Online Appendix: Family Heterogeneity, Human Capital Investment, and College Attainment

Adam Blandin and Christopher Herrington
Virginia Commonwealth University

July 23, 2021

Table of Contents for Online Appendix

A ROBUSTNESS: COLLEGE TRENDS AND FAMILY BACKGROUND

This appendix contains a variety of robustness exercises and sensitivity checks for empirical results presented in Sections 1 and 2.

A.1 Aggregate College Attainment Trends

We first document aggregate college attendance and completion trends from 1960-2015 for the US in several nationally representative data sets. These trends are plotted in Figure A1. We observe similar patterns across the different datasets, although there are some differences.

The first dataset we use is the 1962-2015 Consumer Population Surveys (CPS), harmonized by IPUMS (Flood et al., 2020). Our variable of interest is the educational attainment variable "educ". We categorize as "no college" values between 2–73, which range from "None or preschool" to "High school diploma or equivalent." We categorize as "attending college" values between 80–125, which range from "1 year of college" to "6+ years of college" before 1992 and range from "Some college but no degree" to "Doctorate degree" from 1992-onward. We categorize as "completing college" values between 110–125, which range from "4 years of college" to "6+ years of college" before 1992 and range from "Bachelor's degree" to "Doctorate degree" from 1992-onward. Note that we do observe a somewhat large jump in attendance for the CPS series (about 4 percentage points) between years 1991–1992, which may be attributable to the change in education categories used that occurred at this time. However, we observe no noticeable jump in completion rates between the years 1991-1992.

We plot attainment rates for the CPS for two age groups: 27-29 and 24-26. The former is centered around age 28, which is ten years after most individuals make their initial college decision and is when we measure college attainment throughout the paper. The latter is centered around age 25, which is another common age to measure attainment (see e.g. Bound, Lovenheim and Turner (2010)). Note that it is possible for the 24-26 group to have a higher attendance rate than the 27-29 group because we are plotting attainment at the year observed, not birth cohorts.

The second dataset we use is the 1960-2000 decennial Censuses, combined with the 2001-2015 American Community Surveys (ACS), again harmonized by IPUMS (Ruggles et al., 2021). Our variable of interest again is the educational attainment variable "educ". We categorize as "attending college" values between 80–125, which range from "1 year of college" to "6+ years of college". We categorize as "completing college" values between 110–125, which range from "4 years of college" to "6+ years of college". These series show similar trends in college attainment compared to the CPS throughout the sample period.

The final dataset we use is the Panel Study of Income Dynamics (PSID) (Panel Study of Income Dynamics, 1968–2015). Our complete PSID data extract (Blandin and Herrington, 2021) is available for download separately from the other replication materials at `https://doi.org/10.3886/E145622V1`. Our variable of interest is "highest grade or year of school completed", which is collected for household heads and "spouses." We categorize as "attending college" anyone with at least 13 years of school completed. We categorize as "completing college" anyone with at least 16 years of school completed. Although the PSID does ask about college degrees, it does not do so until 1975, and we require knowledge about college attainment before 1975 when we look at parental education to assign children to family types. To maintain a consistent measure of college attainment, we use "highest grade or year of school completed" throughout our analysis. We are reassured that this variable is an accurate measure of college attendance and completion because we observe similar college trends in the PSID to both the CPS and the Census/ACS in Figure A1. As with the Census/ACS, the PSID displays similar college attainment trends to the other data sets. In particular, college completion increased 47% in the CPS, from 25.0% to 36.8%, and increased 46% in the PSID, from 22.9% to 33.5%; college attendance increased 22% in the CPS, from 54.0% to 65.7%, and increased 25% in the PSID, from 50.5% to 63.2%.

A.2 College Attainment Trends Using Alternative Family Background Measures

Figures A2 and A3 document college attainment trends using alternative approaches to assign individuals to a particular family background. While our baseline procedure categorizes a family as one parent or two based on cohabitation, Figure A2 labels a family as single-parent if the head is legally unmarried, and dual-parent if the head is legally married. The results in Figure A2 are quantitatively similar to our baseline results: completion grows by 14pp for 2H, 12pp for 2L, and 0pp for 1L, compared to 11pp, 8pp, and 2pp, respectively, in Figure 3. Attendance grows by 5pp for 2H, 18pp for 2L, and 7pp for 1L, compared to 2pp, 16pp, and 7pp, respectively, in Figure 3.

While our baseline procedure assigns a child to the lowest resource family type experienced during childhood, Figure A3 assigns a child to the family type the child experienced most often during childhood. The results in Figure A3 are also quantitatively similar to our baseline results: completion grows by 13pp for 2H, 6pp for 2L, and 7pp for 1L, compared to 11pp, 8pp, and 2pp, respectively, in Figure 3. Attendance grows by 7pp for 2H, 12pp for 2L, and 12pp for 1L, compared to 2pp, 16pp, and 7pp, respectively, in Figure 3.

A.3 College Attainment Trends By Gender

Figures A4 and A5 document college attainment trends separately by gender. The qualitative relationship between family type and college attainment continue to hold for both men and women: completion increased more for higher resource families than lower resource families, and attendance tended to increase less for higher resource families. Quantitatively, however, there is much more action among females: the average increase in completion is larger among females, and the gap in completion is larger among females. The completion share increases 15pp overall for females, and 4pp, 13pp, and 19pp for females from 1L, 2L, and 2H families, respectively. By contrast, the corresponding numbers are 7pp overall, and 0pp, 3pp, and 3pp for males from 1L, 2L, and 2H families, respectively. This is consistent with a large existing literature documenting that college attainment has increased much more rapidly for females over the last several decades.

A.4 Robustness of Time Trends in Population Shares of Family Types

Figures A6 and A7 display population shares by family type using alternative approaches to constructing variables. While our baseline procedure categorizes a family as one parent or two based on cohabitation, Figure A6 labels a family as single-parent if the head is legally unmarried, and dual-parent if the head is legally married. Prior to 1976 we do not observe whether a cohabiting couple is legally married or not, so we do not plot these years. The pattern after 1976 is comparable to our baseline results in Figure 2, except that the total single parent share is roughly 5-7pp larger when we use marital status. This is not surprising, as virtually all married couples are cohabiting, while a larger fraction of cohabiting couples are not legally married. These differences are quantitatively fairly small, however, because conditional on having children the vast majority of cohabiting couples are married.

While our baseline Figure 2 plots shares of households with children, Figure A7 plots shares of children. Therefore, in Figure A7 a household with two children receives twice the weight of a household with one child. The broad patterns are consistent with our baseline figure. One difference is that in Figure 2 the single parent share (1L+1H) peaks at 26% in 1992, but when we weight by number of children the single parent share peaks at 33% in 1995. In both figures, the single parent share declines to less than 22% by 2015.

Finally, Figure A8 replicates our baseline PSID results from Figure 2 using the Consumer Population Survey (CPS). The results are quantitatively similar, although there are some modest differences. The 1L share grows from 13% in 1968 to 22% in 2015 relative to growth from 8% to 17% in the PSID. The 2L share decreases from 70% in 1968 to 37% in 2015, relative to declines from 77% to 38% in the PSID. The 2H share grows from 16% in 1968 to 36% in 2015, relative to growth from 15% to 41% in the PSID.

A.5 Robustness of Regression Analysis

In Section I.D we presented regressions of college completion on various measures of family background. That analysis utilized a linear probability model and was unconditional on college attendance. Here we demonstrate robustness to two alternative specifications. First, we conduct the linear probability regression on the population of college attendees to verify that family type is still an important predictor of college completion, even conditional on attendance. Second, since the dependent variable is binary, we use a logit model and show results consistent with both the unconditional and conditional linear probability models.

Table A1 is the analog to Table 1 in the main text, but restricts the population to those who attend at least one year of college. Columns 1–3 show that for the full sample of cohorts turning 28 during 1995–2015, family type was an important predictor of college completion, conditional on attendance and other covariates including parent earnings, parent age, student sex, number of siblings, race, and country region. Specifically, column 3 shows that with all controls included, moving from a 2L to 1L family lowers the predicted probability of college completion by 7.0pp, and moving from 2L to 2H family raises the predicted probability of completion by 22.8pp.

Columns 4 and 5 divide the sample into cohorts turning 28 from 1995–2005 and from 2006–2015. Yet again we find that the differences by family type have increased over this two decade span. The conditional probability of completion for a reference child from a 1L family decreased slightly from 21.2% in the early cohorts to 19.5% in the later cohorts. By contrast, the conditional probability of completion for a reference child from a 2L family increased from 22.2% to 29.9%, and for a reference child from a 2H family increased from 42.6% to 55.1%.

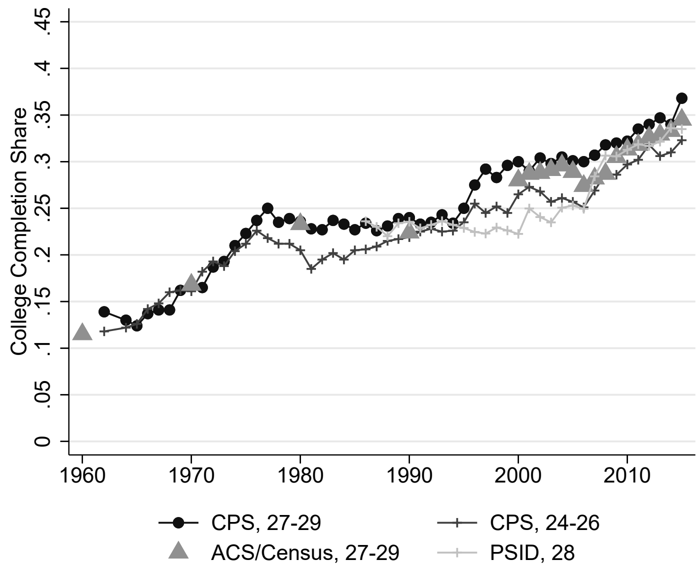
Next, we turn to the logit model results in Table A2, which are for the entire population of high school graduates, and Table A3, which are conditional on college attendance. In both we report the odds ratios rather than the regression coefficients for ease of interpretation.

