COLUMBIA | SIPA
Center for Development Economics and Policy

CDEP-CGEG WORKING PAPER SERIES

CDEP-CGEG WP No. 20

Time Gaps in Academic Careers

Evan Riehl

December 2015

COLUMBIA | SIPA
Center on Global Economic Governance

TIME GAPS IN ACADEMIC CAREERS

EVAN RIEHL*

DECEMBER 2015

Abstract. Forgiving education systems create churning by allowing students to defer the completion of their schooling. This paper asks if time gaps in academic careers can lower educational attainment. I study an academic calendar shift in Colombia that created a one semester gap between high school and potential college entry. This brief gap reduced college enrollment rates relative to unaffected regions. Low SES students were more likely to forgo college, and individuals who did enroll after the gap chose higher paying majors. Thus academic time gaps can affect both the mean and the distribution of schooling, with implications for wage inequality.

* Department of Economics, Columbia University, 420 W. 118th St., #1022, New York, NY 10027 (e-mail: e.riehl@columbia.edu). For useful comments I thank Peter Arcidiacono, Peter Bergman, Nicolás De Roux, Christopher Hansman, Nandita Krishnaswamy, W. Bentley MacLeod, Jonah Rockoff, Juan E. Saavedra, Judith Scott-Clayton, Miguel Urquiola, and participants at the Columbia University colloquia. I am grateful to Luis Omar Herrera Prada and for invaluable data assistance. All errors are my own.

1. Introduction

A defining feature of the U.S. education system is that it permits second chances. Many European countries have national exams and centralized admissions that assign students to educational tracks, often at young ages. In the U.S., graduation standards have historically been left up to local districts, and colleges vary widely in admission criteria. U.S. students can thus more readily switch academic paths, or return to school after dropping out.

A second-chance system helped the U.S. become an early leader in educational attainment, but it may also partly explain the recent stagnation of high school and college graduation rates (Goldin and Katz, 2008). Lenient standards can weaken accountability and reduce school productivity. This concern has motivated research on accountability systems (Kane and Staiger, 2002; Hanushek and Raymond, 2005; Figlio and Loeb, 2011), teacher evaluation (Rivkin et al., 2005; Chetty et al., 2014), and school choice (Hoxby, 2003; Urquiola, 2016).

This paper explores a different channel through which flexible standards affect academic outcomes. Forgiving education systems create churning by allowing students to defer the completion of their schooling. The result is that many students experience time gaps in their academic careers as they cycle in and out of school. In the U.S., for example, roughly one-third of all students who matriculate in college wait more than one semester after high school to enroll. Such gaps, as measured by age differences between secondary graduates and tertiary enrollees, tend to be larger in OECD countries without educational tracking.

I examine whether academic time gaps can lower schooling attainment. This builds on research that asks how the structure of education systems—e.g., school entry or exit ages—affects attainment (Angrist and Krueger, 1992; Bedard and Dhuey, 2006; Dobkin and Ferreira, 2010; Black et al., 2011; Fredriksson and Öckert, 2014). My focus is on the potential gaps that arise in the transition from high school to college. Individuals' experiences during these gaps may affect their decision of whether to acquire further education at all. Joining the labor force can reveal the job satisfaction and earning potential of an existing degree. Time away from school may make it harder to go back.

Isolating the effect of academic gaps on subsequent enrollment requires quasi-random variation in transition timing in an education system. For this I use administrative data from the country of Colombia and a policy that altered its unique system of academic calendars.

Colombian high schools operate on two distinct annual schedules; some schools begin the academic year in January, while others start in September. All public high schools operate on the January calendar except for those in two regions, which historically began in September. From 2008–2010, these regions transitioned public schools to the January calendar. This altered the academic term at nearly 400 high schools that I call *switching schools*, which include all public schools in the affected regions plus some local private schools that followed suit. Other private high schools in these regions did not switch calendars. I use the term *staying schools* to refer to 89 local private schools that stayed on the September calendar.

Colombian colleges offer admission every semester, so students on either high school calendar can typically begin college right after graduation. But the calendar transition led to an unusual gap before potential college entry at both switching and staying schools, for separate reasons. Switching schools transitioned to the new calendar by adding mid-term breaks and delaying graduation by a few months each year. This caused one cohort—2009 graduates—to finish high school just after the start of the September semester at most colleges. Thus, many 2009 graduates could not enter college until January—one semester later than was typical for prior cohorts at switching schools.

Students from staying schools were also affected by the transition. Colombian students apply to both a college and a major, and while many programs are offered twice per year, some are annual. To align with the new public high school calendar, colleges in the affected regions shifted some annual majors from a September to a January start date. Graduates from staying schools who were interested in these programs therefore had to wait an extra semester to enroll. This yielded a second source of post-graduation time gaps, induced in this case by changes in college calendars rather than high school calendars.

At both switching and staying schools, the calendar changes led to sharp declines in the number of students who began college in the first semester after graduation. I show this in a differences in differences analysis with students in unaffected regions as a comparison group. This initial enrollment reduction is a first stage result; the calendar shift provides quasi-random variation in the occurrence of one semester gap between high school and college.

I then consider how this time gap affects subsequent college enrollment. Students who did not enroll in college within one semester could typically have done so within one year. Many did not. Roughly 50 percent of those exposed to the time gap did not enroll in the next possible semester. This occurred at both switching and staying schools, leading to a seven percent reduction in college attendance rates in affected regions. Further, there is little catch-up enrollment in the next two years, suggesting that the decline is not merely temporary.

The decline in college enrollment is evident graphically and in regressions, and it is robust to different comparison groups and to the inclusion of linear trends for each high school. Staying schools provide a clean test of the timing effect because their graduates had few changes in college preparation. Graduates from switching schools were affected by other elements of the transition, such as a longer academic year. But I find little evidence that these other channels are important; the results do not appear to be driven by changes in graduates' entrance exam scores or admission rates at selective colleges.

I explore two sources of heterogeneity that inform the main results. First, I find that low SES and low ability students are more likely to forgo college after the time gap. These students also had lower college enrollment rates in the pre-policy years. Thus time gaps may have a greater effect on individuals who are indecisive about college in the first place.

Second, I show that the enrollment decline comes from students forgoing low-paying majors; those who entered college after the gap tended to choose high-wage majors. This could arise if working or job-hunting in the interim altered students' relative valuations of different degrees. Consistent with this, I use household survey data to show that labor force participation increased among 17 year olds in the affected regions during the gap period.

The bottom line is that time gaps in education systems can affect both the mean and the distribution of schooling, with potential implications for a nation's wage growth and inequality (Goldin and Katz, 2008; Acemoglu and Autor, 2011).

My paper relates to four research areas in labor and education economics. First, it contributes to work on institutional features that affect transitions between education tiers. Some researchers have argued that the U.S.'s disconnected education system, in which K–12 and postsecondary schools function as separate entities, can hinder student outcomes (e.g., Venezia et al., 2003). This concern underlies research on the benefits of educational tracking (Hanushek and Woessmann, 2006; Schütz et al., 2008) and remediation classes (Bettinger and Long, 2009; Martorell and McFarlin Jr, 2011).

I show that timing is another potential disconnect in the transition from high school to college. This adds to work on the "summer melt," a phenomenon in which students who are planning to attend college at the time of high school graduation change their minds by the end of the summer (Castleman and Page, 2014). My findings suggest that tracking or curriculum policies can have additional attainment benefits if they limit time gaps in students' academic careers. Further, policies that promote college access may have a greater impact if they are implemented before students finish compulsory education.

Second, this paper adds to research on the information hurdles that affect schooling choices. Recent work finds that students' decisions respond to information on tuition costs and financial aid (Bettinger et al., 2012; Hoxby and Turner, 2013; Dynarski and Scott-Clayton, 2013), returns to education (Jensen, 2010; Nguyen, 2010; Oreopoulos and Dunn, 2013; Dinkelman and Martínez A., 2014), and school test scores (Hastings and Weinstein, 2008).

In my paper, the information that alters enrollment decisions is acquired not through interventions but by experiences outside of school. This is consistent with Perez-Arce (2015), who finds that applicants with deferred admission to a Mexican university are more likely to forgo college if they work in the interim. Goodman (2013) also shows that the experience of taking the American College Test (ACT) induces some students to attend college. More broadly, my findings contribute to work on how routines/defaults (Scott-Clayton, 2011;

Pallais, 2015) or present biases (Angrist and Lavy, 2009; Angrist et al., 2009; Fryer, 2011) affect educational choices, which may explain the time gap effects.

Third, recent research finds that students' choice of major depends on expected earnings (Arcidiacono et al., 2012; Hastings et al., 2015), initial beliefs (Stinebrickner and Stinebrickner, 2014), and heterogenous tastes (Arcidiacono, 2004; Beffy et al., 2012; Wiswall and Zafar, 2015). I present evidence that program choices change when individuals spend time outside of the education system. This finding parallels that in Zafar (2011), who shows that students update their major preferences based on their performance in college.

Finally, my paper contributes to work on education externalities from labor market conditions. Booms arising from coal prices (Black et al., 2005), trade reforms (Atkin, 2012), and the housing market (Charles et al., 2015) can lead individuals to forgo further schooling. My results suggest that education externalities can also arise from an individual's labor force participation prior to the completion of her academic career.

The next section gives background on the prevalence of time gaps between high school and college in different education systems. Section 3 describes the academic calendar shift in Colombia. Section 4 discusses the resulting effects on schools that switched to the January schedule. Section 5 shows the effects on schools that stayed on the September calendar. Section 6 explores heterogeneity in these results at both school types. Section 7 concludes.

2. Time gaps between high school and college

Every spring roughly three million U.S. students receive their high school diplomas. Approximately two-thirds of them enroll in a two- or four-year college in the fall of the same year. This ratio is sometimes called the *immediate college enrollment rate*, and in the U.S. it has hovered between 60 and 70 percent for the past two decades.[1]

For graduates who do not immediately attend college, the decision to forgo further education is not necessarily a permanent one. The dashed line in Figure 1 illustrates this using the 1997 National Longitudinal Survey of Youth (NLSY). It plots the cumulative fraction of graduates who have ever enrolled in college against the number of years since their high school graduation. The immediate college enrollment rate for NLSY graduates is slightly below 60 percent, as indicated by the data point at 0.5 years after graduation. The fraction ever enrolled rises further after this initial spike, reaching nearly 80 percent after nine years. This value reflects the *overall college enrollment rate*, which would be higher if the data included enrollment at ten or more years.[2]

[1] This statistic is from the National Center for Education Statistics (NCES) (available in December 2015 at http://nces.ed.gov/programs/coe/indicator_cpa.asp).

[2] The overall enrollment could be substantially higher; an NCES report finds that 12 percent of the 1995 incoming cohort waited ten or more years to enroll (available in December 2015 at http://nces.ed.gov/pubs2005/2005152.pdf, Figure I).

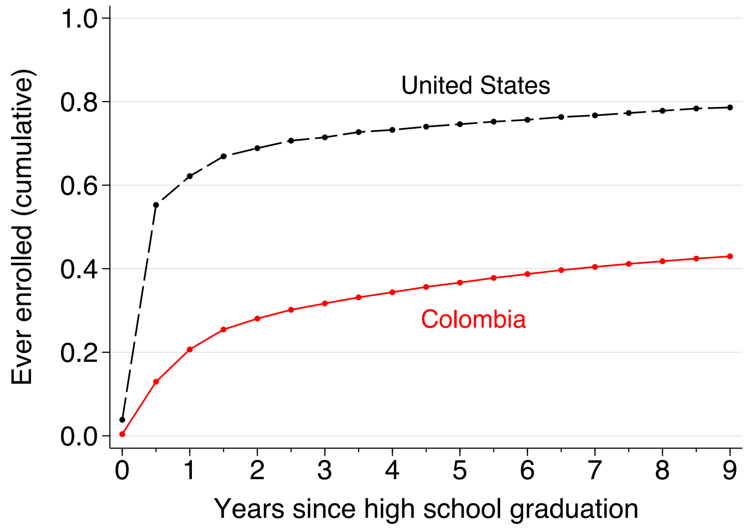


FIGURE 1. College enrollment timing

Notes: The U.S. sample is students in the 1997 NLSY cross-sectional and over-samples with a high school graduation year in 1995–2002. The x-axis is the difference between college enrollment and high school graduation dates, aggregated into six month bins. Averages use 2011 panel weights.

The Colombian sample is any 11^{th} grader who took the national college entrance exam in 2001–2002. The x-axis is the difference between the semesters of college enrollment and the exam. See Section 3 for details on the Colombian data, entrance exam, and academic calendar structure.

For both countries, I count a student as ever enrolled if she enters a two- or four-year college in the dataset within nine years of her graduation month. Enrollment at zero years includes students reported to begin college in or before the month of high school graduation.

The solid line in Figure 1 replicates this curve for Colombian high school graduates using administrative data described below. Both immediate and overall college enrollment are substantially lower in Colombia than in the U.S.

Figure 1 shows that students in both countries experience time gaps between high school and college. In the U.S., 30 percent of students who eventually attended college did not enroll in the first semester after graduation. Delayed enrollment is even more common in Colombia, comprising more than two-thirds of all college-goers. Further, Appendix A.1 shows that these time gaps vary with students' backgrounds. Low SES and low ability students are more likely to delay college entry, as are students who are older than the normative graduation age. This pattern holds in both the U.S. and Colombia.[3]

Figure 2 explores how post-secondary gaps vary across education systems. Since longitudinal data like that in Figure 1 do not exist in some countries, it uses age gaps as a proxy for time gaps. Specifically, it plots the difference between the average ages of new tertiary entrants and secondary graduates for OECD countries with data in 2011 (plus Colombia).[4]

[3] I find that gender is not a significant predictor of delayed enrollment in either country.

