American Economic Journal: Macroeconomics 2021, 13(1): 184–215
https://doi.org/10.1257/mac.20190154

College Quality and Attendance Patterns: A Long-Run View†

By Lutz Hendricks, Christopher Herrington, and Todd Schoellman*

We construct a time series of college attendance patterns for the United States and document a reversal: family background was a better predictor of college attendance before World War II, but academic ability was afterward. We construct a model of college choice that explains this reversal. The model's central mechanism is that an exogenous surge of college attendance leads better colleges to be oversubscribed, institute selective admissions, and raise their quality relative to their peers, as in Hoxby (2009). Rising quality at better colleges attracts high-ability students, while falling quality at the remaining colleges dissuades low-ability students, generating the reversal. (*JEL* I23, J12, N32)

This paper studies how US college entry patterns have evolved over the course of the twentieth century. Our empirical contribution is to document a reversal. In the early twentieth century, college entry was mainly determined by family background, with student abilities playing a lesser role. However, the roles reversed by 1960. Our theoretical contribution is to offer an explanation for the reversal. We argue that it is caused by the stratification of college qualities documented previously by Hoxby (2009). This stratification is, in turn, driven by a surge in college enrollment following World War II that allowed high quality colleges to institute selective admissions.

Our empirical work extends an existing literature that documents the increasing role of student abilities for college entry over the course of the twentieth century (Taubman and Wales 1975, Hendricks and Schoellman 2014). This literature finds that college students have become more selected on measures of academic ability, consistent with the broader sense that college has become more meritocratic. We add to this literature by collecting and harmonizing additional studies that

* Hendricks: Department of Economics, University of North Carolina, Chapel Hill, 6C Gardner Hall, CB 3305, Chapel Hill, NC 27599 (email: hendricksl@protonmail.com); Herrington: Department of Economics, VCU School of Business, Virginia Commonwealth University, Snead Hall, 301 W. Main Street, Richmond, VA 23284 (email: cmherrington@gmail.com); Schoellman: Research Department, Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55401 (email: todd.schoellman@gmail.com). Virgiliu Midrigan was coeditor for this article. We thank Rui Castro, three anonymous referees, and numerous conference and seminar participants for helpful comments and advice. We are grateful to Charles Petersen for bringing additional historical sources to our attention. Data and code for replication available at OpenICPSR (Hendricks, Herrington, and Schoellman 2021). The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

† Go to https://doi.org/10.1257/mac.20190154 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

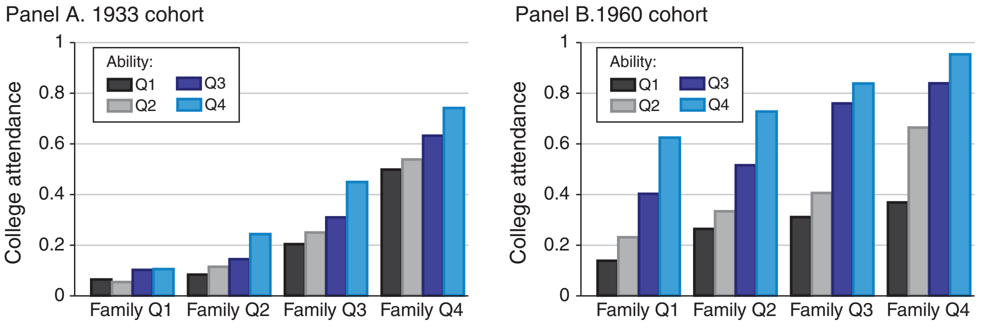

Figure 1. Changing Patterns of College Attendance: Select Cohorts

investigate the role of family background for college attendance. Importantly, ten of our newly harmonized studies tabulate college entry rates as a joint function of academic ability and family background, similar to Belley and Lochner (2007) but for high school graduating classes as early as 1933. We use these studies to estimate the effect of family characteristics on college-going that is conditional on academic ability. We find that the role for family declines at the same time that the role for academic ability rises, consistent with previous evidence from select colleges or later periods.[1]

Taken together, our data reveal a striking reversal in entry patterns that is illustrated in Figure 1. It compares the college attendance rates for two high school graduation cohorts, 1933 and 1960. Students are divided into quartiles according to their academic abilities (measured by test scores) and family background (measured by socioeconomic status). For the 1933 cohort, family background was the main determinant of college attendance; test scores mattered little, particularly for students with below-median family background. The relative importance of these two factors reversed by the 1960 cohort.

In total, we collect and harmonize 40 historical studies that document college entry rates by student abilities and/or family background. We show that the patterns observed in Figure 1 are representative of a broader trend. The reversal appears to be complete by 1960. Thereafter, we do not observe significant changes in entry patterns.

We propose a theory for the reversal that draws on three major structural transformations in the market for higher education that have been documented extensively in the literature. The first is the massive increase in college enrollment after World War II (Goldin and Katz 2008). The second is the emergence of selective college admissions based on standardized testing (Duffy and Goldberg 1998). The third is the increasing stratification of colleges by quality (Hoxby 2009). We discuss these empirical developments in Section II.

[1] Herrnstein and Murray (1994) and Karabel (2006) document similar findings on long-term changes in admissions to select colleges. Hoxby and Long (1999) report a declining role for a broad set of student background characteristics in predicting college attendance for the later period 1958–1981.

The main driving force is the expansion of college enrollment following World War II. We model this as an exogenous, common increase in the value of college for all students. This rise in demand causes high-quality colleges to hit their capacity constraint and institute selective admissions. A feedback mechanism through peer effects endogenously changes the value of college differentially for different students. Top colleges with selective admissions attract high-ability students and make themselves yet more attractive to high-ability students, whose incentives to attend college increase. On the other hand, low-ability students are constrained to nonselective colleges whose quality declines, reducing the students' incentives to attend college. This explanation for the reversal ties it to the facts on increasing college stratification documented by Hoxby (2009). The introduction of college admissions tests interacts with this mechanism. It changes students' perceptions of their own ability and how much they will learn in college, but it also changes how their peers are selected into the various colleges.

We formalize this argument in a quantitative model of college choice. The model allows us to accomplish three objectives. First, we show that a model that incorporates the well-known elements of rising college enrollment and stratification of college quality can also generate a reversal of attendance patterns. Second, we verify that the mechanism generates a quantitatively significant reversal in attendance patterns. Finally, the model allows us to distinguish between ability, which is a latent variable, and observed proxies such as test scores. We use the model to distinguish between the level and trends of sorting along both dimensions.

The key model elements are as follows. There are a large number of locations, each with a single college and a continuum of students. Students are heterogeneous with respect to their academic ability, which affects how much they learn in college, and their family background, which determines the resources they can consume if they attend college. Students cannot borrow to finance college, consistent with evidence that borrowing remained small until after the relevant era. Students decide whether to work after high school, attend their local college, or attend a college outside their local area at an extra cost. Colleges are heterogeneous with respect to their quality, which is determined by their endowment and the average ability of students they attract. Colleges accept students until they hit an enrollment cap; at that point, they adopt selective admissions and accept only the students with the highest ability.

The baseline model features only two time-varying exogenous forces that drive the reversal in college entry patterns. First, the value of college rises over time. This generates the rise in college attendance, which is one important ingredient in our story.[2] Second, standardized college entrance examinations become more common, providing additional information about student abilities. We calibrate the model to match data moments for the 1933 and 1960 high school graduation cohorts, including the attendance patterns shown in Figure 1. We choose 1933 as the earliest year for which high-quality data on college entry rates by student abilities and family background are available. We choose 1960 because by then, the reversal is complete.

[2] An existing literature has proposed several possible explanations for the rise in college attendance (Goldin and Katz 2008, Restuccia and Vandenbroucke 2014, Donovan and Herrington 2019, Castro and Coen-Pirani 2016, Alon 2018). The nature of the underlying driving force is not important for our results.

The calibration shows that the model can generate much of the reversal. We explore a number of alternative or complementary forces in our model, but we find that they play a smaller role.

In the model, the reversal occurs in response to the increasing stratification of colleges. In 1933, college entry rates are low. Since most colleges cannot attract enough students to fill all available seats, their admissions are not selective. This is consistent with admissions policies before World War II (see Section II). Most students, regardless of ability, can attend their local college, and most do so. As a result, most colleges are of fairly similar qualities. This reinforces students' incentives to attend the local college rather than incur the expense of attending a better, nonlocal one.

By 1960, college enrollment has increased substantially. High-quality colleges become oversubscribed and respond by implementing selective admissions. This raises the average ability of their student body, which makes them more attractive to students. As a result, more students, especially those of higher abilities, attend nonlocal colleges. High-ability students match up with the best colleges, raising their quality. Low-ability students are only admitted by less selective colleges, which are therefore of poor quality. Thus, the model endogenously produces the integration of the market for college education. The economy transitions from an equilibrium where all students can choose from a common set of homogeneous colleges to an equilibrium where high-ability students can choose better colleges than low-ability students. This change in the choice set generates the reversal.

We use counterfactual experiments to quantify the exogenous forces driving the reversal. The model implies that the rising value of college and the spread of standardized testing are equally important for generating the rising importance of test scores for college attendance, whereas the spread of standardized testing generates almost all of the declining importance of family background. The stratification of college quality is critical to generating a quantitatively significant reversal. The stratification has implications for the distribution of college quality available to different types of students and for the distribution of human capital formed in college, which we quantify.

We also use the model to distinguish between sorting that occurs based on test scores, which are observed, and actual ability, which is not observed in the data. We show that sorting on actual ability is stronger than sorting on test scores in both periods, but by any measure, student ability becomes more important for college attendance between 1933 and 1960.

The key to understanding these results is that test scores are noisy proxies of student abilities that have become more common over time. The observed sorting by test scores in 1933 is weak in part because of the noise in test scores but also because few students and colleges observed test scores or used them for college admissions. College entrance examinations became nearly universal after the war, providing most students and colleges with a test score signal that they use to help forecast student ability. The result is that students become more sorted on test scores. This change reflects in part more sorting on ability but also sorting on the noise in test scores. The model makes an important and novel contribution by allowing us to differentiate between the two.

The rest of the paper proceeds as follows. Section I introduces our historical data and describes the trends in college attendance patterns. Section II describes the historical context that motivates our model. Section III describes the model, Section IV provides a quantitative assessment, and Section V considers extensions. Finally, Section VI concludes.

I. Historical Data

Our empirical work extends a literature that documents the increased role of academic ability for college attendance over the course of the twentieth century (Taubman and Wales 1975, Hendricks and Schoellman 2014). We collect additional historical studies that add further support to this trend. However, our main empirical contribution is to use these additional studies to document that the role of family background has declined over time.

