Balancing Work, Family and School: Enrollment Pathways and Outcomes of Older Community College Students Compared to Traditional Age Students[1]

Peter Crosta, Juan Carlos Calcagno,
Davis Jenkins, and Tom Bailey

May 2006

Abstract: This paper presents findings from a new study of the experiences and outcomes of older community college students - those who enter college for the first time at age 25 or later. We estimate a discrete-time hazard model using transcript data on a cohort of first-time community college students in Florida to compare the effect of enrollment pathways on educational outcomes of older students with those of traditional age students. Results suggest that reaching milestones such as fall-to-fall retention, obtaining 20 credits or completing 50% of the program is a more important positive factor affecting graduation probabilities for younger students than it is for older students. We also find that although remediation decreases the odds of graduating in any given term, older students who enroll in remediation are less negatively impacted than younger ones who do.

Keywords: graduation rates, older college students, enrollment pathways, remediation

[1] This research was funded through a research grant from the Association for Institutional Research and National Postsecondary Education Cooperative. An earlier version of this paper was presented at the 2005 American Educational Research Association Annual Meeting. We thank Greg Kienzl, Linda Hagedorn for detailed comments and suggestions that have improved the paper. We are grateful to Pat Windham and Judith Thompson for sharing the data and for their suggestions.

I. Introduction

Community colleges are an important entry point to postsecondary education for adults with no previous college education (Cohen & Brawer, 1996). In fall 2002, adults between ages 25 and 64 represented 35 percent of full-time-equivalent (FTE) enrollments at two-year public colleges, compared with only 15 percent of FTE undergraduate enrollments at four-year public institutions (U.S. Department of Education, 2003). Older students are more likely to be working while enrolled, married, caring for children, and less engaged with other more traditional-age students in the college (Choy & Premo, 1995; Horn & Caroll, 1996). They are also more likely to attend part-time, to enroll in occupational programs rather than academic ones, and to seek occupational certificates rather than associate degrees or transfer to a four-year institution (Bailey, Leinbach, et al., 2003). When modeling educational phenomena, age of entrance serves as a proxy for a host of other characteristics that are more common to older individuals than younger ones. Not surprisingly, Adelman contends that "one demographic variable makes an enormous difference in the distribution of virtually any postsecondary outcome or process – age at the time of first entry to postsecondary education" (Adelman, 2005).

All these factors will certainly affect enrollment patterns, enrollment intensity, and the probability of completing a degree (Choy, 2002). In fact, based on analysis of data from the Beginning Postsecondary Students Longitudinal Study (BPS: 96/01), 60 percent of older first-time community college students did not earn any credential or transfer after six years compared with 40 percent of younger first-time students. Successful completion of degrees and certificates is critical because at least some postsecondary training is needed, on average, to advance beyond the wages earned by those with only high school diplomas (Kane & Rouse, 1999, Grubb, 2002; Bailey et al, 2003). In addition, studies focusing on retraining older workers at community colleges have found positive and significant returns to an academic year of schooling at a community college from 7 to 13 percent for both men and women (compared to those who have lost their jobs and did not pursue training), with a higher return to more technically and vocationally oriented courses compared to academic ones (Jacobson, LaLonde, and Sullivan, 2004; 2005).

In this study, we use unit record transcript data on cohorts of first-time community college students in Florida to compare the effect of enrollment pathways on educational outcomes of older and traditional age students. The questions addressed by this study cannot be answered using data from national panel surveys that follow traditional-age students from high school to college. Besides providing insight into an important group of community college students, this study presents a model for analyzing student unit record data from states and individual institutions and employs a method to analyze longitudinal data. The remainder of the paper is organized as follows: The next section presents the basic background and literature review about the topic. Section 3 describes the Florida data and presents the theoretical foundations for the model, Section 4 provides results from single risk discrete-time hazard models, and Section 5 concludes.

II. Review of the literature

Completion and drop out rates have long been a central preoccupation of educators. The general consensus among educators and researchers is that students who have stronger high school records, who come from higher income families, whose parents also went to college, who do not delay college entry after high school, who attend full time and receive some form of aid, and who do not interrupt their college studies are more likely to graduate. The most widely used conceptual frameworks of persistence and completion developed by education researchers are based on Tinto's *Student Integration Model* (1993) and Bean's *Student Attrition Model* (1985). The central implication of their models is that institutions should try to foster the academic and social engagement of their students in and with the college in order to maximize persistence and retention rates.[2] Bean & Metzner (1985) provide a theoretical framework maintaining that nontraditional students (older, part-time and commuter students) are more negatively impacted by environmental factors than positively impacted by social and academic integration, and therefore, are more likely to interrupt and drop out than traditional students.

Research on determinants of college drop out and degree completion performed by economists has just recently begun. The literature on college completion has focused on students of traditional college age who enroll at baccalaureate institutions (Ehrenberg, 2004). Turner (2004) provides a comprehensive analysis of the gap between college enrollment and completion and identifies the universe of possible explanations. Manski and Wise (1983) and Light and Strayer (2000) provide a micro-level analysis of how 4-year students and institutional characteristics affect a student's probability to graduate, while Dynarki (2003) presents suggestive evidence that financial aid has a causal impact on completion. The link between 2-year institutions and the probability of completing a 4-year degree has been studied by Ehrenberg and Smith (2004) and Sandy, et al. (in press). Rouse (1995) and Bailey, et al. (in press) have analyzed the effect of community colleges and their characteristics on degree attainment. None of these studies, however, have looked at enrollment pathways or intermediate educational outcomes such as credit attainment and retention and how these impact the final educational outcomes of older and younger community college students.

[2] These models have generated an immense amount of research that has been thoroughly summarized by Pascarella and Terenzini (2005).

As noted by DesJardins (2003), most studies use longitudinal data to estimate the probability of completing college. However, these studies are methodologically similar and only look at two points in time. In the first period, when students start their postsecondary education, researchers collect a set of relevant covariates presumably associated with completion rates, like gender, race, socio-economic status (SES), test scores and institutional characteristics. After a convenient and arbitrary period of time to allow students to attain a policy-relevant outcome such as graduation, drop-out, or transfer, researchers estimate the direct effect of these factors on reaching the outcome. This strategy masks fundamental variation that explains degree completion because it is not designed to handle the dynamic characteristics of the higher education process. First, researchers must assume that initial conditions are fixed over time. In reality, enrollment patterns or institutional characteristics such as tuition are likely to change as time passes (DesJardins, 2003). Moreover, according to Bean & Metzner (1985) we expect negative effects of environmental factors to have a different impact on enrollment patterns of older and younger students. Unfortunately, the longitudinal patterns of most of these environmental factors are generally unobserved by econometricians. Therefore, empirical studies that only add age of entrance as a predictor tend to underestimate its coefficient because it captures a host of other negative, unobserved factors that are more common to older individuals than younger ones (Crosta et al., 2006). Second, the effect of covariates is likely to change during the enrollment period, but the usual methodology is not designed to handle time-varying coefficients. Finally, this strategy cannot handle right censoring cases – when the educational outcome under study is not determinable for a unit of analysis within the period of time being observed – and therefore, the estimates are likely to be biased (Allison, 1984). Event history models are specifically designed to study the occurrence and timing of events and to handle all the limitations discussed above (Allison, 1984; Singer & Willett, 2003).[3] By using information that is revealed during the educational process, event history models allow us to dynamically measure effects of intermediate outcomes and educational pathways on some final outcome.

