ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 23-713

Cows Don't Give Milk: An Effort Model of College Graduation

Maria Marta Ferreyra
The World Bank

Carlos Garriga
Federal Reserve Bank of St. Louis

Juan David Martin-Ocampo
Banco de la Republica Colombia

Angelica Maria Sanchez Diaz
Georgetown University

This paper estimates a dynamic model of college enrollment, progression, and graduation. A central feature of the model is student effort, which has a direct effect on class completion and an indirect effect mitigating risks on class completion and college persistence. The estimated model matches rich administrative data for a representative cohort of college students in Colombia. Estimates indicate that effort has a much greater impact than ability on class completion. Failing to consider effort as an input to class completion leads to overestimating ability's role by a factor of two or three. It also promotes tuition discounts based on a pre-determined student trait—ability—rather than effort, which can be affected through policy, thus limiting higher education's potential for social mobility.

VERSION: January 2023

Suggested citation: Ferreyra, Maria Marta, Carlos Garriga, Juan David Martin-Ocampo, and Angelica Maria Sanchez-Diaz. (2023). Cows Don't Give Milk: An Effort Model of College Graduation. (EdWorkingPaper: 23-713). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/hyrt-b441

Cows Don't Give Milk: An Effort Model of College Graduation*

Maria Marta Ferreyra
The World Bank

Carlos Garriga
Federal Reserve Bank of St. Louis

Juan David Martín-Ocampo
Banco de la República Colombia

Angelica Maria Sánchez Díaz
Georgetown University

December 26, 2022

Abstract

This paper estimates a dynamic model of college enrollment, progression, and graduation. A central feature of the model is student effort, which has a direct effect on class completion and an indirect effect mitigating risks on class completion and college persistence. The estimated model matches rich administrative data for a representative cohort of college students in Colombia. Estimates indicate that effort has a much greater impact than ability on class completion. Failing to consider effort as an input to class completion leads to overestimating ability's role by a factor of two or three. It also promotes tuition discounts based on a pre-determined student trait—ability—rather than effort, which can be affected through policy, thus limiting higher education's potential for social mobility.

Keywords: Higher Education, academic progression, dropout.
JEL codes: E24, I21

*We thank Kartik Athreya, Raquel Bernal, Adriana Camacho, Stephanie Cellini, Doug Harris, Felicia Elionescu, Oksana Leukhina, CJ Libassi, Alexander Ludwig, B. Ravikumar, Fabio Sanchez, James Thomas, and Susan Vroman for useful comments and suggestions. We thank Andrea Franco and Nicolas Torres for excellent research assistance. We benefited from seminars at The World Bank, the St. Louis Fed, the DC Economics of Education Working Group, and Universitat Autonoma de Barcelona, and by sessions at AEFP 2018, SED 2018, NASMES 2019, and LACEA 2018. Ferreyra acknowledges financial support from The World Bank. The views expressed herein do not necessarily reflect those of The World Bank, the Federal Reserve Bank of St. Louis, the Federal Reserve System, or of Banco de la República Colombia. Email contacts: mferreyra@worldbank.org, carlos.garriga@stls.frb.org, jmartioc@banrep.gov.co, ams668@georgetown.edu. Corresponding author: Maria Marta Ferreyra, Global Engagement and Knowledge, Education Global Practice, The World Bank. Washington, DC 20433. Phone: (202) 473 3166. Email: mferreyra@worldbank.org.

1 Introduction

It is old wisdom among peasant farmers that a cow does not just give milk; rather, the farmer needs to milk the cow every day to extract the milk. Similarly, college students do not obtain a degree just by attending college. Even a highly able student needs to exert some effort to attend classes, study, do homework, and take exams in order to pass classes and graduate. Raising a country's stock of college graduates therefore requires more than college subsidies to expand enrollment; it requires policies to encourage student effort. Since student effort is hard to observe and measure, much of the literature has focused on the role of ability and credit constraints as key drivers of college enrollment, progression, and graduation.

This paper uses a structural model and indirect inference to assess the role of effort in a dynamic model of college enrollment, progression, and graduation using Colombian data.[1] Latin American countries are an excellent case study because their higher education enrollment rates have grown rapidly in recent years, rising from 23 to 52 percent between 2000 and 2017. Since only 14 percent of their working age population has completed higher education, *Mincerian* returns are high—104 percent on average relative to a high school diploma (Ferreyra et al., 2017). Such high returns should attract many individuals to higher education, yet liquidity constraints and the lack of credit markets have severely limited access.

Our structural model incorporates key aspects of students' decisions to attend and graduate from college. Two critical drivers of educational outcomes are the role of student effort and the presence of extreme borrowing constraints.[2] The model features a heterogeneous population of high school graduates who vary in student ability (or academic preparation for college) and parental resources; for some students, the combination of low parental resources and extreme borrowing constraints renders college inaccessible. To graduate, a college student must complete a set number of classes in a given time. Each year she chooses the number of classes she expects to complete and chooses effort according to this target. She also faces shocks which might prevent her from completing a class (performance shock) or from remaining enrolled (dropout shock.) Therefore, student effort has a direct effect on class completion and an indirect effect mitigating risks to class completion and college persistence. These shocks vary across students and over time, and are related to the student's past performance in order to capture the notion that good past performance facilitates good future performance. Labor market wages, in turn, depend on educational attainment (high school graduate, college dropout, college graduate) and experience.

The dynamic model is estimated using Simulated Method of Moments, and captures observed patterns for Colombia's high school class of 2005. In this class, a staggering 70 percent of students come from low-income families and only 32.3 percent of students enroll in college within five years. Among those who attend college, only 46 percent graduate—mostly late. Student ability plays a large role in year 1-survival but a lesser role in subsequent performance.

[1]Our paper relates to a large literature estimating sequential schooling models under uncertainty, with seminal contributions by Keane and Wolpin (2001), Eckstein and Wolpin (1998), and Keane (2002). This literature models college enrollment, performance, and college outcomes, and uncovers structural parameters based on students' observed choices during college.

[2]The conventional view is that, in the absence of credit market frictions, college enrollment should be positively correlated with ability and uncorrelated with parental resources (Cameron and Heckman, 1998, 1999; Carneiro and Heckman, 2002). Lack of family resources, limited opportunities to work during college, and missing credit markets for student loans are clear impediments to college access in countries such as Colombia.

The speed at which students accumulate the completed classes required for graduation (henceforth, "cumulative performance" or "cumulative classes completed") is highly persistent. As a result, early performance is a strong predictor of final outcomes.

The estimates indicate that effort has a much greater impact—about ten times larger—than ability on class completion. If effort were not modeled, ability's role would be overestimated by a factor of two or three. Moreover, a specification without effort does not fit the variation in classes completed or college outcomes among students and over time. In other words, it does not capture the fact that some low-income, low-ability students perform well, or that some high-income, high-ability students do not.

These findings have critical policy implications. When effort is not considered, policies to expand college access tend to alleviate credit constraints—by targeting low-income students—or promote positive selection—by targeting high-ability students. While alleviating credit constraints can help students enroll in college, it will not necessarily help them graduate without specific effort incentives. Targeting high-ability students, in turn, rewards students with a predetermined trait which they cannot affect during college—a strong academic preparation—rather than promoting a choice they *can* control during college—a high effort level. Given the crucial role of effort in college outcomes, college subsidies that do not promote it can only produce disappointing results in terms of human capital accumulation and the ensuing social and intergenerational mobility that highly unequal societies—such as those in Latin America—typically seek through higher education.

Our findings are in line with a growing literature that attempts to measure the contribution of effort to academic outcomes. Some papers exploit a direct or indirect measurement of effort. For example, Zamarro, Hitt, and Mendez (2019) use data from the Program for International Student Assessment (PISA) to show that different effort measures explain about a third of observed cross-country test score variation. Stinebrickner and Stinebrickner (2004) rely on time-use surveys to estimate the effects of study time on grades. Ariely et al. (2009) show that providing incentives can help the average student improve her test performance. Beneito et al. (2018) provide evidence that the tuition increase implemented by Spanish colleges in 2012 boosted student effort. Ahn et al. (2019) model student effort in response to instructor grading policies. Other papers use a complementary approach that exploits random assignments to peers to identify the contribution of effort. Stinebrickner and Stinebrickner (2008) present causal evidence of effort on academic performance using the fact that students are assigned at random to their first-year roommates. Similarly, Mehta, Stinebrickner and Stinebrickner (2019) explore the effects of peers' study time on college outcomes.[3] We contribute to this growing literature by embedding effort in a structural model of college outcomes.

A novel contribution of our paper is the role of effort in mitigating college risk. The idea that higher education is risky is not new (i.e., Levhari and Weiss, 1974; Altonji, 1993; Akyol and Athreya, 2005), but the recent availability of college transcript data in the U.S. has helped estimate the role of risk in student performance. Hendricks and Leukhina (2017, 2018) find that

[3]A related body of research explores the role of learning or acquiring information throughout college as a mechanism to explain outcome variation among students. In this literature, which includes Arcidiacono (2004), Arcidiacono et al. (2016), Ozdagli and Trachter (2011), Stinebrickner and Stinebrickner (2014), and Trachter (2015), students learn about their ability and preferences for college or specific majors as they accumulate college experience. Similar to these papers, students in our model adjust their expected graduation probability based on classes completed and choose effort accordingly.

college transcripts reveal substantial and persistent heterogeneity in students' credit accumulation rates that are strongly related to graduation outcomes, and that more than half of college entrants can predict whether they have at least an 80 percent graduation probability. Stange (2012) finds that the large uncertainty faced by college students makes them place a high value on the ability to drop out at any point rather than pre-commit to completing all graduation requirements. Our paper is similar to these in the use of administrative data to track students' performance, but different in that the risks associated to class completion or college persistence are not fully exogenous—as in these papers—but depend on an endogenous choice—student effort. In our model, students can exert effort in order to mitigate their uncertainty. We develop a measure of anticipated student uncertainty and estimate its response to student effort. The estimated elasticity of anticipated uncertainty with respect to effort is negative and large at -2.34. By responding to greater risk with greater effort, students are self-insuring against uncertainty while also enhancing their class completion and college performance.

The rest of the paper is organized as follows. Section 2 describes data sources, institutional environment, and key data aspects. Section 3 presents the model, and Section 4 describes the estimation strategy and results. Section 5 explores the role of effort in relation to ability, policy implications of omitting effort, and interactions between effort and risk. Section 6 concludes.

2 Data and Stylized Facts

This section begins with a description of our Colombian data sources and institutional environment; further details can be found in Appendix A.1. It then presents key stylized facts on college enrollment, dropout rates, academic performance, and college outcomes.

2.1 Data Sources and Institutional Environment

Our data consists of student- and program-level information drawn from three administrative datasets: Saber 11, SPADIES, and SNIES. Saber 11 contains students' test scores at the national mandatory high school exit exam (also named Saber 11) as well as self-reported socioeconomic information. Saber 11 is a standardized test that measures a student's preparedness for higher education, reflecting not only her talent but also the quality of her primary and secondary education. We use it as a measure of student ability, which we define as academic readiness for higher education. Family income is reported in brackets defined relative to the monthly legal minimum wage (MW), equal to 381,000 Colombian pesos (COP) in 2005 (US$ 1 = 2,321 COP in 2005). All monetary values are expressed in COP of 2005.

SPADIES tracks higher education students throughout college. It records the students' program, number of classes attempted and passed per semester, and graduation or dropout date. It does not record which specific classes were attempted, or the grades obtained. SNIES contains program- and institution-level information, including field of study and tuition.

To analyze college enrollment, we focus on the 2005 cohort of high school graduates. We calculate deciles and quintiles of their ability distribution; in what follows, deciles and quintiles always refer to this distribution. For consistency with the model, we classify students into "student types" defined by combinations of student ability and family income brackets. Table A1 shows the distribution of student types in this cohort. While a remarkable 70 percent of high school graduates come from the lowest two income brackets, less than 5 percent come from the top one. Not surprisingly, income and ability are strongly and positively correlated.

For our analysis of final outcomes and academic progression, we focus on students from the 2006 college entry cohort that enroll in five-year bachelor's programs within five years of finishing high school. Since every program requires a different number of classes for graduation, we normalize the requirement to 100 classes for every program to facilitate exposition. We assume that students are required to complete the same number of classes (20) every year. For a given year, we use the term "classes completed" (or "performance") to denote the number of classes completed in that year, and use "cumulative classes completed" (or "cumulative performance") for the total number of classes completed *over all years*, up to (and including) that one. To graduate on time, the student should complete the on-track requirements of 20, 40, 60, 80, and 100 cumulative classes completed by the end of years 1 through 5, respectively.

It is useful to classify students into *tiers* based on their cumulative performance relative to the corresponding year's on-track requirement. Tiers 1 through 4 correspond to students who complete the following percentage of the year's on-track requirement: 95 percent or more for tier 1; (85,95] percent for tier 2; (65, 85] percent for tier 3; and 65 percent or less for tier 4 (see Table A2 for further details on tier classification).[4] Importantly, a student can change tiers over time—catching up to a higher tier or falling behind to a lower one.

2.2 Stylized Facts

The data has distinctive features that our model seeks to capture. We describe them below.

Fact 1: Students of higher income or ability are more likely to enroll in college. Although the overall enrollment rate is 32 percent, enrollment rates vary widely among student types (Table 1), and rise with income and ability.[5] On average, the enrollment gap between the highest and lowest income brackets is about 50 percentage points (pp), similar to the gap between the highest and lowest ability.

Fact 2: Dropout rates are high, particularly for low-ability students in year 1. Similar to enrollment rates, dropout rates vary widely across student types (Table 2). Conditional on income, higher ability students have lower dropout rates; the average dropout rate gap between the highest and lowest ability quintiles is about 25 pp. Dropout rates vary less by income, suggesting that family resources might affect enrollment more than dropout decisions.

Dropout rates are far from uniform over time. As Figure 1 shows, over a quarter of college students (half of all dropouts) drop out in year 1, and the first two years account for about 70 percent of all dropouts. Only 45.7 percent of students from our entry cohort graduate—15.1 percent graduates on time (in five years) and 30.6 percent graduates late (in 6-8 years).[6] Ability is a strong predictor of final college outcomes (on-time graduation, late graduation, year-1 dropout, year-2 dropout, later dropout.) High-ability students are more likely to graduate,

[4]For example, consider a student who accumulates 16, 35, 42, 50, and 60 classes by the end of years 1 through 5 respectively. This amounts to 80 (=16/20), 88 (=35/40), 70, 62.5, and 60 percent of the corresponding on-track requirements. Thus, the student falls in tiers 3, 2, 3, 4, and 4 in years 1 through 5 respectively.

[5]For comparison, in the U.S. the enrollment rate of individuals ages 16-24 who finished high school in 2005 and started college right away (rather than within five years) is 44.6% (Source: Digest of Education Statistics).

[6]For comparison, in the U.S. 59.2 percent of students from the 2006 cohort graduated within six years—39 percent on time (in four years), and 20.2 percent late ˙Source: Digest of Education Statistics.

Table 1: Enrollment Rates by Income and Ability.

Income Bracket	Ability quintile 1	2	3	4	5	Total
5+ MW	32.85	44.14	58.87	69.23	83.85	73.38
3-5 MW	28.71	39.75	48.41	62.99	79.24	62.03
2-3 MW	20.34	28.72	36.96	48.03	67.88	43.50
1-2 MW	13.94	18.36	23.85	33.84	54.22	28.05
<1 MW	9.05	12.67	17.20	26.56	43.93	17.67
Total	13.43	19.15	26.20	38.93	63.74	32.29

Source: Calculations based on SPADIES and Saber 11 for 2005 high school graduates.
Notes: Each cell reports percent of high school graduates from a given income bracket and ability quintile who enrolled in a bachelor's program between 2006 and 2010. Income is reported in brackets; MW = monthly minimum wage. Ability is reported in quintiles of the standardized Saber 11 scores; quintile 1 is the lowest.

Table 2: Dropout Rates by Income and Ability.

