Who Becomes an Inventor?

The (Important) Role of Knowledge Absorption from Higher Technical Education[*]

Takuya Hiraiwa,[†] Shotaro Yamaguchi,[‡] Serguey Braguinsky,[§] Chad Syverson[**]

Abstract

We examine how individual differences in the absorption of higher technical knowledge at university translate into subsequently becoming inventors. Using detailed individual-level data on the first 40 cohorts of graduates from the Tokyo Imperial University College of Engineering at the turn of the last century, we find that academic achievement relative to peers strongly predicts future invention. Drawing on detailed information about the engineering curriculum and academic rankings at different stages of university education, we show that while classroom learning in the first two years was important, its effect on future patenting becomes significantly larger if integrated with practical engineering training received in the final year of the program. These effects remain strong even after controlling for post-graduation employer. Counterfactual estimations using parameters from the quantitative "human capital production function" model indicate that the probability of becoming an inventor would be reduced by almost 90 percent if all graduates had the same level of academic achievement as the lowest-ranked graduates. Also, while superior academic achievement at earlier stages is generally important as it positively affects subsequent performance, "turning off" superior practical knowledge absorption in the final year has a distinct negative impact on invention outcomes.

March 2026

[*] Acknowledgement: This research was partly supported by the Japan Association for the Promotion of Science Grant-in-Aid for Scientific Research (A) No. 21H04398 (Principal Investigator, Serguey Braguinsky). None of the authors has any financial relationships or potential conflicts of interest related to this research.

[†] Hiraiwa, Robert H. Smith School of Business, University of Maryland, email: thiraiwa@umd.edu
[‡] Yamaguchi, Wisconsin School of Business, University of Wisconsin-Madison, email: shotaro.yamaguchi@wisc.edu
[§] Braguinsky, Robert H. Smith School of Business, University of Maryland, email: sbrag@umd.edu
[**] Syverson, University of Chicago, Booth School of Business

I. Introduction

Interest in human capital as it relates to innovation has been on the rise. As Van Reenen (2022) writes, "*Innovation, after all, begins with people*." Higher technical education serves as a cornerstone in cultivating such human capital, especially important in an environment characterized by the need to adapt to technological change (Nelson and Phelps, 1966). Empirically, the impact of science and engineering education on innovation has mostly been examined at the institutional level and/or using variation in geographical location (Jaffe, 1989; Toivanen and Väänänen, 2016; Hvide and Jones, 2018). Several recent studies incorporating individual-level data (Akcigit et al., 2017; Bell et al., 2019; Bianchi and Giorcelli, 2020) highlight the role of highly educated individuals—particularly college graduates—in generating inventions. What remains largely a black box in all extant studies we are aware of, however, is how the propensity to become an inventor is affected by what the student absorbs while in college.

In this paper, we open this black box by making use of unique individual-level data on the census of bachelor's degree recipients from the Tokyo Imperial University College of Engineering collected from archival sources in Japan. Everybody in our data is a university engineering graduate, and we control in the analyses for different divisions (fields of study) and graduation cohorts. Thus, our focus is not on how differences in years of schooling or education fields affect the probability of becoming an inventor, but on how such probability is affected by the differences in the *absorption* of higher engineering education among individuals, all of whom had been exposed to the same kind of education within the same education field and cohort.

We measure the degree of absorption of the knowledge taught in the university by utilizing individual-level data on academic rankings relative to peers based on their class achievements. These rankings are available both at the time of graduation and at the end of each of the previous two years

of the three-year study program.[1] The during-program rankings reflect relative academic performance within a given academic year based on average test scores in that year. In contrast, the graduation ranking is determined by the cumulative performance throughout the three-year program plus the score for the graduation thesis work as the fourth element. These rankings were published in the Imperial University calendars for the first 40 graduating cohorts between 1879-1918. The universe of the College of Engineering graduates from these 40 cohorts comprises our main sample. Furthermore, curricula published in the same calendars, together with year-specific academic rankings, provide us with information on the relative importance of different types of skills acquired during each year of university education.[2]

We match the above data with the graduates' subsequent domestic and international patents from the Japan Patent Office (JPO) and Google patents. We also utilize our own hand-constructed database of individuals' post-graduation placements to examine how job sorting potentially mediated the effects of knowledge absorption in the university on the likelihood of becoming an inventor. We provide more details about our data and matching process in the data section below and in Appendix D.

Our analysis reveals students' absorption of the knowledge taught to them at all stages of university engineering education (first- and second-year rankings as well as university graduation rankings) was economically and statistically significantly associated with a higher probability of becoming an inventor within 20 years of graduation.[3] This association is particularly strongly pronounced at the right tail of

[1] Japanese higher technical education was composed of two tiers at the time. First, students took three years of general academic education in (old-system) high schools that were designated as preparatory schools for the Imperial University. Second, they studied for three more years at the Imperial University itself to acquire specialized professional knowledge. As described below, we also have students' high school graduation rankings relative to peers in the same school and cohort. We use these to control for students' general academic abilities in our analyses.

[2] While rankings at graduation are available for all 40 cohorts, first- and second-year rankings are only available since the College of Engineering became part of the Imperial University in 1886. Thus, when our analyses include also first- and second-year rankings, the sample comprises the cohorts graduated between 1888-1918.

[3] Our career and patent data extend through 1939 but after 1937, the Japanese economy had been transformed into a centralized government-administered system under the war economy regime, so we opt not to utilize the data past 1937. This leaves 20 years of usable observations for the youngest, 1918 graduating cohort in our sample and motivates us to look at outcomes over the first 20 years after graduation for all cohorts. Most of the findings presented below, however, are not sensitive to using all available observations on earlier cohorts, beyond the first 20 years.

the distribution. Furthermore, later rankings appear more predictive than earlier rankings. According to the curriculum, students completed basic courses in the first year, more specialized/applied courses in the second year, while dedicating most of the third year to practical engineering training and writing the graduation theses. Thus, while the absorption of all types of knowledge was important, the practical training received in the final year was the key predictor of becoming an inventor. Interestingly, however, second-year ranking, which reflect how well the student absorbed knowledge in specialized field courses, was a stronger predictor for the probability of obtaining a global patent (a patent granted outside Japan, such as the U.S., U.K., and so on) than graduation ranking. It thus appears that the theoretical knowledge in specialized fields, acquired during the second-year coursework, was especially important for inventions that had global reach.

A potential identification concern is that higher academic achievement and invention may both simply reflect general inventive ability and/or family background (cf. Nelson and Phelps, 1966). The differential impact of academic achievement across different years of university education appears inconsistent with this explanation. Nevertheless, we probe in several ways for the possibility that unobserved ability is a confounding factor. The evidence is more consistent with a casual interpretation of the school-absorption-to-invention hypothesis.

We first examine whether academic achievement predicts a higher likelihood of patenting in fields closely related to the individual's area of study. Second, we collect data on graduation theses topics and conduct textual analysis to assess the similarity between those topics and post-graduation patenting fields. In both analyses, we find a strong association between academic achievement and the probability of patenting in an area closely related to the individual's field of study or the thesis topic. Furthermore, patenting in division-related fields is more strongly associated with specialized knowledge absorbed in the classroom in the second year, while patenting in theses-related fields is more influenced by practical training received in the third year. These findings corroborate our notion

that the effects of university education achievement reflect the absorption of specialized knowledge rather than innate ability or family background.

Additionally, estimations that include high-school graduation rankings (as a proxy for general academic ability) show that these rankings do not have additional explanatory power once knowledge absorption at the university level is accounted for. Finally, university graduation rankings are more strongly correlated with the probability of becoming an inventor earlier in one's career (*i.e.*, before the median number of years between the graduation and the first patent application in our sample) than later. This is consistent with the idea that the impact of human capital accumulated in school gradually depreciates over time, which is less expected of any innate ability or family background effects.

The relationship between engineering education and the probability of becoming an inventor may be mediated by sorting into jobs where individuals can best utilize the knowledge acquired through education. To examine how much academic rankings affect the probability of becoming an inventor after controlling for post-graduation employer characteristics in the cross-sectional framework in which our baseline analysis is conducted, we identify the first and modal employers (firms, organizations, institutions, etc.) of graduates over the first 20 years of their careers. We then redo the same estimations while absorbing such employer dummies. Depending on the specification, this reduces the impact of university graduation rankings on the probability of invention by as much as 30 percent, but rankings remain a strong and statistically significant predictor of invention.[4]

The empirical findings indicate that the degree of absorption of different types of higher technical education had both distinct and cumulative impact on the formation of human capital that later enabled some graduates to become inventors. We employ a quantitative model of human capital

[4] In the Appendix, we also present the same analyses utilizing panel data to estimate the contribution of academic rankings while controlling for all employer fixed effects in the framework similar to Abowd et al. (1999) as well as the findings from an econometric mediation analysis (Heckman and Pinto, 2015). The findings are qualitatively very similar to those obtained in the cross-section analyses in the main text.

formation in the spirit of Cunha, Heckman, and Schennach (2010) to estimate the parameters of the "inventor production function" to obtain the probabilities of becoming an inventor in various counterfactual scenarios. The estimates indicate that the probability of becoming an inventor would be reduced by almost 90 percent if all students had the same level of academic achievement as the lowest-ranked students. Also, while superior academic achievement at earlier stages is generally important as it positively affects subsequent performance, "turning off" superior practical knowledge absorption in the final year has an outsized negative impact on invention outcomes.

Relation to the Literature. A large and growing literature has examined the relationship between higher education and technical change. Much of the empirical evidence in this literature, however, focuses on correlations between areas with strong science-based universities and private-sector innovation (Jaffe, 1989; Hausman, 2018; Andrews, 2023) or institutional-level reforms to measure the impact of education on innovation (Toivanen and Väänänen, 2016; Bianchi and Giorcelli, 2020). The role of high-level human capital in general, and college education in particular, in generating invention was also examined in some recent studies incorporating individual-level data (Akcigit et al., 2017; Bell et al., 2019). Most studies we are aware of have, however, implicitly or explicitly treat human capital developed through formal education as uniform, with conventional metrics such as years of schooling and attained degrees serving as proxies. However, it is not hard to fathom that considerable—maybe even most of the—heterogeneity in the propensity to innovate remains even conditioning on these observables. That is what we focus on here. We examine how the amount of expertise and proficiency individuals develop over the course of the *same* education impacts their subsequent innovative performance.[5]

[5] In a very different context, Bell et al. (2019) show that math and science test scores in school predict future patents.

Our study is also related to the vast literature on returns to education. One issue still debated in this literature is whether the observed returns to education (measured as life-time earnings) are driven by higher productivity due to human capital accumulated through education or by some other characteristics correlated with educational attainment which simply serves the signaling (sorting) role (Becker, 1964; Stiglitz, 1975; Weiss, 1985; Card and Krueger, 1992; Farber and Gibbons, 1996: Arcidiacono, Bayer, and Hizmo, 2010; Contreras, Rodriguez, and Urzúa, 2024). While the outcomes in our study are not earnings but inventions (patents), the sorting issue is important, as innovation-oriented employers may recruit graduates with high academic achievement, positioning them to succeed in generating inventions. In the main text, we focus on the relative importance of knowledge acquired in the classroom versus practical engineering training for different types of inventions, something that appears to have so far not been examined at all,[6] while controlling for post-graduation job sorting. The sorting issues are addressed in more detail in the mediation analyses in the Appendix.

Finally, our study is related to the emerging literature on the role of knowledge elites in innovation and growth, especially in the historical contexts of the two Industrial Revolutions (Castelló-Climent and Mukhopadhyay, 2013; Squicciarini and Voigtländer, 2015; Akcigit et al, 2017; Maloney and Caciedo, 2022; Hanlon, 2024). This literature has shown that as groups, educational elite, and especially engineers, disproportionately contributed to innovation and innovation-led growth. Due to data limitations, however, this literature could not address individual heterogeneity, so the issue of the importance of what and how much engineers learned during their education could not be examined. Utilizing individual-level academic achievement data, we show that differences in the degree of knowledge absorption matter for innovation outcomes even among highly educated engineers. In a related paper, Bhaskarabhatla et al. (2021) found that inventors' human capital explains a much larger

[6] The closest parallel that we are aware of is the examination of the effects of high school curricula on labor market outcomes in Altonji (1985).

share of variance in innovative output than firm capabilities. Our study highlights the role of knowledge absorption at the stage of higher education as a possible foundation for this.

The rest of the paper proceeds as follows. In Section II, we describe the historical background of Japan's industrialization and the university education systems and curriculum, and in Section III, we provide the brief description of our data. In Section IV, we examine how academic performance at various stages of higher technical education can be associated with the long-term outcomes of invention. In Section V, we present a quantitative model of the "inventor production function" and examine various counterfactual scenarios based on estimating structural equations. Section VI contains the discussion.

II. Historical Context and Engineering Education Curriculum

We first briefly describe the historical context of our study, focusing on the details about higher technical education curriculum in Japan at the time. See Yamaguchi, Braguinsky, and Nakajima (2021) for a general overview of invention by Imperial University graduates in Japan at that time.

II.1. The Emergence of Higher Technical Education in Japan

The first ever institution of higher technical education to emerge in Japan was created by the new Meiji government in 1873, spearheaded by the team of Scottish engineers led by Henry Dyer. After initial struggles due to language and other barriers (*Kyu Kobu Daigakko Shiryo*, 1931), the Imperial College of Engineering was inaugurated in 1877 and merged with the University of Tokyo to become Imperial University College of Engineering in 1886. This remained the only place to receive high engineering education in Japan until 1897 when the second (Kyoto) Imperial University was founded, and even in 1918 (the year of graduation of the final cohort in our sample) more than 70 percent of all university-educated engineers in Japan were graduates of Tokyo Imperial University College of engineering or its predecessor. See Hiraiwa, Braguinsky, and Agarwal (2026) for more details.

II.2. Education Curriculum at the Imperial University College of Engineering

The curriculum at Tokyo Imperial University College of Engineering followed a three-year structure, with three terms in each year.[7] Figure 1 Panel A shows the number of required courses (averaged over different divisions) in each year in the program over time. Even though the university education curriculum combined theory classes with practical exercises from the first year, most of the first two years were dedicated to classroom-based coursework, while the third year was mostly dedicated to practical education, with the number of required courses in the third year less than half of that in the first two years.

[Figure 1 around here]

In the first year, students mostly acquired foundational knowledge. Table A1 in Appendix A shows an example of the course curriculum for the mechanical engineering division in 1918. All courses for the first year are required, and most of those are fundamental subjects including "mathematics," "dynamics," "applied mechanics," and so on. In some divisions, the students were also required to perform experiments in laboratories, though the time allotted to these practical courses was limited during the first year.

In the second year, the students progressed to more advanced and applied courses tailored to specific majors, such as "hydraulic machinery" for those majoring in mechanical engineering, "stability and rolling of ships" for those majoring in naval architecture, and so on. Table A1 shows that in mechanical engineering, the second-year curriculum included courses such as "thermodynamics and heat engines," "metallic materials and metallurgical technology," and "building construction." Students were also required to spend more time performing experiments in laboratories from the

[7] According to the "general regulations for the colleges" in the Annual Calendar of Imperial Universities (*Tokyo Teikoku Daigaku Ichiran*) in 1918, the academic year began in September and ended in July. The first term extended from September to December, the second from January to March, and the third from April to July.

second year. Figure 1 Panel B shows the average number of lab experiment courses in each year in the program and indicates more laboratory work in the second and third years than the first. In mechanical engineering, students were required to take one lab experiment course involving three hours per week in each term of the first year, compared to six hours per week (except the first term in the third year) in the second and third years (see Table A1).

Most importantly perhaps, the number of hours spent in the classroom fell drastically in the third year. Instead, students completed practical training (often internships at private companies or government institutions) and wrote their graduation theses. Figure 1, Panel B shows most divisions offered one course dedicated to practical exercises in the third year, while some divisions required internships already in the second year. In mechanical engineering, for instance, the first term of the third year was entirely dedicated to "practical exercises," and in the remaining two terms, students needed to take just one required six-hour lab experiment course while completing their graduation theses. Although students also selected four elective courses, these involved much less classroom time—fewer than two hours per week per course—compared to the required coursework in the first and second years. Therefore, the third year was mostly dedicated to practical application of the knowledge acquired in the first two years.

The Imperial University's "regulations relating to practical training" stipulated "*the object of practical training is to make students acquainted with the practical application of scientific principles, and for this purpose students shall be trained in practical work or sent out to see various establishments or institutions of industry relating to their respective studies*" and was to be done "*either at the college or outside*" (*Tokyo Imperial University calendar*, 1917-1918, pp. 141-142). "Outside" practical training (internships at private companies or government institutions) often lasted up to half a year or more, with the length varying across students.[8] Students

[8] The aforementioned regulations relating to practical training read, "*The period for practical work or excursion shall be fixed by the Director according to the opinion of the professor in each course of study.*" (*Tokyo Imperial University calendar*, 1917-1918, p. 142). As

in practical exercises were guided by their mentors or faculty. The regulations stipulate, "*Students shall from time to time obtain directions for their work from the professor in charge, but when they by themselves undertake practical work outside the college, or make an excursion, they must previously receive directions from the professor*" (*Tokyo Imperial University calendar*, 1917-1918, p. 143).[9] After the practical exercises, the students were required to "*make out a report of their work and submit it to the professor in charge*" (*ibid.*) Appendix B shows several examples of reports by mechanical engineering students.

At the end of the third year, the students submitted graduation theses to obtain their degrees. While the themes of the theses were diverse, from theoretical essays to details about actual manufactured products or construction, many seemed inspired by students' work, observations, and experiences during practical exercises in the third year. The students often mentioned companies visited during their practical exercises in the acknowledgement section of their theses, and in some cases noted the companies provided materials used in the thesis work. Appendix B shows several examples of these graduation theses. It is noteworthy that almost all the theses during our sample were written in (often not perfect but understandable) English.

All in all, while the first two years of engineering education at Tokyo Imperial University were mostly dedicated to the classroom-based coursework (basic subjects in the first year and more applied subjects in the second year), third-year activities focused on more practical applications of knowledge taught in the classroom. We exploit this structure of the university curriculum to examine how individual-level absorption of different types of technical education influenced individuals' invention outcomes.

