Online Appendix to

Superstar Firms and College Major Choice

Darwin Choi
Hong Kong University of Science and Technology
dchoi@ust.hk

Dong Lou
Hong Kong University of Science and Technology
London School of Economics and CEPR
dlou@ust.hk; d.lou@lse.ac.uk

Abhiroop Mukherjee
Hong Kong University of Science and Technology
amukherjee@ust.hk

In this Online Appendix, we report additional analyses of the relation between superstar firms and college major choice. We then provide further evidence for the role of salience in driving student major choice. Finally, we provide additional details on the college-graduate survey that we conducted on the SurveyMonkey platform.

1 Additional Empirical Evidence

1.1 Robustness of the Main Results

We conduct a number of robustness checks for our main results in Table A2. In the first row, we redo our main tests with biological sciences and health majors included. In the paper, Table 2 excludes these majors because students with biology- and health-related degrees often go to graduate schools before entering the job market (so do not start working in year t). Here we show that the results are similar if we include these majors.

In rows 2 and 3, we show that our results remain highly statistically significant with bootstrapped standard errors and block-bootstrapped standard errors. In the next two rows, we examine the robustness of our results in Table 2 to alternative ways of calculating return measures and other controls at the major level. First, employment-weighted cross-sectional return skewness is calculated within each NAICS 3-digit industry. Then we aggregate across all industries associated with the major. Row 4 takes the equal-weighted average skewness across industries. Row 5 uses the industry with the maximum absolute value. Other industry-level variables are aggregated to the major level in the same way.

Rows 6 to 8 of Table A2 show that our results based on the $tail_N$ measure in Table 6, Panel A are robust to other choices of N (1, 3, and 5), albeit with weaker statistical significance. Next, we show that leaving out the Tech boom years (1990–2002) from our analysis does not change our results materially.

In the final row, we examine the number of graduates with a master's degree (instead of bachelor's) in related fields. The results are similar to those reported in the paper for bachelor's degrees. A one-standard-deviation increase in *Skew* measured in years t-7 to t-3 is associated with an 5.57% increase in the number of students graduating with a master's degree in related fields in year t. Note that we do not have sufficient data here to examine the impact on wages or new hires separately.

In Table A5, we examine the relations between skewness and other independent variables used in the regression of the number of bachelors. Each of the variables is regressed on skewness. We do not observe any significant relation between these variables with skewness.

Table A12 uses an alternative map to calculate return measures and other control variables, instead of the map in Table A1. This map is constructed based on the SurveyMonkey responses. For each major, we select three industries based on the highest number of respondents who choose an industry that corresponds with our map in Table A1. The conclusion remains unchanged.

1.2 Skewness Measured over a Different Horizon

In our main results, we measure industry skewness in years t-7 to t-3 (t being the graduation year) to reflect the fact that college students have to decide their majors by the sophomore year. In this section, we conduct a timing test by measuring *LaggedSkew* in years t-2 and t-1, i.e., the two calendar years right before graduation, but after most students have declared their major. The results of the test are shown in Table A3, Panel A. *LaggedSkew* in these last two years has a much more muted effect on major choice, reflecting the fact that switching majors is much less common than sticking to a declared major. Furthermore, *LaggedSkew* in the last two years of college is unrelated to entry-level wages upon graduation and to the number of new hires.

In Panel B, our results are also robust to other return windows, t-8 to t-3 and t-6 to t-3.

1.3 Major-specific Time Trends and Total Number of Bachelors in Science and Engineering Majors

Table A6 reports the results of regressions of Log Number of Bachelors on skewness measures (measured in years t-7 to t-3, relative to the graduation year t) and other controls, controlling for major-specific linear, quadratic, and logarithmic time trends. The results are similar to Table 2 Column (1) in the main text.

While Table 2 shows that *LaggedSkew* is positively related to the number of bachelors in related fields, it is possible that salient occurrences of superstar firms lead to an overall increase in the number of students who choose science and engineering majors. In Table A10, we show that this is not the case. Neither the maximum skewness nor the average skewness across our majors predicts the total number of bachelors in science and engineering majors.

1.4 Industry Turnover and Future Firm Performance

One concern with our industry skewness measure is that it may reflect unobserved industry performance dynamics. While we cannot rule this out completely, here we check for any footprints of such an association. First, we examine industry job turnover, obtained from BLS and defined as total separations minus hires, as a percentage of employment. Panel A of Table A7 shows that job turnover is not related to *LaggedSkew*. The result suggests that students who are attracted by *LaggedSkew* do not face different job separation risks.

Next, we examine the relation between our skewness measure and various proxies for industry operating performance in years t-3 (Panel B), t (Panel C), t+3 (Panel D), and t+6 (Panel E) (with t being the year of graduation). Columns (1) and (2) analyze *Return_on_Equity* (*RoE*, defined as earnings divided by book equity) and *Return_on_Assets* (*RoA*, defined as earnings divided by firm assets) as measures of industry performance. Columns (3) and

(4) examine Net_Profit_Margin (NPM, defined as earnings divided by firm sales), and $Sales\_Growth$ (year-on-year changes in firm sales). As we see from the table, $Skew$ does not predict any of these industry performance measures in any specification.

In Table A4, Skewness in either t-1 to t-2 or in t-3 to t-7 is not correlated with future stock return or with future return volatility.

1.5 Selection Effects Based on Student Composition/Quality

In Table A8, we use available information from NSCG on various proxies of socio-economic status (SES) to assess whether students drawn to high-skewness majors tend to come from specific backgrounds, associated with differential future earnings profiles.

Specifically, in this table the dependent variable is our (standardized) Skew measure. The key explanatory variables are categorical variables for the respondent's gender, minority status, father's and mother's education level (in levels from NSCG), interaction between father's and mother's education levels, financial aid received (either tuition waivers, loans, work study of other financial aid), gifts to fund UG education received from parents/relatives (not to be repaid), and the total amount borrowed to finance their UG degree. The levels of parents' education, as reported in the survey, are 2: High school diploma or equivalent; 3: Some college, vocational, or trade school; 4: Bachelors degree; 5: Masters degree, Professional degree, or Doctorate. The Amount borrowed to finance UG degree is also coded in the following levels by NSCG—2: None (baseline); 3: 1 - 10000; 4: 10001-20000; 5: 20001-30000; 6: 30001-40000; 7: 40001-50000; 8: 50001-60000; 9: 60001-70000; 10: 70001-80000; 11: 80001-90000; 12: 90001 or more.