Income Bracket	Ability quintile 1	2	3	4	5	Total
5+ MW	81.36	65.83	61.48	52.13	39.04	44.73
3-5 MW	74.23	69.44	62.21	57.86	43.77	51.33
2-3 MW	68.54	67.58	63.73	57.68	46.53	55.09
1-2 MW	71.59	66.64	62.21	57.66	50.56	57.82
<1 MW	69.04	67.95	61.34	55.94	50.30	58.71
Total	70.99	67.44	62.37	56.96	45.84	54.36

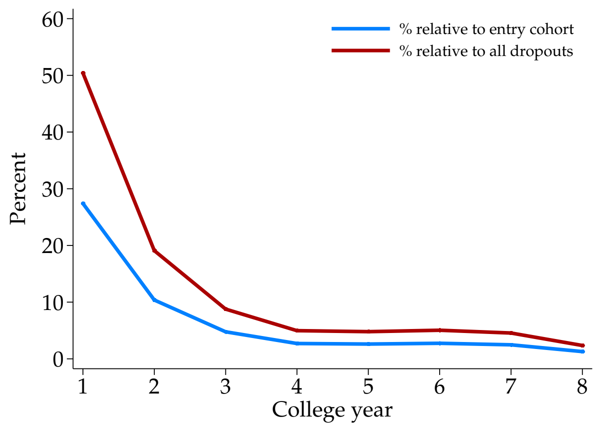
Source: Calculations based on SPADIES for students from the 2006 entry cohort (first semester).
Notes: Each cell reports the percent of students from a given income bracket and ability quintile who drop out of their bachelor's program. A student is classified as a dropout if she does not graduate within eight years. Income is reported in brackets; MW = monthly minimum wage. Ability is reported in quintiles of standardized Saber 11 scores. Quintile 1 is the lowest.

whether on time or late (Figure 2), whereas low-ability students are more likely to drop out, particularly in year 1.

Fact 3: Cumulative performance varies more within than across abilities.
Figure 3's panel a shows average number of classes completed by ability quintile in year 1. The thick black line depicts the average over all students, whereas the color lines depict averages among students conditional on their final outcome. The figure shows that high-ability students complete more classes, on average, than low-ability students in year 1, but also shows a pattern repeated every year: average classes completed vary little *across* abilities—overall and conditional on final outcomes—but greatly *within* abilities. Conditional on ability, by the end of year 1 on-time graduates complete more classes than late graduates, who in turn complete more classes than dropouts. In other words, the number of classes completed—as early as in year 1— is a powerful predictor of final outcomes. This point is further illustrated in Figure 3's panel b, which classifies students at end of the first year into performance tiers and shows their final outcomes conditional on those tiers. Students in the year-1 top tier are indeed more likely than others to graduate. Performance is persistent, as discussed in the next stylized fact.

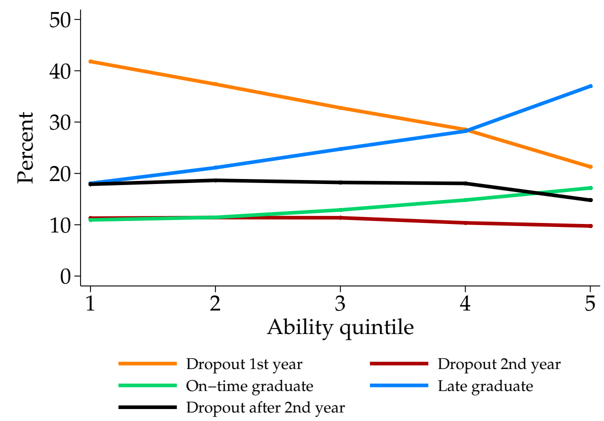
Fact 4: Cumulative performance is highly persistent over time.
To further analyze persistence, Figure 4 depicts, for every year, a student's probability of

Figure 1: Dropout Timing.



Source: Calculations based on SPADIES, for students from the 2006 entry cohort (first semester).
Notes: The blue line shows the percent of students who drop out by college year. The red line shows, among all dropouts, the percent of those who drop out by college year.

Figure 2: College Outcomes by Ability.



Source: Calculations based on SPADIES. Students belong to the 2006 entry cohort (first semester).
Notes: For each quintile, the graph shows the percent of students who attained each outcome. Percents add to 100 by quintile.

attaining each of four outcomes—same-tier persistence, dropout, catch up (to a higher tier), and fall behind (to a lower tier)—by year's end conditional on her previous year's performance tier.[7]

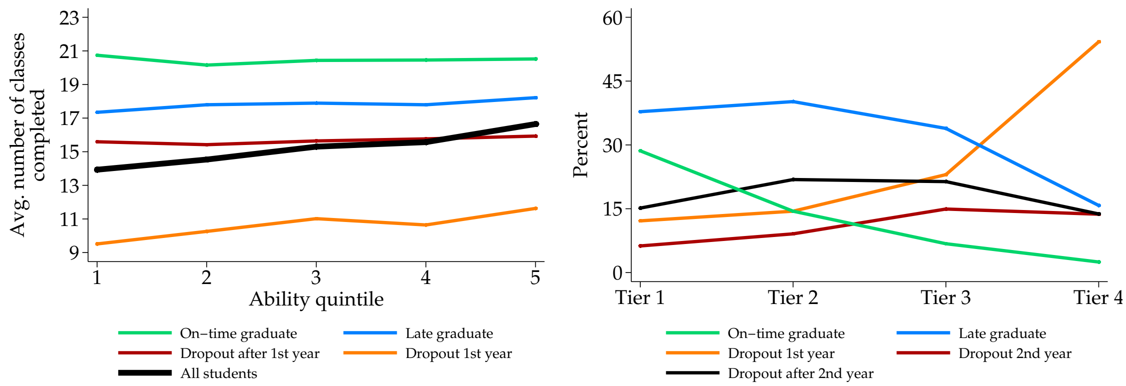
Same-tier persistence rises over time (panel a), partly because dropout rates fall over time (panel b). Although low-performing students are more likely than others to drop out, students from all tiers face a non-zero dropout probability. Some students move across tiers by catching up or falling behind (panels c and d). Higher-performing students are more likely to catch up and less likely to fall behind than others, yet all students face a non-zero probability of falling behind. Despite transitions among tiers, a student in one of the top three tiers is most likely to remain in it the following year. Reaching year 5 does not guarantee graduation. Although most students in the top three tiers graduate (Figure 6), and most tier-1 students graduate on time, most students in the bottom tier do not graduate.

Fact 5: Cumulative performance becomes more concentrated over time.

As low-performing students drop out over time, the distribution of cumulative performance becomes more concentrated at the top tiers. Panel a of Figure 5 shows the distribution of students across performance tiers at the end of year 1 by ability quintile. Although top-tier performance is most likely for the highest-ability students, performance varies greatly within ability quintiles, and a sizable fraction of students from every quintile performs at the middle tiers. Because of this "thick middle," cumulative performance varies little, on average, *across* abilities, echoing Fact 3. Similar patterns hold for year 5 (panel b), although by then all

[7] For example, a student who finished the first year in tier 2 has second-year probabilities of persistence, dropout, catch up, and fall behind equal to 29, 14, 20, and 39 respectively. Beginning in year 5, students can also "transition" into the graduation outcome.

Figure 3: First-Year Classes Completed and College Outcomes.



a. First-year classes completed by college outcome

b. College outcomes by first-year tier of classes completed

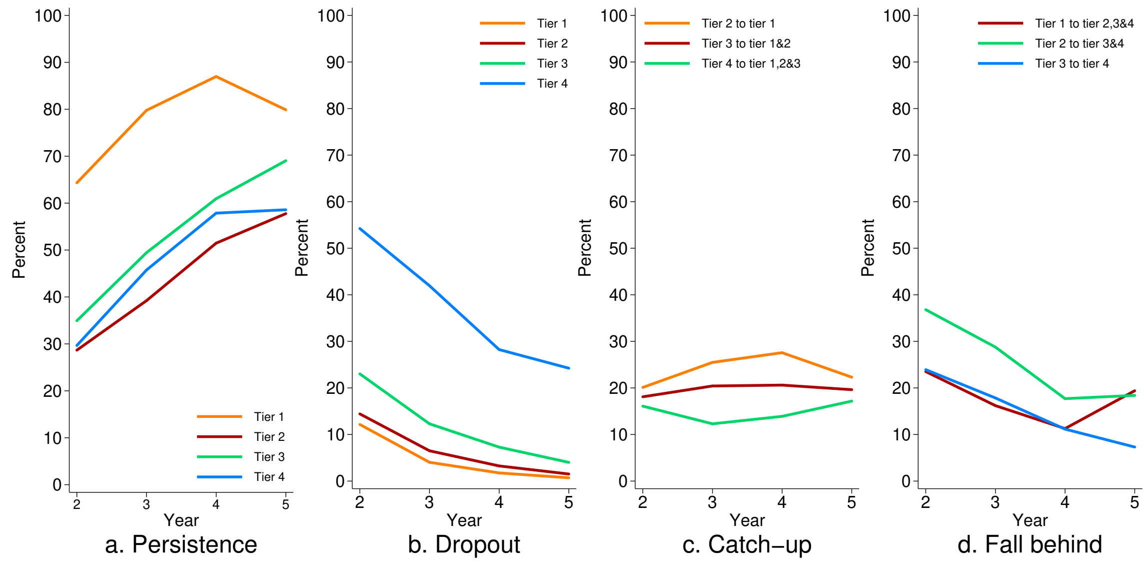
Source: Calculations based on SPADIES for students from the 2006 entry cohort (first semester).
Notes: In panel a, each color represents a college outcome. The green line, for instance, shows the average number of classes completed by the end of year 1 by students of a given ability quintile who went on to graduate on time. The thick black line does the same for all students regardless of college outcome. In panel b, the graph shows the percent of students from a given performance tier by the end of year 1 that attain each college outcome. For the first year, tier 1 corresponds to 19+ classes completed; tier 2 to [17, 19); tier 3 to [13, 17); and tier 4 to [0, 13).

performance distributions are mostly concentrated in the top three tiers because many bottom performers have already dropped out. Further, in both years the performance distribution of low-ability students has more dispersion than that of high-ability students.

Fact 6: Higher-ability students attempt a higher number of classes.
Although a student cannot fully control her performance, she can control the number of classes she attempts in a given year, which is informative of her intended effort. As Figure 7 shows, on average students attempt fewer than the 20 required classes and, on average, high-ability students attempt more classes than low-ability students. We will return to these facts when discussing identification of our empirical model.

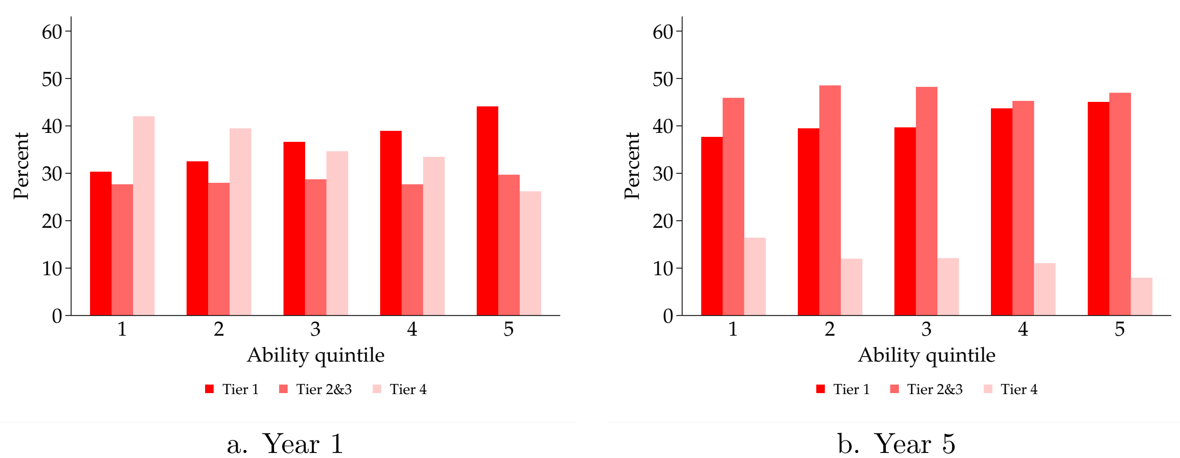
Taking stock. The data shows that higher ability and higher income students are more likely to enroll in college. Conditional on enrolling, lower ability students are substantially more likely to drop out, particularly in year 1. Through this channel, ability serves a strong predictor of graduation. Ability, however, is not a strong predictor of cumulative performance. Cumulative performance varies little, on average, across abilities, yet varies widely within abilities. Further, it is highly persistent over time—students who start on track are more likely to remain on track, less likely to fall behind, and more likely to catch up should they fall behind. As a result, cumulative performance is highly predictive of final outcomes. Our model seeks to capture these data features.

Figure 4: Tiers of Cumulative Classes Completed: Transitions Throughout College.



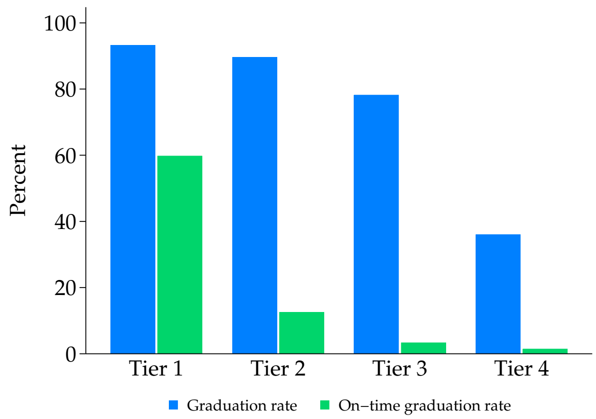
Source: Calculations based on SPADIES for the 2006 entry cohort (first semester).
Notes: Each panel shows the probability that a student who ended the previous year in a given tier experiences the following outcomes in the current year: persist in the tier (panel a), drop out (panel b), catch up (rise) to a higher tier (panel c), or fall behind to a lower tier (panel d). For a given year and tier, probabilities add up to 100 across panels. For example, a student who finished year 1 in tier 3 is depicted in green. In year 2, she is 35, 23, 18, and 24 percent likely to persist in tier 3, drop out, catch up to tiers 1 or 2, and fall behind to tier 4 respectively.

Figure 5: Tiers of Cumulative Classes Completed by Ability.



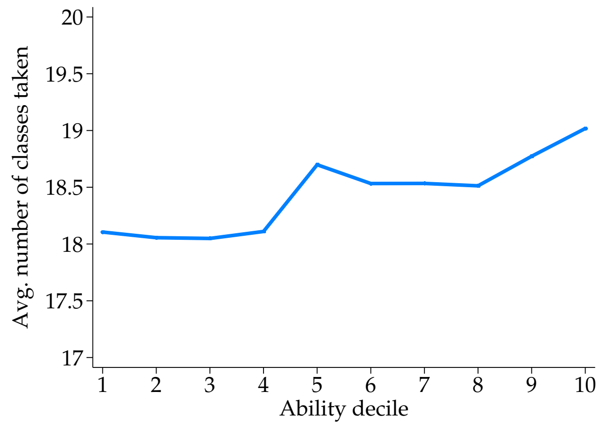
Source: Calculations based on SPADIES for students from the 2006 entry cohort (first semester).
Notes: For students of a given ability quintile who start year 1, panel a shows their classification into tiers of cumulative classes completed by the end of the year. Panel b does the same for students beginning year 5.

Figure 6: Graduation by Year-5 Cumulative Classes Completed.



Source: SPADIES. Students belong to the 2006 entry cohort (first semester).
Notes: The figure refers to students who start year 5 and classifies them into tiers based on their cumulative classes completed by the end of the year. For each tier, the blue bar shows the percent of students who graduate, and the green bar shows the percent of students who graduate on time, relative to all the tier's graduates.

Figure 7: Average Number of Classes Attempted in Year 1, by Ability.



Source: Calculations based on SPADIES for students from the 2006 entry cohort (first semester).
Notes: Figure shows the average number of classes attempted by students in year 1 by ability decile.

3 Model of College Effort

We model a representative cohort of high school graduates who differ in ability, family income, and idiosyncratic preferences for college enrollment, and choose whether to enroll in college or enter the labor force as high school graduates. College graduation requires completing a set number of classes. The combination of ability, effort, and performance and dropout shocks determines class completion and college outcomes. Students can mitigate the risk posed by these shocks by being on track with classes completed. In the labor market, wages depend on educational attainment (high school graduate, college graduate, or college dropout) and work experience. Students can only enter college at the end of high school; once they leave, they cannot return.

It is important to clarify what "ability" and "effort" capture in the model. Ability refers to academic readiness—not innate ability—and therefore captures pre-college elements such as high school quality and home environment. It is predetermined at the time of college enrollment and remains fixed during college. Effort, in contrast, is chosen by the student every year during college. It represents all student inputs that can change over time and are costly to the student, such as quantity and intensity of study time. Exerting greater effort may not mean studying longer but rather more effectively. The critical point is that effort is costly, particularly for students with lower academic readiness (who may not have been given the appropriate study skills, for example). Further, the model captures the notion that, by choosing high effort, low-ability students might perform at the same level as high-ability students who choose low effort.