---

in mechanical engineering, most divisions dedicated one term (often the first) in the third year to practical exercises, so there were no other courses (or lab experiments) that term.

[9] The regulations also read, "*Students shall not be allowed to do anything except what they have been directed to do in the way of practical work; if a student is obliged by any circumstances to alter any point in the directions previously given, he shall submit the matter to the professor for his consideration and approval*" (*ibid.*)

III. Data and Summary Statistics

We first briefly describe our data and explain how we measure individual academic performance in higher technical education. Appendix D has more details.

III.1. Main Data Sources

Our main data source is the (Tokyo) Imperial University annual Calendars (*Tokyo Teikoku Daigaku Ichiran*, 1886-1918), available in open access from the National Diet Library Digital Collections (https://dl.ndl.go.jp/en/).[10] In each year, these calendars list all the students enrolled in each college comprising the university, separately by year and division (such as mechanical engineering, electric engineering, etc., for the College of Engineering). The calendars also list all previous graduates (bachelor's degree recipients) separately by graduation year and division. As mentioned, we use the data on 40 cohorts of bachelor's degree recipients from the College of Engineering of the (Tokyo) Imperial University and its predecessor, Imperial College of Engineering, starting from the first graduation cohort in 1879 and until 1918 (the last year for which individual-level academic achievement data are available, as explained immediately below).

We matched the list of graduates with the data from the Alumni Surveys of Imperial Universities (*Gakushikai Kaiin Shimeiroku*, 1890-1940, "*Gakushikai*"). Those surveys cover almost 90 percent of all graduates in our data and contain self-reported information about jobs and addresses updated on an annual basis, which we cross-checked against various other sources. Please see Appendix D for details.

We then matched the above data with patent data from the Japan Patent Office from 1885 (the year Japan introduced its patent system) until 1940. To obtain data on patents with global reach, we hand-matched the names and addresses of Japanese inventors to Google patents data with application dates prior to 1940. Details about these data and the matching process are in Appendix D.4.

[10] Most of those are in Japanese, but there are English versions called Tokyo Imperial University calendar for some years, from which we take the English course names and rules descriptions.

III.2. Academic Achievement

As mentioned, the data on Tokyo Imperial University graduates until the 1918 graduation cohort provide us with direct measures of students' academic achievement. All graduates through the class of 1918 were listed in the University Calendars in the order of their academic achievement (*sekiji*, in Japanese). Students' achievement scores were composed of average test scores over each of the three years of university education, with the score of their graduation thesis work added as the fourth component (e.g., *Tokyo Teikoku Daigaku Ichiran*, 1907, p. 198).[11] The graduate having the highest total score calculated as above (with all components weighted equally) was listed first among graduates from his division in a given year, the second-highest-scoring graduate was listed second, and so on.[12] While the test and graduation theses scores themselves are not available, the ordering provides us with an ordinal measure of relative academic achievement of each graduate. Note that because the listing in the University Calendars was separate by the academic (graduation) year and division (such as civil engineering, mechanical engineering, etc.), this measure is division-cohort-specific, that is, it reflects how well the student performed relative to his peers in the same division.

Furthermore, Imperial University Calendars (which start in 1886, after the Imperial University was inaugurated) listed currently enrolled students in the order of their average test scores in the previous year. That is, students enrolled in the second year were listed in the order of their average test scores received in the first year, while students enrolled in the third year are listed in the order of their average test scores received in the second year. Thus, starting from the class of 1888, we also have first- and second-year rankings in addition to integrative graduation rankings. We use these three rankings to measure different dimensions of academic performance and their impacts on invention.

[11] According to the "regulations for examinations" in the *Tokyo Imperial University calendar* (1917-1918, p. 137-138), examinations were held at the end of the first and third terms, but if instruction in any subject was completed in the second term, the exam in such subject was held at the end of that term.

[12] We use male pronouns in this paper because only males were admitted to the Imperial University during our sample.

Given that first-year rankings reflected average test scores in first-year courses, which were mostly fundamental subjects as explained in Section II.2, first-year rankings can be thought of as reflective of how well the students absorbed (relative to peers) the basic subjects taught in the classroom. Analogously, second-year rankings reflect how well the students absorbed (relative to peers) the more specialized knowledge taught either in the classroom or in the laboratory. Finally, further changes in rankings by the time of graduation reflect how well students performed in more practical work such as lab experiments, internships, and their graduation theses.

In addition to the above, over 80 percent of the graduates in our sample advanced to the university directly from one of the (old-system) high schools which served as gateways to Imperial Universities. Those three-year high schools were designed to prepare students for rigorous academic courses taught at Imperial Universities. Graduation from one gave automatic university admission. High school graduates were also listed in the order of their test scores, so we utilized high school calendars to measure students' relative academic performance in high school. This can serve as a proxy for general academic ability of students (relative to their peers). To make the rankings (both university and high school) comparable across cohorts and divisions, we compute their within-cohort-division quantile as $Norm_Rank_i = (N_i - R_i)/(N_i - 1)$, where R_i is the raw number of the ranking of graduate i, and N_i is the number of graduates in the cohort-division that graduate i belonged to. Thus, zero is the lowest rank, and one is the highest.

III.3. Graduation Theses and "Excursion" Reports

As mentioned, in the third and final year of university education, engineering students mostly engaged in practical work (internships which the students themselves and even some official university documents called "excursions," so this is the term we use hereafter to describe such internships) and writing their graduation theses. Those theses and excursion reports are partially preserved in the

University of Tokyo Faculty of Engineering libraries (the index can be seen online in the University of Tokyo Repository, https://repository.dl.itc.u-tokyo.ac.jp/about).

For the graduation cohorts in our sample, theses and excursion reports are available for the following College of Engineering divisions: mechanical engineering (including marine engineering, which became a separate subdivision in mechanical engineering from 1907), electric engineering, naval architecture, and mining/metallurgy (single division until 1910, split into mining and metallurgy divisions in 1911). Most of those are only available as hard copies in the University of Tokyo Faculty of Engineering libraries No. 2, 3, and 4, while some select ones are also available for viewing online through the aforementioned University of Tokyo Repository. For the purpose of this paper, we downloaded the titles of all available graduation theses and excursion reports until the 1918 graduation year and hand-matched the authors' names with the list of graduates in our main database. This gives us information about what themes the students chose for their theses work, which we then compared with their future patents (see Section IV.3). We also viewed hard copies of graduation theses and excursion reports written by the graduates in our sample in the Faculty of Engineering libraries to supplement this quantitative analysis with qualitative, historical data.

III.4. Summary Statistics

Table 1 shows the cross-section summary statistics of our main variables. Since the sample comprises engineers across different graduation cohorts, to make the measurement of their long-term outcomes comparable, we define inventors as those who patented at least once within 20 years of their graduation.[13] Among the 3,797 graduates in our sample, 584 (15.4%) were granted at least one patent in this time (and 81 were also granted at least one global patent). The share of inventors varies across graduation divisions. Table 1, Panel B shows the share of inventors is the highest in applied chemistry,

[13] Using all observations to define inventors leads to qualitatively similar results.

followed by electrical and mechanical engineering. This is consistent with Figure A1, where the technology classes of patents by university-educated engineers are mostly concentrated in these fields. Table 1, Panel C shows the trends in patenting by different graduation cohorts up to 1918. While the raw number of inventors increases toward later cohorts, the fractions of graduates who became inventors are not drastically different across cohorts.

Table 1, Panel A also shows the average cumulative number of patents by inventors in their first 20 years after graduation is 3.76, but the distribution is skewed: the standard deviation is 6.26 and mean to median ratio is 1.36. It also shows that academic performance in both university and high school is associated with becoming inventors after graduation. Inventors have higher normalized university graduation rankings, higher second-year and first-year class rankings, and higher high school graduation rankings than non-inventors, and the differences are statistically significant (inventors' average normalized rankings are 9-12 percentage points higher than for non-inventors). This lays the ground for our subsequent analyses where we examine how different types of technical education contribute to the propensity to become inventors.

Figure 2 shows the joint distributions of normalized academic rankings across different years and at graduation. We can see the clear positive associations between each measure, but there are elements off the diagonal. For example, some graduates with higher second-year rankings ended up with lower graduation rankings, and vice versa. In Section V below we present an explicit quantitative model to account for how different stages of education contribute to the overall human capital accumulation leading to potential invention. Overall, we have enough variation to identify the relative impact of various rankings on invention outcomes.

[Figure 2 around here]

IV. Technical Knowledge Absorption and Subsequent Invention

How did the absorption of higher technical knowledge shape graduates' inventive activities? The variation in academic performance at different stages noted above allows us to address this question at several levels, including which particular type of knowledge absorbed during the education process mattered most for future invention. We first examine the relationship between academic performance and the probability of becoming an inventor within the first 20 years after graduation in the whole sample. We then examine the same relationship using the match between patented fields and both graduation division and theses themes to probe potential causal mechanisms.

IV.1. Academic Performance and Invention

We first evaluate the role of academic performance at each stage of higher technical education in the post-graduation propensity to become inventors. Specifically, using all engineers in our sample, we estimate the following linear probability model:

$$y_i = \sum_k \beta_k Acad_Perf_{k(i)i} + d_{c(i)d(i)} + d_{c(i)h(i)} + d_{g(i)} + \varepsilon_i \qquad (1)$$

where y_i is an indicator for whether graduate i became an inventor within 20 years after graduation, and $Acad_Perf_{k(i)i}$ are the normalized academic rankings of graduate i in the $k(i)$-th year in the program (first-year, second-year, and at graduation), as well as his high-school graduation rankings. Since these rankings are cohort $c(i)$- and division $d(i)$-specific, we include cohort-division fixed effects $d_{c(i)d(i)}$ in all specifications as well as cohort-high-school fixed effects $d_{c(i)h(i)}$ whenever we include high-school graduation rankings in the specifications. We also include dummies for graduates' birthplace prefecture $g(i)$.[14]

[14] We also obtained estimates using the Poisson Pseudo Maximum Likelihood (PPML) model. One issue with this model is that since we need to include division-cohort fixed effects, the probability of becoming an inventor is not identified in division-cohorts where there were no inventors, leading to a smaller sample, biased toward inventors. The estimation

Column (1) in Table 2 Panel A shows that when we include only the university first-year rankings in the specification, they significantly predict the likelihood of becoming inventors. Going from the bottom of the first-year ranking distribution to the top (moving from zero to one) increases the propensity of becoming an inventor by 12.4 percentage points.

However, once we add the second-year rankings to the specification, the predictive power shifts from first-year rankings to second-year rankings (column (2)). When we further add the university graduation ranking, neither the first-year rankings nor second-year rankings predict the likelihood of becoming an inventor, only the university graduation rankings do (column (3)).

The result from column (2) may suggest that, while better absorption of basic knowledge in the first year could help graduates to invent in the long run, such absorption is relevant for future invention only if it leads to better absorption of more applied knowledge in the second year. There is no remaining, independent influence of basic knowledge. Similarly, the result from column (3) may indicate better absorption of more applied knowledge is relevant for future invention only if it leads in turn to higher integrated academic performance through practical experience in the third year. We observe a similar pattern when we include high-school graduation rankings (columns (4) to (6)), which embody general academic ability that could be innate or/and acquired before enrollment in the university. In each case, high school rank coefficient is statistically insignificant. Moreover, the magnitudes are small. It is the specialized knowledge absorbed in higher technical education that is associated with graduates' long-term invention outcomes. The specifications in columns (3) and (6) imply the likelihood of a graduate becoming an inventor increases by about 17-18 percentage points over the span of the university graduation ranking distribution. This is a big difference; recall that the unconditional probability of becoming an inventor in the sample is about 16%.

---

results from the PPML model are qualitatively similar to those presented here, while the relative marginal effects of various rankings are more strongly pronounced which is not surprising given that PPML drops division-cohorts with no inventors. Details are available upon request.

[Table 2 around here]

Our dependent variable is dichotomous. One might naturally wonder if knowledge absorption during higher engineering education also influences graduates' invention activity on the intensive margin. Table A2, Panel A in Appendix A shows the results from an alternative estimation where we use the IHS-transformed cumulative number of patents by graduates up to 20 years after graduation as the dependent variable. The qualitative implications are the same: higher university graduation rankings predict a higher number of patents by graduates, even after controlling for general academic ability represented by high-school graduation rankings, and the contributions of the first- and second-year rankings to patenting are muted when including university graduation ranking in the model.

The likelihood of becoming an inventor is particularly associated with graduation rankings at the right tail of the distribution. Indeed, Figure A2 in the Appendix shows that there is almost no difference in the likelihood of becoming an inventor until about the sixth decile, with the likelihood increasing sharply for each decile beyond that. Thus, it appears that truly superior academic performance had the strongest association with future invention.

IV.2. Area of Expertise: Graduation Division

The findings in the previous section indicate the extent to which individuals absorb higher technical knowledge *integrally*, represented by university graduation rankings, could be the main contributing factor in predicting the likelihood of becoming an inventor. If knowledge absorbed at university effectively contributes to patenting by graduates, however, we should expect them to be particularly likely to patent in technology fields closely related to their education background.

To examine this, we create a bridge between patent classes (*rui*) in the Japan Patent Office data and engineering department divisions to identify patent classes most closely related to each field of study (see Appendix D). We construct an indicator equal to one if a graduate patented within 20 years

of graduation at least one invention closely related to the division from which he graduated. We then use this indicator as the dependent variable in estimating otherwise the same regression models as in Section IV.1. By using a dependent variable that specifically captures the relationship between patent technology fields and technology-specific knowledge absorbed in the university, we aim to examine whether higher technical education shapes the *types* and *directions* of inventions by graduates.

The estimation results are in Table 2, Panel B. They suggest the absorption of specialized knowledge is indeed relevant to the types and directions of invention. The results in the full model specification in column (6) indicate university graduation rankings positively and significantly predict patenting in technology classes closely related to graduates' divisions, although the association is somewhat less pronounced than in column (6) in Panel A. Going from the bottom of the division-cohort to the top increases the probability of becoming an inventor in a division-related area by 11 percentage points.

Interestingly, the magnitudes of the second-year ranking coefficients in columns (3) and (6) of Panel B are an order of magnitude larger than their counterparts in Panel A, though their statistical imprecision warrants circumspection. This is consistent with second-year rankings measuring how well graduates absorbed applied knowledge tailored to their specific majors. Perhaps specialized knowledge tailored to each division may play a more important role for division-related inventions.[15]

IV.3. Area of Expertise: Theme of Graduation Thesis

Even though graduation divisions represent the areas of expertise graduates develop through university education, the depth and focus of specialized knowledge can vary significantly among graduates within the same division, depending on their narrower academic interests and, in particular, specific research conducted for their graduation theses. If such granulated specialized knowledge or

[15] Table A2, Panel B in Appendix A shows qualitatively similar estimates when we use as the dependent variable the IHS-transformed cumulative number of patents whose classes are related to inventors' graduation divisions.

research findings play a role in future patenting, graduates might be particularly likely to produce patents closely related to such knowledge.

To explore this, we examine the links between the themes of graduation theses and patent characteristics. As mentioned, graduation theses are available only for a subset of divisions in our sample. Hence, the analysis in this section is limited to graduates from the divisions for which graduation thesis records exist: naval architecture, mechanical engineering, and electric engineering.[16] For a basis of comparison, Table A3, Panels A and B in Appendix A replicate the estimations in Table 2, Panels A and B for the subsample of graduates from these divisions where we can match patents to graduation thesis themes. The estimation results are qualitatively similar to the full sample.

To measure linkages between patents and themes of graduation theses, we compute the sentence embeddings of patent characteristics including titles (*meisho*), classes (*rui*), categories(*dai-bunrui-meisho*), and sub-categories (*sho-bunrui-meisho*), as well as themes of graduation theses. We use the Sentence-BERT, or SBERT (Reimers and Gurevych, 2020), the modification of the *Bidirectional Encoder Representations from Transformer*, or BERT (Devlin et al., 2019). Specifically, we map each sentence to a 384-dimensional vector space by SBERT pre-trained models, taking the semantic meanings of the sentences into account. Given the vector representations of sentences, for each patent observation, we calculate the cosine similarity between each patent characteristic and the theme of the graduation thesis by the patent holder. (More details on the methodology can be found in Appendix C.) We categorize patents as closely related to the patentors' graduation theses if the cosine similarity is above the 75th percentile threshold.[17] We then define an indicator for whether graduates patented at least one invention related to themes of their graduation theses within 20 years after graduation and estimate the same regression models as in Section IV.1.

[16] Graduation theses for the mining division are also available, but their titles are almost universally about designs for specific mines and not usable to measure textual linkages to patent classes.
[17] Table A4 in Appendix A uses the 90th percentile threshold instead, and the results are qualitatively similar.

The estimation results are in Table 2, Panel C. They show similar patterns to those in previous sections. Specifically, while the absorption of classroom-based knowledge in the first and second years is correlated with future inventions, such knowledge must facilitate knowledge absorption in the third year to generate thesis-related future inventions. The size of the effect of the university graduation rankings is estimated to be significantly higher than in the findings on division-related patents in Panel B. Students wrote graduation theses in their third year, so it might be expected that how well they performed in crafting those theses affected future inventions related to the themes of those theses. Table A2, Panel C and Table A4 in the Appendix A also show qualitatively similar estimates when we use (i) the IHS-transformed cumulative number of patents related to thesis themes, (ii) other patent characteristics (classes, categories, and sub-categories) to measure the similarity, and (iii) the 90th percentile threshold to define patents closely related to themes of the patentors' graduation theses.

IV.4. Globally Important Inventions

The analysis so far has investigated patent quantity, but quality (impact) of patents is of course important as well (Lanjouw and Schankerman, 2004). In this section we examine how high-quality inventions were related to academic performance in higher technical education.

Many studies use patent citations as a quality measure. However, patent citations data are not available for our time period. We instead hand-matched Japanese inventors' patents and names with those of inventors/assignees in the Google patent data through 1940. This identifies graduates who were also granted patents outside of Japan. Because it was costly to apply for patents in advanced countries, patents granted outside of Japan can be considered of higher average quality than those granted only in Japan. Eighty-one graduates in our data (14 percent of all inventors) were granted at least one global patent within the first 20 years after graduation.[18]

[18] Some of those patents are still being cited today. See Nicholas (2011) for more about globally important innovations coming out of Japan during that time.