[4] Figure 2 includes only students enrolled in the highest tiers of secondary (general programs) and tertiary (level A, further education/theoretically based) school. It does not therefore capture age gaps for students pursuing technical education, which may vary across countries. The sample also includes only students below

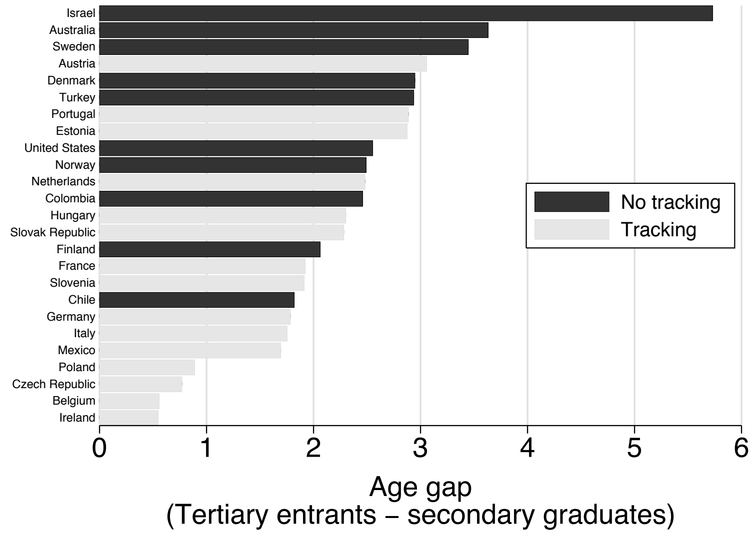


FIGURE 2. OECD countries — Secondary/tertiary age gap

Notes: Data are from the OECD for the year 2011 (available in December 2015 at http://stats.oecd.org). Bars depict the difference between the average ages of new entrants to level A tertiary programs and graduates from upper secondary general programs. In the U.S., upper secondary includes all education programs because there is no distinction between general and vocational/technical programs. In Colombia, the age difference is between entrants to university-level programs and 11^{th} grade college entrance exam takers in 2011 using the data described in Section 3. All secondary and tertiary averages are calculated using students age 29 or below, for which single-year ages are available from the OECD.

Light grey bars depict countries that track students into school types before age 16. This classification follows Hanushek and Woessmann (2006) and their cited source (available in December 2015 at http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/34002216.pdf, Figure 5.20a). Countries not in this source are classified as in Education Policy Outlook 2015 (available in December 2015 at http://www.oecd.org/publications/education-policy-outlook-2015-9789264225442-en.htm).

Following Hanushek and Woessmann (2006), I classify countries based on early age tracking of students into different school types. Light grey bars depict countries with tracking below age 16; black bars indicate countries without tracking by this age.

Age gaps between secondary and tertiary education vary significantly across countries, ranging from below one to nearly six years. Further, countries with tracking systems tend to have smaller age gaps. The smallest gaps occur in countries with early tracking; these include Belgium, the Czech Republic, Germany, and Mexico, all of which begin tracking by age 12. Colombia, the U.S., and Scandinavian countries do not have tracking systems and are near or above the median. Other institutional features also affect age gaps. The large gap in Israel is due in part to its military service requirement. In Ireland—the country with the smallest gap—most secondary students participate in a "transition year" of non-academic subjects and volunteer work before graduating.

age 30; without this restriction the gaps would be larger, and the country ranking may change. Note that the Colombian gap is calculated from the data described in Section 3, which may affect comparability.

Figure 2 suggests that different education systems may lead to substantial variation in the existence of gaps between schooling tiers. These gaps can affect overall tertiary enrollment rates if experiences outside of the education system alter students' views about the returns or costs to further schooling. In the U.S., roughly three-quarters of high school graduates who do not immediately enter college join the labor force by the fall of the same year.[5] Employment or job-hunting may alter the perceived value of students' existing education.

The gaps in Figure 2 are likely correlated with other cultural or labor market factors that affect college attendance rates. Isolating the time gap effect on subsequent enrollment requires exogenous variation in delays induced by an education system. In the next section, I describe a policy that altered Colombia's academic calendars that provides an opportunity to test for the causal effect of time gaps after high school.

3. An academic calendar shift in Colombia

This section describes the high school and college calendars in Colombia, my related data sources, and a government policy that altered these calendars in two regions of the country.

3.1. **Academic calendars and college admissions.** The high school system in Colombia is unique in that students begin the school year at two different times. The large majority of schools start the academic year in January and conclude in November. This schedule yields the longest break during the Christmas season. A small number of private high schools begin in September and finish in June. These schools choose a calendar that aligns with that of U.S. and European colleges.

To enter college, Colombian students are required to take a national standardized admission exam called the Icfes.[6] The exam is generally analogous to the SAT in the U.S., but its results are also used to evaluate high schools. As a result, the vast majority of graduates take the exam, even those who do not plan to attend college.[7] The Icfes is offered semiannually at the end of the last year of high school on each calendar. Scores are returned promptly, and students who apply to college can start in the next semester.

Accordingly, Colombian colleges also enroll students two times per year. Students can begin in either January or September, and nearly all colleges offer admission in both semesters.[8] Unlike the U.S., applicants choose both schools and majors at the time of application, but

[5] This statistic is for 2014 graduates and comes from the Bureau of Labor Statistics (available in December 2015 at http://www.bls.gov/news.release/hsgec.nr0.htm).

[6] Icfes stands for Institute for the Promotion of Higher Education, the former acronym for the agency that administers the exam. The Icfes exam is now named Saber 11°, but I use the name Icfes to match the designation during most of the period covered by my data.

[7] The exam agency estimates that over 90 percent of graduates take the exam.

[8] The January cohort is slightly larger at many colleges because most high schools use the January calendar, but the cohort sizes are more balanced than in the U.S., where almost all undergraduates begin in the fall.

the Colombian market is similar in that there are many public and private colleges with varying selectivity and degree durations. Admissions are also decentralized; students must apply to individual colleges, and each institution controls its own selection criteria.

3.2. **Data sources.** This paper uses two administrative data sources:

(1) Records for all high school seniors (11^{th} graders) who took the Icfes college entrance exam from 2001–2011. These data are from the testing agency and contain each student's high school, background characteristics, and exam scores.
(2) Records for students enrolling in college in 2001–2012. These are from the Ministry of Education and cover almost all higher education institutions in Colombia.[9] The records include each student's enrollment date, program of study, and institution.

I merge these two datasets using national ID numbers, birth dates, and names. This defines the measure of college enrollment for my analysis, which is an indicator for appearing in the Ministry of Education records.[10]

In addition, I use two datasets related to students' labor market outcomes. First, administrative data from the Ministry of Social Protection provides 2008–2012 earnings for all college enrollees employed in the formal sector.[11] I use these data to calculate the average earnings of college graduates by major. Second, I use 2007–2010 Colombian household survey data to measure labor force participation by age and region.[12]

3.3. **An academic calendar shift.** Almost all public high schools in Colombia begin the academic year in January, but in two regions the public school year historically started in September. These regions—Valle del Cauca and Nariño—are the third and eighth largest of Colombia's 33 administrative departments, and the capital of Valle del Cauca, Cali, is the country's third most populous city. I call these two departments the *affected regions*. Anecdotally, the affected regions contained some of Colombia's first private high schools, and thus public schools adopted the same September calendar.

In 2008, the affected regions transitioned public schools to the January calendar to align with the federal fiscal year and with all other public schools. The transition occurred over two years by adding extra instructional periods and mid-term holidays (see Figure 3 below for details). The academic calendar shift was complete by January 2011.

Although the policy affected public high schools, some local private schools also changed from the September to the January calendar. Other private schools chose not to change

[9] See Appendix A.3 for details on data coverage.
[10] I match over 91 percent of college enrollees who took the Icfes during the period covered by the data. Appendix A.2 provides details on the merge and match rates.
[11] My administrative earnings records include only college enrollees, so I cannot use these data to examine labor market outcomes of individuals who do not attend college.
[12] The survey is the *Gran Encuesta Integrada de Hogares* (GEIH).

schedules. I refer to the public and private schools that switched calendars as *switching schools*, and the private schools that stayed on the September calendar as *staying schools*.

Columns (A)–(C) in Table 1 describe switching and staying schools in the affected regions. Switching schools include all 290 public schools plus 109 private schools, while 89 private schools stayed on the September calendar.[13] Staying schools are higher quality by several metrics. 97 percent of staying schools received the exam agency's high or superior rank in any year, compared with half of switching schools. 55 percent of staying schools offer academic-level training, while two-thirds of switching schools provide technical education.

Table 1 also shows the characteristics of 2001–2011 graduates from switching and staying schools. There are roughly 25,000 switching school graduates per year, and their average Icfes exam score is near the median of the national distribution. Staying schools graduate less than 4,000 students per cohort, and their average student performs at the 74$^{\text{th}}$ percentile on the Icfes. Staying school graduates also have higher socioeconomic backgrounds as measured by mother's education.

Lastly, Table 1 shows the fraction of graduates who enroll in college, defined as appearing in any institution in the Ministry of Education records. Roughly 10 percent of switching school graduates enroll in the semester immediately after the Icfes exam. This fraction rises to roughly 40 percent by six years after the exam. Both immediate and overall college enrollment are substantially more common in staying schools.

The next two sections explore the effects of the calendar shift on college enrollment using schools in other regions of Colombia as a comparison group (columns (D) and (E) in Table 1). In particular, the calendar transition created a time gap between high school and college at both switching schools and staying schools, but for different reasons. The next section focuses on switching schools; Section 5 discusses staying schools.

4. Switching schools

This section describes how the calendar transition led to a time gap after graduation for one cohort at switching schools. It then presents the resulting effects on college enrollment.

4.1. **Calendar transition and a time gap.** Figure 3 illustrates the transition from the September to the January academic calendar at switching schools.[14] It depicts the 2007–2012 cohorts, defined by the year of the Icfes college entrance exam. The grey bars represent instructional periods in the last year of high school. Prior to 2009, students began the year

[13] Table 1 does not contain all high schools in Colombia because I include only those with exam takers in all years of my analysis (2001–2011). I also omit schools that ever had a "flexible" calendar, in which students can begin the school year in either semester. See Appendix A.4 for details on the high school sample.

[14] Figure 3 depicts academic calendars from public resolutions by the government in Cali, Valle del Cauca. Resolutions from Nariño show a similar proposed transition.

TABLE 1. Summary statistics by school type

	(A)	(B)	(C)	(D)	(E)
	Affected regions			Other regions	
	Switching schools		Staying schools		
	Public	Private	Private	Public	Private
High school characteristics					
# high schools	290	109	89	2,553	1,156
High/superior rank	0.47	0.58	0.97	0.57	0.90
Academic school	0.38	0.33	0.55	0.50	0.66
2001–2011 graduate characteristics					
Total students per year	19,862	5,247	3,727	174,335	56,360
Icfes percentile	0.51	0.54	0.74	0.49	0.67
Mother attended college	0.12	0.26	0.65	0.14	0.49
Younger than 18	0.66	0.76	0.72	0.72	0.82
College enrollment by years since Icfes					
Enrolled within 0.5 years	0.09	0.13	0.28	0.09	0.31
Enrolled within 3 years	0.26	0.35	0.55	0.29	0.60
Enrolled within 6 years	0.33	0.41	0.59	0.37	0.66

Notes: The sample is 11$^{\text{th}}$ graders who took the Icfes in 2001–2011 and attended high schools with exam takers in every year. Columns (A)–(C) show high schools in Nariño and Valle del Cauca; columns (D)–(E) include high schools in all other regions. Switching schools are those on the January calendar in 2010 and/or 2011, and on the September calendar in all other years. Staying schools are on the September calendar in all years. I omit high schools that are ever listed with a "flexible" academic calendar, affected region schools that change calendars before 2010, and schools in other regions that ever change calendars. Private schools are those that are listed as private in any year; public schools are listed as public in all years.

A school is high/superior rank if it ever received one of the exam agency's top three (of seven) ranks in 2001–2008. A school is academic if it is academic or normal in any of these years, and zero if it is always technical or academic & technical. Averages of these variables are weighted by the number of exam takers.

Icfes percentiles are relative to all 11$^{\text{th}}$ grade exam takers in the same year and are calculated using the average of the scores from the six core components that did not change in 2001–2011: biology, chemistry, language, mathematics, philosophy, and physics. Mother attended college equals one if a graduate's mother has any degree above basic secondary; this variable is only available for the 2008–2011 cohorts. Age is calculated at the end of August in the exam year.

College enrollment within t years of the Icfes is defined as a graduate's first appearance in the Ministry of Education college records. These averages are calculated using 2001–2006 graduates.

See Appendices A.3 and A.4 for further details on the included colleges and high schools.

in September, took the Icfes exam in April/May, and graduated in June. They were eligible to begin college in September of their exam year, indicated by the white boxes.

2009 graduates started senior year and took the Icfes on the typical schedule, but they graduated three months later than usual. This occurred through one extra month of classes and two additional mid-term breaks. 2010 graduates also had an extra month of instruction, but their school year began and ended several months later. The transition was complete by 2011, when the academic calendar matched that for all other Colombian public schools.

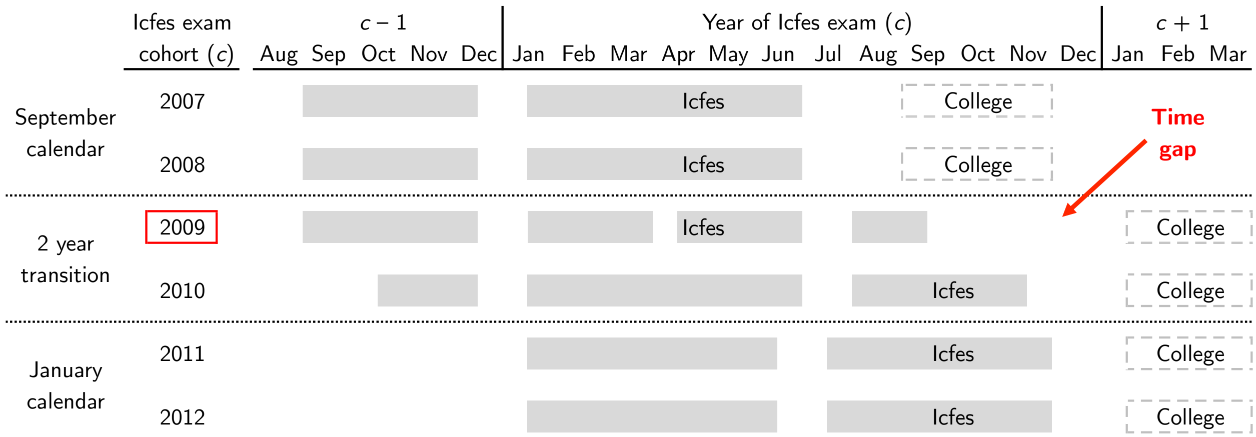


FIGURE 3. Calendar transition and a time gap

Notes: Grey bars are instructional periods in the last year of high school. Gaps are break periods. White boxes represent the first potential college semester. Icfes indicates the timing of the college entrance exam.