To do so, we collect historical studies that characterize college attendance as a function of academic ability (measured by grades or test scores) and/or family background (measured by family income or socioeconomic status) dating back to the high school graduating class of 1919. Our preferred studies tabulate college entry rates as a function of both academic ability and family background. This allows us to estimate the effect of academic ability on college entry *conditional* on family background and vice versa, as Belley and Lochner (2007) does with modern data for recent cohorts. We provide similar statistics that span a much longer period. Before describing the trends, we briefly overview the underlying studies and the data that we use.

Our evidence draws on studies from two different types of sources. For the modern era (high school graduating classes of 1960 onward), we have access to microdata or published results from large nationally representative surveys with multiple measures of family background and academic ability as well as students' postgraduation outcomes. These sources are largely familiar to economists and include most prominently Project Talent and the 1979 National Longitudinal Survey of Youth (NLSY79 hereafter).[3] For students graduating before 1960, our evidence comes from studies conducted by researchers in a variety of fields, including psychology, economics, and education. We have collected and harmonized the results from three dozen such studies, building on the research of Taubman and Wales (1975) and Hendricks and Schoellman (2014) by adding more than a dozen new studies, including many that document patterns of college attendance by family background.

The original microdata from studies before 1957 no longer exist. Instead we rely on their published results, which we have collected from journal articles, dissertations, books, technical volumes, and government reports. The design, sample, and presentation of results are different for each study. Nonetheless, it may be helpful to consider a hypothetical typical study that utilizes the most common elements in order to understand our approach. The online Appendix gives references for the

[3] Summaries and results from Project Talent can be found in Flanagan et al. (1964) and Flanagan et al. (1971), among other volumes. Online Appendix A contains details of the NLSY79 data and analysis used here (Bureau of Labor Statistics 1979–2016).

studies used and summarizes the pertinent metadata and our measures of sorting for each.

In a typical study, a researcher worked with a state's department of education to administer a questionnaire and an aptitude or ability examination to a sample or possibly the universe of the state's high school seniors in the spring, shortly before graduation. Students' academic ability was measured by their performance on the examination or, in some cases, by their rank in their graduating class. The questionnaire inquired about students' family background, with typical questions covering parental education and occupation or estimates of the family's income. These data were used to rank students based on family income or an index of socioeconomic status that would combine several different elements of the data. Finally, the researchers would inquire about students' plans for college or, alternatively, follow up at a later date with the students, their parents, or school administrators to learn about the actual college attendance. Our main data source for this era is published tabulations of these results giving the fraction of students of different academic ability or family background levels (or, ideally, both) that attended college. Most sources cover only whether the students attended college, with little comparable detail about which college they attended; Chetty et al. (2017) has information about this for recent cohorts.

Our preferred studies provide the full cross tabulation of college-going as a function of family background and academic ability. We think of these tabulations as approximations of $C(s,p)$, the function governing the share of students C that attend college as a function of student ability s (test scores, grades) and family background p (family income, index of socioeconomic status). Our goal is to summarize the conditional effect of academic ability or family background on college attendance in a simple way that is easy to compare over time. To do so, we regress $C(s,p)$ on the midpoint of the range of s and p expressed in percentiles.[4]

Figure 2 plots the estimated coefficients β_s and β_p against high school graduation cohort. The role of academic ability (test scores or grades) has risen sharply over time, in line with the previous work of Taubman and Wales (1975) and Hendricks and Schoellman (2014). Our main new finding is that the role of family background (parental income or socioeconomic status) has fallen. Studies conducted before World War II tend to find that family background is more important than academic ability, while studies after World War II tend to find the opposite.

We have highlighted three data points of particular importance. Updegraff (1936) is the first study to cross tabulate college attendance by family background and academic ability. It shows that prior to World War II, family background rather than academic ability was a more important determinant of who attended college. Flanagan et al. (1971) provides results from Project Talent, the first nationally representative study with existing microdata. It shows that sorting patterns had already reversed by 1960; see also Figure 1 in the Introduction. The NLSY79 is the starting point for most of the existing literature. Our data suggest that the level of sorting did not change appreciably between Project Talent and the NLSY79.

[4] Similar results obtain if we implement regressions using standard normalized values of s and p instead of using percentiles; see Figure B1 in the online Appendix.

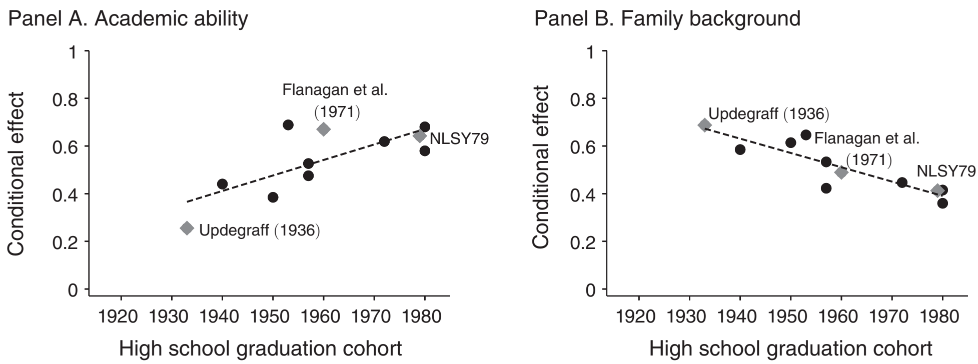

Figure 2. Changing Patterns of College Attendance (Conditional)

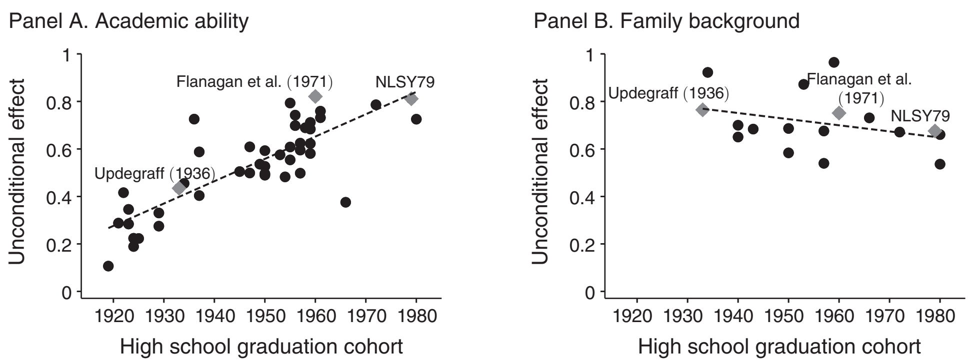

Figure 3. Changing Patterns of College Attendance (Unconditional)

Thus, in our quantitative exercises, we attempt to explain what changed sorting between 1933 and 1960.

In addition, we have many more studies that tabulate college-going as a function of family background or academic ability alone. We use these studies to construct similar time series giving the unconditional estimates of β_s and β_p. This time series allows us to incorporate many more studies covering a longer period. Figure 3 shows the results.

Figure 3, panel A shows that a large number of studies investigate the role of academic ability for college attendance. These studies consistently find that the role of ability increased over time, consistent with previous work. Figure 3, panel B shows that we have fewer studies that investigate family background. They show only a weak decline in β_p over time. A standard omitted variable argument suggests that not controlling for academic ability (which is positively correlated with family background) leads to a positive, growing bias over time as selection on academic ability strengthens. The implication again is that selection on family background must be weakening.

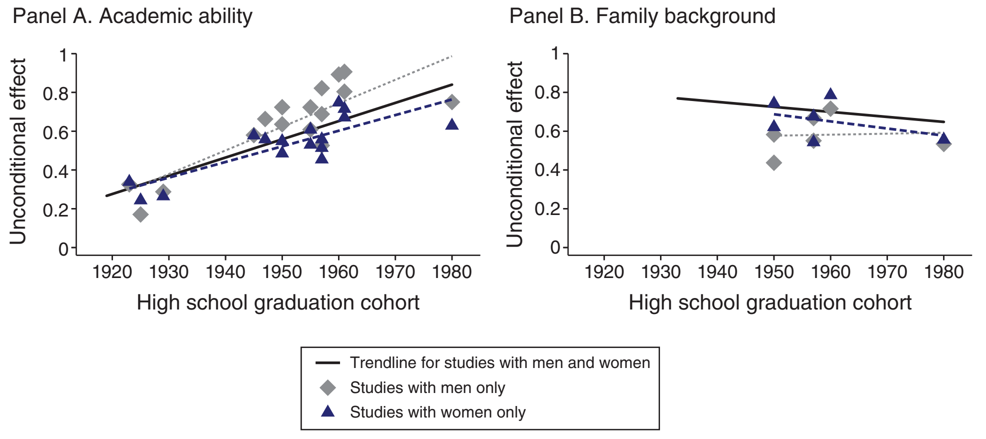

Figure 4. Changing Patterns of College Attendance by Gender

A. *Patterns by Gender*

Our results so far have covered aggregate trends. A large literature has documented important changes in the access of women and minorities to educational and labor market opportunities over this time.[5] Hsieh et al. (2019) argues that these changes may have contributed to aggregate economic growth. About one-third of our historical studies tabulate results separately for men and women, allowing us to study whether the trends differ. We focus on the unconditional effect since none of our sources before 1957 includes the necessary information to estimate the conditional effect separately by gender.

The gender-specific results are shown in Figure 4. We have a large number of studies investigating the role of academic ability by gender, including three studies from the 1920s. Those studies show that academic ability was equally unimportant for both genders in the 1920s and that it became more important for both in the 1940s and 1950s. Academic ability seems to have risen in importance more for men than for women, as indicated by the fact that the data points for men exceed those for women in almost all studies in the 1950s. We have fewer studies investigating the role of family background by gender, and the first such study dates only to 1950. Family background is equally important for men and women in 1980, and it appears from the few available studies to have been more important for women than for men in the 1950s. This is consistent with the conventional wisdom that the college attendance choices of women were more sensitive to family income in the past because it was harder for them to work their way through college, both because they had fewer job opportunities and because they earned lower wages (Greenleaf 1929, Hollis 1957).

Unfortunately, we have little to say about the importance of race. None of our sources from before the 1950s provide separate tabulations by race. In large

[5] See Altonji and Blank (1999) for an overview of labor market differences between men and women, including historical trends.

part, this is because most of these studies were conducted in northern states where Black students would have been much less common. Of the few studies of southern states, several explicitly mention that they restrict attention to schools for white students, and we suspect that the others may have done so implicitly. Hence, our early data sources and overall trends should really be read as applying to white students. We have computed in the NLSY79 that Black and Hispanic students are relatively more sorted by academic ability and less sorted by family background than are white students. Given the absence of earlier race-specific data, we can only speculate about the long-term trends implied by this fact.

