cohort) are disproportionately small. However, Black et al. (2011) find relatively younger students have increased, not decreased, wages, but that the premium disappears by age 30. Yet others have found labor market outcomes to be unrelated to SSA (e.g., Dobkin and Ferreira 2010).

Compared to early schooling and one's labor market years, the time in between has received less attention in this area of research. However, Peña (2020) finds that, in Mexico, relatively older students perform better academically through age 18, in both class and on standardized tests, as well as apply and are more often admitted to better schools. Using data from a private university in Italy, Pellizzari and Billari (2012) find that relatively younger undergraduate students perform better academically, contrary to most other relative age research. Relatively younger students have been shown more likely to attend two-year and/or vocational colleges instead of a four-year or "academic" college in both Italy (Ponzo and Scoppa 2014) and the U.S. state Michigan (Hemelt et al. 2016). Relatively older students have also been shown more likely to take college entrance exams in the U.S. (Hurwitz et al. 2015).

Here, we use a nationwide U.S. sample of students and focus on a singular outcome, college course failure. To be clear, to "fail" a course (i.e., earn a grade of F in the typical U.S. system) means the student did not score high enough to earn credit for it. They must retake the course if it is a requirement for their degree program. There are several reasons to anticipate an association between relative age and college course failure, including differences in human capital investment (e.g., Peña 2020), existing research on primary schooling in the U.S. (e.g., Bedard and Dhuey 2006), and existing research on undergraduates in other nations (e.g., Pellizzari and Billari 2012). Relative age is defined here as the student's relative age in their cohort when they began formal schooling (in kindergarten), which was affected by their birth date and state of residence at that time. By design, our sample only includes students who progressed as intended from kindergarten through high school, entered a four-year (bachelor's) degree program immediately following high school, and ultimately completed this degree program. Thus, we examine if, among those who are on the traditional academic track and ultimately successful in college, is relative age related to college course failure?

Section 2. Data and Methodology

Data come from the Higher Education Research Institute (HERI), which facilitates the Cooperative Institutional Research Program (CIRP) housed at The University of California - Los Angeles. We make use of their The Freshman Survey (TFS) and College Senior Survey (CSS), specifically. The TFS is administered very near a college student's matriculation and captures student demographics and high school outcomes. The CSS is administered very near undergraduate graduation and captures collegiate outcomes. We collect, merge, and analyze all publicly available data from college graduating years 2001-2006 on students who took both surveys, reported their exact birth date (day, month, and year) and geographic history, and entered college the academic year following high school. HERI data after 2006 are not currently available for public use. Student data from this source have been used in one prior study on relative age (Routon and Walker 2023), which examined its relationships with alcohol consumption and partying behaviors. The sample here contains 61,436 students representing 619 U.S. colleges and universities. These data were merged with SSA policy information collected from state government websites. The sample is made up entirely of students who completed a bachelor's (four-year) degree program, because others were not administered the CSS. This means all findings should be interpreted as impacts among college *graduates* as opposed to impacts among college *students*. It also partially explains why the female student sub-sample is larger than the male sample, females having not only higher enrollment but also completion rates (e.g., Reeves and Smith 2021).

For robustness, three different empirical methodologies are used to estimate the impact of relative age on college course failure. First, fixed effects OLS regressions (hereafter, FE) of the form

$$\text{fail}_{i,s,t} = \alpha + \beta \text{RelAge}_{i,s,t} + \gamma \mathbf{X}_{i,s,t} + \mathbf{S}_s + \mathbf{T}_t + \varepsilon_{i,s,t} \tag{1}$$