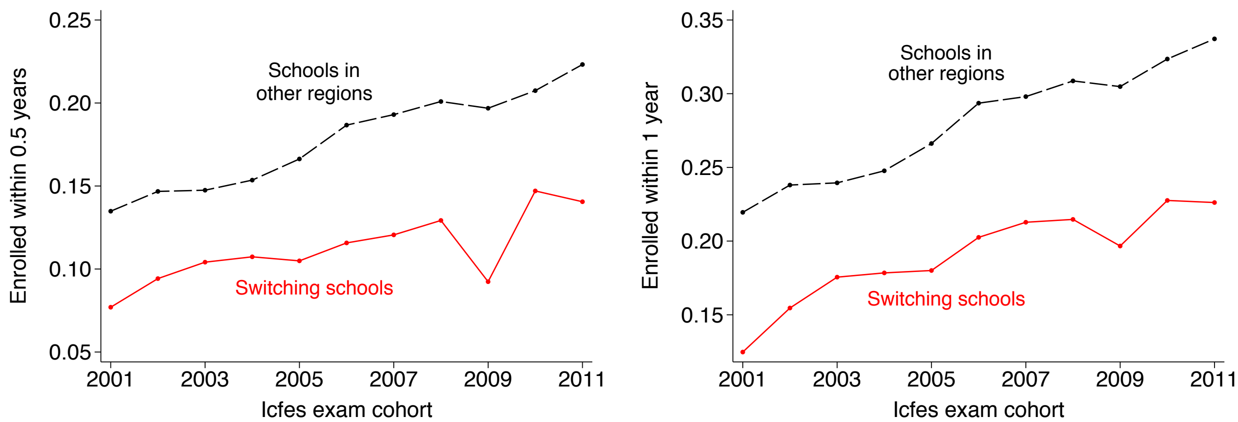
Schedules are approximate based on half-month increments; there are small yearly differences in start/end dates and Icfes timing. This figure is based on 2006–2012 resolutions from the Secretary of Education in Cali, Valle del Cauca. Resolutions from the Security of Education in Nariño show similar schedules.

Figure 3 shows that the transition shifted the first possible college enrollment semester from September to January. This created an unusual four month gap between high school and potential college enrollment for the 2009 cohort, which graduated just after the September college semester began. Below I ask how this time gap altered the decision to enroll in college. I also explore other potential confounding effects such as changes in class time.

4.2. **Benchmark effects on college enrollment.** To study the effects of the time gap, I compare college enrollment among graduates from switching schools (columns (A) and (B) in Table 1) to graduates from other regions of Colombia (columns (D) and (E)).

Figure 4 plots the mean college enrollment rates in switching schools and other schools by Icfes exam cohort. Panel A shows the fraction of graduates entering college one semester after taking the Icfes. There is a sharp drop in the 2009 switching school cohort. This decline in enrollment is in the September college semester, which began before many of these students graduated (see Figure 3). The 2009 enrollment rate is not zero, suggesting that some students or schools graduated early. Nonetheless, many 2009 graduates did not enter college as usual.

The initial decrease in college enrollment in Panel A is a "first stage" result; some 2009 graduates had to wait one extra semester to enroll. Panel B explores the effects of this time gap on enrollment within one year of the exam—the next possible college entrance date. The enrollment decline for the 2009 cohort is less pronounced but still evident; the drop is roughly half of its magnitude from Panel A. This suggests that many students did not enter college following the one semester gap.



Panel A. Enrolled within 0.5 years

Panel B. Enrolled within 1 year

FIGURE 4. Switching schools — College enrollment by years since Icfes

Notes: Switching schools include students from columns (A) and (B) in Table 1. Schools in other regions include students from columns (D) and (E). College enrollment is defined as in Table 1.

Table 2 quantifies these effects using the differences in differences regression

$$y_{hc}^t = \gamma_h + \gamma_c + \beta^t \delta_{hc} + \epsilon_{hc}. \tag{1}$$

The dependent variable, y_{hc}^t, is the fraction of students from high school h and Icfes exam cohort c who enroll in college t years after the exam. The regression includes high school dummies, γ_h, cohort dummies, γ_c, and a treatment indicator, δ_{hc}, which equals one for the 2009 cohort at switching schools. The coefficient of interest, β^t, measures the change in enrollment within t years for 2009 switching school graduates relative to other schools. This is a school-cohort level regression with observations weighted by the number of exam takers.[15]

Panel A of Table 2 displays the β^t coefficients from (1). The columns use dependent variables that measure cumulative enrollment at different durations t. The sample includes only 2001–2009 graduates, for whom I observe enrollment up to $t = 3$ years later.

Column (A) shows that for 2009 switching school graduates, college enrollment in the semester after Icfes declined by about five percentage points. This mirrors the "first stage" result from Panel A of Figure 4. Column (B) shows a three percentage point decline in enrollment within one year. This is consistent with Panel B of Figure 4.

These results suggest that some graduates who would typically have gone to college did not enroll following the time gap. Is this a temporary delay or a permanent decision to forgo further schooling? Columns (C) and (D) measure college enrollment two and three years after the Icfes exam. The point estimates change little relative to column (B). Though the

[15] This yields coefficients identical to those from an individual-level regression but reduces computing time.

TABLE 2. Switching schools — Time gap and college enrollment
Dependent variable: Enrolled in college within t years of the Icfes

	(A)	(B)	(C)	(D)
Panel A. Benchmark specification				
	0.5 years	1 year	2 years	3 years
Switching schools, 2009 cohort (δ_{hc})	-0.049^{***}	-0.030^{***}	-0.028^{***}	-0.027^{***}
	(0.008)	(0.004)	(0.006)	(0.006)
R^2	0.884	0.904	0.911	0.911
Panel B. High school matching				
	0.5 years	1 year	2 years	3 years
Switching schools, 2009 cohort (δ_{hc})	-0.049^{***}	-0.029^{***}	-0.028^{***}	-0.029^{***}
	(0.007)	(0.004)	(0.006)	(0.006)
R^2	0.890	0.909	0.915	0.915
Panel C. Linear high school trends				
	0.5 years	1 year	2 years	3 years
Switching schools, 2009 cohort (δ_{hc})	-0.032^{**}	-0.023^{***}	-0.023^{**}	-0.019
	(0.015)	(0.008)	(0.009)	(0.011)
R^2	0.915	0.932	0.938	0.939
N	36,972	36,972	36,972	36,972
# regions	33	33	33	33
Dependent var. mean	0.161	0.257	0.338	0.382

Notes: The sample is the 2001–2009 Icfes exam cohorts at high schools in columns (A), (B), (D), and (E) of Table 1. The dependent variables are college enrollment within t years of the Icfes, where t is listed in the column header. All columns report coefficients on the treatment variable, δ_{hc}, which equals one for the 2009 cohort at switching schools (columns (A) and (B) in Table 1). All regressions include high school dummies and cohort dummies. Regressions are at the school-cohort level with observations weighted by the number of exam takers. Parentheses contain standard errors clustered at the region level. Dependent variable means are calculated from the 2001–2008 cohorts.

Panel A estimates equation (1). Panel B adds dummies for cells defined by cohort and groups of schools. There are 30 groups based on school ownership (public or private), training (academic/normal, academic & technical, or technical), and Icfes ranking (far superior, superior, high, middle, or low/inferior/far inferior). Each characteristic is time-invariant; a school is private if it is ever private, and I assign each school its highest training level and Icfes ranking. Panel C adds school-specific linear cohort trends to Panel B.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

data do not extend beyond three years for the 2009 cohort, the lack of catch-up enrollment suggests that the decline in college attendance is not merely transitory.

4.3. **Robustness specifications.** The benchmark results compare switching schools to all high schools in other regions. Table 1 shows that public switching schools are mostly similar to other public schools (columns (A) and (D)), but private schools in the two areas differ significantly (columns (B) and (E)). For example, private school students in other regions have higher test scores and socioeconomic backgrounds. These differences arise because the comparison group includes elite private schools in Bogotá, and because the best private schools in the affected regions stayed on the September calendar (column (C)).

The differences between switching and comparison schools raise concerns if they lead to divergent college enrollment trends. Table 2 depicts two robustness specifications that address this possibility. Panel B alters the implicit comparison group by matching high schools using their characteristics. I define 30 high school groups g based on a school's public/private status, its level of academic training, and its Icfes performance ranking.[16] I then add group-cohort dummies, γ_{gc}, to equation (1). This yields a matching estimator (e.g., Abadie and Imbens, 2002; Imbens, 2004) where matches are based on school traits. The coefficients in Panel B are thus identified only from variation in enrollment outcomes within similar types of high schools. The results are essentially unchanged from Panel A, suggesting that the choice of comparison group does not drive the main results.

Panel C adds linear cohort trends for each high school, $\tilde{\gamma}_h \times c$, to the specification for Panel B. This is the standard differences in differences test of adding linear terms in the "time" dimension (Angrist and Pischke, 2009). The coefficient magnitudes fall slightly relative to previous specifications, which reflects a small divergence in enrollment trends between switching schools and other schools (see Figure 4). Nonetheless, the sharp decrease in enrollment for 2009 switching school graduates is distinguishable from the linear trends, and the pattern of coefficients across years matches that in Panels A and B.

Standard errors in Table 2 are clustered at the region level. Colombia has 33 administrative regions, which is below the rule-of-thumb for inference using the typical cluster-robust standard errors (e.g., Angrist and Pischke, 2009). To address this, Appendix A.5 describes additional regressions that are at the region-cohort level and use a $t(33-2)$ distribution for inference. This follows Donald and Lang (2007), who recommend "between-group" estimators in settings with few clusters but a large number of observations per cluster (see also Wooldridge, 2003). These region-cohort level regressions yield larger standard errors but do not measurably alter the point estimates or patterns of statistical significance from Table 2.

4.4. **Potential explanations for the enrollment decline.** Table 2 shows that many 2009 switching school graduates experienced an unusual one semester break after high school and did not subsequently enroll in college. One potential concern in attributing the enrollment decline to the time gap is that these students may have been affected by other elements of the calendar transition. For example, the 2009 school year had fewer instructional days prior to the Icfes exam, and more days afterward (see Figure 3). This may have harmed students' exam performance or reduced their likelihood of graduating.

[16] See the notes to Table 2 for details on the group definitions. Results are similar with alternative definitions.

Appendix A.6 tests these hypotheses. The calendar shift did not have a significant impact on the number students who took the Icfes exam at switching schools. There is also no change in students' average Icfes scores. Thus there is little evidence that the transition reduced students' college preparedness.

Another potential hypothesis is that the transition decrease admission rates at selective colleges. This could arise if colleges did not rebalance their January/September cohort slots in response to the shifting calendar. However, Appendix A.7 shows that the enrollment decline occurs mainly at non-selective colleges; thus the primary effect of the transition was a decline in enrollment at colleges with open admissions, not a shift in access to selective colleges.[17]

It is also hard to explain the enrollment decline by a decrease in graduates' returns to college. The time gap reduced post-college careers by one-half year at at most, and the corresponding loss in discounted returns is small relative to estimates of the college earnings premium (Autor, 2014; Zimmerman, 2014).

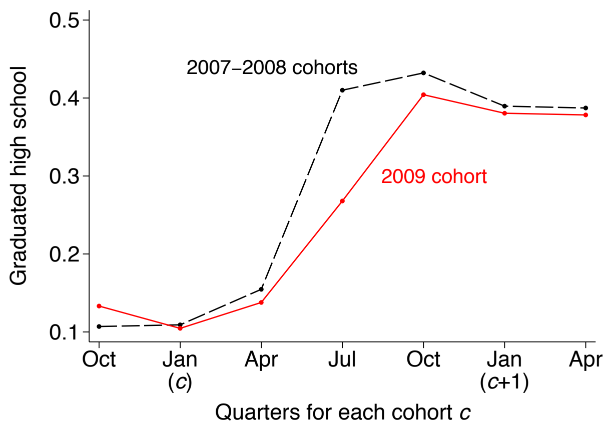
The time gap more likely influenced college preferences through students' experiences while away from school. Figure 5 provides suggestive evidence of this channel using 2007–2010 Colombian household survey data. The graphs depict two cohorts of individuals in the affected regions: those who turned 17 years old in the pre-transition years (2007–2008), and those who turned 17 in the first year of the calendar transition (2009).

Panel A plots the fraction of individuals in each cohort group with a high school degree. The x-axis displays quarters beginning in October–December before the cohort year; this is the start of the final year of high school for 17 year olds with on-time academic progress.[18] The graduation rate for the 2007–2008 cohorts jumps from 10 to 40 percent in the July quarter but does not reach 40 percent until October for the 2009 cohort. This reflects the delayed calendar for 2009 graduates in many switching schools (see Figure 3).

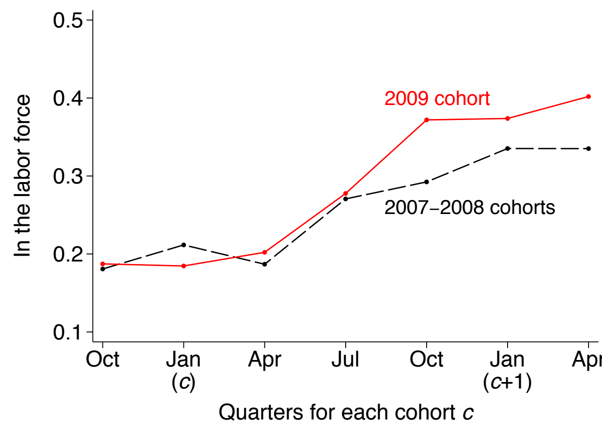
Panel B depicts labor force participation rates for the same cohorts and time periods. Prior to on-time high school graduation, roughly 20 percent of 17 year olds were either employed or looking for work. These rates increase for all cohorts in July–September as students finish high school. In October–December, however, labor force participation is higher in the 2009 cohort than in the 2007–2008 cohorts. October–December is the gap between high school graduation and potential college enrollment for some 2009 graduates. This result suggests that the time gap led some 2009 graduates to work or look for a job. These experiences may

[17] Further, Section 6 shows that the enrollment decline is largest for low SES and low ability students, who are less likely to attend selective colleges.

[18] Age cohort is a noisy measure of graduation cohort because a few students graduate early and many others are behind schedule. This explains why 10 percent of 17 year olds have a high school degree before July and only 40 percent have one after. Figure 5 also cannot separate switching and non-switching schools in the affected regions, but switching school graduates are a large majority of all graduates (see Table 1).