The results from Tables A2 and A3 are entirely consistent with those from the linear probability model in Tables 1 and A1. Even after controlling for relevant observables, we find that family type has a large impact on the odds of college completion. Column 3 of Table

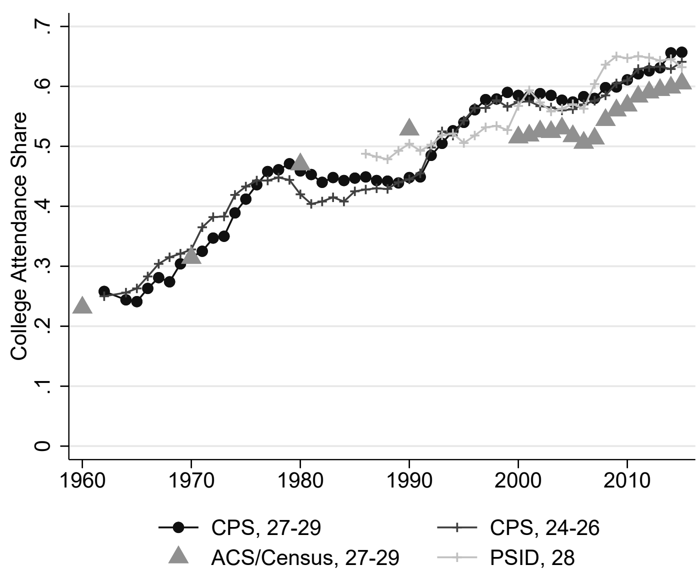
A2 shows that for a reference child from a 1L family the odds of completing college decrease by a factor of 0.69 relative to a reference child from a 2L family. Being a child from a 2H rather than a 2L family increases the odds of college completion by 3.10 times. Moreover, columns 4 and 5 show that these family effects have grown more important between the 1995-2005 and 2006-2015 cohorts. While the odds ratio for 1L children was not significantly different from one in the early cohorts, it decreased to 0.61 in the later cohorts, and became significantly less than one. Similarly, the odds ratio for 2H children increased from 2.98 to 3.29.

The story is quite similar even if we condition on college attendance. As seen in column 3 of Table A3 a reference child from a 1L family relative to a reference child from a 2L family still has lower odds of completion by a factor of 0.72. Likewise, a 2H reference child has odds of college completion 2.58 times larger than a 2L reference child. Columns 4 and 5 again divide the sample into early and late cohorts, and as before we find that family effects have grown stronger over time.

Figure A1: Aggregate College Attendance and Completion Rates in the US, 1960–2015



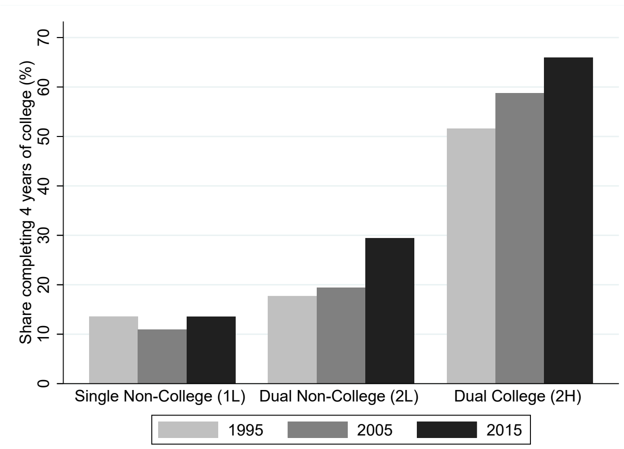
(a) College Completion



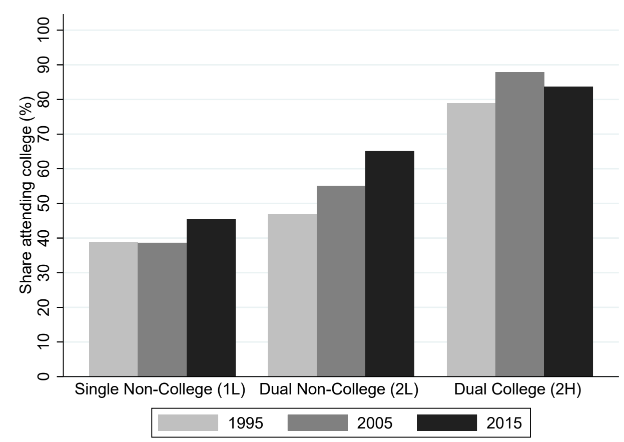
(b) College Attendance

Note: Data comes from three source: the March Consumer Population Survey (CPS), the decennial Census and American Community Surveys (ACS), and the Panel Study of Income Dynamics (PSID). For the CPS we plot college attainment at two different ages, 24-26 and 27-29. For the PSID we use one specific age, not an age range, because we observe individuals in a panel. See text for details on how we classify college attendance and completion.

Figure A2: US College Attendance and Completion Rates by Age 28: Marriage



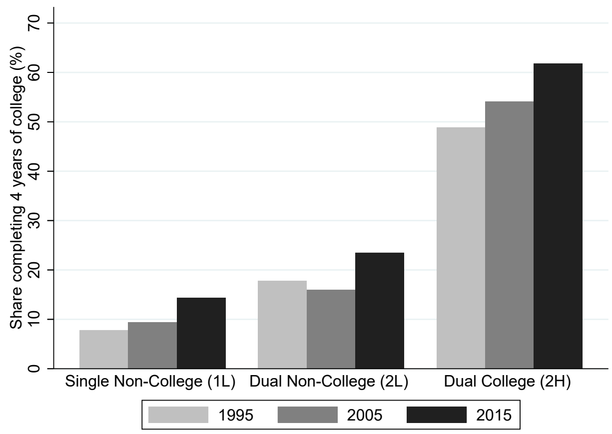
(a) College Completion



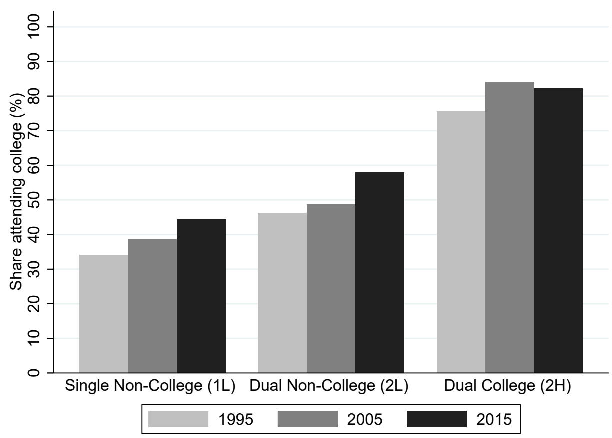
(b) College Attendance

Note: Data source is the PSID. Bars plot college attainment shares by age 28 (5-year moving average centered around year $t \in \{1995, 2005, 2015\}$. The year corresponds to the year individuals turned 28. College attendance corresponds to at least 13 years of education. College completion corresponds to at least 16 years of education. Number of parents is measured using marriage. Individuals assigned to the "minimum resource family type" they experienced from age 0-16. Attainment rates for 1H children are not included due to small sample size. See text for details.

Figure A3: US College Attendance and Completion Rates by Age 28: Mode Family Type



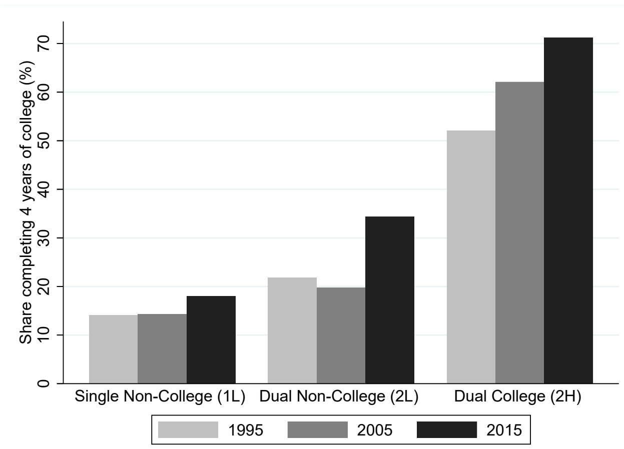
(a) College Completion



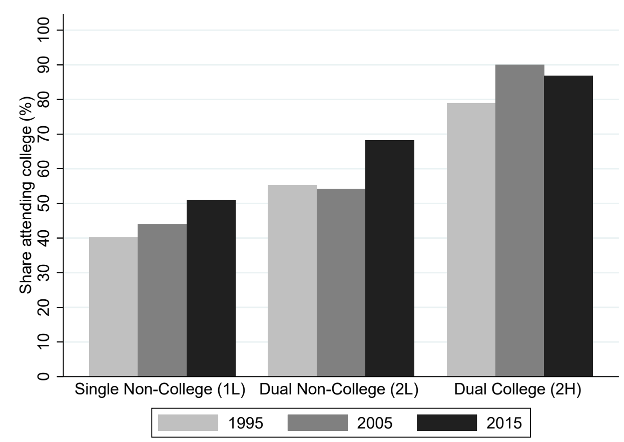
(b) College Attendance

Note: Data source is the PSID. Bars plot college attainment shares by age 28 (5-year moving average centered around year $t \in \{1995, 2005, 2015\}$. The year corresponds to the year individuals turned 28. College attendance corresponds to at least 13 years of education. College completion corresponds to at least 16 years of education. Number of parents is measured using cohabitation. Individuals assigned to the mode family type they experienced from age 0-16. Attainment rates for 1H children are not included due to small sample size. See text for details.

Figure A4: US College Attendance and Completion Rates by Age 28: Females



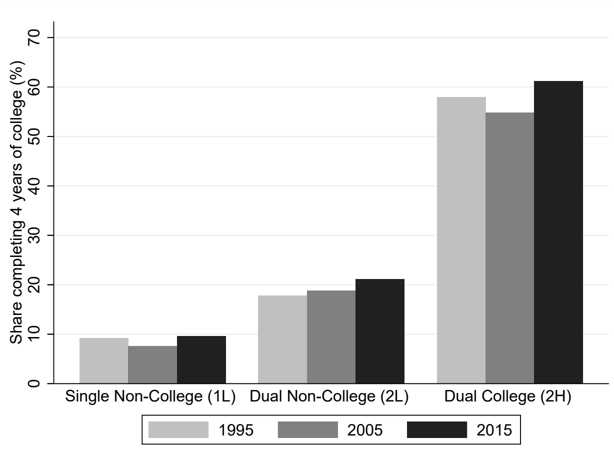
(a) College Completion



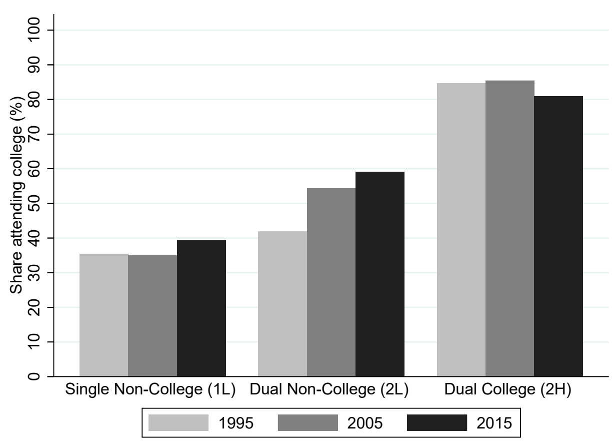
(b) College Attendance

Note: Data source is the PSID. Bars plot college attainment shares by age 28 (5-year moving average centered around year $t \in \{1995, 2005, 2015\}$. The year corresponds to the year individuals turned 28. College attendance corresponds to at least 13 years of education. College completion corresponds to at least 16 years of education. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 0-16. Attainment rates for 1H children are not included due to small sample size. See text for details.

