Late Bloomers: The Aggregate Implications of Getting Education Later in Life*

Zsófia L. Bárány[◇] Moshe Buchinsky[○] Pauline Corblet[⋆]

October 27, 2023

Abstract

It is generally agreed upon that most individuals who acquire a college degree do so in their early 20s. Despite this consensus, we show that in the US from the 1930 birth cohort onwards a large fraction – around 20% – of college graduates obtained their degree after age 30. We explore the implications of this phenomenon. First, we show that these so called late bloomers have significantly contributed to the narrowing of gender and racial gaps in the college share, despite the general widening of the racial gap. Second, late bloomers are responsible for more than half of the increase in the aggregate college share from 1960 onwards. Finally, we show that the returns to having a college degree vary depending on the age at graduation. Ignoring the existence of late bloomers therefore leads to a significant underestimation of the returns to college education for those finishing college in their early 20s.

Keywords: educational choice, late bloomers, college share, wage returns

JEL codes: I21, I26, J31

*The authors thank participants at the CEU Brownbag Seminars, Mannheim Applied Seminars, TU Wien Economics seminar, MEF/ECONtribute-Seminar, University of Bonn, AFSE, and ArmEA Meetings for helpful comments. This paper was previously part of a larger project on late bloomers.

◇Central European University & CEPR, baranyzs@ceu.edu

○Sciences Po & NBER, moshe.buchinsky@sciencespo.fr

⋆New York University Abu Dhabi,pzc208@nyu.edu

1 Introduction

The importance of returning to school later in life is generally understood. Yet, there is a widely held assumption – in economics and elsewhere – that most individuals go to post-secondary education right after high school and finish their education in their early 20s. As a consequence, most papers in economics simply assume that all college graduates follow the standard route of finishing college right after high school without any delays (i.e., Mincer (1974), Becker (1994), Keane and Wolpin (1997), Lee and Wolpin (2006)). While the majority of college graduates indeed follows the standard route (referred to as early college graduates), we show that around a fifth of college graduates complete their college education later in life (referred to as late bloomers). We devote this paper to documenting in detail that a large fraction of individuals return to education later in life. We examine the contribution of late bloomers to the aggregate increase in the share of college graduates and discuss several estimation issues that arise when one assumes away this phenomenon in empirical research.

First, using the decadal American Census combined with the yearly American Community Survey (ACS) and the two samples from the National Longitudinal Survey of Youth (NLSY79 and NLSY97) we document that the continued education of individuals into their 50s is a robust, persistent, and wide-spread phenomenon. For example, at age 24 only 10.2% of the 1936 birth cohort had a college degree, by age 64 this fraction went up to 18.9%. The equivalent numbers for the 1966 birth cohort are 19.8% at age 24 and 31.5% at age 64. This implies that about a third of college graduates by age 64 completed their degree after age 24. Looking at even later bloomers, among individuals born between 1930 and 1970: around 20% of those who graduated from college by age 50 did so after age 30.

Second, we show that the share of late bloomers among college graduates is not constant, but changes non-monotonically across birth cohorts. The share of late bloomers and its evolution differs by gender and by race. Specifically, the share of late graduates is higher among women than among men and is higher among the Black and the Hispanic population than among Whites. We also analyze the contribution of early college graduates and of late bloomers to the changing gender and the racial college share gap across birth cohorts. While the gender gap narrowed and the racial gap widened for early college graduates, late bloomers have contributed to narrowing both the gender and the racial college share gap.

It is a well-known fact that the educational attainment in the population of all advanced countries has been increasing over time. In the United States, the share of individuals with a college degree increased from 8.4% to 34.5% between 1960 and 2019. The literature explains this phenomenon with the *cohort succession model* (see e.g. Ryder (1965), Mare (1995), Cheeseman Day and Bauman (2000)). According to this model, the increase in the share of college graduates stems from successive cohorts obtaining more education. Our third contribution is to decompose the aggregate increase in educational attainment over time into a between-cohort (i.e., cohort succession) and a within-cohort (i.e., late bloomers) component. This decomposition shows that since 1960 more than half of the aggregate increase in the college share comes from the within-cohort component and that this share has been increasing over time.

Finally, we show that the returns to having a college degree differ depending on the age at graduation. Late bloomers – even before graduating from college – earn higher wages than non-graduates. Upon graduation, they receive a college premium, and their wage-experience profile becomes steeper, but their wages stay below the wages of early college graduates. This implies that when conducting regressions failing to account for the exact time of graduation, would lead to biased estimates. We demonstrate that ignoring the existence of late graduates, and assuming that all college graduates follow the standard route, leads to an underestimation of the returns to graduating from college early on average by 27%. The extent of underestimation varies across demographic groups.

Our findings have important implications. The shift-share decomposition shows that to understand the evolution of the aggregate college share, one must pay particular attention to late bloomers. The fact that the share of late bloomers among college graduates changes non-monotonically across birth cohorts suggests that the forces that drive early college graduates and those that drive late bloomers are different. Therefore, to apprehend the drivers of the college share in the aggregate economy, we need to study the forces that drive individuals to acquire education later in life separately from those that drive early college graduation. This need is even more pressing due to labor market returns varying between early and late college graduates.

The fact that the returns to having a college degree are lower for late graduates suggests caution when interpreting college graduation gaps between different demographic groups. In particular, a narrowing of the gap driven by early graduates does not have the same implication as a narrowing of the gap driven by late graduates. Given that we found that late bloomers contribute to the narrowing of both the racial and gender gap, the college graduation gap should be calculated at different ages, and studied jointly.

Our paper is not the first to document that continued education in adulthood is a prevalent phenomenon. Wilson et al. (2011) analyze the processes that explain the trends in average educational attainment. They conclude that the importance of between-cohort changes has decreased and that this was compensated to some extent by an increase in within-cohort changes, especially for women. Bauman (2016) indicates that indeed the CPS data is not consistent with individuals only acquiring education during their childhood, teens, and 20s. However, he does not explore how the fraction of late bloomers has changed over time, their contribution to aggregate increases in college attainment, or the implication for estimates of the gain from having a college degree. Murphy and Topel (2016) document similar patterns for the within-cohort increases in the fraction of college graduates for both women and men in their Figure 5 but do not explore its implications. Altonji et al. (2016) remark that about a quarter of students in their data graduate before 20 or after 24, and exclude them from their sample.

The impact of the timing of schooling on subsequent labor market outcomes has been studied in the literature. Griliches (1980) studies the effect of school interruption and working while in school in the Young men NLS and finds it does not affect subsequent careers. Light (1995) measures returns to education by allowing for discontinuity in age paths before and after school, and finds that education wage boosts are smaller when there has been a discontinuity. Monks (1997) also shows that those who complete college at a later age benefit from smaller returns to education. Ferrer and Menendez (2009) compare the returns to education between the early and late graduates of tertiary education using a unique survey data set from Canada. Contrary to our findings, they find that graduates who delayed their education enjoy a premium relative to graduates who did not.

While our paper does not assess the reasons for returning to college later in life, it suggests that these reasons need to be investigated. The literature addressing this is limited, even though adult education in general is the topic of research in education and psychology. Kuh et al. (2006) provide an overview of the literature on individuals' success in college. They do provide ideas about who are the individuals who may have to drop out of college as well as who might be likely to return to college. Aina et al. (2018) also provide a comprehensive survey on the theoretical approaches used in the literature to study the phenomenon of delayed graduation and university dropout. Zhang and Palameta (2006) study the likelihood of adult schooling and its returns based on demographic characteristics in Canada and find that young men and senior women benefit the most. Jamieson (2007) examines the behavior of older students in the United Kingdom. She finds that a substantial proportion of the 'middle-aged' graduates view their studies as a pathway to enhanced employment opportunities. As noted in Anderson (2008), there is some evidence of the benefits of later life learning to several important elements such as mental health and improved social networks. Yang (2021) uses the NLSY79 data to analyze the socio-economic characteristics that are correlated with the age at which an individual graduates from college. He then constructs an incomplete markets general equilibrium life-cycle model with which he evaluates the welfare implications of the possibility of delayed college education.

The remainder of the paper is organized as follows. In Section 2 we document the existence of 'late bloomers' in the US. This is further expanded upon in Section 3, where we analyze the share of late bloomers among college graduates and their contribution to the changing gender and racial gap in college share. Section 4 is devoted to a decomposition of the increase in the aggregate college share into a between cohort and a within cohort component. In Section 5 we analyze whether the returns to having a college degree differ depending on the age at graduation. Finally, Section 6 provides a summary and concludes.

2 Late bloomers

In what follows we document that the continued education of individuals late into their lives is a robust, persistent, and widespread phenomenon. We first use repeated cross-sectional data – the decadal American Census from 1960-2000 combined with the yearly American Community Survey (ACS) 2001-2019 – to show that the share of college-educated individuals increases within birth cohorts as they age. These data were downloaded from IPUMS (Ruggles et al. (2023)).[1] Next, we use panel data sets – the 1979 and the 1997 releases of the National Longitudinal Survey of Youth (NLSY) – which allows us to follow individuals over time. In these surveys we establish that a large share of individuals obtain college education later in life, that is they are late bloomers.

Figure 1 plots the share of college-educated individuals in the population for different birth cohorts as they age. College education is defined as 16 years or more of completed years of education. Each curve shows the share of college-educated individuals between ages 24 and 64 for one of seven birth cohorts, born every decade from 1926 until 1986. As the data runs from 1960 until 2019 we observe the later-born cohorts for fewer years, i.e. until a younger age.

There are three things to note from this figure. First, the educational attainment is higher for later-born cohorts. Second, the age gradient of the share of college graduates is positive for each cohort, implying that the share of college graduates within a cohort increases as the cohort gets older. Third, the age gradient of the college share is steeper for later-born cohorts.[2] The first observation is expected based on the cohort succession model. Even though there are some research papers documenting the second observation, it is a little-known fact. Finally, to the best of our knowledge, we are

[1] The same holds using the Annual Social and Economic (ASEC) supplement of the Current Population Survey (CPS). We opted to use the decadal Census combined with the ACS due to the larger samples.

[2] These three patterns also hold when breaking down the data into gender and race cells, see Figure B.1 and B.2 in the Appendix.

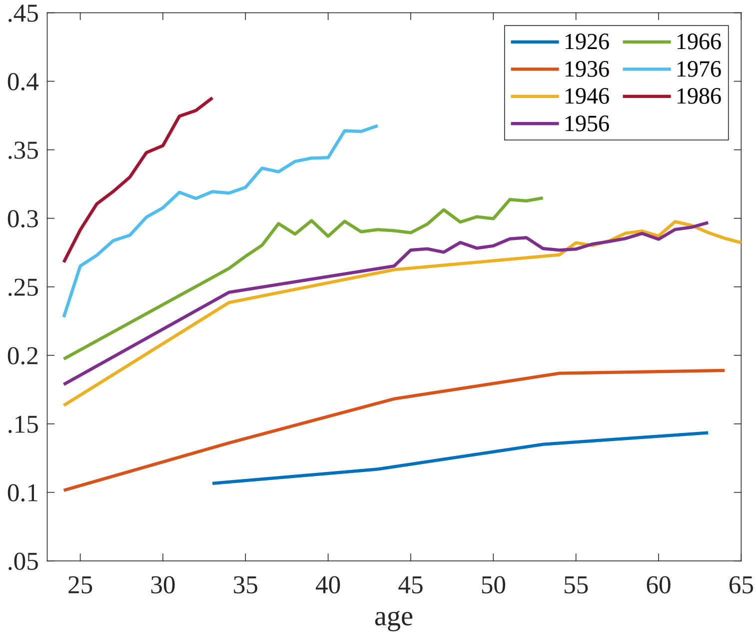


Figure 1: College share by age for different cohorts

Notes: Authors' calculations from the decadal Census for 1960-2010 and ACS data for 2001-2009 and 2011-2019. Each curve shows the share of college graduates for a birth cohort against the age of the birth cohort. The college share for a birth cohort in a given year is calculated as the share of the total sample weight of those with at least 16 years of completed years of schooling relative to the total sample weight of the birth cohort in the given year. The curves show cohorts born in 1926, 1936, 1946, 1956, 1966, 1976, 1986 for ages 24-65 (or widest range available within).