3.1 Student Types and Preferences

High school graduates differ in ability, $\theta \in \{\theta_1, \theta_2, \ldots, \theta_{n_\theta}\}$, and parental resources received if they enroll in college, $y \in \{y_1, y_2, \ldots, y_{n_y}\}$. The combination of ability and parental resources defines the student's type $j = 1, 2, \ldots, J$, with $J = n_\theta \times n_y$. For simplicity we refer to y as "income." Students receive y only if enrolled in college; otherwise they earn a labor market wage as explained below.

A period, t, is equal to a year (and an academic year during college). For college students, instant utility depends on consumption, c, and study effort, e. Effort cost is heterogeneous across students and depends on ability. The instant utility at t is $U(c_t, e_t, \theta) = u(c_t) - g(e_t, \theta)$. Function U satisfies $u' > 0$, $u'' < 0$, $g_1, g_{11} > 0$, $g_2 < 0$, and $g_{12} < 0$. Thus, effort cost is increasing and convex, and total and marginal effort costs are lower for higher-ability students. Once the student leaves school and joins the labor force, utility depends only on consumption. The discount factor for all individuals is $\beta > 0$.

3.2 College Decisions

We model college as a multi-period, risky investment. To graduate, students must complete a required number of classes, h^{grad}. Students must spend at least five years in college in order to graduate, and cannot exceed eight. We use three concepts related to number of classes per year. First is $\bar{x}$, the annual number of classes required to graduate on time. Second is q_t, the expectation formed by the student, at the beginning of the year, of how many classes she will complete (henceforth, "target.") The target can be greater or smaller than $\bar{x}$, and drives the student's effort choice. Third is x_t, the actual number of classes completed in the year, which is a function of the student's effort, ability, and intervening luck. It can be greater or smaller than q_t because of "good" or "bad" luck, respectively.

In the data, we do not observe a student's target; rather, we observe the number of classes she attempts, which is neither x nor q. Since tuition is often constant regardless of the number of classes taken in Colombia, students usually attempt the maximum allowed number of classes even if they do not expect to complete them—namely, their target might be lower than the number of classes taken. We model the target rather than the number of classes attempted because a student's target determines her effort. In estimation we use data on the number of classes attempted, which serves as an upper bound for q.

3.2.1 College Technology

Let h_t denote the cumulative number of classes completed up to the end of t, or $h_t = \sum_{n=1}^{t} x_n$. Students start college with $h_0 = 0$, and by the end of year 1 attain $h_1 = x_1$. While enrolled, students complete classes in t according to the following production function:

$$x_t = H(\theta, e_t, z_t)\ \overline{x}. \tag{1}$$

The function describes the number of classes completed, x_t, as a multiple of the annual requirement for on-time graduation, $\overline{x}$. The scalar $H(.)$ is a function of ability, effort, and a shock to classes completed (or performance shock), $z_t > 0$. We assume H is non-negative and can be lower or greater than one, implying that the student can complete more or fewer classes than the annual requirement. We also assume that, when the student supplies zero effort, she completes zero classes: $H(z_t, \theta, 0) = 0$. The shock z_t, drawn from a continuous distribution known to the student, includes a random *i.i.d.* component as well as a component that depends on her ability, number of cumulative classes completed up to the beginning of t, and year. The shock's dependence on past cumulative classes completed seeks to capture the observed persistence and allows students to affect their future "luck" by completing more classes in the current period. The shock's dependence on ability reflects that ability may be correlated with other, unobserved elements that may systematically affect "luck."[8]

If the student knew z_t when choosing effort, then choosing e_t would be equivalent to choosing classes completed, x_t. Since, as explained below, the student chooses e_t before z_t is realized, choosing effort is equivalent to choosing a target, q_t, where $q_t = E(x_t)$. Thus,

$$q_t = E[H(z_t, \theta, e_t)]\ \overline{x}. \tag{2}$$

Assuming $H(\cdot)$ is linear on z_t so that it can be expressed as $H(z_t, \theta, e_t) = z_t\tilde{H}(\theta, e_t)$, target and effort are functions of $E(z_t)$:

$$q_t = E(z_t)\tilde{H}(\theta, e_t)\ \overline{x} \qquad \text{and} \qquad e_t = \tilde{H}_e^{-1}[\theta, q_t/(E(z_t)\ \overline{x})]. \tag{3}$$

Meanwhile, the actual number of classes completed, x_t, is a function of the effort chosen given the target, and of the realized z_t. Cumulative classes completed by the end of the year, h_t, is

$$h_t = h_{t-1} + x_t. \tag{4}$$

3.2.2 The Student's Dynamic Optimization Problem

The student faces a sequential problem. We distinguish between the *pre-graduation years*, when she cannot yet graduate, and the *graduation years*, when she is eligible to graduate. We divide every year into two sub-periods. In the *first sub-period*, the student chooses her target number of classes and hence effort. At the end of it she receives the performance shock, which determines her actual number of classes completed. In the *second sub-period*, she graduates if she has accumulated the required number of classes; otherwise she draws a shock that determines whether she will remain in college the following year or drop out ("dropout shock"). Thus, as long as she has not completed her graduation requirements, the student draws two shocks per year—one to classes completed in the year and another to college continuity. The two shocks are endogenous in the sense that they depend on cumulative performance, which in turn depends on past effort. In a given year, the student's state vector is (t, h_{t-1}, θ, y). Figure B1 summarizes the timing of events and decisions, described in detail below.

Pre-graduation years (t = 1,...,4). During these years, students have not yet completed the required number of classes for graduation, or $h_t < h^{grad}$. In year 1, all students start with zero cumulative classes completed and are heterogenous only in their type. Since students of a given

[8]For example, lower ability students may choose less selective programs than their more able counterparts, or may have lower levels of the non-cognitive skills necessary to succeed in college. In the first case, $E(z)$ would be higher for lower ability students; in the second case, it would be lower. In our estimation we let the data identify the sign of the relationship between $E(z)$ and θ.

type may receive different performance shocks during the year, they may complete a different number of classes during the year. As a result, from year 2 onward students are heterogenous not only in their type but also in the number of cumulative classes completed with which they begin the year, h_{t-1}.

During these years, at the beginning of the *first sub-period* the student chooses e_t, and at the end z_t is realized and determines the number of cumulative classes completed for the year, $h_t \geq h_{t-1}$. In the *second sub-period* the student receives the dropout shock, $d_t^{drop} = \{0, 1\}$, which determines whether she will remain in college next year or drop out, respectively. The probability that this shock would lead her to drop out is a function of her cumulative performance up to that point, h_t, as well as her type and college year:

$$\Pr(d_t^{drop} = 1 \mid z_t) = \tilde{p}^d(t, h_t, \theta, y). \tag{5}$$

We assume that the student must drop out if her progress is too slow. If she accumulates less than a pre-specified number of classes completed, h_t^{drop}, she must drop out, or $\tilde{p}^d(t, h_t < h_t^{drop}, \theta, y) = 1$. If, in contrast, she has completed $\bar{x}$ classes each year and is on track for on-time graduation, her dropout probability is very low: $\tilde{p}^d(t, t\bar{x}, \theta, y) \approx 0$. In general, $\tilde{p}^d$ depends negatively on cumulative performance and, when choosing effort, the student internalizes its impact on future dropout shocks.

If the student drops out, she joins the labor market the following year as a college dropout and receives the corresponding wage; the value of dropping out is $V^{drop}(t+1)$. Meanwhile, the value of remaining in college is $V^{coll}(t+1, h_t, \theta, y)$.

Graduation years (t = 5,...,7). The *first sub-period* is similar to that of the previous years. In the *second sub-period*, students who have fulfilled the graduation requirements, $h_t \geq h^{grad}$, graduate and enter the labor market, whose value is $V^{grad}(t+1)$, the following year. Remaining students draw the dropout shock to determine whether they will remain enrolled.

Terminal year (t = 8). This is the last year that the student can spend in college. By the end of it, only two outcomes are possible: the student graduates if $h_8 \geq h^{grad}$, or drops out otherwise. Continuation values are equal to $V^{grad}(9)$ and $V^{drop}(9)$ respectively.

We can now present the student's dynamic optimization problem from the *first subperiod* of each college year:

$$\begin{aligned} V^{coll}(t, h_{t-1}, \theta, y) = \max_{e_t} \Bigg\{ & U(c_t, e_t, \theta) + \beta E_z \Big[\mathbf{1}_{\{t \geq 5\}} \Pr\left(h_t \geq h^{grad}\right) V^{grad}(t+1) + \\ & \Pr\left(h_t < h^{grad}\right) \left[\tilde{p}^d(t, h_t, \theta, y) V^{drop}(t+1) + \right. \\ & \left. \left(1 - \tilde{p}^d(t, h_t, \theta, y)\right) V^{coll}(t+1, h_t, \theta, y)\right] \Big] \Bigg\}, \end{aligned} \tag{6}$$

$$
\begin{aligned}
s.t. \quad & c_t = y - T(t, h_{t-1}, \theta, y) \\
& h_t = h_{t-1} + x_t \\
& x_t = H(z_t, \theta, e_t)\overline{x} \\
& c_t > 0.
\end{aligned}
$$

Here, the argument of $E_z[\cdot]$ is the continuation value function. Variable $T(\cdot)$ is tuition, constant regardless of the target, q_t. We write $T(\cdot)$ in general form as the policymaker could make it vary by year, cumulative classes completed, ability, or income. In our baseline and data, however, it varies only by y (see Section 4.1 below). For low-income students tuition might exceed income, which would violate the $c_t > 0$ constraint and make enrollment unfeasible. Note the severe credit constraint: students cannot borrow to pay for tuition nor can they save.[9] The policy function for optimal effort, $e^*(t, h_t, \theta, y)$, solves the dynamic problem defined in (6).

3.3 Workers

An individual can join the labor force after graduating from high school or college, or after dropping out from college. The worker's optimization problem, written in recursive form, is

$$
V^m(t) = \max_{c_t}\{u(c_t) + \beta V^m(t+1)\}, \tag{7}
$$

$$
s.t. \qquad c_t = w_t^m,
$$

where $V^m(t)$ is the value function of a worker with educational attainment $m = \{hs, grad, drop\}$, denoting high school graduate, college graduate, and college dropout respectively.[10] The worker's wage, w_t^m, is specific to educational attainment and varies with t to allow for returns to experience.

3.4 College Enrollment Decision

To decide whether or not to enroll in college, a high school graduate compares the expected payoff of two choices—going to college, and joining the labor force as a high school graduate. The enrollment decision is a discrete choice problem, where the payoff associated to each option is the sum of three components. The *first component* is the expected value of either going to college, $V^{coll}(t = 1, h_0 = 0, \theta, y)$, or entering the labor force as a high school graduate, V^{hs}. The *second component* is a type-specific preference for college enrollment, $\xi_j = \tilde{\xi}(\theta_j, y_j)$. This captures unmeasured elements—such as parental familiarity with college—that might vary across types and affect a student's propensity to enroll in college (we normalize the unobserved preference for joining the labor force as a high school graduate to zero for all types.) The

[9] We do not model student's decision to work during college because our administrative data does not record this information. Further, data from Colombia's National Survey of Time Use (*ENUT*) reveals that high-income college students are more likely to work during college than their lower-income counterparts, suggesting that the primary motivation to work is not necessarily to pay for college (details available upon request). For a model of student workers, see Garriga and Keightley (2007).

[10] We assume that workers consume all their earnings and do not have access to credit markets, which is an accurate representation of developing economies. Since wages rise with experience and workers discount the future at the interest rate, they have no incentives to save.

purpose of this second component is to help us match college enrollment rates by student type; it varies across—but not within—student types.[11] The *third component* is a choice-specific, idiosyncratic shock for each individual, ϵ^{hs} and ϵ^{coll}, corresponding to working as a high school graduate or enrolling in college, respectively. This component varies within and across types, and helps us match the fact that not all individuals of a given type make the same choices. We assume that ϵ^{hs} and ϵ^{coll} are *iid* and distributed Type I Extreme Value with a scaling factor of σ_ϵ.

The individual chooses to attend college if

$$\underbrace{V^{coll}(1,0,\theta_j,y_j)+\xi_j+\sigma_\epsilon\epsilon^{coll}}_{\text{Value of going to college}} \geq \underbrace{V^{hs}+\sigma_\epsilon\epsilon^{hs}}_{\text{Value of working as a high school graduate}} \tag{8}$$

As a result, the probability of college enrollment for an individual of type j is

$$P^{coll}(\theta_j,y_j)=\frac{\exp\{(V^{coll}(1,0,\theta_j,y_j)+\xi_j)/\sigma_\epsilon\}}{\exp\{(V^{coll}(1,0,\theta_j,y_j)+\xi_j)/\sigma_\epsilon\}+\exp\{V^{hs}/\sigma_\epsilon\}}, \tag{9}$$

Its complement, $P^{hs}(\theta,y)=1-P^{coll}(\theta,y)$, is the probability of joining the labor force as a high school graduate. Since P^{coll} varies by ability, θ, and parental transfer, y, predicted enrollment rates vary across student types.

4 Empirical Implementation and Estimation

In this section, first we describe the computational implementation, estimation strategy, and identification of our model. Next, we present parameter estimates and goodness of fit evidence.

4.1 Student Types, Tuition, and Workers

Student types. To construct student types for our computational implementation, we start from the empirical distribution by ability and income of the 2005 cohort of high school graduates. The distribution, shown in Table A1, classifies high school graduates by ability quintile and family income bracket. We refine this distribution to work with ability deciles instead of quintiles. To construct the θ values, which must range between 0 and 1, we standardized the Saber 11 test scores and normalize them between 0 and 1.[12] The 5th, 15th, ...95th percentiles of the distribution of normalized scores are our θ values. To construct the parental transfer values, y, we face the obstacle that our administrative data reports income in brackets rather than levels. In order to assign students in each income bracket a monetary transfer value, we turn to an external data source—Colombia's 2005 household survey data (SEDLAC)—which reports household income and size. We classify households in the survey into the same income brackets as in our administrative data, and calculate average per-capita household income by bracket. This measure of disposable income by household member is our proxy of parental

[11] This component is analogous to the product mean utility modeled by Berry et al. (1995), which allows them to match observed market share. See Appendix C for further details.

[12] The normalized test score is calculated as $(sts-\min(sts))/(\max(sts)-\min(sts))$, where sts denote the standardized test score.

transfers for college students, y. The first two columns of Table A3 show the mapping between income brackets and the resulting parental transfer levels. The final distribution of ability and parental transfers, $\Phi(\theta, y)$, includes $J = 50$ student types and features a positive, strong correlation between θ and y.

Tuition. Student-level data on actual tuition expenses is not available in our administrative data. Therefore, we estimate the tuition paid by students of a given y as the average annual tuition paid by students from the corresponding income bracket at public institutions.[13] Table A3 shows the estimated tuition paid by parental income bracket. As the table shows, average per-capita household income (our proxy for parental transfer) varies greatly across income brackets, but tuition varies much less. Although public institutions provide income-based tuition discounts, the highest-income individuals pay proportionally less than their lower-income counterparts: their per-capita income is about twenty times as large than that of the lowest-income individuals, yet their tuition is only 2.5 times as large.[14]

Workers We use Colombia's household survey (SEDLAC) to compute average wages by educational attainment and experience in 2005. Returns to education are substantive. For workers aged 18-60, the average wage of a college graduate, a college dropout with at least two years of complete college, and a college dropout with up to one year of complete college is 160, 58, and 28 percent higher than the average wage of a high school graduate, respectively (see Table A4).[15] Returns to experience are also large. Among college (high school) graduates, the average wage of experienced workers is 35 (29) percent higher than the average wage of inexperienced workers. Consistent with the data, we assume that the returns to experience of college dropouts are the same as those of high school graduates.

Time periods in the model map onto students' and workers' ages. For example, $t = 1$ corresponds to age 18, whereas the end of work life in period $t = 48$ represents age 65. Regardless of her educational attainment or when she joined the labor force, the individual accrues returns to experience (or becomes "experienced") beginning at age 35 ($t = 28$). College students can drop out and join the labor force after the initial college year starting at age 19 ($t = 2$.)

[13] We use tuition at public institutions because there is always a public institution that the student can attend. Modelling the choice of college type (public or private) is beyond the scope of this paper. We estimate average tuition at public institutions using SNIES and SPADIES.

[14] In Colombia, as in other countries, parental resources matter greatly to college enrollment even controlling for ability. This provides strong evidence for credit constraints limiting college access, as discussed in a large literature. Lochner and Monge-Naranjo (2011) develop a model that helps explain the rising importance of family income for college attendance in the U.S. even in the presence of credit markets. Solis (2017) finds that relaxing credit constraints in Chile had a positive impact on enrollment and college years completed, particularly for low-income students. Although Colombia is a large developing economy, the market for student loans is very limited, covering only 7 percent of students in 2003 (ICETEX 2010).