Of these variables, we think of father's and mother's education levels (and their interaction) as being positively related to SES. Of the variables derived from how the respondents financed their undergraduate education, we think of relying on financial aid (either tuition waivers, loans, through work study programs, or other financial aid) as typically indicative of lower affordability (and hence, SES) given the US education system's need-based financial aid policies. Similarly, we also use information on the amount of student loans taken to indicate different levels of financial need, contingent on there being a need. Finally, we think of those who received gifts to fund UG education from parents/relatives (not to be repaid) as coming from relatively more wealthy families.

As we can see from the Table, we do not find any systematic and significant correlations between the tendency to choose high-skewness majors and these characteristics. For example, in column 1, we find that respondents choosing high-skew majors tend to be male, but this result is not significant in the other columns. Moreover, the economic magnitude of these results is rather small—an average male chooses a major with 0.02 standard-deviation higher skew than an average female. Similar interpretations apply for the rest of the table, including minority status. The only coefficient that seems to be significant statistically (although still limited in economic magnitude) is that those whose father has only high-school education but mother has a post-graduate degree tend to choose high-skewness majors (relative to those whose parents have the lowest level of education). However, this is a very unusual combination of parents' education levels in our sample – only about 1% of our respondents have such combination of parents.

Overall, we do not find any evidence of negative selection on SES accompanying the choice of high skewness majors.

In Table A9, we run a regression with (residual) earnings average and dispersion as dependent variables. Residual wages are first calculated from the individual-level data, using fixed effects for Graduation cohort, Gender, Marital status, Minority status, Region, Highest degree attained, Major-survey year, Industry-survey year, and Industry-major. The key independent variable is the interaction between Lagged Skew relevant to major choice, and the cohort's experience, measured as the number of years that have passed since graduation for that cohort. While the effect of skewness on average earnings gets weaker over time, this decay is gradual. We do not find any significant relationship between wage dispersion and skewness; neither is there any evidence of an increase in dispersion over time for high-skew cohorts.

1.6 Regional Skewness

In Table A11, we use the school-level data provided by IPEDS and construct regional skewness measures. Regional Skew is cross-sectional return skewness calculated using firms in the school's region. The region is one of nine census regions: East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific and US Territories, South Atlantic, West North Central, and West South Central. We control for major-year fixed effects and clustered standard errors at the major-region level. The results are qualitatively unchanged: regional skew is associated with more students in related majors, especially among top schools.

2 More Evidence on the Effect of Industry Salience

Our evidence in the main text on the joint dynamics of quantities (the number of graduates/new hires) and prices (the average wage) points to a relatively larger shift in labor supply than in labor demand with the occurrences of superstar firms. In other words, college students are attracted by superstar firms in deciding their majors, not because they rationally anticipate improved job prospects, but because they are drawn to extreme, salient events. In this section, we provide more evidence on the effect of salience on college major choice.

We exploit structural breaks in industry valuation during the NASDAQ bubble in the late 1990s to identify superstar industries. Our logic is similar to that of Charles, Hurst, and Notowidigdo (2018), who argue that sudden, sharp increases in local house prices in the early 2000s are the result of speculative activity and are unlikely to be caused by changes in local economic conditions. In the same way, our underlying assumption is that abrupt, sharp increases in stock valuations during the Tech Bubble were a result of stock market speculation. In other words, we argue that these sharp price appreciations did not merely reflect changes in rational expectations of industry fundamentals that could affect overall labor demand. We follow the same two-stage estimation procedure as in Charles, Hurst, and Notowidigdo (2018). In the first stage, we estimate industry-specific OLS regressions with a single structural break, and search for the time of the structural break that maximizes the R^2 of the following regression:

$$R_{i,t} = \alpha_i + \tau_i t + \lambda_i (t - t_i^*) \mathbb{1}(t > t_i^*) + \epsilon_{i,t}, \tag{1}$$

where $R_{i,t}$ is the cumulative return of industry i up to quarter t, t_i^* is the date of the structural break in the industry's valuation, restricted to be between 1990Q1 to 1999Q4 (the NASDAQ index peaked in Q1 of 2000). τ_i is the linear time-trend in price appreciation before the structure break, and λ_i is the size of the structural break—reflecting the change in the growth rate at the structural break. This procedure follows standard approaches in time-series econometrics to identify a single break point (e.g., Bai 1997; Bai and Perron 1998).

In the second stage, we conduct an event-time study by comparing the number of college graduates from related majors around the time of the structural break. More specifically, we estimate the following regression:

$$log(bachelor_{i,t}) = \alpha + \beta Post_{i,t} \times \lambda_i + \gamma \mathbf{X}_{i,t-3} + \mu_i + \tau_t + \epsilon_{i,t}, \tag{2}$$

where $Post_{i,t}$ is a dummy variable that equals one if year t is 3 years after the structural break t_i^* (so the structural break occurs by the sophomore year of the year t graduates); we further control for industry and time fixed effects on the right-hand-side of the equation. The difference-in-difference coefficient β then measures the difference in the number of graduates from related major fields before vs. after the structural break weighted by the size of the break.

Table A13 presents these regression results. Panel A shows the result of the first stage. There is significant variation across industries both in terms of the timing of the structural break and the magnitude of the break, consistent with the finding in Campello and Graham (2013) that some non-tech industries also experienced a boom during the tech bubble. Not surprisingly, Computer Science-related industries experience the largest structural break among all science-engineering majors in our sample.[1] Interestingly, Earth and Ocean Sciences-related industries experience a negative structural break, possibly because investors view them as "boring" relative to tech-related industries in this period.

Panel B reports the change in the number of college graduates from related majors around the structural break. As can be seen from Column (1), the size of the structural break is significantly associated with subsequent changes in major enrollment; more specifically, a one-standard-deviation increase in the magnitude of the structural break is associated with a 10.0% (= 0.7758 × 0.1283, where 0.1283 is the standard deviation of λ) increase in the number of graduates in related major fields. Columns (2)–(5) examine changes in industry fundamentals around the same break points; we do not see similar structural breaks in any of the commonly used proxies for industry performance.[2] In sum, our results based on structural breaks in industry valuation provide further evidence for a plausibly causal impact of extreme events on college major choice.

[1] As can be seen from Figure 1, computer science and electrical engineering majors also have very high cross-sectional return skewness during that period. The skewness for these industries was 7.09 during 1997 to 2001; for reference, the average pooled skewness of other majors in the same period was 1.85.

[2] Using the full sample period, Online Appendix Table A7 confirms that *LaggedSkew* in years t-7 to t-3 does not significantly predict these proxies for industry performance measured in years t-3, t, t+3, and t+6.