B. *Controlling for Variation in Historical Study Design*

Our baseline results combine the findings of studies that differ in numerous ways, such as which proxies they use for family background or academic ability, when they measured college attendance, the size of the bins they used for tabulations, and so on. In this section, we explore whether variation in study design systematically affects the estimated trends in β_p and β_s that we document.

We start by investigating the importance of how family background and student ability are measured. Online Appendix Figure B2 documents separately the patterns of sorting by family income versus socioeconomic status. The two show similar trends, but there is a level difference; estimates of β_p are systematically larger when family background is measured as socioeconomic status than when it is measured as parental income. We conjecture that this result may arise because socioeconomic status is a better measure of permanent income than is parental income in one year. We adjust estimates of β_p to eliminate this level difference throughout.[6] Our three main studies of interest (Updegraff 1936, Flanagan et al. 1971, and NLSY79) all use socioeconomic status as the measure of family background, so our quantitative exercises are not affected by this adjustment. For similar results on measures of student abilities, we refer readers to Hendricks and Schoellman (2014).

Our approach for investigating the other dimensions of study design is based on fixing a dataset for which we have the microdata—the NLSY79—and exploring the implications of varying four dimensions of study design. First, studies vary in whether they measure academic ability using test scores or class rank. Within the NLSY, we experiment with using the Armed Forces Qualifying Test (AFQT) score or class rank at high school graduation. Second, studies vary in whether they measure family background using parental income or socioeconomic status. Within the NLSY, we experiment with using family income at the time of the student's high school graduation or creating an index of socioeconomic status. Third, studies vary in whether they measure college attendance plans or actual college attendance. Within the NLSY, we experiment with using whether high school seniors planned for one or more years of college (versus zero) and using the longitudinal aspect of the NLSY to track whether they actually attended college. Finally, historical

[6] Specifically, we regress β_p on cohort and a dummy variable for whether the study used family income instead of socioeconomic status. The intercept difference is -0.23 for conditional estimates and -0.30 for unconditional estimates. We adjust upward all estimates of β_p derived from studies with family income to correct for this difference.

studies grouped academic ability and family background into bins of various sizes. We do the same within the NLSY. Details on sample selection and measurement are available in online Appendix A.

We vary these four dimensions systematically within the NLSY and study how they affect the resulting estimates β_s and β_p. In doing so, we reconfirm that estimates of β_p are systematically higher when family background is measured using socioeconomic status as compared to family income. The other dimensions have little effect on our findings. To show this point, we conduct a falsification test. We mimic each of our historical studies by taking the NLSY data and setting the four dimensions of interest to match those of the original study so that we have comparable measures of college attendance, family background, and academic ability. We then estimate the counterfactual β_s and β_p that we would have found using a fixed dataset but varying methodologies.

In Figure 5, we recreate Figure 2 with our counterfactual estimates of β_s and β_p plotted against high school graduation cohort (for the original study).[7] It is clear from this figure that variation in study design induces noise in our estimates of β_s and β_p. Given the same NLSY79 data, we can find a range of possible results depending on what proxies we use and how we format the data. However, the main message is that this variation seems to be uncorrelated with time and hence likely does not bias our estimates of the underlying trends.

II. The Growth and Integration of the Market for College Education

Our empirical results show that college attendance patterns changed sharply in the 1940s and 1950s. In the next section, we formulate a model that is grounded in two important changes that affected colleges after the war: the growth and integration of the market for college education (Hoxby 2009). The model takes the expansion of college and the introduction of standardized test scores as exogenous driving forces and endogenously produces the integration of the market for college education. The latter differentially affects the quality of colleges available to high-ability and high-income students, which affects their attendance decisions and generates the reversal. Here we document some of the relevant facts that motivate our model setup.

We start with attendance. Figure 6 shows the dramatic increase in college enrollment using statistics on high school graduates and new college enrollment by year from the United States Office of Education (1918–1958) and National Center for Education Statistics (1962–2013). We show complementary statistics derived from census data in online Appendix C. Panel A of Figure 6 shows total new college enrollment by year. Enrollment hovered around 400,000 students per year during the Great Depression and fell during World War II. There was a large spike after the war associated in large part with the GI Bill. There was also a long upward trend until around 1970. Our historical data and our model focus on the college attendance decisions of high school graduates. Panel B of Figure 6 shows college enrollment relative to high school graduation rates. These figures were low

[7] Similar results apply for the unconditional correlations; see Figure B3 in the online Appendix.

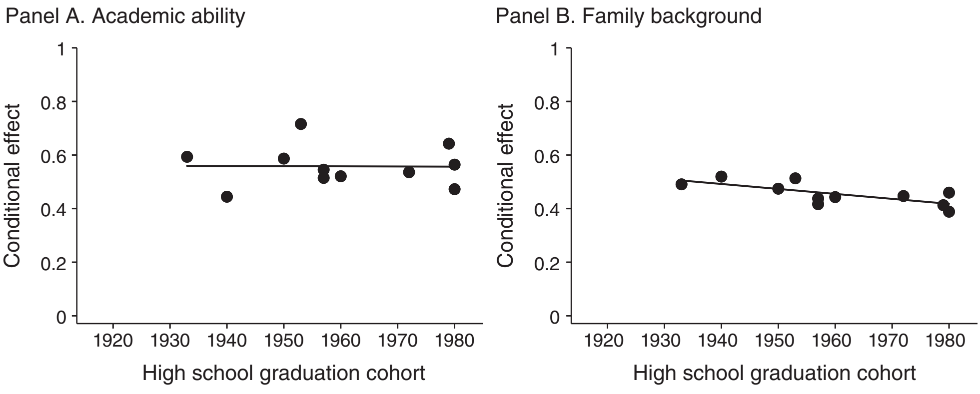

Figure 5. Counterfactual Changes in Patterns of College Attendance

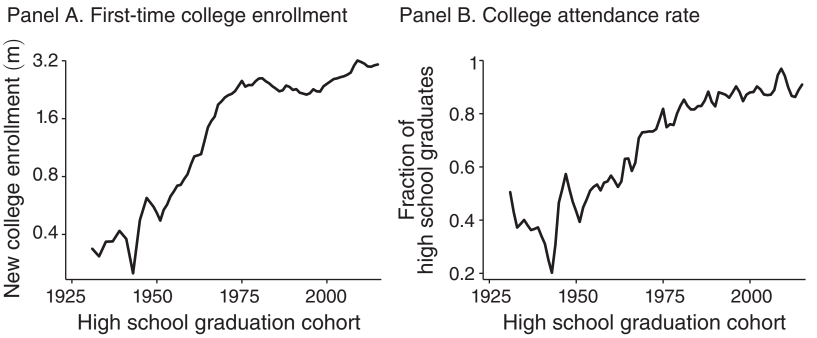

Figure 6. Increase in College Attendance

Notes: Panel A plots the number of first-year college students by year. Panel B plots the ratio of first-year college students to high school graduates.

during the Great Depression and fell during World War II. They spiked after the war but also show a sustained long-term increase to around 80–85 percent.

Our model takes the rise in the demand for college itself as an exogenous driving force. Nonetheless, it is useful to note that there are several plausible candidates for this trend in the literature. One is the declining cost of college (Donovan and Herrington 2019). We document in online Appendix C that the cost of a year of college relative to income fell by three-fourths between the Great Depression and the postwar period, reaching an all-time low in 1947. Alternatively, Alon (2018) argues that changes in high school and college curricula around this time made college more valuable. Several papers in the literature suggest that the success of the GI Bill may have triggered widespread changes in beliefs about the benefits of college (Bound and Turner 2002, Goldin and Katz 2008).[8] The exact source of the rise in demand for college is not important for our results.

[8] The rising college wage premium is often considered an important driver of the long-run expansion of college attendance (Goldin and Katz 2008, Restuccia and Vandenbroucke 2014). However, it is challenging to attribute the postwar surge of attendance to wages because the college wage premium was at its lowest around 1950 before

We now turn to the integration of the market for college education. An important driving force for this change is that college applications and admissions procedures became standardized and streamlined after World War II. Prior to World War II, college admissions decisions were based on whether students had demonstrated mastery of certain knowledge. The subjects to be mastered, level of knowledge required, and mechanism for demonstrating mastery varied widely by college and year, with many colleges offering multiple paths to achieve admissions (Kurani 1931). Given the idiosyncratic nature of college requirements and admissions processes, college guides from the 1930s recommended that students choose a college as early as possible and then work with its admissions department to demonstrate compliance with the relevant standards (Halle 1934). In many states, high schools would form a relationship with a local college. The high school tailored its curriculum to the college's requirements, while the college agreed to certify and accept the high school's graduates for admissions.

This system was replaced by a homogeneous system based on standardized college admissions exams (the SAT and later its competitor, the ACT) after the war. The real cost of these tests fell by two-thirds after the introduction of machine scoring in 1937, leading them to become an attractive option for assessing the rapidly growing number of applicants after the war; see online Appendix C for details. Figure 7 shows the main takeaway: an explosion of test taking took place from 1950 to 1965. At the peak, there were more tests taken than college freshmen, and roughly three-quarters of high school seniors took a test.[9] This rise in test taking is the second exogenous input into our model.

The standardization of admissions and the surge of demand for college had two important implications that will act as mechanisms for our model. First, they led students to apply to more colleges over a larger geographic area. Hoxby (2009) documents some geographic facts and cites the fall in transportation and communication costs. Before the war, students applied to multiple colleges only rarely because of the difficulty of complying with multiple admissions requirements.[10] College guides from after the war already recommend applying to "three or four" colleges (Dunsmoor and Davis 1951). Just under three-fourths of applicants applied to a single college in 1947, only one-half did so by 1959, and less than one-third did so by 1979 (Roper 1949, Flanagan et al. 1964, Pryor et al. 2007). This "plague" or "specter" of multiple applications was a recurring topic of discussion among admissions officers in the 1950s.[11]

Second, the growth in applications allowed better colleges to switch from recruitment to selective admissions. Before the war, the typical college accepted

subsequently rising. High school graduates would have to predict the future increase in the wage premium and have a very low discount rate for the timing to work.

[9] Figures include ACT test taking from its introduction in 1959 onward. The discontinuity reflects a break in how the SAT reports test taking; until 1971 it reports tests taken, while from 1972 it reports unique test takers.

[10] Partridge (1925) provides figures from a large urban high school with a large majority of students attending college, which was rare at the time. Even at this evidently advantaged high school, only 11 percent of students applied to more than one college.

[11] See Duffy and Goldberg (1998, 37–39) and Bowles (1967, 117).