[15] This creates, in effect, four college attainments: high school, college, some college (one year), some college (two or more years). The two "some college" categories correspond to college dropouts. We work with two rather than one dropout category because their wages are quite different and can hence affect dropout timing.

4.2 Functional Forms

Production function of classes completed. The technology to complete classes, $H(z_t, \theta, e_t)$, has constant returns to scale in ability and effort:

$$x_t = \overline{x}H(z_t, \theta, e_t) = \overline{x}z_t(\theta^{\alpha} e_t^{1-\alpha}), \tag{10}$$

where $\alpha \in (0, 1)$ is the elasticity of classes completed with respect to ability. Consistent with the model, we set the required number of classes per year, $\bar{x}$, to 20 classes, and the minimum number of classes required for graduation, h^{grad}, to 98 (rather than 100, since in the data some students graduate with slightly fewer than 100 classes, perhaps due to measurement error).

Performance shock. The shock that determines "luck" in class completion, z_t, is parameterized with a flexible functional form:

$$z_t = \exp\{-\exp\{-(\kappa_0 + \kappa_1 d_1 + \kappa_h \tilde{h}_{t-1} + \kappa_\theta \theta + (\sigma + \sigma_1 d_1 + \sigma_\theta \theta)\nu_t)\}\}, \tag{11}$$

where $\tilde{h}_{t-1}$ is a measure of past cumulative number of classes completed, with $\tilde{h}_{t-1} = \ln(h_{t-1})$ for every $t > 1$ and $\tilde{h}_0 = 0$ for $t = 1$. The terms associated with d_1 allow the shock distribution to be different in year 1, when $d_1 = 1$, than in other years. The shock depends on an *iid* component, ν_t, drawn from the uniform distribution $U(0, 1)$. The functional form in (11) ensures that the shock is bounded, $z_t \in (0, 1)$, for any combination of parameter values and for all $\tilde{h}, \theta \in \mathbb{R}$. Importantly, all the parameters in (11) affect both the mean and variance of z_t. In Section 4.5 we discuss the effect of $\tilde{h}_{t-1}$ and θ on this mean and variance at our parameter estimates.

Dropout shock. The functional form for the probability that the dropout shock forces the student to drop out depends on student characteristics characteristics (ability and income), college year, and cumulative performance:

$$\tilde{p}^d(t, h_t, \theta, y) = \frac{\exp\{\delta(t, \theta, y) + \pi\tilde{h}_t\}}{1 + \exp\{\delta(t, \theta, y) + \pi\tilde{h}_t\}}, \tag{12}$$

where $\delta(t, \theta, y)$ is a year-, ability- and income-specific fixed effect, and $\tilde{h}_t$ measures cumulative performance up (and including) to the current academic year. The student internalizes the effect of effort on the likelihood of a good cumulative performance through a higher $\tilde{h}_t$, which decreases the dropout probability. The shock's dependence on individual characteristics captures the fact that shocks outside the student's control, which might force her to drop out, are more likely for some students than others. For example, low-income students may be more likely to drop out when a family member suffers an adverse health shock or losses a job, as they may need to work in support of the family or care for other members.

The specification in (12) is flexible, and encompasses the *exogenous dropout probability* as a special case. This can be calculated by evaluating $\tilde{p}^d(t, h_t, \theta, y)$ at $\pi = 0$, and is the dropout probability for students of a given type even if cumulative performance does not affect dropout chances. It is exogenous because it is independent of effort. For example, low-income, low-ability students may have a high exogenous probability in year 1—perhaps due to little parental guidance to navigate the transition to college—yet a lower one in subsequent years.

Preferences. The preference specification for college students assumes separability of consumption and effort allowing different curvatures for each one. Formally,

$$U(c,e,\theta) = \frac{c^{1-\rho}-1}{1-\rho} - \mu\frac{e^{\gamma}}{(1+\theta)^{k}}, \tag{13}$$

where the curvature with respect to consumption and effort is represented by ρ and γ, respectively.[16] This formulation allows student ability, θ, to shape the marginal cost of effort with a scope determined by k. Workers' preferences are a special case of students' preferences, setting $\mu = 0$, and are given by

$$u(c) = \frac{c^{1-\rho}-1}{1-\rho}. \tag{14}$$

The discount factor is set to $\beta = 0.04$, which corresponds to an implicit real discount rate of 4 percent. The parameter σ_{ϵ}, which is the scaling factor of the Type I Extreme Value distribution of idiosyncratic preferences for college and work, is set equal to 1.

4.3 Estimation Strategy

The model does not have a closed-form solution and requires a numerical algorithm (described in Appendix C) to solve the dynamic optimization problem for a given set of parameters, $\tilde{\Theta} = (\Theta, \boldsymbol{\xi}, \boldsymbol{\delta})$, where

$$\Theta = (\rho, \mu, \gamma, k, \alpha, \kappa_0, \kappa_1, \kappa_h, \kappa_\theta, \sigma, \sigma_1, \sigma_\theta, \pi) \tag{15}$$

is the vector of parameters common across individuals; $\boldsymbol{\xi}_{J\times 1}$ contains the type-specific unobserved preferences for college, ξ_j (see equation (9)); and $\boldsymbol{\delta}_{(J*8)\times 1}$ contains the exogenous dropout probability fixed effects, $\delta(t, \theta_j, y_j)$, for the J types and 8 years (see equation (12)).

We estimate the model using Simulated Method of Moments (SMM).[17] Our estimation searches for the value of Θ whose predicted moments, $\hat{\mathbf{M}}(\Theta)$, best match the observed ones, $\mathbf{M}$. Formally, the SMM parameter estimates solve the following problem:

$$\arg\min_{\Theta} \quad (\hat{\mathbf{M}}(\Theta) - \mathbf{M})' W^{-1} (\hat{\mathbf{M}}(\Theta) - \mathbf{M}), \tag{16}$$

where Θ is a 13×1 vector of parameters, $\mathbf{M}$ and $\hat{\mathbf{M}}$ are 585×1 vectors of sample and predicted moments, respectively, and W is a diagonal weighting matrix whose diagonal contains the standard error of the sample moments. We compute numerically the predicted values, $\hat{\mathbf{M}}$, for every value of Θ.

The 585 moments we match are listed in Table 3. They capture aggregate and distributional patterns associated with the stylized facts listed in Section 2.2: final college outcomes; dropout rates; academic progression and class completion; persistence; and target number of classes. Note that enrollment rates are matched by construction and are not a moment targeted in estimation (see Appendix C.4 for further details.)

[16] When students enroll in college, we assume they receive a government stipend equal to 1,000,000 COP in order to guarantee a sufficiently high consumption after paying tuition, particularly for the lowest-income students. In reality, this stipend can be viewed as the multiple subsidies (to transportation, food, and so forth) provided by Colombia's government to college students.

[17] The estimation of $\boldsymbol{\delta}$ and $\boldsymbol{\xi}$ is nested within the model solution for a given of value of Θ, in the spirit of Berry et al. (1995). See Appendix C for further details.

Table 3: **Moments Matched in Estimation.**

Data aspect	Moments	Number of Moments
Dropout rates	Dropout rate by year.	8
	Dropout rate by ability quintile and income.	25
	Dropout rate by ability decile.	10
College outcomes	College outcomes by ability quintile.	25
	Fraction of students that graduate by year (years 5-8).	4
Cumulative classes completed	Average number of cumulative classes completed by year, ability quintile, and college outcome.	140
	Average number of cumulative classes completed by year and ability decile (years 1-5).	50
	Distribution of students into tiers of cumulative classes completed, by year.	24
	Distribution of students into tiers of cumulative classes completed, by ability quintile and year (years 1-5).	75
Transition probabilities	Pr(tier Y in $t+1$|tier X in t) for years 1-7.	112
	$\Pr(d_t^{drop} = 1 \vert \text{tier } X \text{ in } t)$ for years 1-8.	32
Target number of classes	Average target number of classes by ability decile and year.	80
Total		585

Source: Own estimation.
Notes: Moments per year are computed for years 1-8 unless otherwise specified. Tiers are 1-4, based on cumulative classes completed. Under "College outcomes," outcomes include on-time graduate, late graduate, drop out first year, drop out second year, drop out after second year. Under "Cumulative classes completed," which are calculated by year, outcomes include on-time graduate (until year 5), late graduate, drop out this year, drop out later (until year 7); "this year" and "later" refer to the year under consideration. Under "Transition probabilities", t refers to the year; tiers X and Y are 1,...,4. Observed data for target number of classes is average number of classes attempted by the corresponding students.

4.4 Identification

A critical challenge is identifying the role of ability, effort, and performance shocks in the production of classes completed. For a given student type, we observe its ability by construction. Section 2.2 shows that average classes completed varies little by ability but greatly within abilities, suggesting that ability alone cannot explain the observed variation in classes completed. Similarly, college outcomes vary substantially within abilities as well. These data features create a role for the non-ability determinants of performance in our model, namely effort and the performance shock. Model assumptions and functional forms help to separately identify the roles of these two elements. Below we provide more intuition for identification by describing each parameter's role in the model—namely, how parameter changes affect model predictions and therefore the fit of the data. Section 5 complements this discussion.

Effort (e_t)**:** If effort had no role in the number of classes completed ($\alpha = 1$) or were costless ($\mu = 0$), then all students would take the required number of classes per year. The fact that students take, on average, a lower number of classes than required (see Section 2.2) indicates that effort does have a role in classes completed and helps identify μ. An increase in μ leads to lower targets, effort, and number of classes completed. It also leads to lower college enrollment—particularly for low-income students, who have a lower consumption than others to compensate for a given effort. Speed of accumulation of classes completed, as well as transitions among tiers over time, helps identify γ. A high γ penalizes high effort levels and makes it costly to catch up. The fact that higher ability students take more classes, on average, than their lower-ability counterparts indicates they have lower effort costs and identifies k. An increase in k raises the variance of average target, effort, and classes completed across abilities.

Performance shock (z_t)**:** In the data, average number of classes completed varies widely within abilities. Some of this variation can be explained by income, as the model predicts that, conditional on ability, effort varies by income. Nonetheless, income alone does not explain all the within-ability variation of classes completed. The remainder of this variation, then, is explained by the performance shock, z. Since θ is between 0 and 1, we restrict z to this range to pin down the scale for effort. An increase in κ_0 makes shocks more favorable for all students, thus raising the number of classes completed and lowering dropout rates across the board. Parameter κ_1 is an intercept shifter that makes the scale of z comparable across years. An increase in κ_θ makes the expected shock relatively more favorable for high-ability students and raises the dispersion in average classes completed and college outcomes across abilities. Parameter σ is identified by the overall variation of classes completed conditional on ability, whereas σ_θ is identified by the greater variation of classes completed among low- than high-ability students (see Section 2.2.) The higher overall variation of classes completed in year 1 relative to other years identifies σ_1. After year 1, κ_h is identified by the persistence of students in their performance tiers.

Ability (θ_t)**:** Given the role of effort and performance shocks in the production of classes completed, the variation of average classes completed across abilities identifies α. This variation rises with an increase in α.

Other parameters: An increase in ρ raises the aversion to consumption variations over time and decreases the propensity to college enrollment. It also lowers the speed of class accumulation and increases time-to-degree. Finally, the sensitivity of dropout rates with respect to current classes completed, conditional on student income and ability, identifies π.

A sufficient condition for local identification is full rank for the matrix of first derivatives of the moments' predicted values with respect to the parameter vector when evaluated at the true parameter point. The evaluated matrix at the estimated parameters has full column rank.

4.5 Parameter Estimates

Table 4 shows the estimates for the 13 parameters common across individuals, Θ. A key parameter is the elasticity of classes completed with respect to ability, α, which ranges between 0 and 1. The estimated α is low (0.085), consistent with the low variation in average classes completed across abilities. As a result, the estimated elasticity of credits completed with respect to effort $(1 - \alpha)$ is high (0.915). Since our ability measure captures college academic readiness rather than innate ability, these estimates indicate that student effort has a much greater role (about 10 times larger) than college academic readiness in class completion. A direct implication of these estimates is that, in order to affect human capital accumulated and classes completed, tuition subsidies must seek to affect effort rather than target high-ability students.

The estimated preference curvature with respect to effort, γ, is equal to 4.73. This indicates a very high marginal cost of effort—exceeding typical quadratic costs—which prevents large catch-up efforts. The estimate of the parameter relating effort cost and ability, k, indicates that effort cost has a strong, negative relationship with ability. When $k = 0$, effort cost is the same for all abilities, whereas when $k = 1$ effort cost falls with ability, at a decreasing rate. Our estimate of 1.23 yields the same qualitative pattern as $k = 1$ but with an even greater gap in effort cost among high- and low-ability students.

The estimate for ρ, equal to 0.882, indicates that students have a high tolerance to consumption changes over time (or a high intertemporal elasticity of substitution), which raises their willingness to attend college and gives them a lower risk aversion.

Table 4: Parameter Estimates.

Parameter	Symbol	Estimate
Utility function		
Consumption curvature	ρ	0.882
Effort weight	μ	0.062
Effort curvature	γ	4.727
Effort cost w.r.t. ability	k	1.225
Number of classes completed		
Elasticity w.r.t. ability	α	0.085
Performance shock		
Constant	κ_0	-4.207
Year 1 shifter	κ_1	3.534
Persistence component	κ_h	1.304
Ability component	κ_θ	0.407
Std. dev. of *iid* shock	σ	1.789
Std. dev. of *iid* shock - Year 1 shifter	σ_1	0.317
Std. dev. of *iid* shock - Ability shifter	σ_θ	-1.282
Dropout shock		
Cumulative performance component	π	-2.951

Source: Own estimation.

To interpret the parameter estimates related to the performance shock, z, we consider their implications on the shock's mean and variance. According to the estimates, $E(z_t)$ is higher for students with higher past performance. In other words, good past performance creates a future "good luck" and good future performance, helping us match the persistence patterns observed in the data. In addition, low-ability students have higher $E(z_t)$ and $Var(z_t)$ than high-ability students. Their (slightly) higher mean of z is consistent with the observed fact that low-ability students enroll in less selective (and presumably less demanding) programs than their abler counterparts.[18] Together with effort, this higher mean helps us match the good performance and graduation of some low-ability students. The greater shock variance for low-ability students, in turn, helps us match their greater performance variance.

From a student's perspective, an important question is whether future "luck" depends more on ability or past performance. To examine the relative impact of past performance and ability

[18] We measure program selectivity as the average Saber 11 test score of the program's students. We find a negative correlation between student ability and program selectivity. Results are available upon request.

on z given our estimates, we consider two hypothetical students. Student A is much more able than B, with $\Delta_\theta = \theta_A - \theta_B = 0.22$. This is a large ability difference, equal to the difference between the 55th and the 5th ability percentile, or between the 95th and the 75th percentile. Student A has an additional advantage over B, because by the beginning of t she has completed one more class than B. Since A is abler than B, her $E(z_t)$ should be lower, yet because she has completed more classes, her $E(z_t)$ should be higher. Based on our estimates, having completed that one additional class gives her the same $E(z_t)$ as B's, even though B is much less able. In other words, $E(z_t)$ is much more sensitive to h_{t-1} than to θ. This makes z highly persistent and more dependent on something the student can control—her performance—than on ability, which she cannot control. Once again, the policy implication is clear in favor of seeking to affect effort—rather than targeting ability—through tuition subsidies.

Another important question, for a student, is how effort (through its impact on cumulative effort) affects the dropout risk. The estimated value for π, equal to -2.95, indicates that an additional class completed by the end of the year, on average, decreases the probability of dropping out by about 5 pp, or approximately 10 percent of the sample average dropout rate. Since $\tilde{p}^d(\cdot)$ has a logistic functional form, this marginal effect is stronger for students with intermediate values of the dropout probability rather than values close to zero or one.

How important is the exogenous dropout risk component relative to the endogenous one in the dropout probability? Recall that our full set of parameter estimates includes the dropout probability fixed effects in Equation (12), $\hat{\delta}(t, \theta, y)$. To illustrate the relative magnitude of $\hat{\pi}$ and $\hat{\delta}(t, \theta, y)$, consider the average number of additional classes that a student from the second ability quintile ("Q2 student") must complete to attain the same dropout probability as a student from the top ability quintile ("Q5 student"). In year 1, she must complete more than 4 additional classes—20 percent of the annual requirements—reflecting a high exogenous dropout probability. In year 5 she only needs one additional class completed, as the Q2 students reaching year 5 are approximately on par with Q5 students. The important point is that, early on, low-ability students face a high exogenous dropout probability, which they can only counter through high initial effort or favorable performance shocks.