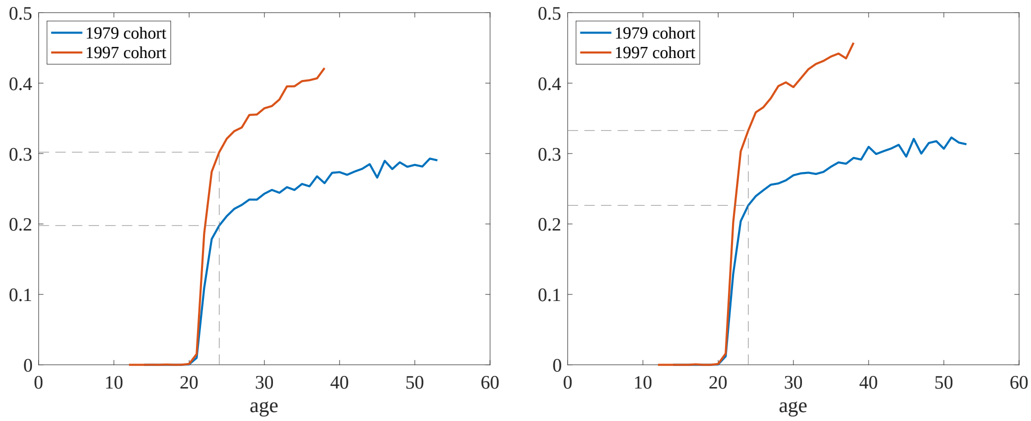
the first to document the third observation. We next investigate whether these patterns hold by gender and by race.

The most natural explanation for new observations 2 and 3 is that in all birth cohorts individuals continue to acquire college degrees in their 30s, 40s, and 50s and that the rate at which individuals get college degrees later in life increases across successive cohorts. While the repeated cross-sectional nature of the Census does not allow us to verify whether individuals acquire education later in life, the available information on enrollment in school is in line both with the increases in the college share within birth cohort over age and with the increases in the age-gradient across birth cohorts. See Appendix section B.2 for details.

There are other potential explanations for our new observations. First, it is possible that the fraction of individuals with a college degree among immigrants is higher for older individuals, and that the age-gradient of the college share of immigrants increased over time. We can rule this explanation out, as the same patterns hold when we only consider individuals born in the US, see Figure B.4 in the Appendix. A second explanation could be that there are differences in the mortality rate of individuals with and without a college degree. The available data shows that the mortality rate is higher among individuals without a college degree and that the difference is larger for older individuals. These facts imply that as a given birth cohort ages, a larger share of non-college individuals in that cohort die, and hence the college share within the birth cohort increases. However, the available data suggests that the impact of mortality differences by education level on the evolution of the college share of a birth cohort is tiny. See Appendix B.4 for details. A third explanation is that perhaps the survey (non-)response rates are higher for those without a college degree and this difference increases within a cohort as it ages and is also larger for later-born cohorts. By definition, we cannot rule this explanation out from repeated cross-sectional data.

The best way to rule out alternative explanations is to use panel data, which follows individuals over time. The two releases of the NLSY follow individuals from a young age, between 14 and 22 in the 1979, and between 12 and 15 in the 1997 release. The NLSY79 is annual until 1994, and bi-annual until 2020, while the NLSY97 is annual until 2011 and bi-annual after that. These surveys provide information on school enrollment, highest grade completed, employment, and wages, as well as a rich set of descriptors. As in the Census data, we define college graduates as those who completed at least 16 years of schooling.

In Figure 2 each curve represents individuals surveyed in one release of the NLSY. The blue curve, corresponding to the NLSY79, shows individuals born between 1957 and 1964, while the red curve, corresponding to the NLSY97, shows individuals born between 1980 and 1984. The left panel shows the college share at each age among all individuals surveyed at that age, while in the right panel the share at each age is calculated on a fixed sample, keeping only individuals who responded in all rounds of the survey. These graphs show that (1) in the NLSY97, individuals acquire more education, (2) in both NLSY79 and NLSY97, individuals continue to acquire education as long as the survey follows them, and (3) the age-gradient of the NLSY97 is steeper. Since the left and the right panel are similar we can rule out that the age-gradient in the college share is driven by differential survey non-response rates by education level.



(a) Full sample (b) Stayer sample

Figure 2: College share by age in the two NLSY cohorts

Notes: Authors' calculations from the two releases of the NLSY. In each panel the two curves show the share of college graduates against age for individuals from the NLSY79 (in blue) and the NLSY97 (in red). The left panel shows this in the full sample, whereas the right panel shows these shares among those who responded in all rounds of the respective survey.

As in the Census, the same patterns hold for both men and women and all race and ethnicity groups, see Figure B.6 and B.7 in the Appendix. Therefore the facts that we established based on the Census data also hold in the NLSY and are due to individuals acquiring education later in life.

Thus far we have established that a large fraction of individuals are late bloomers, who acquire education late in their life. Figure 2 demonstrates that an increasing share of individuals acquire education later in life (the NLSY97 has a steeper slope than the NSLY79), and that also an increasing share of individuals acquire education early in life (the NLSY97 has a higher intercept, or age 24 share). Given the rising shares of both early and late graduates across birth cohorts, it is not clear whether these shares are shifting simultaneously or at different rates. If the latter is true, then it implies that the factors driving early and late graduation are different, which is the reason why it is crucial to gain insight into this question. This is what we investigate in the next section.

3 Share of late bloomers among all college graduates

Table 1 shows – for different birth cohorts in each row, and for different subgroups of the cohort in each column – the share of those who get a college education after 30 among those who get a college education by age 50.[3] We chose the cutoff age of 50 to be able to include the 1970 cohort. The patterns are the same if we set the cutoff age at 60. If the share of early and late graduates changed together we would expect

cohort	all	men	women	Whites	Blacks
1930	22.6	20.9	27.7	21.5	40.0
1940	28.1	25.6	31.8	26.2	48.9
1950	17.3	12.2	23.3	16.3	32.4
1960	16.5	12.3	20.3	14.8	34.3
1970	18.7	18.1	18.9	18.1	37.0

Table 1: Share of late college graduates among college graduates, Census

Notes: Authors' calculations from the decadal Census for 1960-2010 and ACS data for 2019. The table shows the share of those who obtained a college degree after the age of 30 among those who obtained a college degree by age 50 for different groups of the population in each column and for different birth cohorts in each row.

these shares to be roughly constant. This is clearly not the case. Table 1 shows that the share of late bloomers among college graduates decreased monotonically until the 1960 birth cohort and slightly increased thereafter. Among men and Blacks, the increase started from the 1950 birth cohort, for Whites from the 1960 birth cohort, whereas for women the share decreased monotonically across these 5 birth cohorts. The highly non-monotonic evolution of the share of late graduates among all college graduates suggests that distinct forces are driving early and late graduation, while the different evolution of this share across demographic groups suggests that these forces are of different importance depending on the demographic group.

Table 2 illustrates a similar trend using the NLSY data. The respondents of the NLSY79 were between 55 and 63 years old in 2020 during the last round of the survey. The respondents in the NLSY97 were between 30 and 40 years old during the last year of that survey, administered in 2019 and 2020. To get comparable shares of late bloomers, we truncate the NLSY79 in 1998 when the respondents were between the ages of 31 and

[3]See Table C.1 in the Appendix for the share of graduates at different ages for these birth cohorts.

sample	all	men	women	Hispanics	Blacks	others
1979	16.7	12.6	20.6	27.3	30.0	14.9
1979 truncated	9.3	7.5	11.0	15.6	13.6	8.7
1997	12.0	12.3	11.8	20.1	22.4	9.8

Table 2: Share of late graduates among college graduates, NLSY

Notes: Authors' calculations from the NLSY79 and the NLSY97. The table shows the share of those who obtained a college degree at or after the age of 30 among those who obtained a college degree during the 1979 survey in the top row (i.e., by 2020), during the truncated 1979 survey in the middle row (i.e., by 1998), and in the 1997 survey in the bottom row (i.e., by 2019) for different groups of the population in each column.

41. Importantly the age distributions in the last year of the truncated NLSY79 and that of the NLSY97 are similar, with a mean age of around 36. Table 2 shows the share of those who obtained a college degree at or after age 30 among those who obtained a college degree at any point during the survey for different samples in each row and different subgroups of the population in each column.[4] Given the birth years of NLSY79 respondents the first row should be–and indeed it is–close to the 1960 row of Table 1. Comparing the first two rows of the table we see that by truncating the NLSY79 in 1998 we have removed around 40-50% of late bloomers for each demographic group. The comparison of the last two rows of Table 2 shows that the share of late bloomers when calculated only until individuals reach their mid-30s has increased between those who were born between 1957 and 1964 and those who were born between 1980 and 1984. The increase is the largest for Blacks and men and is the lowest for women and for the non-Hispanic non-Black population. If these increases in the share of late bloomers measured until their mid-30s are indicative of increases until the ages of 50 and 60, then putting together with Table 1 we expect the share of late bloomers to increase after 1970 in all demographic groups of the population. Moreover, not only the share of late bloomers among college graduates but also the share of college graduates at different ages substantially increased for all demographic groups between the NLSY79 and NLSY97, see Table C.2 in the Appendix.

An implication of the large share of late bloomers – together with its different levels across demographic groups, and its differential evolution across birth cohorts – is that

[4] In the NLSY the 'other' race category includes the non-Hispanic, non-Black individuals, and is close to the 'White' category in the Census.

late bloomers could have played a role in the narrowing or the widening of the college share gap between different demographic groups. In Table 3 we show the college share gap between men and women (also by race) and between Whites and Blacks (also by gender) for different birth cohorts at age 30 and age 50.

	gender gap			White-Black gap		
	all	Whites	Blacks	all	men	women
	college share gap at age 30					
1930	7.31	3.19	-0.01	7.72	7.37	3.38
1940	6.98	2.93	0.00	10.14	9.31	6.37
1950	7.08	2.63	0.01	13.77	12.20	10.09
1960	0.87	-0.54	-0.01	11.28	10.89	10.48
1970	-3.09	-2.45	-0.06	13.82	16.27	13.03
	college share gap at age 50					
1930	8.32	8.83	0.32	8.15	12.53	4.02
1940	8.00	8.52	-0.52	10.53	15.19	6.15
1950	4.63	4.91	0.34	13.06	15.37	10.81
1960	-1.53	-1.36	-4.21	9.00	10.53	7.68
1970	-4.13	-4.31	-8.46	11.13	13.38	9.23

Table 3: College share gap at different ages in percentage points

Notes: Authors' calculations from the decadal Census for 1960-2010 and ACS data for 2019. The table shows the differences in the share of college graduates between men and women (in the first three columns) and between Whites and Blacks (in the second three columns), also broken down by race and gender respectively. The top panel shows these college share gaps measured at age 30 for different birth cohorts in each row, while the bottom shows them measured at age 50.

Looking at the gender gap in college education, the narrowing from the 1930 to the 1970 birth cohort is more pronounced at 12.45 percentage points if measured at age 50 (-4.13-8.32), compared with the 10.4 percentage points at age 30 (-3.09-7.31). The difference in narrowing is especially pronounced if we look at the gender gaps by race: more than half of the narrowing for Whites, and all of the narrowing for Blacks happens after age 30. While the racial college share gap widened across these birth cohorts (though not monotonically), the widening is less pronounced if measured at a later age, implying that the racial college share gap actually narrowed (also by gender) after age 30. Therefore late bloomers narrow both the gender and the racial college share gaps. If late bloomers have lower earnings than early college graduates, then a narrowing of the gap after age 30 has a different implication than its narrowing before age 30. We explore the earning profiles of late bloomers in Section 5.

4 Contribution to aggregate college share increase

Thus, we have established that a significant fraction of individuals acquire a college degree later in their lives and that the share of both early and late graduates is larger for later-born birth cohorts. Therefore, both of these channels must contribute to the increase in the aggregate college share. To quantify the contribution of late-bloomers to the change in the aggregate college share over time, we conduct a standard shift-share decomposition that also accounts for the exit of older cohorts and the entry of younger cohorts.[5]

As a first step, we express the (aggregate) college share in the economy at time t among 24-65-year-olds as

$$\lambda_t = \sum_{b=t-65}^{t-24} \omega_{b,t}\lambda_{b,t},$$

where b indexes the birth cohort, $\lambda_{b,t}$ is the college share among birth cohort b in year t, and $\omega_{b,t}$ is the share of birth cohort b among all birth cohorts who are between the age of 24 and 65 in year t (such that $\sum_{b=t-65}^{t-24} \omega_{b,t} = 1$).