Table 2, Panel D shows the results from the same specification as in Section IV.1 using as the dependent variable an indicator for whether graduates obtained at least one global patent within 20 years after graduation. In contrast to our previous findings, there is still substantial predictive power from second-year rankings even controlling for integrative graduation rankings. This suggests that the degree to which the graduates absorbed theoretical knowledge taught in the second year had as much or even more influence on graduates becoming inventors on the global stage, while most domestic inventions were more driven by the knowledge absorbed through practical training at the end of the curriculum. Table A2, Panel D in Appendix A shows qualitatively similar estimates when we use the IHS-transformed cumulative number of global patents for 20 years after graduation.

IV.5. Causality

An obvious potential identification threat to the causal interpretation of the absorption of higher technical knowledge on future inventions is that graduates might have innate or acquired inventive ability before their enrollment in Imperial University, and academic achievement might just reflect the amount of such ability. It is worth noting that the evidence above suggests that the type of knowledge the graduates absorbed better at university predicts the direction of future inventions the graduated engineers made. It is unlikely that *innate* differences in inventive ability would govern differences in patenting in particular technology classes (cf. Bell et al., 2019).

It is conceivable, however, that graduates might have already nurtured abilities to invent *in specific technology areas* (e.g., through exposure during childhood as in Bell et al., 2019) before they enroll, and this might also be reflected in their absorption of related knowledge after enrollment. That direction-specific endogeneity concern seems farther-fetched, however, in the case of students' graduation theses. Any such pre-existing, direction-specific inventive ability investment would have had to occurred at least three years before the student would even know where they would go for their third-year internship and what theme they would choose for their thesis work. Moreover, even if students

were already interested in specific technology areas that would later become the themes of their graduation theses, it would have been hard for them to acquire enough skills to invent in such specific technology areas without specialized knowledge taught in the university and faculty-directed internships. Thus, even in the case where there is sorting by general or even division-related inventive ability potential at enrollment, such ability must have been combined with specific knowledge absorbed in the university.

IV.6. Common-Sense Validation: Early- vs Late-Career Inventors and a Placebo Test

Our findings so far indicate that students' absorption of higher technical knowledge taught to them in the classroom and in practical training ("excursions") affected their likelihood of becoming an inventor. Graduates in our data patent at different points after graduation, however. Almost nobody applies for a patent in the first year after graduation, and the median (and average) length of time to first patent application in our sample is 12 years after graduation. A standard model of human capital accumulation with depreciation over time suggests that knowledge absorption during university education would be relatively more important earlier in their careers rather than later. If that is the case, we would likely observe stronger relationships between university education and patenting (in education-related fields) for those who patented earlier in their careers.

To examine the above conjecture, we define "early inventors" as graduates whose first patent application happened before the median number of years (12 years) to first patent after graduation. "Late inventors" are those whose first patent application happened after 12 years (up to year 20). We then estimated the same regression models as in Section IV.1-IV.3 using these indicators.

Table 3 presents the estimation results. In the first two columns we show results for all early and late inventors, as in Column (6) of Table 2, Panel A. In the next four columns the dependent variables are early and late inventors who had at least one patent closely related to their field of study, and at least one patent closely related to their theses themes, as in Columns (6) of Table 2, Panels B and C,

respectively. In all cases, university graduation rankings are more strongly related to the probability of becoming an inventor for those whose first patent application was before the median number of years to such application, with the difference between the coefficients on graduation rankings having p-values of 0.060 between columns (1) and (2), 0.099 between columns (5) and (6), and 0.048 between columns (3) and (4).[19]

[Table 3 around here]

As mentioned, a concern in terms of identification is that both in-program and graduation rankings and future invention may both be simply explained by higher innate ability. We argued above that this is not likely to be the case for several reasons. Here we go one step further by conducting a "placebo test" where we examine whether higher graduation rankings were also associated with the promotion to executive positions in the private sector. The idea is that if higher graduation rankings just reflect higher overall ability, we should find positive association not just with patents, but with such promotion as well.

Table 4 shows the results of our standard regression specification with the outcome variable replaced by the probability of becoming an executive of a private company in the first 20 years after graduation.[20] In columns (1)-(3) the outcome includes executive positions in all private-sector employers, while in columns (4)-(6) we limit the outcome to executive positions in firms affiliated with the four largest business groups in Japan—*Mitsui*, *Mitsubishi*, *Sumitomo*, and *Yasuda*—commonly referred to as (the four largest) *zaibatsu*.

[19] Note that reported pairwise *t*-tests treat the two standard errors as independent, while in fact early and late patent outcomes are likely to be positively correlated across individuals. Thus, the true standard errors of the difference are likely to be smaller and the *p*-values might be overstated.

[20] We identify executives in our data based on their job titles; an individual is an executive in a given year if he is the company president ("*shacho*" or "*kaicho*"), a board member ("*torishimariyaku*," including also executive directors "*jomu torishimariyaku*" and "*senmu torishimariyaku*"), or an auditor ("*kansayaku*").

In columns (1) and (4), the outcome is the dummy equal to one if the individual became an executive at some point during the first 20 years after graduation, regardless of whether he also had a patent or not.[21] The estimated coefficient on the graduation ranking variable is positive, although statistically not significant at conventional levels. In columns (2)-(3) and (5)-(6), the outcome variables are dummies equal to one if an individual who had (in columns (2) and (5)) or did not have (in columns (3) and (6)) at least one patent in the first 20 years after graduation became an executive. Comparing the estimated coefficients in columns (2) versus (3) and (5) versus (6), it is clear any impact of graduation rankings on the probability of being promoted to an executive position was significant only for inventors. There is no relationship whatsoever for those who were not inventors.

To sum up, graduation rankings may be mediating promotion through their impact on patenting (indeed, we confirmed in unreported regressions that inventors are more likely than non-inventors to be promoted to executive positions, although not in *zaibatsu*-affiliated companies). However, we find no evidence that they directly affected the probability of promotion, in stark contrast with the findings in Table 2.

IV.7. Controlling for post-graduation jobs

The relationship between engineering education and the likelihood of becoming an inventor may be mediated by sorting into jobs where individuals can most effectively apply the knowledge acquired through education. To examine how much of the effect of academic rankings on the probability of becoming an inventor remains after controlling for post-graduation employer characteristics in our baseline cross-sectional analyses, we identify the first and modal employers (firms, organizations, institutions, etc.) of the graduates in our sample over the first 20 years of their careers. We then redo the same estimations as in Table 2 above, augmenting specifications with fixed effects for both the

[21] In the regressions in Table 4 the sample includes only the "population at risk," that is, those graduates who worked for a private company at some point during their careers.

student's first employer after graduation and their modal employer during the first 20 years after graduation.[22] (In the Appendix, we present complementary findings from a similar exercise using panel data, as well as results from a mediation analysis where we decompose the effect of graduation rankings on the probability of becoming an inventor into direct and measured indirect effects through various mediators, following the framework of Heckman and Pinto, 2015, and Heckman et al., 2013).

Table 5 presents the results. Comparing the estimated coefficients across the corresponding panels of Tables 2 and 5, we can see that controlling for employer characteristics reduces the estimated impact of university graduation rankings on the probability of invention within 20 years by as much as 30 percent in some specifications. Nonetheless, graduation rankings remain a strong predictor overall. It is also worth noting that including first-employer fixed effects generally leads to a larger reduction in the magnitude of the graduation ranking coefficient than absorbing modal-employer fixed effects. This is consistent with graduation rankings affecting graduates' sorting to their first jobs and subsequent job changes being driven more by career events.

Furthermore, comparing the coefficients in Panels C of Tables 2 and 5 shows that including modal-employer fixed effects has no impact on how graduation rankings affect the probability of invention in the individual's graduation thesis-related field. This is consistent with the causal interpretation of the effect of knowledge absorption on the likelihood of becoming an inventor as discussed above, rather than employer sorting. The findings from panel data analyses while controlling for a full set of employer fixed effects with a framework similar to Abowd et al. (1999) are qualitatively similar to those obtained in the cross-sectional setting here. The direct effect of education rankings also remains strong in econometric mediation analysis (Heckman and Pinto, 2015); these results are also presented in the Appendix.

[22] The correlation coefficient between the first and model employer is 0.45. In cases of ties between two modal employers, we assigned the modal employer based on randomly assigned employer ids.

V. Human Capital Accumulation and Counterfactual Analyses

The empirical findings in the previous section indicate graduates' absorption of different types of higher technical education during their three years of university education exerted both separate and cumulative impacts on their human capital formation, which later enabled some of them to become inventors. In this section, we present a parsimonious quantitative model to organize and interpret the reduced-form estimates in the previous section.

Specifically, we adapt to our context the dynamic human capital production process model from Cunha, Heckman, and Schennach (2010), where earlier stages of the accumulation process affect subsequent stages, and the resulting human capital stock then affects the outcome (invention, in our case). We recover the parameters of this "inventor production function" from observed outcomes (patenting) and academic rankings at various stages of the education process. We then utilize those parameters to conduct counterfactual estimations of the probability of becoming an inventor under various counterfactual scenarios—for instance, what sort of invention outcomes to expect if all students had the same level of academic achievement as lowest-ranked students at each stage of their education and across all three years of university education.

V.1. General Formulation

Consider latent outcomes describing the absorption of different knowledge during university engineering education. Let ω_{1i}, ω_{2i}, and ω_{3i} represent the level of absorption of basic (first-year), applied (second-year), and practical (third-year) knowledge by graduate i, respectively. In each year, the knowledge is absorbed by combining knowledge accumulated in preceding years and idiosyncratic shocks, represented by η_{1i}, η_{2i}, and η_{3i}, respectively:

$$\omega_{1i} = g_1(\omega_{0i}, \eta_{1i}),$$

$$\omega_{2i} = g_2(\omega_{0i}, \omega_{1i}, \eta_{2i}),$$

$$\omega_{3i} = g_3(\omega_{0i}, \omega_{1i}, \omega_{2i}, \eta_{3i}),$$

where ω_{0i} is the general academic ability and $g_1(\cdot)$, $g_2(\cdot)$, and $g_3(\cdot)$ represent the "policy functions" for knowledge absorption in each year. The nature of the reliance on knowledge absorbed in previous years is not explicitly specified, but we presume that better absorption in preceding years can help the absorption of more advanced knowledge, or that graduates could make more efforts or investments in learning based on their previous performances (cf. Cunha et al., 2010). Assume that the absorption of knowledge in each year is reflected in the corresponding test scores, denoted by $s_{0i} = s_0(\omega_{0i})$, $s_{1i} = s_1(\omega_{1i})$, $s_{2i} = s_2(\omega_{2i})$, and $s_{3i} = s_3(\omega_{3i})$, respectively. While we do not observe these test scores directly, we observe graduates' rankings, r_{ki}, based on these scores in year k in the program ($k = 0, 1, 2$). Graduation rankings, r_{Gi}, are determined by the graduation score, $s_{Gi} = w_1 s_{1i} + w_2 s_{2i} + w_3 s_{3i}$, where $w_1, w_2, \text{and } w_3$ are the appropriate (known) weights. Finally, we can formulate the "inventor production function" as follows:

$$y_i = f(\omega_0, \omega_1, \omega_2, \omega_3, O_i, \varepsilon_i)$$

where y_i is a dummy equal to one if graduate i had a patent in the first 20 years after graduation, O_i is a vector of other observable characteristics (such as the first or modal employer), and ε_i is an idiosyncratic "productivity" shock.

V.2. Parametric Specifications

For the practical specification, we assume that (i) (additively invariant) test scores reflect each logged knowledge stock ($s_k(\cdot) = \ln(\cdot)$), (ii) the policy functions are Cobb-Douglas, and (iii) the inventor production function is also Cobb-Douglas. Formally, we can specify the model as:

$$s_{0i} = \ln \omega_{0i} = \eta_{0i}, \tag{1}$$

$$s_{1i} = \ln \omega_{1i} = \rho_{10} \ln \omega_{0i} + \eta_{1i}, \tag{2}$$

$$s_{2i} = \ln \omega_{2i} = \rho_{20} \ln \omega_{0i} + \rho_{21} \ln \omega_{1i} + \eta_{2i} \tag{3}$$

$$s_{3i} = \ln \omega_{3i} = \rho_{30} \ln \omega_{0i} + \rho_{31} \ln \omega_{1i} + \rho_{32} \ln \omega_{2i} + \eta_{3i} \tag{4}$$

$$y_i = \beta_0 \ln \omega_{0i} + \beta_1 \ln \omega_{1i} + \beta_2 \ln \omega_{2i} + \beta_3 \ln \omega_{3i} + \Gamma(O_i) + \varepsilon_i, \tag{5}$$

where $\Gamma(\cdot)$ is some function. Because the test scores are arbitrarily scaled and so are the parameters β_k in (5), we normalize the scores' standard deviation to one. To capture differences in knowledge accumulation across different university and high-school cohorts, we assume η_{ki} can be decomposed into a cohort-specific effect and a zero-mean error term:

$$\eta_{0i} = \lambda_{0h} + e_{0i} \quad \text{and} \quad \eta_{ki} = \lambda_{kc} + e_{ki} \quad \text{for } k = 1, 2, 3,$$

where λ_{0h} and λ_{kc} are the mean effects that are specific to high school h or university cohort c that graduate i belongs to, respectively. We also assume each e_{ki} is independently normally distributed, and therefore the marginal distribution of each score follows normal distribution.

V.3. Estimation of the Parameters

Our goal is to estimate the structural parameters given the observations of inventor outcomes, y_i, rankings, r_{0i}, r_{1i}, r_{2i}, and r_{Gi}, and other observable characteristics, O_i. While multiple approaches exist, here we take the naïve but simplest approximation. Assuming the rankings reflect the exact percentile points in test score distributions, we estimate the equations (2)-(5) as if we know the latent test scores and thus knowledge stocks.[23] Let u_r and b_r denote the maximum and minimum of the percentiles in test score distributions. Given the number of students in university cohort c or high-

[23] Alternatively, we can use SMM (Simulated Method of Moments) where we treat test scores as latent variables and see if the moments of observed variables, that is, y_{it} and r_{ki}, are consistent with our estimated model, or SML (Simulated Maximum Likelihood or Gibbs Sampler) where the conditional probability distribution of latent knowledge stocks is estimated using simulated data or Bayesian approach. These will be tried later.

school cohort h that graduate i belongs to, we can transform the rankings into the inferred percentiles as follows:

$$z_{ki} = b_r + \frac{N_c - r_{ki}}{N_c - 1} \cdot (u_r - b_r),$$

where z_{ki} represents the percentile of graduate i's test score in year k in the program, N_c is the number of students in university cohort c, which we can replace with the number of students in high-school cohort h, N_h, for the high-school ranking.

Given these percentile point estimates, we construct the normalized test scores by applying the inverse of the normal distribution function:

$$\tilde{s}_{ki} = \Phi^{-1}(z_{ki}).$$

We then estimate the parameters in equations (1)-(5) using the normalized scores and including appropriate fixed effects. Formally, using these normalized test scores, we can rewrite the equations (1)-(5) as follows:

$$\tilde{s}_{0i} = \ln \omega_{0i} - E_h(\ln \omega_{0i}) = e_{0i}, \tag{6}$$

$$\tilde{s}_{1i} = \ln \omega_{1i} - E_c(\ln \omega_{1i}) = \rho_{10}\tilde{s}_{0i} + \varphi_{1h} + \varphi_{1c} + e_{1i}, \tag{7}$$

$$\tilde{s}_{2i} = \ln \omega_{2i} - E_c(\ln \omega_{2i}) = \rho_{20}\tilde{s}_{0i} + \rho_{21}\tilde{s}_{1i} + \varphi_{2h} + \varphi_{2c} + e_{2i} \tag{8}$$

$$\tilde{s}_{3i} = \ln \omega_{3i} - E_c(\ln \omega_{3i}) = \rho_{30}\tilde{s}_{0i} + \rho_{31}\tilde{s}_{1i} + \rho_{32}\tilde{s}_{2i} + \varphi_{3h} + \varphi_{3c} + e_{3i} \tag{9}$$

$$y_i = \beta_0\tilde{s}_{0i} + \beta_1\tilde{s}_{1i} + \beta_2\tilde{s}_{2i} + \beta_3\tilde{s}_{3i} + \Gamma(O_i) + \varphi_{yh} + \varphi_{yc} + \varepsilon_i, \tag{10}$$

where φ_{kh} and φ_{yh} are high-school cohort fixed effects, and φ_{kc} and φ_{yc} represent the university cohort fixed effects ($k = 1, 2, 3$).

Equation (10) is basically the same as our regression specification employed in Section IV, except for using the inferred test scores instead of the normalized rankings. We can estimate these equations separately, so just running each regression will recover the underlying structural parameters.

One issue is that we do not directly observe third-year rankings, r_{3i}, so we cannot directly obtain z_{ki} from those. Instead, we observe first- and second-year rankings, r_{1i}, and r_{2i}, as well as graduation rankings, r_{Gi}, which are the weighted sum of the three. Without knowing underlying variance of graduation scores and correlations between the third-year scores and other test scores, we cannot immediately obtain $\tilde{s}_{3i}$ from those. However, we can still algebraically recover those third-year normalized scores from the observed variables as detailed in Appendix F. Using those normalized scores, we can estimate the inventor production function and recover the structural parameters.

V.4. Estimation Results and Counterfactual Scenarios

The estimation results from equations (7)-(10) using inferred high-school, first-, second-, and third-year scores are presented in Table A5 in the Appendix. The first three columns respectively present the results of regressing first-year scores on high-school scores; second-year scores on high-school and first-year scores; and third-year scores on high-school, first-year, and second-year scores. The estimates confirm that higher scores in each later stage are positively associated with higher scores in all previous stages, but especially in the immediately preceding stage, as in Cunha et al. (2010).

The last column shows the inventor production function. While estimated coefficients in this column are not directly comparable to those in Table 2, the overall picture remains the same: third-year scores, which reflect students' performance in practical engineering training and in graduation theses, are most predictive of future invention. Importantly, however, the previous column makes clear that third-year scores are not independent of absorption in earlier stages. Indeed, those reflect the cumulative process of human capital accumulation, starting at least from high school.