Panel A. High school graduation



Panel B. Labor force participation

Figure 5. 17 year olds in the affected regions

Notes: Data are from the 2007–2010 monthly urban (*cabecera*) and rural (*resto*) GEIH household surveys. The sample is 1990–1992 birth cohorts in the affected regions; this defines cohorts of individuals who turned 17 in 2007–2009. For February 2009 the sample is current 16 year olds because birthdates are missing. The sample includes only children, grandchildren, or other relatives of the household head.

The x-axis combines monthly surveys into three-month quarters. It begins in October before the cohort year and ends in June after the cohort year (e.g., 10/2008 through 6/2010 for the 2009 cohort). 2006 surveys are not available, so Oct–Dec 2006 values for the 2007 cohort are equal to Oct–Dec 2007 values for the 2008 cohort plus the average difference between the 2007 and 2008 cohorts in Jan–Jun of the cohort year.

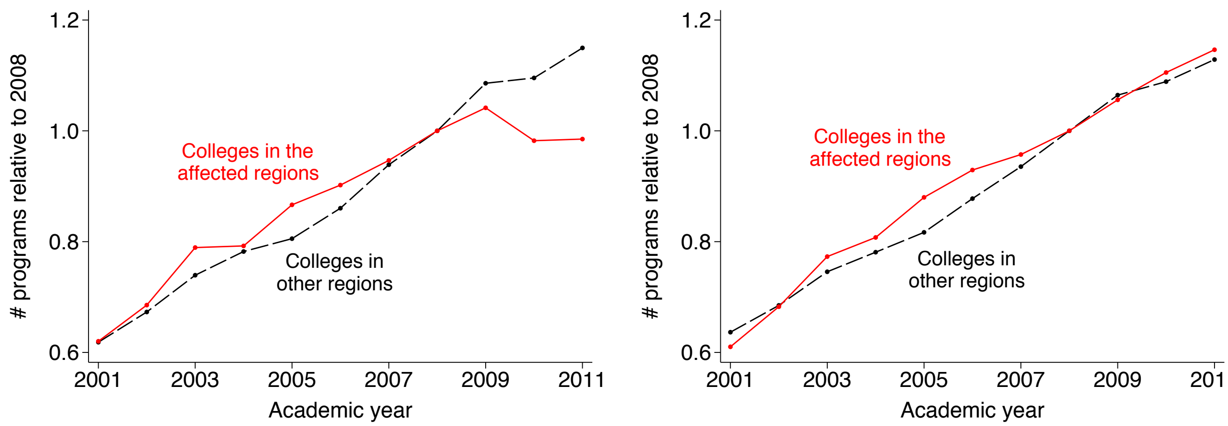
Panel A shows the fraction of each cohort group with a high school degree or above. Panel B shows the fraction in the labor force, defined as appearing in the employed (*ocupados*) or unemployed (*desocupados*) survey. Panel B omits months with mid-term breaks (1/2007, 1/2008, and 4/2009; see Figure 3) when labor force participation temporarily increases, and it is lagged one month because survey questions refer to labor force activity in the prior 1–4 weeks (e.g., October values are from the November survey).

Calculations use weights that are fixed over time, which are the sum of survey weights across all months within non-missing cells defined by region, age, gender, urban/rural survey, and secondary educated mother.

have prompted some students to forgo further schooling, as suggested by the results in Table 2. Consistent with this, Panel B in Figure 5 shows that labor force participation remains higher for the 2009 cohort through the start of the next year.[19]

Though the calendar transition had multiple components, Figure 5 suggests the decline in college attendance among 2009 graduates is most consistent with a causal effect of the post-graduation gap. The next section corroborates this finding by analyzing high schools in the affected regions that did not change calendars, which yield an arguably cleaner test of the effect of time gaps on college enrollment.

[19] Appendix A.8 shows that the findings in Figure 5 are statistically significant in regressions that also include 17 year olds in other regions as a comparison group.



Panel A. September cohort

Panel B. Both cohorts

FIGURE 6. Growth in college programs relative to 2008

Notes: This graph shows the number of programs with at least one enrollee in each academic year of the Ministry of Education records. Programs are defined by an institution and a non-missing program name. Academic year includes September of the listed year and January of the following year. For both affected regions and other regions, the y-axis is the number of programs in a given year divided by the number of programs in 2008. Panel A shows the number of September programs. Panel B shows the number of September programs plus the number of January programs.

5. Staying schools

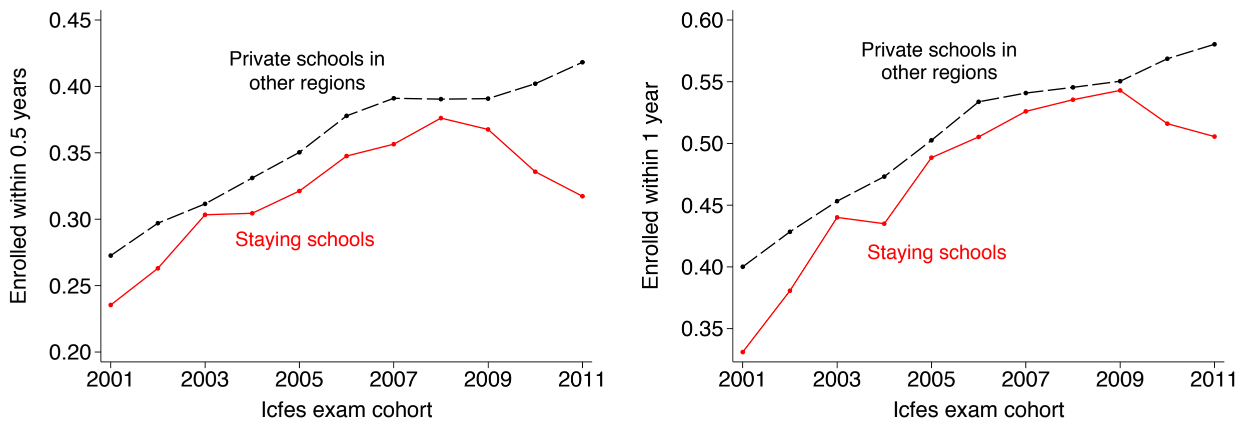
This section shows that the academic calendar shift also affected college entry timing for graduates from schools in the affected regions that did not change calendars, which I call *staying schools*. It then asks how the resulting time gap affected college enrollment.

5.1. **A time gap at staying schools.** Colombian students apply to both a college and a major. Many programs are offered twice per year, while others begin annually. In the affected regions, colleges historically offered their annual programs in the September cohort because nearly all local high school students graduated in June. With the calendar shift, the large majority of students now finish in November. As a result, colleges in the affected regions shifted some annual programs to the January cohort.

Figure 6 illustrates this shift in college program timing. Panel A plots the number of programs in the September cohort at colleges in the affected regions and in other regions. The y-axis is normalized to represent the change in programs relative to 2008. In both areas the quantity of September programs increased by more than two-thirds from 2001–2009.[20] In 2010–2011, however, the number of September programs declined by about six percent in the affected regions, falling behind program growth in other regions.

Panel B shows the total number of programs offered across the September and January cohorts. The two areas exhibit similar annual program growth in 2010 and 2011. Figure 6 thus shows that colleges in the affected regions shifted some programs from the September

[20] Program growth reflects new majors at existing colleges and the opening of new colleges.



Panel A. Enrolled within 0.5 years

Panel B. Enrolled within 1 year

FIGURE 7. Staying schools — College enrollment by years since Icfes

Notes: Staying schools include students from column (C) in Table 1. Private schools in other regions include students from column (E). College enrollment is defined as in Table 1.

to the January cohort without altering the total annual quantity. This affected the college enrollment timing of the 2010–2011 cohorts at staying schools, which finished in June. Graduates who wanted to apply to programs that moved to the January cohort had to wait an extra semester to do so. This created another time gap after graduation, induced in this case by changes in college calendars rather than high school calendars.

5.2. **Effects on college enrollment.** I explore the effects of the college program shift by comparing graduates from staying schools (column (C) in Table 1) to graduates from private schools in other regions (column (E)).

Figure 7 shows the college enrollment rates for these two school groups by Icfes exam cohort. Panel A displays the fraction of graduates who began college one semester after taking the Icfes. These initial college entry rates for staying schools and other schools follow similar upward trends through 2009. In 2010–2011, the immediate enrollment rate drops sharply at staying schools. This matches the timing of the college program shift in Figure 6.

Panel A of Figure 7 is a "first stage" result; some graduates could not immediately enter programs that moved to the January cohort. Panel B shows college enrollment rates within one year of the exam, which includes entry into both the September and January cohorts. Since Figure 6 shows no change in annual program offerings, staying school graduates who were waiting for January programs could have enrolled within one year. Panel B suggests that many did not. The enrollment decline in the 2010–2011 cohorts is slightly smaller in magnitude but still evident. This suggests the time gap after high school altered graduates' college attendance decisions.

Table 3. Staying schools — Time gap and college enrollment
Dependent variable: Enrolled in college within t years of the Icfes

	(A)	(B)	(C)
Panel A. Benchmark specification			
	0.5 years	1 year	1.5 years
Staying schools, 2010–2011 cohorts (δ_{hc})	-0.070^{***}	-0.050^{***}	-0.039^{***}
	(0.007)	(0.009)	(0.010)
R^2	0.831	0.857	0.866
Panel B. High school matching			
	0.5 years	1 year	1.5 years
Staying schools, 2010–2011 cohorts (δ_{hc})	-0.067^{***}	-0.048^{***}	-0.037^{***}
	(0.005)	(0.007)	(0.008)
R^2	0.839	0.866	0.875
Panel C. Linear high school trends			
	0.5 years	1 year	1.5 years
Staying schools, 2010–2011 cohorts (δ_{hc})	-0.072^{***}	-0.080^{***}	-0.077^{***}
	(0.007)	(0.006)	(0.007)
R^2	0.874	0.899	0.909
N	13,695	13,695	13,695
# regions	25	25	25
Dependent var. mean	0.338	0.482	0.550

Notes: The sample is the 2001–2011 Icfes exam cohorts at high schools in columns (C) and (E) of Table 1. The dependent variables are college enrollment within t years of the Icfes, where t is listed in the column header. All columns report coefficients on the treatment variable, δ_{hc}, which equals one for the 2010–2011 cohorts at staying schools (column (C) in Table 1). All regressions include high school dummies and cohort dummies. Regressions are at the school-cohort level with observations weighted by the number of exam takers. Parentheses contain standard errors clustered at the region level. Dependent variable means are calculated from the 2001–2008 cohorts.

Panel A estimates equation (1). Panel B adds dummies for cells defined by cohort and groups of schools (see Table 2 for details). Panel C adds school-specific linear cohort trends to Panel B.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3 presents the regression version of Figure 7. This table is analogous to Table 2, and the regressions are also based on equation (1). In Table 3, however, the sample is 2001–2011 graduates from staying schools and private high schools in other regions, and the treatment indicator, δ_{hc}, equals one for the 2010–2011 cohorts at staying schools. Since the college enrollment records extend through 2012, this table shows cumulative enrollment at a maximum duration of $t = 1.5$ years after the Icfes.

Panel A shows the benchmark results from equation (1). Columns (A) and (B) match Figure 7. There is a seven percentage point decline in immediate college enrollment for

2010–2011 staying school graduates relative to graduates from other regions. The enrollment decline within one year is slightly smaller at five percentage points.

Column (C) depicts cumulative enrollment within 1.5 years. The magnitude of the coefficient falls slightly relative to column (B), but there is still an enrollment shortfall of four percentage points in staying schools. Data constraints do not allow me to see if this catch-up enrollment continued, but Table 3 shows that many staying school graduates were not enrolled in college one full year after the time gap.[21]

Panels B and C present the same two robustness specifications as in Table 2. Panel B matches treated and comparison high schools by adding dummies for cells defined by cohort and school traits (see Section 4.3). These dummies lower the standard errors but have little effect on the point estimates. Panel C adds high-school-specific linear cohort trends to the specification for Panel B. These terms have little effect on the immediate enrollment coefficient, but they substantially increase the magnitudes of the estimates in columns (B) and (C). This reflects non-parallel enrollment trends in early Icfes cohorts (2001–2005) at staying schools and comparison schools (see Panel B of Figure 7). Nonetheless, these trends converge in the cohorts prior to the college program shift (2005–2009), and the estimates in Panel C of Table 3 still suggest a large negative impact of the time gap on enrollment.[22]

The results in Table 3 are not attributable to changes in instructional time because staying schools did not alter their academic calendars. This is corroborated by balance tests in Appendix A.6 that show no differential changes in the number of Icfes exam takers or their average scores. Further, the results are not likely due to decreases in college admission rates as there were no alterations in annual program offerings. Instead, these results are evidence of a causal effect of post-high school gaps on college enrollment.

6. Heterogeneous responses to the time gap

Sections 4 and 5 showed that switching school and staying school graduates experienced a time gap after high school, but for different reasons. Both mechanisms, however, led some graduates to forgo further schooling. The net effect was a seven percent decline in college enrollment rates in affected regions.

This section asks whether time gaps can also affect the distribution of schooling outcomes by exploring heterogeneity in these effects at both school types. I first examine whether responses to the gap vary by student characteristics. I then ask if students were more likely to forgo low-paying majors.

[21] Unlike Table 3, Table 2 shows little evidence of catch-up enrollment at switching schools. This is consistent with evidence in Section 6 that low SES/ability students are more likely to forgo college after the gap.
[22] Appendix A.5 shows that the results in Table 3 are also robust to the stricter inference procedures based on Donald and Lang (2007) (see Section 4.3).

6.1. **Heterogeneity by student characteristics.** To study how time gap responses vary across individuals, I estimate coefficients like those in Tables 2 and 3 for different groups of students. However, student types may vary in both baseline college enrollment rates and in their exposure to the time gap. For example, high ability graduates are more likely to enroll in college, and they may also have been more likely to choose college programs that shifted from the September to the January cohorts. This could lead to larger estimates even absent any heterogeneity in responsiveness to the time gap.