The decadal change in the college share can be written as:

$$\begin{aligned}
\Delta\lambda_t \equiv \lambda_t - \lambda_{t-10} &= \sum_{b=t-65}^{t-24} \omega_{b,t}\lambda_{b,t} - \sum_{b=t-75}^{t-34} \omega_{b,t-10}\lambda_{b,t-10} \\
&= \sum_{b=t-33}^{t-24} \omega_{b,t}\lambda_{b,t} - \sum_{b=t-75}^{t-66} \omega_{b,t-10}\lambda_{b,t-10} + \sum_{b=t-65}^{t-34} \omega_{b,t}\lambda_{b,t} - \sum_{b=t-65}^{t-34} \omega_{b,t-10}\lambda_{b,t-10}.
\end{aligned}$$

This formulation highlights that different birth cohorts are included in the aggregate college share of different years. The cohorts born between $t-65$ and $t-34$ are included in the aggregate college share of both years t and $t-10$. The older cohorts born between $t-75$ and $t-66$ are only included in the college share of year $t-10$, as they retire or

[5]See for example Acemoglu and Autor (2011) for a standard shift-share decomposition of occupational changes into within industry and between industry changes, and Melitz and Polanec (2015) for a discussion of how to treat (firm) entry and exit in shift-share decompositions of aggregate productivity changes.

exit the sample by year t. Similarly, the younger cohorts, born between $t-33$ and $t-24$ are only included in the year t's college share, as in year $t-10$ they were too young to have entered the sample. Based on this logic, it is useful to define the following age groups: young are the current 24-33 year olds, middle-aged are the current 34-55 year olds, and old are the current 56-65 year olds. We define within age-group weights: $\widetilde{\omega}_{b,t}^{m,o} = \omega_{b,t} / \sum_{b=t-65}^{t-34} \omega_{b,t}$ and $\widetilde{\omega}_{b,t}^{y,m} = \omega_{b,t} / \sum_{b=t-55}^{t-24} \omega_{b,t}$, such that the weights in the indicated groups sum to 1, i.e. $\sum_{b=t-65}^{t-34} \widetilde{\omega}_{b,t}^{m,o} = 1$ and $\sum_{b=t-55}^{t-24} \widetilde{\omega}_{b,t}^{y,m} = 1$. We further define the following two within age-group average college shares: one among middle-aged and old $\overline{\lambda}_t^{m,o} = \sum_{b=t-65}^{t-34} \widetilde{\omega}_{b,t}^{m,o} \lambda_{b,t}$, and another among the young and middle-aged $\overline{\lambda}_t^{y,m} = \sum_{b=t-55}^{t-24} \widetilde{\omega}_{b,t}^{y,m} \lambda_{b,t}$. With these definitions, we can further rewrite the decadal change in the college share as:

$$\Delta\lambda_t = \underbrace{\sum_{b=t-33}^{t-24} \omega_{b,t}(\lambda_{b,t} - \overline{\lambda}_t^{m,o})}_{\text{entry}} + \underbrace{\sum_{b=t-75}^{t-66} \omega_{b,t-10}(\overline{\lambda}_{t-10}^{y,m} - \lambda_{b,t-10})}_{\text{exit}}$$
$$+ \underbrace{\sum_{b=t-65}^{t-34} (\widetilde{\omega}_{b,t}^{m,o} - \widetilde{\omega}_{b,t-10}^{y,m})\lambda_{b,t-10}}_{\text{shift}} + \underbrace{\sum_{b=t-65}^{t-34} \widetilde{\omega}_{b,t}^{m,o}(\lambda_{b,t} - \lambda_{b,t-10})}_{\text{within}}. \tag{1}$$

The *entry* component captures changes due to the difference in the college share in year t between new entrants and birth cohorts that were already active in the previous decade, in year $t-10$. The *exit* component captures changes due to the difference in the college share in year $t-10$ between the retiring cohorts and the cohorts that will be active also in the next decade, in year t. The *shift* component is due to changes in the relative weights of birth cohorts that are active in both periods, essentially capturing the gradual shift of the population towards later-born cohorts with a different college share. Given that cohorts born later tend to have higher educational attainment, as emphasized by the cohort succession model, each of these three components should be positive. The sum of these three components captures the between cohort changes and accounts for the cohort succession channel, as the included changes are solely driven by changes in the share of different birth cohorts among the currently active population. Finally, the *within* component is due to changes in the college share within birth cohorts over time, it thus captures changes coming from individuals acquiring education later in life, i.e. from late bloomers.

Table 4 shows the above decomposition by decade, as well as for two 30-year periods. As expected, the share of college-educated workers increased in all the decades considered. The cohort succession mechanism, whereby consecutive cohorts acquire more and more education, has significantly contributed to this overall increase, as captured by the between component. This itself is composed of entry, exit, and shift. New entrants with higher college shares are driving the succession channel in the first half of the sample, whereas retiring cohorts with lower college shares are the largest contributors to the between component in the later periods, and shifts among birth cohorts active in both periods hardly have any effect. The within-birth cohort component, which is the focus of our paper, is an important contributor to the aggregate increase of the college share, with its contribution monotonically increasing from around 40% in 1960-1970 to more than 80% in 2010-2019. Comparing the 1960-1990 period with the 1990-2019 period, the within-cohort increases explain 48% of the change in the earlier period, whereas they explain 72% of the increase in the later period. Thus, late bloomers have been significantly contributing to the increase in the aggregate supply of skilled workers since the 1960s, and their contribution has grown over time.

	aggregate change	entry	exit	shift	between cohort	within cohort
1960-1970	3.5	1.4	0.7	0.1	2.1	1.4
1970-1980	6.1	2.4	0.9	0.1	3.4	2.6
1980-1990	4.5	-0.1	1.7	0.1	1.8	2.7
1990-2000	4.0	0.3	1.3	0.0	1.6	2.4
2000-2010	3.4	0.3	0.8	0.0	1.1	2.3
2010-2019	4.7	0.6	0.0	0.0	0.7	4.0
1960-1990	14.0	3.7	3.3	0.3	7.3	6.7
1990-2019	12.0	1.2	2.1	0.0	3.4	8.6

Table 4: Aggregate college share change decomposition

Notes: Authors' calculations from the decadal Census for 1960-2010 and ACS data for 2019. The decomposition in (1) is implemented using sample weights to calculate the weight of each birth cohort b in year t among those aged 24-65, $\omega_{b,t}$. The college share for birth cohort b in year t is as in Figure 1. The 'between cohort' is calculated as the sum of 'entry', 'exit' and 'shift'. The sum of 'between cohort' and 'within cohort' is equal to the aggregate change. All changes are shown in percentage points.

The same decomposition done separately by gender and by race shows that – while there are quantitative differences – the role of within-cohort increases in educational attainment became more pronounced for all demographic groups. Table 5 shows that for all demographic groups considered the within cohort increases in educational attainment explain more than half of the aggregate increase between 1990-2019.[6] In terms of magnitude there are marked differences, for men and Blacks (almost) everything, for Whites around 70% and for women slightly more than 50% is due to within cohort changes. Table 5 also shows that there are different patterns in terms of changes in aggregate college shares across demographic groups. For men and for Whites the increase in the aggregate college share is smaller in 1990-2019 than in 1960-1990, while for women and Blacks the increase is larger in the second half.

	aggregate change	between cohort	within cohort
		men	
1960-1990	14.3	6.6	7.7
1990-2019	7.1	-1.3	8.4
		women	
1960-1990	13.8	7.8	6.0
1990-2019	16.8	7.9	8.9
		Black	
1960-1990	9.1	2.9	6.1
1990-2019	11.3	0.8	10.5
		White	
1960-1990	14.4	7.9	6.5
1990-2019	11.2	3.2	8.0

Table 5: Aggregate college share change decomposition by gender & race

Notes: Authors' calculations from the decadal Census for 1960-2010 and ACS data for 2019. The decomposition in (1) is implemented separately for different demographic groups. The 'between cohort' is calculated as the sum of 'entry', 'exit' and 'shift'. The sum of 'between cohort' and 'within cohort' is equal to the aggregate change. All changes are shown in percentage points.

To summarize, we have shown that late bloomers explain a large and growing part of the aggregate increase in the college share. These findings imply that late bloomers play an important role in the evolution of the aggregate college share, and we need

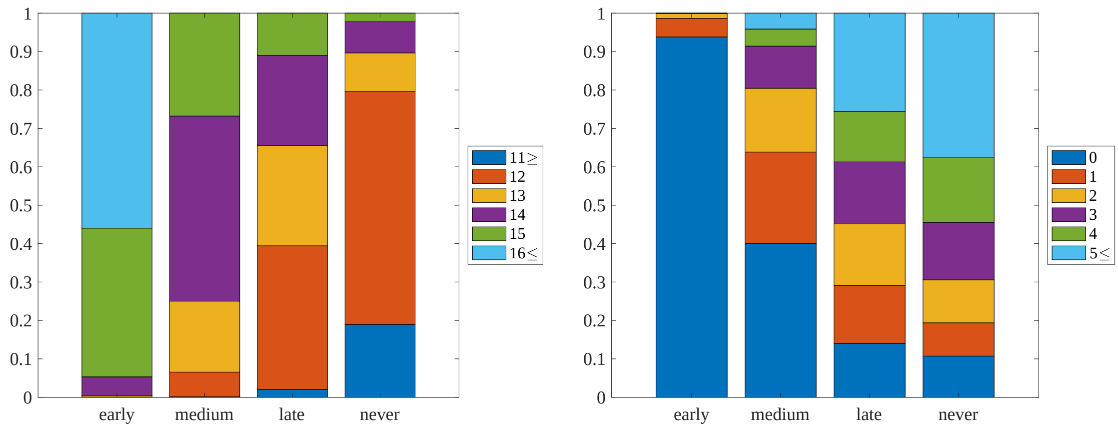
[6] See Table D.3 in the Appendix for the decade-by-decade decomposition.

to study the forces that drive individuals to acquire education later in life separately from those that drive early college graduation.

In what follows we first document that the returns to college education are different depending on the age at graduation. We then show that these differences are important, especially since in most data we do not observe the age at graduation. In such data, the returns to college education are underestimated for the majority of college graduates, who complete their college degree straight after high school.

5 The returns to college education by age at graduation

In this section, we aim to estimate the returns to college education by age at graduation. There are two issues to consider. First, to conduct such an analysis, we need to observe the age at which the individual obtained a college degree. This implies that we need to use the NLSY since this type of information is not available in the Census and more generally in cross-sectional datasets. Second, since the NLSY has a relatively small sample we cannot do a quasi-continuous analysis by age at graduation, but we need to create a few groups based on the age at graduation. We split respondents in the NLSY79 into four groups: (i) early college, those who graduate from college at or before age 24, (ii) medium college, those who graduate from college between age 25 and 29, (iii) late college, those who graduate from college at or after age 30, and (iv) never college, those who do not get a college degree during the time we observe them. While the cutoff ages between these groups are somewhat arbitrary, the groups we construct are markedly different in terms of the highest grade completed (HGC) and in terms of the years of work experience individuals accumulate in their 20s before college graduation. Figure 3 presents the distribution of HGC at age 22 for the 4 groups on the left, and on the right it presents the distribution of pre-college work experience at age 26. At age 22 more than 95% of the early college group have at most one year left to complete college, and almost 75% of the medium college group have at most 2 years left. In complete contrast almost 40% of the late college group and 80% of the never college group did not even start college. In terms of pre-college work experience, more than 90% of the early college group have none. In contrast, about



(a) Highest grade completed at 22

(b) Work experience at 26

Figure 3: Distribution of highest grade completed and work experience

Notes: Authors' calculations from NLSY79 data. The panel on the left shows the distribution of highest grade completed (measured in years) at age 22 among those who end up in the early, medium, late, and never college groups. The right panel shows the distribution of pre-college work experience (in years) at age 26 for the same four groups.

60% of the medium college group have at most 1 year, whereas around 60% of the late and 70% of the never college group have 3 years or more work experience at age 26.[7]

A standard way to assess the lifetime returns to college education is to run separate wage-experience regressions for those with and without a college degree. This is done for instance in the literature on unlucky cohorts and the scarring effect to compare the outcomes of post-secondary education graduates and less-educated workers (Genda et al. (2010), Brunner and Kuhn (2014), Schwandt and von Wachter (2019)), or more generally to measure the returns to experience by education level (Connolly and Gottschalk (2006)). The equation is as follows:[8]

$$\log w_i = \alpha + \beta_1 \text{exp}_i + \beta_2 \text{exp}_i^2 + \gamma' X_i + \epsilon_i, \tag{A}$$

where $\log w_i$ is individual i's real log hourly wage, exp_i is individual i's labor market experience, and X_i is a set of individual controls. In such regressions, the differences between the estimates for α for those without and with a college degree capture the

[7]See Appendix section E.1 for further descriptive statistics on HGC and work experience for the 4 groups.