are estimated, where $\text{fail}_{i,s,t}$ is an indicator for students who failed one or more college courses during their undergraduate program (the exact number of courses failed is unavailable from this data source); α a constant; $\text{RelAge}_{i,s,t}$ the relative age at the start of kindergarten (measured in months) of student i at institution s in graduating cohort t, with β its corresponding parameter; $\mathbf{X}_{i,s,t}$ the set of student-specific control variables, with γ its corresponding parameter; $\mathbf{S}_s$ institution (or "school") fixed effects; $\mathbf{T}_t$ time (graduating year) fixed effects; and $\varepsilon_{i,s,t}$ the usual error term. The fixed effects help address unobservable heterogeneity at the institution level and across time, such as any national trends in higher education during the sample period. In the interest of data disclosure, summary statistics for the set of student-specific control variables in $\mathbf{X}_{i,s,t}$ are presented in Table 1. All are self-explanatory save the last variable, which is somewhat unique. Students were asked to report whether they primarily chose their institution (college or university) for academic reasons (degree programs offered, institutional prestige, etc.) or for other reasons (a friend attends, a desired social club is present, etc.). Versions of this first empirical strategy have been incorporated in several studies within the relative age literature (e.g., Routon and Walker 2022).

Table 1. Control variable summary statistics

Continuous variable	Mean	Std. dev.
High school grade point average (GPA)	3.590	0.428
Parental education (years; avg. of parents)	16.561	2.989
Real family income ($0,000; at matriculation)	48.088	30.720
Absolute age at college graduation	22.204	0.820
Average age in college cohort	22.178	0.539
Indicator variable	**Rel. freq.**	
First-generation college student	0.150	
Two-parent household (at matriculation)	0.802	
Black/African-American	0.028	
Hispanic/Latino	0.034	
Other non-white race or multiracial	0.092	
Campus resident (at any point)	0.903	
Fraternity/sorority member (at any point)	0.206	
Chose institution for non-academic reasons	0.290	

Notes: n = 61,436. Colleges/universities represented = 619.

As the second methodology, we use a difference-in-differences two-stage least squares approach (hereafter, DiD 2SLS).[1] In the first stage, the dependent variable is an indicator for students who "missed" their SSA cutoff and therefore are among the oldest in their schooling cohorts, with indicators for birthdays near their state's cutoff date used as instruments. Then, the second stage can be described by

$$\text{fail}_{i,s,t} = \alpha + \beta \text{MissCutoff}_{i,s,t} + \gamma \mathbf{X}_{i,s,t} + \mathbf{S}_s + \mathbf{T}_t + \varepsilon_{i,s,t}, \tag{2}$$

where $\text{MissCutoff}_{i,s,t}$ is the indicator for students who missed their cutoff from the first stage, with all other terms the same as before. Now, β is the estimated course failure impact of being the oldest within one's cohort relative to being the youngest.

As the final methodology, we use a fuzzy regression discontinuity design (hereafter, RDD). Like the DiD 2SLS approach, this method compares those at the extreme ends of the relative age distribution. In equation form, this method can be described by

$$\text{fail}_{i,s,t} = \alpha + \beta \text{MissCutoff}_{i,s,t} + \delta(z_{i,s,t} - c) + \lambda[\text{MissCutoff}_{i,s,t}(z_{i,s,t} - c)] + ... + \varepsilon_{i,s,t} \tag{3}$$

where $(z_{i,s,t} - c)$ is the gap between a student's birth date and their SSA cutoff in days and all other terms are the same as before (control variables, $\mathbf{X}_{i,s,t}$, and fixed effects, $\mathbf{S}_s + \mathbf{T}_t$, still included). Optimal bandwidths around the SSA cutoff were calculated within-sample using a well-known RDD procedure (*rdwinselect*) and were determined to be $\pm$57 days for

[1] In previous versions of this manuscript, we referred to this approach simply as two-stage least squares (2SLS). Per a suggestion from a referee, we now refer to it as difference-in-differences two-stage least squares (DiD 2SLS).

male students, ±45 days for female students, and ±73 days when using a combined sample. These models pass the typical RDD diagnostics. For brevity and since the RDD results are largely insignificant, these diagnostic tests are not presented here but are available from the authors upon request.

Table 2. Estimated course failure impacts of relative age

Method (measure)	Sample		
	Full	Female	Male
FE (monthly impact)	-0.00430***	-0.00380***	-0.00487***
	(0.00045)	(0.00042)	(0.00074)
	[61,436]	[39,070]	[22,366]
DiD 2SLS (missed cutoff)	-0.128***	-0.106***	-0.155***
	(0.013)	(0.011)	(0.022)
	[61,436]	[39,070]	[22,366]
RDD (missed cutoff)	-0.070	-0.075	-0.082
	(0.058)	(0.061)	(0.253)
	[20,461]	[10,433]	[4,959]

Notes: Values are marginal effects with robust standard errors in parentheses and sample sizes in brackets. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$. All regressions include all control variables in Table 1 plus time & school fixed effects. See Section 2 for methods.