3 Details On Our Own Survey Using SurveyMonkey

In this section, we provide further details on our own survey of College graduates. We used SurveyMonkey to conduct this survey. This involved selecting College graduates employed in the United States. SurveyMonkey sent the survey to 1200 people enlisted on their platform; our target sample size was primarily motivated by research budget constraints.

Below is the information on sample composition that we received from SurveyMonkey:

Age Balancing: Basic

Gender Balancing: Census

Incidence rate: 35-49%

Employment Status: Employed full-time

COUNTRY: United States

AGE: 22-65

Education: 4-year college, Graduate degree

Important Notes: We will send your survey to up to 1200 respondents to get 420 completed responses at your estimated incidence rate of 35-49%. If your actual incidence rate is lower than the incidence rate you selected, more respondents will be disqualified from your survey and you may get fewer completed responses than you ordered for the same total cost.

We screened our respondents on SurveyMonkey using the following criteria: at the time of the survey, the respondent had to be a) a US college graduate with one of the NSF majors that we examine in the paper, b) between 22 and 65 years old, and c) employed full time for at least one full year. In addition, we also screened respondents based on whether they cared/worried about job market outcomes when they chose their majors, leaving out those that chose majors based purely on academic interest.[3] Eventually, we ended up with 394 respondents with complete information on age and gender who graduated before 2017 (so

[3] Specifically, we asked, "How important was the availability of jobs or future income prospects in related industries (where people with this major typically worked) in your major choice decision?" and only kept those who answered "Most Important" or "Somewhat Important" (screening out those that chose "Least Important").

that we have at least one full year of earnings information on them).

We do not think that this sample is representative of the population of US college graduates; after all, these are people willing to fill out surveys for a few dollars. Further details about this sample, based on descriptive statistics calculated from survey responses, are in Panel A of Table A14. The survey instrument is available at `https://personal.lse.ac.uk/loud/ChoiLouMuk_SurveyQuestions.pdf`.

In this survey, we asked specific questions regarding the industry of the respondent's first job after college (as opposed to their current job, as in NSCG), their target industry when they chose their major, the year (high school or freshman/sophomore year of college) when they made their major choice (allowing us to more precisely measure the relevant timing for our skewness variable), reasons for switching if they changed the industry they worked in after college, as well as their beliefs about job opportunities and preferences for skewness when they were in college.

In Panel B of Table A14, we show descriptive statistics for key survey questions. First, when we examine when respondents decided on their major, we find that 72.8% decided sometime between years t-7 and t-3. Further, in order to control for financial need, we asked students directly about student loan status. 70% of our sample had such a loan. Out of this group of students, 18.9% said that they were not worried at all about paying it back, 32% said they were "somewhat worried" about it, with the remaining 19.1% indicating they were "very worried". We control for these differences using fixed effects in our empirical specification.

In order to understand respondents' expectations regarding the industry they wished to join after graduating, we construct a starting industry (*startind*) variable. For respondents with a single self-reported target industry, *startind* equals 1 if respondent joined her target industry at graduation; for respondents with multiple target industries, *startind* is set to equal 1 if respondent joined any of their target industries, *startind* equals 0 otherwise. We also create an expected industry (*exp_ind*) variable to (to use as a fixed effect and for clustering): for respondents with a single target industry, *exp_ind* equals that industry; for respondents with multiple target industries, *exp_ind* is set to ”multiple”.

Next, we find that 40.1% of our respondents did not start working in the industry they targeted in college, and 73.8% of respondents stayed in the same industry where they started working. Finally, a majority of respondents confirms making errors in expectations, but far fewer (29.6%) have lottery preferences when it comes to income streams.

References

Bai, Jushan, 1997, “Estimation of a change point in multiple regression models,” Review of Economics and Statistics 79, 551–563.

Bai, Jushan, and Pierre Perron, 1998, “Estimating and testing linear models with multiple structural changes,” Econometrica 66, 47–78.

Campello, M. and J.R. Graham, 2013, “Do stock prices influence corporate decisions? Evidence from the technology bubble,” Journal of Financial Economics 107, 89–110.

Charles, Kerwin Kofi, Erik Hurst and Matthew J. Notowidigdo, 2018, “Housing Booms and Busts, Labor Market Opportunities, and College Attendance,” American Economic Review, forthcoming.

Table A1

Industries and Majors

This lists the science and engineering majors used in the paper and a map between majors and 3-digit NAICS industry codes (we exclude NAICS codes that start with 92, which correspond to Public Administration and are not covered in economic census).

1	Aeronautical and astronautical engineering
2	Biological sciences*
3	Chemical engineering
4	Civil engineering
5	Computer sciences
6	Earth and ocean sciences
7	Economics
8	Electrical engineering
9	Health*
10	Industrial and manufacturing engineering
11	Materials science
12	Mechanical engineering

*We exclude Biological sciences and Health majors in our aggregate-level analysis, because many biology- and health-related jobs require an advanced degree and students often go to graduate schools before entering the job market. In the individual-level analysis, we include these two majors because these surveys are not based on entry-level employees, but also contain respondents that graduated 20 or more years earlier.

(The NSF data include these other majors as well: Agricultural sciences, Astronomy, Atmospheric sciences, Chemistry, Engineering technology, Mathematics, Physics, Political science, Psychology, Sociology)

3-digit NAICS	Industry	Major(s)
113	Forestry and Logging	Earth and ocean sciences
115	Support Activities for Agriculture and Forestry	-
211	Oil and Gas Extraction	Chemical engineering Earth and ocean sciences
212	Mining (except Oil and Gas)	Chemical engineering Earth and ocean sciences
213	Support Activities for Mining	Chemical engineering Earth and ocean sciences
236	Construction of Buildings	Civil engineering
237	Heavy and Civil Engineering Construction	Civil engineering
238	Specialty Trade Contractors	-
311	Food Manufacturing	-
312	Beverage and Tobacco Product Manufacturing	-
313	Textile Mills	Industrial and manufacturing engineering Materials science Mechanical engineering
314	Textile Product Mills	Industrial and manufacturing engineering Materials science Mechanical engineering
315	Apparel Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering
316	Leather and Allied Product Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering
321	Wood Product Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering
322	Paper Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering

3-digit NAICS	Industry	Major(s)
323	Printing and Related Support Activities	-
324	Petroleum and Coal Products Manufacturing	Chemical engineering Industrial and manufacturing engineering Materials science Mechanical engineering
325	Chemical Manufacturing	Chemical engineering Industrial and manufacturing engineering Materials science Mechanical engineering
326	Plastics and Rubber Products Manufacturing	Chemical engineering Industrial and manufacturing engineering Materials science Mechanical engineering
327	Nonmetallic Mineral Product Manufacturing	Industrial and manufacturing engineering Materials science
331	Primary Metal Manufacturing	Chemical engineering Industrial and manufacturing engineering Materials science Mechanical engineering
332	Fabricated Metal Product Manufacturing	Chemical engineering Industrial and manufacturing engineering Materials science Mechanical engineering
333	Machinery Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering
334	Computer and Electronic Product Manufacturing	Computer sciences Electrical engineering
335	Electrical Equipment, Appliance, and Component Manufacturing	Computer sciences Electrical engineering
336	Transportation Equipment Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering
337	Furniture and Related Product Manufacturing	Industrial and manufacturing engineering Materials science Mechanical engineering
339	Miscellaneous Manufacturing	Industrial and manufacturing engineering Mechanical engineering
423	Merchant Wholesalers, Durable Goods	-
424	Merchant Wholesalers, Nondurable Goods	-
425	Wholesale Electronic Markets and Agents and Brokers	-
441	Motor Vehicle and Parts Dealers	-
442	Furniture and Home Furnishings Stores	-
443	Electronics and Appliance Stores	-
444	Building Material and Garden Equipment and Supplies Dealers	-
445	Food and Beverage Stores	-
446	Health and Personal Care Stores	-
447	Gasoline Stations	-
448	Clothing and Clothing Accessories Stores	-
451	Sporting Goods, Hobby, Book, and Music Stores	-
452	General Merchandise Stores	-
453	Miscellaneous Store Retailers	-
454	Nonstore Retailers	Computer sciences Electrical engineering
481	Air Transportation	Aeronautical and astronautical engineering
482	Rail Transportation	-
483	Water Transportation	-
484	Truck Transportation	-
485	Transit and Ground Passenger Transportation	-
486	Pipeline Transportation	-
488	Support Activities for Transportation	-
491	Postal Service	-
492	Couriers and Messengers	-

3-digit NAICS	Industry	Major(s)
493	Warehousing and Storage	-
511	Publishing Industries (except Internet)	Computer sciences Electrical engineering
512	Motion Picture and Sound Recording Industries	Computer sciences Electrical engineering
515	Broadcasting (except Internet)	-
516	Internet Publishing and Broadcasting	Computer sciences Electrical engineering
517	Telecommunications	Computer sciences Electrical engineering
518	Internet Service Providers, Web Search Portals, and Data Processing Service	Computer sciences Electrical engineering
519	Other Information Services	Computer sciences Electrical engineering
521	Monetary Authorities - Central Bank	Economics
522	Credit Intermediation and Related Activities	Economics
523	Securities, Commodity Contracts, and Other Financial Investments and Related Activities	Economics
524	Insurance Carriers and Related Activities	Economics
525	Funds, Trusts, and Other Financial Vehicles	Economics
531	Real Estate	Economics
532	Rental and Leasing Services	Economics
533	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	Economics
541	Professional, Scientific, and Technical Services	Computer sciences Electrical engineering
551	Management of Companies and Enterprises	-
561	Administrative and Support Services	-
562	Waste Management and Remediation Services	-
611	Educational Services	-
621	Ambulatory Health Care Services	Biological sciences Health
622	Hospitals	Biological sciences Health
623	Nursing and Residential Care Facilities	Biological sciences Health
624	Social Assistance	Biological sciences Health
711	Performing Arts, Spectator Sports, and Related Industries	-
712	Museums, Historical Sites, and Similar Institutions	-
713	Amusement, Gambling, and Recreation Industries	-
721	Accommodation	-
722	Food Services and Drinking Places	-
811	Repair and Maintenance	-
812	Personal and Laundry Services	-
813	Religious, Grantmaking, Civic, Professional, and Similar Organizations	-
814	Private Households	-

Table A2
Robustness Tests

This table repeats the main tests with various robustness tests and reports the coefficient of Skew or proxies that replace Skew. The dependent variables are Log Number of Bachelors, Log Annual Wage, and Net New Hires. The robustness tests are: including Biological sciences and Health majors; calculating standard errors using bootstrap with 10,000 times and block-bootstrap with 10,000 times; aggregating industries within a major to calculate skewness and other variables, by equal-weighting the industries or picking the industry measures with the highest absolute value; defining Top_N ($Bottom_N$) as the average return of the top (bottom) 5, 3, or 1 firms among all firms that are mapped to the major minus the median return, divided by the difference between the 90th and 10th return percentile, after dropping firms in the lowest 50th size percentile; excluding bachelors who graduate from 1990 to 2002; and using the number of masters instead of bachelors.

Robustness Checks	Log Number of Bachelors	Log Annual Wage	Net New Hires
	(1)	(2)	(3)
Including Biological Sciences and Health Majors	0.0934***	-0.0102***	0.0288
	(0.0282)	(0.0024)	(0.0192)
Bootstrapped Standard Errors	0.1136***	-0.0165***	0.0426
	(0.0293)	(0.0033)	(0.0365)
Block-Bootstrapped Standard Errors	0.1136***	-0.0165***	0.0426
	(0.0278)	(0.0033)	(0.0322)
Equal-Weight Industries Within the Same Major	0.1574***	-0.0144***	0.0089
	(0.0322)	(0.0029)	(0.0292)
Pick the Industry with the Highest Absolute Value	0.1682***	-0.0086**	0.0052
	(0.0381)	(0.0035)	(0.0275)
Defining Tail Using Top and Bottom 5 Firms	0.1070	-0.0204**	0.0383
	(0.0553)	(0.0088)	(0.0536)
Defining Tail Using Top and Bottom 3 Firms	0.0728	-0.0139	0.0557
	(0.0548)	(0.0081)	(0.0602)
Defining Tail Using Top and Bottom 1 Firm	0.0369	-0.0026	0.0557
	(0.0448)	(0.0057)	(0.0566)
Excluding Bachelors Who Graduate Between 1990 and 2002	0.1499***	-0.0107***	0.0339
	(0.0364)	(0.0033)	(0.0512)
Using Masters Instead of Bachelors	0.0557***		
	(0.0200)		

Table A3
Skewness Measures of Different Horizons

This table reruns regressions of Log Number of Bachelors, Log Annual Wage, and Net New Hires, all in year t. In Panel A, the return measures (skewness, mean, and standard deviation of return) are measured in years t-2 to t-1 or years t-7 to t-3. In Panel B, the return measures (skewness, mean, and standard deviation of return) are measured in years t-8 to t-3 or years t-6 to t-3. All other variables are the same as Table 2 in the main text. Standard errors are clustered at the year level. All independent variables are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