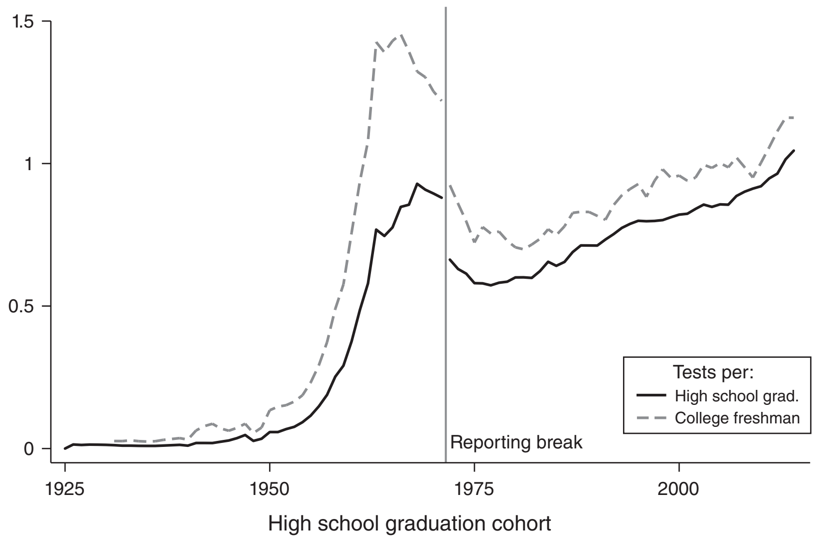

Figure 7. Rise of College Entrance Examinations

all students who met the posted requirements.[12] The surge of attendance after the war was sufficiently large and rapid that more desirable colleges found it infeasible to expand enrollment in proportion to their applications. College entrance exam scores emerged as a key metric of college quality and selectivity. The result was the "fanning out" of colleges documented in Hoxby (2009): average student test scores have risen at top colleges but fallen for median and below-median colleges since at least 1962.

In contrast to these changes, the way students financed college remained mostly constant over this period. Three surveys provide a very similar picture of how students financed college throughout the 1950s (Hollis 1957; Iffert and Clarke 1965; Lansing, Lorimer, and Moriguchi 1960). The main source of financing was students and their families, accounting for 80–87 percent of the total. Loans were rare and small: only 1.9–3.3 percent of students and 14 percent of families report borrowing from any source, with the total borrowed accounting for a tiny fraction of total expenditures (see also Harris 1962). These facts lead us to model students and parents making college financing choices in autarky.

The lack of borrowing reflects in part the fact that significant federal government involvement in college financing via grants and loans started only in 1959 with the National Defense Education Act and did not become quantitatively important until the 1960s. These programs had large effects for cohorts after our period of

[12] From Duffy and Goldberg (1998, 35): "[S]tudents tended to apply only to their first-choice college, and they were usually accepted," and "Admissions officers visited selected high schools, interviewed candidates for admissions, and then usually offered admission to students on the spot." Less politely, this was the "warm body, good check" stage of admissions (34). Admission was certainly implied under the widely used certificate system (Wechsler 1977).

interest. By 1969–1970 the share of college expenses paid for by families had fallen below three-quarters, with loans taking up much of the shortfall (Haven and Horch 1972). The federal government did intervene in other ways in earlier periods, mainly through the GI Bill, which was enormous (accounting for one-quarter of all college income at its peak) but also short-lived and applied only to men, and so is unlikely to drive our lasting changes. Online Appendix C has further details.

The main change to college financing during our period of interest was a gradual increase in the frequency and importance of scholarships, from being negligible before World War II to 8 percent of college financing by 1957. Although they were not large, scholarships were targeted toward students based on academic ability and family background, giving them an outsized influence. We study their role for college attendance patterns in Section VA.

III. Model

We develop a model of college choice and admissions that captures the forces described in Section II. The economy contains a discrete number of locations (islands) indexed by $i \leq I$. Each location is home to a single college and a measure 1 of new high school graduates per year. Locations are heterogeneous with respect to the quality of the local college but are otherwise identical. Each college sets an admissions policy that specifies the expected ability needed for admission. Students with heterogeneous family backgrounds and expected abilities decide whether to attend the local college, attend college elsewhere, or work straight out of high school.

The model is static: it covers the college attendance decisions of a single high school graduation cohort in isolation. Our goal in the next section is to show that the model can generate a quantitatively significant reversal of who attends college, consistent with the data. When we do so, we simulate two equilibria of the model, corresponding to the equilibrium of the 1933 and 1960 cohorts. Most parameters will be held fixed, but we will allow two to vary over time; we denote these parameters with a t subscript to highlight their particular role in the analysis.

A. *Colleges*

Colleges have endowments $\bar{q}_i$ spaced uniformly on the interval $[\underline{q}, \bar{q}]$. This represents the literal endowment of the college: the land, buildings, and financial accounts that a college possesses. The college's quality q_i depends on both its endowment and the mean ability of its students $\bar{a}_i$, $q_i = \bar{q}_i + \bar{a}_i$.

Colleges set an admissions criterion, which is specified as a minimum expected ability for acceptance, $\underline{a}_i$. Their objective is lexicographic. Their first priority is to maximize enrollment e_i until it hits capacity E. Keeping enrollment high is important for colleges because they need to finance large fixed costs associated with building maintenance. For colleges that are at capacity, their goal is to maximize quality, which leads them to set the highest value of $\underline{a}_i$ that maintains full enrollment.

We hold endowments and capacity fixed in the baseline model. Although there are interesting dynamics in the mean and the distribution of endowments, they begin around roughly 1980, after our period of interest.[13] Our motivation for holding capacity fixed is that enrollment rose quickly after the war, leaving colleges little time to build classrooms or dormitories. For example, first-year enrollment in 1947 was 150 percent larger than in 1943 and 50 percent larger than the prewar peak. However, we also explore extensions where capacity expands in online Appendix D.

B. *Students*

High school graduates have heterogeneous endowments (a,p,z,s,l). Ability a affects how much they learn in and benefit from college. Family (parental) background p determines the resources that students can access to finance consumption if they attend college. It can be thought of as including transfers from parents plus income from work while in college, minus payments for tuition. Children from richer families can access more funding and enjoy higher consumption while in college, making it more enjoyable. Students are endowed with two noisy signals of their ability, z and s. Finally, l is their endowed location, which determines the quality of their local college. Endowments are drawn from a distribution $F(a,p,z,s)$ that is constant across locations and (in the baseline analysis) over time.

Ability is unobservable to students and to colleges when application and admissions decisions are made. Instead, students and colleges form expectations about the student's ability. Below we assume that p, z, and s are all correlated with a and hence are potentially useful for forming expectations. Our first time-varying driving force is the subset of this information $\mathcal{I}_t$ that is observed by cohort. We assume that prewar cohorts had information sets $\mathcal{I}_t = (p,z)$, while postwar cohorts had more information, $\mathcal{I}_t = (p,z,s)$. The variable z represents the set of information that is available in the absence of test scores. Empirically, it can be thought of as a student's transcript (courses taken, grades, rank in class) and letters of recommendation. The variable s represents the information provided by scores on standardized college admissions tests, which are available only to postwar cohorts. We denote by $\mathbb{E}(a|\mathcal{I}_t)$ the expected ability given available information.

Given this time-varying information set, graduates make an irrevocable decision whether to work as a high school graduate or attend college. High school graduates who enter the labor force directly obtain a continuation value V_t^{HS}. This value varies over time to capture changing wages or nonpecuniary aspects of working as a high school graduate.

Alternatively, graduates can choose to attend a college. We start by defining the value of attending the local college, which is feasible as long as the student's expected ability exceeds the college's cutoff, $\mathbb{E}(a|\mathcal{I}_t) \geq \underline{a}_l$. The student finances consumption while in college using family resources p, which gives them flow utility $\log(p)$. Students are restricted from borrowing against their future income, although they would wish to do so, consistent with the financial environment through the mid-1960s. Upon graduation, they acquire human capital given by a CES production function that takes the student's ability and college quality as

[13] Details and figures available upon request.

inputs, $h(a,q) = \left[\phi q^{\gamma} + (1-\phi)a^{\gamma}\right]^{\alpha/\gamma}$. The parameter ϕ is the weight on quality in the production of human capital, γ governs the elasticity of substitution between quality and ability, and α is the overall curvature of human capital formation.

College graduates enjoy a continuation value that depends on $\log\left[h(a,q)\right]$, which captures the idea that period utility is logarithmic and future consumption is likely to be proportional to human capital. It also depends on a general term V_t^C, that varies over time to capture changing wages and nonpecuniary aspects of working as a college graduate. The total value of attending the local college is then given by

$$(1)\qquad V(p,\mathcal{I}_t,l) = \log(p) + \alpha\mathbb{E}_a\left[\log\left(\left[\phi q_l^{\gamma} + (1-\phi)a^{\gamma}\right]^{1/\gamma}\right)\Big|\mathcal{I}_t\right] + V_t^C.$$

Finally, students can pay a financial cost κ to apply to and attend nonlocal colleges. This cost represents transportation costs, application costs, out-of-state tuition fees, and so on. Once this cost is paid, students can attend any college where their expected ability meets the admissions criteria. On the other hand, it reduces their resources for consumption while in college to $p-\kappa$, which makes applying to nonlocal colleges particularly expensive for low-p students. These trade-offs are embedded in the value function for nonlocal applicants:

$$(2)\qquad W(p,\mathcal{I}_t,l) = \mathbb{E}_{a,\zeta_i}\left\{\max_{i\neq l:\mathbb{E}(a|\mathcal{I}_t)\geq \underline{a}_i} V(p-\kappa,\mathcal{I}_t,i) + \bar{\zeta}\zeta_i\right\},$$

where ζ_i is an i.i.d. type-I extreme value taste shock for college i. It is revealed to students only after they choose to apply outside their local area. Its primary purpose is to make the model more tractable computationally by smoothing students' application behavior across the parameter space. The parameter $\bar{\zeta}$ controls the dispersion of the shocks, which in turn controls the relative importance of taste versus human capital formation for college choices.

Students choose among these three options (work as high school graduate, attend local college, search among all colleges) to maximize lifetime utility:

$$(3)\qquad \max\left\{V_t^{HS} + \bar{\eta}\eta_{HS}, V(p,\mathcal{I}_t,l) + \bar{\eta}\eta_V, W(p,\mathcal{I}_t,l) + \bar{\eta}\eta_W\right\},$$

where the ηs are again i.i.d. type-I extreme value taste shocks scaled by $\bar{\eta}$ and introduced mainly for computational tractability. Since only the relative utility of the three choices matters, we normalize $V_t^{HS} \equiv 0$, which is equivalent to reinterpreting V_t^C as shifting the relative value of college versus high school for cohort t. Thus, for example, students prefer attending the local college to working as a high school graduate if

$$\log(p) + \alpha\mathbb{E}_a\left[\log\left(\left[\phi q_l^{\gamma} + (1-\phi)\,a^{\gamma}\right]^{1/\gamma}\right)\Big|\mathcal{I}_t\right] + V_t^C + \bar{\eta}\eta_V > \bar{\eta}\eta_{HS}$$

and similarly for the remaining comparisons.