Section 3. Results

Table 2 presents the initial set of estimates. Because some previous studies have found that relative age impacts students differently across gender (e.g., McEwan and Shapiro 2008), regressions for the female and male sub-samples were estimated in addition to the full sample. Two patterns are immediately apparent. First, in all nine cases, relatively older students are shown less likely to fail courses. Second, the estimated impact is larger for male students than for female students. Within the FE and DiD 2SLS approaches, all estimates are statistically significant at the 99% level. None of the RDD estimates are statistically significant, however, though they exhibit the same negative sign.

Regarding the FE methodology, an additional month of relative age is found to reduce the probability of course failure by 0.43 percentage points for the average student, or about 0.49 and 0.38 for male and female students, respectively. A year of relative age would therefore be estimated to have a 5.16 (12×0.43) percentage point differential. Comparing only those of extreme ages, the DiD 2SLS approach finds the oldest students are an impressive 12.8 percentage points less likely to fail a course than their youngest peers, or 15.5 and 10.6 percentage points for male and female students, respectively.[2] While also statistically

[2] As robustness checks, these models were also estimated with a similar identification strategy to that of Fumarco and Baert (2019) and Peña and Duckworth (2018), where an individual's month of birth is used as an instrument for relative age. The results are similar to the DiD 2SLS estimates presented in Table 2.

insignificant, the RDD approach points toward a smaller average differential at about seven percentage points for the average student.

Table 3. Robustness checks for the full samples

Variation	Method (measure)		
	FE (monthly impact)	DiD 2SLS (missed cutoff)	RDD (missed cutoff)
(a) Birth month dummies added	-0.00234***	-0.161***	-0.134**
	(0.00010)	(0.029)	(0.056)
(b) Excluding possible endogenous variables	-0.00086***	-0.016***	-0.020
	(0.00030)	(0.006)	(0.062)
(c) Both (a) and (b)	-0.00174***	-0.048**	-0.081
	(0.00010)	(0.021)	(0.077)
n	61,436	61,436	20,461

Notes: Values are marginal effects with robust standard errors in parentheses. $^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$. All regressions include all control variables in Table 1 plus time & institution fixed effects. See Section 2 for methods.

Additional robustness checks are presented in Table 3. First, birth month (January, February, etc.) indicator variables are added to the set of controls to hold constant any related seasonality of birth effects. Next, the control variables that might possibly be endogenous (high school GPA and age at graduation) are excluded from the regressions. Finally, regressions where both of these variations (birth month indicators included and possibly endogenous variables excluded) are incorporated were estimated. While these variations change the magnitude of relative age's estimated impact on course failure, they do not change the direction of the estimated impact. Within the FE and DiD 2SLS methodologies, high statistical significance remains. Within RDD, insignificance holds largely constant, though is at the 95% level when only birth month dummies are added. Similar results are found when these variations are performed within the student gender sub-samples.

Using the DiD 2SLS approach, Table 4 shows how these estimates change when the sample is not only split by student gender, but also by race/ethnicity. Among females, the estimated impact for Black students (23.1 percentage points) is over twice as large as that for White students (9.9 percentage points), with Hispanic (14.9 percentage points) and the aggregate group of all other students (11.5 percentage points) falling in between. For males, the estimates for Black (27.2 percentage points) and the aggregate group of other students (27.5 percentage points) are similar, with white students exhibiting an impact that is approximately half as large (14.1 percentage points). For the first and only time under the DiD 2SLS approach, an estimate is statistically insignificant, with the course failure

Additionally, we estimated versions of these models where samples are limited to the same time windows used in the RDD approach. Those results are also similar to the DiD 2SLS estimates presented in Table 2. While not presented here for brevity, these robustness checks are available from the authors upon request.