We now use the model's estimates to examine five counterfactual scenarios: (i) all high school students had the same scores as the students at the bottom of the distribution, (ii) all first-year university students had the same scores as the students at the bottom of the distribution, (iii) all second-year university students had the same scores as the students at the bottom of the distribution, (iv) all third-year university students had the same scores as the students at the bottom of the distribution, and (v) all students across all three years at university had the same scores as the students at the bottom of the distribution. Figure 3 shows the estimated probability of becoming an inventor in the first 20 years after graduation in the observed data as well as in each of the above five counterfactual scenarios.

A first clear pattern is the importance of cumulative human capital acquired through all three years at the university. Under scenario (v)—depicted at the very right of the graph in Figure 3—the estimated likelihood of becoming an inventor falls to 1.8 percent, compared to 16 percent in the actual data. The impact of "turning off" superior performance at just one education stage is more modest; the negative impact is relatively larger when all students are reduced to the bottom of the distribution in high school than in the first university year, and larger in the first year than in the second. This difference is not surprising, as earlier superior performance affects outcomes not only directly but also indirectly through its influence on performance at subsequent stages. Importantly, however, this monotonicity breaks down at the third year. Confining all students in the third year to the same scores as bottom students results in a larger drop in future invention than the same counterfactual scenario applied to the second year. This is despite the fact that, as the final year of study, third-year scores have no influence on subsequent scores. This finding once again confirms the distinctive role of knowledge absorbed through practical training ("excursions") and graduation thesis work in shaping future invention activities.

VI. Discussion

Using unusually detailed individual-level data in the setting of Japan's higher engineering education in the late 19th-early 20th century, we found that even within the same cohorts of engineers, all of whom took the same classes and were exposed to the same education during their university years, those with higher academic rankings were more likely to become inventors. These findings complement the existing literature that showed the importance for predicting innovation of educational *attainment* (years of schooling or degrees), especially in science and engineering. We argue and find evidence that how much students absorbed in the classroom and in practical internship-like training (as reflected in their academic ranking at graduation) was also important.

While data availability led our empirical focus to be on higher engineering education at Japan's Imperial University at the turn of the last century, we believe our setting and results offer broader lessons. Even though the subjects and content may be different, the gradual progression from more general to more practical knowledge over the course of education, as well as individual heterogeneity in the "absorptive capacity," are not too different in higher science and engineering education today. The mechanisms we discover here could easily operate in other countries and time periods; they might just be difficult to isolate in standard datasets. Furthermore, the data span the era of industrialization for Japan, accompanied by setting up the modern higher technical education system for the first time in the country's history. Information as detailed as our data is unusual even in today's advanced countries, to say nothing of developing countries whose situation might be more similar to that of Japan at the time of our analysis. Hence, we believe that one can draw important lessons regarding the development of a higher technical education system and how that translates into invention and innovation.

Our findings indicate that knowledge absorbed in higher technical education, especially through practical engineering training, had a strong and lasting impact on future invention. Utilizing similarity

measures between graduation theses' themes and technological classes of future patents, we demonstrate a substantial impact of the content of education at the university on future invention, alleviating identification concerns about unobserved ability. We also show graduation rankings did not predict promotion to executive positions, which might have been expected if rankings reflected higher general ability rather than better engineering knowledge absorption. Controlling for post-graduation sorting to different employers reduces the magnitude of the estimated impact of graduation rankings on invention outcomes by 10-30 percent, depending on specifications, but the effect remains strong and statistically significant in basically all cases.

We then estimate the parameters of the "inventor production function" in a quantitative human capital accumulation model of a framework similar to Cunha et al. (2010). We evaluate various counterfactual scenarios where we assume that all students only had the same achievement as lowest-ranked students. Forcing the students to only accumulate human capital accumulated by bottom students in all three years at the university results in almost 90 percent reduction in the probability of becoming an inventor, while forcing them to acquire the same human capital as bottom students in the third and final year only results in a 31 percent reduction in such probability.

Our study has important economic and policy implications. From an economic perspective, our study implies that not just years of education, but also the degree to which educational content is absorbed by students, is a key determinant of their future productivity as inventors. At the same time, part of the knowledge absorption effect on future invention is mediated by employer characteristics, underscoring the importance of managing talent in a way that facilitates inventive activity by high-achieving engineers and creates effective complementarities. From a policy perspective, our findings highlight the importance of links between theory and practice in higher technical education, especially the importance of internships.

References

Abowd, J., Kramarz, F., Margolis, D. (1999). High wage workers and high wage firms. *Econometrica*, 67(2), 251-333.

Akcigit, U., Grigsby, J., and Nicholas, T. (2017). The rise of American ingenuity: Innovation and inventors of the golden age. NBER Working Paper No. 23047, Cambridge, MA: National Bureau of Economic Research.

Altonji, J. (1995). The effects of high school curriculum on education and labor market outcomes. *Journal of Human Resources, 30*(3), 409-438.

Andrews, M. (2023). How do institutions of higher education affect local invention? Evidence from the establishment of US colleges. *American Economic Journal: Economic Policy, 15*(2), 1-41.

Arcidiacono, P., Bayer, P., and Hizmo, A. Beyond signaling and human capital: Education and the revelation of ability. (2010). *American Economic Journal: Applied Economics, 2,* 76-104.

Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. National Bureau of Economic Research, Cambridge, MA.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. (2019). Who becomes an inventor in America? The importance of exposure to innovation. *Quarterly Journal of Economics*, *134*(2), 647-713.

Bhaskarabhatla, A., Cabral, L, Hedge, D., and Peeters, T. (2021). Are inventors of firms engines of innovation? *Management Science*, 67(6), 3321-3984.

Bianchi, N., and Giorcelli, M. (2020). Scientific education and innovation: from technical diplomas to university STEM degrees. *Journal of the European Economic Association*, *18*(5), 2608-2646

Card, D., and Krueger, A. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy, 100*(1), 1-40.

Castelló-Climent, A., and Mukhopadhyay, A. (2013). Mass education or a minority well educated elite in the process of growth: The case of India. *Journal of Development Economics*, *105*, 303-320.

Contreras, D., Rodriguez, J., and Urzúa, S. (2024). Is private education worth it? Evidence from school-to-work transitions in Chile. *Labour Economics, 90*, 1-14.

Cunha, F., Heckman, J., and Schennach, S. (2010) Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3), 883-931.

Devlin, J., Chang, M-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Fagereng, A., Mogstad, M., and Rønning,, M. (2021). Why do wealthy parents have wealthy children? *Journal of Political Economy, 129*(3), 703-756.

Farber, H., and Gibbons, R. (1996). Learning and wage dynamics. *Quarterly Journal of Economics, 111*(1), 1007-1047.

Frake, J., Hagemann, A., Uribe, J. (2024). Collider bias in strategy and management research: An illustration using women CEO's effect on other women's career outcomes. *Strategic Management Journal, 45*(7), 1393-1419.

Gakushikai Kaiin Shimeiroku (Listings of Members of the Alumni Association of Imperial Universities, in Japanese), 1890-1940. Gakushikai, Tokyo.

Grant, R. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal, 17,* 109-122.

Hanlon, W. (2024). The rise of the engineer: Inventing the professional inventor during the Industrial Revolution. NBER Working Paper 29751, Cambridge, MA.

Hausman, N. (2022). University innovation and local economic growth. *Review of Economics and Statistics, 104*(4), 718-735.

Heckman, J., and Pinto, R. (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric Review, 34*(1-2), 6-31.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review, 103*(6), 2052-2086.

Hiraiwa, T., Braguinsky, S., and Agarwal, R. (2026). If we build it, we will come: Institution development in academia and the evolution of career choices by top talent during Japan's industrialization. NBER Working Paper 33471, Cambridge, MA.

Hvide, H. K., and Jones, B. F. (2018). University innovation and the professor's privilege. *American Economic Review*, *108*(7), 1860-1898.

Jaffe, A. (1989). Real effects of academic research. *American Economic Review*, *79*(5), 957-970.

Kyu Kobu Daigakko Shiryo (Materials of the Former Imperial College of Engineering, in Japanese). (1931). Toranomon-kai, Tokyo. Available at https://dl.ndl.go.jp/pid/1464148/1/1

Lanjouw, Jean, and Mark Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *Economic* Journal, *114*(495), 441-465.

Maloney, W.F. and Caicedo F. V. (2022). Engineering Growth. *Journal of the European Economic Association*, *20*(4), 1554-1594.

Metcaffe, R., Sollaci, A., and Syverson, C. (2023). Managers and productivity in retail. NBER Working Paper, w31192, Cambridge, MA.

Nelson, R., and Phelps, E.S. (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review*, *56*(1/2), 69-75.

Nicholas, T. (2011). The origins of Japanese technological modernization. *Explorations in Economic History*, 48: 272-291.

Reimers, N., and Gurevych, I. (2020). Making monolingual sentence embeddings multilingual using knowledge distillation. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 4512-4525, Online. Association for Computational Linguistics.

Romer, P.M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5, Pt. II), S71-S102.

Rosenkopf, L., and Almeida, P. (2003). Overcoming local search through alliances and mobility. *Management Science, 49*(6), 751-766.

Squicciarini, M. P., and Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics*, *130*(4), 1825-1883.

Stadler, C., Helfat, C., and Verona, G. (2022). Transferring knowledge by transferring individuals: Innovative technology use and organizational performance in multiunit firms. *Organization Science, 33*(1), 253-274.

Stiglitz, J. (1975). The theory of "screening," education, and the distribution of income. *American Economic Review, 65*(3), 283-300.

Toivanen, O., and Väänänen, L. (2016). Education and invention. *Review of Economics and Statistics*, *98*(2), 382-396.

Tokyo Teikoku Daigaku Ichiran (aka *Tokyo Teikoku Daigaku Yoran* in some years) (Tokyo Imperial University Calendar), 1886-1939. Tokyo Imperial University. Available at https://ndlsearch.ndl.go.jp/search?cs=bib&display=panel&from=0&size=20&keyword=東京帝国大学一覧&q-title=東京帝国大学一覧&r-year=1886,1939&f-ht=ndl&f-ht=library

Van Reenen, J. (2022). Innovation and human capital policy. In Goolsbee, A. and Jones B., eds. *Innovation and Public Policy*, National Bureau of Economic Research, Cambridge, MA, pp. 61-83.

Weiss, A. (1995). Human capital vs. singalling explanation of wages. *Journal of Economic Perspectives, 9*(4), 133-154.

Yamaguchi, S., Braguinsky S., Okazaki, T., and Yuki, T. (2023). Resource allocation and growth strategies in a multi-plant firm: Kanegafuchi spinners in the early 20th century. *Strategic Management Journal,* https://doi.org/10.1002/smj.3567

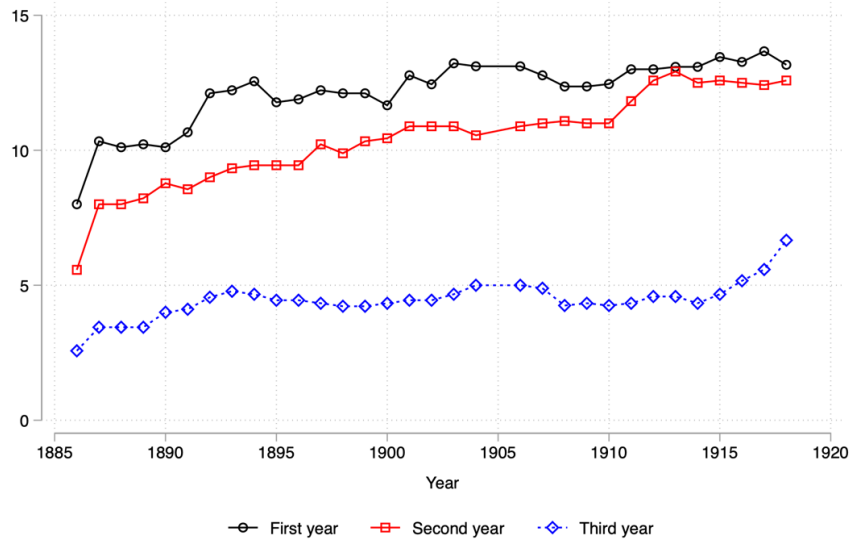
Yamaguchi, S., Braguinsky S., and Nakajima, K. (2021). Science and engineering education and invention in Japan's industrialization. University of Maryland Ed Snider Center Working Paper No. 3, available at https://drive.google.com/file/d/16YrncMlmeq4NZ7KRhNeEWGRAQjwYA7op/view

Zucker, L., and Darby, M. (1996). Star scientists and institutional transformation: Patents of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences, 93,* 12709-12716.

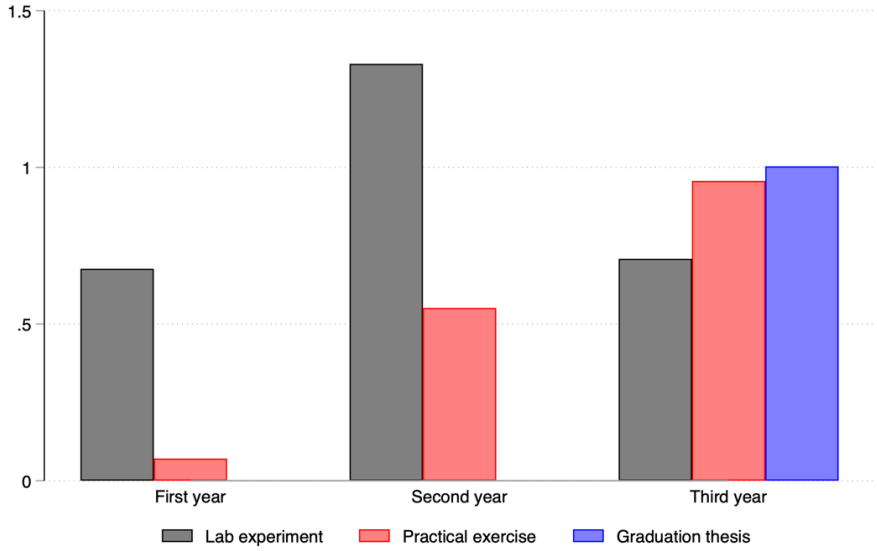
Figures and Tables

Figure 1. Numbers of Courses in the Engineering Department at Tokyo Imperial University

Panel A. Trends in Average Number of Required Courses by Year in the Program

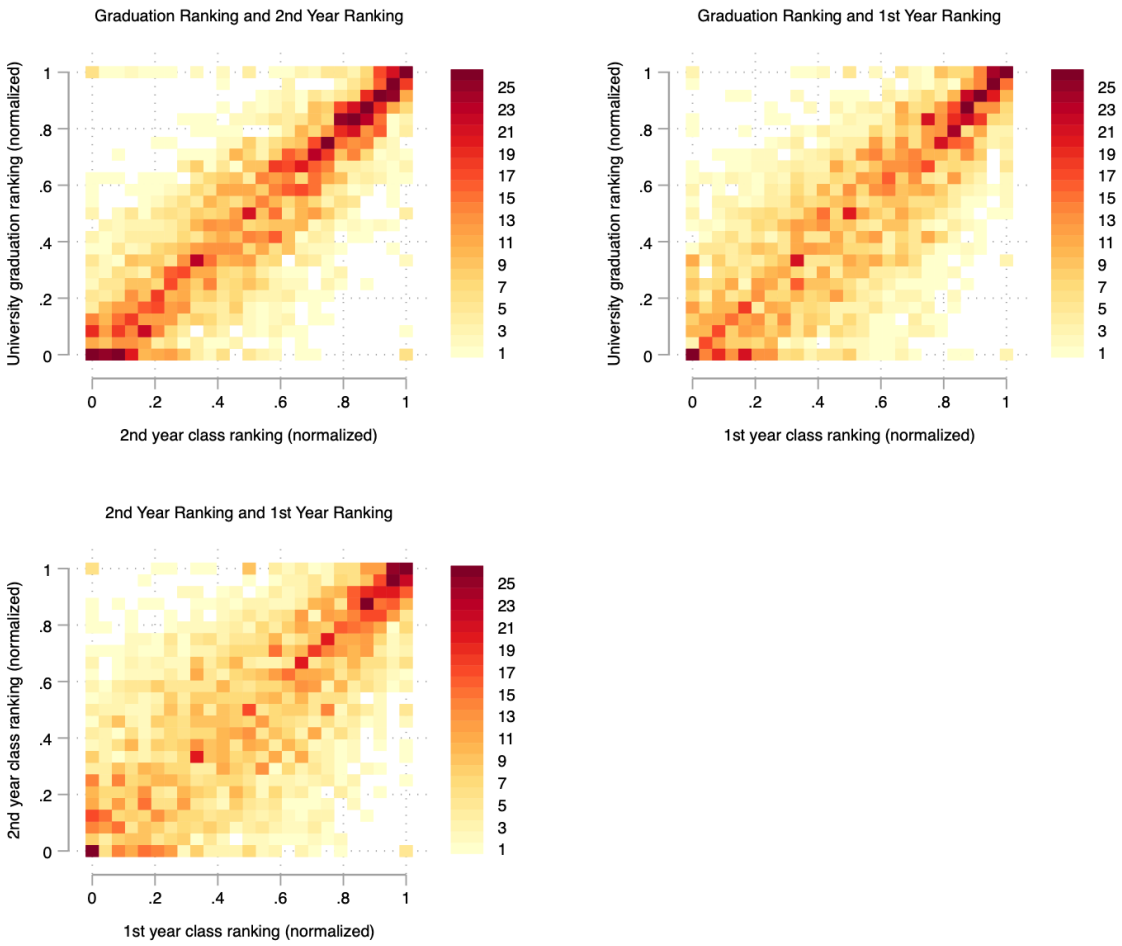


Panel B. Average Number of Practical Courses by Year in the Program



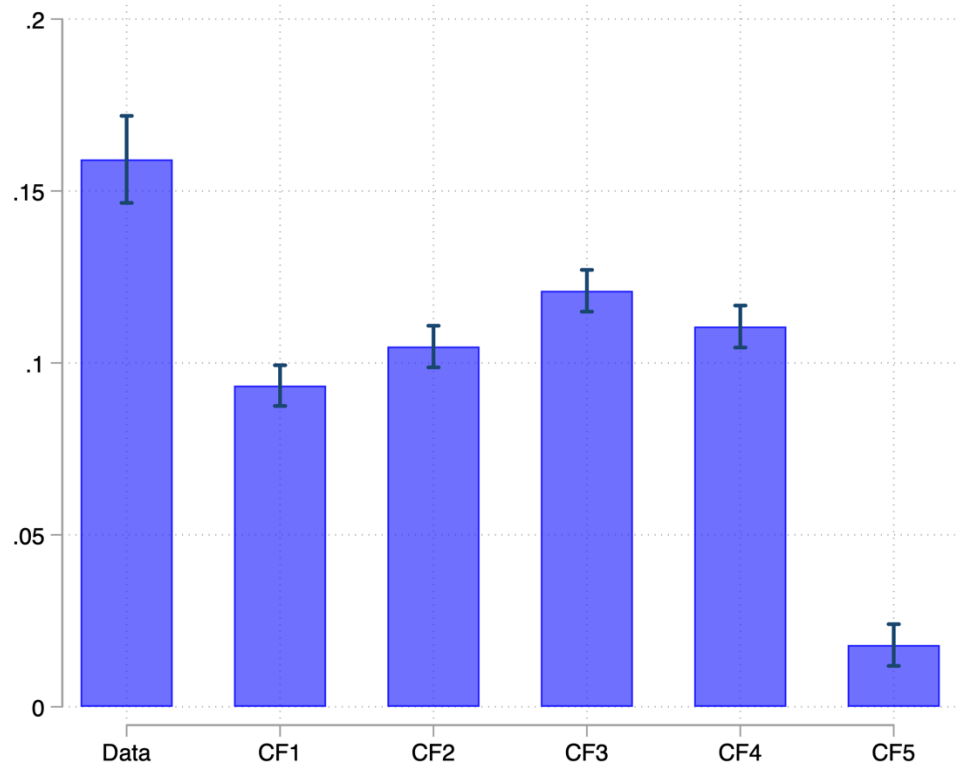
Note: The figures are created based on *Tokyo Imperial University calendar*. In Panel A, the number of the courses required in each year in the program is averaged over divisions in each calendar year. In Panel B, the number of the courses is averaged over both divisions and calendar years.