We have two driving forces that we will vary as we simulate the choices of different cohorts: V_t^C, which we use to fit the fraction of each cohort that attends college, and $\mathcal{I}_t$, which captures the improved signals of students' abilities after the introduction of standardized testing.

C. *Equilibrium and Equilibrium Selection*

An equilibrium in this model consists of college choices for students (whether to attend and, if so, which college), admissions cutoffs for colleges, and college qualities. The choices need to maximize the lifetime utility of each student (equation (3)) and the lexicographic objective of the colleges. The equilibrium quality of each college also has to be consistent with the set of students who actually attend the college.

As in most models with peer effects, we face the possibility of multiple equilibria. For example, if we take an equilibrium and rank colleges from highest to lowest quality, it may be the case that we can switch the student bodies of the highest- and lowest-quality colleges and obtain a new equilibrium. The extent of multiplicity depends on the relative importance of peer effects as compared to differences in college endowments in the overall production of college quality. We follow the approach of Epple et al. (2017) and focus on what they call a "hierarchical adherence" equilibrium, which requires the college quality hierarchy to follow the endowment hierarchy.[14] This produces what we (and they) view as the most natural equilibrium. We verify computationally that such an equilibrium exists. Extensive experimentation with different (weakly increasing) initial guesses of college quality as a function of college endowment $q_i(\bar{q}_i)$ suggest that there is a unique equilibrium in the parameter region of interest.

IV. Quantitative Assessment

In this section, we calibrate the model and study its implications for the time series patterns of sorting. We simulate two equilibria of the model, corresponding to the 1933 and 1960 cohorts. We calibrate the model to fit the fraction of students of different types who attend college in the two cohorts as well as the application behavior of students by cohort. As emphasized in the last section, most of our parameters are time invariant. Our calibration exercise is thus judged on whether we can generate a quantitatively large reversal in college attendance patterns using two time-varying driving forces: a change in the relative value of college for all students and an increase in information about students' abilities. We show that the model is capable of doing so. We explore the mechanism, which is that the endogenously generated change in application and admissions behavior differentially affects the quality of college available to students of different types. We disentangle the role of the two exogenous driving forces as well as show the importance of the stratification of colleges as an endogenously generated mechanism in the model.

A. *Calibration*

The model has a number of parameters that need to be calibrated for a quantitative assessment. We start with the parameters relevant to colleges. We assume that

[14] They distinguish between private and public colleges when taking their model to the data; our historical data do not allow us to do so.

colleges have endowments spaced uniformly on the interval $[\underline{q}, \bar{q}]$. These parameters fix the mean and range of college endowments' contributions to overall college quality. We also need to choose the capacity of each college, E.

The second set of parameters govern students' endowments. We assume that $(a, \log(p))$ are drawn from a bivariate normal distribution with mean (μ_a, μ_p), standard deviations $(1, \sigma_p)$, and correlation ρ.[15] We assume that the signals z and s are unbiased draws from a normal distribution with standard deviations σ_z and σ_s.[16] Since all variables are jointly normal, we can solve analytically for $\mathbb{E}(a|\mathcal{I}_t)$.

The third set of parameters govern human capital formation and its labor market returns. The human capital production function has three parameters, ϕ, γ, and α, that govern the relative weight on quality versus ability in the production of human capital, the elasticity of substitution between the two, and the overall curvature of human capital production. The parameter κ is the extra cost to apply to nonlocal colleges. The relative value of college (compared to high school) for cohort t is V_t^C.

Finally, we have two preference parameters, $\bar{\eta}$ and $\bar{\zeta}$, that provide a scale to the type-I i.i.d. extreme value shocks for the three broad choices (work as a high school graduate, attend local college, attend nonlocal college) and for specific nonlocal colleges, respectively. All told, this gives us 17 parameters, which are summarized in Table 1.

We choose these parameters to fit a weighted quadratic loss function with 32 moments from each cohort, or 64 in total. Our main targets are the share of students in each (s,p) quartile and the share of each (s,p) quartile that attends college for each cohort. We map the test scores and indices of socioeconomic status in the data into the model objects s and p. Note that for the 1933 cohort, we match the model and the data on the basis of test scores, even though we have assumed that agents in the model do not know test scores. The idea is that although we have access to test scores from Updegraff (1936), and students covered by this study likely did as well, test scores—particularly standardized college admissions test scores—were generally rare at the time.

Finally, we fit a measure of how nationally integrated the market for higher education is. Before World War II, most students applied to only a single college, typically one with a close relationship with their high school. Our best estimate for the 1933 cohort is that 85 percent apply to just one college, which is a midpoint between the estimate of 89 percent from the 1920s and 75 percent from 1947 (see Section III for sources). By contrast, about one-half of students in the 1960 cohort applied to multiple colleges (Flanagan et al. 1964). We calibrate the share of students attending nonlocal colleges in the model to fit the share of students who apply to multiple colleges in the data. Our underlying idea is that students who apply to only a single college are probably choosing a college with a close relationship with their high school and a high probability of acceptance, which is how we think of the local college in our model. Submitting multiple applications indicates a broader search.

[15] Our human capital production function requires a to be positive. We truncate the distribution and replace all nonpositive values with a small positive value.

[16] We also explored allowing for a more general structure of correlations between (a, p, z, s) but found that doing so does not substantially improve the model fit or change its predictions.

Table 1—Calibrated Parameters

Parameter	Description	Value
Colleges		
$\underline{q}$	Lower bound on college endowments	0.61
$\bar{q}$	Upper bound on college endowments	2.26
E	College capacity	0.55
Endowments		
μ_p	Mean log parental transfer	−0.08
μ_a	Mean ability	0.90
σ_p	Standard deviation of log transfer	0.10
ρ	Correlation of parental transfers and ability	0.43
σ_z	Noise in information signal	0.74
σ_s	Noise in test score signal	1.50
Human capital production		
γ	Substitution between ability and quality	−0.26
ϕ	Weight on quality	0.74
α	Curvature of human capital production	0.71
κ	Application cost	0.41
V_t^C	Relative value of college	(−0.37, 0.66)
Preferences		
$\bar{\eta}$	Scale of taste shocks among broad education choices	0.08
$\bar{\zeta}$	Scale of taste shocks among colleges	0.08

Note: This table gives model parameters, a brief description of their role, and the calibrated value.

B. *Model Fit*

Table 1 describes the calibrated parameters. We highlight two areas of special interest. First is the human capital production function. This function puts a large weight on college quality ($\phi = 0.74$). It also finds that college quality and student ability are complementary inputs to the formation of human capital ($\gamma < 0$). This calibrated production function implies that students, particularly high-ability students, have incentives to seek out high-quality colleges.

Second, we are interested in the evolution of the parameters that vary by cohort. The relative value of attending college V_t^C rises substantially. The level of V_t^C governs whether a worker who will acquire $h = 1$ units of human capital prefers high school (negative) or college (positive). The rise in V_t^C generates a large increase in college attendance. We allow for additional information about students' abilities in later cohorts in the form of s (test scores). The large variance of s relative to z suggests that these test scores are a less precise signal than the sum of other information already available on a student's transcript or in their letters of recommendation, which is consistent with the literature that estimates the marginal value of test scores for predicting freshman year grades (Morgan 1989, Kobrin et al. 2008). Nonetheless, we show below that this change does help the model fit the reversal in sorting patterns.

Table 2—Summary of Model Fit, 1933 and 1960

	1933 cohort		1960 cohort	
	Data	Model	Data	Model
College attendance	0.29	0.29	0.53	0.52
Local college attendance	0.85	0.85	0.51	0.51
β_s	0.23	0.29	0.71	0.78
β_p	0.69	0.67	0.48	0.60

Notes: Columns compare the model to the data for the 1933 and 1960 high school graduation cohorts. The rows provide four moments: the share of graduates who attend college, the share of college students who attend a local college, and the importance of test scores and family background for determining who attends college.

The model delivers a good fit to the data. Table 2 briefly summarizes the four main moments that we target for the 1933 and 1960 cohorts: the fraction of high school graduates who attend college, the fraction of college enrollees who choose the local college, and sorting by test score and family background. For the table, we focus on the conditional correlations estimated by regressing the share of students attending college on (percentiles of) s and p, exactly as in Section I. The model fits the targets well, with the main challenge being that it captures only about one-third of the decline in the importance of family background for college attendance. We consider an extension in Section V that goes some way toward improving the fit along this dimension. Figure 8 shows the full pattern of college entry by (s,p) quartiles from the data and the model for the 1933 and 1960 cohorts.[17] Family background dominates attendance patterns for the 1933 cohort, but academic ability does for the 1960 cohort, consistent with the data. The main area where the model struggles is with the increase in attainment of students with low test scores, particularly those with both low test scores and below-median family background.

We focus on the model's implied changes in sorting by test scores (s) and family background (p) because this is what we observe in the data. However, the model also allows us to construct sorting when ability is measured directly (a) or proxied for by expected ability $(\mathbb{E}(a|\mathcal{I}_t)$, constructed using the information available to students and colleges). Table 3 compares the sorting in 1933 and 1960 when ability is proxied for by test scores, actual ability, or expected ability. In each case, we measure sorting using the coefficients from a regression of college attendance on the percentiles of the respective ability proxy and family background, as in Section I.

Table 3 offers two main lessons. First, there are large differences in the implied patterns of sorting depending on which ability proxy is used. "Ability" sorting is weakest when measured by test scores because our calibration implies that test scores are a noisy proxy for ability. It is much stronger when measured using actual ability. Finally, it is stronger still when measured using expected ability because that is the information available for college attendance and admissions decisions. In some cases, students are sorting into college based on noise in their expectations.[18] The

[17] See online Appendix D for fit on remaining dimensions.