This paper will make a unique contribution to the higher education literature. To summarize, it focuses on students in community colleges rather than four-year institutions and disentangles age from other nontraditional student characteristics. The focus will be on testing if educational pathways, milestones, and enrollment patterns have the same effects on the conditional probability of graduating for older students as they do for younger ones. Specifically, we will determine: 1) how the attainment of certain educational milestones affects the probability of graduation for older community college students and how this might differ from the effect on traditional-age students at community colleges, 2) to what extent is developmental education a barrier to completion for older students compared to younger ones, and 3) if interruptions in enrollment affect older students differently than traditional age community college students.

III. Data and Empirical Model

Dataset and Variables

The data for this study are drawn from the unit record transcript data of nearly 42,641 first-time, degree-seeking[4] Florida college students who enrolled in a college-credit course at one

---

[3] Although not prevalent in the higher education literature, and almost absent from the community college literature, the method has been used on a number of occasions to model phenomena related to education. DesJardins, et al. (1999) use event history to model student departure from a large research university to determine when students are most likely to be at risk of leaving college. Scott and Kennedy (2005) modeled outcomes of associate degree students, dropout, transfer, and completion, as competing risks.

[4] A student is considered degree-seeking if the college deems her studying at any of the following categories of program level: Associate in Arts Degree, Associate in Science Degree, Vocational Certificate, General Freshman, Linkage (programs in more than one school), Associate in Science Certificate, Associate in Applied Science Degree, Applied Technology Diploma, or Advanced Technical Certificate.

of the 28 Florida community colleges in the fall of the academic year, 1998-99.[5] The dataset tracks enrollment by students at Florida community colleges through the spring of 2004.[6] It includes information on the demographic characteristics of all entering students including age, gender, race/ethnicity, previous education, and college placement test scores. It also includes basic transcript information for all students who enrolled in a college-level course at a community college in Florida, with credits attempted and completed by semester, full- or part-time enrollment status, program of study, course grades, credentials earned, and amount and type of financial aid received in the first semester.[7]

We have removed any students who were formerly in dual-enrollment programs and now out of high school and officially in college. The last restriction concerns the age of students. We first restrict the sample to those students who are between the ages of 17 and 65 on September 1 of their first trimester of college (Fall 1998). We have made an additional modification in this study to limit the younger student cohort to those who enter college between the ages of 17 and 20 while the older cohort remains those who enter between age 25 and 65.[8] This restriction addresses the fact that students who begin at, for example, age 23 would technically be older students by the third year of study. Therefore, our analysis compares students who enroll at a traditional age and remain traditional-age students for 15 of the 17 trimesters to those who enroll at an older age and remain nontraditional-age students for the entire span of data collection. All

[5] Three of these 28 colleges currently award bachelor's degrees and are often not considered community colleges. The colleges were all considered community colleges at the time of this study.
[6] This amounts to 17 trimesters, where a trimester is a fall, summer, or spring term. This time span will be referred to as the event period.
[7] The dataset does not include longitudinal data on financial aid. Our regressions include an indicator for receiving federal aid or not in the first term and then we allow the effect to persist throughout a student's higher education.
[8] We lose 3,619 observations with this restriction.

of these modifications result in a sample with 29,421 traditional-age students and 5,652 older students.[9]

Table 1 summarizes the variables that we will use in the analysis. Clearly, the older cohort of students is comprised of more females and has a larger proportion of Black and White students. Hispanic students are more likely to be in the younger cohort than the older one. Most striking are the differences seen in English proficiency and attainment of a high school diploma or GED, with older students more likely to have a first language other than English and to have received a nontraditional secondary credential. There are interesting differences in the math and verbal test scores of older and younger students.[10] Traditional age college students, on average, scored about 87 points higher than older students on mathematics placement exams, but scored about 29 points lower on tests of verbal skills. This discrepancy could be due to older students' being away from formal mathematics education for an extended period of time, whereas verbal scores may improve over time as vocabulary and language skills advance with age.

[Table 1 here]

As the first term of college is decidedly important, we also present in Table 1 a set of student characteristics relative to the first trimester of college enrollment.[11] We first note that older students were more likely to receive federal financial aid. They may be better informed about their financial aid eligibility and application processes than younger students, which may result in their having an advantage in gaining financial assistance. Although community college students of all ages often hold full-time jobs while they attend school, it seems reasonable that

---

[9] Nationally, older students comprise about 43.8% of full- and part-time public community college students (US Dept. of Education 2002). Florida is below this national average.

[10] Students report SAT or ACT test scores upon enrollment, and if they have not taken these, they must take a college placement test at their institution. These scores were all converted to an SAT scale (200-800) using the test-makers' formulae.

[11] As *defined* in Table 1, these variables are all time-invariant, restricted to the first trimester of college. However, a selection of them will be used in the survival analysis as time varying covariates. These variables – tuition, full-time status, and program length – will be allowed to vary over time, along with the pathways variables discussed below.

older students are significantly more often part-time students as they may be more entrenched in careers, financial obligations, and families.[12] Older students are expected to be enrolling in community colleges more often to upgrade jobs skills, and this would result in a disproportionate number of certificate seekers with shorter program lengths among this cohort. Table 1 indicates that the difference in average program length between the cohorts is slightly positive, suggesting that older students are well represented in associate degree programs.

Lastly, there exist differences in first-term enrollment characteristics between older and younger students. Following from the fact that older students are more likely to be part-time, they attempted fewer degree credits and developmental credits during the first term of study than did the traditional age students. We also find that in this first trimester, the older cohort was less likely than younger students to enroll in any developmental (remedial) classes. Non-traditional age students also earned about 78% of the credits they attempted in the first term compared to younger students, who successfully completed about 72% of credits.

The last set of variables that will be of use in our analysis concerns enrollment pathways and outcomes. The top of Table 2 shows the main outcome event that we consider: completion. Completion is defined as receiving a degree or certificate in one of the 17 trimesters of the event period. Those who do not complete a degree in the time period and who are still enrolling are "right censored" in their last enrollment term. As seen in Table 2, on average, traditional age students were more likely to graduate than non-traditional age students in 17 terms.

[Table 2 here]

Several time-varying covariates that represent enrollment pathways and milestones are the main focus of our model. These can be classified into 4 categories: nominal credit

[12] Though not shown here, older students are in fact more likely to be part-time in each of the 17 semesters of the event period.

milestones, percentage of program completion milestones, retention, and remediation. The first set of milestones identifies when a student has earned 10 or 20 credits. Adelman (2004) contends that a major milestone in undergraduate education is the attainment of 10 credits. Our time-varying indicator remains 0 until a student has earned 10 or 20 credits; then it shifts to 1 and remains there for the rest of event time. This allows us to see the shift in risk between students who have and have not reached this educational milestone. In our sample, 62% of older and 79% of younger students obtained at least 10 credits during the observation period. If we exclude remedial credits from this measure, 54% of older and 71% of younger students attained at least 10 non-remedial credits during event time. The disparities noted here may be due to differences in enrollment intensity, with older students having a more difficult time accumulating credits if they enroll primarily part-time. Similarly, a greater proportion of younger students, on average, reaches the 20-credit milestone as well, whether counting all credits or only non-remedial credits.