Figure A5: US College Attendance and Completion Rates by Age 28: Males



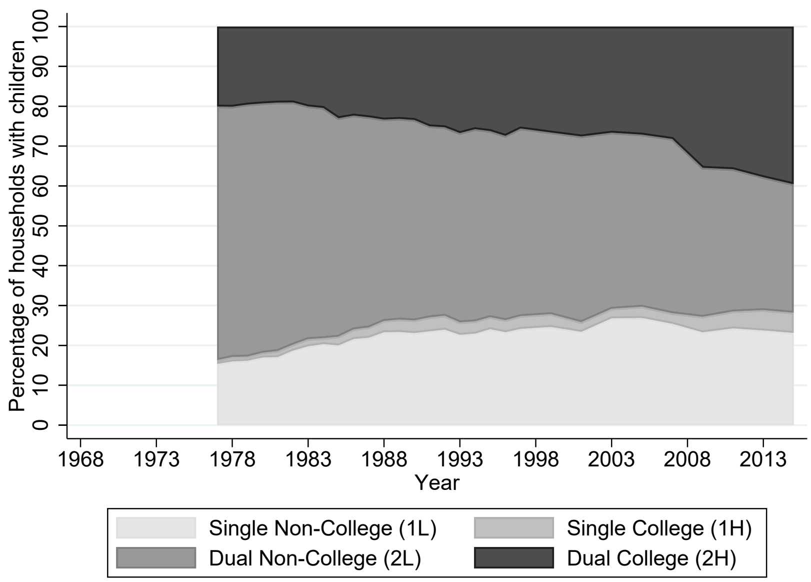
(a) College Completion



(b) College Attendance

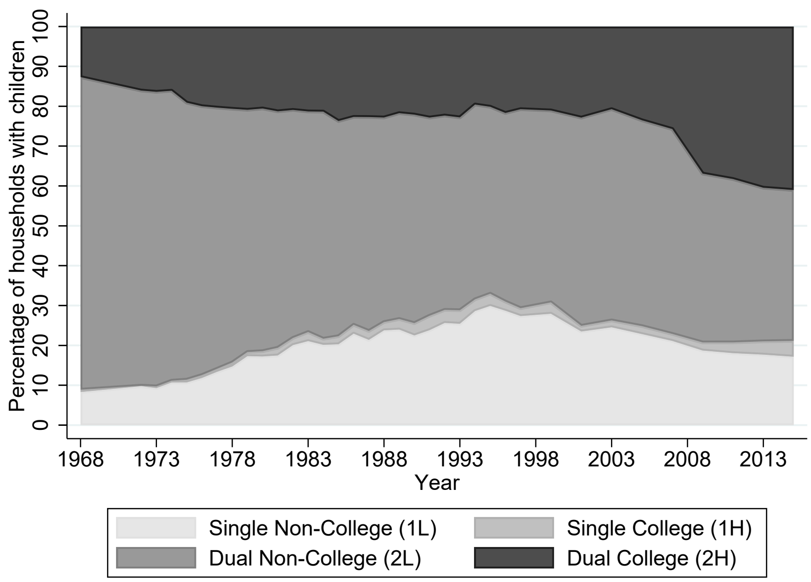
Note: Data source is the PSID. Bars plot college attainment shares by age 28 (5-year moving average centered around year $t \in \{1995, 2005, 2015\}$. The year corresponds to the year individuals turned 28. College attendance corresponds to at least 13 years of education. College completion corresponds to at least 16 years of education. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 0-16. Attainment rates for 1H children are not included due to small sample size. See text for details.

Figure A6: Population Shares of Family Types, 1968-2015: Using Marital Status



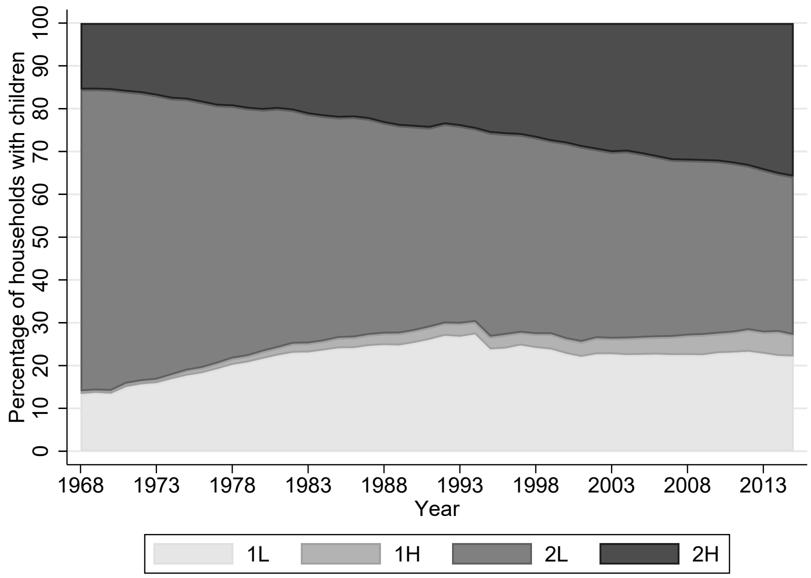
Note: Data source is the PSID. Family type is composed of number of parents (1 or 2) and education of parents (no college graduate, L, or at least one college graduate, H). Number of parents is measured using marital status. Parents are labeled college graduates if they report completing at least 16 years of education. We cannot distinguish legal marriage from cohabitation before 1976, so we do not include these years. See text for details.

Figure A7: Population Shares of Family Types, 1968-2015: Weighting By Numer of Children



Note: Data source is the PSID. Family type is composed of number of parents (1 or 2) and education of parents (no college graduate, L, or at least one college graduate, H). Number of parents is measured using cohabitation. Parents are labeled college graduates if they report completing at least 16 years of education. We exclude years 1969-71 because the education of the "spouse" was not reported for those years. See text for details.

Figure A8: Population Shares of Family Types, 1968-2005: Consumer Population Survey



Note: Data source is the Consumer Population Survey. Family type is composed of number of parents (1 or 2) and education of parents (no college graduate, L, or at least one college graduate, H). Number of parents is measured using cohabitation. Parents are labeled college graduates if they report completing at least 16 years of education. See text for details.

Table A1: Linear Probability Model for College Completion, Conditional on Attendance

	Full Sample: Cohorts age 28 from 1995–2015			1995–2005	2006–2015
	(1)	(2)	(3)	(4)	(5)
Constant	0.445	0.414	0.338	0.222	0.299
	(0.034)	(0.034)	(0.040)	(0.038)	(0.030)
1L	−0.123	−0.068	−0.070	−0.010	−0.104
	(0.021)	(0.022)	(0.022)	(0.036)	(0.028)
1H	0.126	0.110	0.094	0.175	0.051
	(0.038)	(0.038)	(0.038)	(0.066)	(0.046)
2H	0.302	0.243	0.228	0.204	0.252
	(0.020)	(0.021)	(0.021)	(0.034)	(0.027)
Parent earnings	–	0.119	0.125	0.164	0.102
		(0.014)	(0.015)	(0.027)	(0.018)
Parent Age	–	–	0.003	0.001	0.005
			(0.001)	(0.002)	(0.002)
Female	–	–	0.093	0.039	0.123
			(0.017)	(0.028)	(0.021)
# Siblings	–	–	0.002	0.036	−0.014
			(0.010)	(0.017)	(0.012)
Black	–	–	−0.040	−0.095	−0.008
			(0.040)	(0.069)	(0.049)
Region effects	–	–	X	X	X
Year effects	X	X	X	–	–
Obs.	3417	3377	3303	1249	2054
R^2	0.12	0.13	0.15	0.11	0.17

Data source is the population of college attendees, derived from the full PSID sample from 1995-2015 for columns 1–3, cohorts turning 28 from 1995–2005 for column 4, and cohorts turning 28 from 2006–2015 for column 5. College completion corresponds to at least 16 years of education. Reference group is a child from a 2L family. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 1–16. Parent earnings are the mean of log parent earnings (summing over the family head and the spouse) over ages 1–16. Parent age is the mean age of the parent(s) who report living with the child in each survey from age 1–16. Number of siblings is mean siblings in the household from age 1–16. See text for details.

Table A2: Odds Ratios from Logit Model for College Completion

	Full Sample: Cohorts age 28 from 1995–2015			1995–2005	2006–2015
	(1)	(2)	(3)	(4)	(5)
Constant	0.363	0.296	0.178	0.102	0.151
	[0.275, 0.480]	[0.218, 0.402]	[0.125, 0.253]	[0.072, 0.144]	[0.116, 0.197]
1L	0.473	0.696	0.689	0.832	0.608
	[0.392, 0.572]	[0.572, 0.847]	[0.564, 0.840]	[0.610, 1.135]	[0.470, 0.786]
1H	2.318	1.959	1.896	2.753	1.518
	[1.664, 3.230]	[1.392, 2.757]	[1.344, 2.674]	[1.437, 5.275]	[1.034, 2.229]
2H	5.088	3.199	3.099	2.981	3.287
	[4.332, 5.977]	[2.680, 3.819]	[2.583, 3.719]	[2.233, 3.980]	[2.604, 4.149]
Parent Earnings	–	2.614	2.731	3.670	2.338
		[2.216, 3.083]	[2.306, 3.235]	[2.729, 4.935]	[1.923, 2.842]
Parent Age	–	–	1.015	1.005	1.023
			[1.002, 1.028]	[0.987, 1.024]	[1.006, 1.041]
Female	–	–	1.839	1.496	2.080
			[1.578, 2.144]	[1.168, 1.916]	[1.715, 2.523]
# Siblings	–	–	1.001	1.065	0.986
			[0.917, 1.092]	[0.928, 1.222]	[0.882, 1.102]
Black	–	–	1.030	0.907	1.132
			[0.668, 1.588]	[0.468, 1.758]	[0.648, 1.976]
Region Effects	–	–	*X*	*X*	*X*
Year Effects	*X*	*X*	*X*	–	–
Obs.	5969	5882	5732	2381	3351
Pseudo R^2	0.13	0.16	0.18	0.14	0.19
Wald χ^2	671.62	704.86	744.93	242.27	464.74

Numbers in square brackets below odds ratios are the 95% confidence interval. Data source is the full PSID sample from 1995-2015 for columns 1–3, cohorts turning 28 from 1995–2005 for column 4, and cohorts turning 28 from 2006–2015 for column 5. College completion corresponds to at least 16 years of education. Reference group is a child from a 2L family. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 1–16. Parent earnings are the mean of log parent earnings (summing over the family head and the spouse) over ages 1–16. Parent age is the mean age of the parent(s) who report living with the child in each survey from age 1–16. Number of siblings is mean siblings in the household from age 1–16. See text for details.