Panel A: Adding t-2 to t-1 Measures						
	Log Number of Bachelors		Log Annual Wage		Net New Hires	
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Skew (t-2 to t-1)	0.0579**	0.0578**	0.0001	-0.0021	-0.0102	0.0034
	(0.0242)	(0.0219)	(0.0037)	(0.0033)	(0.0648)	(0.0642)
Lagged Skew (t-7 to t-3)		0.1098***		-0.0172***		0.0433
		(0.0276)		(0.0029)		(0.0299)
Lagged Mean Return (t-2 to t-1)	-0.0473	-0.0377	-0.0074	-0.0025	-0.0206	-0.0191
	(0.0334)	(0.0292)	(0.0065)	(0.0049)	(0.0414)	(0.0484)
Lagged Mean Return (t-7 to t-3)		0.1141***		0.0052		0.0266
		(0.0226)		(0.0042)		(0.0346)
Lagged Standard Dev of Return (t-2 to t-1)	-0.0098	-0.0032	0.0013	-0.0051	-0.0188	-0.0174
	(0.0311)	(0.0294)	(0.0075)	(0.0065)	(0.0872)	(0.0930)
Lagged Standard Dev of Return (t-7 to t-3)		-0.0661		0.0005		0.0232
		(0.0327)		(0.0056)		(0.0890)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
# Observations	513	513	210	210	200	200
Adj. R-Squared	0.84	0.85	0.96	0.96	0.28	0.28

Panel B: Other Horizons						
	Log Number of Bachelors		Log Annual Wage		Net New Hires	
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Skew (t-8 to t-3)	0.1285***		-0.0185***		0.0590	
	(0.0320)		(0.0027)		(0.0411)	
Lagged Skew (t-6 to t-3)		0.1042***		-0.0138***		0.0601
		(0.279)		(0.0031)		(0.0466)
Lagged Mean Return and Standard Deviation of Return	t-8 to t-3	t-6 to t-3	t-8 to t-3	t-6 to t-3	t-8 to t-3	t-6 to t-3
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
# Observations	503	503	210	210	200	200
Adj. R-Squared	0.85	0.85	0.97	0.96	0.30	0.30

Table A4

Mean Return, Standard Deviation of Return, and Return Skewness

This table reruns regressions of Mean Return and Standard Deviation of Return, both in year t. The independent variables include return measures (skewness, mean, and standard deviation of return) measured in years t-2 to t-1 or years t-7 to t-3. All other variables are the same as Table 2 in the main text. Standard errors are clustered at the year level. All independent variables are standardized with zero mean and unit standard deviation. ** p < .05; *** p < .01.

	Mean Return at t			Standard Deviation of Return at t		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Skew (t-2 to t-1)	-0.0017		-0.0003	0.0172		0.0154
	(0.0075)		(0.0080)	0.0150		(0.0155)
Lagged Skew (t-7 to t-3)		-0.0039	0.0031		-0.0103	-0.0059
		(0.0112)	(0.0078)		(0.0090)	(0.0080)
Lagged Mean Return (t-2 to t-1)	0.1754***		0.1767***	0.0452***		0.0347**
	(0.0184)		(0.0188)	(0.0123)		(0.0145)
Lagged Mean Return (t-7 to t-3)		-0.0179	-0.0093		-0.0154	-0.0172
		(0.0208)	(0.0169)		(0.0213)	(0.0227)
Lagged Standard Dev of Return (t-2 to t-1)	-0.0264		-0.0278	0.0058		0.0205
	(0.0274)		(0.0269)	0.0119		(0.0201)
Lagged Standard Dev of Return (t-7 to t-3)		-0.0407	0.0130		-0.0621	-0.0544
		(0.0308)	(0.0256)		(0.0305)	(0.0371)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
# Observations	503	503	503	503	503	503
Adj. R-Squared	0.81	0.66	0.81	0.52	0.52	0.54

Table A5

Relations Between Skewness and Other Variables

This table examines the relations between skewness and other independent variables in Column (1) in Table 2. In each column, one variable is selected as the dependent variable and the others are independent variables. All variables are the same as Table 2 in the main text. Standard errors are clustered at the year level. All independent variables are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

	Lagged Mean Return	Lagged Standard Deviation of Return	Lagged Log Average Wage	Lagged Log Average Market Cap	Lagged Log Average Book-to-Market	Lagged Log Average Firm Age
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Skew	-0.0146	0.0092	-0.0479	0.0207	0.0244	-0.0462
	(0.0489)	(0.0242)	(0.0488)	(0.0240)	(0.0226)	(0.0337)
Lagged Mean Return		0.1738***	0.1290	0.0511	-0.3523***	-0.0603
		(0.0572)	(0.0743)	(0.0269)	(0.0436)	(0.0659)
Lagged Standard Deviation of Return	0.3588***		0.2089**	-0.2390***	0.0769	0.3702***
	(0.1045)		(0.0802)	(0.0294)	(0.0769)	(0.0946)
Lagged Log Average Wage	0.0746	0.0585**		0.0014	0.0992**	-0.0968
	(0.0499)	(0.0274)		(0.0150)	(0.0416)	(0.0526)
Lagged Log Average Market Cap	0.3124	-0.7080***	0.0144		-0.1791	0.6916***
	(0.1690)	(0.1263)	(0.1588)		(0.1439)	(0.0889)
Lagged Log Average Book-to-Market	-0.6088***	0.0644	0.2965**	-0.0506		0.0649
	(0.0846)	(0.0653)	(0.0880)	(0.0401)		(0.0769)
Lagged Log Average Firm Age	-0.1165	0.3464***	-0.3234**	0.2185***	0.0725	
	(0.1297)	(0.0628)	(0.1348)	(0.0350)	(0.0841)	
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
# Observations	513	513	513	513	513	513
Adj. R-Squared	0.69	0.85	0.47	0.95	0.82	0.80

Table A6
Number of Bachelors Regressions Controlling for Major-Specific Time Trends

This table reports the results of regressions of Log Number of Bachelors on skewness measures (measured in years t-7 to t-3, relative to the graduation year t) and other controls, controlling for major-specific linear, quadratic, and logarithmic time trends. Log Number of Bachelors is the log annual number of bachelor degrees awarded for a major. Skew is the employment-weighted cross-sectional skewness of annual returns in all firms that are mapped to the major. Other controls are the same as Table 2 Column (1) in the main text. Standard errors are clustered at the year level. All independent variables are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

	Log Number of Bachelors	Log Number of Bachelors	Log Number of Bachelors
	(1)	(2)	(3)
Lagged Skew	0.0806***	0.0784***	0.0825***
	(0.0247)	(0.0243)	(0.0249)
Lagged Mean Return	0.1609***	0.1615***	0.1603***
	(0.0230)	(0.0229)	(0.0231)
Lagged Standard Deviation of Return	0.0180	0.0186	0.0176
	(0.0243)	(0.0242)	(0.0245)
Major-Specific Time Trends	Linear	Quadratic	Logarithmic
Other Controls	Yes	Yes	Yes
# Observations	513	513	513
Adj. R-Squared	0.84	0.84	0.84

Table A7
Industry Turnover and Operating Performance Measures

This table reports the results of regressions of Industry Turnover (in year t) and operating performance measures (in year t-3, t, t+3, or t+6) on skewness measures (measured in years t-7 to t-3) and other controls. Industry Turnover is the total separations minus total hires (both as % of total employment). RoA is the return on assets, defined as earnings divided by total assets. NPM is the net profit margin, that is, earnings divided by sales. Sales growth is the percentage growth in sales. Skew is the employment-weighted cross-sectional skewness of annual returns in an industry.