If older students are overrepresented in programs that require less than 10 or 20 credits, then these credit accumulation differences may not be such useful measures. Therefore, we suggest an alternate way to gauge progress: percentage of program completion. This is simply the proportion of credits earned relative to the number of credits required for the student's program. We consider five milestones for program completion at 5, 15, 25, 50, and 75 percent. Presumably students who complete increasing amounts of their program will have much better odds of graduating and fewer odds of dropping out. There may also be percentage milestones that are particularly important for our two cohorts. In all five measures, the percentage of older students is less than the percentage of younger students with respect to these progress measures.

Remediation effects will be tested using two different variables, one that is time-invariant and one that is time-varying. The time-invariant measure determines whether or not a student ever enrolled in developmental classes. In our sample, there is little difference between older and younger students in the percentage that enroll in remedial classes – about 60% each. Our second remediation measure captures the effects of developmental education throughout event period. If a student is taking remedial courses in a particular term, a dummy indicator is turned on. The indicator is set to zero for those terms in which the student is not enrolling in developmental education. Table 2 reports that in the first term, a higher proportion of younger students were enrolled in remediation than older ones.

Our next pathways variable will aid in determining how younger and older students respond to interruptions and persistence in enrollments. An interruption is defined as a two-consecutive-trimester stop in enrollment. A student will be in a post-interruption period in all terms after the first interruption but will only be observed as an interrupter if she returns after the interruption. Table 2 provides statistics for people who will be considered interrupted due to missing the second and third trimesters. At least 25% percent of younger students and 35% of older students are susceptible to a post-interruption period of enrollment.

Fall-to-fall and fall-to-spring retention are often considered crucial milestones that will move students toward degree completion. A student will be considered retained under these two circumstances for all of the trimesters following the terms of retention. If a student enrolls in the first spring trimester (term 2), she is considered a fall-to-spring retained student for the remainder of event period. A student who then also enrolls in the following fall trimester will be considered a fall-to-fall retained student. Sixty percent of older students and 70% of younger students enroll in the first spring term. The next level of retention is not as common and decreases markedly for older students to 36%.

Our list of explanatory variables has some key omissions that should be noted. As discussed, studies of completion rates show, not surprisingly, that students who come from higher income families and whose parents also went to college, tend to have higher probabilities of graduation. The Florida unit record dataset does not include information on any of these socioeconomic (SES) characteristics of students. Our study provides two strategies to overcome this limitation. First, we add an indicator variable for students receiving federal aid. This measure, which is primarily comprised of Pell Grants awarded to low- and middle-income students, acts as a proxy for the relative income level of students. Second, we include comprehensive information on test scores and we assume these variables are highly correlated with unobserved SES. To the extent that long-term family and environmental factors are reflected in measures of scholastic ability, we accurately control for SES.[13,14]

Econometric Analyses

The statistical method we will use to model outcomes of community college students is the single risk discrete-time hazard model (Allison, 1984; Willet & Singer, 2003; DesJardins, 2003). To facilitate discussion of this model, it is useful for the reader to understand how the data set is organized. Rather than one observation per student, we have a person-period data set with a maximum of 17 observations per student – one for each trimester, and each student is only observed when she is enrolled. Time-invariant variables remain constant for each person in each

[13] The effect of unobserved heterogeneity is thoroughly analyzed in Crosta, et al. (2006).

[14] Noticeably absent from this discussion is the outcome of transferring to a baccalaureate institution. Those students who fail to obtain degrees or certificates but do transfer out of the Florida Community College System - whether to another state's 2-year college or a private or public 4-year college - will be right censored observations; we do not know what happens to them. Unfortunately, we have limited information on students who transfer, only knowing if they transferred to the Florida State University System, and we therefore miss transfers to other 4-year schools. This paper focuses on the completion outcome, but it may certainly be worthy addressing transfers in a competing risk framework. For now, we are considering transferring without a 2-year credential the same as non-completion.

period, and time-varying variables can take on different values in different time periods. Our event is completion of a degree from a community college in Florida, and in our framework, we will say that each student is "at risk" of completing when she is enrolled. Once a student experiences an event, her observations in later time periods are discarded, effectively removing her from reentering the risk set. Since we have 17 trimesters of data (5 years plus 2 trimesters), we can only observe outcomes for these 17 terms, and these periods are referred to as "event time" or the "event period". The beginning of event time is the first semester of enrollment in a Florida community college (fall 1998) and the end is spring 2004. Students who have not completed by the 17th term have unknown outcomes at the end of analysis period and thereafter.[15]

Formally, we are modeling the risk of completion in each trimester, called the hazard. This is the conditional probability that an individual will obtain an outcome in time period *j* given that she did not do so in an earlier time period and given that she is in the risk set. To be in the risk set in a given trimester, the student must be enrolled. We can write this basic discrete-time hazard function as:

$$h(t_j) = \Pr[C_k = j \mid C_k \geq j] = \frac{n_j}{S_j} \tag{1}$$

where $C_k = j$ indicates student *k*'s outcome in term *j*, n_j is the number of students who completed in term *j*, and S_j indicates the number of students who could potentially complete in term *j*. The condition $C_k \geq j$ ensures that an outcome for student *k* has not occurred before time period *j* and the student is enrolled in (observed in) time period *j*. This initial specification of the model

[15] We know whether students have or have not completed a degree or certificate, but if they do not complete and continue to enroll, we do not know whether or not they will complete or not after the end of event time. In order to produce unbiased analyses, we must assume that the censoring due to the end of event time or due to non-enrollment is noninformative (Willet & Singer, 2003). Noninformative censoring maintains that the individuals who still have not completed by the end of the data collection period are actually still capable of graduating. Essentially, this means that censoring is independent of event occurrence. Some additional assumptions about these students are required to perform this analysis.

assumes that every student has the same risk of attaining an outcome in each time period if enrolled; that is, there is no observed heterogeneity among students. A convenient way to introduce the basic form of the hazard function is by looking at life tables.

Table 3 presents life tables that describe our completion outcome in temporal terms for the entire sample and for the sub-samples of younger and older community college students, respectively. The column labeled “At risk” denotes the number of students who are eligible to graduate in that term because they have enrolled. It is clear that the fall and spring terms have more students enrolled than the summer terms. The column labeled “Completed” indicates the number of students who earned a degree or certificate in each term. In the first term, 185 students earned a certificate or degree. Since 35,073 were in the risk set, the hazard function (conditional probability of graduating) is 0.53%. In the third trimester, only 12,497 students enrolled and 84 completed a degree or certificate. Although fewer completed than in the first term, the hazard function is greater because the probability is conditional on being in the risk set, or enrolling. As students vary enrollment patterns, the composition of the risk set changes.

[Table 3 here]

The rightmost column in each panel - the hazard function from equation (1) - is of particular interest and is the empirical function we will comprehensively model later. For example, student’s have the greatest risk of completion in the 9th trimester, where the conditional probability is roughly 13.4% overall. We can now note the basic shape of the risk profiles by analyzing the hazard function columns. Over time, the hazard increases and reaches local peaks during the summer terms; it then gradually decreases. The summer spikes are due to relatively small risk sets, as mentioned above. Comparing younger and older cohorts (the second and third panels, respectively), we see that the hazard function for older students is larger in magnitude than the one for younger students in the beginning of event time, but by the fifth term, younger students begin to have a greater risk of completion in each period. This could be due to older students enrolling in more short-term occupational certificate programs and younger students enrolling in longer associate degree programs. However, we noted earlier that there was little difference in the average program lengths between the two cohorts.