4.6 Goodness of Fit

The estimated model fits the data well and replicates the patterns described in Section 2.2. In terms of graduation, the model matches closely the observed overall graduation rate (45.64 percent) with a predicted rate of 45.02 percent. The model replicates the distribution of dropouts by academic year (Table 5). In both the observed and predicted data the dropouts in the first two (three) years account for 70 (78) percent of all dropouts.

The observed and predicted distribution of dropout rates across student types are presented in Table 6. The model is consistent with the observed patterns, whereby low-ability and low-income students are more likely to drop out. Further, the model captures the academic progression patterns observed in the data. Table 7 shows that the model replicates the observed patterns of persistence, drop out, catch-up, and fall-behind.

As shown in Figure 8, the model fits extremely well the distribution of college outcomes by ability. As a result, the observed on-time graduation rate of 15.1 percent is predicted perfectly, and the observed fraction of 2005 high school graduates that complete college (14.7 percent) is

Table 5: Goodness of Fit: Dropout by Academic Year.

Year	Observed	Predicted
1st	27.41	27.92
2nd	10.38	10.43
3rd	4.77	4.94
4th	2.7	2.83
5th	2.61	2.81
6th	2.74	2.92
7th	2.48	1.94
8th	1.28	1.19
Total	54.36	54.98

Source: SPADIES for observed data; fitted values for predicted data.
Notes: Table shows the observed and predicted percent of students who drop out each year.

Table 6: Goodness of Fit: Dropout Rates by Income and Ability.

	Ability quintiles											
Income	Observed values						Predicted values					
Bracket	1	2	3	4	5	Total	1	2	3	4	5	Total
5+ MW	81.4	65.8	61.5	52.1	39.1	44.7	84.4	67.8	63.4	51.9	40.7	46.7
3-5 MW	74.2	69.4	62.2	57.9	43.8	51.3	81.6	68.3	62.2	57.3	44.4	52.5
2-3 MW	68.5	67.6	63.7	57.7	46.5	55.1	70.8	67.5	61.9	60.1	44.4	55.6
1-2 MW	71.6	66.6	62.2	57.7	50.6	57.8	69.9	66.5	60.9	56.2	46.9	56.7
<1 MW	69.0	67.9	61.3	55.9	50.3	58.7	69.6	66.3	58.6	53.7	46.2	57.5
Total	71.0	67.4	62.4	60.0	45.8	54.4	71.3	66.9	61.0	56.6	44.8	55.0

Source: Source: SPADIES for observed data; fitted values for predicted data.
Notes: Values are expressed in percentages (%). Income is reported in brackets; MW = monthly minimum wage. Ability is reported in quintiles of standardized Saber 11 scores. Quintile 1 is the lowest.

predicted closely (14.5 percent). Further, the model captures the progressive concentration of students in the upper tiers over time (Figure 9).

The model replicates the wide variation in classes completed within abilities as well as the low variation across abilities (see Figure D1), and the patterns of class completion by year and ability as well as conditional on final college outcome (see Table D1.)[19] Finally, Figure D2 shows the average predicted target and average number of classes taken by the student. Recall that the latter is theoretically an upper bound for the former (see Section 3.2.) As such, our average predicted target is lower than the average observed number of classes. Overall, then, our model fits the data well and allows us to further investigate the role of effort in college outcomes.

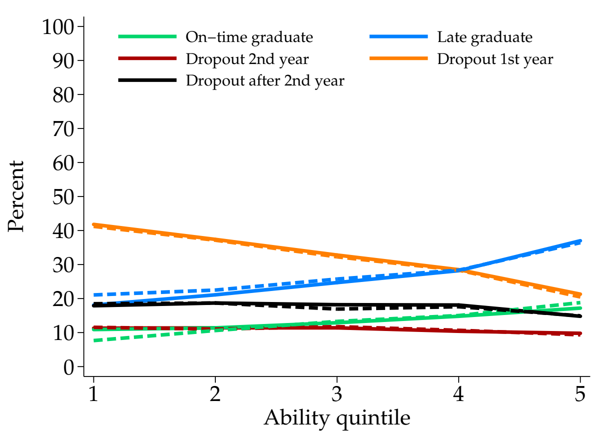
[19]When evaluating the fit of individual moments, note that, by using a weighting matrix in (16), moments with a greater number of underlying observations—such as those from early years, or for higher-ability students—attain a better fit.

Table 7: Goodness of Fit: Transitions Among Tiers of Cumulative Classes Completed.

	Observed values				Predicted values			
	Year 2	Year 3	Year 4	Year 5	Year 2	Year 3	Year 4	Year 5
Persistence								
Tier 1	64.33	79.78	86.98	79.89	63.70	75.80	75.10	74.30
Tier 2	28.67	39.23	51.47	57.76	22.80	38.40	53.70	60.10
Tier 3	34.94	49.44	60.94	69.04	35.10	59.20	71.20	78.80
Tier 4	29.67	45.75	57.87	58.59	37.10	42.50	60.70	76.00
Dropout rate								
Tier 1	12.16	4.03	1.75	0.71	11.60	4.20	3.00	2.20
Tier 2	14.43	6.50	3.25	1.52	16.20	6.40	4.60	3.00
Tier 3	23.02	12.30	7.29	4.02	25.10	9.20	7.00	5.20
Tier 4	54.23	41.95	28.24	24.24	55.20	50.70	35.20	21.40
Prob. of Catch up								
Tier 3 to Tiers 1 & 2	18.10	20.42	20.61	19.63	26.60	20.30	13.10	9.40
Tier 4 to Tiers 1 & 2	4.26	0.55	0.22	0.14	0.82	0.00	0.00	0.00
Prob. of Fall behind								
Tier 1 to Tiers 3 & 4	11.01	5.41	2.54	1.66	4.30	1.20	0.30	0.00
Tier 2 to Tiers 3 & 4	36.8	28.77	17.70	18.40	26.00	25.60	22.80	24.40

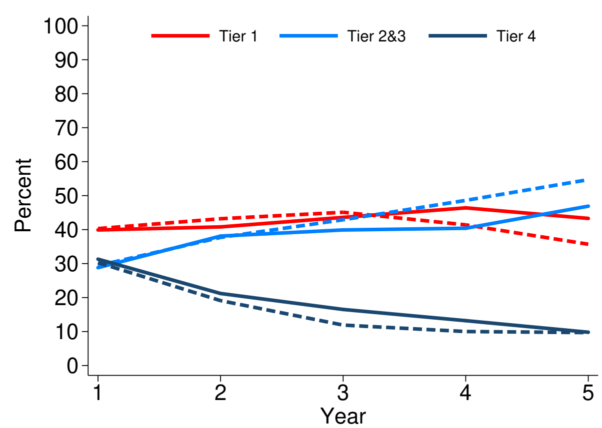
Source: SPADIES for observed data; model simulations for predicted data.
Notes: Every cell shows the observed and predicted percent of students of a given tier who persist in the same tier, drop out, catch up, or fall behind for years 1-5 (patterns are similar for years 6-8; not shown.)

Figure 8: Goodness of Fit: College Outcomes.



Source: SPADIES for observed data; fitted values for predicted data.
Notes: Whole lines show observed values; dashed lines show fitted values.

Figure 9: Goodness of Fit: Tiers of Cumulative Classes Completed, by Year.



Source: SPADIES for observed data; fitted values for predicted data.
Notes: For each year, the figure shows the percent of students by tier of cumulative classes completed. Whole and dashed lines show observed and predicted percentages, respectively.

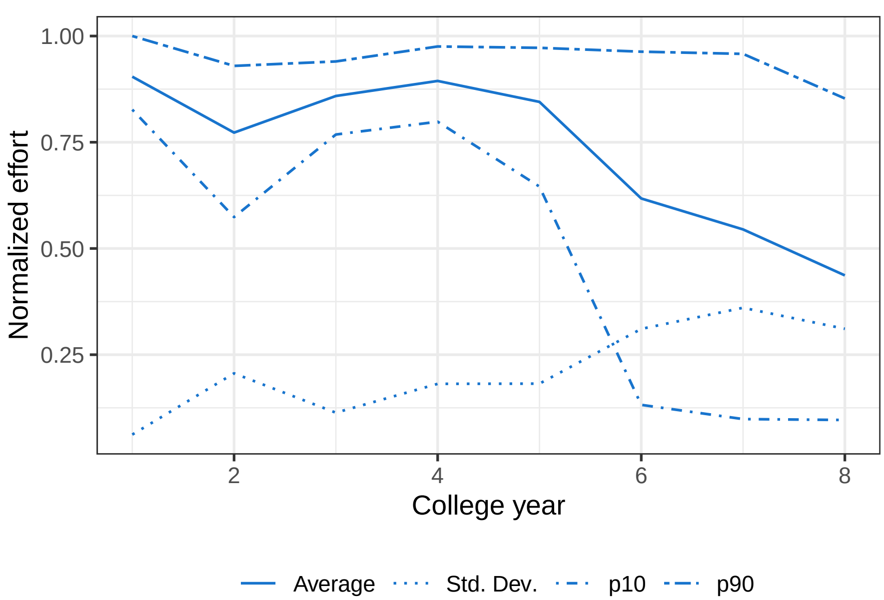
5 The Role of Effort and the Implications of Omitting it

In this section we investigate the role of effort in college outcomes. We begin by showing moments of the distribution of student effort in the baseline (i.e., the model evaluated at the parameter estimates). To highlight the role of effort, we compare our model with another one which does not include effort, and quantify the bias in production functions estimated on administrative datasets, which usually do not include effort measures. We conclude by investigating the relationship between effort and uncertainty.

5.1 Effort Across Students and Over Time

Our baseline model allows us to retrieve the (simulated) effort of every student in each college year, which in turn allows us to characterize effort over time and across students. Figure 10 shows some moments of the effort distribution by year, conditional on the student being enrolled. The solid line depicts average student effort. On average, this is relatively high in years 1 and 4. In year 1, students work hard to complete classes and mitigate the dropout risk, which is highest that year. In year 4, they increase effort in order to graduate. For the students that do not graduate on time, average effort drops after year 5 because they do not have, on average, many classes left. The solid dash lines show the 10th and 90th percentile of the effort distribution, and illustrate the overall effort variation among students, particularly after year 5. Similarly, the standard deviation of effort (bottom series in the chart) is somewhat limited in the first five years, but nearly doubles after the first batch of students graduates. While this figure includes all students, regardless of whether and when they graduate, patterns are similar for on-time graduates.

Figure 10: Predicted Effort by Academic Year



Source: Model baseline predictions.
Notes: The figure shows the following moments of the effort distribution by year: average, 10th percentile, 90th percentile, and standard deviation. Individual efforts are normalized by the 95th percentile of the year's effort distribution.

To further examine the variation of effort across students and over time, Table 8 uses simulated data from the model to show correlates of student effort based on reduced-form regressions. In the model, a student's optimal effort in a given year is the solution of the dynamic optimization problem in (6). It is a function of the student's ability, income, target number of classes, cumulative classes completed up to that year, expected shock to classes completed, and graduation probability. Thus, in Table 8 the dependent variable is log student effort. All regressions control for college year and student income. Column 1 controls for target number of classes, whereas columns 2, 3, and 4 control for variables that determine target and effort—past cumulative classes completed (proxied by average classes completed per year in past years), expected dropout probability, and expected shock to classes completed. Column 1 shows that higher ability students exert lower effort than their less able counterparts, whereas column 2, 3, and 4 show the opposite. The reason for the opposing results is that column 1 shows the effect of ability controlling for target, by which higher ability students need less effort for a given target, whereas the other columns show *total* ability effects—including those on target choice—by which higher ability individuals choose more demanding targets and therefore higher effort levels.

Table 8: Correlates of Effort Choice.

	(1)	(2)	(3)	(4)
ln(ability)	-0.015^{***}	0.076^{***}	0.107^{***}	0.067^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
ln(target)	0.787^{***}			
	(0.004)			
ln(average classes completed)		-0.173^{***}		
		(0.008)		
ln(expected dropout probability)			0.022^{***}	
			(0.000)	
ln(expected shock to classes completed)				-0.159^{***}
				(0.006)
Constant	-1.537^{***}	1.019^{***}	0.764^{***}	0.519^{***}
	(0.011)	(0.018)	(0.003)	(0.006)
Adj. R^2	0.766	0.285	0.557	0.286
Num. obs.	127,044	127,044	127,044	127,044

Source: OLS estimation using model's simulated baseline values.
Notes: The dependent variable is ln(effort), or $\ln(e_t^*)$ in the model. An observation is a student-year. Ability is θ, the target is q_t, and cumulative classes completed is $\bar{h}_{t-1}$. Expected shock to classes completed is $E[z_t]$; it varies across students and over time. Expected dropout probability is $E[\tilde{p}^d(t,\cdot)]$ as calculated by the individual prior to the realization of z_t. All regressions include year and income fixed effects (not shown). Standard errors (in parentheses) are clustered by student. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Column 2 shows that students with better past performance exert less effort. This is the outcome of a direct and indirect effects of past performance on effort, both moving in the same direction. Because of the direct effect, students with better past performance are closer to completing graduation requirements and need less effort. Because of the indirect effect, they expect better shocks to classes completed and a lower dropout probability, both of which make effort less necessary (columns 3 and 4). The effects of past performance illustrate the importance of a strong beginning: high-performing students in the early years can expect better "luck" in the future, which eliminates the need for a costly catch-up. Columns 3 and 4 show that students exert effort to compensate either for a high expected dropout probability or negative "luck" in class completion. As these elements vary across students and over time, students vary in their effort choices, giving rise to the effort variation captured by our simulated data.

5.2 What if the Model did not Include Effort?

One way of investigating the role of effort is estimating a model without effort as an input in the production function, and comparing it with our baseline model. To do so, we replace $H(z_t, \theta, e_t) = z_t(\theta^{\alpha} e_t^{1-\alpha})$ in (10) with $H(\tilde{z}_t, \theta)=\tilde{z}_t\theta^{\alpha}$. This is a standard technology in the literature, where academic progression is entirely determined by student ability, θ, and shocks, $\tilde{z}$. This formulation is a special case that results from setting the weight on effort in (10), $(1-\alpha)$, equal to zero.

To estimate the no-effort model, we postulate two flexible specifications that allow for different levels of heterogeneity in class completion among students. The first specification is similar to a general class of models that do not include effort at all, and assumes a homogeneous productivity for all students and years, which is calibrated to the overall average effort in the baseline. This yields $\tilde{z}_t = z_t\bar{e}^{1-\alpha}$. The second specification is akin to adding (year, student type) fixed effects to models with no effort, thereby assuming a heterogeneous productivity factor that varies across (student type, year) combinations. For a given (student type, year) combination, this factor is calibrated to the baseline average effort for that combination, or $\tilde{z}(t, \theta, y) = z_t\bar{e}(t, \theta, y)^{1-\alpha}$. In both specifications, the z_t shock still varies across individual students and over time.[20]

The parameter estimates and objective function values for these specifications and the baseline model are presented in Table 9. Not surprisingly, the specification with heterogeneous productivity factors fits the data better than that with homogeneous factors. However, both specifications fit the data worse than the baseline model, as the objective function value—which measures the distance between the observed and predicted data—rises 5-8 times relative to the baseline. In other words, the worsening of the data fit due to omitting effort from the model cannot cannot be solved by adding (student type, year) fixed effects to a no-effort model.

Removing effort from the production of classes completed places greater weight on ability and the performance shocks. The value for α increases to 0.21-0.26, yielding an elasticity of credits completed with respect to ability that is 2.5-3 times larger than the baseline's. Since higher ability students have lower effort costs and exert greater effort, the no-effort model fully attributes to ability an effect that is mediated, at least partly, by effort. As for the performance shock, in the no-effort model the implied $E(z)$ rises steeply with ability, thereby amplifying ability's role, and z becomes more dispersed in an attempt to match within-ability performance variation (see Figure 11.) Moreover, the estimate for the risk aversion parameter, ρ, becomes twice as large to reflect that students, who can no longer mitigate risk through effort, behave as if they were more risk averse.

As we saw above, the no-effort models provide a worse fit of the data than the baseline model. Although they predict the overall graduation rate quite well (at 45-47 percent), relative to the baseline model they over predict the variation of college outcomes across abilities. As

[20] In ((11)), the ν_t values for every student and year are the same as in the baseline model.

Table 9: The Role of Effort: Parameter Estimates with and without Effort.