Table A3: Odds Ratios from Logit Model for College Completion, Conditional on Attendance

	Full Sample: Cohorts age 28 from 1995–2015			1995–2005	2006–2015
	(1)	(2)	(3)	(4)	(5)
Constant	0.820 [0.595, 1.129]	0.687 [0.488, 0.967]	0.464 [0.311, 0.693]	0.273 [0.187, 0.399]	0.386 [0.284, 0.524]
1L	0.569 [0.460, 0.704]	0.736 [0.590, 0.918]	0.721 [0.574, 0.904]	0.943 [0.663, 1.340]	0.620 [0.463, 0.831]
1H	1.674 [1.166, 2.405]	1.556 [1.078, 2.245]	1.466 [1.009, 2.131]	2.165 [1.069, 4.387]	1.197 [0.788, 1.819]
2H	3.561 [2.956, 4.290]	2.704 [2.213, 3.304]	2.582 [2.103, 3.170]	2.293 [1.661, 3.168]	2.940 [2.257, 3.829]
Parent Earnings	–	1.877 [1.575, 2.236]	1.955 [1.633, 2.340]	2.408 [1.761, 3.294]	1.731 [1.400, 2.140]
Parent Age	–	–	1.015 [1.001, 1.031]	1.004 [0.982, 1.027]	1.022 [1.002, 1.042]
Female	–	–	1.563 [1.315, 1.857]	1.201 [0.913, 1.578]	1.822 [1.461, 2.271]
# Siblings	–	–	1.015 [0.920, 1.120]	1.193 [1.020, 1.395]	0.941 [0.831, 1.067]
Black	–	–	0.827 [0.515, 1.328]	0.637 [0.298, 1.363]	0.946 [0.519, 1.723]
Region Effects	–	–	X	X	X
Year Effects	X	X	X	–	–
Obs.	3417	3377	3303	1249	2054
Pseudo R^2	0.09	0.10	0.12	0.09	0.13
Wald χ^2	311.29	340.68	372.28	112.87	247.01

Numbers in square brackets below odds ratios are the 95% confidence interval. Data source is the population of college attendees, derived from the full PSID sample from 1995-2015 for columns 1–3, cohorts turning 28 from 1995–2005 for column 4, and cohorts turning 28 from 2006–2015 for column 5.. College completion corresponds to at least 16 years of education. Reference group is a child from a 2L family. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 1–16. Parent earnings are the mean of log parent earnings (summing over the family head and the spouse) over ages 1–16. Parent age is the mean age of the parent(s) who report living with the child in each survey from age 1–16. Number of siblings is mean siblings in the household from age 1–16. See text for details.

B ROBUSTNESS: ADDITIONAL HOUSEHOLD TYPES

In the main text we partition families with children into four categories based on the number and education of parents. In particular, we group together all two parent households where at least one parent has a college degree (2H). This appendix examines the robustness of the empirical results documented in Sections I–II to a finer partition of household types. In particular, here we distinguish between 2-parent households with one college degree (2HL) and those where both parents have a college degree (2HH).

Figure B1 reveals that the share of 2HH families grew somewhat faster than the share of 2HL families. From 1968–2015, 2HH families increased by 18pp, from 5% to 23%, while 2HL families increased by 8pp, from 10% to 18%. For both types, about half of the growth occurred prior to 2005, when our final cohort of interest turned 18. Because of this, one might worry that observed changes in pre-college investment, college preparedness, or college attainment among 2H households are being driven by composition effects rather than behavioral changes within relatively homogeneous groups over time. However, we document here that composition effects are not the primary driver of changes in these variables.

B.1 College Attainment Trends By Family Background

To begin, Figure B2 reveals that college completion increased by 6pp for children from 2HL families and by 11pp for children from 2HH families, compared with an 11pp increase among 2H children overall. Meanwhile, college attendance fell by 2pp among 2HL children and increased by 2pp among 2HH children, compared with a 2pp increase among 2H children overall. Thus, even after decomposing 2H households into 2HL and 2HH types, we find similar patterns for each. Namely, the change in college completion was significantly larger than the change in attendance, implying that increased completion was largely due to better preparation among 2HL and 2HH children, not simply increased attendance.

B.2 Regression Analysis: Predictors of College Completion

We provide further evidence for this point in Table B1, which repeats the regressions from Table 1 after dividing 2H children into 2HL and 2HH types. Columns 1–3 cover the full 1995–2015 sample period, regressing college completion on progressively more controls including parent earnings, parent age, sex, family size, and race. The main finding is that children from 2HL and 2HH families are much more likely to complete college than their 1L and 2L counterparts. In particular, column 3 shows that 2HL children are about 24pp more likely to complete than 2L children, and 2HH children are about 37pp more likely. For comparison, recall from Table 1 that 2H children overall held a 30pp advantage relative to 2L. It is also worth noting the completion gap between children from 1H and 2HL families. Even with all controls in place, 2HL children are 9.5pp more likely to complete college than 1H children. In both cases these children have the advantage of one college-educated parent, yet the child from a two-parent household is much more likely to complete college.

Columns 4 and 5 subdivide the sample period into two parts, and we find that the effects of family type have grown stronger over time, consistent with our evidence on growing gaps in pre-college investment and preparation. For 1995–2005 cohorts, 1L children were not significantly less likely than 2L children to complete college, but for 2006–2015 cohorts we find an 8pp gap. On the other hand, 2HL and 2HH children were 24.4pp and 30.5pp more likely than 2L children to complete college during 1995–2005. The 2HL-2L gap declined slightly to 23.5pp during 2006–2015, while the 2HH-2L gap increased to 40.4pp. We also note that the gap between 1H and 2HL children increased from 4.5pp to 12.5pp between the two time periods, further indicating the growing importance of a second parent, even if that parent was not college-educated.

B.3 Diverging Pre-College Expenditures on Children

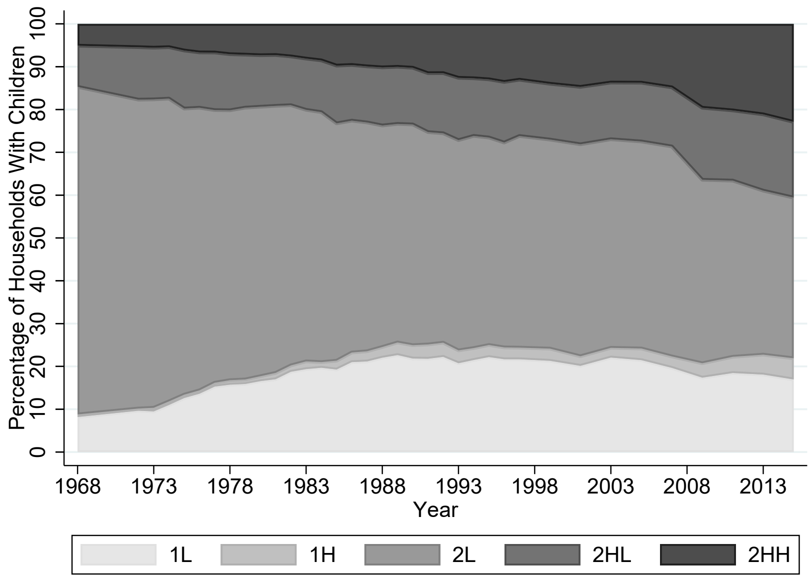
Figures B3 and B4 are comparable to Figures 4 and 5 in the main paper, except we again divide 2H children in to 2HL and 2HH types. While there are apparent level differences in mean expenditures per child between 2HL and 2HH families (as would be expected given average income differences), the crucial point for our argument is that they show similar trends over time. In particular, Figure B3 shows that mean expenditures per child grew by 42% in 2HL families and by 51% in 2HH families. There is a small composition effect, therefore, contributing to the growth rate of 53% among 2H families overall; however, even if the composition within 2H had not changed, the 2H growth rate would have still been roughly three times faster than either 1L or 2H.

Figure B4 decomposes total expenditures per child into three categories: education, childcare, and goods. As with Figure 5 in the main paper, we confirm again that education and childcare spending are primarily responsible for the growing gaps between high and low resource households, even after splitting 2H types into 2HL and 2HH. For the goods component, real expenditures per child increased only 7% for 2HH families, and was essentially unchanged for 2HL. By contrast, real education expenditures per child increased by 73% and 56% for 2HH and 2HL, respectively. Both types realized even greater growth in childcare expenditures, 99% for 2HH and 142% for 2HL. Among both groups, education and childcare account for at least 90% of the increased expenditure gap relative to 1L households, indicating that the changes for 2H overall are indeed behavioral and not simply a result of composition effects.

B.4 Diverging College Preparedness

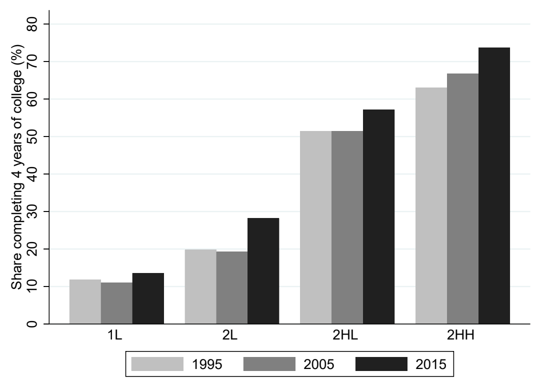
Finally, Table B2 replicates Table 2 from the main paper but with 2H children split into 2HL and 2HH types. The top panel of Figure B2 shows that both 2HL and 2HH children held large test score advantages over 1L and 2L children in the NLSY79. Additionally, the gaps relative to 1L were larger for 2HH than 2HL children throughout the distribution. (This is also true for the gaps with respect to 2L children, but to conserve space we do calculate this as a separate column.) Finally, the 2HL-1L and 2HH-1L gaps both increased from NLSY79 to NLSY97, particularly at the top of the test score distribution. At the 75th percentile, in particular, we observe the 2HL-1L gap grow from 47 to 69 points, and the 2HH-1L gap grew from 80 to 115 points. While the growth in test scores is greater among 2HH than 2HL children, which would create some composition effect within 2H overall, we also observe large and growing differences between 2HL children and those from lower resource families.

Figure B1: Population Shares of Family Types, 1968-2005

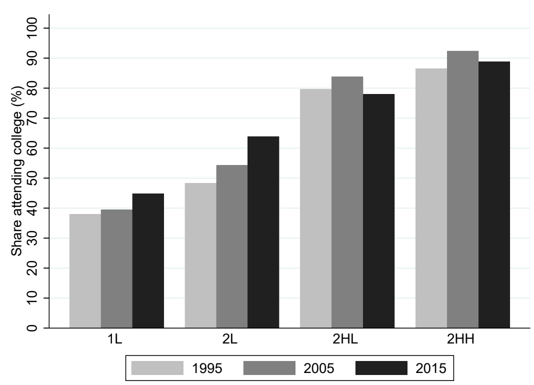


Note: Data source is the PSID. Cross-sectional distribution of family types in each year consists of the following types: single parent with no college degree (1L), two parents with no college degree (2L), two parents with at least one college degree between them (2HL), or two college-educated parents (2HH). Number of parents is measured using cohabitation. Parents are labeled college graduates if they report completing at least 16 years of education. We exclude 1969-71 because the education of the "spouse" was not reported for those years. See text for details.