Figure 2. Joint Distributions of Academic Achievement



Note: The figures represent the distributions of the number of graduates across different pairs of ranking measures. The colors represent the number of graduates. In the cells with the darkest color, the number of graduates in a cell is greater than 25. The top left panel shows the joint distribution of the university graduation rankings and second-year ranking, the top right panel shows the joint distribution of the university graduation rankings and first-year rankings, and the bottom left panel shows the joint distribution of the second-year rankings and first-year rankings.

Figure 3. Counterfactual Predicted Probabilities of Becoming an Inventor



Note: The graph presents the probability of becoming an inventor in the first 20 years after graduation in the data in five counterfactual scenarios (CF1-CF5) estimated from the structural equations. The five counterfactual scenarios are as follows. CF1: if all high school students had the same scores as the students at the bottom of the distribution, CF2: if all first-year university students had the same scores as the students at the bottom of the distribution, CF3: if all second-year university students had the same scores as the students at the bottom of the distribution, CF4: if all third-year university students had the same scores as the students at the bottom of the distribution, and CF5: if all students in all three years at university had the same scores as the students at the bottom of the distribution.

Table 1. Cross-section Summary Statistics

Panel A. Inventors vs. Non-inventors

	(1) Inventors	(2) Non-inventors	(3) Diff.
Cumulative number of patents	3.74	-	-
	(6.23)	(-)	(-)
University graduation ranking (normalized)	0.60	0.49	0.12***
	(0.32)	(0.31)	(7.98)
2nd year class ranking (normalized)	0.60	0.50	0.09***
	(0.31)	(0.30)	(6.62)
1st year class ranking (normalized)	0.60	0.51	0.09***
	(0.31)	(0.30)	(6.302)
High-school graduation ranking (normalized)	0.61	0.51	0.10***
	(0.29)	(0.30)	(7.07)
Observations	584	3213	3797

Panel B. Number of Inventors by Graduation Division

	Division Graduated From							
	Civil Engineering	Mechanical Engineering	Naval Architecture	Electric Engineering	Architecture	Applied Chemistry	Metal and Mining	Others
Non-Inventors	792 (95.54)	665 (81.10)	324 (90.76)	419 (75.50)	236 (92.55)	223 (62.46)	471 (90.40)	83 (85.29)
Inventors	37 (4.46)	155 (18.90)	33 (9.24)	136 (24.50)	19 (7.45)	134 (37.54)	50 (9.60)	19 (18.63)
Total	829	820	357	555	255	357	521	102

Panel C. Number of Inventors by Graduation Cohorts

	Graduation Cohort							
	-1886	1886-1890	1891-1895	1896-1900	1901-1905	1906-1910	1911-1915	1916-1918
Non-Inventors	169 (90.37)	107 (86.99)	140 (86.02)	406 (86.02)	391 (87.28)	828 (84.66)	704 (81.48)	468 (83.13)
Inventors	18 (9.63)	16 (13.01)	22 (13.58)	66 (13.98)	57 (12.72)	150 (15.34)	1609 (18.52)	95 (16.87)
Total	187	123	162	472	448	978	864	563

Note: ***, **, and * in Panel A refer to statistical significance using two-sided *t*-test at the 1%, 5%, and 10% levels, respectively. The numbers in brackets in Panel A are standard errors, while those in Panels B and C are shares in total.

Table 2. Academic Performance and Invention

Panel A. All Inventions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.124***	0.0520**	-0.0132	0.0924***	0.0325	-0.0269
	(0.0217)	(0.0252)	(0.0288)	(0.0295)	(0.0302)	(0.0324)
2nd year class ranking		0.102***	-0.00971		0.0986***	-0.00310
		(0.0278)	(0.0432)		(0.0304)	(0.0460)
University graduation ranking			0.179***			0.170***
			(0.0477)			(0.0485)
High-school graduation ranking				0.0555*	0.0366	0.0259
				(0.0292)	(0.0314)	(0.0317)
Mean of DV	0.160	0.161	0.161	0.158	0.159	0.159
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3379	3350	3346	3240	3213	3210
Adj R-Squared	0.105	0.108	0.113	0.0922	0.0947	0.0988

Panel B. Inventions Related to Graduation Divisions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.109**	0.0245	-0.0129	0.0750***	0.00911	-0.0295
	(0.0202)	(0.0217)	(0.0264)	(0.0262)	(0.0254)	(0.0289)
2nd year class ranking		0.121***	0.0561		0.111***	0.0443
		(0.0230)	(0.0369)		(0.0264)	(0.0392)
University graduation ranking			0.104**			0.111**
			(0.0444)			(0.0427)
High-school graduation ranking				0.0389	0.0156	0.00860
				(0.0251)	(0.0270)	(0.0272)
Mean of DV	0.119	0.120	0.120	0.119	0.119	0.119
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3379	3350	3346	3240	3213	3210
Adj R-Squared	0.118	0.125	0.127	0.0975	0.103	0.105

Note: Estimation method: OLS. The dependent variable is the dummy equal to one if at least one patent in the first 20 years after graduation, zero otherwise. Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2. Academic Achievement and Invention (Continued)

Panel C. Inventions Related to Themes of Graduation Theses: Patent Titles and 75th Percentile Threshold (Only Naval Architecture, Mechanical Engineering, and Electrical Engineering Divisions)

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0695***	0.0130	-0.0499	0.0752**	0.0215	-0.0424
	(0.0232)	(0.0270)	(0.0368)	(0.0308)	(0.0332)	(0.0424)
2nd year class ranking		0.0795**	-0.0275		0.0883***	-0.0181
		(0.0316)	(0.0376)		(0.00298)	(0.0371)
University graduation ranking			0.173***			0.179**
			(0.0501)			(0.0478)
High-school graduation ranking				-0.00234	-0.0232	-0.0313
				(0.0291)	(0.0292)	(0.0292)
Mean of DV	0.0756	0.0765	0.0765	0.0774	0.0783	0.0784
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0346	0.0379	0.0455	0.0167	0.0199	0.0281

Panel D. Global Inventions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0424***	0.0116	0.00246	0.0364***	0.0150	0.0103
	(0.0100)	(0.0100)	(0.0112)	(0.0134)	(0.0122)	(0.0141)
2nd year class ranking		0.0439***	0.0285*		0.0367***	0.0285*
		(0.0116)	(0.0150)		(0.0123)	(0.0167)
University graduation ranking			0.0249*			0.0138
			(0.0142)			(0.0159)
High-school graduation ranking				0.01331	0.00442	0.00355
				(0.00963)	(0.0101)	(0.0101)
Mean of DV	0.0234	0.0236	0.0236	0.0238	0.0240	0.0240
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3379	3350	3346	3240	3213	3210
Adj R-Squared	0.00195	0.00607	0.00624	-0.00945	-0.00756	-0.00790

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Academic Performance and Invention: Early versus Late Inventors

	All		Division-Related		Thesis-Related	
	Early	Late	Early	Late	Early	Late
	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	-0.030	0.004	-0.034	0.005	-0.048	0.005
	(0.027)	(0.021)	(0.023)	(0.018)	(0.040)	(0.018)
2nd year class ranking	-0.024	0.020	0.009	0.035	-0.015	-0.003
	(0.033)	(0.034)	(0.028)	(0.028)	(0.035)	(0.018)
University graduation ranking	0.130***	0.040	0097***	0.013	0.131***	0.048**
	(0.033)	(0.034)	(0.029)	(0.031)	(0.045)	(0.022)
High-school graduation ranking	0.009	0.017	-0.001	0.010	-0.019	-0.013
	(0.023)	(0.020)	(0.022)	(0.018)	(0.024)	(0.016)
Mean of DV	0.091	0.069	0.069	0.050	0.053	0.025
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3210	3210	3210	3210	1506	1506
Adj R-Squared	0.075	0.033	0.072	0.020	0.002	0.052

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. "Early Inventors" are those who applied for their first patent within the median number of years from graduation until the first patent in our sample, while "Late Inventors" are those who applied for their first patent after the median number of years from graduation until the first patent. Please see the main text for details. The sample in columns (5)-(6) is limited to graduates from Naval Architecture, Mechanical Engineering, and Electrical Engineering for whom we have graduation theses data.

Table 4. Placebo Test: Academic Performance and Promotion to Executive

	All executives	Inventors	Non-Inventors	*Zaibatsu* executives	*Zaibatsu* executives-inventors	*Zaibatsu* executives-non-inventors
	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	-0.0514	-0.0112	-0.0401	-0.0226	-0.0131	-0.0124
	(0.0497)	(0.0272)	(0.0423)	(0.0253)	(0.0107)	(0.0223)
2nd year class ranking	-0.00643	-0.0145	0.00805	-0.00597	-0.0117	-0.00415
	(0.0570)	(0.0384)	(0.0563)	(0.0221)	(0.0152)	(0.0151)
University graduation ranking	0.0808	0.0870**	-0.00621	0.0463	0.0265	0.0318
	(0.0562)	(0.0413)	(0.0529)	(0.0317)	(0.0164)	(0.0241)
High-school graduation ranking	-0.0100	-0.0209	0.0108	0.000992	0.00844	-0.00631
	(0.0416)	(0.0277)	(0.0350)	(0.0187)	(0.00536)	(0.0172)
Mean of DV	0.160	0.0541	0.106	0.0283	0.00596	0.0209
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2013	2013	2013	2013	2013	2013
Adj R-Squared	-0.0145	-0.00452	-0.0232	-0.00848	-0.0585	0.0109

Note: Standard errors are clustered at the cohort-division level and cohort-high-school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Academic Performance and Invention, Absorbing Employer Fixed Effects

Panel A. All Inventions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0823***	0.0367	-0.00126	0.0620**	0.00211	-0.0494
	(0.0280)	(0.0316)	(0.0365)	(0.0277)	(0.0296)	(0.0326)
2nd year class ranking		0.0785**	0.0100		0.109***	0.0217
		(0.0323)	(0.0503)		(0.0331)	(0.0495)
University graduation ranking			0.114**			0.150***
			(0.0545)			(0.0519)
High-school graduation ranking	0.0265	0.0135	0.00876	0.00640	-0.0129	-0.0207
	(0.0332)	(0.0349)	(0.0349)	(0.0312)	(0.0331)	(0.0330)
Mean of DV	0.155	0.156	0.157	0.147	0.148	0.148
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
First employer FEs	Yes	Yes	Yes	No	No	No
Mode employer FEs	No	No	No	Yes	Yes	Yes
Observations	2858	2829	2824	2705	2682	2679
Adj R-Squared	0.142	0.145	0.146	0.182	0.183	0.186

Panel B. Inventions Related to Graduation Divisions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0609**	0.00837	-0.0196	0.0380	-0.0164	-0.0437
	(0.0248)	(0.0262)	(0.0309)	(0.0250)	(0.0259)	(0.0303)
2nd year class ranking		0.0923***	0.0417		0.0984***	0.0526
		(0.0282)	(0.0443)		(0.0295)	(0.0432)
University graduation ranking			0.0841*			0.0794
			(0.0486)			(0.0497)
High-school graduation ranking	0.0126	-0.00542	-0.00901	0.00828	-0.0108	-0.0147
	(0.0296)	(0.0319)	(0.0319)	(0.0288)	(0.0307)	(0.0308)
Mean of DV	0.117	0.117	0.118	0.114	0.114	0.115
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
First employer FEs	Yes	Yes	Yes	No	No	No
Mode employer FEs	No	No	No	Yes	Yes	Yes
Observations	2858	2829	2824	2705	2682	2679
Adj R-Squared	0.144	0.148	0.149	0.176	0.178	0.179

Note: The dependent variable is the dummy equal to one if at least one patent in the first 20 years after graduation, zero otherwise. Columns (1)-(3) include the first employer fixed effects, while columns (4)-(6) include the mode employer fixed effects Standard errors are in parentheses and clustered at the cohort-division level and cohort-high-school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Academic Performance and Invention, Absorbing Employer Fixed Effects (Continued)

Panel C. Inventions Related to Themes of Graduation Theses: Patent Titles and 75th Percentile Threshold (Only Naval Architecture, Mechanical Engineering, and Electrical Engineering Divisions)

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0438	-0.00786	-0.0492	0.0322	-0.00451	-0.0663
	(0.0292)	(0.0325)	(0.0395)	(0.0355)	(0.0378)	(0.0452)
2nd year class ranking		0.0887**	0.0203		0.0602	-0.0446
		(0.0344)	(0.0439)		(0.0379)	(0.0445)
University graduation ranking			0.115**			0.182***
			(0.0453)			(0.0525)
High-school graduation ranking	-0.0177	-0.0370	-0.0400	0.00860	-0.00223	-0.0105
	(0.0291)	(0.0281)	(0.0282)	(0.0293)	(0.0303)	(0.0299)
Mean of DV	0.0760	0.0769	0.0769	0.0798	0.0807	0.0808
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
First employer FEs	Yes	Yes	Yes	No	No	No
Mode employer FEs	No	No	No	Yes	Yes	Yes
Observations	1315	1300	1300	1241	1227	1225
Adj R-Squared	0.0525	0.0543	0.0571	0.118	0.115	0.123

Panel D. Global Inventions

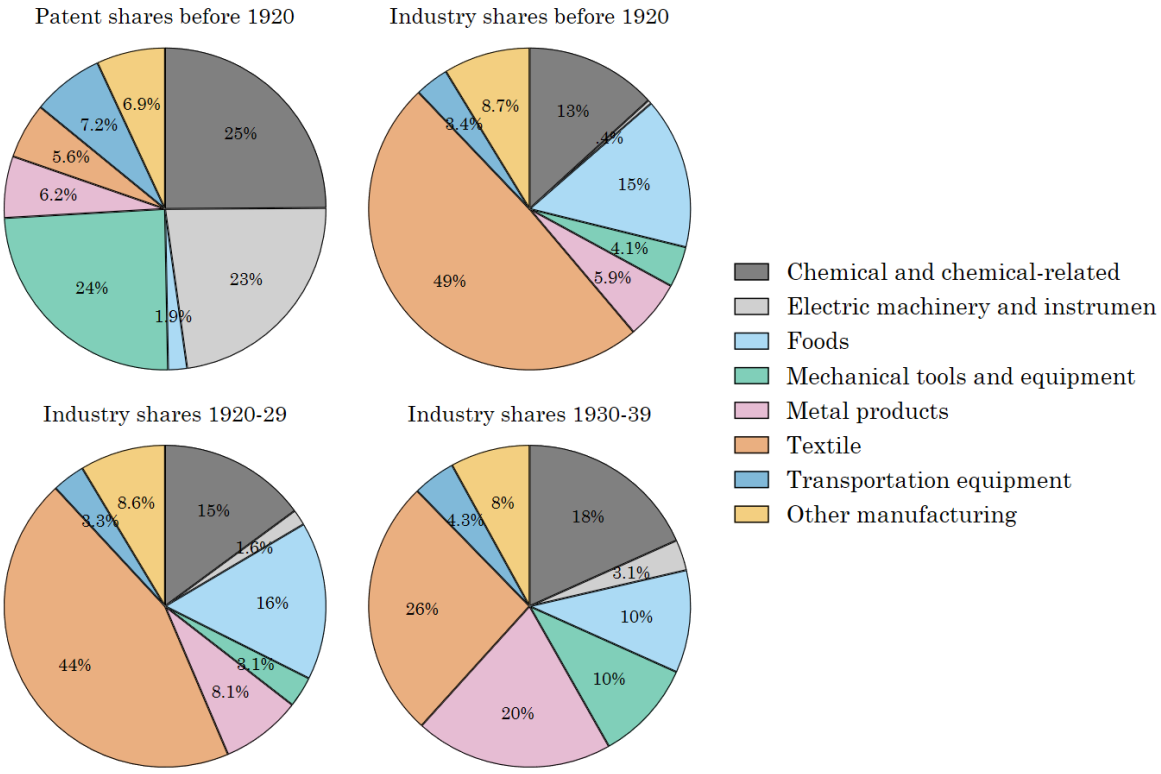
	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0310	0.00224	-0.0155	0.0174	-0.0108	-0.0265
	(0.0248)	(0.0275)	(0.0294)	(0.0216)	(0.0263)	(0.0295)
2nd year class ranking		0.0517**	0.0223		0.0470	0.0201
		(0.0256)	(0.0338)		(0.0321)	(0.0397)
University graduation ranking			0.0496			0.0467
			(0.0308)			(0.0392)
High-school graduation ranking	0.00953	-0.00352	-0.00481	0.0154	0.00604	0.00393
	(0.0172)	(0.0186)	(0.0187)	(0.0180)	(0.0192)	(0.0194)
Mean of DV	0.0357	0.0362	0.0362	0.0363	0.0367	0.0367
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
First employer FEs	Yes	Yes	Yes	No	No	No
Mode employer FEs	No	No	No	Yes	Yes	Yes
Observations	1315	1300	1300	1241	1227	1225
Adj R-Squared	0.00502	0.00542	0.00575	0.00413	0.00339	0.00361

Note: The dependent variable is the dummy equal to one if at least one patent in the first 20 years after graduation, zero otherwise. Columns (1)-(3) include the first employer fixed effects, while columns (4)-(6) include the mode employer fixed effects Standard errors are in parentheses and clustered at the cohort-division level and cohort-high-school level. *** p<0.01, ** p<0.05, * p<0.1.