[8]The equation is a generalization of the Mincerian regression as presented by Card (2001), where the additivity assumption between years of schooling and experience is relaxed.

college premium. The differences in β_1 and β_2 allow for wages to increase with experience differently for these two groups. Using these estimates one can calculate the returns to college education as the difference between the present value of lifetime wages for those with and without a college degree.

Regression (A) can be run on cross-sectional datasets, in which only the education level and labor market experience of individuals are recorded. However, we have shown that late bloomers accumulate work experience before completing college (as seen in Figure 3) and that a significant fraction of individuals are late bloomers. These two facts imply that equation (A) is misspecified unless two assumptions hold. First, that the returns to pre-college and post-college years of experience are the same for late bloomers once they graduate. Second, it has to hold that the wage paths for early and late college graduates are the same after graduation. Otherwise, the estimates obtained with regression (A) are biased.

Using the panel aspect of the NLSY dataset, we can check if the two assumptions of same returns and same wage paths hold. We proceed by assessing whether the returns to college education depend on the age at graduation in the NLSY79.[9] We modify equation (A) and estimate the following separate regressions for each of the four groups:

$$\log w_{it} = \alpha + \beta_{1,n}\text{exp}_{n,it} + \beta_{2,n}\text{exp}^2_{n,it} + \theta D_{it} + \beta_{1,c}\text{exp}_{c,it} + \beta_{2,c}\text{exp}^2_{c,it} + \gamma' X_i + \epsilon_{it}. \quad \text{(B)}$$

By running the regression separately for each group, we can see whether the wage profile of late graduates is similar to the never graduates before obtaining a college degree and whether it is similar to the early graduates once they obtain their degree.

There are two key differences between equation (A) and equation (B). The first is that we allow for the college graduation premium, θ, to vary by the age at graduation, through the dummy variable D_{it}, which takes value 1 if i is a college graduate at time t. The second key difference is that in equation (B) we explicitly account for years of experience before completing college $\text{exp}_{n,it}$ and years of experience after completing college $\text{exp}_{c,it}$. By including $\text{exp}_{n,it}$ and $\text{exp}_{c,it}$ separately we allow for college gradua-

[9]Note that as we use panel data, observations are at the individual-time level.

tion to change the slope of the wage-experience profile.

The four groups are quite different in terms of their initial conditions, endowments of skills, family background, etc. This, in fact, is the reason why in X_i we control for a set of time-invariant individual characteristics. These include gender, race, height, weight, and body mass index measured in 1981, and the number of times the individual used drugs in 1979. We also control for the score in Numerical Operations and in Coding Speed from the Armed Services Vocational Aptitude Battery (ASVAB) test completed in 1980.[10] Controlling for these characteristics allows us to capture the impact of initial individual skills on the level of wages.[11]

In Figure 4 we depict the trajectory of a representative individual from each of the four groups. In Table E.6 in the Appendix we provide the regression results.[12] We compute this trajectory for each group by setting the covariates at the level of the group's average, and by assuming that the college degree is obtained at the average of the maximum years of non-college experience across all individuals in the group. As expected, the never college group has the lowest and the early college group has the highest wage trajectory. The medium group has wages close to the never college group before graduation while receiving a big college premium upon graduating. They don't quite catch up with the early college group. The difference between the early and the medium college group becomes more pronounced in the first 15 years of experience, as early graduates have a steeper slope. Interestingly late graduates earn significantly more than the never college and medium college group before graduation. Their college premium upon graduating is smaller than that for the medium college group, but their trajectory after graduation is steeper than for all other groups, thereby gradually narrowing the wage gap with the early and medium college groups. The wage paths for these four groups look qualitatively similar when we run regression (B) by gender and by race, see Figures E.8 and E.9 in the Appendix.

[10]We selected this set of characteristics as it is also available in the NLSY97 sample and allows us to keep a relatively large sample from both data sets. The results for 1979 are very similar when using a broader set of controls.

[11]This does not account for selection into the different college groups, a process that we do not model here.

[12]We also run a regression similar to (B) with pre-college and post-college experience dummies. The results show that the quadratic in experience captures the shape of the wage trajectories well. See Figure E.10 in the Appendix.

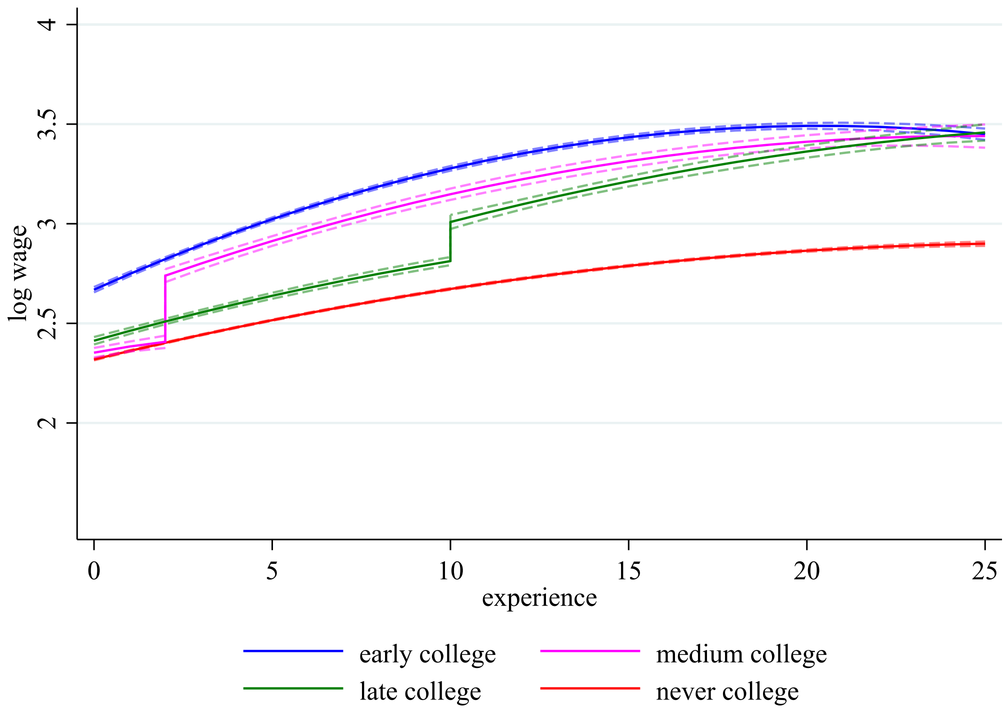
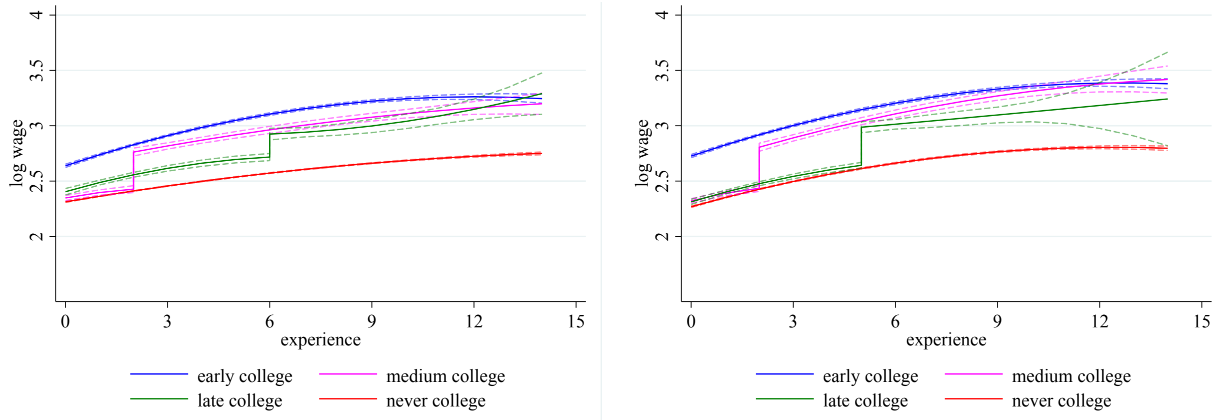


Figure 4: Average wage trajectories by group

Notes: Authors' calculations from NLSY79 data. Each curve shows the predicted wage path against experience from regression (B) for the representative individual from the early (in blue), medium (in pink), late (in green) and never college group (in red). The dashed lines show the 95% confidence intervals.

To see whether the patterns look similar for the four college groups across different birth cohorts, we run similar wage regressions in the truncated NLSY79 data and the NLSY97 data. Figure 5 provides these curves: in the left panel for individuals born in 1957-1964, and in the right panel for those born in 1980-1984. Crucially, Figure 5 indicates that the wage curves for the four groups are qualitatively very similar across the two samples.

In most datasets, even when it is known whether the individual has a college degree, the age at which the individual graduated from college is not recorded. This is true in particular of cross-sectional datasets on which regressions similar to (A) are run. Consequently, the years of experience before and after college graduation are also unknown. Therefore, it is only possible to estimate the returns to college education by running a regression of type (A) for two groups, those who do not have a college degree in the cross-section and those who have a college degree in the cross-section. We have shown in Figures 4 and 5 that the two assumptions needed for the estimates to be



(a) NLSY79 truncated

(b) NLSY97

Figure 5: Average wage trajectories in NLSY79 and NLSY97

Notes: Authors' calculations from NLSY79 and NLSY97 data. In each panel the curves show the predicted wage paths against experience from regression (B) as in Figure 4. The left panel shows these from the regression on the truncated NLSY79 data, while the right panel shows this from the NLSY97 data.

unbiased, namely that the returns to pre- and post-college experience are the same for late bloomers and that the wage paths of early college graduates and late bloomers are the same after graduation, are actually false in the NLSY79. Therefore both estimates for college and non-college graduates from regression (A) are biased. The resulting wage paths for college graduates in the cross-section would be a weighted average of (parts) of 3 of the above 4 curves. The medium and late college graduates would contribute to the non-college group before they graduate and would contribute to the college group after graduation.[13] This would lead to a flatter college wage curve in the cross-section than for the early college graduates, and likely an upward-shifted cross-sectional non-college curve relative to the never college group. This implies that, in general, the returns to college education estimated in the cross-section will be below the correctly estimated returns for the early college group, who constitute the majority of college graduates.

To gauge the magnitude of this error, we estimate (A) for college and non-college workers and (B) for the early college and the never college groups. To mimic what

[13]The shape of the cross-sectional college wage path – even if the wage generating process for each group is unchanged – depends on the share of late graduates at different ages. Therefore the different share of late bloomers across birth cohorts or years may lead to large changes in the cross-sectional estimates of the college wage curve across birth cohorts or years. Our simulations show that the shape of the estimated wage curve differs both across birth cohorts and across years, but the magnitude of these differences is small.

can be estimated in most data, we use potential experience instead of experience and we do not control for any skill measures.[14] We then predict the age-income paths, $\widehat{y}_{i,j,a}$, for college ($j = c$) and non-college workers ($j = n$) both from regression (A) ($i = A$) and regression (B) ($i = B$).[15] Assuming that college graduates start working at age 22, and everyone else at 18, the present value of lifetime earnings until age 48 for college and non-college workers from both these regressions ($i \in \{A, B\}$) is given by:[16]

$$PV_{i,c} = \sum_{a=22}^{48} \beta^{a-18} \widehat{y}_{i,c,a},$$
$$PV_{i,n} = \sum_{a=18}^{48} \beta^{a-18} \widehat{y}_{i,n,a}.$$

The estimated returns to college education based on regression $i \in \{A, B\}$ is then

$$R_i^c \equiv \frac{PV_{i,c} - PV_{i,n}}{PV_{i,n}}. \tag{2}$$

Table 6 provides the estimated returns to college in percentage terms based on the two estimation methods, as well as the extent to which the predictions based on (A) underestimate the returns for early college graduates predicted from (B) (also in percentage terms). As expected the estimated returns based on (A) in the top row are lower than the estimated returns from (B) in the middle row. The extent of underestimation (in the bottom row) is on average 27% among all individuals and varies between 23% for women and other races (non-Black and non-Hispanic), 28% for men, and 35-40% for Blacks and Hispanics. These patterns are in line with the predicted wage paths for the different demographic groups, as shown in Figures E.8 and E.9 in the Appendix. The differences between early graduates and the medium and late graduates are much more pronounced for men than for women and are also more pronounced for Blacks and Hispanics than for non-Black, non-Hispanic individuals. These results suggest that estimating the returns to college education from cross-sectional data, where the age at graduation is not recorded leads to biased estimates. Specifically, the returns

[14]Where potential experience is measured as age $-\ 6\ -$ years of education.