Figure B2: Trends In College Completion and Attendance by Family Background



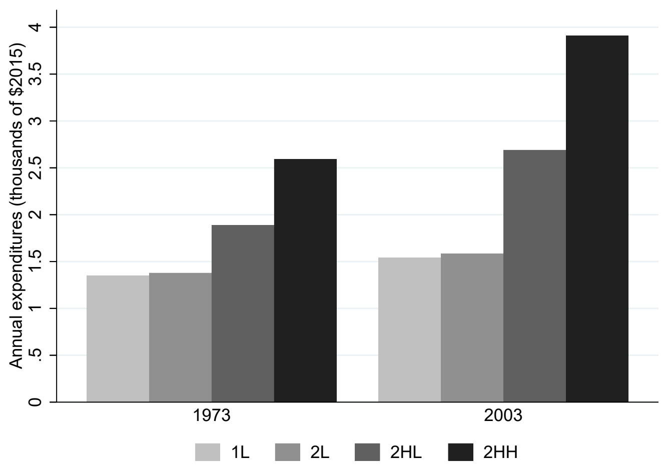
(a) College completion



(b) College attendance

Note: Data source is the PSID. Bars plot college attainment shares at age 28 by the type of family that the child grew up in: a single parent with no college degree (1L), two parents with no college degree (2L), two parents with at least one college degree between them (2HL), or two college-educated parents (2HH). Attainment rates for 1H children are not included due to small sample size. The year corresponds to the year individuals turned 28; we plot a 5-year moving average centered around year $t \in \{1995, 2005, 2015\}$. College attendance corresponds to at least 13 years of education. College completion corresponds to at least 16 years of education. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 0-16. See text for details.

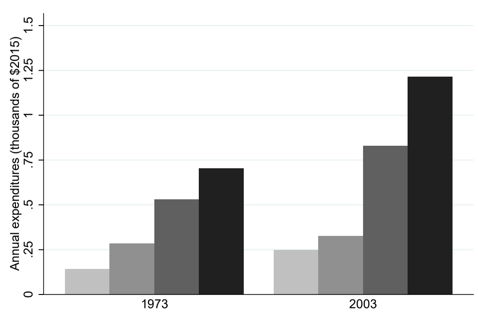
Figure B3: Trends in Annual Expenditures Per Child, By Family Type



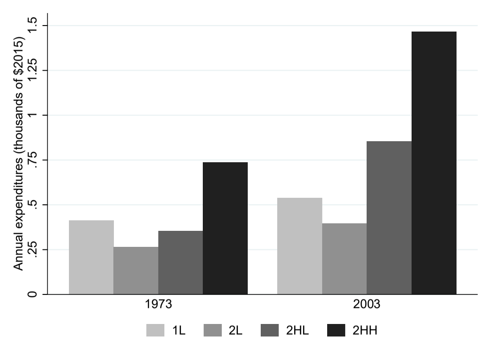
Note: Figure shows mean expenditures per child by family type in the 1973 and 2003 waves of the CEX. For details see Appendix D.

Figure B4: Trends in Categories of Child Expenditures, By Family Type

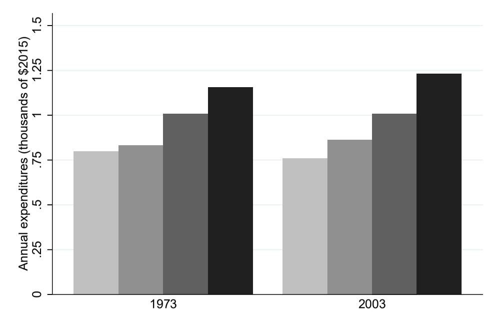
(a) Education



(b) Childcare



(c) Goods



Note: Figure shows mean expenditures per child by family type within subcategories in the 1973 and 2003 waves of the CEX. For details see Appendix D.

Table B1: Linear Probability Model for College Completion

	Full Sample: Cohorts age 28 from 1995–2015			1995–2005	2006–2015
	(1)	(2)	(3)	(4)	(5)
Constant	0.259	0.251	0.183	0.100	0.155
	(0.023)	(0.023)	(0.026)	(0.022)	(0.019)
1L	−0.102	−0.052	−0.053	−0.011	−0.080
	(0.012)	(0.013)	(0.013)	(0.020)	(0.018)
1H	0.177	0.154	0.144	0.199	0.110
	(0.028)	(0.028)	(0.028)	(0.046)	(0.035)
2HL	0.289	0.248	0.239	0.244	0.235
	(0.018)	(0.018)	(0.018)	(0.027)	(0.025)
2HH	0.468	0.388	0.373	0.305	0.404
	(0.020)	(0.021)	(0.021)	(0.038)	(0.025)
Parent earnings	–	0.099	0.101	0.136	0.083
		(0.008)	(0.008)	(0.015)	(0.010)
Parent Age	–	–	0.002	0.001	0.003
			(0.001)	(0.001)	(0.001)
Female	–	–	0.093	0.056	0.117
			(0.011)	(0.017)	(0.015)
# Siblings	–	–	−0.002	0.009	−0.006
			(0.006)	(0.009)	(0.008)
Black	–	–	−0.006	−0.015	0.007
			(0.025)	(0.039)	(0.032)
Region effects	–	–	*X*	*X*	*X*
Year effects	*X*	*X*	*X*	–	–
Obs.	5969	5882	5732	2381	3351
R^2	0.16	0.19	0.20	0.14	0.23

Data source is the full PSID sample from 1995-2015 for columns 1–3, cohorts turning 28 from 1995–2005 for column 4, and cohorts turning 28 from 2006–2015 for column 5. College completion corresponds to at least 16 years of education. Reference group is a child from a 2L family. Number of parents is measured using cohabitation. Individuals assigned to the "minimum resource family type" they experienced from age 1–16. Parent earnings are the mean of log parent earnings (summing over the family head and the spouse) over ages 1–16. Parent age is the mean age of the parent(s) who report living with the child in each survey from age 1–16. Number of siblings is mean siblings in the household from age 1–16. See text for details.

Table B2: Trends in SAT Scores, By Family Type

NLSY79						
Percentile	1L	2L	2HL	2HH	2HL-1L	2HH-1L
25	389	400	450	471	61	82
50	444	457	505	533	61	89
75	505	512	552	585	47	80
NLSY97						
Percentile	1L	2L	2HL	2HH	2HL-1L	2HH-1L
25	391	413	450	489	59	98
50	445	473	509	545	64	100
75	498	533	567	613	69	115

This table displays SAT scores among individuals who attended college in the NLSY79 and NLSY97. We display scores at the 25^{th}, 50^{th}, and 75^{th} percentiles by family type. Scores are the average of the math and verbal sections. For individuals who do not report SAT scores we impute SAT scores from ASVAB scores.

C ADDITIONAL QUANTITATIVE RESULTS

C.1 Alternative Counterfactual Exercise

Our model focuses on the role of preparation for college *completion*, so it is not particularly surprising that we do not perfectly match attendance trends in the data in the main counterfactual exercise described in Section V.A. For example, our model abstracts from changes to government loan programs, changes in the variation in college tuition within family types, and changes to social norms about the importance of college.

To evaluate how our model performs conditional on matching changes in attendance, we conduct an alternative counterfactual which is identical to our main counterfactual except that we also introduce a *residual* term for each family type so that the model exactly replicates the change in attendance from the 1995 to 2015 cohorts. Specifically, we allow the mean of the college preference shock, μ_ζ, to vary by family type. For example, in this exercise the change in μ_ζ from 1995–2015 is negative for 2H children, which dampens the too large attendance growth predicted in the main counterfactual exercise.

Table C1 shows that the changes in college completion from this exercise remain broadly in line with the changes in the data. Aggregate completion increases by 9.9pp, 93% of the increase in aggregate completion observed in the data. By family type, completion increases by 3.8pp for 1L children and 9.4pp for 2L children, and 3.9pp for 2H children. We emphasize that the exercise generates slightly higher completion gains for 2H children than 1L children despite the fact that attendance increases only 2.0pp for 2H children versus 7.0pp for 1L children.

We note that, despite matching the *change* in attendance for each family type, aggregate attendance increases 13.6pp versus 12.7pp in the data. This is because the baseline model does not exactly match the 2015 attendance rates by family type (see Figure 10c), which implies that the impact of changes in family composition from 1995-2015 differs somewhat between model and data.

C.2 Future Projections of Family Composition and College Completion

In Section V.C we showed that changing family composition alone could account for up to one-fourth of the aggregate increase in college completion observed for the cohorts turning 28 between 1995 and 2015. In this section, we ask whether the composition of family types is likely to continue impacting aggregate college completion for cohorts following 2015. To do so, we first estimate a trendline for each of our four family types using a multinomial logit regression to guarantee that at each point the sum of family type shares add up to one. We then project forward the family shares up to cohort turning 28 in 2035. Finally, using these projected family shares we can project the aggregate college completion rate under the assumption that completion rates by family type remain unchanged from their 2015 levels.

Figure C1 displays the fitted trendlines for family shares. Before discussing the projected change in college completion, there are several things to note. First, the trendlines in this figure differ somewhat from patterns documented in Figure 2 because here we are assigning each child to a single family type based on their experience from age 0-16, while in Figure 2 we are simply reporting the cross-section of family shares for each year. Second, note that the final cohort in our main analysis turned 28 in 2015. Because we assign childhood type at age 16, we observe the family shares for an additional 12 cohorts after the 2015 cohort (though we do not yet observe their educational attainment). Therefore, we only need to project family shares for the final 8 cohorts, from 2028–2035.

Relative to the 2015 cohort, we project that the 2L share will decline by 13.1pp in the 2035 cohort, while the shares of 1L, 1H, and 2H cohorts will rise by 4.4pp, 1.0pp, and 7.6pp, respectively. Absent further changes in college completion rates by family type, this change in the composition of family types would increase aggregate college completion by 2.4pp. For comparison, this is nearly as large as the 2.8pp increase due to family composition we computed over the 1995–2015 period. In summary, it is quite plausible that the aggregate college completion rate will continue to increase in the near future purely due to changes in family composition, even in the absence of further behavioral changes.

C.3 Robustness: Alternative Values for the Substitution Elasticity of Investments

In Section IV.A we discussed the range of estimates in the literature for ρ, which governs the substitutability between time and market investments. Given the CES functional form of the human capital production function in (18), the elasticity of substitution between market and time investments is $1/(1-\rho)$. When $\rho < 0$ the two inputs are complements; when $\rho > 0$ the two are substitutes. Our baseline calibration assumed the Cobb-Douglas case of $\rho \to 0$.

Table C2 compares the results of our baseline counterfactual with calibrations using two alternative values from the literature. Zhu and Vural (2013) estimate a value of $\rho = -0.9$, implying a substitution elasticity of 0.53, while Sommer (2016) uses a value of $\rho = 0.7$, which implies a substitution elasticity of 3.3. A key takeaway is that, in response to the increase in the college premium, families with college-educated parents (2H) substitute towards market investments more intensively when ρ is larger. This is intuitive, since a rise in the college premium increases the wage rate of college-educated parents, which increases their opportunity cost of time investments. By contrast, this pattern is not present for non-college parents who do not experience any change in the opportunity costs of time investments.