Appendix to: Who Becomes an Inventor? The (Important) Role of Knowledge Absorption from Higher Technical Education

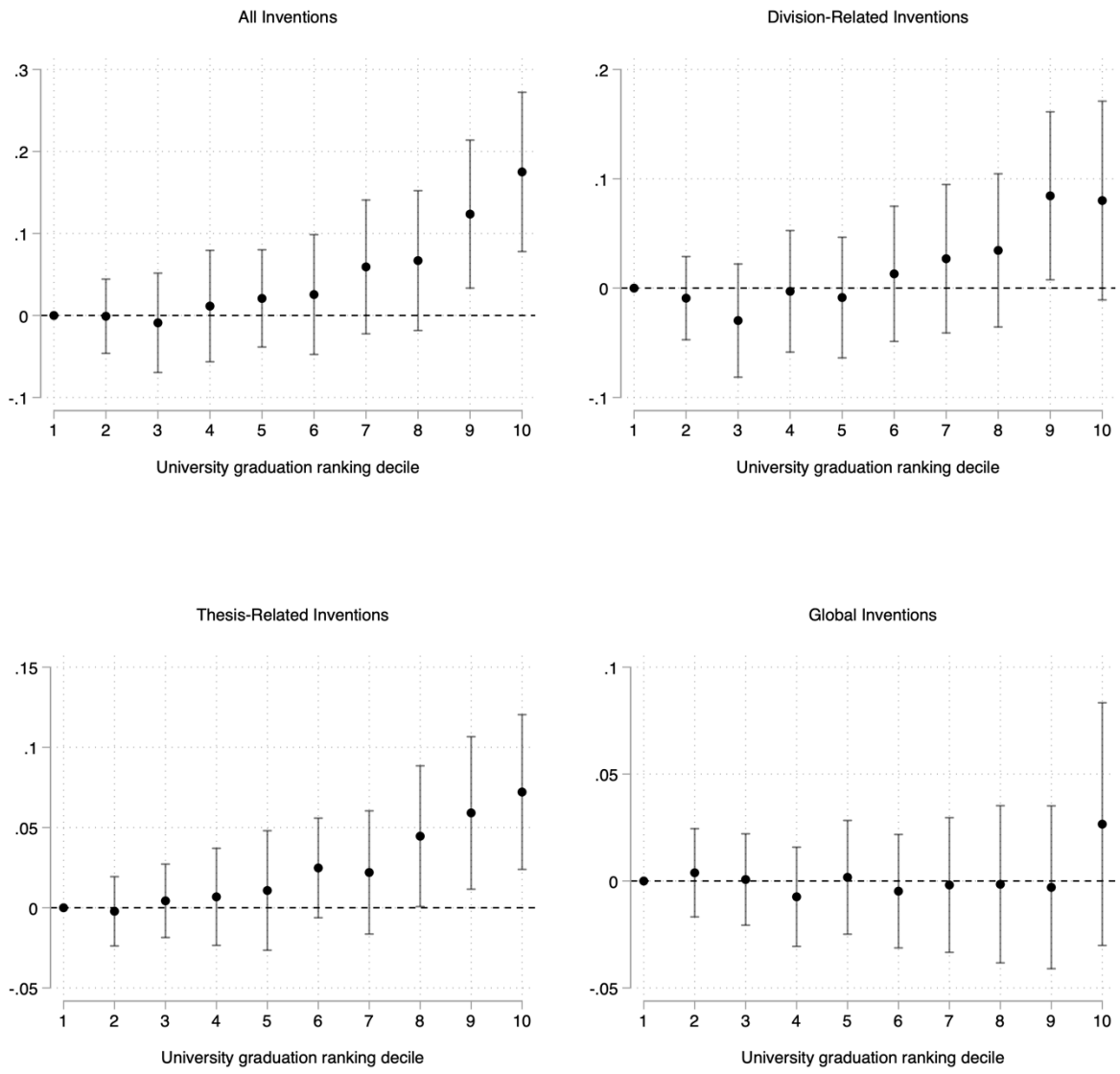
A. Figures and Tables Cited in the Main Text

Figure A1. Shares of Patent Classes and the Dynamics of Corresponding Industries



Note: "Patent shares" are the shares of patents by S&E inventors in corresponding patent classes in the total number of patents by S&E inventors with application dates between 1885-1919. "Industry shares" are the shares of corresponding industries in the total manufacturing output in Japan. Source: Yamaguchi et al. (2021).

Figure A2. Estimated Probability of Becoming an Inventor by Graduation Rankings Deciles



Note: Estimated coefficients on university graduation rankings by decile, controlling for high school graduation rankings as well as first- and second-year rankings at the university, and absorbing university division-cohort, high school-cohort, and birthplace prefecture fixed effects. Standard errors are clustered at the cohort-division level and cohort-high-school level. Bars represent 95-percent confidence intervals.

Table A1. Course Curriculum for the Mechanical Engineering Division in 1918

First Year	Hours per Week		
	1st Term	2nd term	3rd Term
Required Subjects			
Mathematics	3	3	-
Dynamics	1	1	1
Applied Mechanics	3	3	2
Heat and Heat Engines	3	2	3
Mechanism	1	1	1
Hydraulics	-	2	2
Mechanical Technology	1	2	2
Machine Design and Materials	2	2	2
Electrical Engineering	2	2	2
Exercises in Applied Mechanics	3	3	3
Physical Laboratory	3	3	3
Workshop Practice	3	3	3
Design and Drawing	14	10	15

Second Year	Hours per Week		
	1st Term	2nd term	3rd Term
Required Subjects			
Thermodynamics and Heat Engines	3	3	3
Dynamics of Machinery	1	1	1
Hydraulic Machinery	1.5	2	2
Electrical Engineering	2	2	2
Testing of Electrical Machines	1	1	1
Metallic Materials and Metallurgical Technology	3	2	2
Building Construction	2	2	-
Machine Design	2	1	-
Industrial Economy	1.5	1.5	1.5
Experimental Engineering	1	1	1
Mechanical Engineering Laboratory	3	3	3
Design and Drawing	12	12	15
Electrical Engineering Laboratory	-	3	3
Elective Subjects (one of the subjects to be chosen)			
Machine Tools	2	1	-
Machine Engines	2	2	2
Spinning and Weaving	1.5	1.5	1.5
Railway Vehicles	2	2	2

Note: The table is created based on *Tokyo Imperial University calendar*. If the hour per week is missing in the table (-), then there was no course in that term. If the hour per week is not specified, the course existed, but there was no specific hour per week in that course. In the second year, students must take one elective subject among the list of the courses, while in the third year, students must take four elective subjects among the list of the courses.

Table A1. Course Curriculum for the Mechanical Engineering Division in 1918 (Continued)

Third Year	Hours per Week		
	1st Term	2nd term	3rd Term
Required Subjects			
Mechanical Engineering Laboratory	-	6	6
Practical Exercises	Not specified	-	-
Graduation Design or Essay	-	Not specified	Not specified
Elective Subjects (four of the subjects to be chosen)			
Heating and Ventilation	-	1.5	1.5
Refrigeration and Refrigeration Machinery	-	1.5	1.5
Lifting Machinery	-	1	1
Compressed Air Machinery	-	1	1
Workshop Management	-	1	1
Design of Workshop and Power Plant	-	2	2
Electrical Engineering	-	1.5	1.5
Special Lectures	-	2	2

Note: The table is created based on *Tokyo Imperial University calendar*. If the hour per week is missing in the table (-), then there was no course in that term. If the hour per week is not specified, the course existed, but there was no specific hour per week in that course. In the second year, students must take one elective subject among the list of the courses, while in the third year, students must take four elective subjects among the list of the courses.

Table A2. Academic Performance and IHS-Transformed Cumulative Number of Patents

Panel A. IHS-Transformed Number of Patents

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.230***	0.0875*	-0.0143	0.172***	0.0570	-0.0333
	(0.0444)	(0.0488)	(0.0605)	(0.0619)	(0.0620)	(0.0715)
2nd year class ranking		0.204***	0.0293		0.191***	0.0359
		(0.0549)	(0.0736)		(0.0587)	(0.0806)
University graduation ranking			0.2816***			0.260***
			(0.0830)			(0.0861)
High-school graduation ranking				0.0988*	0.0613	0.0449
				(0.0590)	(0.0619)	(0.0619)
Mean of DV	0.255	0.257	0.257	0.255	0.256	0.256
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3379	3350	3346	3240	3213	3210
Adj R-Squared	0.0905	0.0949	0.0982	0.0714	0.0747	0.0774

Panel B. IHS-Transformed Number of Patents Related to Graduation Divisions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.207***	0.0610	-0.0112	0.150***	0.0373	-0.0323
	(0.0403)	(0.0417)	(0.0514)	(0.0544)	(0.0514)	(0.0594)
2nd year class ranking		0.210***	0.0856		0.190***	0.0695
		(0.0465)	(0.0617)		(0.0487)	(0.0661)
University graduation ranking			0.199***			0.201***
			(0.0713)			(0.0692)
High-school graduation ranking				0.0816	0.0414	0.0287
				(0.0521)	(0.0539)	(0.0537)
Mean of DV	0.189	0.191	0.191	0.190	0.191	0.191
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3379	3350	3346	3240	3213	3210
Adj R-Squared	0.104	0.110	0.112	0.0854	0.0900	0.0920

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2. Academic Performance and IHS-Transformed Cumulative Number of Patents (Continued)

Panel C. IHS-Transformed Number of Patents Related to Themes of Graduation Theses: 75th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.125***	0.0403	-0.0368	0.130***	0.0509	-0.0303
	(0.0389)	(0.0411)	(0.0511)	(0.0495)	(0.0510)	(0.0594)
2nd year class ranking		0.119***	-0.0122		0.130***	-0.00567
		(0.0430)	(0.0500)		(0.0392)	(0.0495)
University graduation ranking			0.212***			0.228***
			(0.0666)			(0.0630)
High-school graduation ranking				-0.000218	-0.0301	-0.0405
				(0.0407)	(0.0415)	(0.0419)
Mean of DV	0.100	0.101	0.102	0.103	0.104	0.104
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0289	0.0332	0.0382	-0.00915	-0.00538	0.000419

Panel D. IHS-Transformed Number of Global Patents

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0684***	0.0214	0.00353	0.0607**	0.0284	0.0169
	(0.0183)	(0.0163)	(0.0194)	(0.0243)	(0.0199)	(0.0236)
2nd year class ranking		0.0667***	0.0368		0.0554***	0.0355
		(0.0185)	(0.0256)		(0.0210)	(0.0301)
University graduation ranking			0.0487*			0.0334
			(0.0279)			(0.0308)
High-school graduation ranking				0.0219	0.00852	0.00640
				(0.0161)	(0.0171)	(0.0172)
Mean of DV	0.0375	0.0379	0.0379	0.0383	0.0386	0.0387
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3379	3350	3346	3240	3213	3210
Adj R-Squared	-0.00264	0.000492	0.000795	-0.0118	-0.0107	-0.0108

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Academic Performance and Invention:
Naval Architecture, Mechanical Engineering, and Electrical Engineering Divisions

Panel A. Invention in General

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.158***	0.0503	-0.0474	0.141***	0.0448	-0.0452
	(0.0364)	(0.0383)	(0.0514)	(0.0508)	(0.0459)	(0.0555)
2nd year class ranking		0.151***	-0.0158		0.152**	0.00193
		(0.0418)	(0.0663)		(0.0515)	(0.07863)
University graduation ranking			0.269***			0.253***
			(0.0823)			(0.0850)
High-school graduation ranking				0.0601	0.0310	0.0194
				(0.0466)	(0.0502)	(0.0506)
Mean of DV	0.194	0.195	0.195	0.192	0.194	0.194
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0347	0.0402	0.0485	0.00131	0.00720	0.0147

Panel B. Logged Invention Related to Graduation Divisions

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.130***	0.00216	-0.0681	0.0944**	-0.00404	-0.0742
	(0.0335)	(0.0353)	(0.0518)	(0.0465)	(0.0444)	(0.0556)
2nd year class ranking		0.179***	0.0588		0.157***	0.0402
		(0.0371)	(0.0630)		(0.0463)	(0.0707)
University graduation ranking			0.193**			0.197**
			(0.0868)			(0.0811)
High-school graduation ranking				0.0565	0.0253	0.0163
				(0.0440)	(0.0460)	(0.0467)
Mean of DV	0.145	0.147	0.147	0.145	0.146	0.146
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0418	0.0538	0.0589	0.00264	0.0125	0.0179

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Academic Performance and Thesis-Related Invention

Panel A. Invention Related to Themes of Graduation Theses: Patent Titles and 75th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0695***	0.0130	-0.0499	0.0752**	0.0215	-0.0424
	(0.0232)	(0.0270)	(0.0368)	(0.0308)	(0.0332)	(0.0424)
2nd year class ranking		0.0795**	-0.0276		0.0883***	-0.0181
		(0.0316)	(0.0376)		(0.0298)	(0.0371)
University graduation ranking			0.173***			0.179***
			(0.0501)			(0.0478)
High-school graduation ranking				-0.00234	-0.0232	-0.0313
				(0.0291)	(0.0292)	(0.0292)
Mean of DV	0.0756	0.0765	0.0765	0.0774	0.0783	0.0784
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0346	0.0379	0.0455	0.0167	0.0199	0.0281

Panel B. Invention Related to Themes of Graduation Theses: Patent Classes and 75th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0675***	0.0140	-0.0641*	0.0698**	0.0169	-0.0542
	(0.0226)	(0.0299)	(0.0366)	(0.0312)	(0.0355)	(0.0404)
2nd year class ranking		0.0768**	-0.0559		0.0826**	-0.0358
		(0.0344)	(0.0452)		(0.0319)	(0.0494)
University graduation ranking			0.214***			0.199***
			(0.0521)			(0.0544)
High-school graduation ranking				0.0177	0.00326	-0.00582
				(0.0337)	(0.0345)	(0.0337)
Mean of DV	0.0730	0.0733	0.0733	0.0734	0.0737	0.0737
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	-0.000914	0.00319	0.0157	-0.0144	-0.00870	0.00242

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Academic Performance and Thesis-Related Invention (Continued)

Panel C. Invention Related to Themes of Graduation Theses: Patent Categories and 75th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0779***	0.0208	-0.0564	0.0630**	0.00861	-0.0661
	(0.0233)	(0.0284)	(0.0383)	(0.0302)	(0.0338)	(0.0442)
2nd year class ranking		0.0825**	-0.0487		0.0853***	-0.0387
		(0.0331)	(0.0397)		(0.0309)	(0.0425)
University graduation ranking			0.212***			0.209***
			(0.0525)			(0.0548)
High-school graduation ranking				0.0317	0.0171	0.00770
				(0.0325)	(0.0330)	(0.0323)
Mean of DV	0.0794	0.0797	0.0797	0.0800	0.0803	0.0803
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0271	0.0336	0.0449	0.000694	0.00912	0.0203

Panel D. Invention Related to Themes of Graduation Theses: Patent Sub-Categories and 75th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0798***	0.0168	-0.0501	0.0684**	0.0173	-0.0434
	(0.0262)	(0.0283)	(0.0318)	(0.0325)	(0.0319)	(0.0350)
2nd year class ranking		0.0885***	-0.0253		0.0825**	-0.0186
		(0.0320)	(0.0416)		(0.0324)	(0.0447)
University graduation ranking			0.184***			0.170***
			(0.0464)			(0.0469)
High-school graduation ranking				0.0448	0.0260	0.0183
				(0.0299	(0.0317)	(0.0318)
Mean of DV	0.0775	0.0784	0.0785	0.0793	0.0803	0.0803
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0212	0.0258	0.0342	-0.0138	-0.0114	-0.00423

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Academic Performance and Thesis-Related Invention (Continued)

Panel E. Invention Related to Themes of Graduation Theses: Patent Titles and 90th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0451**	0.00942	-0.0226	0.0381*	0.00412	-0.0237
	(0.0193)	(0.0193)	(0.0242)	(0.0229)	(0. 0229)	(0.0258)
2nd year class ranking		0.0505**	-0.00340		0.0553**	0.00950
		(0.0211)	(0.0263)		(0.0223)	(0.0288)
University graduation ranking			0.0872***			0.0773***
			(0.0307)			(0.0289)
High-school graduation ranking				0.0183	0.00542	0.00198
				(0.0183)	(0.0190)	(0.0188)
Mean of DV	0.0394	0.0398	0.0399	0.0400	0.0405	0.0405
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	0.0218	0.0250	0.0281	0.0231	0.0253	0.0274