probabilities of Hispanic male students appearing to be unaffected by relative age. Similar models using the FE approach show the same significance levels, while RDD continues to produce estimates that are largely statistically insignificant.[3]

Table 4. DiD 2SLS estimates across race/ethnicity

	Race/ethnicity			
	Black	Hispanic	White	Other
Female students	-0.231**	-0.149**	-0.099***	-0.115***
	(0.093)	(0.067)	(0.012)	(0.041)
	[1,174]	[1,379]	[32,802]	[3,715]
Male students	-0.272**	-0.064	-0.141***	-0.275***
	(0.126)	(0.083)	(0.024)	(0.061)
	[566]	[691]	[19,154]	[1,955]

Notes: Values are marginal effects with robust standard errors in parentheses & sample sizes in brackets. All regressions include all control variables in Table 1 plus time & school fixed effects. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Finally, Table 5 displays results across three important distributions: high school GPA, college entrance exam (SAT) scores, and family income. All findings in this table come from the DiD 2SLS methodology and are statistically significant at the 99% level. The same is true for otherwise similar models under the FE methodology, and as before, RDD produces estimates that are largely statistically insignificant. For both male and female students, those who had the least academic success in high school (as measured by their grades) are the ones most affected by relative age in college, in terms of course failure probabilities. A similar pattern is shown for SAT scores, with the lowest performers on these standardized tests being the most affected. For the family income distribution, estimates are noticeably different across gender, however. For females, it is those from the relatively poorest families who are most affected, while for males the opposite is true. The latter finding breaks with expectations, given the overall set of results, and warrants attention in future research. In the following section, we offer concluding remarks.[4]

Section 4. Conclusions

SSA laws and norms result in academic cohorts where the oldest students are approximately a full year older than their youngest peers. With age comes maturity and experience, and for young children, an additional year of age often makes a significant difference. Thus, relatively

[3]As with the DiD 2SLS models from Table 2, the models presented in Tables 4 and 5 were also estimated where an individual's birth month was used as an instrument for relative age and/or where the samples were reduced to their RDD versions (students close in age to their cutoff date), with similar results to those shown here. All these robustness checks are available from the authors upon request.

[4]Throughout the analysis, we find that a male student's *absolute* age (a control variable used in all regressions here) also appears negatively related to college course failure. For female students and within the full sample, the relationship is statistically insignificant.

older students within a peer group begin school with these advantages and may be set on different (better) trajectories. Here, we find evidence relative age impacts the probability of college course failure, with younger students being more likely to fail. Estimated differentials are generally larger for male students, minorities, and those who experienced lower levels of academic success pre-college. These findings contrast with those of Pellizzari and Billari (2012), who found that, at a private Italian college, relatively younger students perform better academically.

Table 5. Additional DiD 2SLS sub-sample results

Quartile:	1	2	3	4
		Female students		
HS GPA quartiles	-0.134***	-0.107***	-0.108***	-0.077***
	(0.030)	(0.018)	(0.022)	(0.028)
	[9,087]	[10,565]	[9,276]	[10,142]
SAT score quartiles	-0.169***	-0.089***	-0.105***	-0.063***
	(0.029)	(0.017)	(0.020)	(0.030)
	[10,105]	[10,479]	[10,220]	[8,265]
Family income quartiles	-0.135***	-0.099***	-0.088***	-0.088***
	(0.018)	(0.019)	(0.019)	(0.026)
	[11,349]	[9,635]	[9,349]	[8,737]
		Male students		
HS GPA quartiles	-0.268***	-0.140***	-0.123***	-0.152***
	(0.046)	(0.039)	(0.046)	(0.048)
	[4,519]	[5,602]	[5,721]	[6,524]
SAT score quartiles	-0.214***	-0.147***	-0.142***	-0.149***
	(0.042)	(0.040)	(0.046)	(0.060)
	[4,581]	[5,298]	[6,133]	[6,345]
Family income quartiles	-0.124***	-0.156***	-0.166***	-0.182***
	(0.033)	(0.036)	(0.041)	(0.048)
	[5,208]	[5,235]	[6,049]	[5,874]

Notes: Values are marginal effects with robust standard errors in parentheses and sample sizes in brackets. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$. All regressions include all control variables in Table 1 plus time & institution fixed effects.