To address this, I estimate two stage least squares (2SLS) regressions that relate the long-term enrollment effect to the immediate effect of the transition. Consider the regressions:

$$y_{hc}^{0.5} = \gamma_h + \gamma_c + \beta^{0.5}\delta_{hc} + \epsilon_{hc}, \tag{2}$$

$$y_{hc}^{max} = \alpha_h + \alpha_c + \theta\hat{y}_{hc}^{0.5} + \upsilon_{hc}. \tag{3}$$

Equation (2) is a first stage regression identical to equation (1). The dependent variable, $y_{hc}^{0.5}$, is the fraction of students from high school h and Icfes cohort c who begin college 0.5 years after the exam. The regression includes high school dummies, γ_h, cohort dummies, γ_c, and a treatment indicator for schools and cohorts affected by the calendar shift, δ_{hc}.

Equation (3) is the second stage regression. The dependent variable, y_{hc}^{max}, is cumulative college enrollment at the maximum observable duration since the Icfes exam. The regression also includes school and cohort dummies, but the main independent variable is enrollment within 0.5 years of the exam, $y_{hc}^{0.5}$. The coefficient of interest, θ, measures the effect of a one percentage point change in the immediate college enrollment rate on longer-term enrollment.

College enrollment is endogenous, so the second stage (3) uses predicted values from the first stage (2), $\hat{y}_{hc}^{0.5}$. In other words, the treatment variable, δ_{hc}, is an instrument for $y_{hc}^{0.5}$. The 2SLS coefficient, θ, equals the ratio of the longer-term enrollment effect to the immediate effect ($\beta^{max}/\beta^{0.5}$). Thus θ measures responsiveness to the time gap—the reduction in longer-term enrollment that is attributable to the decline in initial enrollment from the transition.

Column (A) in Table 4 shows the estimate of θ from equation (3). Panel A displays the result for schools that switched to the January calendar. The sample is the same as in Table 2, and the dependent variable is cumulative college enrollment within three years of the Icfes. The coefficient shows that a one percentage point decline in initial enrollment during the transition led to a 0.55 percentage point decrease in the three-year enrollment rate.[23]

Columns (B)–(D) use this 2SLS procedure to compare time gap responsiveness across student types. The columns divide the sample into two groups defined by socioeconomic status, academic ability, and age, respectively.[24] These regressions are similar to column (A), but all terms are interacted with a dummy for one of the groups. In Table 4 this

[23] This equals the ratio of the estimates from columns (D) and (A) in Panel A of Table 2.

[24] I do not find heterogeneous responses by gender that are consistent across specifications.

Table 4. Heterogeneity in responsiveness to the time gap

Panel A. Switching schools
Dependent variable: Enrolled in college within 3 years of the Icfes

	(A)	(B)	(C)	(D)
		Definition of Disadvantaged group		
	All students	Mother no college	Icfes below median	Age 18 or older
Time gap ($\hat{y}^{0.5}_{hcd}$)	0.553***	0.295*	0.385**	0.524***
	(0.165)	(0.174)	(0.177)	(0.183)
Time gap ($\hat{y}^{0.5}_{hcd}$) × Disadvantaged$_d$		0.344***	0.464***	0.250**
		(0.114)	(0.131)	(0.100)
N	36,972	14,672	73,074	73,178
R^2	0.324	0.253	0.261	0.289
# regions	33	33	33	33

Panel B. Staying schools
Dependent variable: Enrolled in college within 1.5 years of the Icfes

	(A)	(B)	(C)	(D)
		Definition of Disadvantaged group		
	All students	Mother no college	Icfes below median	Age 18 or older
Time gap ($\hat{y}^{0.5}_{hcd}$)	0.564***	0.632***	0.524***	0.449***
	(0.093)	(0.059)	(0.071)	(0.123)
Time gap ($\hat{y}^{0.5}_{hcd}$) × Disadvantaged$_d$		0.282***	0.054	0.235***
		(0.102)	(0.176)	(0.079)
N	13,695	9,583	26,660	26,717
R^2	0.451	0.332	0.397	0.398
# regions	25	25	25	25

Notes: In Panel A, the sample is as in Table 2. The dependent variable is college enrollment within 3 years of the Icfes. The treatment variable, δ_{hc}, equals one for the 2009 cohort at switching schools.

In Panel B, the sample is as in Table 3. The dependent variable is college enrollment within 1.5 years of the Icfes. δ_{hc} equals one for the 2010–2011 cohorts at staying schools.

Column (A) is the 2SLS regression (3), where enrollment within 0.5 years of the Icfes, $y^{0.5}_{hc}$, is instrumented by δ_{hc}. I report only the coefficient on $\hat{y}^{0.5}_{hc}$. The regression includes high school dummies and cohort dummies. The regression is at the school-cohort level with observations weighted by the number of exam takers.

Columns (B)–(D) are the same 2SLS regression, but all terms are interacted with a dummy for the "Disadvantaged" group in the column header. The reported variables, $\hat{y}^{0.5}_{hcd}$ and $\hat{y}^{0.5}_{hcd}$ × Disadvantaged$_d$, are instrumented by δ_{hc} and δ_{hc} × Disadvantaged$_d$. The regression is at the school-cohort-Disadvantaged level with observations weighted by the number of exam takers.

The variables that define Disadvantaged are calculated as in Table 1. Column (B) excludes the 2001–2007 cohorts because mother's education is not available. Median Icfes is defined in the sample for each regression.

Parentheses contain standard errors clustered at the region level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

dummy is called "Disadvantaged$_d$" and equals one for: low SES students (column (B)), low ability students (column (C)), and students older than the on-time age (column (D)).[25]

[25] Regressions in columns (B)–(D) are at the school-cohort-Disadvantaged level (hcd); they calculate enrollment rates separately for Advantaged and Disadvantaged students in the same high school and cohort. The

In column (B) of Panel A, the coefficient on $\hat{y}^{0.5}_{hcd}$ is the effect of time gap for switching school graduates whose mothers attended college. The coefficient on the $\hat{y}^{0.5}_{hcd} \times \text{Disadvantaged}_d$ interaction is the difference in the time gap effects for students without and with college educated mothers. The effect for graduates without college educated mothers (0.295 + 0.344) is more than double that for high SES students (0.295). In other words, low SES graduates who experienced a time gap after high school were more likely to forgo college.[26]

Columns (C) and (D) in Panel A show similar patterns for academic ability and age. Column (C) splits the sample by Icfes scores, and the time gap effect is more than twice as large for graduates with below-median scores. Column (D) compares students with on-time and delayed academic progress; the time gap effect is about 50 percent larger for students who are 18 years old or older.

Panel B of Table 4 shows analogous results for schools in the affected regions that did not change calendars. The sample is the same as in Table 3, and the dependent variable is college enrollment within 1.5 years of the Icfes. Column (A) shows that the mean time gap effect at staying schools is nearly identical to that at switching schools.[27] Columns (B)–(D) show heterogeneity using the same characteristics as in Panel A. The results mirror those for switching schools. "Disadvantaged" staying school graduates are more responsive to the time gap, although the heterogeneity by ability is statistically insignificant.

Table 4 thus shows that low SES, low ability, and older students are more likely to forgo college after the time gap. Another characteristic these groups share is that each has a low overall college enrollment rate. Figure 8 illustrates this using 16 student types defined by mothers' college attendance, age above/below 18, and quartiles of Icfes scores. The y-axis displays time gap coefficients, θ, from separate estimations of equation (3) for each of the 16 types, with larger symbols indicating more precise coefficients. The x-axis depicts the three-year college enrollment rates of each group in the pre-transition cohorts (2001–2008). Panel A displays coefficients from switching schools; Panel B depicts staying school estimates.

The scatterplots in both panels of Figure 8 have a downward shape; student types with low baseline enrollment rates are also more responsive to the time gap.[28] The group definitions in Figure 8 are analogous to those in Appendix A.1, which shows an inverse relationship between overall enrollment rates and the likelihood of delayed enrollment. Taken together,

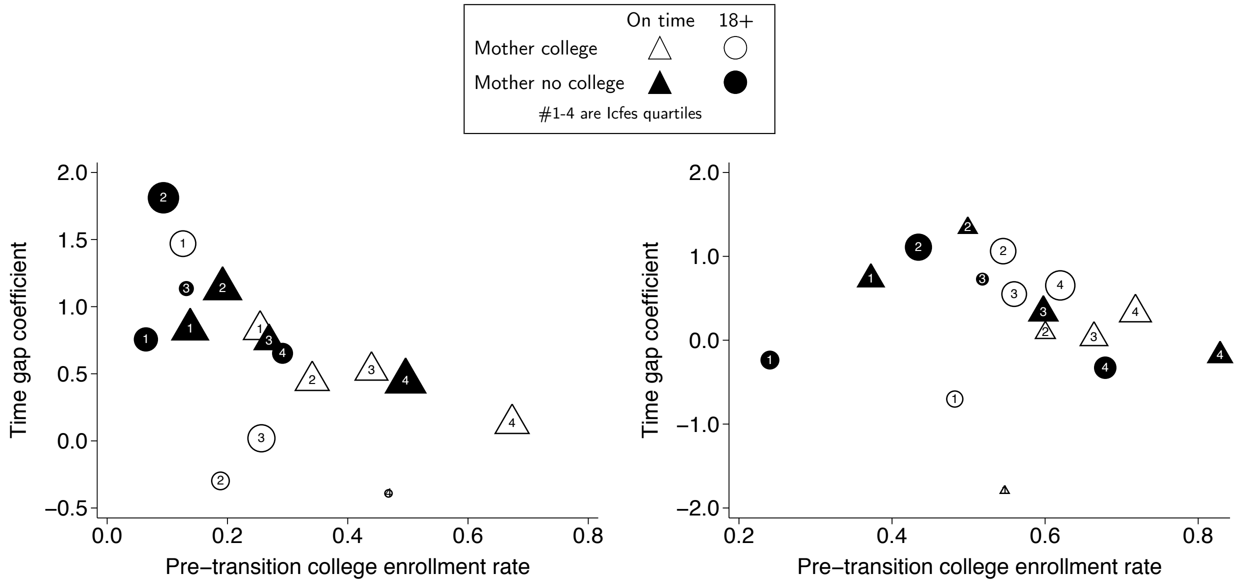
full second-stage specification, of which Table 4 displays only the θ and θ^D coefficients, is

$$y^{max}_{hcd} = \alpha_{hd} + \alpha_{cd} + \theta \hat{y}^{0.5}_{hcd} + \theta^D \left(\hat{y}^{0.5}_{hcd} \times \text{Disadvantaged}_d\right) + \upsilon_{hcd}.$$

[26] Column (B) includes only the 2008–2011 cohorts for which mother's education is available. Results are similar if I define SES by mean mother's education at the high school level and include all 2001–2011 cohorts.

[27] This coefficient also equals the ratio of the estimates from columns (C) and (A) in Panel A of Table 3.

[28] Some time gap coefficients in Figure 8 are greater than one, suggesting that the gap also reduced the long-term enrollment of students who would not have initially enrolled anyway.



Panel A. Switching schools

Panel B. Staying schools

FIGURE 8. Time gap responsiveness and enrollment rates by SES, ability, and age

Notes: The samples for Panels A and B are the same as in Panels A and B of Table 4, respectively. This figure defines 16 student groups based on the same definitions of SES (two types), ability (four types), and age (two types) as in Table 4, with two exceptions. First, it uses Icfes quartiles rather than the median. Second, mother's education is only available in the 2008–2011 cohorts, so this figure defines SES at the high school level to include all cohorts. Mother no college indicates high schools above the sample median fraction of 2008 graduates without college educated mothers; mother college are high schools below the median.

For each of the 16 student groups, the pre-transition college enrollment rate is the fraction of 2001–2008 affected region graduates who enter college within three years of the Icfes. The time gap coefficient is the estimate of θ from separate regressions (3) for each student group. Regressions are as in column (A) of Table 4. Symbol size is proportional to $\log(1 + 1/\hat{\sigma}_g^2)$, where $\hat{\sigma}_g^2$ is the standard error on $\hat{\theta}_g$ for each group g.

these results suggest that time gaps may have a doubly negative effect on college enrollment; not only are disadvantaged students more likely to take time off after high school, they are also more likely to forgo college after these time gaps.

A recent literature explores the mechanisms underlying socioeconomic gaps in college attendance (Bailey and Dynarski, 2011; Hoxby and Avery, 2013; Black et al., 2015; Clotfelter et al., 2015). Heterogeneity in transition timing may be another factor that contributes to these disparities. The results in Table 4 and Figure 8 are also consistent with a greater influence of time gaps on individuals who are *ex ante* indecisive about college.

6.2. **Heterogeneity by average program earnings.** The previous section shows that time gap effects vary across types of students. This section asks if the time gap also influences students' program choices.

Specifically, I explore whether graduates' post-gap program choices were correlated with their expected earnings in those majors. I calculate program earnings using administrative social security records for 2003–2008 college graduates. Programs are defined by the Ministry of Education's 54 college major groups. For each program, I calculate the average earnings of college graduates with a given Icfes score and gender.[29] This yields a measure of a potential enrollee's predicted earnings in each program based on college graduates with similar characteristics. For each gender and Icfes score combination, I then calculate the median predicted earnings across all 54 programs. I define "high wage" programs as those above the median predicted earnings; "low wage" programs are those below the median.

Table 5 shows regressions similar to those in Tables 2 and 3, but the dependent variables are indicators for enrolling in high or low wage programs. In Panel A, columns (A) and (B) show the effect of the switching school time gap on enrollment in each program type. The time gap led to a similar enrollment decline one semester after the Icfes for both high earnings and low earnings programs. Columns (C) and (D) show the effect on enrollment within three years of the exam. For low wage programs, the three-year enrollment decline is similar to the initial effect. The enrollment effect mostly disappears at high wage programs, and the three-year coefficient is not statistically significant. After the time gap, students were thus more likely to enter programs that yield high earnings, and more likely to forgo programs that deliver low earnings.

Panel B shows that this pattern also holds for graduates from staying schools. The coefficients in columns (A) and (B) are similar, suggesting that the initial decline in enrollment is relatively earnings-neutral. Columns (C) and (D) show that the longer-term reduction in enrollment is driven by students forgoing programs with low expected earnings.

Table 5 thus provides suggestive evidence that labor market returns affect students' decisions to return to school after the time gap. This analysis does not study the causal returns to program choice as in Hastings et al. (2015) but rather the descriptive returns, which may be more salient. Related research finds that students' schooling experiences and preferences affect their major choices (Zafar, 2011; Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2014; Wiswall and Zafar, 2015). My results suggest that time outside of the education system can also shape schooling decisions.

7. Conclusion

An influential literature in labor economics argues that college access lowers wage inequality (Katz and Murphy, 1992; Goldin and Katz, 2008). Recent work shows that one potential barrier to obtaining a college degree is a lack of information. Thus there has been a push by both academics and policymakers to make college costs and outcomes more transparent.