Parameter	Symbol	Baseline With Effort	No Effort 1. Homog. Prod.	No Effort 2. Heter. Prod.
Utility function				
Consumption curvature	ρ	0.882	1.722	1.702
Effort weight	μ	0.062		
Effort curvature	γ	4.727		
Effort cost w.r.t. ability	k	1.225		
Number of classes completed				
Elasticity w.r.t. ability	α	0.085	0.260	0.206
Performance shock				
Constant	κ_0	-4.207	-4.440	-4.478
Year 1 shifter	κ_1	3.534	2.765	2.781
Persistence component	κ_h	1.304	1.206	1.257
Ability component	κ_θ	0.407	1.473	1.395
Std. dev. of *iid* shock	σ	1.789	3.196	3.016
Std. dev. of *iid* shock - Year 1 shifter	σ_1	0.317	0.450	0.178
Std. dev. of *iid* shock - Ability shifter	σ_θ	-1.282	-0.005	-0.045
Dropout shock				
Cumulative performance component	π	-2.951	-2.557	-2.644
Objective function value		8.929	69.524	44.019

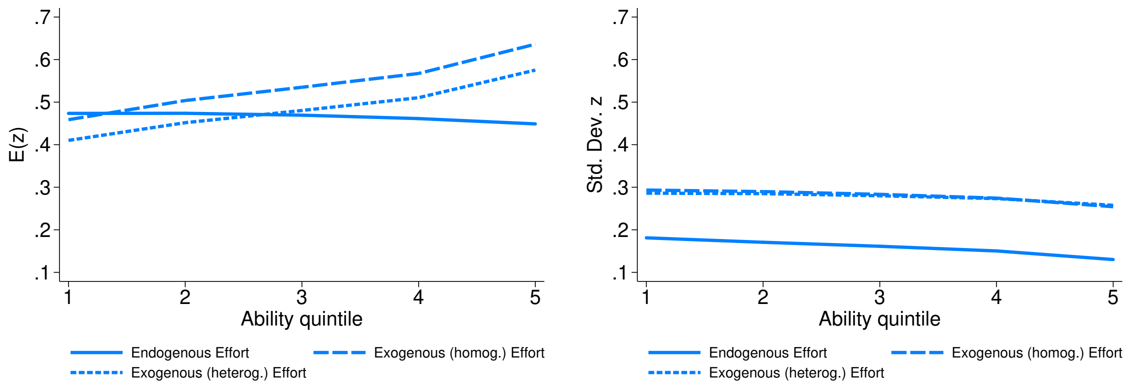
Source: Own estimation. The baseline model includes effort. "Homog. Prod." and "Heter. Prod." denote the no-effort specifications with homogeneous and heterogeneous productivity factors, respectively. See the text for a description of these specifications.

shown by Figure 12's panel a, they make the dropout rate fall more steeply with ability and the on-time graduation rate rise much faster than the baseline model. This is a consequence of over estimating the role of ability in class completion. Similarly, the no-effort models over predict performance variation across abilities but under predict it within abilities (compare Figure 12's panel b with Figure 5's panel a). As a result, the no-effort models over predict the academic performance and on-time graduation of high-ability students, and under predict it for low-ability students.[21]

Modeling effort, then, allows us to match the fact that some low-ability students perform better (due to effort) than predicted by their ability while some high-ability students perform worse (due to lack of effort) than predicted by their ability. The no-effort specifications are able to replicate the observed aggregate graduation rate well, but not the variation of outcomes across student types and over time. This has serious distributional implications—in order to raise the achievement of low-ability or low-income students, modeling it correctly is a critical first step. It also has stark policy implications. When only ability and luck matter to performance, selection might be the policymaker's only tool to raise the fraction of college students—subsidizing, for instance, the highest-ability students. When effort matters, in contrast, a new tool appears

[21]Due to space limitations, in this figure we compare the baseline model to the no-effort specification with heterogeneous productivities, but the conclusions hold for the other no-effort specification as well. Since the baseline model fits well the distribution of students across performance tiers, we compare this prediction from the no-effort model directly to the observed data. Similar patterns arise when comparing predictions for subsequent college years.

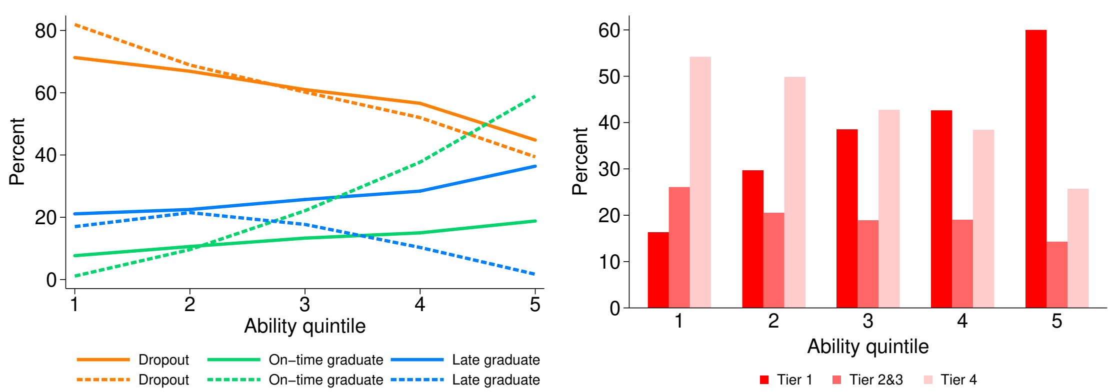
Figure 11: The Performance Shock in the Baseline and No-Effort Specifications.



a. Average performance shock   b. Standard deviation of performance shock

Source: Own estimations.
Notes: For each quintile, panel a and b show the implied *E(z)* and standard deviation of z respectively for the three effort models: baseline ("endogenous effort"), no-effort with homogenous productivities ("exogenous (homog.) effort") and no-effort with heterogeneous productivities ("exogenous (heterog.) effort."

Figure 12: Effort and Model Predictions.



a. College Outcomes   b. Tiers of Cumulative Classes Completed (Year 1)

Source: Own estimations.
Notes: In panel a, for each ability quintile the graph shows the predicted percent of students who attain each outcome. Solid line indicates baseline model and dotted line indicates no-effort model (with heterogeneous productivities). In panel b, for each ability quintile, the figure shows the predicted distribution into performance tiers by the end of year 1 for the no-effort model (with heterogeneous productivities.)

for the policymaker—affecting student behavior—and a new set of subsidies, based on student performance rather than income or ability, become relevant.

5.3 Estimations Using Administrative Data: The Omitted Effort Bias

Despite the importance of effort in performance, in general administrative data sets do not provide effort measures. Our insights from the no-effort models indicate that, when these data sets are used to estimate the relationship between classes completed and ability through linear regressions, lack of effort measures might bias the estimates upwards. One advantage of our structural model is that it recovers the variables—effort and shocks—not observed in the administrative data, allowing us to measure the size and direction of the omitted effort bias.

In Table 10, the dependent variable is log classes completed, $\ln(x_t)$. We regress it on log ability using observed data in column 1 and simulated data in columns 2-4. The coefficients on log ability in columns 1 and 2 are very close, as expected from our good fit. In both cases, log ability explains about 20 percent of the variation in log classes completed.

Table 10: Classes Completed Per Year.

	Actual data	Simulated Data		
	(1)	(2)	(3)	(4)
ln(ability)	0.166***	0.156***	0.090***	0.085***
	(0.015)	(0.005)	(0.005)	(0.000)
ln(effort)			0.854***	0.915***
			(0.004)	(0.000)
ln(shock to classes completed)				1.000***
				(0.000)
Constant	2.060***	2.748***	2.197***	2.996***
	(0.012)	(0.004)	(0.004)	(0.000)
R^2	0.213	0.204	0.518	1.000
Num. Obs.	123,101	127,044	127,044	127,044

Source: OLS estimation using SPADIES for actual data and model's baseline predictions for simulated data. *Notes*: Dependent variable is ln(classes completed per year), or $\ln(x_t)$ in the model. An observation is a student-year; years 1-8 are included. Ability is θ, effort is e_t^*, and shock to classes completed is z_t. All regressions include year fixed effects (not shown). Standard errors (in parentheses) are clustered by student. * p < 0.10, ** p < 0.05, *** p < 0.01.

The simulated data allows us to expand column 2's specification in order to gauge the relative roles of effort, performance shocks, and ability. By adding effort (column 3), we are able to explain an additional 30 percent of the variation in classes completed. As expected, we explain the full variation in classes completed when we also add the performance shock (column 4), with a coefficient on log ability equal to our point estimate for α. Further, in column 4 effort and the shock account for 47 and 52 percent of the variation in classes completed, respectively, leaving a mere 1 percent explained by log ability. In other words, in our simulated data almost the whole variation in classes completed is due to effort and performance shocks.

Consistent with our estimates from the no-effort specifications, the regressions that do not control for effort (in columns 1 and 2) overestimate the coefficient on log ability almost by a factor of 2. In other words, regressions using administrative datasets, which typically lack effort measures, are likely to overestimate the role of ability in academic performance. Policies based on these biased estimates might place too much weight on a fixed student trait—ability—rather than a student choice—effort—that could, in principle, be affected through policy.

5.4 Effort and Uncertainty

The typical rationale for tuition subsidies is that students face financial constraints which prevent them from undertaking an investment with positive net present value. Our model highlights another, novel rationale: college students face *investment uncertainty*. They are uncertain about their ability to complete classes and remain enrolled in college due to shocks that might be particularly important in developing economies, where a family health shock or job loss often force students to leave college.

In light of this uncertainty, an important question is whether student effort can mitigate it. We define a measure of anticipated uncertainty and assess whether effort reduces it. Our measure of anticipated uncertainty for student i at the beginning of period t is the coefficient of variation of her college payoffs:

$$CV_{it} = \frac{\sqrt{\text{Var}_z\left[\tilde{V}^{coll}(t, h_{it-1}, \theta, y; z_{it}, e^*_{it})\right]}}{E_z\left[\tilde{V}^{coll}(t, h_{it-1}, \theta, y; z_{it}, e^*_{it})\right]}, \tag{17}$$

where the right-hand side is the ratio between the standard deviation and expected value of college payoffs, evaluated at the student's optimal effort, e^*_{it}. Given e^*_{it}, the randomness in $\tilde{V}^{coll}(\cdot)$ comes from the performance and dropout shocks, z and d^{drop} respectively, associated with effort via cumulative classes completed. Anticipated uncertainty varies across students and college years and, most importantly, depends on student optimal effort.[22]

We compute this measure using our simulated baseline data, and investigate the relationship between effort and anticipated uncertainty. A simple regression of anticipated uncertainty on optimal effort would suffer from endogeneity because optimal effort would appear in both sides of the regression. Our model, however, provides an instrument for effort. Since e^*_{it} is a function of the state variables, t, h_{it-1}, θ_i, and y_i, we instrument for effort using the state

[22]In ((6)), the value function $V^{coll}(t, h_{t-1}, \theta, y)$ is calculated for the student's optimal effort, e^*_t, while taking expectation over z_t. We can also define the college payoff from *any* effort and realization of z_t as

$$\begin{aligned}\tilde{V}^{coll}(t, h_{t-1}, \theta, y; z_t, e_t) = U(c_t, e_t, \theta) + \beta\Big[&\mathbf{1}_{\{t\geq 5\}}\Pr\left(h_{t-1} + H(z_t, \theta, e_t)\,\bar{x} \geq h^{grad}\right)V^{grad}(t+1) + \\ &\Pr\left(h_{t-1} + H(z_t, \theta, e_t)\,\bar{x} < h^{grad}\right)\Big(\tilde{p}^d(t, h_{t-1} + H(z_t, \theta, e_t)\,\bar{x}, \theta, y)V^{drop}(t+1) + \\ &\left(1 - \tilde{p}^d(t, h_{t-1} + H(z_t, \theta, e_t)\,\bar{x}, \theta, y)\right)V^{coll}(t+1, h_{t-1} + H(z_t, \theta, e_t)\,\bar{x}, \theta, y)\Big)\Big].\end{aligned} \tag{18}$$

It then follows that the denominator of ((17)), $E_z[\tilde{V}^{coll}(t, h_{t-1}, \theta, y; z_t, e^*_t)]$, is indeed the value function, $V^{coll}(t, h_{t-1}, \theta, y)$.

variables (see Table E1 for first-stage results), and run the second-stage regression reported in Table 11. Estimates indicate that, in our baseline model, effort indeed lowers anticipated uncertainty, with a negative and large elasticity of -2.34. Student effort, therefore, is highly effective at mitigating performance and dropout risks. In other words, students can mitigate their investment uncertainty by exerting greater effort. By incentivizing student effort, tuition subsidies have the potential of lowering students' investment uncertainty.

Table 11: Anticipated Uncertainty.

	Dependent variable ln(Anticipated risk)
ln(effort)	-2.340^{***}
	(0.089)
ln(average classes completed)	-1.767^{***}
	(0.043)
Constant	-0.772^{***}
	(0.068)
Num. obs.	116,761

Source: 2SLS estimation using model's simulated baseline values.
Notes: The dependent variable is ln(anticipated risk), or $\ln(CV_{it})$ for the baseline. An observation is a student-year. Upper 5% tail of uncertainty has been trimmed. Effort is e^*_{it}, instrumented with the following variables: classes completed, year, income, and ability fixed effects. ln(average classes completed) is $\tilde{h}_{it-1}$, as defined in the model. The regression includes year fixed effects (not shown). Standard errors (in parentheses) are clustered by student. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusions

Just as cows do not give milk, neither do colleges give degrees. Without student effort, class completion and college graduation are simply not possible. In this paper we have developed and estimated a dynamic model of college enrollment, performance, and graduation. A central piece of the model, student effort, has a direct effect on the completion of classes and an indirect effect mitigating risks on class completion or college persistence. We have estimated the model using rich administrative data from Colombia. According to our estimates, effort has a much greater impact than ability on class completion. Failing to model effort as an input to class completion leads to overestimating the role of ability by a factor or two or three. In terms of policy, it leads to subsidizing college for low-income or high-ability students rather than based on student performance, which is highly dependent on student effort.

Similar to other investments, college is risky. Our model highlights that effort provides an insurance against performance and persistence risks. As a result, policies that incentivize effort not only promote better student performance; they also promote an insurance against the inherent uncertainty of this investment.

A critical message from our paper is that, given the quantitative importance of effort to college outcomes, tuition subsidies that do not promote effort will not yield the human capital accumulation or intergenerational mobility that disadvantaged students and their families long for. Along these lines, Ferreyra et al. (2022) apply the model developed in this paper to study alternative free college policies, including universal, targeted (by income or ability), and performance-based free college. At this time of severe fiscal constraints for many countries yet heightened expectations of individual and social progress, the time is ripe to build human capital policies based on the wisdom that cows do not give milk; you need to milk them.

References

[1] Ahn, T., Arcidiacono, P., Hopson, A., & Thomas, J. R. (2019). Equilibrium grade inflation with implications for female interest in stem majors. NBER Working Paper No. w26556.

[2] Akyol, A. & Athreya, K. (2005). Risky higher education and subsidies. *Journal of Economic Dynamics and Control, 29(6)*, pp. 979-1023.

[3] Altonji, J. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics, 11(1),* pp. 48-83.

[4] Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics, 121(1-2),* pp.343-375

[5] Arcidiacono, P., E. Aucejo, A. Maurel, & T. Ransom (2016). College attrition and the dynamics of information revelation. NBER Working Paper No. 22325.

[6] Ariely, D., U. Gneezy, G. Loewenstein, & N. Mazar (2009). Large stakes and big Mistakes. *Review of Economic Studies, 76(2),* pp. 451-469.

[7] Beneito, P., J. Bosca & J. Ferri (2018). Tuition fees and student effort at university. *Economics of Education Review, 64(C),* pp. 114-128.

[8] Berry, S., J. Levinsohn, A. Pakes (1995). Automobile prices in market equilibrium, *Econometrica,* pp: 841–890

[9] Cameron S.V. & Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *The Journal of Political Economy, 106(2),* pp. 262-333.

[10] Cameron S.V. & Heckman, J. J. (1999).The dynamics of educational attainment for blacks, hispanics and whites (Working Paper 7249). Cambridge, MA: National Bureau of Economic Research.

[11] Carneiro, P. & Heckman, J. J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal, 112(482),* pp. 705-734.

[12] Dynarski, S., & Scott-Clayton, J. (2013). Financial aid policy: Lessons from research (No. w18710). National Bureau of Economic Research.

[13] Eckstein, Z., & Wolpin, K. I. (1998). Youth employment and academic performance in high school. CEPR Discussion Paper Series No. 1861

[14] Ferreyra, M., Avitabile, C., Botero Alvarez, J., Haimovich Paz, F., & Urzua, S. (2017). *At a crossroads: higher education in Latin America and the Caribbean.* The World Bank.