Our results are relevant in the following contexts. First, they help connect some existing research. Studies have shown consistent support for the idea that relative age matters for young students and mixed support it matters in adult labor markets. Results here show that relative age likely does matter in college, at least in terms of course failure. Second, it is important to understand the determinants of academic success, generally. Relative age appears to be one such determinant and its impact can be seen even among a sample comprised only of baccalaureate graduates. Third, these results may be considered partial justification for some parents who wish to delay the SSA of their children. Lastly, the findings have policy relevance since it is the set of SSA laws that form these inequalities. If one desires to "level the playing field," or more generally diminish the importance of age in academics, perhaps age should not always be the primary determining factor of the date of a child's first day of school.

More research is needed to understand the importance of relative age among young adults. The sample here came from hundreds of institutions and did not suffer from the external validity concerns that arise from single institution samples often used in education research. However, it only includes students who graduated from college and did so during the period 2001-06. A more recent sample of students and/or one that includes drop-outs would shed additional light on the importance of relative age during one's college years. Other collegiate outcome variables are worth examination, including GPAs and exact counts of failed courses. We leave these tasks for future research.

References

[1] Bedard, K. and E. Dhuey (2006) "The persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects" *The Quarterly Journal of Economics* **121**, 1437-1472.

[2] Black, S.E., P.J. Devereux, and K.G. Salvanes (2011) "Too Young to Leave the Nest? The Effects of School Starting Age" *The Review of Economics and Statistics* **93**, 455-467.

[3] Dhuey, E. and S. Lipscomb (2010) "Disabled or Young? Relative Age and Special Education Diagnoses in Schools" *Economics of Education Review* **29**, 857-872.

[4] Dobkin, C. and F. Ferreira (2010) "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" *Economics of Education Review* **29**, 40-54.

[5] Du, Q., H. Gao, and M.D. Levi (2012) "The Relative-Age Effect and Career Success: Evidence from Corporate CEOs" *Economics Letters* **117**, 660-662.

[6] Fredriksson, P. and B. Ockert (2014) "Life-Cycle Effects of Age at School Start" *The Economic Journal* **124**, 977-1004.

[7] Hemelt, S.W. and R.B. Rosen (2016) "School Entry, Compulsory Schooling, and Human Capital Accumulation: Evidence from Michigan" *The BE Journal of Economic Analysis & Policy* **16**, 1-29.

[8] Hurwitz, M., J. Smith, and J.S. Howell (2015) "Student Age and the Collegiate Pathway" *Journal of Policy Analysis and Management* **34**, 59-84.

[9] McEwan, P.J. and J.S. Shapiro (2008) "The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates using Exact Birth Dates" *Journal of Human Resources* **43**, 1-29.

[10] Pellizzari, M. and F.C. Billari (2012) "The Younger, the Better? Age-Related Differences in Academic Performance at University" *Journal of Population Economics* **25**, 697-739.

[11] Peña, P.A. (2020) "Relative Age and Investment in Human Capital" *Economics of Education Review* **78**, 102039.

[12] Fumarco, L. and S. Baert (2019) "Relative Age Effect on European Adolescents' Social Network" *Journal of Economic Behavior and Organization* **168**, 318-337.

[13] Peña, P.A. and A.L. Duckworth (2018) "The Effects of Relative and Absolute Age in the Measurement of Grit from 9th to 12th Grade" *Economics of Education Review* **66**, 183-190.

[14] Ponzo, M. and V. Scoppa (2014) "The Long-Lasting Effects of School Entry Age: Evidence from Italian Students" *Journal of Policy Modeling* **36**, 578-599.

[15] Reeves, R.V. and E. Smith (2021) "The Male College Crisis is Not Just in Enrollment, but in Completion" Brookings Institution article. Retrieved July 3, 2023 from https://www.brookings.edu/articles/the-male-college-crisis-is-not-just-in-enrollment-but-completion/

[16] Routon, P. W. and Walker, J. K. (2023) "(More) Party Time in School? Relative Age and Alcohol Consumption" *Applied Economics* **55**, 5048-5064.