[29] Specifically, for each of the 54 programs, I regress log average daily earnings in 2008–2012 on a gender dummy, Icfes percentile, and a quadratic in years since college graduation. Predicted earnings are the sum of the gender and Icfes terms, each multiplied by their coefficients from this regression.

TABLE 5. Heterogeneity by program earnings
Dependent variable: Enrolled in college within t years of the Icfes

Panel A. Switching schools

	(A)	(B)	(C)	(D)
	Enrolled within 0.5 years		Enrolled within 3 years	
	Low wage programs	High wage programs	Low wage programs	High wage programs
Switching schools, 2009 cohort (δ_{hc})	-0.027^{***}	-0.022^{***}	-0.023^{***}	-0.004
	(0.007)	(0.002)	(0.004)	(0.006)
N	36,972	36,972	36,972	36,972
R^2	0.758	0.845	0.794	0.860
# regions	33	33	33	33
Dependent var. mean	0.076	0.086	0.183	0.199

Panel B. Staying schools

	(A)	(B)	(C)	(D)
	Enrolled within 0.5 years		Enrolled within 1.5 years	
	Low wage programs	High wage programs	Low wage programs	High wage programs
Staying schools, 2010–2011 cohorts (δ_{hc})	-0.037^{***}	-0.033^{***}	-0.030^{***}	-0.009
	(0.006)	(0.004)	(0.011)	(0.006)
N	13,695	13,695	13,695	13,695
R^2	0.636	0.768	0.665	0.787
# regions	25	25	25	25
Dependent var. mean	0.146	0.192	0.244	0.307

Notes: The samples for Panels A and B are the same as in Tables 2 and 3, respectively. All columns report coefficients on the treatment variable, δ_{hc}, which is defined as in those tables. Regressions are identical to Tables 2 and 3 except for the dependent variables.

The dependent variables measure the fraction of students from each high school and cohort enrolling in high or low wage programs within t years of the Icfes, as listed in the column header. High/low wage programs are defined using social security records for 2003–2008 college graduates, with programs defined by the Ministry of Education's 54 college major groups. For each program (plus one for missing values), I regress log average daily earnings in 2008–2012 on a gender dummy, Icfes percentile, and a quadratic in years since college graduation. I define predicted earnings as the sum of the gender and Icfes terms, each multiplied by their coefficients from this regression. I then calculate the median predicted earnings across all 54 programs for each gender and Icfes score combination in the samples for Panels A and B. "High wage" programs are those above the median predicted earnings; "low wage" programs are those below the median.

Parentheses contain standard errors clustered at the region level. Dependent variable means are calculated from the 2001–2008 cohorts.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This research and policy agenda focuses on information that is at least potentially attainable prior to the enrollment decision through *search* (Stigler, 1961). For example, Jensen (2010) and Hoxby and Turner (2013) provide students with information on search traits like average returns and tuition, finding substantial effects on their educational choices.

This paper has explored a type of information that is less amenable to information interventions: individuals' *experiences* (Nelson, 1970). In particular, it asked whether breaks after high school graduation affect the decision of whether to enroll in college. To isolate this channel, it exploited a policy that altered academic calendars in some regions of Colombia, which created a one semester gap before potential college entry for some students.

This brief time gap led to a persistent reduction in college enrollment rates in affected regions. The decline was largest for students who were less likely to attend college in the first place. There is also suggestive evidence that labor force participation increased during the time gap, and that individuals who subsequently enrolled chose higher paying majors. These results suggest that students' out-of-school experiences altered their college choices.

The idea that individuals' non-academic experiences matter for their schooling attainment has implications for the design of education systems. Critics of early tracking argue that it forces students into career paths based on limited information from high-pressure exams. Yet such systems may also ease transitions between schooling tiers, which can create educational churning in more flexible systems. Policies that clarify enrollment standards and simplify the application process can also limit academic gaps. The success of initiatives to increase college attainment rates, such as the recent proposal for free community college in the U.S., may depend in part on the degree to which they discourage enrollment delays, with potential implications for wage inequality.

References

Abadie, A. and G. Imbens (2002). Simple and bias-corrected matching estimators for average treatment effects. National Bureau of Economic Research.

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics 4*, 1043–1171.

Angrist, J., D. Lang, and P. Oreopoulos (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics 1*(1), 136–163.

Angrist, J. and V. Lavy (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *The American Economic Review 99*(4), 1384–1414.

Angrist, J. D. and A. B. Krueger (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association 87*(418), 328–336.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*(1), 343–375.

Arcidiacono, P., V. J. Hotz, and S. Kang (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics 166*(1), 3–16.

Atkin, D. (2012). Endogenous skill acquisition and export manufacturing in Mexico. National Bureau of Economic Research.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the 'other 99 percent'. *Science 344*(6186), 843–851.

Bailey, M. J. and S. M. Dynarski (2011). Gains and gaps: Changing inequality in U.S. college entry and completion. National Bureau of Economic Research.

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*(4), 1437–1472.

Beffy, M., D. Fougere, and A. Maurel (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? *Review of Economics and Statistics 94*(1), 334–347.

Bettinger, E. P. and B. T. Long (2009). Addressing the needs of underprepared students in higher education does college remediation work? *Journal of Human Resources 44*(3), 736–771.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Black, D. A., T. G. McKinnish, and S. G. Sanders (2005). Tight labor markets and the demand for education: Evidence from the coal boom and bust. *Industrial and Labor Relations Review 59*(1), 3–16.

Black, S. E., K. E. Cortes, and J. A. Lincove (2015). Apply yourself: Racial and ethnic differences in college application. National Bureau of Economic Research.

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*(2), 455–467.

Castleman, B. L. and L. C. Page (2014). A trickle or a torrent? understanding the extent of summer "melt" among college-intending high school graduates. *Social Science Quarterly 95*(1), 202–220.

Charles, K. K., E. Hurst, and M. J. Notowidigdo (2015). Housing booms and busts, labor market opportunities, and college attendance. National Bureau of Economic Research.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review 104*(9), 2633–2679.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor (2015). Public universities, equal opportunity, and the legacy of Jim Crow: Evidence from North Carolina. National Bureau of Economic Research.

Dinkelman, T. and C. Martínez A. (2014). Investing in schooling in Chile: The role of information about financial aid for higher education. *Review of Economics and Statistics 96*(2), 244–257.

Dobkin, C. and F. Ferreira (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review 29*(1), 40–54.

Donald, S. G. and K. Lang (2007). Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics 89*(2), 221–233.

Dynarski, S. and J. Scott-Clayton (2013). Financial aid policy: Lessons from research. *The Future of Children 23*(1), 67–91.

Figlio, D. and S. Loeb (2011). School accountability. In E. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 3, pp. 383–421.

Fredriksson, P. and B. Öckert (2014). Life-cycle effects of age at school start. *The Economic Journal 124*(579), 977–1004.

Fryer, R. G. (2011). Financial incentives and student achievement: Evidence from randomized trials. *The Quarterly Journal of Economics 126*(4), 1755–1798.

Goldin, C. D. and L. F. Katz (2008). *The Race Between Education and Technology*. Harvard University Press.

Goodman, S. F. (2013). Learning from the test: Raising selective college enrollment by providing information. Board of Governors of the Federal Reserve System.

Hanushek, E. A. and M. E. Raymond (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management 24*(2), 297–327.

Hanushek, E. A. and L. Woessmann (2006). Does educational tracking affect performance and inequality? differences-in-differences evidence across countries. *The Economic Journal 116*(510), C63–C76.

Hastings, J., C. A. Neilson, and S. D. Zimmerman (2015). The effects of earnings disclosure on college enrollment decisions. National Bureau of Economic Research Working Paper No. 21300.

Hastings, J. S. and J. M. Weinstein (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics 123*(4), 1373–1414.

Hoxby, C. and S. Turner (2013). Expanding college opportunities for high-achieving, low income students. Stanford Institute for Economic Policy Research Discussion Paper.

Hoxby, C. M. (2003). School choice and school productivity. could school choice be a tide that lifts all boats? In C. M. Hoxby (Ed.), *The Economics of School Choice*, pp. 287–342.

University of Chicago Press.
Hoxby, C. M. and C. Avery (2013). Missing one-offs: The hidden supply of high-achieving, low-income students. Brookings Papers on Economic Activity.
Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics 86*(1), 4–29.
Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics 125*(2), 515–548.
Kane, T. J. and D. O. Staiger (2002). The promise and pitfalls of using imprecise school accountability measures. *The Journal of Economic Perspectives 16*(4), 91–114.
Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.
Martorell, P. and I. McFarlin Jr (2011). Help or hindrance? the effects of college remediation on academic and labor market outcomes. *The Review of Economics and Statistics 93*(2), 436–454.
Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy 78*(2), 311–29.
Nguyen, T. (2010). Information, role models and perceived returns to education: Experimental evidence from Madagascar. Unpublished manuscript.
Oreopoulos, P. and R. Dunn (2013). Information and college access: Evidence from a randomized field experiment. *The Scandinavian Journal of Economics 115*(1), 3–26.
Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics 33*(2), 493–520.
Perez-Arce, F. (2015). Is a dream deferred a dream denied? college enrollment and time-varying opportunity costs. *Journal of Labor Economics 33*(1), 33–61.
Rivkin, S. G., E. A. Hanushek, and J. F. Kain (2005). Teachers, schools, and academic achievement. *Econometrica 73*(2), 417–458.
Schütz, G., H. W. Ursprung, and L. Woessmann (2008). Education policy and equality of opportunity. *Kyklos 61*(2), 279–308.
Scott-Clayton, J. E. (2011). The shapeless river: Does a lack of structure inhibit students' progress at community colleges? Community College Research Center Working Paper.
Stigler, G. J. (1961). The economics of information. *Journal of Political Economy 69*(3), 213–225.
Stinebrickner, R. and T. R. Stinebrickner (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies 81*(1), 426–472.
Urquiola, M. (2016). Competition among schools: Traditional public and private schools. In E. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education.*
Venezia, A., M. Kirst, and A. Antonio (2003). Betraying the college dream: How disconnected K-12 and postsecondary education systems undermine student aspirations. Technical report, Palo Alto, CA: Stanford University.
Wiswall, M. and B. Zafar (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.
Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review 93*(2), 133–138.

Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics 29*(2), 301–348.
Zimmerman, S. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics 32*(4), 711–754.

A. Appendix — Not for publication

A.1. **Delayed college enrollment by SES, ability, and age.** Figure 1 in Section 2 shows that delayed enrollment is an important phenomenon in both the U.S. and Colombia. Figure A1 explores the relationship between overall and delayed college enrollment. To do this it divides the samples for Figure 1 into 16 student types based on age, socioeconomic status, and academic ability. For each type, the graph depicts both the overall college enrollment rate (x-axis) and the prevalence of delayed enrollment (y-axis), defined as the fraction of all enrollees who wait more than one semester.

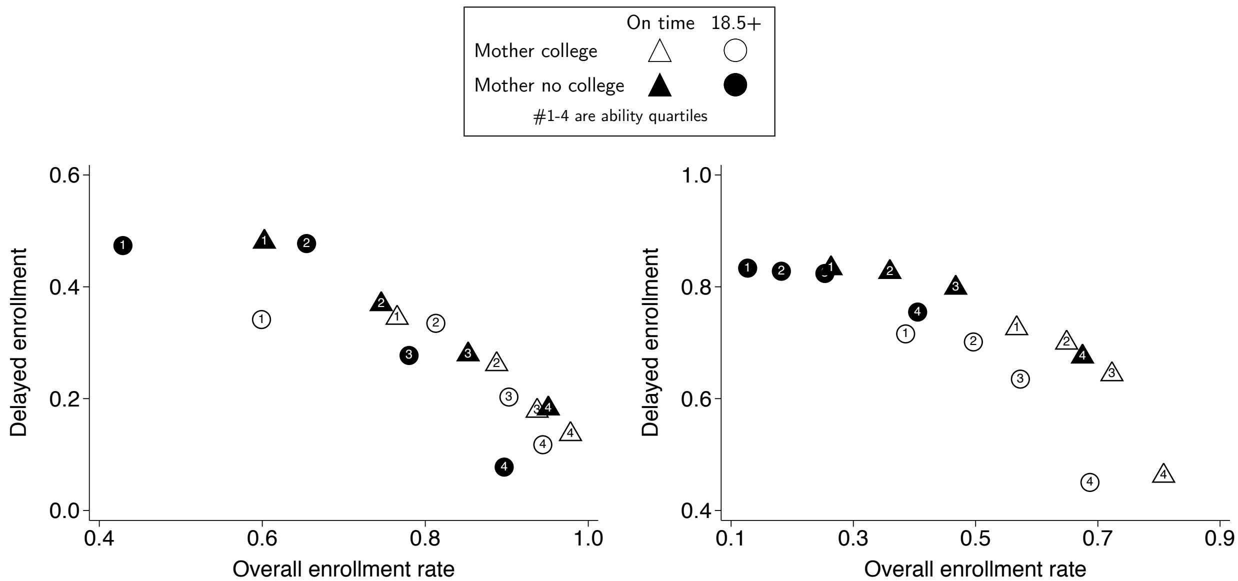
Panel A shows U.S. high school graduates from the NLSY. There is a strong negative relationship between overall and delayed enrollment; students who are less likely to attend college at all are more likely to delay conditional on enrolling. Both delaying and forgoing college are more common among students who are "disadvantaged" in their academic progression, as defined by age, socioeconomic status, and ability.

Panel B shows a similar relationship between delayed and overall enrollment for Colombian high school graduates.

A.2. **Data merging.** The main analysis of this paper uses two administrative datasets: 1) records of Icfes high school exam takers; and 2) records of students enrolled in colleges tracked by the Ministry of Education. I merge these datasets using national ID numbers, birth dates, and names. Nearly all students in both datasets have national ID numbers, but Colombians change ID numbers around age 17. Most students in the Icfes records have the below-17 ID number (*tarjeta*), while the majority of students in the college enrollment records have the above-17 ID number (*cédula*). Merging using ID numbers alone would therefore lose a large majority of students. Instead, I merge observations with either: 1) the same ID number and a fuzzy name match; 2) the same birth date and a fuzzy name match; or 3) an exact name match for a name that is unique in both records.