Panel A: Turnover	
Lagged Skew	-0.0004
	(0.0010)
Lagged Mean Return and Standard Deviation	Yes
Other Controls	Yes
Year and Major Fixed Effects	Yes
# Observations	153
Adj. R-Squared	0.82

Panel B: Operating Performance Measures 3 Years Before Graduation				
	RoE	RoA	NPM	Sales Growth
	(1)	(2)	(3)	(4)
Lagged Skew	-0.0001	-0.0027	0.0007	-0.0059
	(0.0064)	(0.0026)	(0.0036)	(0.0073)
Lagged Mean Return and Standard Deviation	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
# Observations	481	481	481	481
Adj. R-Squared	0.26	0.52	0.46	0.27

Panel C: Operating Performance Measures Upon Graduation				
	RoE	RoA	NPM	Sales Growth
	(1)	(2)	(3)	(4)
Lagged Skew	-0.0059	-0.0026	0.0023	-0.0059
	(0.0070)	(0.0021)	(0.0024)	(0.0053)
Lagged Mean Return and Standard Deviation	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
# Observations	502	502	502	502
Adj. R-Squared	0.27	0.43	0.39	0.38

Panel D: Operating Performance Measures 3 Years After Graduation				
	RoE	RoA	NPM	Sales Growth
	(1)	(2)	(3)	(4)
Lagged Skew	-0.0006	-0.0009	0.0026	0.0035
	(0.0100)	(0.0021)	(0.0032)	(0.0070)
Lagged Mean Return and Standard Deviation	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
# Observations	474	474	474	474
Adj. R-Squared	0.18	0.35	0.32	0.34

Panel E: Operating Performance Measures 6 Years After Graduation				
	RoE	RoA	NPM	Sales Growth
	(1)	(2)	(3)	(4)
Lagged Skew	0.0084	0.0021	0.0030	0.0019
	(0.0161)	(0.0035)	(0.0037)	(0.0091)
Lagged Mean Return and Standard Deviation	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
# Observations	447	447	447	447
Adj. R-Squared	0.13	0.32	0.30	0.32

Table A8

Socio-Economic Status Imputed from NSCG

In this table the dependent variable is Skew, the return skewness associated with the chosen major of each respondent. The key explanatory variables are categorical variables for gender, minority status, father s and mother s education level (in levels from NSCG), interaction between father s and mother s education levels, financial aid received (tuition waivers, loans, or work study of other financial aid), gifts to fund UG education recived from parents/relatives (not to be repaid), and amount borrowed to finance UG degree. The levels of parents education are 2--High school diploma or equivalent; 3--Some college, vocational, or trade school; 4--Bachelors degree; 5--Masters degree, Professional degree, or Doctorate. The amount borrowed to finance UG degree is also coded in the following levels by NSCG: 2--None (baseline); 3--1 - 10000; 4--10001-20000; 5--20001-30000; 6--30001-40000; 7--40001-50000; 8--50001-60000; 9--60001-70000; 10--70001-80000; 11--80001-90000; 12--90001 or more. Each regression controls for fixed effects for age, graduation cohort, major and and survey cohort. Standard errors are clustered by graduation cohort. ** p < .05; *** p < .01.

Who Chooses High-Skew Majors?					
		(1)	(2)	(3)	(4)
Male		0.021	0.02	0.02	0.019
		(0.011)	(0.016)	(0.015)	(0.015)
Minority		0.018	0.019	0.019	0.019
		(0.011)	(0.012)	(0.012)	(0.012)
Father s education level					
	2	-0.0001	-0.012	-0.012	-0.012
		(0.023)	(0.032)	(0.032)	(0.031)
	3	0.01	0.01	0.011	0.011
		(0.04)	(0.056)	(0.056)	(0.055)
	4	0.013	0.026	0.026	0.027
		(0.053)	(0.072)	(0.072)	(0.071)
	5	0.092	0.142	0.142	0.144
		(0.087)	(0.112)	(0.112)	(0.111)
Mother s education level					
	2	-0.031	-0.051	-0.051	-0.05
		(0.023)	(0.031)	(0.031)	(0.031)
	3	-0.019	-0.015	-0.015	-0.014
		(0.032)	(0.042)	(0.042)	(0.042)
	4	-0.039	-0.026	-0.026	-0.024
		(0.047)	(0.058)	(0.058)	(0.058)
	5	-0.07	-0.099	-0.099	-0.099
		(0.053)	(0.061)	(0.061)	(0.06)
Father s X Mother s education level					
	2 2	0.017	0.042	0.042	0.042
		(0.022)	(0.03)	(0.031)	(0.031)
	2 3	0.002	0.005	0.005	0.004
		(0.034)	(0.043)	(0.043)	(0.043)
	2 4	0.001	-0.013	-0.013	-0.015
		(0.054)	(0.063)	(0.063)	(0.062)
	2 5	0.104	0.168**	0.168**	0.168**
		(0.062)	(0.079)	(0.08)	(0.079)
	3 2	0.015	0.015	0.015	0.014
		(0.042)	(0.061)	(0.061)	(0.061)