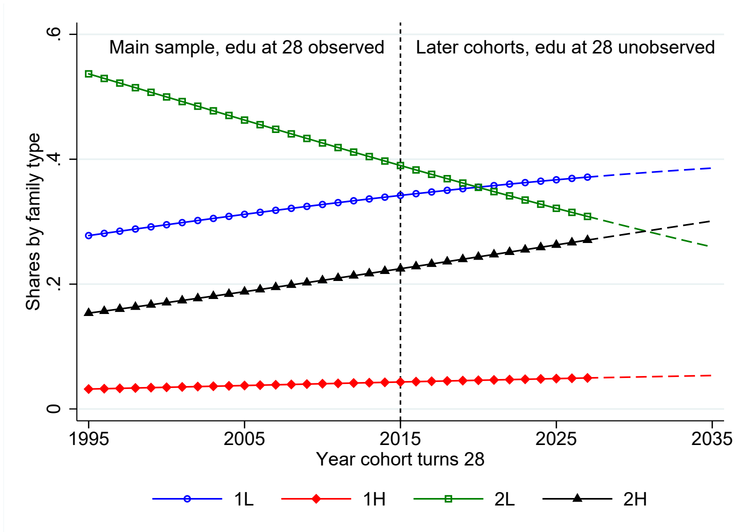
Notably, the resulting change in college completion is fairly stable for different values of ρ. The intuition for this result is that all other parameters in Table 4 are not held constant as we vary ρ, but are instead recalibrated according to the same procedure as in Section IV.B. In particular, the calibrated value of θ_m, the CES weight on market investments, is declining in ρ: it is 0.041, 0.037, and 0.030 when ρ is -0.9, 0, and 0.7, respectively. If θ_m were held constant as ρ increased, 2H families quickly begin to substitute towards market investments, which produces an overall mean level of market investments that is too high relative to the calibration target. Therefore, the joint calibration routine responds by lowering θ_m as ρ increases. This lower weight on market investments in turn limits the extent to which 2H families can substitute market investments for time investments, even for larger values of ρ.

Table C1: Counterfactual Changes in Investments, Preparation, and College Attainment, $1995-2015$

			Counterfactuals	
Variable	Family Type	Data	Main	Alternative
Δ Completion share	1L	0.018	0.047	0.038
	2L	0.084	0.067	0.094
	2H	0.111	0.124	0.039
	Agg	0.106	0.094	0.099
Δ Attendance share	1L	0.070	0.093	0.070
	2L	0.155	0.100	0.155
	2H	0.020	0.147	0.020
	Agg	0.127	0.121	0.136
Δ Mean Market Investments	1L	$106	$35	$33
	2L	$43	$99	$108
	2H	$429	$319	$287
	Agg	$208	$163	$165
Δ Mean Time Investments (hours/week)	1L	1.75	0.80	0.74
	2L	3.10	1.16	1.27
	2H	4.41	1.63	1.35
	Agg	3.09	1.33	1.36
Δ Mean Human Capital Gap	2L−1L	13.3	2.9	4.4
	2H−1L	7.1	7.8	5.7

Note: This table displays changes in college attainment by family type in the data and model. For the data column, changes in completion and attendance are from the PSID; changes in mean human capital (SAT score) are from the NLSY79-97. For the model columns, changes are expressed in attainment shares relative to the baseline economy. See text for details.

Figure C1: Family type shares, fitted values from multinomial logit regression



Note: This figure displays the fitted trendlines for population shares by family type. Trendlines are constructed using empirical population shares for cohorts turning 28 from 1995–2027. We then use these trendlines to extrapolate population shares until the 2035 cohort.

Table C2: Change in Investments and College Completion for Different Substitution Elasticities

Variable	Family Type	$\rho = -0.9$	$\rho \to 0$ (Baseline)	$\rho = 0.7$
Δ Completion share	1L	0.047	0.047	0.049
	2L	0.067	0.067	0.070
	2H	0.125	0.124	0.128
	Agg	0.094	0.094	0.096
Δ Mean Market Investments	1L	\$36	\$35	\$38
	2L	\$109	\$99	\$102
	2H	\$247	\$319	\$1109
	Agg	\$138	\$163	\$447
Δ Mean Time Investments (hrs/wk)	1L	0.75	0.80	0.80
	2L	1.15	1.16	1.18
	2H	1.66	1.63	1.54
	Agg	1.33	1.33	1.29

Note: This table displays changes in pre-college investments and college completion by family type in the model for different values of the substitution parameter ρ. Changes are expressed relative to the baseline. See text for details.

D DATA DETAILS

D.1 National Longitudinal Surveys of Youth (NLSY79 and NLSY97)

In both the NLSY79 and NLSY97 (Bureau of Labor Statistics, 1979–2016, 1997–2017), our sample consists of individuals who met two criteria. The first was that individuals remained in the survey until at least age 27 or 28, so that we could measure educational attainment by age 28. Consistent with our PSID methodology, we say that an individual attended college if they report completing at least 13 years of school by age 28. We say that an individual completed college if they report completing at least 16 years of school by age 28.

The second sample criteria is that we are able to assign the individual to a single family type. While the PSID allowed us to observe the family type during each year of an individual's childhood, the NLSY does not begin interviewing children until age 12 or later. Instead, family background information is primarily communicated through retrospective questions. We do our best to replicate our PSID methodology using information available in the NLSY. Specifically, in the NLSY79, we say that an individual had two parents if they report that they resided with two biological parents until at least 16 years of age. The NLSY97 does not contain this exact question, but it does ask about family background at three ages of childhood: 2, 6, and 12. For this sample we say that an individual had two parents if they report residing with two parents in all three of these questions.

Table D2 displays the share of individuals assigned to each family type in the NLSY79 and NLSY97. Individuals in these surveys turned 18 between 1976-1983 and 1998-2002. For comparison, Table D2 also lists the analagous information for the PSID cohorts who turn 18 in 1985 and 2005. Note that our first PSID cohort turns 18 in 1985, so the NLSY79 cohort turned 18 earlier than the PSID 1985 cohort. The NLSY and PSID shares are largely consistent with each other.

Note that the shares of individuals assigned to a given family type are different from family type shares in the cross section (see Figure 2). The reason is that individuals are assigned to the "lowest resource" type they experience during their childhood. Therefore, the share of children assigned to 1L will be larger than the cross sectional 1L share for any given year, and the share of children assigned to 2H will be smaller than the cross sectional 2H share for any given year.

Table D3 displays the share of individuals who attended and completed college by age 28, by family type, in the NLSY79 and NLSY97. Again, for comparison, Table D2 also lists the analagous information for the PSID cohorts who turn 18 in 1985 and 2005. Again, the NLSY and PSID attainment rates are fairly consistent with each other. One significant difference is that attendance rates are lower in the NLSY79 sample than the PSID sample. This is unsurprising: the NLSY79 cohort on average turned 18 in 1979, and aggregate attendance increased by several percentage points between this cohort and the cohort that turned 18 in 1985.

Our main purpose in utilizing the NLSY data is to access information on college preparedness, college attendance decisions, and college completion rates. Our preferred measure of college preparedness is a student's average score on the math and verbal sections of the Scholastic Aptitude Test (SAT). One potential problem, however, is that SAT test takers are highly selected, i.e., they tend to be students who are more prepared and therefore more likely to attend and complete college. In the NSLY79 data only 11 percent of high school graduates in our sample have SAT scores, and for the NLSY97 this figure is 20 percent. Estimating the college graduation probability function $\gamma(h)$ only on the population of SAT test takers, therefore, would overstate the probability of completion for the average student.

To combat this issue of selection, we impute SAT scores for survey respondents who did not take the test. For this task we utilize ASVAB test scores, which are reported for 94 percent of high school graduates in our NLSY79 sample and 80 percent in our NLSY97 sample. For NLSY79, we impute SAT math scores from the ASVAB arithmetic reasoning and mathematics knowledge score, and we impute SAT verbal scores from the word knowledge and paragraph comprehension score. For NLSY97, we imputed a combined average SAT math and verbal score from the ASVAB combined math-verbal percentile score.

As described in Section II.B, we estimate the following weighted logit regression on individuals who attend college in NLSY79 and NLSY97 data:

$$college\_grad = \beta_0 + \beta_1 SAT\_score + \varepsilon, \tag{D1}$$

where *college_grad* represents the binary outcome of either graduating college or not. Table D1 shows these regression results.

Table D1: Logit regression results

	NLSY97	NLSY79
β_0	-4.589	-5.543
	(4e-4)	(4e-4)
β_1	0.010	0.012
	(8e-7)	(8e-7)
Pseudo R^2	0.121	0.142

To verify the validity of our approach to imputing SAT scores from ASVAB scores, we offer two pieces of evidence. First, we note that SAT and ASVAB scores are highly correlated among individuals for whom both are available. Among our sample of high school graduates in the NLSY79 the correlation is 0.89, and in the NLSY97 it is 0.85. Second, we estimate logit regressions of college graduation (conditional on attendance) on ASVAB scores, just as we did for SAT scores above. We then predict graduation probabilities for students who do not have SAT scores using ASVAB scores instead. By comparing the probability of graduation given their ASVAB score against the probability of graduation using their imputed SAT score (as in Figure 8), we can verify whether the predicted probabilities are similar. Figure D1 shows for both NLSY79 and NLSY97 that the probability of college completion is nearly identical using both methods for students who do not take the SAT. Hence, whether we use ASVAB scores or a combination of actual and imputed SAT scores to estimate the graduation probability does not affect the results.