Figure 1 is a graph of the hazard functions for the two groups, and it depicts in which time periods students are at the greatest and least risk of graduating. Both curves have an inverted "U" shape, which indicates that students are at less risk of graduating at the beginning and end of event time and at greatest risk of completing a degree or certificate somewhere in the middle of the 17 trimester period[16]. Younger students are at greatest risk of graduating in the 9th term, which is the summer semester in the third year of study. Nontraditional-age students have a peak in the 12th term, and their curve is flatter than the one that describes younger students. It is evident, however, that in almost every time period, the older cohort has a lower conditional probability of completing a degree.

[Figure 1 here]

The model thus far has produced little more than a description of univariate data over time.[17] To develop a more comprehensive model of completion for Florida community college students, we accept the prospects of observed heterogeneity in our sample. This basically means that students with different characteristics – both time varying and time-invariant – will have different hazard functions. The general population discrete-time hazard can be conceptualized as:

$$h(t_j) = \Pr\left[C_k = j \mid C_k \geq j, G, \mathbf{X}, \mathbf{Z}\right] \tag{2}$$

[16] It should be noted that, due to censoring, we cannot say for sure if a trends continue downward after the 17th term. All statements made only pertain to the specified and observed event time.

[17] This is often called the baseline hazard.

where the hazard now is the conditional probability that student *k* completes in term *j* given: that she has not completed before *j* and is in the risk set, an indicator variable, *G*, for being an older student, a vector **X** of student characteristics, and a vector **Z** of enrollment pathways and milestones. Algebraically, we can write the relationship in (2) as:

$$\text{logit h}(t_j) = \mathbf{D}'_j \alpha_j + G'\delta + \mathbf{X}'\beta + \mathbf{Z}'\gamma \qquad (3)$$

In equation (3), we have taken the logit of the hazard[18] and defined a linear relationship between the conditioning data and logit hazard, where $\mathbf{D}_j$ is a vector of dummy variables indexing each trimester; *G*, **X,** and **Z** are as defined above; and α, δ, β and γ are parameters to be estimated. Taking an inverse transformation of both sides, we derive:

$$h(t_j) = \frac{1}{1 + e^{-\left[\mathbf{D}'_j \alpha_j + G'\delta + \mathbf{X}'\beta + \mathbf{Z}'\gamma\right]}} \qquad (4),$$

which is now a nonlinear relationship between the predictors and the hazard and analogous to the standard logistic regression routine (Singer & Willet, 2003). Once the data is put in a person-period dataset, we can estimate parameters that maximize the likelihood of observing the sample data assuming a logistic distribution.

Since survival analysis has a large temporal component, it is important to discuss our treatment of time in the equations presented above. Our models have assumed a general non-parametric specification for time that is entered as a series of *t* dummy variables indicating each trimester as $\mathbf{D}'_j \alpha_j$. That is, there are no explicit functional restrictions placed on how time affects the probability of completion. This allows for the baseline hazard (the hazard modeled by equation (1)) to take on any shape and thus capture the effect of time, or the profile of hazard

[18] A discrete-time hazard model using a logit transformation assumes proportional odds rather than proportional hazards. A complementary log-log link could have also been used, but it did not seem necessary to assume proportional hazards (and implicitly assume that the data process is occurring in continuous time that is interval-censored). See Singer & Willet 2003, Ch. 12.

over time. Often, researchers can justify entering the time component into the model as a linear, quadratic, or cubic function of time, for example, due to an observed relationship with the hazard event and to reduce degrees of freedom used in the analysis. However, the sample size of our dataset allows us to add seventeen new parameters (one for each time period) to estimate without compromising the reliability of results. Moreover, after some experimentation, we have determined that the general specification is the best fit for our data.[19]

IV. Empirical Results

In this section, we estimate several hazard models to test how milestone attainment, remediation, and interruption differently affect older and younger cohorts of students and their probability of graduating. We approach hazard modeling – estimating equations in the form of equation (4) – by first beginning with a simple model, which includes only the effects of time and the older student dummy variable. Then covariates and substantive predictors are introduced into the model to compare the older and younger cohorts. This is followed by entering our intermediate outcomes (pathways) as both time-varying and time-invariant predictors (Allison, 1984).

Table 4 presents our initial odds ratios and standard errors derived from maximum likelihood estimation of logistic regression parameters.[20] In column (1), we present a simple baseline hazard model with the time-invariant dummy variable indicating whether the student is in the older-student cohort or not. The odds ratio of 0.925 indicates that in any given period, an older (younger) student is 0.925 (1.08) times as likely to complete a degree or certificate as a student in the younger (older) cohort. In this model, the effects of time represent the odds of

[19] Model fits were determined using deviance statistics and the Akaike Information Criterion. The general non-parametric specification allows us to capture summer spikes as is discussed below.

[20] $Odds = e^{\beta}$, where β represents the estimated coefficient estimated in equation (3). An odds ratio will not be statistically significant if it is very close to 1.

graduating in each period for a younger student. Note that because of the parameterization and logit link, the odds for the subgroups in each period are proportional with a ratio of 0.925. We shall also note that the test of the change in deviance indicates that adding the indicator variable for the older cohort improves the fit of the model. Figure 2 presents these conditional hazard probabilities for each cohort in graphical form. These functions, like the hazard estimates of Figure 1 are risk profiles for each group. To maintain proportional odds, each function is a diminution or magnification of the other function. We can see clear spikes in the 9th trimester when both groups are at the greatest risk of completing a degree. The graphic also shows the proportionally lower odds of the older cohort in each term.[21]

[Figure 2 here]

Now that we have established an idea of baseline hazard, we add controls for sex, race (White is the reference group), receiving federal aid in the first term, US citizenship, secondary credential, measured ability, tuition, program length, and full-time status in column (2) and (3). The parameter estimates again represent the covariate's effect on the conditional probability of completion in any given trimester. Column (2) of Table 4 suggests that those who are female and high school diploma holders are more likely to graduate each period, whereas Blacks, American Indians, and Hispanics are less likely to complete a degree or certificate during the event period. Adding this first wave of controls pushes the effect of being an older student upward, so that older students are expected to be 0.94 times as likely to graduate in each time period, all other factors held constant. When we add in measures of ability and ability squared to the model (as

[21] The spikes in the summer terms give an odd shape to the hazard. However, the overall parabolic shape of the function is clearly seen (as shown in the hazards of Figure 1).

shown in column (3)), the effect of being an older student reverses in sign.[22] An older student is 1.31 times as likely to complete a degree as a younger student, ceteris paribus.[23]

[Table 4 here]

We next analyze the difference in older and younger student responses to our enrollment pathways. To test for differences in the effects of the pathways variables, we include in the regression the older student indicator and the pathway variable as well as an interaction term that specifies the joint effect. This interaction term will tell us if there is a difference in the effect of the odds of graduating in any given period between the two groups. We utilize the coefficient on older students and the interaction term to compute the impact of the pathway that is specific to older students. Models for each pathway variable are estimated individually as extensions of the model in Table 4, column (3). Table 5 reports exponentiated coefficients for the substantive portion of the individual models.

[Table 5 here]

The first column in Table 5 indicates the milestone or pathway that we are testing, and each row represents a separate regression that independently tests effects of the specific pathway. The next column (a) presents coefficients and standard errors for the older student dummy that indicates the direct effect of being an older student. Column (b) is the direct effect of the pathways, which is also the effect specific to younger students since we remove the effect specific to older students through an interaction term between the dummy for older students and the pathway as shown in column (c). Column (d) computes the impact of the pathway that is specific to older students by multiplying the odds ratios associated with the pathway (column

---

[22] Ability and tuition are entered in the model as standardized variables with mean zero and standard deviation one.