	3 3	-0.017	-0.037	-0.037	-0.037
		(0.054)	(0.074)	(0.074)	(0.073)
	3 4	0.006	-0.019	-0.019	-0.022
		(0.065)	(0.084)	(0.084)	(0.084)
	3 5	0.024	0.038	0.038	0.038
		(0.081)	(0.089)	(0.089)	(0.087)
	4 2	0.026	0.029	0.029	0.027
		(0.052)	(0.074)	(0.074)	(0.074)
	4 3	0.001	-0.025	-0.025	-0.027
		(0.061)	(0.083)	(0.083)	(0.083)
	4 4	0.039	0.004	0.004	0.002
		(0.063)	(0.082)	(0.082)	(0.081)
	4 5	0.029	0.05	0.05	0.05
		(0.077)	(0.093)	(0.094)	(0.093)
	5 2	-0.007	-0.02	-0.02	-0.02
		(0.091)	(0.119)	(0.119)	(0.118)
	5 3	-0.072	-0.142	-0.142	-0.141
		(0.096)	(0.125)	(0.124)	(0.123)
	5 4	-0.069	-0.14	-0.14	-0.141
		(0.098)	(0.123)	(0.124)	(0.122)
	5 5	-0.033	-0.069	-0.068	-0.067
		(0.099)	(0.118)	(0.117)	(0.116)
Financial Aid			0.012	0.012	0.001
			(0.009)	(0.008)	(0.006)
Gifts to fund College Education				-0.001	0.0002
				(0.01)	(0.01)
UG student loan level					
	3				0.024
					(0.016)
	4				0.021
					(0.018)
	5				0.024
					(0.023)
	6				0.013
					(0.02)
	7				0.012
					(0.025)
	8				0.003
					(0.037)
	9				0.051
					(0.035)
	10				-0.038
					(0.039)
	11				0.033
					(0.041)
	12				0.034
					(0.043)
Effective # Observations		169993	95452	95452	95452
R-Squared		0.65	0.66	0.66	0.66

Table A9

Major-Cohort Level Average Wage and Its Dispersion

This table reports results from a regression with major-survey_year-graduation_year level (residual) earnings average and dispersion as dependent variables. Residual wages are first calculated from the individual-level data, using fixed effects for Graduation cohort, Gender, Marital status, Minority status, Region, Highest degree attained, Major-survey year, Industry-survey year, and Industry-major. The key independent variable is the interaction between Lagged_skew relevant to major choice, and the cohort s experience, measured as the number of years that have passed since graduation for that cohort in Columns (1) and (2), and (logarithm of the) number of years in Columns (3) and (4). In all columns, the following fixed effects are included: Major-survey_year, Major-years out of college, and Years out of college-Survey_year. Standard errors are clustered by major-survey_year. ** $p < .05$; *** $p < .01$.

	Average earnings	Earnings Dispersion	Average earnings	Earnings Dispersion
	(1)	(2)	(3)	(4)
Lagged Skew	-0.0531***	-0.0026	-0.077**	0.0068
	(0.0176)	(0.0155)	(0.035)	(0.0275)
Lagged Skew * Years out of college	0.0026**	0.0001		
	(0.0012)	(0.001)		
Lagged Skew * log (Years out of college)			0.0266	-0.0029
			(0.0147)	(0.0117)
Lagged Mean Return and Standard Deviation of Return	Yes	Yes	Yes	Yes
Effective # Observations	2597	2554	2593	2553
Adj. R-Squared	0.14	0.31	0.14	0.31

Table A10
Total Number of Bachelors

This table reports the results of regressions of Log Change in Total Number of Bachelors on skewness measures (measured in years t-7 to t-3, relative to the graduation year t). Log Total Number of Bachelors is the log annual total number of bachelor degrees across all the 10 majors in Table A1. Skew is the employment-weighted cross-sectional skewness of annual returns in all firms that are mapped to the major. Our control variables are all measured at year t-3 and include Log Average Wage, the 3-year average wage obtained from Compustat (up to 1998) or from BLS (1999 and onward); Mean Return and Standard Deviation of Return, both are employment-weighted. Other controls are Log Average Market Cap, Log Average Book-to-Market, and Log Average Firm Age, weighted by employment. All measures are aggregated using two methods: averaging across all majors (weighted by the number of bachelors in year t-3) in Column (1) and picking the measure with the maximum absolute value in Column (2). Standard errors are clustered at the year level. All independent variables are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

Log Total Number of Bachelors		
Aggregating Method	Bachelor-weighted Average	Max Absolute Value
	(1)	(2)
Lagged Skew	0.0007	-0.0083
	(0.0498)	(0.0593)
Lagged Mean Return	0.2586***	0.1436***
	(0.0543)	(0.0515)
Lagged Standard Deviation of Return	0.0275	-0.0144
	(0.0582)	(0.0585)
Other Controls	Yes	Yes
Time Trend	Linear and Quadratic	Linear and Quadratic
# Observations	49	52
Adj. R-Squared	0.94	0.90

Table A11
School-level Regressions with Regional Skewness

The dependent variable is School-level Log Number of Bachelors, the log annual number of bachelor degrees awarded for a major at the school-level. There are 336 schools in total. These are 4-year universities in the US, offering at least 5 out of our 10 majors. Regional Skew is Skew calculcated using firms in the school's region. The region is one of nine census regions: East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific and US Territories, South Atlantic, West North Central, and West South Central. Skew is the employment-weighted cross-sectional skewness of annual returns in all firms that are mapped to the major. Top School is a dummy variable, indicating that the school is in the top 50 in US News Rankings 4 years prior to graduation. Other controls are the same as those in the corresponding regressions in Table 4 in the main text. Standard errors are clustered at the major-region level. All independent variables are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

	School-level Log Number of Bachelors	
	(1)	(2)
Lagged Regional Skew	0.0858***	0.0569**
	(0.0216)	(0.0266)
Lagged Regional Skew * Top School		0.1355***
		(0.0477)
Lagged Mean Return and Standard Deviation of Return	Yes	Yes
Other Controls	Yes	Yes
Fixed Effects	Major-year	Major-year
# Observations	31792	31792
Adj. R-Squared	0.22	0.22

Table A12
Using an Alternative Major-Industry Map Based on *SurveyMonkey*

This table reruns regressions of Log Number of Bachelors, Log Annual Wage, and Net New Hires. When mapping firms into majors, we use an alternative map based on the SurveyMonkey responses instead of Table A1. For each major, we select three industries based on the highest number of respondents who choose an industry that corresponds with our map in Table A1. All variables are the same as Table 2 in the main text. Standard errors are clustered at the year level. All independent variables are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

	Log Number of Bachelors	Log Annual Wage	Net New Hires
	(1)	(2)	(3)
Lagged Skew	0.0625***	-0.0076***	0.0381
	(0.0204)	(0.0025)	(0.0389)
Lagged Mean Return	0.1086***	0.0023	0.0143
	(0.0353)	(0.0028)	(0.0290)
Lagged Standard Deviation of Return	-0.1035	-0.0177***	0.0453
	(0.0546)	(0.0061)	(0.0707)
Lagged Log Average Wage	0.0850**	0.0062***	-0.0265
	(0.0328)	(0.0013)	(0.0286)
Lagged Log Number of Bachelors	No	Yes	Yes
Lagged Male/Female Ratio	No	Yes	Yes
Other Controls	Yes	Yes	Yes
Year and Major Fixed Effects	Yes	Yes	Yes
# Observations	448	186	178
Adj. R-Squared	0.84	0.98	0.16