[15]We calculate $\widehat{y}_{i,j,a} = 2080\widehat{w}_{i,j,a}$, where $\widehat{w}_{i,j,a}$ is the predicted hourly wage.

[16]We truncate at age 48, as there are few observations after this age in the NLSY79, and hence our predictions for hourly wages at those ages are less reliable.

are underestimated for the college graduates who follow the standard timing of graduation (our early college graduates), and who are the bulk of college graduates.

	all	men	women	Black	Hispanic	other
Reg. (A)	25.9	31.1	23.0	21.8	27.7	23.7
	(.7)	(1.5)	(2.3)	(1.2)	(.9)	(.9)
Reg. (B)	35.2	43.1	29.8	33.6	45.8	30.8
	(.9)	(1.8)	(3.3)	(1.4)	(1.0)	(1.1)
% underestimation	26.6	27.9	23.0	35.3	39.5	23.1
	(3.2)	(7.0)	(8.7)	(4.3)	(4.4)	(4.8)

Table 6: Returns to college, percent

Notes: Authors' calculations from NLSY79 data. The table shows the estimated returns to college (in percent) defined in (2) based on (A) in the top row and based on (B) in the middle row, and for different demographic groups in the columns. Note that regressions (A) and (B) are run with potential experience instead of experience and without any control for skills. The bottom row shows by how much the predicted returns from (A) underestimate the predicted returns of early college graduates from (B), as a percentage the latter. Standard errors in brackets.

6 Summary and Conclusion

It is a generally well-known fact that most individuals who go to college complete their education in their early 20s. However, in this paper we document that a large fraction of individuals does not follow that route and completes their college education well into their 30s or even later. We carefully investigate the extent of this phenomenon and the contribution of late bloomers to various measures of interest.

We show that the share of late bloomers among college graduates is large – around 20% – and changes non-monotonically across birth cohorts. We document that the differentially changing share of late bloomers across cohorts and across demographic groups contributed to narrowing both the gender and the racial college share gap, despite the fact that for early college graduates the racial gap widened. We also quantify the contribution of cohort succession – whereby successive cohorts acquire more-and-more education – and of late bloomers to the increase in the aggregate college share. We find that the contribution of late bloomers increased from around 50% for the 1960-1990 period to around 70% for the 1990-2019 period.

Regarding the value of obtaining a college degree later in life, we show that the returns to having a college degree differ depending on the age at graduation. Late bloomers receive a college premium. Moreover, their wage-experience profile becomes steeper. Nevertheless, their wages stay below the wages of their early college counterparts. Importantly, we demonstrate that ignoring the existence of late graduates leads to an underestimation of the returns on average by 27% for those who follow the standard college graduation timing. Also, the extent of underestimation varies across demographic groups.

Our findings have important implications. The decomposition of the aggregate increase in the college share indicates that in order to understand the evolution of the aggregate college share, one must pay particular attention to late bloomers. More importantly, the fact that the share of late bloomers among college graduates changes non-monotonically across birth cohorts suggests that there might be vastly different forces that drive early college graduates than those that drive late bloomers. Specifically, we need to examine the factors that affect the decision making process of individuals over their life cycle. There seem to be different factors affecting the decisions of different groups, especially by race and gender, for returning to college. Furthermore, we need to better understand the changing forces over time, and by group, that lead to the very distinct patterns of returning to college across the various sub-populations. This is the topic of further research that we are are engaged in.

Increases in educational attainment are generally viewed as an advantage for innovation, economic competitiveness, and social mobility. However we show in this paper that the returns to educational attainment vary depending on age at graduation. This advocates strongly for a better awareness of late bloomers when designing educational public policy.

References

ACEMOGLU, D. AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by D. Card and O. Ashenfelter, Elsevier, vol. 4, 1043–1171.

AINA, C., E. BAICI, G. CASALONE, AND F. PASTORE (2018): "The Economics of University Dropouts and Delayed Graduation: A Survey," *IZA Discussion Paper Series*.

ALTONJI, J. G., L. B. KAHN, AND J. D. SPEER (2016): "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success," *Journal of Labor Economics*, 34, S361–S401.

ANDERSON, S. (2008): "Later Life Learning: A Review of the Literature," *The Association for Education and Ageing, The Latest on Later Life Learning conference paper*.

BAUMAN, K. (2016): "College Completion by Cohort, Age and Gender, 1967 to 2015," *Census Working Papers*.

BECKER, G. S. (1994): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, Third Edition*, The University of Chicago Press.

BROWN, J., J. B. LIEBMAN, AND J. POLLET (2002): "Appendix: Estimating Life Tables That Reflect Socioeconomic Differences in Mortality," in *The Distributional Aspects of Social Security and Social Security Reform.*, ed. by M. Feldstein and J. Liebman, University of Chicago Press.

BRUNNER, B. AND A. KUHN (2014): "The Impact of Labor Market Entry Conditions on Initial Job Assignment and Wages," *Journal of Population Economics*, 27, 705–738.

CARD, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69, 1127–1160.

CHEESEMAN DAY, J. AND K. BAUMAN (2000): "Have We Reached the Top? Educational Attainment Projections of the U.S. Population." *Census Working Papers*, 43.

CONNOLLY, H. C. AND P. GOTTSCHALK (2006): "Differences in Wage Growth by Education Level: Do Less-Educated Workers Gain Less from Work Experience?" *IZA Discussion Paper Series*.

FERRER, A. M. AND A. MENENDEZ (2009): "The Returns to Flexible Postsecondary Education: The Effect of Delaying School," *CLSSRN working papers*.

GENDA, Y., A. KONDO, AND S. OHTA (2010): "Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States," *The Journal of Human Resources*, 45, 157–196.

GRILICHES, Z. (1980): "Schooling Interruption, Work While in School and the Returns from Schooling," *The Scandinavian Journal of Economics*, 82, 291–303.

JAMIESON, A. (2007): "Higher Education Study in Later Life: What Is the Point?" *Ageing & Society*, 27, 363–384.

KEANE, M. P. AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473–522.

KUH, G. D., J. KINZIE, J. BUCKLEY, B. BRIDGES, AND J. C. HAYEK (2006): "What Matters to Student Success: A Review of the Literature," in *Commissioned Report*.

LEE, D. AND K. I. WOLPIN (2006): "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, 74, 1–46.

LIGHT, A. (1995): "The Effects of Interrupted Schooling on Wages," *The Journal of Human Resources*, 30, 472–502.

MARE, R. D. (1995): "Changes in Educational Attainment and School Enrollment," *State of the Union: America in the 1990s*, 1, 155–213.

MELITZ, M. J. AND S. POLANEC (2015): "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit," *The RAND Journal of Economics*, 46, 362–375.

MINCER, J. A. (1974): *Schooling, Experience, and Earnings*, National Bureau of Economic Research.

MONKS, J. (1997): "The Impact of College Timing on Earnings," *Economics of Education Review*, 16, 419–423.

MURPHY, K. M. AND R. H. TOPEL (2016): "Human Capital Investment, Inequality, and Economic Growth," *Journal of Labor Economics*, 34, S99–S127.

ROY, B., C. I. KIEFE, D. R. JACOBS, D. C. GOFF, D. LLOYD-JONES, J. M. SHIKANY, J. P. REIS, P. GORDON-LARSEN, AND C. E. LEWIS (2020): "Education, Race/Ethnicity, and Causes of Premature Mortality Among Middle-Aged Adults in 4 US Urban Communities: Results From CARDIA, 1985-2017," *American Journal of Public Health*, 110, 530–536.

RUGGLES, S., JS. FLOOD, M. SOBEK, D. BROCKEN, S. RICHARDS, AND M. SCHOUWEILERK (2023): "IPUMS USA," Tech. rep., Minneapolis, MN.

RYDER, N. B. (1965): "The Cohort as a Concept in the Study of Social Change," *American Sociological Review*, 30, 843–861.

SCHWANDT, H. AND T. VON WACHTER (2019): "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets," *Journal of Labor Economics*, 37, S161–S198.

WILSON, J., C. ZOZULA, AND W. GOVE (2011): "Age, Period, Cohort and Educational Attainment: The Importance of Considering Gender," *Social Science Research*, 40, 136–149.

YANG, G. (2021): "Best Time for College? A Tale of Two Endowments," *Working Paper*.

ZHANG, X. AND B. PALAMETA (2006): "Participation in Adult Schooling and Its Earnings Impact in Canada," *Analytical Studies Branch Research Paper Series - Statistics Canada*.

Appendix

A Data selection, variable definitions

The following Census and ACS data was downloaded from IPUMS: 1960 1% sample, 1970 1% form 2 state sample, 1980 1% metro sample, 1990 1% metro sample, 2000 1% sample, and yearly ACS surveys (1%) for 2001-2019. We keep individuals between 18 and 65 years old at the time of the survey and calculate their birth cohort from the year of the survey and age. We define education levels based on the detailed highest grade of school attended or completed variable *'higraded'* for 1960-1980, and based on the detailed highest year of school or degree completed *'educd'* from 1990 onwards. College graduates are individuals who completed at least 4 years of college (*'higraded'*), or received a bachelor's degree or more (*'educd'*). We calculate the share of college-educated workers as the sum of person weights (*'perwt'*) among college-educated individuals divided by the sum of person weights among all individuals. We calculate this by birth cohort and year, and potentially one or more of the following categories: gender, race, foreign-born. We use the single race identification variable (*'racesing'*) to determine whether an individual is White, Black, or other. Note that since the share of individuals in the other category is very small, but increasing over time even within cohort, we only analyze trends for Black and White categories separately. To identify foreign and local-born individuals, we use the detailed birthplace variable (*'bpld'*) and code anyone born outside the US as foreign-born. We rely on the attending school variable (*'school'*) to determine whether an individual is attending school in a given year. We calculate the share of individuals attending school among all individuals using person weights by birth cohort, year, and also by level of completed education.

We use two releases from the NLSY: from 1979 to 2014 and from 1997 to 2019. Each release follows a cohort over time. We refer to each of the two cohorts as the NLSY79 and the NLSY97. Individuals in the same cohort may be of different ages: the youngest individual in the NLSY79 is 14 in 1979 and the oldest is 22, while in the NLSY97 cohort, the youngest individual is 12 in 1997, and the oldest is 18. The NLSY79 cohort is interviewed every year between 1979 and 1994, and every two years between 1994 and 2014. The NLSY97 is interviewed every year between 1997 and 2013, and every two years between 2013 and 2019. We exclude from the sample individuals who are not interviewed every year for the first four years, as well as individuals who are interviewed less than four years in total. The NLSY79 and NLSY97 are not interviewed over the same time length: the NLSY79 covers 36 years, but the NLSY97 only covers 23 years. To compare the two cohorts – in terms of the share of late bloomers in Section 3 or in terms of their returns to education by age of graduation in Section 5, we truncate the NLSY79 to obtain comparable age distributions in the final year of the truncated NLSY79 and the NLSY97. We truncate the NLSY79 after 1998; respondents are between 33 and 41 in 1998 in the NLSY79 and are between 34 and 41 in the NLSY97 in 2019, with an average age of 36 in both.