[23] This result is discussed in length in Crosta, et al. (2006).

(b)) and the interaction term (column (c)). Columns (e) and (f) display deviance statistics and the change in deviance that indicate improvements in the model fit compared to Table 4, column (3).

The first group of estimates concerns credit attainment. These include time-varying indicator variables that “turn on” when a student reaches the credit milestone and stay on in all periods thereafter. Column (a) in Table 5 reinforces our earlier claim that on average, older students are more likely to graduate in any given period as their younger counterparts, everything else held constant. Looking at the direct effect of the credit milestones (column (b)), we note that the effects are all positive and significant and the largest effect is associated with earning 20 non-remedial credits as suspected. The results suggest that, on average, earning 20 non-remedial credits increases a younger student’s odds of graduating in any given trimester by a factor of 7.6 over a similar younger student who has not reached this credit milestone. The interaction coefficients yield information on whether the milestone affects older and younger students differently (column (c)). If this term is close to one, it indicates that there is not much of an observed difference between the two groups of students. Younger and older students seem to respond similarly to reaching the 10-credit milestones, and this is noted by statistically insignificant interaction terms. However, earning 20 credits has different effects on older and younger students. This milestone is significantly less important for older students than younger ones. For example, a younger student who receives 20 non-remedial credits is 7.6 times as likely to graduate as a younger student who does not receive 20 non-remedial credits. An older student who receives 20 non-remedial credits is only 4.9 times as likely to graduate as an older student without 20 credits. There is still a significant boost in odds of graduating for both age groups, but the milestone has a more positive effect on younger students. In fact, older students are less sensitive to all the credit milestones than their younger counterparts.

Although we have been controlling for program length, using the credit count may not be the best way to measure progress milestones. The next set of estimates in Table 5 illustrates the impact on odds from completing certain percentages of the program. Like the credit milestones, the direct effect of the pathway grows as program percentage increases, as expected. We see interaction terms for all models that are similar in magnitude to the interaction terms estimated with the credit milestones, but only the term indicating that a student has finished 50% of her program is significantly different from no effect. Here we see again that the impact of finishing 50% of the program on odds of graduating in any given term is smaller for older students than for younger ones, though the impact is clearly positive for both. An older student who completes 50% of her program increases her odds of graduating in any given term by a factor of 11.5 whereas a younger student increases her odds by a factor of 15.5.

The two measures related to remediation are summarized next. Our first remediation measure is a time-invariant dummy indicating whether or not a student enrolled in any developmental education classes. As expected, any stint in remediation significantly lowers the odds of graduating in any trimester since the student is not obtaining credit that counts towards the degree. The interaction term indicates that remediation does affect the two cohorts differently. Younger students who enroll in remedial courses are 0.58 times as likely to graduate as a younger student who does not take these courses. Older student who need remediation change their odds of graduating in any term by 0.77 compared to older student who do not enroll in remediation. This tells us that the impact of taking developmental classes is less detrimental (and significantly so) to the probability of graduating for an older student than for a younger student. This may reflect the varying motivations and goals of older students. Older students may not let academic challenges deter them as much as younger students. Also, it may be the case that, since they have been out of school for longer, older students are more likely to need remediation because their basic skills are merely "rusty" rather than grossly deficient.

A second variable addressing the effects of remediation - this one time-varying - was also tested. This dummy variable turns on only if the student is enrolled in remediation in a particular time period. Like our other remediation measure, this one suggests that although the chances of someone graduating in a given term while taking remedial classes are significantly lowered, the impact is more negative for younger students than older students. The variable offers a slightly different graphical interpretation as shown in Figure 3. Older students who do not take remediation are represented by the top curve, but any term in which the student enrolls in remedial classes relegates the student to the third curve from the top. Students in both groups can bounce back and forth between curves depending on their choice of classes in each term. Remediation, as expected, puts both younger and older students on the lower set of curves due to its negative effects on completion. The interaction term indicates that the difference between older student profiles is smaller than the difference for younger students.

[Figure 3 here]

Our final pathways variables concentrate on the effects of retention. To accomplish this we include a post-interruption indicator and then separate indicators specifying fall-spring and fall-fall retention as defined above. Parameter estimates shown in Table 5 indicate that although older students seem to be less impacted by our retention measures, only the interaction with fall-to-fall retention is significant. The interruption measure significantly lowers the odds of graduating in any given time period (by a factor of 0.25), but there does not seem to be a significantly different effect for older students. Fall-to-spring retention is also seen to almost double the odds of graduating thereafter but also does not have a significant effect on older students. We concentrate here on fall-fall retention, which impacts odds of completion for younger students by a factor of 2.6. Older students do not see so much of a benefit as their odds of graduating only increase by a factor of 2. Figure 4 reports the fall-to-fall retention results graphically. The general pattern we have seen thus far is preserved, and the distance between the functions is greater for younger students than for older ones as indicated by the interaction term less than on in magnitude.

[Figure 4 here]

V. Discussion and Final Remarks

In this paper we have conducted an analysis that tests for differences between age cohorts of factors that affect community college completion. Using a single-risk discrete-time survival methodology, we have found that younger and older students do in fact respond differently to reaching credit milestones, taking remedial courses, and interrupting enrollment. Reaching 20-credit milestones more positively impacts the probability of graduation for a younger student than for an older student. We also find that although remediation decreases the odds of graduating in any given term, older students who enroll in remediation are less negatively impacted than younger ones who do the same. Our final analysis of educational pathways shows that fall-to-fall retention positively impacts the odds of graduating, but is not as important for older students as it is for younger students. In addition to the above, we should reiterate our somewhat strange finding that after controlling for ability, older students have a higher probability of graduating in any given trimester. Knowledge of age effects and how educational milestones and pathways affect the cohorts differently can help colleges derive policies that will better accommodate both groups of students.

References

Adelman, C. (2004). *Principal Indicators of Student Academic Histories in Postsecondary Education, 1972-2000*. Washington, D.C.: U.S. Department of Education.

Adelman, C. (2005). *Moving into town – and moving on. The community college in the lives of traditional-age students.* Washington, DC: U.S. Department of Education.

Allison, P. (1984). *Event history analysis: Regression for longitudinal event data*. Thousand Oaks, CA: Sage Publications.

Bailey, T., Calcagno, J., Jenkins, D., Keinzl, G. and Leinbach, T. (2005). *Community college student success: What institutional characteristics make a difference?* New York: Columbia University, Teachers College, Community College Research Center.

Bean, J., and Metzner, B. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research, 55*, 485–540.

Cohn, A. and Brawer, F. (1996). *The American community college*: Third Edition. Jossey-Bass Inc., Publishers. San Francisco.

Choy, S. (2002). *Nontraditional undergraduates: Findings from the condition of education, 2002.* Washington, D.C.: National Center for Education Statistics.

Choy, S. and Premo, M. (1995). *Profile of older undergraduates, 1989-90.* Washington, D.C.: National Center for Education Statistics.

Crosta, P., Calcagno, J., Bailey, T. and Jenkins, D. (2006). *Does age of entrance affect community collage completion probabilities? Evidence from a discrete-time hazard model*. New York: Columbia University, Teachers College, Community College Research Center.