Merge quality is important because my main dependent variable—enrolling in college—is an indicator for a student's appearance in the enrollment dataset. 41 percent of the 2001–2011 Icfes exam takers appear in the enrollment records, which is broadly comparable to the higher education enrollment rate in Colombia during the same time period.[30] A better indicator of merge success is the percentage of college enrollees that appear in the Icfes

[30] The gross tertiary enrollment rate grew from 25 percent to 43 percent between 2001 and 2012 (World Bank World Development Indicators, available at http://data.worldbank.org/country/colombia in December 2015). This rate is not directly comparable to my merge rate because not all high school aged Colombians take the Icfes exam. Roughly 20 percent of the secondary school aged population is not enrolled in high school. This would cause my merge rate to be higher than the World Bank's tertiary enrollment rate.



Panel A. United States

Panel B. Colombia

FIGURE A1. Delayed college enrollment by SES, ability, and age

Notes: The samples for both panels are students from Figure 1 with non-missing values of the age, SES, and ability variables described below. College enrollment and years since graduation are defined as in Figure 1. Overall enrollment rate is the fraction of graduates who enter college within nine years of graduation. Delayed enrollment is the fraction of all enrollees who began more than six months after graduation.

In Panel A, age is calculated at the end of the month prior to graduation. Mother college means a graduate's mother has some college education. Ability is defined by quartiles of the Armed Services Vocational Aptitude Battery math/verbal percentile within the sample. All calculations use 2011 panel weights.

In Panel B, age is calculated at the end of August in the college entrance exam year. Mother college means a graduate's mother has any degree above basic secondary. Ability is defined by quartiles of the aggregate entrance exam percentile within the sample. See Table 1 for details on these variables.

records because all domestic college students must take the exam. Among enrollees who took the Icfes exam between 2001 and 2011, I match 91 percent.[31]

A.3. **Colleges in the Ministry of Education records.** This section describes the colleges that are included in the Ministry of Education records. For this I use another administrative dataset from a college exit exam called *Saber Pro* (formerly *ECAES*). This national exam is administered by the same agency that runs the Icfes entrance exam. The exit exam became a requirement for graduation from any higher education institution in 2009.

Column (A) in Table A1 the depicts the 310 colleges that have any exit exam takers in these administrative records in 2009–2011. These colleges are categorized into the Ministry of Education's five types of higher education institutions, which are listed in descending

[31] The enrollment records contain age at time of Icfes for some students, which allows me to calculate the year they took the Icfes exam. Approximately 16 percent of students in the enrollment dataset have missing birth dates, which accounts for the majority of observations I cannot merge. Some duplicate matches arise because students took the Icfes exam more than once, though I erroneously match a small number of students with the same birth date and similar names.

Table A1. Higher education institutions in Ministry of Education records

	(A)	(B)	(C)
Institution category	# colleges	# exit exam takers/year	% colleges in records
University	122	134,496	1.00
University Institute	103	53,338	0.88
Technology School	3	2,041	1.00
Technology Institute	47	15,092	0.82
Technical/Professional Institute	35	11,408	0.99
Total	310	216,375	0.96

Notes: Column (A) depicts the number of colleges that have Saber Pro exit exam takers in 2009–2011 using administrative records from the testing agency. Colleges are categorized into the Ministry of Education's five higher education institution types. Column (B) shows the number of 2009–2011 exam takers per year. Column (C) shows the proportion of colleges that appear in the Ministry of Education records, where colleges are weighted by the number of exit exam takers.

order of their normative program duration.[32] Column (B) shows the number of exit exam takers per year. The majority of exam takers are from university-level institutions, with fewer students from technical colleges.

Column (C) shows the fraction of these 310 colleges that appear in the Ministry of Education records that I use in my analysis. These proportions are weighted by the number of exam takers depicted in column (B). Column (C) shows that the Ministry of Education records included all Universities but are missing a few colleges that provide more technical training.[33] Overall, 96 percent of exit exam takers attend colleges that appear in the Ministry of Education records.

Another potential issue is that the Ministry of Education's institution coverage has been increasing over time. This could affect the main results if there are differential changes in coverage across regions. Panel B of Figure 6, however, suggests that this is not an issue. This panel depicts the number of college-program pairs that appear in the Ministry of Education records in each academic year. The number of programs is increasing over time due to both program growth and increasing data coverage. But there is no evidence of differential increases between affected and unaffected regions.

In sum, the main results in the paper are likely driven by students forgoing college altogether rather than switching to institutions that are not tracked by the Ministry of Education.

[32] Most programs at universities required 4–5 years of study, while programs at Technical/Professional Institutes typically take 2–3 years.

[33] The largest omitted institutions are the national police academy (*Dirección Nacional de Escuelas*) and the Ministry of Labor's national training service (*Servicio Nacional de Aprendizaje*). I also omit one university, *Pontificia Universidad Javeriana*, which has significant variation in the number of enrollees across years in the records. This omission does not affect my main results.

Table A2. Construction of high school sample

	(A) Unbalanced panel only	(B) Flexible calendar	(C) Final sample
# high schools	9,548	433	4,197
Missing high school	0.06	0.00	0.00
Proportion of all students	0.40	0.05	0.54
# students per school & year	20.1	58.2	61.8
Icfes percentile	0.46	0.41	0.54

Notes: The sample is 11$^{\text{th}}$ graders who took the Icfes in 2001–2011. Column (A) includes high schools that are missing exam takers in any year. It also includes schools for which I cannot cleanly merge in location and academic calendar information, schools that are listed in different departments or municipalities over time, and observations with missing school information. Column (B) includes high schools that are listed with a "flexible" academic calendar in any year, affected region schools that change calendars before 2010, and schools in other regions that ever change calendars. Column (C) includes the remaining high schools that have exam takers in every year.

# of students per school & year is the total number of students in 2001–2011 divided by the number of high schools divided by 11. Icfes percentiles are relative to all 11$^{\text{th}}$ grade exam takers in the same year and are calculated using the average of the scores from the six core components that did not change in 2001–2011: biology, chemistry, language, mathematics, philosophy, and physics.

A.4. **Sample of high schools.** This section describes how I select the sample of high schools for my analysis. My sample excludes two categories of schools. First, I exclude high schools that have zero Icfes exam takers in any year between 2001 and 2011. This includes schools for which I cannot cleanly merge in location and academic calendar information.[34] I also drop high schools that are listed in different departments or municipalities over time. This first set of excluded schools includes the 9,548 schools shown in column (A) of Table A2. It also includes six percent of all exam takers who have no high school information.

Second, I exclude high schools that are listed as having a "flexible" academic calendar in any year in 2001–2011. A flexible calendar means that students can begin the school year in either semester. I also omit schools in the affected regions that change calendars before 2010, and schools in other regions that change calendars in any year. These schools were likely more able to adapt to the academic calendar shift. This excludes the 433 schools shown in column (B) of Table A2.

My final sample includes the remaining 4,197 high schools in column (C) of Table A2 (see also Table 1). These schools have Icfes exam takers in every year from 2001–2011, and they contain 54 percent of all high school graduates during this time period. Schools in my sample have 62 students per cohort on average and are larger than excluded schools. Their students also perform better on the Icfes entrance exam.

[34] I identify high schools by numeric school IDs, but the Icfes records do not contain these IDs in 2008–2009. I must therefore merge in location and academic calendar data by high school name in these two years. This causes me to drop some high schools that have the same name and time of day (complete day/morning/afternoon) as another school in these years.

A.5. **Region-cohort level regressions.** Standard errors in my main analysis are clustered at the region level. Colombia has 33 administrative regions, which is below the rule-of-thumb for inference using the typical cluster-robust standard errors (e.g., Angrist and Pischke, 2009). This section addresses this potential issue by running region-cohort level regressions that mirror the school-cohort level regressions in Tables 2 and 3. This follows Donald and Lang (2007), who recommend "between-group" estimators in settings with few clusters but a large number of observations per cluster (see also Wooldridge, 2003).

I run the differences in differences regression

$$\bar{y}^t_{rc} = \gamma_r + \gamma_c + \beta^t \delta_{rc} + \epsilon_{rc}. \tag{A1}$$

This is similar to the benchmark specification (1), but equation (A1) is at the region-cohort level rather than the school-cohort level. The dependent variable, $\bar{y}^t_{rc}$, is the mean college enrollment rate t years after the Icfes exam in region r and exam cohort c. The regression includes region dummies, γ_r, cohort dummies, γ_c, and a treatment indicator, δ_{rc}, which equals one for regions and cohorts affected by the transition. Observations are weighted by the number of exam takers. Following the advice in Angrist and Pischke (2009), I use the maximum of OLS and robust standard errors. I use a $t(33-2)$ distribution to calculate statistical significance as suggested by Donald and Lang (2007).

Panel A in Table A3 shows the β^t coefficients from equation (A1) using switching schools and the same comparison group as in Table 2, and with δ_{rc} as an indicator for the 2009 affected region cohort. The coefficients are broadly similar to those in Table 2, but the standard errors are substantially larger in these region-cohort level regressions. Nonetheless, all estimates except for the $t = 3$ coefficient are statistically significant, even with the use of a $t(33-2)$ distribution for inference.

Panels B–D in Table A3 also estimate the region-cohort level regression (A1), but they use dependent variables $\bar{y}^t_{rc}$ that more closely reflect the three panels of Table 2. Specifically, I calculate the residuals from three regressions that mirror the panels of Table 2:

$$y^t_{hc} = \gamma_h + \epsilon_{hc}, \tag{A2}$$

$$y^t_{hc} = \gamma_h + \gamma_{gc} + \epsilon_{hc}, \tag{A3}$$

$$y^t_{hc} = \gamma_h + \gamma_{gc} + \tilde{\gamma}_h \times c + \epsilon_{hc}. \tag{A4}$$

γ_h are high school dummies, γ_{gc} are dummies for cells defined by high school groups g and cohort, and $\tilde{\gamma}_h \times c$ are school-specific linear cohort trends (see Table 2 for details).[35] I then calculate the mean of the residuals from these regressions for each region r and cohort c. The region-cohort mean residuals from (A2)–(A4) are the dependent variables $\bar{y}^t_{rc}$ in Panels

[35] Observations in regressions (A2)–(A4) are also weighted by the number of exam takers.

TABLE A3. Switching schools — Region-cohort level regressions
Dependent variable: Enrolled in college within t years of the Icfes

	(A)	(B)	(C)	(D)
Panel A. Regional means				
	0.5 years	1 year	2 years	3 years
Affected regions, 2009 cohort (δ_{hc})	-0.046^{***}	-0.026^{**}	-0.023^{*}	-0.022
	(0.009)	(0.011)	(0.013)	(0.013)
R^2	0.905	0.913	0.909	0.914
Panel B. Benchmark specification mean residuals				
	0.5 years	1 year	2 years	3 years
Affected regions, 2009 cohort (δ_{hc})	-0.049^{***}	-0.030^{**}	-0.028^{**}	-0.027^{**}
	(0.009)	(0.011)	(0.012)	(0.012)
R^2	0.782	0.815	0.799	0.800
Panel C. High school matching mean residuals				
	0.5 years	1 year	2 years	3 years
Affected regions, 2009 cohort (δ_{hc})	-0.047^{***}	-0.028^{**}	-0.027^{**}	-0.027^{**}
	(0.009)	(0.012)	(0.012)	(0.013)
R^2	0.093	0.024	0.019	0.019
Panel D. Linear high school trend mean residuals				
	0.5 years	1 year	2 years	3 years
Affected regions, 2009 cohort (δ_{hc})	-0.031^{***}	-0.023^{**}	-0.023^{**}	-0.019^{*}
	(0.010)	(0.011)	(0.010)	(0.010)
R^2	0.044	0.018	0.020	0.015
N	297	297	297	297
# regions	33	33	33	33

Notes: The sample is as in Table 2. All columns report coefficients on the treatment variable, δ_{rc}, which equals one for the 2009 cohort in the affected regions. All regressions include region dummies and cohort dummies. Regressions are at the region-cohort level with observations weighted by the number of exam takers. Parentheses contain the maximum of OLS and robust standard errors.

In Panel A, the dependent variables are mean region-cohort level college enrollment within t years of the Icfes, where t is listed in the column header. The dependent variables in Panels B, C, and D are the mean region-cohort residuals from equations (A2), (A3), (A4), respectively. See Table 2 for details.

Inference uses a $t(33-2)$ distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B–D of Table A3. The regressions and inference procedures are otherwise identical to those for Panel A.

The coefficients in Panels B–D of Table A3 closely correspond to those in Table 2, but the standard errors are in most cases substantially larger. However, all estimates are still statistically significant at the ten percent level.

Table A4 repeats these region-cohort level regressions for staying schools and the same comparison group as in Table 3. The table and methods are otherwise identical to those in Table A3 except the treatment variable, δ_{rc}, equals one for the 2010–2011 cohorts in the affected regions. The comparison group for these regressions excludes public schools, and private schools appear in only 25 regions in my sample. Thus inference in Table A4 is based on a $t(25-2)$ distribution.

The results in Table A4 mirror those in Table 3. The point estimates are similar in both tables, and despite larger standard errors in the region-cohort level regressions, nearly all coefficients are statistically significant.

In sum, the results in Tables A3 and A4 show that the main results of this paper are robust to stricter inference methods that address the relatively small number of clusters.

A.6. **Balance tests.** This section tests for effects of the calendar transition on the number of Icfes exam takers or their scores. This is a potential concern at switching schools, where students may have been affected by other elements of the transition such as changes in instructional time. I also test for effects at staying schools, although other transition elements are less of a concern as these schools did not change their academic calendars.

Panel A of Table A5 shows the results of these balance tests for switching schools. The sample is the same as in Table 2. I use the same benchmark regression (1) with different dependent variables. In column (A), the dependent variable is the number of Icfes exam takers in each high school and cohort. The differences in differences coefficient shows that the number of exam takers declined by two students in the 2009 cohort at switching schools relative to schools in other regions. This change is small relative to the mean of 62 students per school-cohort and is not statistically significant.

Column (B) tests for effects of the transition on students' entrance exam scores by using Icfes percentile as the dependent variable. The point estimate of the transition effect is one-half of one percentile in magnitude and is not statistically significant.

Panel B replicates these regressions for staying schools. The sample and regressions are identical to those in Table 3. In both columns, the point estimates are smaller than those in Panel A and are insignificant. There is no evidence of differential changes in the number of exam takers or their performance in the 2010–2011 cohorts at staying schools.

Thus, Table A5 suggests that the decline in college enrollment at switching schools and staying schools is not driven by changes in the student population or their college preparation.