We use the demographic (race and gender) variables provided in the surveys. We define education levels based on the highest grade completed: if the individual reaches 16 years of education or more, we consider them a college graduate. To define years of labor market experience, we rely on the number of hours and weeks worked every year, as well as the number of months in which the individual reported to be at school or in college. All three variables are available also when interviews are bi-annual. An individual gains one year of experience in year t if the four following conditions are met in year $t-1$: they were at least 18, worked at least 800 hours, at least 25 weeks, and were in education (school or college) for less than 6 months. Non-college experience is the labor market experience accumulated before the individual graduates from college, and the college experience is accumulated after they graduate from college. College experience is set to zero in the year individuals graduate from college. Non-college graduates have no college experience by construction. We set early college graduates' non-college experience to zero, to avoid counting high school summer jobs as labor market experience. We use the hourly wages provided in the surveys whenever available and construct them as yearly wages divided by the total number of hours worked in the year when necessary. Finally, we use the following controls provided in the surveys: height in centimeters, weight in kilograms, BMI (calculated from height and weight), the ASVAB numerical operations and coding speed scores, and a measure of drug usage. All controls are from the first year of the survey. Drug usage for the NLSY79 is on a scale of 0 (never used drugs) to 6 (used drugs more than 50 times). The NLSY97 reports a scale of 0 (never used drugs) to 3 (uses drugs often). Although ASVAB scores are available in more sections, we choose not to use these because their values are missing for a large share of the NLSY97 respondents.[17]

When measuring the returns to education by age of graduation in section 5, we impose additional constraints on the data. We run the regression on individuals whose highest grade completed is known, who work, and who earn a nominal hourly wage above the federal minimum wage. We deflate nominal wages to perform the analysis on real wages. We exclude individuals whose height, weight, ASVAB numerical operations and coding speed scores, and drug usage are unknown in 1979 or 1997. When performing the analysis on NLSY79 only, we cap observations at 25 years of total labor experience, for reasons of data scarcity. When performing the analysis on the truncated NLSY79 and NLSY97, we cap at 14 years of experience.

B Late bloomers

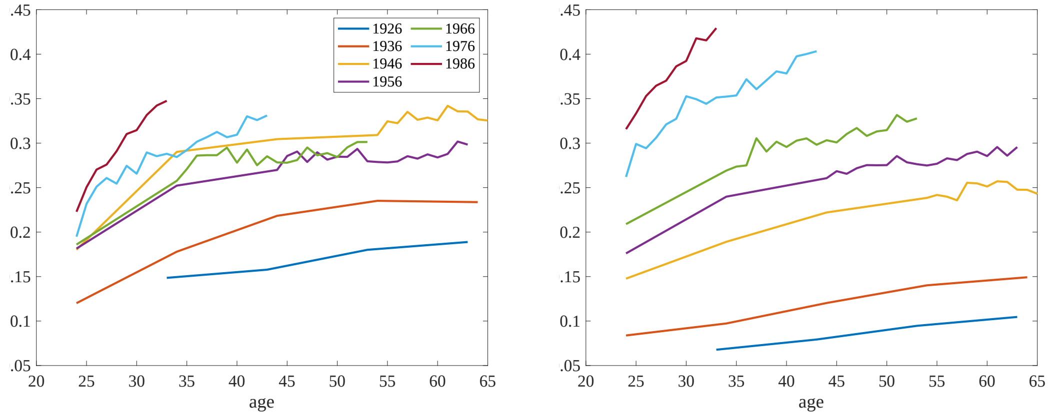
B.1 College share by age for different demographic groups – Census

Figure B.1 shows that the same three patterns hold for both men (on the left) and women (on the right). The two panels also show that the cross-cohort level differences in college share are much larger for women than for men. In fact, for men, there are hardly any differences at age 24 across the cohorts born at or after 1946. This shows how women caught up and overtook men in terms of college education.

Figure B.2 shows that the three patterns hold by gender-race as well.[18] Women increased their education more than men among the White and the Black population. Interestingly, the college share among Black women is higher than among Black men for all cohorts.

[17] These sections are general science, arithmetic reasoning, word knowledge, paragraph comprehension, auto and shop information, mathematics knowledge, mechanical comprehension, and electronics information.

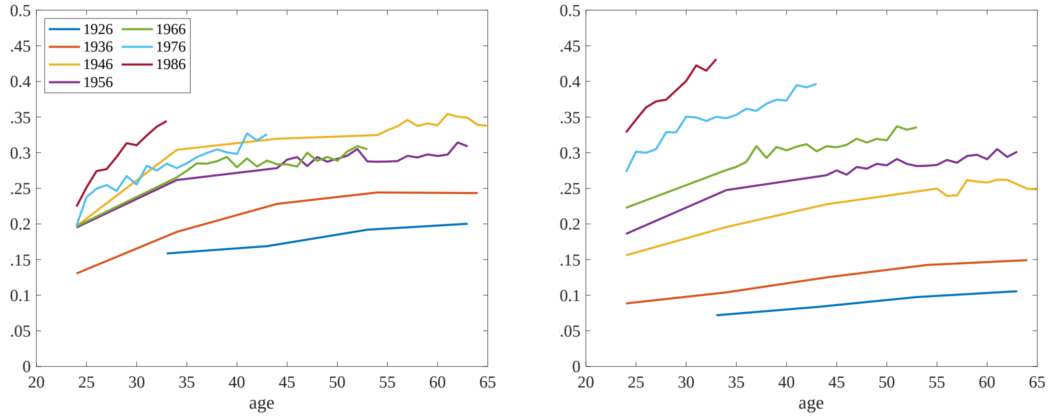
[18] In our sample period the only race categories that can be consistently defined are White, Black, and other. The fraction of individuals in the other category is very small, and their share *within* a cohort is increasing over time, implying that the classification of the same individuals must have changed over time. For this reason, we do not include individuals in the other race category when breaking down the data by race.



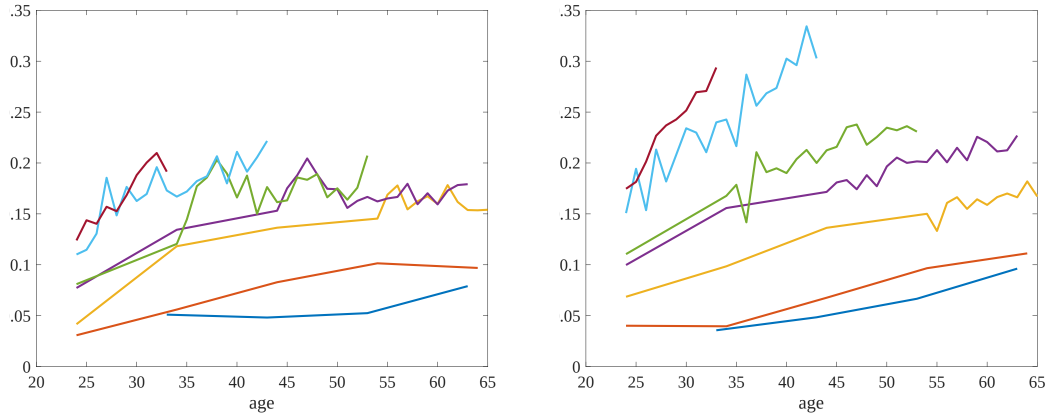
(a) Men (b) Women

Figure B.1: College share by age and gender for different cohorts

Notes: Authors' calculations from Census 1960-2010 and ACS 2001-2009 and 2011-2019. Each curve shows the share of college graduates for a birth cohort against age for men and for women.



(a) White men (b) White women



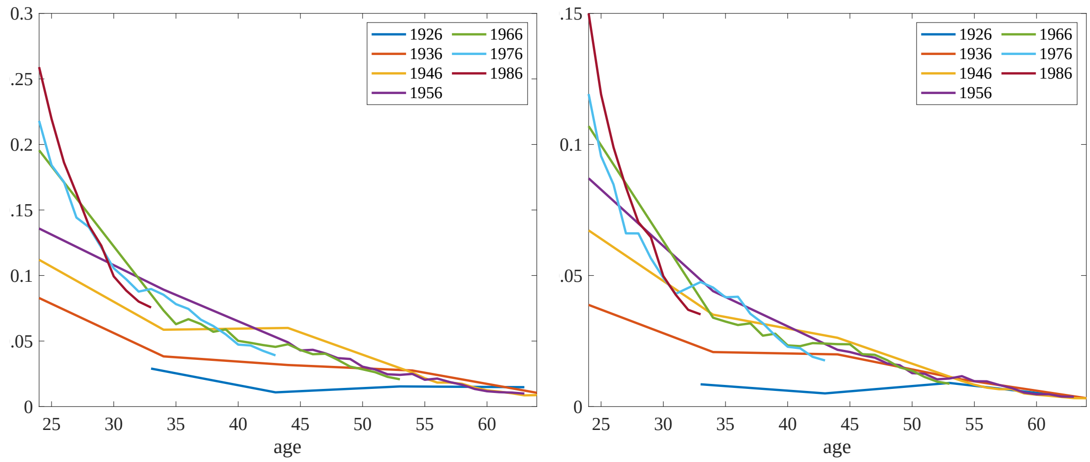
(c) Black men (d) Black women

Figure B.2: College share by age, gender, and race for different cohorts

Notes: Authors' calculations from the Census 1960-2010 and ACS 2001-2009 and 2011-2019. Each curve shows the share of college graduates for a birth cohort against age by gender and race.

B.2 School enrollment in the Census

Figure B.3 provides some direct evidence for the fact that individuals continue to acquire education later in life. The left graph shows the share within a birth cohort enrolled in school against the age of the cohort. The graph on the right shows the share within a birth cohort of those who are enrolled in school and already have some college education. These graphs clearly demonstrate that a significant share of 30, 40, and even 50 year olds are enrolled in school and that these fractions increase across birth cohorts (higher fraction at a given age in later-born cohorts).



(a) Share in school

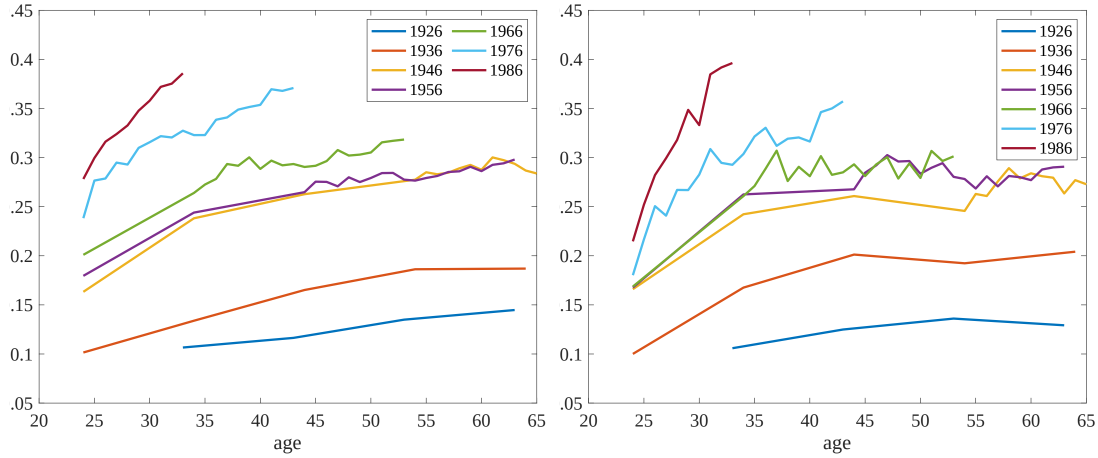
(b) Share in school with some college

Figure B.3: Share enrolled in school by age for different cohorts

Notes: Authors' calculations from the Census 1960-2010 and ACS 2001-2009 and 2011-2019. Each curve shows the share enrolled in school for a birth cohort against age. The left panel shows everyone who is enrolled in school, the right only those who already have some college education.

B.3 College share by age and place of birth

Figure B.4 shows the share of college-educated individuals in the Census and ACS for different birth cohorts as they age, among local-born individuals (on the left) and foreign-born (on the right). The main takeaway is that the patterns we document hold among the local-born population, implying that the increasing share of college-educated individuals as a birth cohort ages is not due to immigration. Interestingly, similar patterns hold for foreign-born individuals of the later cohorts, whereas the age gradient flattens out after age 35 for cohorts born before 1970.



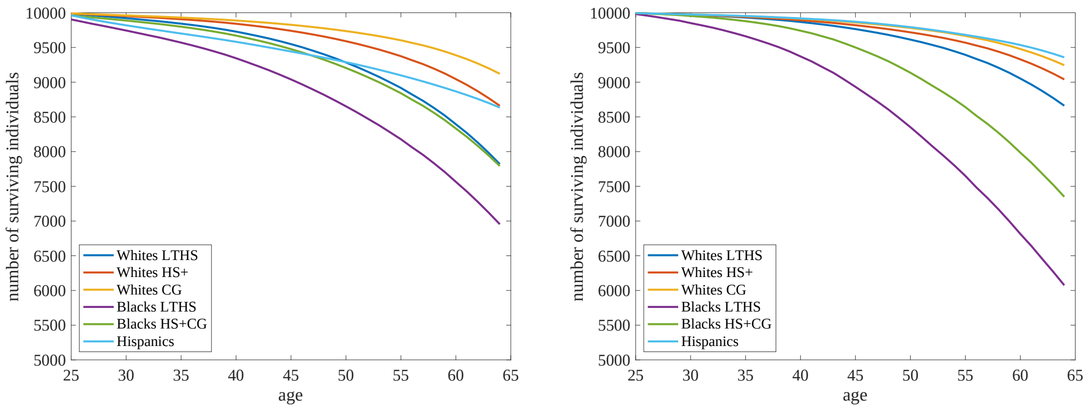
(a) Local (b) Foreign born

Figure B.4: College share by age and place of birth for different cohorts

Notes: Authors' calculations from the Census 1960-2010 and ACS 2001-2009 and 2011-2019. Each curve shows the share of college graduates for a birth cohort against the age of the cohort for local-born in the left panel and for foreign-born in the right panel.