DesJardins, S. 2003. Event history methods: Conceptual issues and an application to student departure from college. In J. Smart (ed.), *Higher Education: Handbook of Theory and Research* 18: 421-471.

DesJardins, S., Ahlburg, D., and McCall, B. (1999). An event history model of student departure. *Economics of Education Review, 18,* 375-390.

Dynarski, S. (2003). Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review,* 93(1): 279-288.

Ehrenberg, R. (2004). Econometric studies in higher education. *Journal of Econometrics*, *121*(1), 19 – 37.

Ehrenberg, R. and Smith, C. (2004). Analyzing the success of student transitions from 2- to 4-year institutions within a state. *Economics of Education Review,* 23(1): 11-28.

Grubb, W. (2002). Learning and earning in the middle, part I: national studies of pre-baccalaureate education. *Economics of Education Review, 21*(4), 299-321.

Horn, L., and Carroll, C. 1996. *Nontraditional undergraduates: Trends in enrollment from 1986 to 1992 and persistence and attainment among 1989-90 beginning postsecondary students.* Washington: National Center for Education Statistics, U.S. Department of Education.

Jacobson, L., LaLonde, R., and Sullivan, D. (2004). Estimating the returns to community college schooling for displaced workers. *Journal of Econometrics*, 125, 271-304.

Kane, T. and Rouse, C. (1999). The community college: Educating students at the margin between college and work. *Journal of Economic Perspectives*, 13(1), 63-84.

Light, A. and Strayer, W. (2000). Determinants of college completion: School quality or student ability*? Journal of Human Resources, 35*(2), 299-332.

Manski, C. and Wise, D. (1983). *College choice in America*, Boston, MA: Harvard University Press.

Rouse, C. (1995). Democratization or diversion? The effect of community colleges on education attainment. *Journal of Business and Economic Statistics,* 13(2): 217-224.

Sandy, J., Gonzalez, A. and Hilmer, M. (in press). Alternative paths to college completion: Effect of attending a 2-year school on the probability of completing a 4-year degree. *Economics of Education Review*

Scott, M. and Kennedy, B. 2005. Pitfalls in pathways: Some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics*, 30(4): 413-442.

Turner, S. (2004). Going to college and finishing college. In C. Hoxby (ed), *College choice: The economics of where to go, when to go, and how to pay for It*. Chicago: University of Chicago Press.

U.S. Department of Education, National Center for Education Statistics. (2003). *Integrated postsecondary education data system—Fall enrollment survey: 2002* [Data File].

Singer, J. and Willett, J. 2003. *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Tables and Figures

Table 1: Descriptive Statistics Demographics and First Trimester Student Characteristics

Characteristic	Younger	Older	Difference
Female	52.38	59.34	-6.96*
Age	18.25	33.93	-15.68*
Race			
Black	16.59	18.84	-2.25*
Asian or Pacific Islander	2.86	2.25	0.61*
Hispanic	19.25	15.82	3.43*
American Indian	0.46	0.65	-0.19
White	60.48	62.31	-1.84*
Unknown Race	0.36	0.12	0.24*
US citizen	88.79	82.61	6.18*
High School Credential			
HS Diploma	85.67	68.00	17.67*
GED	6.06	26.85	-20.78*
Other HS credential	0.57	0.09	0.48*
Placement test scores (200-800 scale)			
Mathematics	415.17	327.68	87.50*
Verbal	447.38	476.00	-28.62*
Received federal aid (term 1)	26.73	35.93	-9.20*
Tuition (term 1)	1309.68	1316.47	-6.79*
Full-time (term 1)	65.55	31.00	34.55*
Program length (term 1)	60.38	58.88	1.50*
Credits enrolled (term 1)	7.66	5.58	2.09*
Credits earned (term 1)	5.76	4.50	1.26*
Developmental credits enrolled (term 1)	3.60	2.54	1.06*
Developmental credits earned (term 1)	2.52	1.85	0.67*
Total credits earned (term 1)	8.28	6.35	1.93*
Ratio total credits earned/credits enrolled (term 1)	0.72	0.78	-0.62*
Number of Observations†	29,421	5,652	

* denotes significance at 0.01, two-tailed test, unequal variances
† Sample sizes are slightly smaller for High School Credential variables due to missing information.

Table 2: Outcome and Educational Pathway Descriptive Statistics

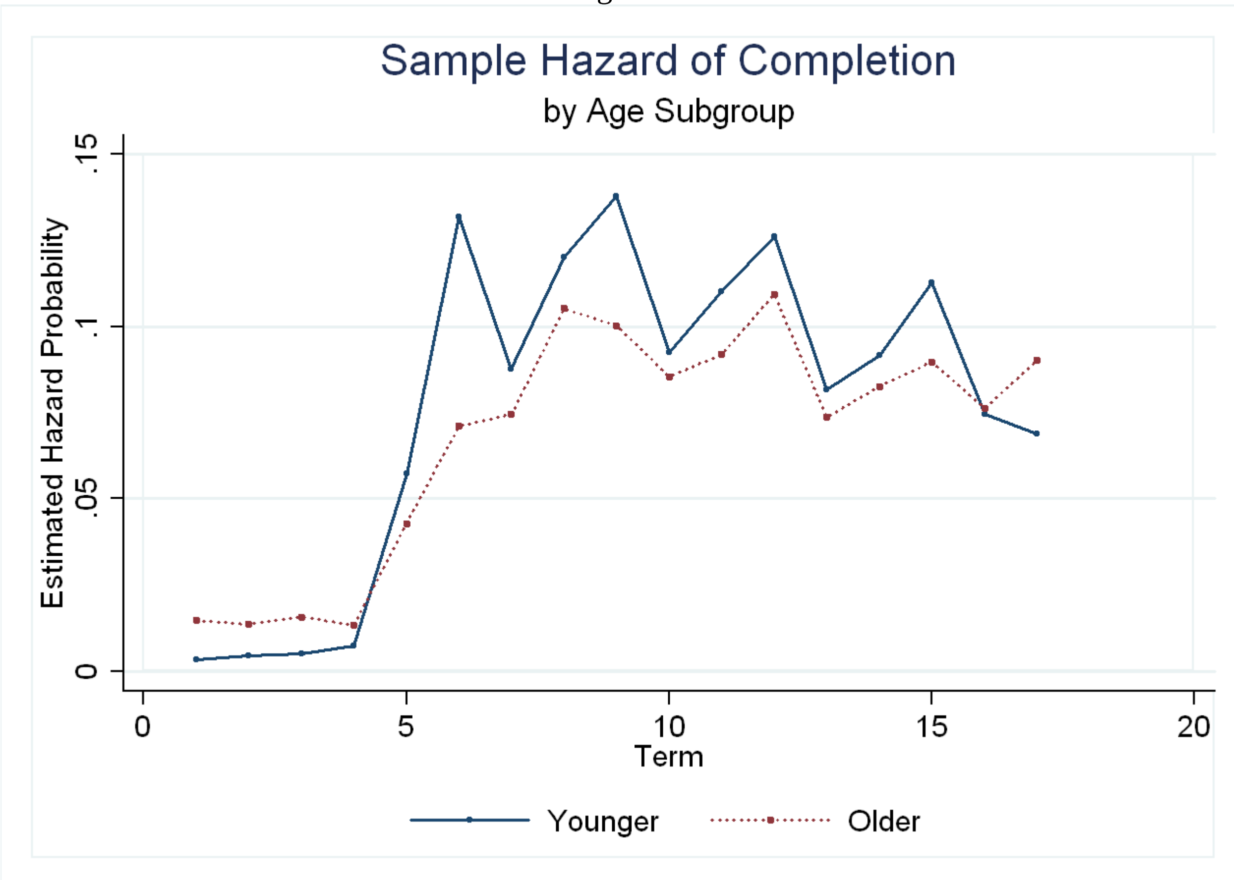
Outcomes (events) / Enrollment Pathways	Younger	Older	Difference
Outcome			
Completion in 17 terms	29.92	19.04	10.88*
Nominal Credit Milestones			
Earned 10 credits	78.71	61.50	17.21*
Earned 10 non-remedial credits	70.98	53.72	17.27*
Earned 20 credits	65.47	44.55	20.92*
Earned 20 non-remedial credits	59.13	39.61	19.52*
Percentage of Program Completion			
Finished 5% of program	88.11	80.91	7.20*
Finished 15% of program	77.17	65.55	11.61*
Finished 25% of program	69.12	56.74	12.38*
Finished 50% of program	55.46	43.97	11.49*
Finished 75% of program	46.70	36.75	9.95*
Remediation			
Enrolled in remediation	61.43	60.19	1.24
Enrolled in remediation in term 1	55.73	48.66	7.07*
Retention			
Missed 2nd and 3rd term (Post-interruption for term 4)	24.37	35.05	-10.68*
Fall to spring	70.44	59.66	10.78*
Fall to spring to fall	52.13	36.25	15.88*
Number of Observations	29,421	5,652	