[15] Ferreyra M., Garriga, C., Martin-Ocampo, J. & Sanchez-Diaz, A. (2022). The limited impact of free college policies. [Unpublished manuscript]

[16] Garriga, C. & M.P. Keightley (2007). A general equilibrium theory of college with education subsidies, in-school labor supply, and borrowing constraints. Federal Reserve Bank of St. Louis Working Paper 2007-051A.

[17] Hendricks, L. & Leukhina, O. (2017). How risky is college investment? *Review of Economic Dynamics, 26,* pp. 140-163.

[18] Hendricks, L., & Leukhina, O. (2018). The return to college: selection and dropout risk. *International Economic Review, 59(3),* pp. 1077-1102.

[19] ICETEX (2010). Informe de gestión 2003-2010. Instituto Colombiano de Crédito Educativo y Estudios Técnicos en el Exterior (ICETEX).

[20] Keane, M. P. & Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review, 42(4),* pp. 1051-1103.

[21] Keane, M. P. (2002). Financial aid, borrowing constraints, and college attendance: Evidence from structural estimates. *The American Economic Review*, *92(2)*, pp. 293-297.
[22] Levhari, D. & Weiss, Y (1974). The effect of risk on the investment in human capital. *The American Economic Review, 64(6),* pp. 950-963.
[23] Lochner, L.J. & A. Monge-Naranjo (2011). The nature of credit constraints and human capital. *American Economic Review, 101(6),* pp. 2487-2529.
[24] Londoño, J., Rodriguez, C., & Sanchez, F. (2020). Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser Pilo Paga in Colombia. *American Economic Journal: Economic Policy, 12(2),* pp. 193-227.
[25] Londoño, J., Rodriguez, C., Sanchez, F. & Alvarez L. (2022). Financial Aid and Social Mobility: Evidence from Colombia's *Ser Pilo Paga*. [Unpublished manuscript]
[26] Mehta, N., R. Stinebrickner, and T. Stinebrickner. 2019. "Time-Use and Academic Peer Effects in College." Economic Inquiry 57, no. 1: 162-171.
[27] Ozdagli, A. K., & Trachter, N. (2011). On the distribution of college dropouts: Household wealth and uninsurable idiosyncratic risk. Working Papers No. 11-8.
[28] Solis, A. (2017). Credit access and college enrollment. *Journal of Political Economy, 125(2),* pp. 562-622.
[29] Stange, K.M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics 4 (1),* pp. 49-84
[30] Stinebrickner, R. & T. R. Stinebrickner (2004). Time-Use and college outcomes. *Journal of Econometrics, 121(1-2),* pp 243-269.
[31] Stinebrickner, R., and T. Stinebrickner. 2008. "The causal effect of studying on academic performance." The BE Journal of Economic Analysis Policy 8, no. 1.
[32] Stinebrickner, R. & T. R. Stinebrickner (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics, 32(3),* pp. 601 - 644.
[33] Trachter, N. (2015). Stepping stone and option value in a model of post-secondary education. *Quantitative Economics, 6(1),* pp. 223-256.
[34] Zamarro, G., C. Hitt, & I. Mendez (2019). When students don't care: Reexamining international differences in achievement and student Effort. *Journal of Human Capital, 13(2),* pp. 519-552.

For Online Publication: Appendices

Appendix

A Data

A.1 Students and Graduation Requirements

We use data from Colombia in the early 2000s. At that time, 80 percent of higher education students attended bachelor's programs (lasting four to six years), while the remaining 20 percent attended two- or three-year programs (akin to associate's degree programs in the U.S). For several decades now, Colombia has implemented a mandatory high school exit exam, Saber 11. This is a standardized test covering multiple academic fields that measures academic readiness for higher education. We average scores over fields and standardize the average by semester-year.

To study academic progression and college outcomes, it is customary in the literature to focus on students from the same college entry cohort who study programs of the same duration. We focus on the 2006 college entry cohort of students who pursue five-year programs. These students, however, may have finished high school at different times. This is because, in Colombia in the early 2000s, only 10 percent of high school graduates would enroll in college right away but about a third would do it within a five-year window. Using this window therefore provides a more accurate enrollment rate. Since the composition of high school cohorts is relatively stable during the early 2000s, we view the 2005 cohort of high school graduates as representative of cohorts from the 2000-2005 period, and use it to study college enrollment decisions. We therefore use data from the 2005 high school cohort to study college enrollment decisions, and from the 2006 college entry cohort to study college academic progression and outcomes.

The 2005 high school cohort comprises 415,269 students ages 15-22 who took Saber 11 in 2005. We classify a student as having enrolled in college if she did so between 2006 and 2010. The 2006 college entry cohort of students pursuing five-year programs includes 27,344 students. Five-year programs capture about three-quarters of enrollment in bachelor's programs. Dropout rates and academic progression correspond to the student's first program in higher education. Some students drop out in the first semester, without completing the corresponding academic year. Since a period in our model is a school year rather than a semester, for those students we impute a number of classes completed in the dropout year equal to twice their number of classes completed in their last semester. This imputation is reasonable because, in their dropout year, second-semester dropouts complete approximately twice as many classes, on average, as first-semester dropouts.

Since we do not observe the number of classes that the student in a given program must complete in order to graduate, we assume it is equal to the average number of classes completed by the program's graduates, which we do observe. This number varies across programs. To facilitate calculations and exposition, we normalize the total number of classes required by a program to 100. This allows us to describe academic progress in terms of the cumulative number of classes completed by the student at given time. This is, of course, equivalent to describing academic progress in terms of the *percent* of cumulative classes completed relative to the number of required classes. For example, if a program's graduation requirement is 50 classes, completing

10 classes is equivalent to completing 20 percent of the graduation requirements and, in our normalization, equivalent to completing 20 classes. Lacking data on how many classes students are supposed to complete *per year*, we assume it is the same (20 classes) across the five years. This assumption is supported by the data, which shows that students who graduate on time complete classes at a fairly uniform pace.

A.2 Tables

Table A1: Family Income and Ability Distribution of High School Graduates.

Income	Ability quintile					
Bracket	1	2	3	4	5	Total
5+ MW	0.21	0.31	0.48	0.90	3.15	5.05
3-5 MW	0.88	1.08	1.37	1.94	3.43	8.69
2-3 MW	2.72	2.94	3.30	3.69	3.95	16.60
1-2 MW	8.47	8.99	9.16	8.69	6.64	41.95
<1 MW	7.95	6.89	5.80	4.58	2.49	27.71
Total	20.23	20.21	20.11	19.80	19.65	100.00

Source: Calculations based on Saber 11. The distribution refers to 415,269 high school graduates from 2005.
Notes: Family income is reported in brackets; MW = monthly minimum wage. Ability is reported in quintiles of standardized Saber 11 scores. Quintile 1 is the lowest.

Table A2: Cumulative Performance Tiers: Lower Bound by Year.

Tier	Year 1	Year 2	Year 3	Year 4	Year 5
Tier 1	19	38	57	76	95
Tier 2	17	34	51	68	85
Tier 3	13	26	39	52	65
Tier 4	0	0	0	0	0
Required cumulative classes	20	40	60	80	100

Source: Own classification.
Notes: This table shows the lower bound number of cumulative classes completed for each tier, by year. For a given year, tiers are defined relative to the required number of cumulative classes completed (equal to 20, 40, 60, 80, and 100 in years 1 through 5, respectively). Tier 1: 95 percent of the requirement or more; Tier 2: (85, 95] percent; Tier 3: (65, 85] percent; Tier 4: 65 percent or less. For example, in year 2 (when 40 cumulative classes are required), the tier lower bounds (expressed in cumulative classes completed) are as follows: 38 classes = 0.95×40 for Tier 1; 34 classes = 0.85×40 for Tier 2; and 26 classes = 0.65×40 for Tier 3.

Table A3: Family Income and Tuition.

Income Bracket	Avg. Per-Capita Household Income (y)	Avg. Per-Capita Tuition (T)
5+ MW	21,027,690	2,195,972
3-5 MW	9,191,642	1,826,386
2-3 MW	5,337,010	1,177,543
1-2 MW	2,952,288	978,690
<1 MW	1,119,633	855,493

Source: Calculations based on Saber 11 and SEDLAC (household surveys) for average per-capita household income; SNIES and SPADIES for annual tuition. Income and tuition are expressed in Colombian pesos (COP) of 2005.
Notes: Since Saber 11 provides income brackets rather than actual income, we use SEDLAC (household surveys) data on household income and size to calculate average per-capita income for households in a given bracket; this is our proxy for parental transfer or *income* (y). To calculate average tuition by bracket, we average over the tuitions paid by student from the corresponding income bracket at public institutions. MW = monthly minimum wage. Both income (y) and tuition (T) are expressed in annual terms.

Table A4: Average Hourly Wage by Age Bracket and Educational Attainment.

	Age bracket			
	18-60	18-22	23-35	36-60
College graduates	6,308	3,636	5,305	7,171
HS graduates	2,424	1,845	2,213	2,864
College dropouts; completed 1 year or less	3,091	2,349	2,897	3,619
College dropouts, completed 2 years or more	3,824	2,459	3,340	4,451

Source: Household surveys for Colombia (SEDLAC); year 2005.
Notes: Wages are expressed in Colombian pesos (COP) of 2005. Calculations include males and females who work. Attainment reflects an individual's highest completed level of schooling.

B Model timeline

The figure below summarizes the timing of events and student decisions:

Figure B1: Summary of Timing of Events and Individuals' Decisions

Enrollment Decision $(t = 0)$
- College
- Labor force

College $(t = 1{:}4)$	State	i. Choice	ii. Classes completed shock	iii. Accumulated classes completed	iv. Dropout shock	v. Payoffs
	(t, h_{t-1}, θ, y)	e_t	z_t	h_t	d_t^{drop}	$V^{drop}(t+1)$ $V^{coll}(t+1, h_t, \theta, y)$

College $(t = 5{:}7)$	State	i. Choice	ii. Classes completed shock	iii. Accumulated classes completed	iv. Dropout shock	v. Payoffs
	(t, h_{t-1}, θ, y)	e_t	z_t	h_t	d_t^{drop}	$V^{drop}(t+1)$ $V^{coll}(t+1, h_t, \theta, y)$ $V^{grad}(t+1)$

College $(t = 8)$	State	i. Choice	ii. Classes completed shock	iii. Accumulated classes completed	iv. Dropout shock	v. Payoffs
	$(8, h_7, \theta, y)$	e_8	z_8	h_8	d_8^{drop}	$V^{drop}(9)$ $V^{grad}(9)$

C Computation and Estimation

In this appendix we first summarize the computational solution of the model, and then provide detail on its main steps.

C.1 Solving the Model: A Summary

Recall that the state vector is (t, h_{t-1}, θ, y). We discretize the state space for a total of 40,400 points. We simulate N = 100,000 high school graduates from the empirical distribution of ability and income (or parental transfer), $\Phi(\theta, y)$. For each simulated high school graduate, we draw one *i.i.d.* shock per year, $\{\nu_{it}\}_{t=1}^{8}$. These shocks enter the performance shock, z (see equation (11) in the main text). For a given simulated high school graduate, these shocks are the same across parameter vectors during estimation.

Among the simulated high school graduates of a given type, a fraction of them receives a college enrollment shock equal to 1 and enrolls in college (thus becoming the "simulated college students"); the fraction is equal to the type's observed college enrollment rate.

To compute the model's predictions for a given value of Θ, the algorithm proceeds as follows:

1. For each point in the state space, use backward induction to solve for the sequence of optimal efforts (the policy function) and the value function, $e^*(t, h_{t-1}, \theta, y)$ and $V^{coll}(t, h_{t-1}, \theta, y)$ respectively.

2. For each simulated college student, and for every year she is enrolled, combine her optimal effort with the corresponding ν_t shock to determine the performance shock and the probability of dropping out, $\tilde{p}^d(t, h_t, \theta, y)$. Draw the binary dropout shock; the shock is equal to 1 with probability $\tilde{p}^d(t, h_t, \theta, y)$.

3. Based on step 2, aggregate the simulated dropout decisions to obtain a predicted dropout rate for each of the 400 (t, θ, y)-combinations.

4. Find the vector $\boldsymbol{\delta}$ that minimizes the distance between the predicted and observed dropout rate for each (t, θ, y)combination, using the contraction mapping algorithm described in Appendix C.3.

5. By comparing the value of going and not going to college for each type, $V^{coll}(1, 0, \theta, y)$ and V^{hs}, respectively, find the type-specific college enrollment shock, ξ_j, that renders the type indifferent between going and not going to college. Further details are provided in Appendix C.4.

Solving steps 1-5 of the dynamic optimization problem for $100,000$ simulated high school graduates and $40,400$ states takes approximately 8 minutes in a 1.4 GHz Intel Core i5 processor. Since the model does not have a closed-form solution, in estimation the problem must be solved anew for each value of Θ. The estimation of $\boldsymbol{\delta}$ and $\boldsymbol{\xi}$ is nested within the model solution for a given of value of Θ, in the spirit of Berry et al. (1995), as described in Appendix C.3.

C.2 Further Details on Model Solution

The solution of the dynamic problem for a given value of Θ involve three steps: calculating the value of working by educational attainment, solving for the policy and value functions, and simulating college students.

Calculating the Value of Working by Educational Attainment Since we solve the student's dynamic programming problem by backward induction, we begin by calculating the final value of the individuals' finite horizon problem. We calculate the value of the future discounted payoffs of working as a college graduate, high school graduate, college dropout with one year of college, and college dropout with two or more years of college. In the timing of the model, $t = 1$ when the individual is 18 years old, at which point she either starts college or joins the labor force as a high school graduate. The value of working as a high school graduate since $t = 1$ onward is

$$V^{hs} = \sum_{t=1}^{L} \beta^{t-1} u(w_t^{hs}), \tag{19}$$

where w_t^{hs} is the average wage for a high school graduate in year t, and L is retirement age (65 years old, or $L = 48$). For this and the other educational attainments, we allow the wage to vary over time to incorporate returns to experience (which accrue after age 35, that is for $t > 17$). Similarly, the value of working as a college dropout who has completed n years of college is

$$V^{drop}(n+1) = \sum_{t=n+1}^{L} \beta^{t-n-1} u(w_t^{drop}), \tag{20}$$

where w_t^{drop} is the wage received by this individual in year t. Finally, the value of working as a college graduate who took n years to graduate is:

$$V^{grad}(n+1) = \sum_{t=n+1}^{L} \beta^{t-n-1} u(w_t^{grad}), \tag{21}$$

Solving for the Policy and Value Functions Since the state vector, (t, h_{t-1}, θ, y), has four elements, we build two-four dimensional grids—one for the policy function, $e^*(t, h_{t-1}, \theta, y)$, which contains the optimal effort choice by state, and another for the optimal payoffs, $V^{coll}(t, h_{t-1}, \theta, y)$, associated with $e^*(t, h_{t-1}, \theta, y)$. The grid includes 8 points (years) for t, 101 points for h (to represent $0, 1, 2, \ldots, 100$ credits completed), ten points (ability deciles) for θ, and five points for y, for a total of 40,400 points.

We use backward induction to solve the Bellman equation for each period. Starting from the last year of college, $t = 8$, when students must either graduate or drop out, the Bellman equation is:

$$V^{coll}(8, h_7, \theta, y) = U(c_8, e_8, \theta) + \beta E_z \Big[\Pr\left(h_8 \geq h^{grad}\right) V^{grad}(9) + \Pr\left(h_8 < h^{grad}\right) V^{drop}(9)\Big]. \tag{22}$$

By choosing e_8 to maximize V^{coll} for every state $(8, h_7, \theta, y)$, we solve for the policy and value functions, $e^*(8, h_7, \theta, y)$ and $V^{coll}(8, h_7, \theta, y)$.

Moving backwards to $t = 7$, we proceed analogously:

$$\begin{aligned} V^{coll}(7, h_6, \theta, y) = U(c_7, e_7, \theta) \ + \beta E_z \Bigg[& \Pr\left(h_7 \geq h^{grad}\right) V^{grad}(8) + \\ & \Pr\left(h_7 < h^{grad}\right) \Big(\tilde{p}^d(7, h_7, \theta, y) V^{drop}(8) + \\ & \left(1 - \tilde{p}^d(7, h_7, \theta, y)\right) V^{coll}(8, h_7, \theta, y) \Big) \Bigg], \end{aligned} \tag{23}$$

where the year-8 payoffs, $V^{grad}(8)$, $V^{drop}(8)$ and $V^{grad}(8, \cdot)$, are already known. We continue this procedure for $t = 6, \ldots, 1$ in order to complete the calculation of $e^*(t, h_{t-1}, \theta, y)$ and $V^{coll}(t, h_{t-1}, \theta, y)$ for all possible states.