B.4 The impact of mortality differences

While there is some evidence that mortality is higher for individuals with lower educational attainment, the difference is not large enough to account for the pattern observed in the data. Roy et al. (2020) use as their primary outcome of interest years of potential life lost (YPLL). Using a rather innovative multivariable model, they conclude that indeed, each level of education achieved is associated with 1.37 fewer YPLL (P = .007). They also conclude that race was also associated with YPLL. Nevertheless, the magnitude indicates that the differences in education outcomes we observe in the data cannot possibly be explained by mortality differences. To illustrate the magnitude of mortality differences we provide Figure B.5. This figure is computed using information from three different sources. First, we use the life tables from the National Vital Statistics Report for 2011, in which we have information about mortality rates by race. Second, we use the annual life table statistics provided by the National Vital Statistics Reports for all years. Finally, we use relative mortality rates by race, ethnicity, and education reported in Brown et al. (2002). These three sources make it possible to calculate the annual survival rate for several important groups, namely Whites with less than high school education (LTHS), Whites with high school degree, but did not complete a four-year college degree (HS+), and Whites with a college degree (CG). For Blacks we can calculate these probabilities for those with less than high school education and those with more than high school education, including college (HS+CG). For Hispanics we are not able to distinguish between the various levels of education. Finally, we can compute the probabilities for women and men.[19] We then simulate the number of surviving individuals for each of the groups described above starting from a population of 10,000 individuals at age 25. The surviving population against age for these groups is depicted in the left panel of Figure B.5 for men and in the right panel of Figure B.5 for women. We can see that there are large differences across the groups. For example, there is a huge difference between the White college graduate and the Black with less than high school group. Nevertheless, these differences are far too small to raise the percentage of college graduates in the data to the extent that is observed in the data. To illustrate this, consider a population of 5,000 individuals who are 25 years old, 4,000 of whom are White women, while 1,000 are Black women. Also, assume that among the White women 30% are college graduates, and among the Black women 15% are college graduates. The differential in survival rates would lead to the fact that after 20 years among White women, 30.6% would be college graduates, while among Black women 15.6% would be college graduates. These are all very small changes relative to the changes observed in the data. Below we provide the



(a) Men

(b) Women

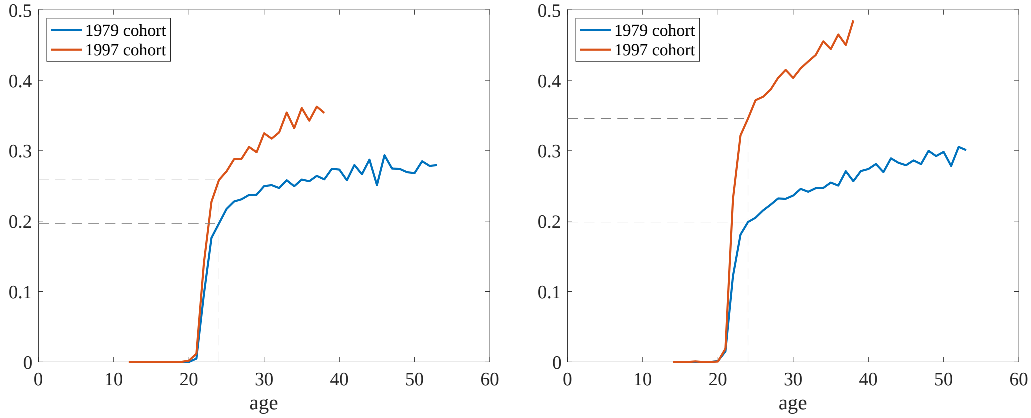
Figure B.5: Simulated surviving population by gender, race and education

Notes: Authors' simulations based on several data sources, as described in the text.

[19]The resulting mortality rate tables by gender, education, and race are available on request.

B.5 College share by age for different demographic groups – NLSY

Figure B.6 shows that these facts hold both for men and for women, and that, as in the Census, the college share among women increased both due to a higher initial college share (age 24) and due to a steeper age-gradient in educational attainment.

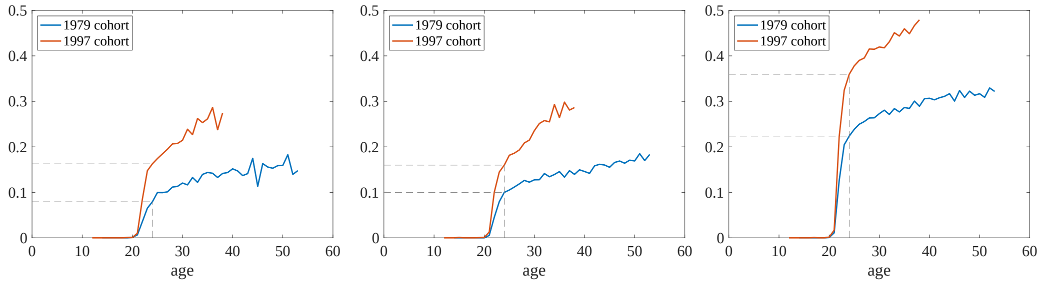


(a) Men (b) Women

Figure B.6: College share by age and gender in the two NLSY cohorts

Notes: Authors' calculations from the two releases of the NLSY. In both panels the two curves show the share of college graduates against age for individuals from the NLSY79 (in blue) and the NLSY97 (in red). The left panel shows this among men, the right among women.

Figure B.7 shows that these facts hold both among the Black, the Hispanic, and the non-Black non-Hispanic population.



(a) Hispanics (b) Blacks (c) others

Figure B.7: College share by age and race in the two NLSY cohorts

Notes: Authors' calculations from the two releases of the NLSY. In all panels the two curves show the share of college graduates against age for individuals from the NLSY79 (in blue) and the NLSY97 (in red). The left panel shows this among Hispanics, the middle among Blacks, and the right among the non-Hispanic non-Black population.

C Share of college graduates by age

Table C.1 shows – for different birth cohorts in each row, and for different subgroups of the cohort in each column – the share of those who get a college education by age 30 in the top, by age 50 in the middle, and the difference between the two in the bottom panel. The bottom panel of the table shows that the share of college graduates has increased almost monotonically across these birth cohorts, except for the 1960 birth cohort, which has a lower college share than the 1950 birth cohort for Whites and for men, and a slower rate of increase for Blacks and for women. Table C.2 shows the same

	all	men	women	Whites	Blacks
cohort	college share at age 30				
1930	11.2	14.9	7.6	12.0	4.3
1940	14.6	18.2	11.2	15.5	5.3
1950	24.3	27.9	20.8	25.6	11.8
1960	22.6	23.0	22.1	23.5	12.2
1970	28.9	27.4	30.5	29.5	15.7
	college share at age 50				
1930	14.5	18.8	10.5	15.3	7.1
1940	20.3	24.5	16.4	20.9	10.4
1950	29.4	31.8	27.1	30.6	17.5
1960	27.0	26.2	27.8	27.5	18.5
1970	35.5	33.4	37.6	36.0	24.9
	share graduated between age 30 and 50				
1930	3.3	3.9	2.9	3.3	2.9
1940	5.7	6.3	5.2	5.5	5.1
1950	5.1	3.9	6.3	5.0	5.7
1960	4.4	3.2	5.6	4.1	6.4
1970	6.7	6.1	7.1	6.5	9.2

Table C.1: Share of college graduates by age, gender and race

Notes: Authors' calculations from the decadal Census for 1960-2010 and ACS data for 2019. The table shows the share of college graduates for different demographic groups in each column and for different birth cohorts in each row. The top panel shows the share of college graduates at age 30, the middle panel at 50, and the bottom panel is the difference between the top two panels.

measures in the NLSY79, the NLSY97, and the truncated NLSY79, which is truncated in 1998, to make its age range comparable with that of the NLSY97.

NLSY sample	all	men	women	Hispanics	Blacks	others
	share at age 24					
1979	19.4	19.3	19.5	7.8	9.7	22.0
1997	28.7	23.8	33.7	15.8	15.8	33.8
	share between age 25 and 29					
1979	3.5	3.6	3.4	3.5	2.8	3.6
1997	5.9	5.3	6.6	5.7	5.7	6.0
	share at or after 30					
1979	4.6	3.3	5.9	4.2	5.3	4.5
1979 truncated	2.1	1.7	2.5	1.9	1.7	2.2
1997	4.7	4.1	5.4	5.4	6.2	4.3

Table C.2: Share of college graduates by age, gender and race, NLSY

Notes: Authors' calculations from NLSY79 and NLSY97. The share of college graduates is shown for different demographic groups (columns), for different samples (rows), and at different ages (panels).

D Shift-share decomposition by decade and demographic groups

	aggregate change	entry	exit	shift	between cohort	within cohort
			men			
1960-1970	4.3	1.5	0.9	0.2	2.6	1.7
1970-1980	6.9	1.8	1.4	0.2	3.4	3.5
1980-1990	3.1	-1.3	1.8	0.1	0.6	2.5
1990-2000	2.2	-0.7	0.8	-0.1	0.1	2.2
2000-2010	1.6	-0.6	0.2	-0.1	-0.5	2.1
2010-2019	3.3	0.1	-0.9	0.0	-0.8	4.2
			women			
1960-1970	2.7	1.3	0.4	0.0	1.6	1.1
1970-1980	5.3	2.9	0.5	0.1	3.4	1.9
1980-1990	5.7	1.1	1.6	0.1	2.8	3.0
1990-2000	5.6	1.2	1.7	0.1	3.0	2.7
2000-2010	5.2	1.2	1.4	0.0	2.7	2.5
2010-2019	6.0	1.2	1.0	0.1	2.3	3.7
			Blacks			
1960-1970	1.4	0.4	0.3	0.0	0.7	0.8
1970-1980	4.4	1.4	0.3	0.0	1.7	2.7
1980-1990	3.2	-0.3	0.7	0.1	0.5	2.7
1990-2000	2.9	-0.1	0.6	0.0	0.5	2.4
2000-2010	3.8	-0.1	0.4	0.0	0.3	3.4
2010-2019	4.7	-0.2	0.2	0.0	0.0	4.7
			Whites			
1960-1970	3.6	1.5	0.7	0.1	2.2	1.4
1970-1980	6.1	2.6	1.0	0.1	3.7	2.5
1980-1990	4.6	0.0	1.9	0.1	2.0	2.6
1990-2000	3.9	0.2	1.4	0.0	1.6	2.3
2000-2010	3.0	0.2	0.9	0.0	1.1	1.9
2010-2019	4.3	0.6	-0.2	0.0	0.5	3.8

Table D.3: Aggregate college share change decomposition by gender & race

Notes: Authors' calculations from the Census for 1960-2010 and ACS data for 2019. The table shows the decomposition in (1) by gender and by race.

E Returns to college by age at graduation

E.1 HGC and work experience by college groups

Table E.4 reports the average highest grade completed and work experience by college group at different ages for the NLSY79, truncated NLSY79 and NLSY97 cohorts. The earlier the graduation, the higher the grade completed at every age, and the less the work experience.

	highest grade completed					work experience		
at age	20	22	24	26	28	22	26	30
				NLSY79				
early	13.90	15.51	16.29	16.57	16.75	0.05	0.08	0.08
medium	13.19	13.95	14.63	15.52	16.09	0.48	1.29	1.44
late	12.65	13.02	13.26	13.40	13.60	0.92	2.92	5.03
never	11.72	11.88	11.95	12.00	12.04	1.13	3.58	6.15
				NLSY97				
early	13.86	15.66	16.62	17.12	17.42	0.11	0.14	0.14
medium	12.91	13.69	14.26	15.51	16.40	0.77	1.72	1.84
late	12.38	12.76	12.93	13.19	13.92	1.10	2.79	4.02
never	11.58	11.77	11.83	11.88	11.96	1.47	3.72	6.02

Table E.4: Work experience and highest grade completed by age and college group

Notes: Authors' calculations from NLSY79 and NLSY97. Columns 1-5 show the average of the highest grade completed, columns 6-9 show the average years of work experience before completing college at different ages, in the 4 college groups in each row. Both are computed using sampling weights.