[18] The measured sorting on family background follows an inverse pattern. This finding can be understood primarily as a result of using noisy, correlated regressors. For example, when ability is proxied using test scores in

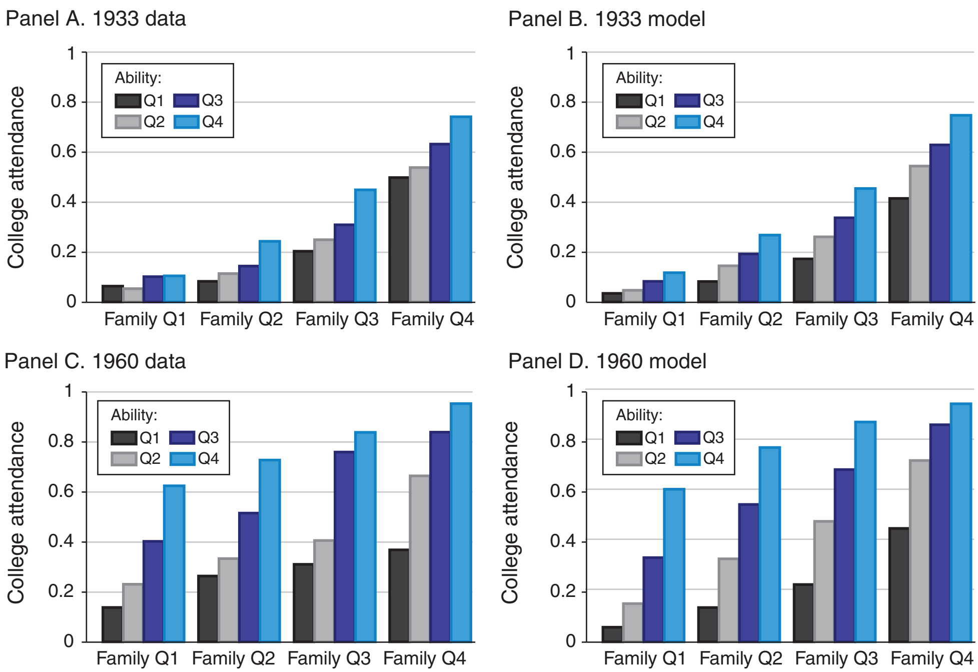

Figure 8. College Attendance Patterns

Table 3—Sorting by Cohort for Alternative Proxies for Ability

	1933 cohort		1960 cohort	
	Ability proxy	Family	Ability proxy	Family
Test scores (s)	0.29	0.67	0.78	0.60
Ability (a)	0.61	0.51	1.08	0.39
Expected ability ($\mathbb{E}(a \mid \mathcal{I}_t)$)	0.84	0.37	1.46	0.17

Notes: Columns give estimated coefficients from a joint regression of college attendance on an ability proxy and family background in the calibrated 1933 and 1960 equilibria of our model. Rows give different ability proxies: test scores (as in the baseline), ability, and expected ability given available information.

findings suggest a more nuanced view of the historical trends. The model implies that ability has always been more important for college attendance than family background. Focusing on sorting by test scores can obscure this fact. These findings are consistent with the results from Cooper and Liu (2019), which finds that much of the apparent mismatch between students and colleges on the basis of test scores is due to noise in test scores.

The second main lesson of this table is that students become more sorted on ability and less sorted on family background over time regardless of which proxy we

the regression, then the coefficient on family background is inflated because family background is correlated with expected ability, which is only imperfectly controlled for by test scores.

use to measure ability. In fact, the increase in sorting on ability is about as large as the increase in sorting on test scores. Thus, our findings do still support that college has become more "meritocratic" over time. In the next section, we explain how the model is able to generate this change.

C. *Model Mechanisms*

The model generates a large reversal in college attendance patterns. The calibrated 1933 equilibrium features a local market for college: few students attend college, and most who do attend their local college. The exogenously higher V_t^C in the 1960 equilibrium increases the share of students who wish to attend college. For colleges, this implies that many of the best colleges are oversubscribed, and so selective admissions are more common. For students, it implies that many more students apply to and attend colleges outside their local area. In equilibrium, this integration of the market for college education leads to a different menu of colleges and college qualities available to students of different types, which in turn generates different college attendance patterns. Although there are important feedback effects between college and student behavior, we consider each in turn.

For colleges, the main effect of the expansion of enrollment is that many more colleges are capacity constrained and practice selective admissions. Whereas in the 1933 equilibrium, only 8 percent of colleges have selective admissions, in the 1960 equilibrium, 86 percent do. Recall that our definition of selective admissions is minimal: it means only that a college is at capacity and imposes any floor on expected ability for admission.

The widespread adoption of selective admissions leads colleges to be much more differentiated by student ability. This change can be understood as the result of three differences between the 1933 and 1960 equilibria. First, colleges that practice selective admissions in the 1933 equilibrium are even more selective in the 1960 equilibrium. Second, many more colleges are selective in the 1960 equilibrium. Finally, the fact that most students in the 1933 equilibrium attend their local college implies that even low-quality colleges have some high-ability students. Many fewer students attend local colleges in the 1960 equilibrium, which further reduces the average student ability in these low-quality colleges.

Hoxby (2009) identifies growing quality heterogeneity as one of the central features of the integration of the market for college education. She constructs a figure that ranks colleges by median test score (e.g., SAT test score) of their student bodies, with test score again acting as the empirical proxy for expected ability. She shows that test scores have risen at the top colleges but fallen for below-median colleges. While we cannot adopt her data as a formal calibration target, we can construct the same figure using our model and compare the two.[19] Figure 9 shows the same figure implied by our model. Here, we rank colleges by test score, then compute the average test score of the top decile of colleges, the second decile, and the bottom four quintiles, where each decile has an equal share of enrollment. We plot the points

[19] Unfortunately, Hoxby's (2009) data stretch back only to the 1950s, not the 1930s, and it provides only a figure (Figure 1), not the data plotted in the figure.

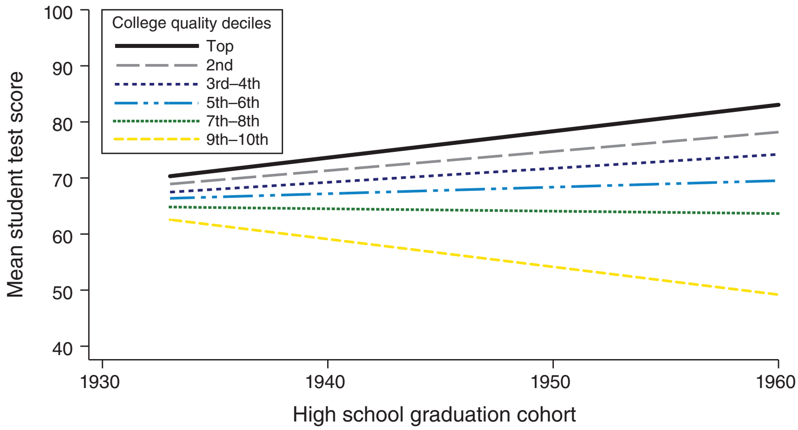

Figure 9. Fanning Out of Colleges

against time to mimic the same figure in Hoxby (2009), although of course we have only two equilibria.

In the 1933 equilibrium, only the very top decile of colleges is selective, so the gap in test scores between top and bottom colleges is small, less than 10 percent. In the 1960 equilibrium, college quality is much more dispersed. Mean test scores are higher for above-median colleges but lower for below-median colleges. The gap between top and bottom colleges in the 1960 equilibrium is around 30 percent. This figure matches the earliest figures in Hoxby (2009) quite well. Hoxby (2009) finds that the gap in 1962 was 40 percent and suggests based on spotty earlier evidence that the gap in the 1950s was probably around 20 percent. Hence, both the level and the trend in college quality heterogeneity are consistent with existing evidence.

For students, the main changes are higher college attendance (which is delivered by the exogenous rise in V_t^C) and lower rates of local college attendance. The model is calibrated to fit each change. The higher dispersion of colleges by quality and the lower rates of local college attendance have important implications for the menu of colleges available to each student. One metric that speaks to this changing menu is the fraction of students who have access to their first-choice college, meaning the college they would attend if students were individually exempted from admissions standards. In the 1933 equilibrium, 99 percent of students can do so. This finding is explained by the fact that few colleges are selective, but also by the fact that quality gaps are generally small enough that most students prefer to attend their local, unselective college.

In the 1960 equilibrium, only 55 percent of students can attend their first-choice college. The share of students who can attend their first-choice college varies strongly in characteristics such as test score. For example, Figure 10 plots the fraction of students who can attend their first-choice college by (s, p) quartile. While most top-quartile test score students can attend their most preferred college, few bottom-quartile test score students can.

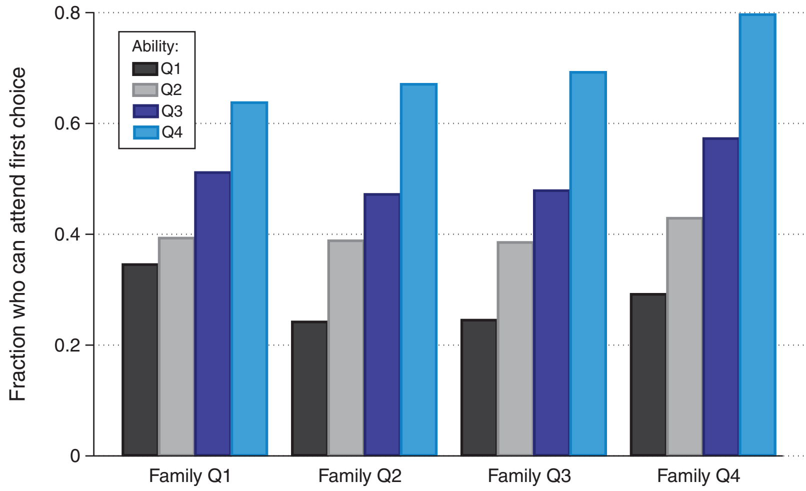

Figure 10. Access to First-Choice College

Table 4—Human Capital Formation

	Δ mean $\log(h)$	Share higher h	Share higher q
Baseline	−0.12	0.27	0.36
No change in sorting	−0.11	0.00	0.00
No change in attainment	−0.01	0.63	0.48

Notes: Columns give the change in mean human capital of college graduates between 1933 and 1960, the share of college graduates with higher human capital in 1960, and the share of colleges with higher quality in 1960. Rows give the baseline model and counterfactual models that hold either sorting patterns or college attendance rates fixed.

A second metric to gauge the changing menu of college qualities is to examine the changing distribution of human capital and college quality. Table 4 provides several statistics that summarize these changes. Focusing on the first row, we see that the average human capital of college graduates declines over time. The distribution also becomes more dispersed because of increased stratification. Students in the top 27 percent of the 1960 human capital distribution have more human capital than those in the top 27 percent of the distribution in 1933, while students in the bottom 73 percent have less. Quality drops at 64 percent of colleges, again suggesting growing dispersion.