Panel F. Invention Related to Themes of Graduation Theses: Patent Classes and 90th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0317*	0.0135	-0.0349	0.0340	0.0112	-0.0276
	(0.0180)	(0.0211)	(0.0257)	(0.0237)	(0.0261)	(0.0303)
2nd year class ranking		0.0273	-0.0538*		0.0325	-0.0313
		(0.0225)	(0.0324)		(0.0215)	(0.0338)
University graduation ranking			0.132***			0.108***
			(0.0353)			(0.0360)
High-school graduation ranking				-0.00569	-0.00789	-0.0127
				(0.0213)	(0.0212)	(0.0209)
Mean of DV	0.0362	0.0360	0.0360	0.0367	0.0365	0.0365
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	-0.00610	-0.00363	0.00531	0.00419	0.00827	0.0140

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Academic Performance and Thesis-Related Invention (Continued)

Panel G. Invention Related to Themes of Graduation Theses: Patent Categories and 90th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0205	-0.000378	-0.0274	0.0151	-0.0114	-0.0356
	(0.0164)	(0.0169)	(0.0218)	(0.0211)	(0.0200)	(0.0235)
2nd year class ranking		0.0315	-0.0139		0.0412	0.00132
		(0.0221)	(0.0279)		(0.0251)	(0.0288)
University graduation ranking			0.0735**			0.0673**
			(0.0295)			(0.0274)
High-school graduation ranking				-0.00224	-0.00823	-0.0112
				(0.0186)	(0.0202)	(0.0205)
Mean of DV	0.0317	0.0315	0.0315	0.0321	0.0319	0.0319
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	-0.00231	0.00129	0.00396	-0.00671	-0.00390	-0.001982

Panel H. Invention Related to Themes of Graduation Theses: Patent Sub-Categories and 90th Percentile Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.0359**	0.00132	-0.0394*	0.0337	0.00183	-0.0385*
	(0.0180)	(0.0195)	(0.0221)	(0.0218)	(0.0212)	(0.0221)
2nd year class ranking		0.0496**	-0.0195		0.0535**	-0.0132
		(0.0214)	(0.0308)		(0.0215)	(0.0309)
University graduation ranking			0.112***			0.112***
			(0.0361)			(0.0358)
High-school graduation ranking				0.00627	-0.00674	-0.0118
				(0.0181)	(0.0189)	(0.0191)
Mean of DV	0.0381	0.0386	0.0386	0.0393	0.0398	0.0398
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	No	No	No	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1575	1556	1555	1525	1507	1506
Adj R-Squared	-0.00327	-0.00502	0.00531	-0.0274	-0.0262	-0.0205

Note: Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5. Academic Performance and Invention, Estimated from the Structural Equation

	(1)	(2)	(3)	(4)
High school score	0.553***	0.231***	0.105***	0.00466
	(0.0181)	(0.0167)	(0.0198)	(0.0102)
1st year score		0.597***	0.109***	0.0146
		(0.0148)	(0.0303)	(0.00915)
2nd year score			0.567***	0.0135
			(0.0333)	(0.0112)
3rd year score				0.0338***
				(0.00887)
Mean of DV	0.0871	0.0697	-0.0401	0.159
Division X cohort FEs	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes
Birthplace FEs	No	No	No	Yes
Observations	3250	3223	3220	3210
Adj R-Squared	0.278	0.539	0.405	0.103

Note: The Table shows results from estimating equations (7)-(10) in the main text, using inferred high-school, first-, second-, and third-year scores. The first three columns present the results of regressing first-year scores on high-school scores; second-year scores on high-school and first-year scores; and third-year scores on high-school, first-year and second-year scores, respectively. The final column shows the estimated "inventor production function." Standard errors are clustered at the cohort-division and cohort-high-school levels. *** p<0.01, ** p<0.05, * p<0.1.

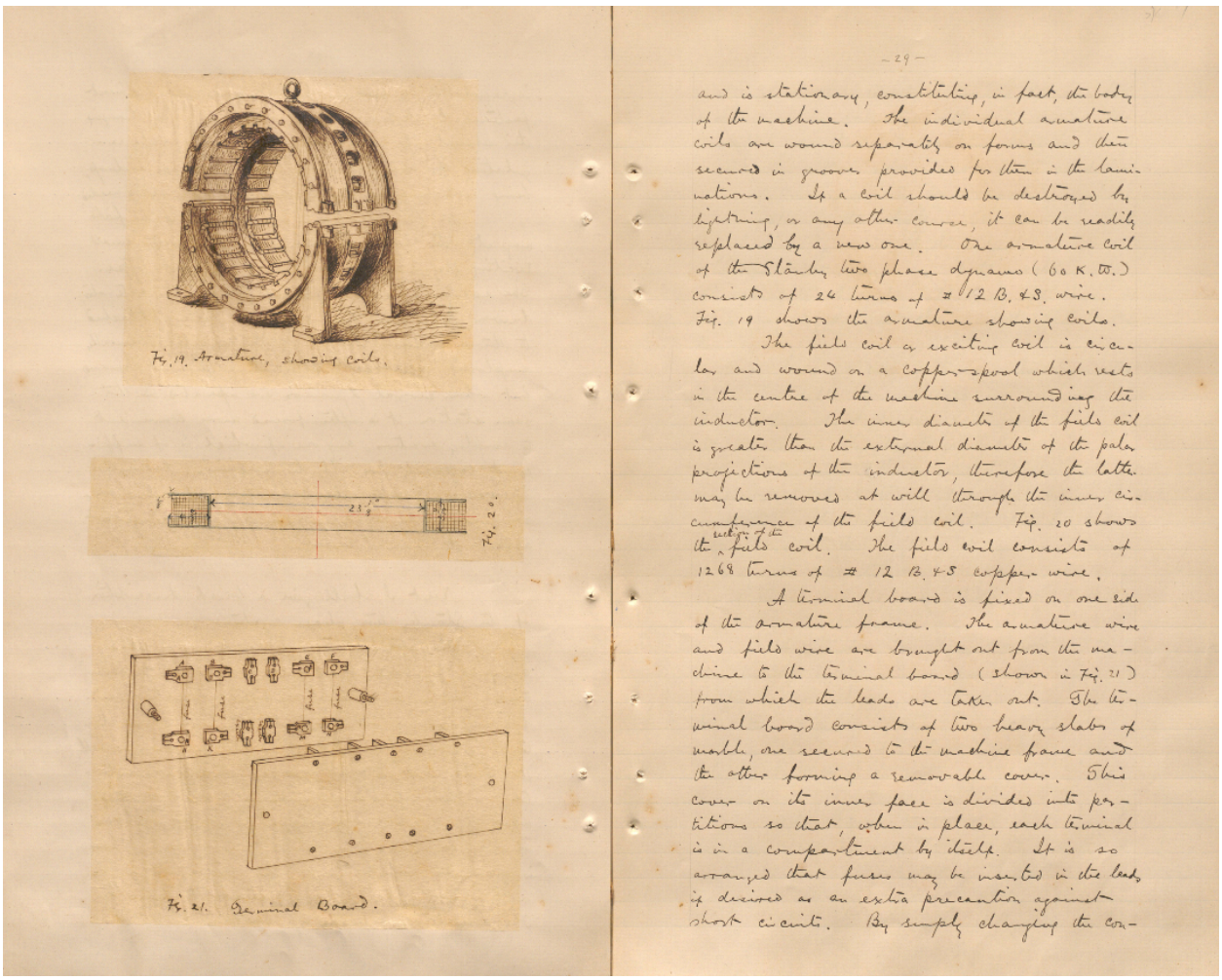
B. Illustrative Examples

This appendix describes some illustrative examples of excursion reports and graduation theses.

B.1. Keijiro Kishi

Keijiro Kishi (岸敬二郎) graduates from electrical engineering division in 1895 and ranked at the second in the same division at the time of graduation. After he graduated from Tokyo Imperial University, he worked in Shibaura Seisakujo (today's Toshiba) for his whole life, raising to a chief engineer and an executive director. He also had had 16 patents altogether.

Figure B1. A Sample Page from Kishi's Report on Practical Work



Source: U Tokyo Repository, originally coming from Engineering Building 2 Library at University of Tokyo, "Report on Practical Work" by Keijiro Kishi

Kishi also went to and worked in Shibaura Seisakujo for practical exercise in his third year when he was a university student. Figure B1 shows a sample page from his report on practical work. The report is basically filled with the detailed description of machines installed in Shibaura Seisakujo (he observed during practical work) and has more than 100 pages in total. In his report, he also wrote the following preface:

"From the beginning of July of the last year, I have gone on a journey to Ashio, Nikko, and the Kansai district for the practical work of my third year session, and returned to Tokyo at Nov. 28 of the same year. From the 2nd of December up to the present time, I have engaged in a continuation of my practical work in the Shibaura Engineering Works [Toshiba]*, Tokyo.*

Teacher, I beg your pardon, I can not write a minute report on my practical work, because I am now very busy in my work in the Shibaura Engineering Works — I am in a labour of above 10 hours every day. — June 23, 1895"

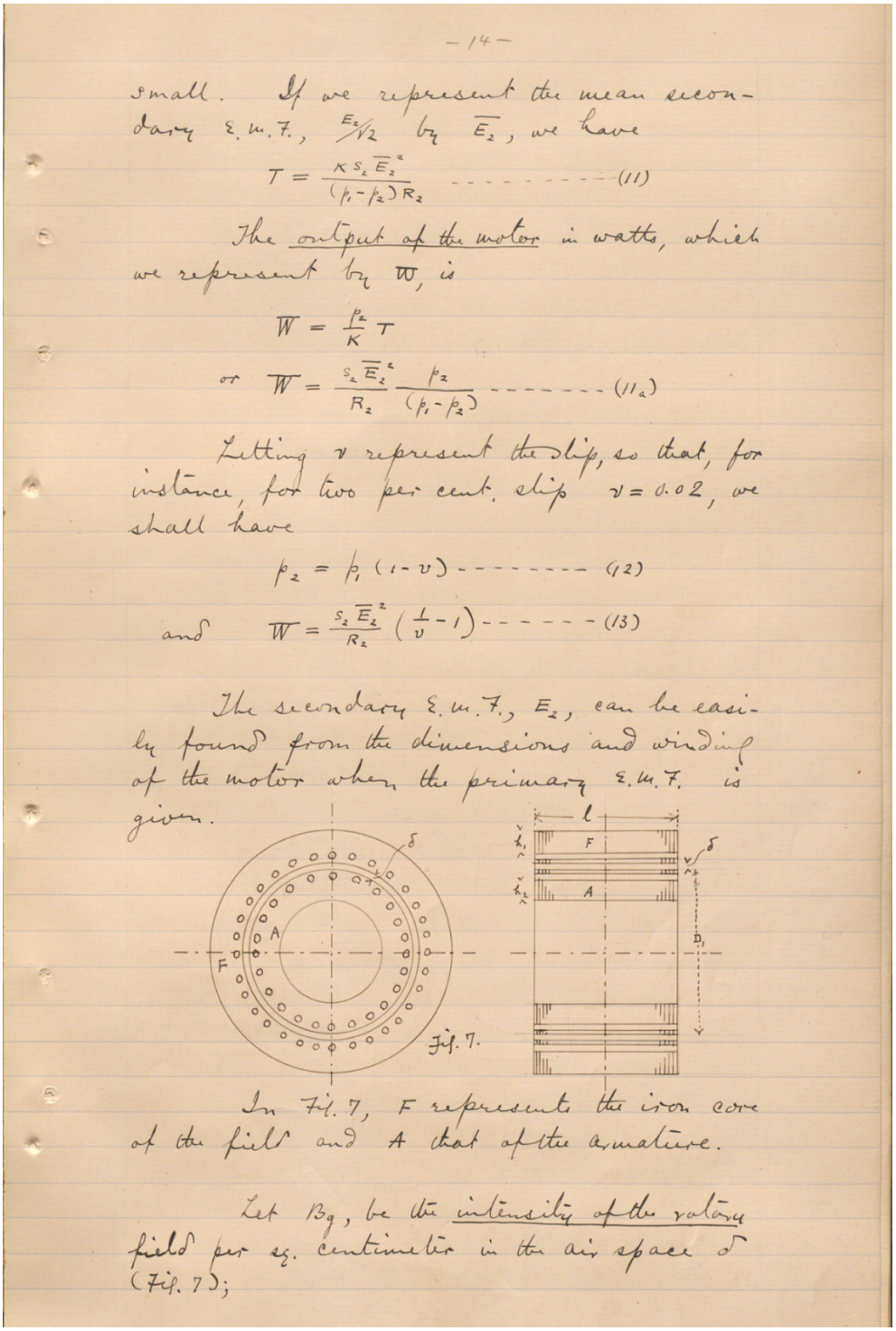
From the above description, there are two observations. First, Kishi went to multiple organizations for practical exercise at different points in time, including Shibaura Seisakujo where he eventually worked after graduation. Second, it seems like he spent many hours working for practical exercise in his third year, and practical exercise was, at least in his case, not a supplementary but core activity in the university education curriculum.

Figure B2 shows a sample page from Kishi's graduation thesis. Compared with the excursion report, he not just provided the description of a machine, but also applied scientific knowledge using mathematical and physical theories and advanced the method that had been used in practice. According to his preface:

"During the last 7 months of my third year, I have engaged in practical work in the Shibaura Engineering Works [Toshiba]*, Tokyo. Among many other designs, I have chosen the design of a 25 HP two phase motor as the thesis of my graduation essay. The motor has been designed for the Ade Copper Mine, Kaga, to work a pump. The motor is being now constructed in the same works, and it will be perfectly built up in the next month. — June 23, 1895"*

First of all, it seems like Kishi's theme of graduation theses was inspired by his work, observations, and experiences in practical work described above. In fact, in his report on practical work, he described a 5 HP two phase motor used in Shibaura Seisakujo, and therefore, his 25 HP two phase motor is an upgraded version of it. Second, his work had already started to be installed in practice when he wrote the graduation theses. This means that graduation theses in higher technical education during that time were high quality and close to frontier technologies even used in the process of industrialization.

Figure B2. A Sample Page from Kishi's Graduation Thesis



— 14 —

small. If we represent the mean secondary E.M.F., $E_2/\sqrt{2}$ by $\overline{E}_2$, we have

$$T = \frac{K s_2 \overline{E}_2^2}{(p_1 - p_2) R_2} \qquad \text{---------- (11)}$$

The output of the motor in watts, which we represent by W, is

$$W = \frac{p_2}{K} T$$

or $$W = \frac{s_2 \overline{E}_2^2}{R_2} \frac{p_2}{(p_1 - p_2)} \qquad \text{-------- (11a)}$$

Letting v represent the slip, so that, for instance, for two per cent. slip $v = 0.02$, we shall have

$$p_2 = p_1 (1 - v) \qquad \text{--------- (12)}$$

and $$W = \frac{s_2 \overline{E}_2^2}{R_2} \left(\frac{1}{v} - 1\right) \qquad \text{------ (13)}$$

The secondary E.M.F., E_2, can be easily found from the dimensions and winding of the motor when the primary E.M.F. is given.

Fig. 7.

In Fig. 7, F represents the iron core of the field and A that of the armature.

Let B_g, be the intensity of the rotary field per sq. centimeter in the air space δ (Fig. 7);

Source: U Tokyo Repository, originally coming from Engineering Building 2 Library at University of Tokyo, "The Calculations of Multi Phase Motors, with a Design of a 25HP Two Phase Motor" by Keijiro Kishi

C. Text Analysis Details

In this appendix, we describe the details on text similarity analysis in Section IV.

C.1. Bidirectional Encoder Representations from Transformer (BERT)

The Bidirectional Encoder Representations from Transformer, or BERT, is a natural language processing (NPL) model introduced in 2018 by Google. BERT is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers (Devlin et al., 2019). The BERT model is used to interpret texts and translate them into word embeddings. Embeddings are vector representations of terms derived from co-occurrence of words in texts. The word embedding models, such as GloVe and Word2vec, can produce a vector representation for any given term capturing semantic relationships between terms, but these models work without context to generate a single representation for each word in the sequence. For example, the word "bat" would be represented the same way whether referring to a piece of sporting gear or a night-flying animal. The context-dependent embedding models, such as BERT, ELMo, and GPT, take the surrounding context into account to generate the embeddings of each word. Among them, BERT is especially designed for tasks that require the understanding of the meanings and context of words. GPT is a unidirectional model, and it only processes language only from left to right when interpreting the context. ELMo is, on the other hand, a bidirectional model, and it interprets the context both left-to-right and right-to-left directions, but it considers the left-to-right and right-to-left paths independently instead of as a single unified view of the entire context. The Transformer architecture enables BERT to process words in relation to all the other words in a sentence at once rather than individually, using attention mechanism to gather information about the relevant context of the word, and to encode the context in a rich vector.

C.2. Sentence-BERT (SBERT)

While BERT encodes the meanings of words taking the surrounding context into account, one of the disadvantages of the BERT network structure is that independent sentence embeddings are not computed, which makes it difficult to derive integrated sentence vector representations. Therefore, it does not perform well when it comes to comparing similarities of entire sentences (Filippi-Mazzola et al., 2023). We employ the sentence-BERT, or SBERT, to derive semantic textual similarity (Reimers and Gurevych, 2020).[24] The main difference with the regular BERT is that it encodes the semantic meaning of whole sentences instead of individual words. The *sentence-transformers* package in python gives access to pre-trained models from the Hugging Face repository that encodes sequences into fixed-sized vectors. From this library, we downloaded one of the models that encodes texts into vectors of size 384.[25]

[24] See also Carlson (2023) for the application of SBERT to differentiation measure.
[25] The specific pre-trained model used here is *paraphrase-MiniLM-L12-v2*.