Table A13
Structural Breaks in Industry Valuation

This table uses the NASDAQ bubble period in the 1990s to identify structural breaks in industry valuation. In Panel A, time series regressions are run for every major-related industry using the cumulative quarterly industry return from 1990 to 1999. Time Trend is the base time trend of the period, and Lambda is the change in time trend after the structural break. The structural break is identified by the time series regression that has the maximum adjusted R^2. The t-stats of the Lambda estimates are also reported. In Panel B, Post is a dummy variable indicating the time is 3 years after the structural break of the major. The dependent variables are Log Number of Bachelors, RoE, RoA, NPM, and Sales Growth. Log Number of Bachelors is the log annual number of bachelor degrees awarded for a major. RoE is the return on equity, defined as earnings divided by equity. RoA is the return on assets, defined as earnings divided by total assets. NPM is the net profit margin, that is, earnings divided by sales. Sales growth is the percentage growth in sales. The sample period for this panel is from 1990 to 2002. Other controls are all measured at year t-3 and include Log Average Wage, the 3-year average wage obtained from Compustat (up to 1998) or from BLS (1999 and onward); Log Average Market Cap, Log Average Book-to-Market, and Log Average Firm Age, weighted by employment. Standard errors are clustered at the year level. All independent variables (except Post and Lambda) are standardized with zero mean and unit standard deviation. ** $p < .05$; *** $p < .01$.

Panel A: Identifying Structural Break

Major	Max Adj. R^2	Time Trend	Lambda	t-stat	Break YearQtr
Aeronautical and astronautical eng.	84.32%	0.0181	0.0475	(3.39)	199404
Chemical engineering	95.22%	0.0301	0.0449	(6.77)	199402
Civil engineering	82.53%	0.0197	0.0242	(2.44)	199501
Computer sciences	97.62%	0.0740	0.3235	(8.95)	199701
Earth and ocean sciences	11.93%	0.0049	-0.0283	(-2.35)	199703
Economics	97.13%	0.0615	0.1854	(10.88)	199502
Electrical engineering	97.62%	0.0740	0.3235	(8.95)	199701
Industrial and manufacturing eng.	92.97%	0.0336	0.0322	(4.04)	199404
Materials science	93.42%	0.0324	0.0376	(4.80)	199404
Mechanical engineering	92.57%	0.0330	0.0319	(3.90)	199404

Panel B: Regressions on Structural Break

	Log Number of Bachelors	RoE	RoA	NPM	Sales Growth
	(1)	(2)	(3)	(4)	(5)
Post x Lambda	0.7758**	-0.0923	-0.0712	-0.1086	-0.0918
	(0.2875)	(0.0897)	(0.0451)	(0.0982)	(0.0928)
Lagged Log Average Wage	0.0564	-0.0177	-0.0047	-0.0065	-0.0168**
	(0.0267)	(0.0083)	(0.0046)	(0.0088)	(0.0060)
Lagged Log Average Market Cap	-0.1117	-0.0799	-0.0450	-0.0639	-0.0207
	(0.0844)	(0.0492)	(0.0219)	(0.0450)	(0.0627)
Lagged Log Average Book-to-Market	-0.0104	0.0065	-0.0017	-0.0176	0.0188
	(0.0638)	(0.0205)	(0.0100)	(0.0250)	(0.0228)
Lagged Log Average Firm Age	-0.0722	0.0730	0.0374	0.0748	0.0152
	(0.0949)	(0.0411)	(0.0207)	(0.0405)	(0.0318)
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes
# Observations	130	130	130	130	130
Adj. R-Squared	0.97	0.29	0.28	0.12	0.12

Table A14
Descriptive Statistics from SurveyMonkey Sample

This table shows descriptive statistics for the SurveyMonkey sample. Panel A describes variables given to us by SurveyMonkey, while Panel B shows statistics for key questions we asked in the survey.

Panel A

Age	Freq.	Percent	Cum.
18-29	102	25.89	25.89
30-44	148	37.56	63.45
45-60	110	27.92	91.37
> 60	34	8.63	100

Gender	Freq.	Percent	Cum.
Female	200	50.76	50.76
Male	194	49.24	100

Household Income	Freq.	Percent	Cum.
$0-$9,999	4	1.02	1.02
$10,000-$24,999	12	3.05	4.06
$100,000-$124,999	54	13.71	17.77
$125,000-$149,999	33	8.38	26.14
$150,000-$174,999	22	5.58	31.73
$175,000-$199,999	14	3.55	35.28
$200,000+	32	8.12	43.4
$25,000-$49,999	46	11.68	55.08
$50,000-$74,999	82	20.81	75.89
$75,000-$99,999	76	19.29	95.18
Prefer not to answer	19	4.82	100

Region	Freq.	Percent	Cum.
East North Central	53	13.59	13.59
East South Central	15	3.85	17.44
Middle Atlantic	62	15.9	33.33
Mountain	21	5.38	38.72
New England	17	4.36	43.08
Pacific	81	20.77	63.85
South Atlantic	75	19.23	83.08
West North Central	28	7.18	90.26
West South Central	38	9.74	100

Panel B

When did you decide on your major?	Freq.	Percent	Cum.
After sophomore year	80	20.15	20.15
Freshman Year	70	17.63	37.78
had decided earlier in grades 8-10 in school	26	6.55	44.33
I had decided even earlier	28	7.05	51.39
had decided in the last 2 years of school	94	23.68	75.06
Sophomore Year	99	24.94	100

How worried were you about paying back your student loan..	Freq.	Percent	Cum.
I did not have/was not going to have loans	119	29.97	29.97
not worried at all	75	18.89	48.87
somewhat worried	127	31.99	80.86
very worried	76	19.14	100

Starting job in target industry	Freq.	Percent	Cum.
0	159	40.05	40.05
1	238	59.95	100

Works in the same industry as first job	Freq.	Percent	Cum.
0	104	26.2	26.2
1	293	73.8	100

Incomplete information collection: Gathered information only about a small number of	Freq.	Percent	Cum.
0	219	55.16	55.16
1	178	44.84	100

Recalling expectation errors: Wrong about others	Freq.	Percent	Cum.
0	173	43.58	43.58
1	224	56.42	100

Could major choice have been better with more research?	Freq.	Percent	Cum.
0	139	35.1	35.1
1	257	64.9	100

Would you choose a lottery-like payoff over a stable, average income?	Freq.	Percent	Cum.
0	278	70.38	70.38
1	117	29.62	100