The next two rows in Table 4 give the results from counterfactual experiments that explain these findings. The second row shows the same statistics for the case in which, for each student, we take the decision of whether or not to attend college from the 1960 equilibrium but the decision of which college to attend from the 1933 equilibrium. This row shows that the expansion of college lowers the mean human capital of college graduates, primarily because the students who enter college in the 1960 equilibrium but not the 1933 equilibrium have lower expected ability. By itself, this change implies that all students should have lower human capital and all colleges should have lower quality. The third row shows results from the reverse case: for each agent, we take the decision of whether or not to attend college from the 1933 equilibrium but the decision of which college to attend from the 1960 equilibrium. This row shows that sorting improves outcomes for about two-thirds of students and about one-half of colleges. For the most part, these are high-expected-ability students who sort into selective colleges with their peers. Overall, the results in the first row combine the effects of an expansion of education, which lowers average human capital and quality, and a change in sorting, which raises human capital and college quality for selective colleges and the high-ability students who attend them.

Thus, the model endogenously produces changes in application and admissions behavior consistent with the integration of the market for college education and the facts documented in Section II. These changes combine to imply very different college qualities available to students of different academic abilities and family backgrounds, because colleges are more selective and more differentiated by quality, and students are more willing to apply to nonlocal colleges. The change in college qualities available to students drives the change in sorting patterns. In the next section, we consider which of the driving forces is most responsible for our results.

D. *Decomposing Results*

Next we decompose the results to highlight the role of three essential ingredients: the rising value of college, changing information, and the stratification of colleges by quality. We start by taking our calibrated model with the parameters from Table 1. These parameters fit the 1933 and 1960 data as well as possible. We then construct two alternative 1960 equilibria, which hold $\mathcal{I}_t$ or V_t^C fixed at the 1933 level. We show the results in Table 5. The rows are the same fit statistics as in Table 2, as well as the degree of sorting that we would estimate if we regressed college attendance on actual ability rather than proxies, which we denote by β_a, and two summary statistics for the model mechanism: the share of college students who can attend their first-choice college and the share of colleges that are selective. The columns show results for the 1933 and 1960 baseline calibrations and the two counterfactual 1960 equilibria.

We start with the third column of results, which shuts off the test scores and focuses on the rising value of college. The rising value of college accounts for nearly all of the rise in college attendance and the decline in local college attendance. It also produces essentially all of the national integration of the market for college and of the switch to selective admissions. However, the results for sorting patterns are more subtle. The rising value of college accounts for about one-third of the rise in sorting on test scores and none of the decline in sorting on family background.

The last column of results shuts off the rising value of college and focuses on the new information provided by test scores. The introduction of standardized college admissions tests accounts for little of the rise in college or the national integration of college. It accounts for a substantial portion of the change in sorting patterns: about one-third of the rise in sorting on test scores and the entirety of the decline in sorting

Table 5—Decomposing Model Results

		1960 cohort		
	1933 cohort	Baseline	No test scores	Constant V^c
College attendance	0.29	0.52	0.52	0.29
Local college attendance	0.85	0.51	0.51	0.81
β_s	0.29	0.78	0.46	0.47
β_p	0.67	0.60	0.68	0.61
β_a	0.61	1.08	1.02	0.65
Access to first choice	0.99	0.56	0.55	0.98
Fraction selective	0.08	0.86	0.86	0.12

Notes: Columns compare results from the model for the baseline 1933 calibration, the baseline 1960 calibration, and alternative 1960 equilibria where the information set $\mathcal{I}_t$ or V_t^C is held fixed. Rows display the share of graduates who attend college; the share of college students who attend a local college; and the importance of test scores, family background, and ability for determining who attends college. The last two rows contain moments related to how the model works: the share of students who can attend their first-choice college and the share of colleges that are selective.

on family background. The mechanism is a straightforward information story: when test scores become available, students and colleges' forecasts of student ability put more weight on test scores and less weight on family background.[20]

The model also includes an interaction effect between the two forces in explaining the rise in the importance of test scores for college admissions. The two driving forces interact through selective admissions. For example, consider a high s, low p student. In the 1933 equilibrium, this student is highly unlikely to go to college (only about 10 percent of such students do in the data and the calibrated equilibrium). Rising V_t^C makes this student more likely to attend college. Allowing the student to observe s makes them more likely to attend college because it raises their expected ability and hence their expected gains from college. The interaction between the two comes through the college choice set: as other high-ability students become more likely to go to college, college quality at top colleges rises, which makes college yet more attractive to this student.[21]

We can use the model to infer the sources of the rise in sorting on academic ability documented in Table 3. The surprising conclusion is that although sorting by test scores and ability both increased, they did so for very different reasons. While sorting by test scores is explained roughly equally by the rising value of college, new information, and the interaction between the two, sorting by ability is explained entirely by the rising value of college. The intuition for this finding relies on the fact that we find test scores to be quite noisy signals. The introduction of this noisy

[20] Since ability and all available signals are jointly distributed as a multivariate normal, we can characterize in closed form the constant effect of a one standard deviation increase in each signal on expected ability. Using this, we find that the effect of higher p declines from 0.15 to 0.13 between 1933 and 1960, while the effect of z declines from 0.76 to 0.66. The effect of test scores goes from 0 (by construction) to 0.23, helping generate the reversal.

[21] Similarly, low s, high p students become less likely to attend college because the best colleges set admissions criteria that they cannot meet. While they can attend college (14 percent of colleges are below capacity and have nonselective admissions even in the 1960 equilibrium), the quality of colleges available to them is sufficiently low that they choose not to.

Table 6—Model Results with Constant College Quality

		Model	
	Data	Baseline	Constant quality
College attendance	0.24	0.22	0.23
Local college attendance	−0.34	−0.34	−0.29
β_s	0.48	0.49	0.38
β_p	−0.21	−0.07	0.05
Access to first choice	-	−0.44	−0.51
Fraction selective	-	0.78	0.48

Notes: Columns compare results from the data (where available), the baseline model, and an alternative, recalibrated model where education quality is held fixed at the 1933 level. The rows give the difference in each moment $m_{1960} - m_{1930}$, where the moments m are the share of graduates who attend college, the share of college students who attend a local college, the importance of test scores and family background for determining who attends college, the share of students who can attend their first-choice college, and the share of colleges that are selective.

signal leads students to become more sorted on test scores, including the noise in test scores, but has little impact on the sorting by ability.

Finally, we show that the stratification of colleges is an important mechanism for our results. Figure 9 showed that the calibrated model generates a quantitatively important stratification. Here, our goal is to show that this stratification is an important ingredient for generating the reversal in attendance patterns. To do so, we consider a version of the model where college quality is held fixed over time, meaning that it is both exogenous and independent of mean student ability. We recalibrate the model parameters to fit the targets as well as possible.

We study the results in Table 6, which gives results for the data (where available), the baseline model, and the recalibrated model with fixed college quality. The model where quality is fixed can generate a rise in college attendance through V_t^C, but it falls somewhat short on the decline in local college attendance. Most importantly, it can fit only three-fourths of the increase in the importance of test scores, and it generates an increase rather than a decrease in the importance of family background. It struggles on both dimensions because it shuts down the mechanism of worsening college quality available to low-ability students that generates this reversal. We conclude that there is a strong link between our empirical findings and those previously documented by Hoxby (2009).

V. Extensions

In this section, we consider two extensions to the baseline model. First, we allow for growth in the availability of scholarships. Second, we consider changes in the pool of high school graduates. A sensitivity analysis is available in online Appendix D.

A. *Scholarships*

Our baseline analysis fits the data well in most dimensions, but it generates a decline in the importance of family background for college attendance that is one-third ($= 0.07/0.21$) the size of the actual decline. This result leads us to ask whether the baseline model is missing any features that might allow the model to better fit the data. One important, plausible candidate is the expansion of scholarships and other forms of financial assistance for students.[22] This candidate is plausible because scholarships and other similar programs became viewed as useful tools in the search for talent in the late 1950s. For example, the National Merit Scholarship Program (now combined with the PSAT) was established and scholarships were awarded starting in 1955. Scholarships are important in the context of the model if they are targeted toward students of particular academic abilities or family backgrounds.

We review the evidence on the availability and distribution of scholarship awards in online Appendix D. We combine data from several sources to document two main features of scholarship programs. First, scholarships were growing, from negligible amounts before World War II to 8.4 percent of college income/expenses around 1960. Second, scholarships were targeted, being more likely to go to students with below-median family income or higher test scores (online Appendix Figure D2).

Given these promising findings, we experiment with including scholarships in the model. Scholarships affect the distribution of p, which represents income and transfers minus tuition payments. Relative to the baseline model, we replace the distribution of p in 1960 with an altered one, $\log(p') = \log(p + \iota g(s,p))$. We take $g(s,p)$ to be a piecewise linear function that governs how scholarships vary with student characteristics and ι to be a parameter that governs their level. We recalibrate the model, adding as new moment the distribution of scholarships by s and p described above as well as the share of scholarships in overall student income.

Table 7 shows the results of this experiment in the "scholarships" column. The main result is that introducing scholarships allows the model to generate larger changes in sorting patterns. The model actually overshoots slightly in terms of sorting on test scores, while it generates a larger share of the decline in sorting on family background: 57 percent, as compared to 33 percent in the baseline model. The larger change in sorting turns out to be driven by the fact that scholarships are targeted toward students with low family income. On the other hand, targeting scholarships toward students with high test scores actually works in the opposite direction, because test scores are positively correlated with family income.[23] The remaining model results are affected little, if at all.

B. *Time-Varying High School Graduation Patterns*

For our baseline analysis, we assume that the distribution of students $F(a,p)$ is the same for both cohorts, and we calibrate the correlation parameter ρ between a and p to fit the observed distribution $F(s,p)$ as well as possible in the two cohorts. We do so because most of our empirical studies from Section II, including notably Project Talent, concern the college-going behavior of high school graduates—the most common study design involves surveying students shortly before high school

[22] We are indebted to an anonymous referee for this suggestion.
[23] Results that isolate each dimension are available upon request.

Table 7— Model Results with Extensions

	Data	Model: Baseline	Model: Scholarships	Model: Time-varying graduation
College attendance	0.24	0.22	0.22	0.22
Local college attendance	−0.34	−0.34	−0.35	−0.33
β_s	0.48	0.49	0.52	0.48
β_p	−0.21	−0.07	−0.12	−0.04
Access to first choice	-	−0.44	−0.28	−0.44
Fraction selective	-	0.78	0.80	0.78

Notes: Columns compare results from the data (where available), the baseline model, a model augmented to allow for scholarships, and a model augmented to allow time variation in the composition of students who graduate high school. The rows give the difference in each moment $m_{1960} - m_{1930}$, where the moments m are the share of graduates who attend college, the share of college students who attend a local college, the importance of test scores and family background for determining who attends college, the share of students who can attend their first-choice college, and the share of colleges that are selective.

graduation. However, the expansion of high school over this period raises the concern that changes in the set of students who graduate high school may contribute to or confound the reversal in sorting patterns that we document.