Figure D1: Predicted graduation probabilities using SAT and ASVAB scores

(a) NLSY79 (b) NLSY97

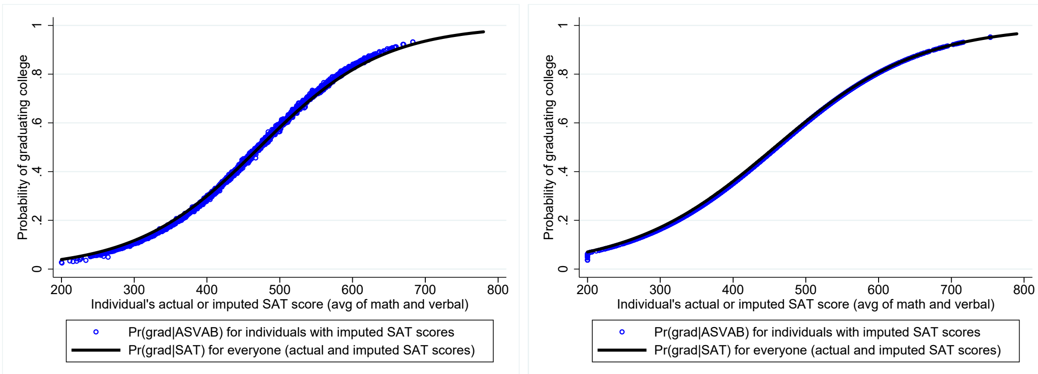


Table D2: Share of children assigned to each family type in the PSID and NLSY

Family type	1979 NLSY	1986 PSID	1997 NLSY	2000 PSID	Δ NLSY	Δ PSID
1L	0.262	0.269	0.358	0.327	0.096	0.058
1H	0.020	0.024	0.088	0.047	0.068	0.023
2L	0.545	0.565	0.337	0.404	−0.208	−0.160
2H	0.173	0.142	0.217	0.221	0.044	0.079

Table D3: Attendance and completion rates by family type in the PSID and NLSY

ATTENDANCE RATE						
Family type	1979 NLSY	1986 PSID	1997 NLSY	2000 PSID	Δ NLSY	Δ PSID
Agg	0.451	0.518	0.616	0.647	0.164	0.129
1L	0.321	0.385	0.433	0.521	0.112	0.136
2L	0.383	0.498	0.601	0.603	0.218	0.105
2H	0.831	0.829	0.895	0.884	0.064	0.055
COMPLETION RATE						
Family type	1979 NLSY	1986 PSID	1997 NLSY	2000 PSID	Δ NLSY	Δ PSID
Agg	0.226	0.225	0.334	0.313	0.108	0.088
1L	0.107	0.116	0.147	0.127	0.041	0.012
2L	0.163	0.196	0.290	0.250	0.127	0.053
2H	0.579	0.531	0.682	0.684	0.103	0.152

D.2 Consumer Expenditure Survey (CEX)

We utilize CEX data from the 1973, 1983, 1993, and 2003 interview surveys (Bureau of Labor Statistics, 1973–2003). Because we only examine pre-college investment expenditures, we drop households that have no children age 0–18. We also drop households that are missing data on marital status or parent education, else we are unable to classify these family types as 1L, 1H, 2L, or 2H.

Our measure of total expenditures on children consists of three subcategories: goods, childcare, and education. "Goods" includes: clothing and shoes for boys, girls, and infants; infant furniture and equipment; bicycles/tricycles, toys, and games; playground equipment; musical instruments, supplies, accessories; and computers, software and repair (non-business use). Following Kornrich and Furstenberg (2013) we also include the service, rental, and repair for these goods where available, for consistency with later data. "Childcare" includes tuition, books, supplies, equipment, and other supplies for daycare, nursery, and other schools, as well as babysitting. "Education" includes: newspapers, magazines, books, encyclopedias, and reference books; tuition, books, supplies and equipment, for primary/secondary schools. Notably, we exclude post-secondary education expenses because our model and empirical analysis focus on pre-college investments. In all categories we exclude expenditures that are categorized as gifts for individuals outside of the household.

Unfortunately, the Universal Classification Codes (UCC) used to identify spending categories can change slightly from year to year as items are added, deleted, and sometimes aggregated or disaggregated. We make every attempt to select UCC codes in each survey year so that our measures are as consistent as possible across surveys. Listed below are the codes we include in each subcategory for each survey year. Codebooks for each year describe the precise goods and services included for each UCC code.
UCC codes for 1973:

- Goods: 6066 – 6068, 6076 – 6077, 6079, 7053 – 7057, 7271 – 7273, 7281 – 7282, 7284, 7286, 10001 – 10079, 12001 – 12079, 13001 – 13091, 21020 – 21025, 21050 – 21051, 21057 – 21059, 21069, 21083 – 21086.
- Childcare: 6007, 6009, 23001, 23008, 23015, 23016.
- Education: 7284, 21070, 22001 – 22005, 23002, 23005, 23006, 23009, 23012, 23013, 23017, 23018, 23025, 23026, 23027, 23029.

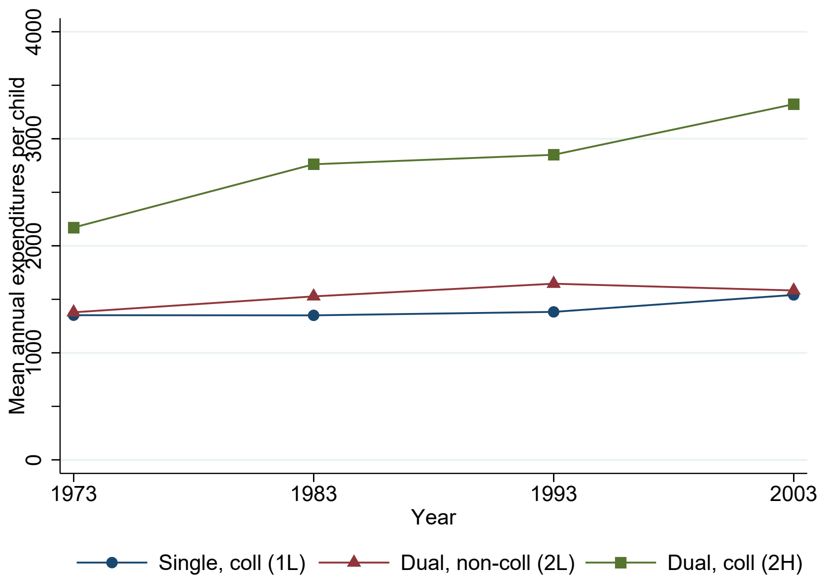
UCC codes for 1983:

- Goods: 290420, 320130, 370110 – 370902, 390110 – 390902, 400210, 400220, 410111 – 410904, 600310, 610110, 610120, 610130, 620904, 690110.
- Childcare: 340210, 660900, 670310.
- Education: 590111 – 590230, 660310, 660210, 670210, 670901, 670902.

UCC codes for 1993:

- Goods: 290420, 320130, 370110 – 370902, 390110 – 390902, 400210, 400220, 410110 – 410901, 600310, 610110, 610120, 610130, 620904, 690111 – 690113.
- Childcare: 340210 – 340212, 660900, 670310.
- Education: 590110 – 590230, 660310, 660210, 670210, 670901, 670902.

Figure D2: Average annual expenditures per child by family type



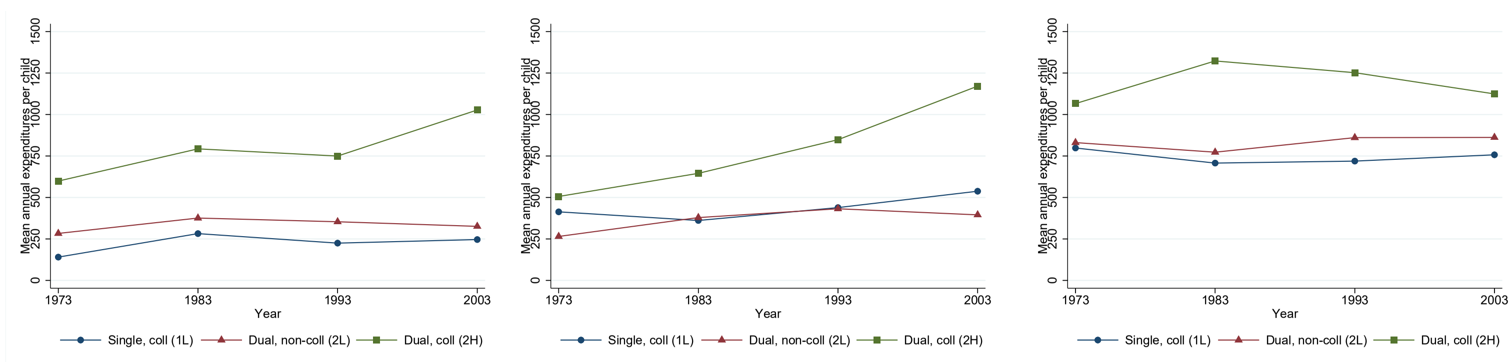
UCC codes for 2003:

- Goods: 290420, 320130, 370110 – 370904, 390110 – 390902, 400210, 400220, 410110 – 410901, 600310, 610110, 610120, 610130, 620904, 690111 – 690113.
- Childcare: 340211, 340212, 660900, 670310.
- Education: 590111 – 590230, 660310, 660210, 670210, 670901, 670902.

In the main paper we show results for the 1973 and 2003 surveys to illustrate the long-run changes. Figures D2 and D3 include data for 1983 and 1993 to demonstrate that intervening years are consistent with the longer run trends. Figure D2 shows that growth in total expenditures per child in 2H families was relatively steady and exceeded that for 1L and 2L families throughout this 30 year span. Figure D3 breaks total expenditures into the three subcategories we consider – education, childcare, and goods. Across all three we find that 2H households have the highest levels and long-run growth in real spending per child.

Figure D3: Categories of child expenditures by family type

(a) Education (b) Childcare (c) Goods



D.3 American Heritage Time Use Study (AHTUS)

Our time use analysis utilizes the harmonized AHTUS/ATUS data available from IPUMS (Fisher et al., 2018). We include individuals age 18 to 64 who report having at least one coresident child under age 18. Parents are categorized as married if they are legally married or cohabiting (available post-1998). Parents are categorized as college-educated if they report having a college degree or higher. We exclude individuals who are missing data on age of youngest child, marital status, sex, education, and number of children under 18. We also exclude "low quality" data indicated by the variable LOWQUAL.

IPUMS aggregates time use activities across all survey years into 98 categories. Our "narrow" measure used in the main paper mirrors the same time use categories as the "ACT_CHCARE" variable created by IPUMS. Specifically, it includes the following (numbered by IPUMS activity code):

- 33: Care of infants
- 34: General care of older children
- 35: Medical care of children
- 36: Play with children
- 37: Supervise child or help with homework
- 38: Read to, talk with child
- 39: Other child care

We also construct a "broader" measure that includes all of the above plus:

- 65: Physical activity, sports with child
- 96: Travel related to child care

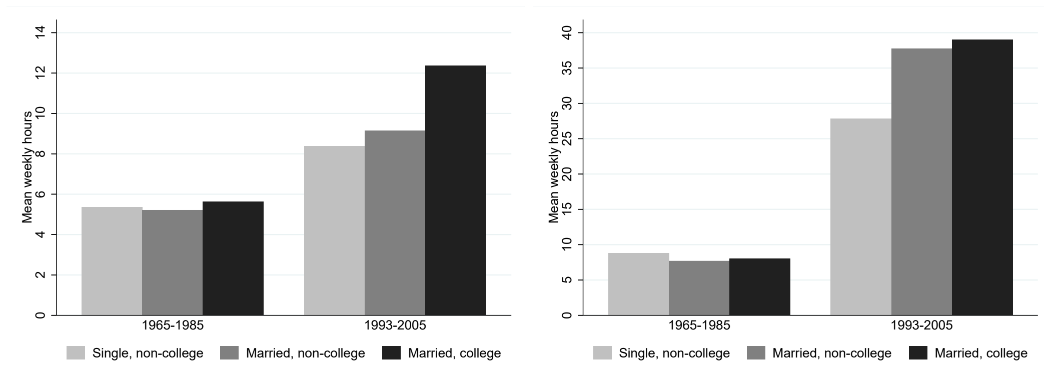
Finally, our "broadest" measure of child time investment includes all of the above, plus *any* time spent doing any activity when at least one child under 18 was present.