* denotes significance at 0.01, two-tailed test, unequal variances

Table 3: Life Tables for Full Sample and Age Cohorts

		Full Sample			Younger			Older		
Term	Date	At risk	Completed	Hazard Function	At Risk	Completed	Hazard Function	At risk	Completed	Hazard Function
1	Fall-98	35073	185	0.0053	29421	101	0.0034	5652	84	0.0149
2	Spring-99	24077	142	0.0059	20711	96	0.0047	3366	46	0.0136
3	Summer-99	12497	84	0.0067	10599	54	0.0051	1898	30	0.0158
4	Fall-99	20608	167	0.0081	18172	134	0.0073	2436	33	0.0136
5	Spring-00	18402	1026	0.0558	16326	937	0.0574	2076	89	0.0429
6	Summer-00	9795	1221	0.1246	8642	1139	0.1318	1153	82	0.0711
7	Fall-00	13897	1200	0.0864	12370	1086	0.0878	1527	114	0.0746
8	Spring-01	11422	1352	0.1183	10139	1217	0.1201	1283	135	0.1052
9	Summer-01	6148	821	0.1336	5440	750	0.1378	708	71	0.1003
10	Fall-01	8402	771	0.0917	7455	690	0.0926	947	81	0.0856
11	Spring-02	7048	762	0.1081	6232	687	0.1102	816	75	0.0919
12	Summer-02	3992	496	0.1243	3535	446	0.1262	457	50	0.1094
13	Fall-02	5388	435	0.0807	4764	389	0.0816	624	46	0.0737
14	Spring-03	4525	410	0.0906	3993	366	0.0917	532	44	0.0827
15	Summer-03	2755	303	0.11	2421	273	0.1128	334	30	0.0898
16	Fall-03	3706	277	0.0748	3287	245	0.0745	419	32	0.0764
17	Spring-04	3163	226	0.0714	2786	192	0.0689	377	34	0.0902

Figure 1



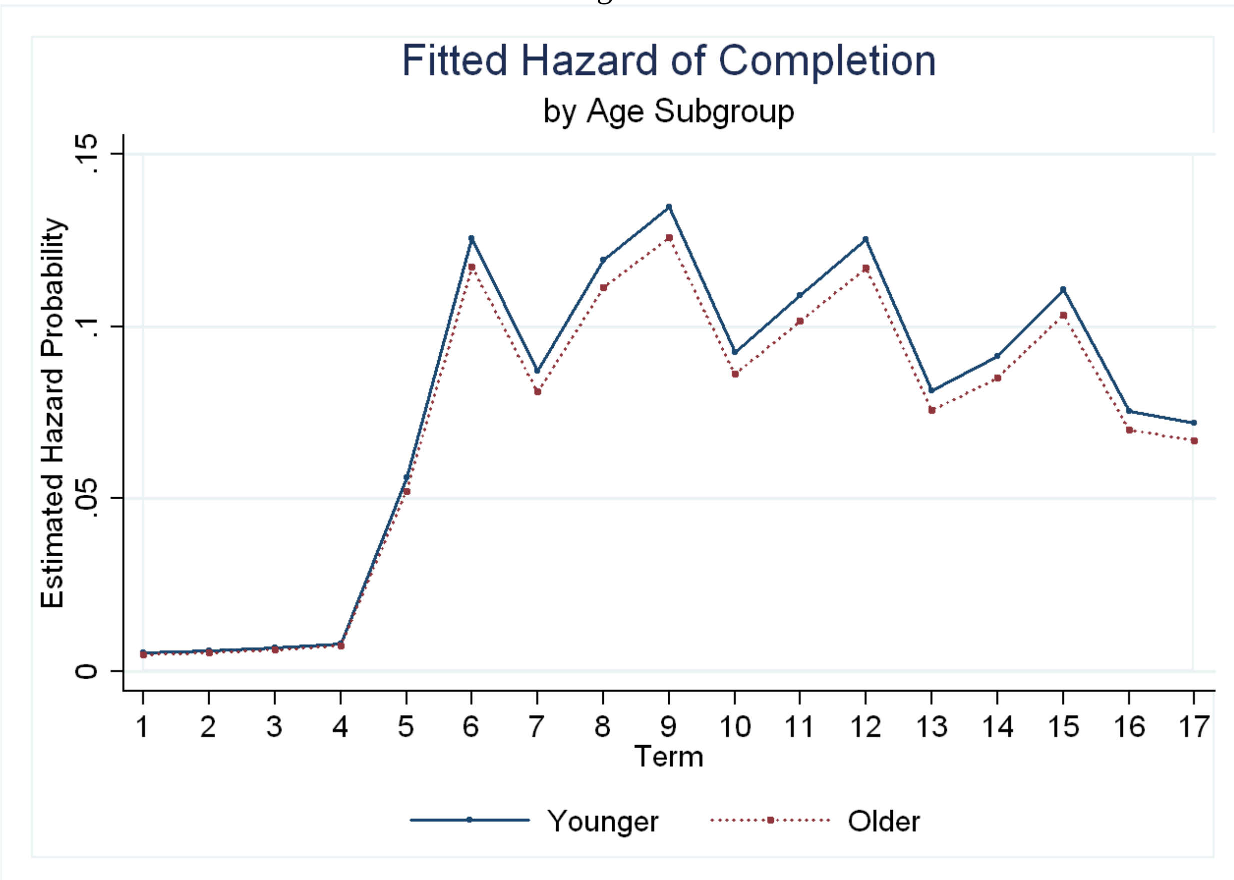
Sample Hazard of Completion
by Age Subgroup
Estimated Hazard Probability
0
.05
.1
.15
0
5
10
15
20
Term
Younger
Older