A.7. **Heterogeneity by college selectivity.** This section explores the hypothesis that the observed college enrollment decline for switching school graduates was driven by an increased difficulty in gaining admission to selective colleges. This could arise if affected region colleges

Table A4. Staying schools — Region-cohort level regressions
Dependent variable: Enrolled in college within t years of the Icfes

	(A)	(B)	(C)
Panel A. Regional means			
	0.5 years	1 year	1.5 years
Affected regions, 2010–2011 cohorts (δ_{hc})	-0.060^{***}	-0.039^{**}	-0.028
	(0.014)	(0.016)	(0.017)
R^2	0.932	0.931	0.928
Panel B. Benchmark specification mean residuals			
	0.5 years	1 year	1.5 years
Affected regions, 2010–2011 cohorts (δ_{hc})	-0.069^{***}	-0.050^{***}	-0.039^{**}
	(0.015)	(0.016)	(0.016)
R^2	0.847	0.865	0.847
Panel C. High school matching mean residuals			
	0.5 years	1 year	1.5 years
Affected regions, 2010–2011 cohorts (δ_{hc})	-0.072^{***}	-0.053^{***}	-0.042^{**}
	(0.014)	(0.016)	(0.016)
R^2	0.164	0.128	0.112
Panel D. Linear high school trend mean residuals			
	0.5 years	1 year	1.5 years
Affected regions, 2010–2011 cohorts (δ_{hc})	-0.074^{***}	-0.082^{***}	-0.079^{***}
	(0.015)	(0.014)	(0.014)
R^2	0.106	0.133	0.128
N	275	275	275
# regions	25	25	25

Notes: The sample is as in Table 3. All columns report coefficients on the treatment variable, δ_{hc}, which equals one for the 2010–2011 cohorts in the affected regions. All regressions include region dummies and cohort dummies. Regressions are at the region-cohort level with observations weighted by the number of exam takers. Parentheses contain the maximum of OLS and robust standard errors.

In Panel A, the dependent variables are mean region-cohort level college enrollment within t years of the Icfes, where t is listed in the column header. The dependent variables in Panels B, C, and D are the mean region-cohort residuals from equations (A2), (A3), (A4), respectively. See Table 3 for details.

Inference uses a $t(25-2)$ distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

did not rebalance their January/September cohort slots in response to the shifting academic calendar.

To test this hypothesis, I run regressions that are similar to the benchmark specification in Table 2, but I use dependent variables that differ according to the selectivity of the college of enrollment. I use aggregate Ministry of Education data on the number of applicants

Table A5. Balance tests

Panel A. Switching schools

	(A)	(B)
	Dependent variable	
	# exam takers	Icfes percentile
Switching schools, 2009 cohort (δ_{hc})	−2.281	−0.005
	(1.816)	(0.009)
N	36,972	36,972
R^2	0.885	0.889
# regions	33	33
Dependent var. mean	61.537	0.527

Panel B. Staying schools

	(A)	(B)
	Dependent variable	
	# exam takers	Icfes percentile
Staying schools, 2010–2011 cohorts (δ_{hc})	−1.696	−0.002
	(1.832)	(0.004)
N	13,695	13,695
R^2	0.892	0.913
# regions	25	25
Dependent var. mean	48.277	0.664

Notes: The samples for Panels A and B are the same as in Tables 2 and 3, respectively. All columns report coefficients on the treatment variable, δ_{hc}, which is defined as in those tables. Regressions are at the school-cohort level with observations unweighted in column (A) and weighted by the number of exam takers in column (B). Regressions are otherwise identical to Tables 2 and 3 except for the dependent variables.

In column (A), the dependent variable is the number of Icfes exam takers in each high school and cohort. The dependent variable in column (B) is the school-cohort mean Icfes percentile (see Table 1).

Parentheses contain standard errors clustered at the region level. Dependent variable means are calculated from the 2001–2008 cohorts.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and admitted students to calculate the admission rate for each college in the years before the calendar transition (2007–2008). I define "less selective" colleges as those above the median admission rate, and "more selective" colleges as those below the median. I then run regressions as in Table 2 with dependent variables that measure only enrollment in less selective or more selective colleges.

Panel A in Table A6 shows the results for switching school graduates using the same sample as in Table 2. Columns (A) and (B) show the enrollment effect one semester after the Icfes exam. The decline in enrollment at switching schools occurs at both less selective and more selective colleges with a similar magnitude. Columns (C) and (D) show the effect on enrollment within three years of the exam. For less selective colleges, the three-year

Table A6. Heterogeneity by college selectivity
Dependent variable: Enrolled in college within t years of the Icfes

Panel A. Switching schools

	(A)	(B)	(C)	(D)
	Enrolled within 0.5 years		Enrolled within 3 years	
	Less selective colleges	More selective colleges	Less selective colleges	More selective colleges
Switching schools, 2009 cohort (δ_{hc})	-0.024^{***}	-0.025^{**}	-0.020^{*}	-0.007
	(0.004)	(0.010)	(0.010)	(0.007)
N	36,972	36,972	36,972	36,972
R^2	0.816	0.829	0.854	0.857
# regions	33	33	33	33
Dependent var. mean	0.080	0.081	0.186	0.196

Panel B. Staying schools

	(A)	(B)	(C)	(D)
	Enrolled within 0.5 years		Enrolled within 1.5 years	
	Less selective colleges	More selective colleges	Less selective colleges	More selective colleges
Staying schools, 2010–2011 cohorts (δ_{hc})	-0.025^{***}	-0.044^{***}	0.009	-0.048^{***}
	(0.007)	(0.004)	(0.012)	(0.006)
N	13,695	13,695	13,695	13,695
R^2	0.713	0.856	0.760	0.875
# regions	25	25	25	25
Dependent var. mean	0.154	0.183	0.258	0.292

Notes: The samples for Panels A and B are the same as in Tables 2 and 3, respectively. All columns report coefficients on the treatment variable, δ_{hc}, which is defined as in those tables. Regressions are identical to Tables 2 and 3 except for the dependent variables.

The dependent variables measure the fraction of students from each high school and cohort enrolling in less or more selective colleges within t years of the Icfes, as listed in the column header. College selectivity is defined using aggregate application data from the Ministry of Education (available in December 2015 at http://www.mineducacion.gov.co/sistemasdeinformacion/1735/w3-article-212400.html). The sample for this calculation includes technical- and university-level programs with a non-zero number of applicants and admitted students in the pre-transition periods for which data are available (2007–2008). I calculate the admission rate for each college by dividing the total number of admitted students by the total number of applicants over this time period. I then calculate the median admission rate across all colleges in the samples for Panels A and B. "Less selective" colleges are those above the median admission rate or that do not appear in the aggregate application data; "more selective" colleges are those below the median.

Parentheses contain standard errors clustered at the region level. Dependent variable means are calculated from the 2001–2008 cohorts.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

enrollment decline is similar to the initial effect. The enrollment decline mostly disappears at selective colleges and is not statistically significant.

Panel A thus shows that the results in Table 2 are driven by students forgoing colleges with open enrollment and not by an enrollment decline at selective colleges. This argues against the hypothesis that changes in admission rates are driving the main results. The results in Panel A of Table A6 are also consistent with evidence in Section 6 showing that low SES and low ability students were more likely to forgo college during the transition, as these students less frequently attend selective colleges.

Panel B of Table A6 replicates the regressions in Panel B for the staying school sample from Table 3. Unlike the results for switching schools, both the initial and longer-term enrollment declines at staying schools are driven primarily by more selective colleges. This reflects the fact that time gap for staying school graduates was due to a shift in annual programs from the September to the January calendar (see Figure 6). Annual programs typically cover specialized subjects and are thus offered primarily by selective colleges; non-selective colleges mainly offer semiannual programs in only the most popular subjects. Thus one would expect the staying school enrollment decline to occur primarily at more selective colleges that shifted their academic calendars, consistent with Panel B.

A.8. **Colombian household survey regressions.** Figure 5 in Section 4 provides suggestive evidence that many 2009 switching school graduates joined the labor force during the time gap created by the calendar transition. This section explores the statistical significance of this finding in a regression framework.

For this I use the same 2007–2010 waves of the Colombian household survey as in Figure 5. As in that figure, I define three cohorts of individuals denoted by c: those who turned 17 years old in the pre-transition years ($c \in$ 2007–2008), and those who turned 17 in the first year of the calendar transition ($c = 2009$).

Figure 5 plots dependent variables against time since the start of each cohort's last year of high school. In this section I use t to denote months with $t = 0$ representing the beginning of the last high school grade for students with on-time progression. The regressions below include the first seven quarters following the start of this last grade ($t \in$ 1–21). Further, while Figure 5 includes only 17 year olds in the affected regions, below I also include individuals in other regions as an additional comparison group. I use r to denote regions.

I estimate the regression

$$\bar{y}_{rct} = \gamma_{rc} + \gamma_{rt} + \gamma_{ct} + \beta_p \delta_{rc} + \epsilon_{rct}. \tag{A5}$$

The dependent variable, $\bar{y}_{rct}$, is a mean outcome in a region-cohort-month cell. The regression includes region-cohort dummies, γ_{rc}, region-month dummies, γ_{rt}, and cohort-month dummies, γ_{ct}. The treatment indicator, δ_{rc}, equals one for the 2009 cohort in the affected regions. I allow the treatment coefficient, β_p, to vary with time periods p that capture the quarter of typical high school graduation ($t \in$ 10–12), the time gap quarter for the 2009 cohort ($t \in$ 13–15), and the period following the time gap ($t \in$ 16–21).

Equation (A5) is thus a triple differences regression. The "time" dimension is months since the start of the last year of high school, t, and it uses both age cohorts c and regions r as comparison groups. Treatment is defined for the 2009 cohort in the affected regions as the months of delayed graduation and the periods following it ($t \geq 10$). β_p measures how the change in the outcome from previous months compares to analogous changes in other regions and cohorts. I weight observations by the sum of survey weights in each region-cohort-month cell. Since equation (A5) is a "between-groups" regression, I use the same inference procedures as in Appendix A.5.[36]

Table A7 shows the results from regression (A5). The dependent variable in column (A) is the fraction of individuals with a high school degree. The table reports the three β_p coefficients from this regression, which corresponds to the time periods listed in the beginning of each row for the 2009 cohort in the affected regions. July–September is the quarter in which affected region students typically finished high school in the years before the calendar transition. The first coefficient in Table A7 shows that in July–September 2009, the fraction of affected region 17 year olds reporting a high school degree falls by 15 percentage points relative to other cohorts and regions. This graduation effect disappears by the last quarter of 2009, as the other two coefficients in column (A) are near zero. Column (A) thus replicates the finding from Panel A in Figure 5 that the academic calendar transition led to a one-quarter graduation delay for some students in the 2009 cohort.

Column (B) of Table A7 also estimates equation (A5), but the dependent variable is an indicator for labor force participation. I define individuals as in the labor force if they report being either employed or looking for work in the household survey data. The first coefficient shows that there is no differential change in labor force participation in July–September 2009. However, affected region labor force participation rates increase by 11 percentage points in October–December 2009 relative to other regions and cohorts. This is the time gap between high school graduation and potential college enrollment for some 2009 graduates in the affected regions. Like Panel B of Figure 5, this result suggests that the time gap led some 2009 graduates to work or look for a job, but Table A7 shows that this effect is statistically significant at the ten percent level.

The last coefficient in column (B) shows that the labor force participation increase for the 2009 cohort persists into the first half of 2010 and is statistically significant. The magnitude of this effect falls by roughly 50 percent, consistent with the finding in Table 2 that the initial effect of the time gap on college enrollment declines by one-half in subsequent periods.

[36] Specifically, I use the maximum of OLS and robust standard errors (Angrist and Pischke, 2009), and I use a $t(24-4)$ distribution for inference (Donald and Lang, 2007) given 24 regions in the survey data and three coefficients of interest.

TABLE A7. 17 year olds in Colombian household survey data

	(A)	(B)
	Dependent variable	
Affected regions, 2009 cohort (δ_{rc}) in ...	Graduated high school	In the labor force
July−September 2009	-0.149^{***}	-0.011
	(0.036)	(0.036)
October−December 2009	-0.054	0.109^{*}
	(0.057)	(0.053)
January−June 2010	-0.037	0.057^{*}
	(0.036)	(0.029)
N	1,506	1,504
R^2	0.868	0.816
# regions	24	24
Dependent var. mean	0.372	0.244

Notes: Data are from the 2007–2010 monthly urban (*cabecera*) and rural (*resto*) GEIH household surveys. The sample is the 1990–1992 birth cohorts; this defines cohorts of individuals who turned 17 in 2007–2009. For February 2009 the sample is current 16 year olds because birthdates are missing. The sample includes only children, grandchildren, or other relatives of the household head.

The columns estimate the region-cohort-month level regression (A5) with t defined as months since the start of 11$^{\text{th}}$ grade for most high schools in the region. This means that $t = 0$ in September before the cohort year in the affected regions, and $t = 0$ in February of the cohort year in other regions. The sample includes months $t = 1$ to $t = 21$ for each cohort. 2006 surveys are not available, so values at $t \leq 3$ are missing for the 2007 cohort in the affected regions.

The dependent variable in column (A) is the fraction of each region-cohort-month cell with a high school degree or above. In column (B), the dependent variable is the fraction of each region-cohort-month cell in the labor force, defined as appearing in the employed (*ocupados*) or unemployed (*desocupados*) survey. This variable is lagged one month because survey questions refer to labor force activity in the prior 1–4 weeks (e.g., October values are from the November survey).

All columns report coefficients β_p on the treatment variable, δ_{rc}, which equals one for the 2009 cohort in the affected regions. This variable is interacted with dummies for three time periods p defined by $t \in$ 10–12, $t \in$ 13–15, and $t \in$ 16–21. This corresponds to the time periods listed in the beginning of each row for the 2009 cohort in the affected regions.

All regressions include region-cohort dummies, region-month dummies, and cohort-month dummies. Regressions use weights that are fixed over time, which are the sum of survey weights across all months within non-missing cells defined by region, age, gender, urban/rural survey, and secondary educated mother. Parentheses contain the maximum of OLS and robust standard errors. Dependent variable means are calculated from the 2007–2008 cohorts.

Inference uses a $t(24-4)$ distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In sum, Table A7 shows that the effects of the calendar transition on high school graduation and labor force participation documented in Figure 5 are statistically significant in a regression framework.