Figure D4 depicts the two alternative measures of child time investment just described. Read in tandem with Figure 6, these graphs show some level differences but otherwise nearly identical results. Regardless of how broadly one measures time spent with children, parents of all types spent roughly equal amounts of time (per parent) with children during the 1965–1985 period. Furthermore, while child investment time has grown among all parents, that growth has been greatest among married, college-educated parents and least among single, non-college parents.

Finally, we make a brief point about which activities have experienced decreased time use in order for parents to allocate additional time to childcare. Toward this end, we categorized all time use into five main groups: childcare (our "narrow" measure from the main paper), home production (cleaning, cooking, and other unpaid domestic work), paid work, leisure (recreational activities, sleep, personal care, etc.), and education (attending classes, studying, doing homework, etc.). Notably, this categorization intentionally departs from some previous work such as Ramey and Francis (2009) because they include various childcare activities as either leisure or home production, whereas we necessarily need to separate childcare from other uses of time.

Figure D4: Alternative measures of child time investment per parent, by marital status and education

(a) "Broader" measure of time with children (b) "Broadest" measure of time with children



We have three main findings from this exercise. First, average daily time spent on education is essentially unchanged. Second, the average daily time spent on home production decreased by 0.37 hours, which almost exactly offsets the increase of 0.39 average daily hours spent doing paid work. Third, the increase of 0.63 average daily hours of childcare time is almost exactly offset by a decrease in leisure time of 0.67 average daily hours. Notably, these aggregate results are also remarkably similar across family types. We conclude that the increase in childcare time is largely coming at the expense of leisure, and the increase in paid work time likely represents substitution with home production through increased outsourcing of domestic services like cooking, cleaning, and lawn care.

D.4 National Postsecondary Student Aid Study (NPSAS)

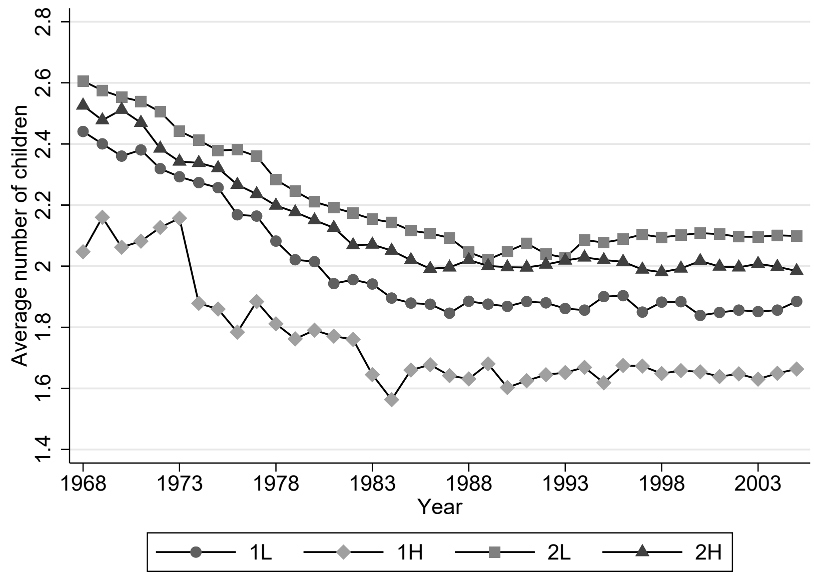
This section describes our use of the National Postsecondary Student Aid Study to calculate the tuition costs of college utilized in the quantitative model.

D.4.1 College Tuition

Our measure of college tuition cost in the NPSAS (U.S. Department of Education, 1987–2008) data is tuition and fees minus all grants for full-time, full-year students attending four-year public and private not-for-profit universities. For each year the data is available (1987, 1990, 1993, 1996, 2000, 2004, and 2008) we regress the tuition cost on dummies for parent's education (college or non-college) and marital status (married or not married). Given these regression coefficients we then predict average college cost for students of each family type: 1L, 1H, 2L, and 2H. We linearly interpolate for years between 1987 and 2008 without data observations, and we extrapolate back to 1985 to estimate costs for our first birth cohort.

For the benchmark model, which is calibrated to match the cohort turning 18 in 2005, we set the annual college tuition cost τ_m equal to the average of the college costs we estimate for 2005–2008. For the counterfactual exercise covering the cohort turning 18 in 1985, we set the annual college tuition cost τ_m equal to the average college costs we estimate for

Figure D5: Average number of children by family type



1985–1988. In the main paper, tuition values for the benchmark calibration were listed in Table 3, and those for the counterfactual appeared in footnote 25.

D.5 Additional Information on Family Heterogeneity

This section briefly describes data on additional dimensions of family heterogeneity, namely trends in number of children by family types and the gender composition of single parents.

D.5.1 Trends in Number of Children

Fertility rates vary by the number and education of parents, and they have changed over time among all family types. Figure D5 shows the average number of children under 18 by family type in the CPS between 1968 (the year our first cohort of interest was born) and 2005 (the year our last cohort of interest turned 18). The average number of children has decreased by about 0.5 among all family types, but the differences between family types has been remarkably stable. 1L, 2L, and 2H families show very similar and parallel trends in average number of children, yet they show very different investment trends, as discussed in Section II. This is the main reason we do not view changing fertility as a likely candidate to explain the investment or preparedness trends. Furthermore, we control for number of siblings in the empirical analysis (see Tables 1, A1, A2, A3, and B1), and with few exceptions this has no significant effect on the likelihood of college completion.

D.5.2 Gender of Single Parents

Our model does not distinguish the gender of parents, but this is a potentially important dimension, especially in the case of single parents if male and female single parents invest differently in their children. However, this is a challenging issue to study given that the overwhelming majority of single parents are female. CPS data on the historical living arrangements of children (`https://www.census.gov/data/tables/time-series/demo/families/children.html`) indicate only 1.1% of children under 18 in 1968 lived with a single father (this is 9.3% of all children living with a single parent), and this increased to about 4.8% living with a single father by 2005 (16.9% of all children living with a single parent). We find very similar patterns in our PSID sample. With respect to the education of single fathers, we also calculate that 97% of single fathers were non-college educated in 1968, and this share decreased to 77% by 2005.

Nevertheless, to study the potential impact of single parent gender, we ran a version of the regression in Table 1 (full results available upon request) in which we split 1L and 1H children into four groups – 1LM and 1LF, 1HM and 1HF – where the additional M or F indicates whether the child lived with a Male or Female parent during the majority of their childhood. The point estimates indicate that children of single male parents (1LM and 1HM) have lower probability of college completion than comparable children of single female parents (1LF and 1HF). However, when testing equality of coefficients on 1LM versus 1LF, we cannot reject the null hypothesis that the two are equal at the 5% level (F-stat = 3.00, p = 0.083). We do find that 1HM and 1HF are significantly different (F-stat = 7.99, p = 0.005), but we also note that 1HM children account for less than 0.5% of children in most cohorts and thus do not contribute significantly to the macroeconomic trends we study in this paper.

E DEFINITION OF EQUILIBRIUM FOR QUANTITATIVE MODEL

The state vector for parents during childhood is $x = (P, S, H, a)$ and after childhood is $x = (P, S, H, K)$. The state vector for children at $j = 2$ prior to the college decision is $x = (P, S, h, \zeta)$ and at $j = 2$ after completing all education is $x = (P, S, s, h, \mathbb{I}_a)$. After children enter the labor market ($j = 3, ..., J$) the state vector is $x = (s, h, k)$. An equilibrium for the model economy is: a collection of parental decisions during childhood ($j = 1$) $\{C_1(x), K_2(x), N_1(x), m_1(x), i_1(x)\}$; parental consumption, saving, and labor decisions after childhood ($j > 1$) $\{C_j(x), K_{j+1}(x), N_j(x)\}$; child college attendance decisions at college age ($j = 2$) $s(x)$; and child consumption, saving, and labor decisions from college age-on ($j > 2$) $\{c_j(x), k'_j(x), n_j(x)\}$ such that:

1. During childhood parents solve the decision problems defined in Section III.B.
2. After childhood parents solve the decision problem defined in Section III.E.
3. College aged children solve the decision problems defined in Section III.C.
4. From college age on, children solve the decision problems defined in Section III.D.

References

Blandin, Adam, and Christopher Herrington. 2021. "Supplementary data for: Family Heterogeneity, Human Capital Investment, and College Attainment." Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI. `https://doi.org/10.3886/E145622V1` (accessed July 20, 2021).

Bound, John, Michael F Lovenheim, and Sarah Turner. 2010. "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources." *American Economic Journal: Applied Economics*, 2(3): 129–157.

Bureau of Labor Statistics. 1973–2003. "Consumer Expenditure Survey." ICSPR. `https://www.icpsr.umich.edu/web/ICPSR/series/20` (accessed June 2018).

Bureau of Labor Statistics. 1979–2016. "National Longitudinal Survey of Youth 1979." NLS Investigator. `https://www.nlsinfo.org/investigator/pages/login.jsp` (accessed June 2017).

Bureau of Labor Statistics. 1997–2017. "National Longitudinal Survey of Youth 1997." NLS Investigator. `https://www.nlsinfo.org/investigator/pages/login.jsp` (accessed June 2017).

Fisher, Kimberly, Jonathan Gershuny, Sarah M. Flood, Joan Roman Garcia, and Sandra L. Hofferth. 2018. "American Heritage Time Use Study." IPUMS, Minneapolis, MN. `https://www.ahtusdata.org` (accessed January 2019).

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and Robert J. Warren. 2020. "Current Population Survey." IPUMS, Minneapolis, MN. `https://cps.ipums.org/cps/` (accessed February 2021).

Kornrich, Sabino, and Frank Furstenberg. 2013. "Investing in children: Changes in parental spending on children, 1972–2007." *Demography*, 50(1): 1–23.

Panel Study of Income Dynamics. 1968–2015. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. `https://simba.isr.umich.edu/default.aspx` (accessed October 2017).

Ramey, Valerie A, and Neville Francis. 2009. "A century of work and leisure." *American Economic Journal: Macroeconomics*, 1(2): 189–224.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. "IPUMS USA." IPUMS, Minneapolis, MN. `https://usa.ipums.org/usa/` (accessed August 2018).

Sommer, Kamila. 2016. "Fertility choice in a life cycle model with idiosyncratic uninsurable earnings risk." *Journal of Monetary Economics*, 83: 27–38.

U.S. Department of Education. 1987–2008. "National Postsecondary Student Aid Study." National Center for Education Statistics, Washington, DC. `https://nces.ed.gov/surveys/npsas/index.asp` (accessed December 2019).

Zhu, Guozhong, and Gulfer Vural. 2013. "Inter-generational effect of parental time and its policy implications." *Journal of Economic Dynamics and Control*, 37(9): 1833–1851.