E.2 Descriptive statistics

Table E.5 presents the sample means of the variables used in the regression analysis for the NLSY79, the truncated NLSY79, and the NLSY97 cohorts. In the NLSY79, the difference in average maximum non-college experience is especially stark between medium and late college graduates: medium college graduates have little experience before they finish college (1.32 years), and accumulate college experience similarly to early college graduates (15.99 years and 17.69 years). Meanwhile, late college graduates accumulate much more non-college experience (10.13 years), and less post-college experience (8.89 years). This difference in human capital accumulation is the reason why we find it important to distinguish between these two groups. Similar observations can be made about the truncated NLSY79 and the NLSY97. Note that in the last year of both the truncated NLSY79 and the NLSY97, individuals are at least 33 and at most 41 years old, and on average 36 years old. This implies that all early and medium graduates have already graduated, but some late graduates from the NLSY79 still appear in the never college group in the truncated NLSY79. Also, note that individuals who do not report a wage over the time span of each cohort are not accounted for in this table. This time span is 20 years in the truncated NLSY79 and 36 years in the

	early	medium	late	never
	NLSY79			
# individuals	1,607	329	446	8,159
# observations	22,954	5,686	8,168	115,538
mean max exp non-college	-	1.32	10.13	2.13
mean max exp college	17.69	15.99	8.89	-
share women	0.53	0.49	0.65	0.55
share Black	0.16	0.23	0.32	0.20
share Hispanic	0.07	0.17	0.16	0.10
mean ASVAB coding speed	53.40	48.48	47.89	51.69
mean ASVAB numerical operations	40.78	36.82	36.54	39.44
mean weight (kg)	65.49	65.30	64.38	65.25
mean height (m)	1.72	1.71	1.69	1.71
mean BMI	22.12	22.18	22.45	22.19
median drug use	0	0	0	0
	NLSY79 truncated			
# individuals	1,606	328	204	8,353
# observations	16,781	4,282	2,709	91,140
mean max exp non-college	-	1.32	5.14	0.75
mean max exp college	9.29	6.65	2.73	-
share women	0.53	0.49	0.61	0.53
share Black	0.16	0.23	0.25	0.18
share Hispanic	0.07	0.16	0.16	0.09
mean ASVAB coding speed	53.42	48.55	48.60	52.21
mean ASVAB numerical operations	40.78	36.84	37.05	39.82
mean weight (kg)	65.48	65.28	65.05	65.41
mean height (m)	1.72	1.71	1.70	1.71
mean BMI	22.11	22.18	22.33	22.15
median drug use	0	0	0	0
	NLSY97			
# individuals	1,594	359	291	3,417
# observations	14,544	4,603	3,448	4,0167
mean max exp non-college	-	1.80	4.98	1.04
mean max exp college	8.84	6.01	1.63	-
share women	0.59	0.55	0.59	0.58
share Black	0.15	0.23	0.34	0.19
share Hispanic	0.11	0.19	0.20	0.14
mean ASVAB coding speed	7828.15	6623.88	6620.64	7478.90
mean ASVAB numerical operations	19889.14	17455.93	17204.02	19151.67
mean weight (kg)	57.55	58.21	59.68	57.93
mean height (m)	1.65	1.64	1.64	1.65
mean BMI	21.02	21.38	22.03	21.21
median drug use	0	0	1	0

Table E.5: Descriptive statistics by college group

Notes: Authors' calculations from the NLSY79 and the NLSY97. The sample is the same as the one used in Table E.6, see appendix A.

NLSY79, which is why the latter contains more individuals than the former.

E.3 Regression results

Table E.6 presents the estimation results of equation (B) on the NLSY79, the truncated NLSY79, and the NLSY97. These are the estimated coefficients we use to construct the wage paths presented in Figures 4 and 5.

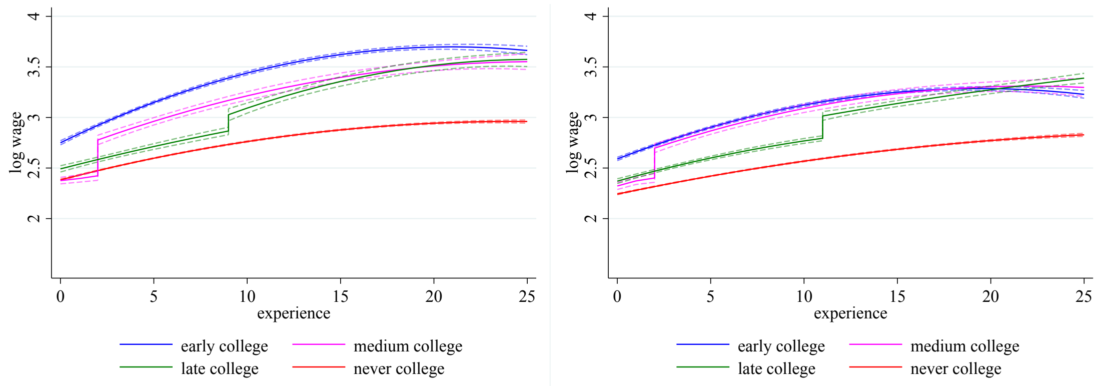
	log wage			
	early	medium	late	never
		NLSY79		
const	4.216	2.479	1.798	3.169
	(.444)	(.760)	(.617)	(.128)
$\exp_n$	-	.034	.050	.043
		(.012)	(.003)	(.001)
$\exp_n^2$	-	-.004	-.001	-.001
		(.002)	(.000)	(.000)
D	-	.332	.196	-
		(.019)	(.019)	
$\exp_c$	.081	.062	.046	-
	(.002)	(.004)	(.005)	
$\exp_c^2$	-.002	-.001	-.001	-
	(.000)	(.000)	(.000)	
Nb. obs	22,954	5,686	8,168	115,538
R^2	0.274	0.390	0.305	0.227
		Truncated NLSY79		
const	4.595	2.027	1.955	3.174
	(.495)	(.821)	(.989)	(.144)
$\exp_n$	-	.058	.089	.052
		(.013)	(.009)	(.001)
$\exp_n^2$	-	-.009	-.006	-.002
		(.003)	(.001)	(.000)
D	-	.335	.208	-
		(.019)	(.029)	
$\exp_c$	.104	.057	.012	-
	(.003)	(.009)	(.023)	
$\exp_c^2$	-.004	-.002	.004	-
	(.000)	(.001)	(.004)	
Nb. obs	16,781	4,282	2,709	91,140
R^2	0.192	0.280	0.200	0.170
		NLSY97		
const	4.087	.081	2.824	2.302
	(.377)	(.612)	(.636)	(.146)
$\exp_n$	-	.080	.093	.087
		(.012)	(.007)	(.002)
$\exp_n^2$	-	-.012	-.005	-.003
		(.002)	(.001)	(.000)
D	-	.371	.344	-
		(.020)	(.026)	
$\exp_c$	.104	.087	.027	-
	(.004)	(.010)	(.027)	
$\exp_c^2$	-.004	-.003	.000	-
	(.000)	(.001)	(.005)	
Nb. obs	14,544	4,603	3,448	40,167
R^2	0.187	0.352	0.259	0.220

Table E.6: Estimates from equation (B)

Notes: All regressions contain gender and race fixed effects as well as individual time-invariant controls.

E.4 Returns to college education by gender and race

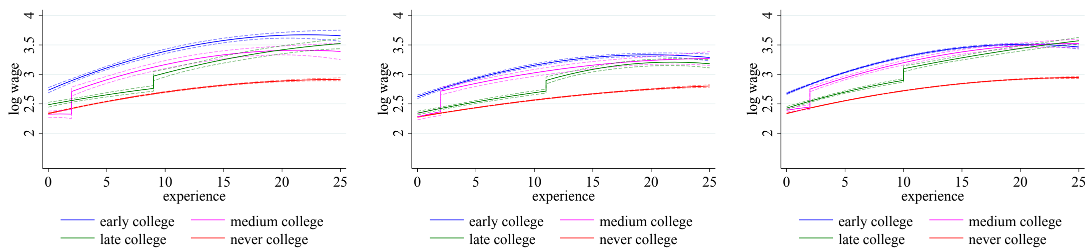
Figure E.8 shows the wage trajectory for the representative individual from each college group when equation (B) is estimated separately by gender on the NLSY79. They reveal different dynamics for women and men. First female early graduates' returns to experience are flatter and start lower than male early graduates'. Second, female medium graduates do catch up with female early graduates, while this is not true of male medium graduates. Finally, male late graduates experience a steeper increase in their returns to experience after they obtain their college degree than female late graduates do.



(a) Men (b) Women

Figure E.8: Average wage trajectories by college group for men and women

Notes: Authors' calculations from the NLSY79. The two panels show for men and women the predicted wage paths against experience from regression (B) for the representative individual of the early (in blue), medium (in pink), late (in green) and never college group (in red). The dashed lines show the 95% confidence intervals.



(a) Hispanics (b) Blacks (c) others

Figure E.9: Average wage trajectories by college group and race

Notes: Authors' calculations from the NLSY79. Each panel shows for the given race/ethnicity the predicted wage paths against experience from regression (B) for the representative individual of the early (in blue), medium (in pink), late (in green) and never college group (in red). The dashed lines show the 95% confidence intervals.

Figure E.9 shows representative wage paths for Hispanics, Blacks, and non-Hispanic, non-Black individuals in each of the four college groups from the NLSY79. We also observe disparities between races. First, the gap between medium and early graduates is wider among Hispanics, and then Blacks. Second, Black late graduates accumulate more labor market experience than other races before obtaining a college degree, although it does not shift their wage paths higher once they obtain their college degree. Figure E.10 shows the predicted wage paths for each college group from the following regression:

$$\log w_{it} = \alpha + \sum_{s=0}^{30} \beta_{s,n} \mathbf{I}\{\text{exp}_{n,it} == s\} + \sum_{s=0}^{30} \beta_{s,c} \mathbf{I}\{\text{exp}_{c,it} == s\} + \gamma' X_i + \epsilon_{it}. \quad (3)$$

This equation is very similar to (B) except that instead of assuming that experience affects wages in a quadratic way, we allow for each year of experience to have a different effect on wages. Note also that in this case the college completion dummy can be omitted, as it will be subsumed in the relevant $\beta_{s,c}$. The predicted wage paths are very similar to those in Figure 4, suggesting that the assumption of a quadratic effect of experience on log wages is not overly restrictive.

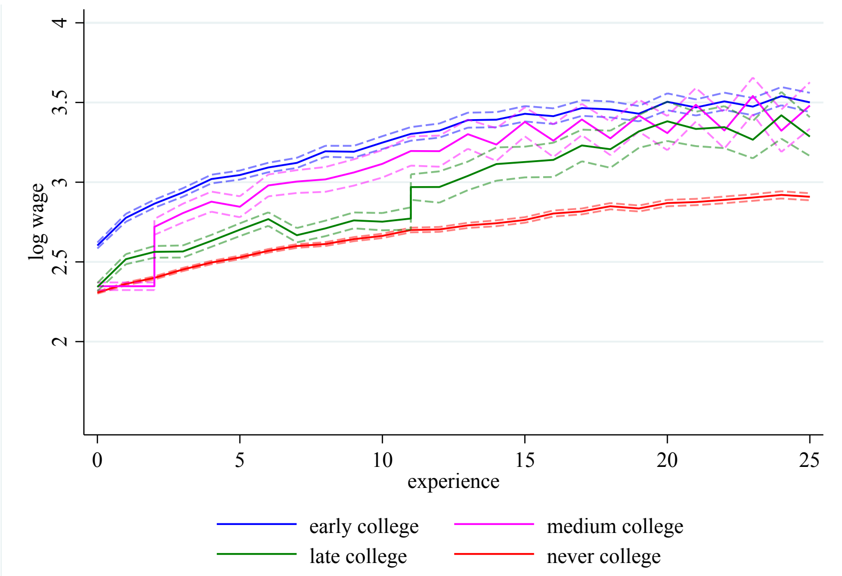


Figure E.10: Average wage trajectories by group

Notes: Authors' calculations from the NLSY79. Each curve shows the predicted wage path against experience from regression (3) for the representative individual of the early (in blue), medium (in pink), late (in green) and never college group (in red). The dashed lines show the 95% confidence intervals.