To make progress on this question, we need information on the selection of high school graduates over time. Fortunately, Updegraff (1936) is a rare example of a historical study with extra information. It records outcomes for all students with at least a sixth-grade education, which we take to cover all students. As noted above, similar data do not exist for Project Talent. Instead, we explore using data from the NLSY79. Since this is a later cohort with a higher high school graduation rate than Project Talent, we hypothesize that substituting NSLY79 data overstates the importance of rising high school graduation rates and changing high school graduate composition.

We recalibrate the model. We now choose ρ to fit the observed distribution over (s,p) quartiles for all students (not just high school graduates) in these two cohorts and explicitly feed in the high school graduation rate by (s,p) quartiles for each cohort, measured from Updegraff (1936) and the NLSY79. The rest of the calibration procedure remains the same. We study the results in the column labeled "time-varying graduation" in Table 7. The model captures slightly less of the change in sorting patterns, but overall we conclude that variation in who graduates high school does not have a first-order effect on our results. The underlying intuition is that the model already fits college attendance conditional on (s,p); changing somewhat the distribution of students across cells $F(s,p)$ has second-order effects on our results.

VI. Conclusion

This paper documents large changes in the patterns of college attendance in the United States during the twentieth century. Prior to World War II, family income and socioeconomic status were more important predictors of who attended college, whereas academic ability was more important afterward. We provide a quantitative theory of these changes: rising demand for college and the introduction of standardized test scores generated the increasing college stratification documented by Hoxby (2009), which made college more attractive to high-ability students but less attractive to low-ability students.

Our theory includes a stylized model of labor markets to focus attention on changes in the nature of college. Enriching the model of labor markets is a promising avenue for future work. Doing so would make it possible to quantify the effect of changing college attendance patterns on the distribution of wages and income. It would also permit discussion of possible feedback effects from the labor market on college attendance decisions. For example, existing work suggests that there was an increasing role for ability and a decreasing role for family background in determining labor market outcomes (Herrnstein and Murray 1994; Comerford, Rodríguez Mora, and Watts 2017; Hsieh et al. 2019). Such changes could indirectly affect the incentives of different types of students to pursue college. They could also change the incentives of endowment-minded colleges to accept students of different types.[24]

Our analysis stops in 1960 for two reasons. First, the federal government introduced and expanded college grant and loan programs after this time, rendering our assumption of self-financing college unpalatable. The subsequent period is better thought of in a framework such as Lochner and Monge-Naranjo (2011), where access is affected by a race between the expanding generosity of federal loan programs and rising college tuition. Second, the reversal in sorting patterns appears to be complete by this time, with Belley and Lochner (2007) showing that the trend even reversed for later cohorts. Increased demand for college and college stratification appear to have affected other margins for this later era, including the college preparatory behavior of high school students and the amount of time parents spend with their children (Bound, Hershbein, and Long 2009; Ramey and Ramey 2010; Blandin and Herrington 2020).

REFERENCES

Alon, Titan. 2018. "Earning More by Doing Less: Human Capital Specialization and the College Wage Premium." Unpublished.

Altonji, Joseph G., and Rebecca M. Blank. 1999. "Chapter 48—Race and Gender in the Labor Market." In *Handbook of Labor Economics*, Vol. 3, edited Orley C. Ashenfelter and David Card, 3143–3259. Amsterdam: North-Holland.

Belley, Philippe, and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital* 1 (1): 37–89.

Blandin, Adam, and Christopher Herrington. 2020. "Family Heterogeneity, Human Capital Investment, and College Attainment." https://drive.google.com/file/d/12fZXyNij_8nwDLNfxbWdg_tLIfvGA4xL/view. Unpublished.

Bound, John, Brad Hershbein, and Bridget Terry Long. 2009. "Playing the Admissions Game: Student Reactions to Increasing College Competition." *Journal of Economic Perspectives* 23 (4): 119–46.

Bound, John, and Sarah Turner. 2002. "Going to War and Going to College: Did World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans?" *Journal of Labor Economics* 20 (4): 784–815.

Bowles, Frank H. 1967. *The Refounding of the College Board, 1948–1963—An Informal Commentary and Selected Papers*. New York: College Entrance Examination Board.

[24] We thank an anonymous referee for this idea.

Bureau of Labor Statistics. 1979–2016. "National Longitudinal Survey of Youth, 1979." NLS Investigator. https://www.nlsinfo.org/investigator/pages/login.jsp (accessed June 9, 2015).

Castro, Rui, and Daniele Coen-Pirani. 2016. "Explaining the Evolution of Educational Attainment in the United States." *American Economic Journal: Macroeconomics* 8 (3): 77–112.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017. "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." NBER Working Paper 23618.

Comerford, David, José V. Rodríguez Mora, and Michael J. Watts. 2017. "The Rise of Meritocracy and the Inheritance of Advantage." Unpublished. https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=SED2017&paper_id=1644.

Cooper, Russell, and Huacong Liu. 2019. "Mismatch in Human Capital Accumulation." *International Economic Review* 60 (3): 1291–1328.

Donovan, Kevin, and Christopher Herrington. 2019. "Factors Affecting College Attainment and Student Ability in the U.S. since 1900." *Review of Economic Dynamics* 31: 224–44.

Duffy, Elizabeth A., and Idana Goldberg. 1998. *Crafting a Class—College Admissions and Financial Aid, 1955–1994*. Princeton, NJ: Princeton University Press.

Dunsmoor, Clarence C., and Oliver C. Davis. 1951. *How to Choose that College—A Guide for Students and Parents*. Boston: Bellman Publishing.

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg. 2017. "A General Equilibrium Analysis of State and Private Colleges and Access to Higher Education in the U.S." *Journal of Public Economics* 155: 164–78.

Flanagan, John C., Frederick B. Davis, John T. Dailey, Marion F. Shaycoft, David B. Orr, Isadore Goldberg, and Clinton A. Neyman Jr. 1964. *The American High-School Student: The Identification, Development, and Utilization of Human Talents*. Pittsburgh, PA: Project TALENT Office, University of Pittsburgh.

Flanagan, John C., Marion F. Shaycoft, James M. Richards Jr., and John G. Claudy. 1971. *Five Years after High School*. Pittsburgh, PA: American Institutes for Research and University of Pittsburgh. http://www.projecttalent.org/wp-content/uploads/2017/02/Five_Years_After_High_School_(1971).pdf.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race between Education and Technology*. Cambridge, MA: Harvard University Press.

Greenleaf, Walter J. 1929. *Self-Help for College Students*. Washington, DC: US Government Printing Office.

Halle, Rita S. 1934. *Which College?* New York: Macmillan.

Harris, Seymour Edwin. 1962. *Higher Education: Resources and Finance*. New York: McGraw-Hill.

Haven, Elizabeth W., and Dwight H. Horch. 1972. *How College Students Finance Their Education: A National Survey of the Educational Interests, Aspirations, and Finances of College Sophomores in 1969–70*. New York: College Entrance Examination Board.

Hendricks, Lutz, Christopher Herrington, and Todd Schoellman. 2021. "Replication data for: College Quality and Attendance Patterns: A Long-Run View." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E115791V1.

Hendricks, Lutz, and Todd Schoellman. 2014. "Student Abilities during the Expansion of US Education." *Journal of Monetary Economics* 63: 19–36.

Herrnstein, Richard J., and Charles Murray. 1994. *The Bell Curve—Intelligence and Class Structure in American Life*. New York: Free Press.

Hollis, Ernest V. 1957. *Costs of Attending College: A Study of Student Expenditures and Sources of Income*. Washington, DC: US Government Printing Office.

Hoxby, Caroline M. 2009. "The Changing Selectivity of American Colleges." *Journal of Economic Perspectives* 23 (4): 95–118.

Hoxby, Caroline M., and Bridget Terry Long. 1999. "Explaining Rising Income and Wage Inequality among the College-Educated." NBER Working Paper 6873.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87 (5): 1439–74.

Iffert, Robert E., and Betty S. Clarke. 1965. *College Applicants, Entrants, Dropouts*. Washington, DC: US Department of Health, Education, and Welfare.

Karabel, Jerome. 2006. *The Chosen—The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Boston: Mariner Books.

Kobrin, Jennifer L., Brian F. Patterson, Sandra M. Barbuti, Krista D. Mattern, and Emily J. Shaw. 2008. *Validity of the SAT for Predicting First-Year College Grade Point Average*. New York: College Board.

Kurani, Habib A. 1931. *Selecting the College Student in America: A Study of Theory and Practice*. New York: Teachers College, Columbia University.

Lansing, John B., Thomas Lorimer, and Chikashi Moriguchi. 1960. *How People Pay for College*. Ann Arbor, MI: Survey Research Center, University of Michigan.

Lochner, Lance J., and Alexander Monge-Naranjo. 2011. "The Nature of Credit Constraints and Human Capital." *American Economic Review* 101 (6): 2487–2529.

Morgan, Rick. 1989. *Analysis of the Predictive Validity of the SAT and High School Grades from 1976 to 1985*. New York: College Entrance Examination Board.

National Center for Education Statistics. 1962–2013. *Digest of Educational Statistics*. Washington, DC: US Department of Health, Education, and Welfare, Office of Education.

Partridge, E.A. 1925. "College-Entrance Credentials from the High-School Point of View." *School Review* 33 (6): 422–27.

Pryor, John H., Sylvia Hurtado, Victor B. Saenz, José Luis Santos, and William S. Korn. 2007. *The American Freshman: Forty Year Trends*. Los Angeles: Higher Education Research Institute, UCLA.

Ramey, Garey, and Valerie A. Ramey. 2010. "The Rug Rat Race." *Brookings Papers on Economic Activity* 40 (1): 129–76.

Restuccia, Diego, and Guillaume Vandenbroucke. 2014. "Explaining Educational Attainment across Countries and over Time." *Review of Economic Dynamics* 17 (4): 824–41.

Roper, Elmo Burns. 1949. *Factors Affecting the Admission of High School Seniors to College: A Report for the Committee on a Study of Discriminations in College Admissions*. Washington, DC: American Council on Education.

Taubman, Paul, and Terence Wales. 1975. "Mental Ability and Higher Educational Attainment in the Twentieth Century." In *Education, Income, and Human Behavior*, edited by F. Thomas Juster, 45–70. Cambridge, MA: NBER.

United States Office of Education. 1918–1958. *Biennial Survey of Education in the United States*. Washington, DC: US Government Printing Office.

Updegraff, Harlan. 1936. *Inventory of Youth in Pennsylvania*. Washington, DC: American Youth Commission of the American Council on Education.

Wechsler, Harold S. 1977. *The Qualified Student—A History of Selective College Admission in America*. New York: John Wiley and Sons.