Table 4: Estimated Odds Ratios for Hazard Models, Outcome is Completion

Variables	Odds (SE) (1)	Odds (SE) (2)	Odds (SE) (3)
Older student	0.925 (.031)	0.94 (.03)	1.31 (.06)*
Female		1.08 (.02)*	1.28 (.03)*
Black		0.48 (.02)*	0.65 (.03)*
Asian		0.93 (.06)	0.94 (.07)
Hispanic		0.58 (.02)*	0.68 (.02)*
American Indian		0.61 (.10)*	0.64 (.12)
No Race		0.78 (.14)	0.89 (.17)
US Citizen		1.04 (.04)	0.95 (.04)
High School Diploma		1.07 (.03)	1.08 (.04)
Received Federal Aid in term 1		1.02 (.03)	1.08 (.03)*
Verbal score			1.21 (.02)*
Verbal score^2			0.93 (.01)*
Math score			1.58 (.03)*
Math score^2			0.96 (.01)*
Tuition in term *j*			0.99 (.01)
Full-time in term *j*			1.44 (.04)*
Program length in term *j*			0.98 (.00)*
Number of Observations	190898	186266	159845
Number of Groups	35073	34004	27730
Deviance (-2*Log Likelihood)	67850	65422	54085
Change in Deviance†	-5.34**	-2428*	-11337*

* denotes significance at 0.01, Chi-squared test, 10 and 15 d.f., respectively. Standard errors are in parenthesis. All models include 17 time dummies. † Column (1) difference tested from a model with only time-dummies.

Figure 2



Fitted Hazard of Completion
by Age Subgroup
Estimated Hazard Probability
0
.05
.1
.15
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
Term
Younger
Older

Table 5: Estimated Odds Ratios for Hazard Models, Pathways and Interactions

Dependent Variable: Completion	Older Cohort (S.E.)	Pathway (Impact for Younger Students) (S.E.)	Interaction (Older * Pathway) (S.E.)	Impact for Older Students	Model Fit	
					Deviance	Change in Deviance†
	(a)	(b)	(c)	(d) = (b) * (c)	(e)	(f)
Credits						
Earned 10 credits	2.019 (.621)*	4.447 (.536)*	0.640 (.198)	2.846	53827	257.9*
Earned 10 non-remedial credits	1.746 (.418)*	4.911 (.457)*	0.730 (.177)	3.585	53582	503.5*
Earned 20 credits	2.142 (.362)*	6.737 (.561)*	0.593 (.103)*	3.995	53111	974.1*
Earned 20 non-remedial credits	1.938 (.292)*	7.595 (.538)*	0.645 (.100)*	4.899	52530	1554.9*
Program percentage						
Finished 5% of program	2.080 (.816)	2.621 (.415)*	0.625 (.246)	1.638	54033	52.0*
Finished 15% of program	1.947 (.481)*	4.108 (.404)*	0.655 (.163)	2.691	53751	334.2*
Finished 25% of program	1.725 (.333)*	6.151 (.495)*	0.728 (.143)	4.478	53183	901.9*
Finished 50% of program	1.684 (.246)*	15.463 (1.020)*	0.743 (.112)*	11.487	50211	3874.3*
Finished 75% of program	1.559 (.185)*	31.496 (1.725)*	0.813 (.102)	25.601	44630	9455.3*
Remediation						
Enrolled in remediation ever	1.073 (.086)	0.575 (.018)*	1.333 (.124)*	0.766	53781	304.8*
Enrolled in remediation in term *j*	1.240 (.057)*	0.153 (.017)*	1.702 (.431)*	0.260	53491	594.8*
Retention						
Interruption	1.230 (.061)*	0.254 (.009)*	0.934 (.109)	0.237	52217	1867.9*
Fall-spring retention	1.488 (.163)*	2.042 (.081)*	0.873 (.102)	1.783	53695	390.6*
Fall-fall retention	1.598 (.141)*	2.643 (.091)*	0.773 (.076)*	2.043	53115	970.5*

* denotes significance at 0.05 level. Standard errors are in parenthesis. All models include 17 time dummies
† Deviance tested against Table 4, column (3)
Note: Each line is a separate regression

Figure 3

Estimated Probability of Completion
By Age and Time-Varying Remediation

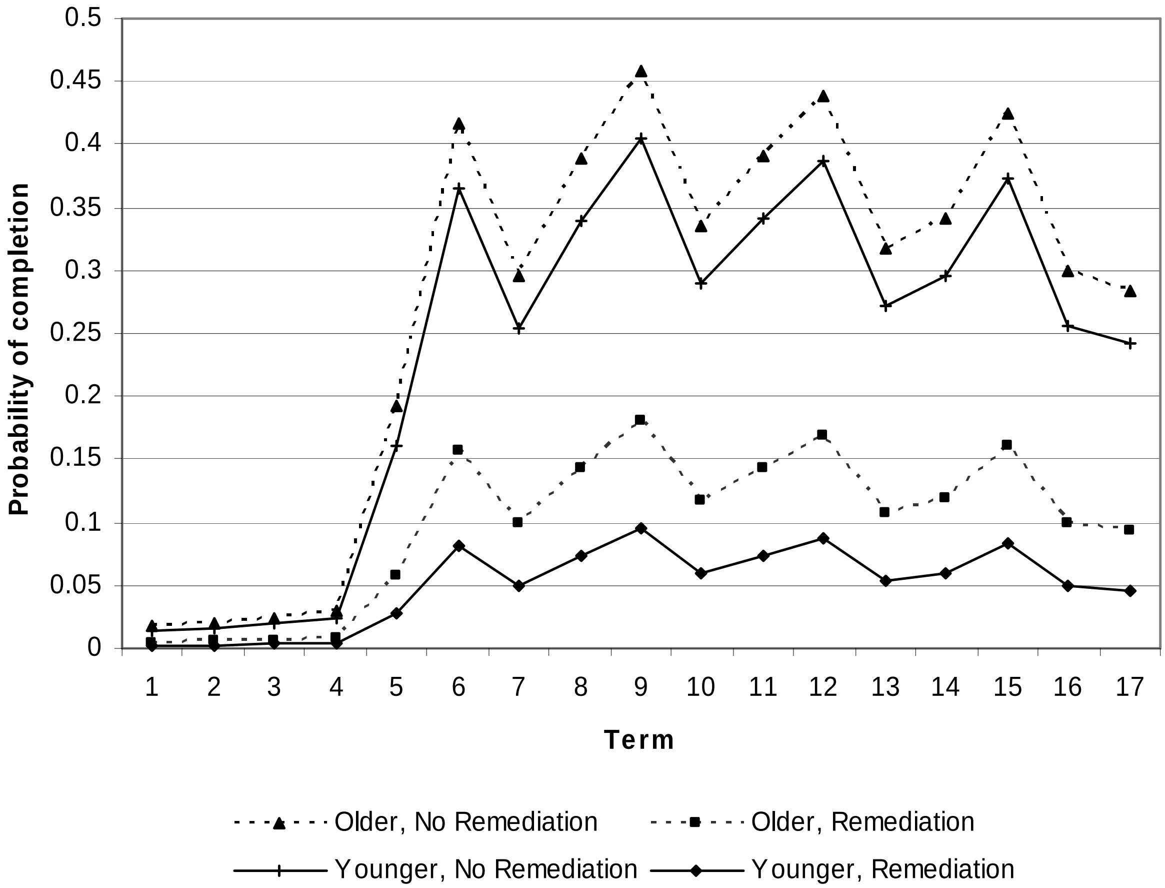


Figure 4

Estimated Probability of Completion
By Age and Fall-to-fall Retention