We use the resulting policy function, $e^*(t, h_{t-1}, \theta, y)$, whenever we need to retrieve a student's optimal effort during estimation or baseline calculations. In addition, we use the resulting value function at $t = 1$, that is, $V^{coll}(1, 0, \theta, y)$, to calculate the value of enrolling in college. This is to be compared with the value of joining the workforce as a high school graduate, V^{hs}, following equation (8).

Simulating College Students Recall that a student type j is given by a (θ_j, y_j) combination. We have $J = 50$ types. For each type, let $P^{coll}(\theta_j, y_j)$, equal to the actual, observed share of individuals *of that type* that enrolls in college. Note that $P^{coll}(\theta_j, y_j)$ varies across types, as illustrated by Table 1. Consider individual i who belongs to type j. For each simulated individual, we draw a binary variable, d_i^{enr}, to determine whether the student goes to college or not. More specifically,

$$d_i^{enr} = \begin{cases} 1, & i \text{ goes to college, with probability } P^{coll}(\theta_j, y_j) \\ 0, & i \text{ does not goes to college, with probability } 1 - P^{coll}(\theta_j, y_j) \end{cases} \tag{24}$$

Simulated students who receive $d_i^{enr} = 1$ are those who enroll in college. In other words, the proportion of simulated students of a given type who receive $d_i^{enr} = 1$ is the same as the proportion of actual students of that type who enroll in college.[23] For students who do not enroll in college, the value function is V^{hs}. For those who enroll in college, we simulate classes completed and dropout shocks as described below.

For $t = 1$, we use the policy function, $e^*(1, 0, \theta_j, y_j)$, corresponding to every student type j. Since all students start at $t = 1$ with zero accumulated credits, $h_0 = 0$, the policy function assigns the same effort to all students of a given type j. Then, we draw the *iid* shock ν_{i1} for each student; this, in turn, yields a value for the z_{i1} shock. The combination of the student's ability, effort, and z_{i1} shock yields the number of completed credits by the end of the year, h_{i1}. Because of the z shock, individuals of a given type may attain different values of h.

For student i, we use the realized h_{it} to establish whether the student drops out before the second period. The student receives a draw of the binary variable d_{it}^{drop}; if the draw is equal to 1, she drops out. The probability that $d_{it}^{drop} = 1$ is a function of student type, year, and average

[23]For a large number of simulations such as ours, this is asymptotically equivalent to simply assigning $d_i^{enr} = 1$ to a fraction of simulated students of a given type equal to the type's observed enrollment rate.

performance up to (and including) the corresponding year:

$$d_{it}^{drop} = \begin{cases} 1, & i \text{ drops out of college, with probability } \tilde{p}^d(t, h_{it}, \theta_j, y_j) \\ 0, & i \text{ continues in college, with probability } 1 - \tilde{p}^d(t, h_{it}, \theta_j, y_j) \end{cases} \tag{25}$$

where $\tilde{p}^d$ is defined as in (12) in the main text.

Another binary variable, d_{it}^{grad}, indicates whether a student graduates. The graduation requirement is $h^{grad} = 98$ rather than 100 because some students in the data graduate with fewer than 100 classes. Whenever $t \geq 5$ and $h_{it} \geq h^{grad}$, we set $d_{it}^{grad} = 1$ and $d_{it}^{drop} = 0$. In other words, a student in year 5 or beyond who has completed at least 98 credits is no longer subject to the dropout risk and automatically graduates. In addition, a student who reaches $t = 8$ without having completed at least 98 classes cannot graduate ($d_{it}^{grad} = 0$) and must drop out ($d_{it}^{drop} = 1$).

The final outcome of the simulation is a "dataset" with $N = 100{,}000$ simulated high school graduates, some of whom enroll in college. For those who enroll, we obtain their simulated number of classes completed by year and final college outcome (graduation or drop out), along with their graduation or dropout year. This dataset mimics the observed student-level administrative data.

C.3 Estimation of Fixed Effects in the Dropout Probability

We now describe the estimation of the time- and type-specific fixed effects that enter in the dropout probability, $\delta(t, \theta, y)$. This estimation is nested within the estimation Θ because it must take place for every possible value of Θ.

From the simulation of college students described above, we compute the predicted dropout rates by year and student type. By definition, this is the predicted fraction of students of type j who drop out in every t, or $\hat{p}_{jt}^{drop} = f^d(\delta(t, \theta, y); \Theta)$. We compare these predicted rates with the observed ones, denoted by p_{jt}^{drop}, and compute a measure of the distance between them.

For each pair of predicted and observed dropout rates, we calculate the fixed effects $\delta(t, \theta, y)$ that minimize this distance. We do this through an iterative contraction mapping, in the spirit of Berry et al. (1995). While Berry et al. (1995) uses a contraction mapping to find the unobserved product characteristics that make predicted market shares for each product equal to their observed counterparts, we search for the time- and type-fixed effects that bring the observed the dropout rates as close as possible to their observed counterparts for each period and student type.

Formally, we use a contraction mapping algorithm to find the vector $\boldsymbol{\delta} = [\delta(t, \theta, y)]_{J(8)\times 1}$ that fulfills the following condition:

$$\|\mathbf{f}^d(\boldsymbol{\delta}; \Theta) - \mathbf{p}^{drop}\| \leq \epsilon^d, \tag{26}$$

where ϵ^d is our chosen tolerance level. Below are the algorithm steps; recall that they are conditional on a given parameter point, Θ:

1. Establish an initial guess for the fixed effects vector, $\boldsymbol{\delta}^{(0)}$.

2. Solve the dynamic optimization problem (see Appendix C.)
3. Compute the predicted vector of drop out rates $\mathbf{f}^d(\boldsymbol{\delta}^{(0)}; \Theta) = \hat{\mathbf{p}}^{drop}$.
4. Using the observed drop out rates, compute the updated fixed effects vector, $\boldsymbol{\delta}^{(1)}$, as follows:

$$\delta^{(1)}(t, \theta_j, y_j) = \ln\left(\frac{p_{jt}^{drop}}{f^d(\delta^{(0)}(t, \theta_j, y_j); \Theta)}\right). \tag{27}$$

5. Using $\boldsymbol{\delta}^{(1)}$ as the new initial guess, repeat steps 1 through 4 until either condition (26) is satisfied or a predetermined maximum number of iterations are completed.

For some parameter values, the algorithm may not be able to meet (26) due to model non-convexities. For instance, dropping out at $t = 8$ in the model is a deterministic function of the number of credits completed. However, in the data we observe some individuals graduate without having completed all credits, likely due to measurement error in number of classes completed. Another non-convexity arises because students in the model must meet a minimum number of cumulative credits per period, h_t^{drop} in order to remain enrolled. In the data, in contrast, some students remain enrolled even though they do not meet this requirement.

C.4 Recovering Type-Specific Preferences for College Enrollment

Recall that $\xi_j = \tilde{\xi}(\theta_j, y_j)$ is the type-specific unobserved preference shock for enrolling in college. For a given value of Θ, we recover it as follows. As described above, we compute the value function, $V^{coll}(\cdot)$, for every state (t, h_t, θ, y). This allows us to compare the value of going to college, $V^{coll}(1, 0, \theta_j, y_j)$, with the value of working as a high school graduate, V^{hs}. Thus, ξ_j takes on the value that makes the predicted probability of enrolling to college be equal to the observed one. Under the assumption that $\sigma_\epsilon = 1$ (see Section 4.2), we solve for ξ_j in equation (9):

$$\xi_j = \ln\left(\frac{P^{coll}(\theta_j, y_j)}{1 - P^{coll}(\theta_j, y_j)}\right) - (V^{coll}(1, 0, \theta_j, y_j) - V^{hs}). \tag{28}$$

D Goodness of Fit: Additional Evidence

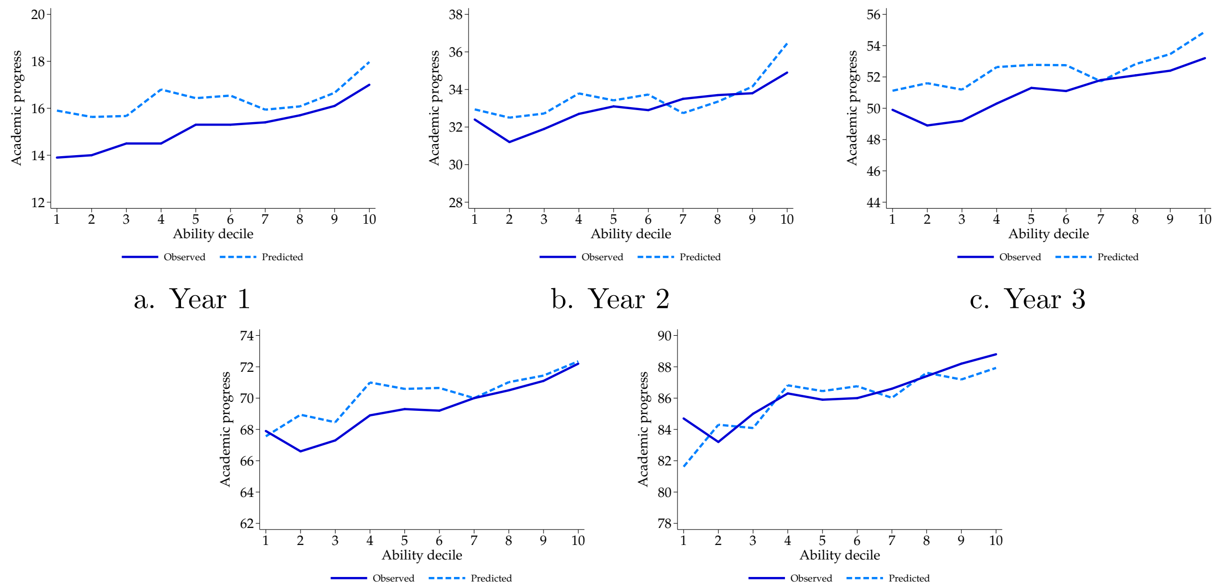
Table D1: Goodness of Fit: Cumulative Classes Completed by Year.

	Year 1		Year 2		Year 3		Year 4	
	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted
Ability quintile								
1	13.9	15.7	31.7	32.7	49.3	51.4	67.2	68.3
2	14.5	16.3	32.3	33.3	49.8	52.0	68.2	69.9
3	15.3	16.5	33.0	33.6	51.2	52.8	69.2	70.6
4	15.6	16.0	33.6	33.1	52.0	52.3	70.3	70.6
5	16.6	17.4	34.5	35.5	52.9	54.3	71.8	72.0
On-time graduate	20.5	20.7	41.2	41.5	62.0	62.0	82.8	82.1
Late graduate	18.0	19.4	35.4	37.8	52.5	55.4	70.1	72.0
Dropout later	15.8	14.9	30.4	29.0	44.5	42.9	59.3	53.9
Dropout this year	10.9	13.2	23.8	21.7	36.4	37.5	46.4	55.6
Total	15.8	16.7	33.8	34.2	52.1	53.2	70.6	71.0

	Year 5		Year 6		Year 7		Year 8	
	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted
Ability quintile								
1	83.9	83.1	88.8	90.4	93.6	91.3	96.0	92.2
2	85.7	85.7	91.3	92.0	95.3	92.3	97.2	91.6
3	85.9	86.6	91.5	92.9	95.8	93.1	98.0	94.3
4	87.1	86.9	92.7	92.6	96.6	92.3	98.4	90.5
5	88.6	87.6	94.0	93.3	97.6	93.5	99.0	91.4
On-time graduate	99.0	99.5						
Late graduate	86.6	87.1	95.6	96.8	98.3	98.4	98.9	99.3
Dropout later	72.4	64.4	78.5	68.8	82.7	75.7		
Dropout this year	69.7	64.5	84.5	77.3	86.2	76.7	88.2	81.5
Total	87.5	86.9	93.0	92.8	96.8	92.9	98.5	91.6

Source: SPADIES for observed data; model simulations for predicted data.
Notes: The table shows the observed and predicted cumulative number of classes completed by year—overall, by ability quintile, and by final college outcome. Ability quintile 5 is the highest. (%).

Figure D1: Goodness of Fit: Cumulative Classes Completed by Ability Decile and Year.

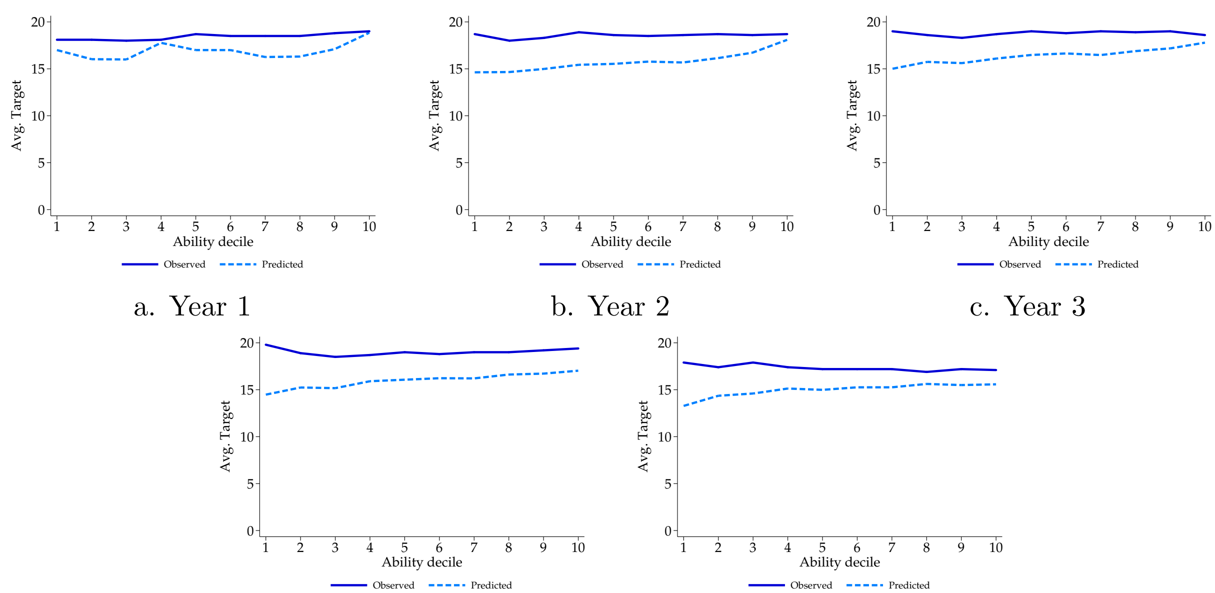


a. Year 1 b. Year 2 c. Year 3

d. Year 4 e. Year 5

Source: SPADIES for observed data; model simulations for predicted data.
Note: For each year, the panels depict observed and predicted cumulative number of classes completed by ability decile. Figures correspond to students who begin each year.

Figure D2: Goodness of Fit: Target Number of Classes by Ability Decile and Year



a. Year 1 b. Year 2 c. Year 3

d. Year 4 e. Year 5

Source: SPADIES for observed values; model's own simulations for predicted values.
Note: Observed values correspond to the average number of classes attempted; predicted values correspond to average target number of classes as defined in the model.

E Anticipated Uncertainty

Table E1: Determinants of Effort.

	Dependent variable ln(Optimal Effort)
ln(average classes completed)	-0.060^{***}
	(0.013)
Income 1-2 MW	-0.001
	(0.003)
Income 2-3 MW	-0.008^{**}
	(0.003)
Income 3-5 MW	0.001
	(0.003)
Income 5+ MW	-0.031^{***}
	(0.003)
Year=2	0.090^{**}
	(0.039)
Year=3	0.099^{***}
	(0.037)
Year=4	0.036
	(0.038)
Year=5	-0.014
	(0.037)
Year=6	-0.256^{***}
	(0.037)
Year=7	-0.289^{***}
	(0.034)
Year=8	-0.321^{***}
	(0.039)
Ability Q2	0.007
	(0.006)
Ability Q3	0.035^{***}
	(0.005)
Ability Q4	0.055^{***}
	(0.005)
Ability Q5	0.107^{***}
	(0.005)
Constant	0.707^{***}
	(0.005)
R^2	0.226
Num. obs.	116,761

Source: First-stage of 2SLS estimation based on model's predicted baseline values.
Notes: The dependent variable is ln(optimal effort), or $\ln(e_t^*)$. An observation is a student-year. Upper 5% tail of risk has been trimmed. Independent variables are state variables at t. Ability is θ; income is y; year is t; average classes completed is $\bar{h}_{t-1}$. The regression includes year fixed effects (not shown). Standard errors (in parentheses) are clustered by student. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.