C.3. Sentence Similarity Measures

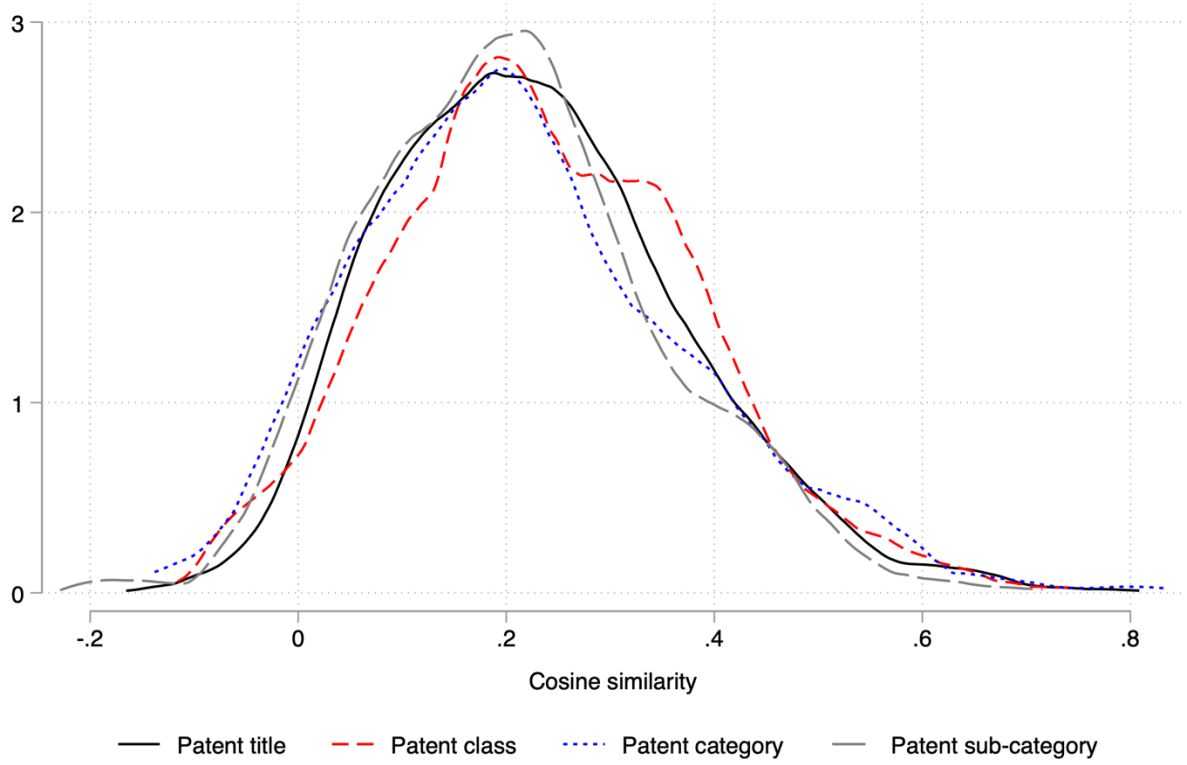
We use the SBERT approach to measure the similarity between patent characteristics (titles, classes, categories, and sub-categories) and themes of graduation theses. While graduation theses by students were written in English and thus the titles of theses are also English, patent characteristics in JPO are described in Japanese. Although the *sentence-transformers* package provides the multilingual versions of pre-trained models, in order to improve precisions, we first translate the words and sentences in patent characteristics into English by the Google Translate API available from python. Then, we apply the *sentence-transformers* package to generate the vector representations of patent characteristics and themes of graduation theses, and given the vector representations of sentences, for each patent observation, we calculate the cosine similarity between each of patent characteristics and theme of graduation thesis by the patent holder. The cosine similarity takes values between -1.0 and 1.0, and higher values mean that patent descriptions are more related to themes of graduation theses, while negative values mean that they are oppositely related.

C.4. Summary Statistics

Figure C1 shows the kernel densities of the calculated cosine similarity. Most of the patents have cosine similarity greater than 0, since it is rare that patents characteristics are oppositely related to the patentors' graduation theses. Although the broader technology categories tend to have a little bit higher similarity than the narrower categories, the shapes of the distributions are relatively similar across different measures: the medians are around .2, the 75-percentile is around .3, and the 90-percentile is around .4. This may be mainly because we can only compare the titles of the theses and those of patents and the information of actual contents is limited. Having that said, there are large variations in the similarity measures, and some patents are perceived as patents so closely related to graduation theses with the similarity scores around .5.

Using these measures, we can come up with several ways to define to what extent graduates produced patents closely related to their graduation theses. For example, one can define patents with the similarity scores above a certain threshold as patents highly related to graduation theses. To be conservative, we use two thresholds: 75-percentile and 90-percentile points for each measure. Table C1 shows the summary statistics for those patents. As in the main text, we define both the number of such patents and an indicator for whether graduates had at least one such patent within 20 years after graduation. Alternatively, one can define the cumulative values of cosine similarity for each graduate as the extent to which they produced patents closely related to their graduation theses. Table C1 also shows its summary statistics. As we show in the main text and the Appendix A, the findings are robust across different measures. We also conducted independent manual matching based on our own visual evaluation of the degree of similarity between patent and graduation theses titles for part of the data and the correlation between SBERT measures of similarity and our own manual matching measure of similarity was quite high.

Figure C1. Distributions of the Textual Similarity Measures



Note: The figure plots the kernel densities of the cosine similarity measures between the theme of graduation thesis and each of patent characteristics. The black solid line shows the density for patent titles, the red dashed line shows the density for patent technology classes, the blue short-dashed line shows the density for patent categories, and the gray long-dashed line shows the density for patent sub-categories. The textual similarity measures are calculated for patents by graduates from naval architecture, mechanical engineering, and electrical engineering divisions.

Table C1. Summary Statistics of Variables Related to Similarity Measures

Panel A. Patent Titles

	N	Mean	SD	Min	Max
1(at least one patent with top 25% similarity score)	1736	0.07	0.25	0.00	1.00
1(at least one patent with top 10% similarity score)	1736	0.04	0.19	0.00	1.00
Logged cumulative number of patents with top 25% similarity score	1736	0.09	0.38	0.00	4.09
Logged cumulative number of patents with top 10% similarity score	1736	0.05	0.26	0.00	3.74
Logged cumulative similarity score	1736	0.15	0.70	-0.10	16.77

Panel B. Patent Technology Classes

	N	Mean	SD	Min	Max
1(at least one patent with top 25% similarity score)	1736	0.07	0.25	0.00	1.00
1(at least one patent with top 10% similarity score)	1736	0.03	0.18	0.00	1.00
Logged cumulative number of patents with top 25% similarity score	1736	0.09	0.40	0.00	4.03
Logged cumulative number of patents with top 10% similarity score	1736	0.04	0.26	0.00	2.78
Logged cumulative similarity score	1736	0.16	0.70	-0.09	11.16

Panel C. Patent Technology Categories

	N	Mean	SD	Min	Max
1(at least one patent with top 25% similarity score)	1736	0.07	0.26	0.00	1.00
1(at least one patent with top 10% similarity score)	1736	0.03	0.17	0.00	1.00
Logged cumulative number of patents with top 25% similarity score	1736	0.10	0.40	0.00	4.06
Logged cumulative number of patents with top 10% similarity score	1736	0.04	0.27	0.00	3.91
Logged cumulative similarity score	1736	0.15	0.69	-0.18	16.31

Panel D. Patent Technology Sub-categories

	N	Mean	SD	Min	Max
1(at least one patent with top 25% similarity score)	1736	0.07	0.25	0.00	1.00
1(at least one patent with top 10% similarity score)	1736	0.03	0.18	0.00	1.00
Logged cumulative number of patents with top 25% similarity score	1736	0.10	0.39	0.00	3.87
Logged cumulative number of patents with top 10% similarity score	1736	0.04	0.25	0.00	3.40
Logged cumulative similarity score	1736	0.14	0.64	-0.46	13.12

Note: Each table shows the summary statistics of variables measuring to what extent graduates produced patents closely related to their graduation theses. Panel A, B, C, D show the summary statistics when we use patent titles, technology classes, categories, and sub-categories to measure the similarity with graduation theses, respectively. These variables are calculated for graduates from naval architecture, mechanical engineering, and electrical engineering divisions.

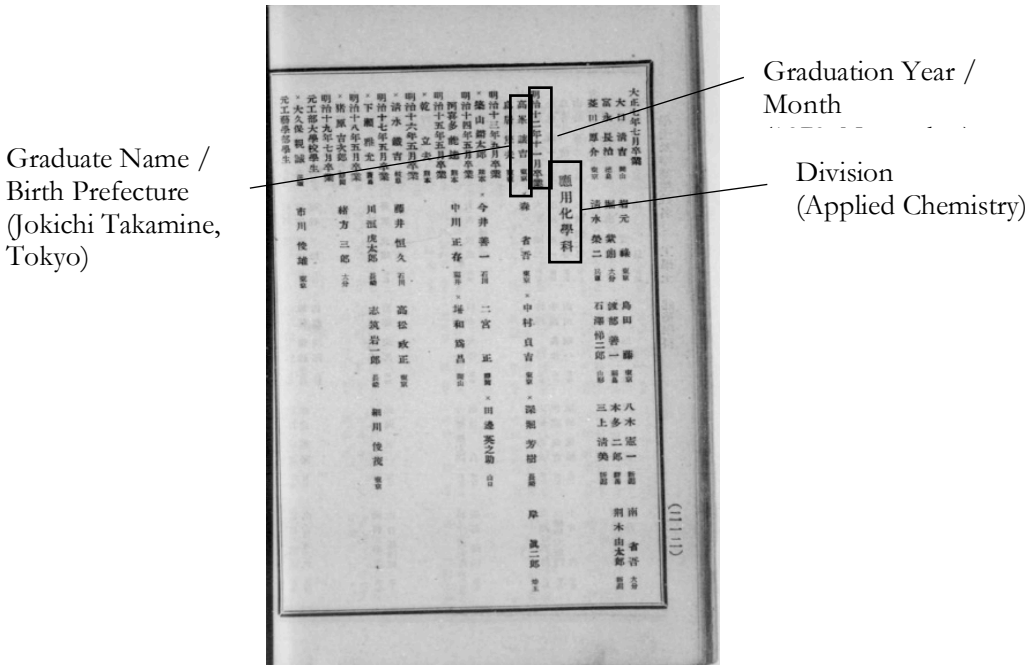
D. Data and Matching Process

D.1. Annual Calendars of Imperial Universities (*Ichiran*)

As mentioned in the main text, our first data source is annual Calendars of Imperial Universities (such as *Tokyo Teikoku Daigaku Ichiran*, also called *Tokyo Teikoku Daigaku Yoran* in some years) available in open access from the National Diet Library Digital Collections (https://dl.ndl.go.jp/en/). These calendars contain names, birth prefectures, and divisions of all students during their years at the university as well as graduation years and months after graduation. Importantly, up until 1918, Tokyo Imperial University Calendars listed the students in the order (*sekiji*, in Japanese) of the average of their test scores in each year, while the graduates were listed in the order of average test scores over all three years of university education, with the score of their graduation thesis added as the fourth component. Thus, the ordering of names is reflective of academic performance relative to peers (in the same division and cohort). As an example, Figure B1 shows that Jokichi Takamine (the discoverer of adrenaline) graduated from the applied chemistry division at the top of the class of six people in 1879.

From these calendars, we digitized the data on all bachelor's degree recipients in engineering from Japan's first (Tokyo) Imperial University and its predecessor (Imperial College of Engineering), starting from the first graduation cohort in 1879 and until the 1918 graduation cohort which gave us the census of all engineering graduates from those cohorts for whom we have information about their graduation rankings .

Figure D1. A sample page of graduates lists in Tokyo Imperial University *Ichiran*



Source: Annual Catalog of Tokyo Imperial University (*Ichiran*) Taisho 7 (1918).

D.2. Public High School Registry

Most students attending Imperial Universities during the industrialization period entered directly from one of the nine public high schools (*Kyusei Koko*) known as No.1-8 high schools and the non-numbered one in Yamaguchi prefecture. These three-year high schools were specifically designed to prepare students for the rigorous academic curriculum offered at Imperial Universities and graduating from these high schools granted automatic admission to the universities.[26] That said, the fierce competition among students existed at high-school admission rather than university admission. The images of high-school own catalogs are also available through the National Diet Library Online (https://ndlonline.ndl.go.jp/#!/), and around 88 percent of university graduates in our sample were matched with their public high-school records.[27] Each public high school maintained annual registries that listed graduates in order of academic achievement, grouped by their chosen departments at the university. The graduation rankings of the high schools followed the same structure as university graduation rankings, allowing us to obtain a proxy for the graduates' academic achievement at the time of their university entrance.

D.3. Annual Alumni Surveys of Imperial Universities' Graduates ("*Gakushikai Kaiin Shimeiroku*")

We hand-matched the data on all engineering graduates from the (Tokyo) Imperial University described in the previous section with *Gakushikai Kaiin Shimeiroku*, the Alumni Surveys of Imperial Universities' Graduates. Images of those surveys for every year from 1890 to 1939 (with three missing years, 1893, 1894, and 1923) were provided to us by *Gakushikai* (the Alumni Association), and they contain information about addresses and workplaces for each graduate belonging to the association, updated annually. We digitized information for all the graduates in our sample of engineering graduates who were found in these surveys (more than 90% of the universe of all graduates) to construct longitudinal data on their employers and addresses. Figure B2 shows how the information for each graduate appears in these surveys.

While some parts of the alumni surveys were used in past studies by Japanese historians (Uchida, 1979; Uemura, 2017a,b; Sawai, 2020), we are the first to use *all* the data available on engineering graduates from those surveys. One of the most challenging aspects in creating the full longitudinal database from the alumni surveys was to keep track of name changes that were common in Japan at that time, mostly due to frequent cases of adult adoption.[28] We also took into account that writing of characters in the names varied from year to year to make sure we do not miss individuals because of differences in how characters are written.

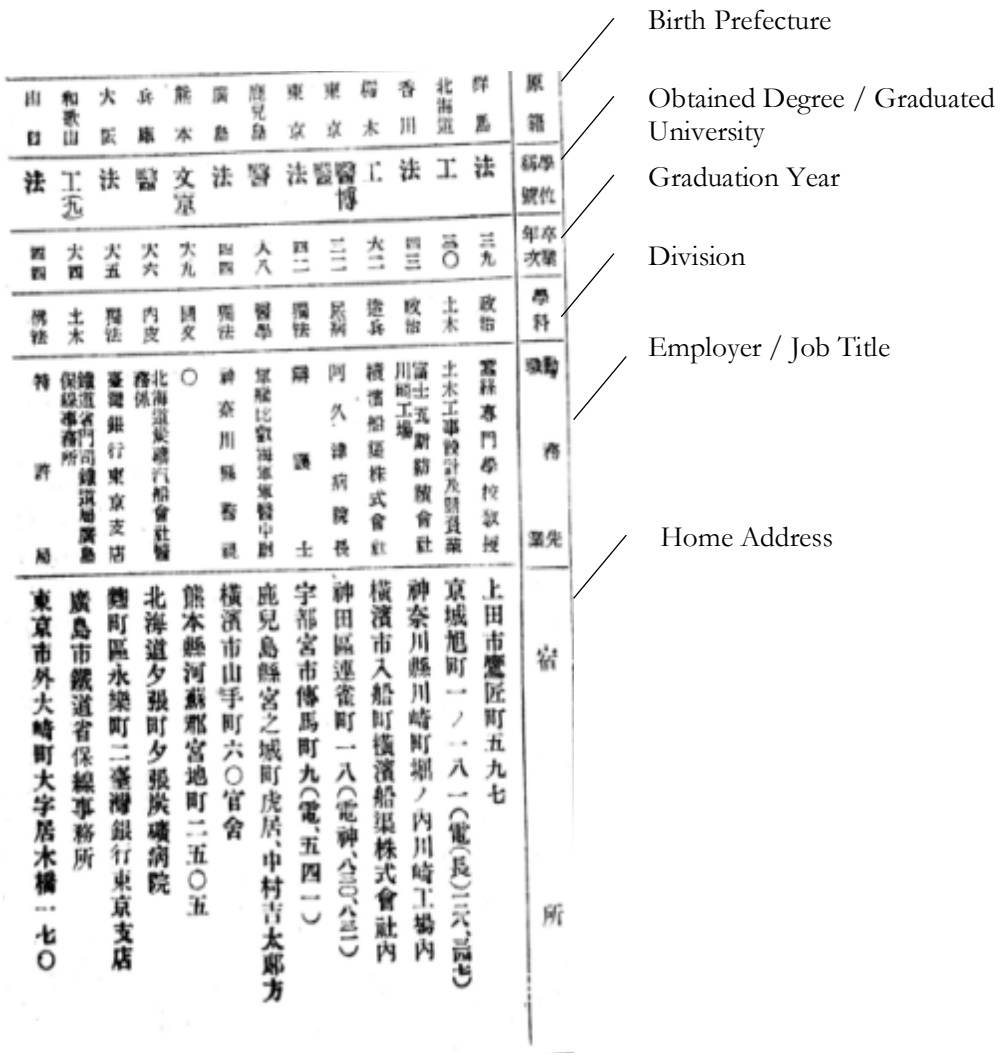
---

[26] While the admission of public high-school graduates into Imperial Universities was promised, they may not have necessarily been enrolled in the division they chose. Some divisions within the universities were highly competitive, and would hold admission exams if demand exceeded the number of available slots.

[27] While almost all the "*Kyusei Koko*" graduates entered an Imperial University, graduating from a public high school was not the only way to admission. Slightly less than 20 percent of graduates-inventors in our sample entered Imperial Universities by taking entrance exams.

[28] It was common in the Edo and Meiji era that people occasionally had multiple names and changed their own names. Although the government enforced laws prohibiting multiple names and renaming in 1873, it relaxed the regulations due to a strong backlash (Yamanushi, 1962). As a result, we observe many cases of changes in both family and given names

Figure D2 A sample page from the annual alumni surveys of Imperial Universities' graduates



Source: Annual Alumni Surveys of Imperial Universities' Graduates ("*Gakushikai Kaiin Shimeiroku*") Taisho 9 (1920). The graduate names are not displayed.

D.4. Matching with Patent Publication Records and Google Patents

The patent law was first enacted in Japan in 1885 and patent data since that year are available from the from the Patent Information Platform (J-PlatPat) operated by the Industrial Property Information and Training Institute (INPIT, https://www.j-platpat.inpit.go.jp/). The collaborating team in Japan shared with us the digitized bibliographic information that recorded all patent specifications for patents granted between 1885 and 1940 (around 126,000 patents), which include patent numbers and titles, technology classes, inventors' and

for the same individuals. For most cases, however, both old and new names were listed for the same individual in separate sections of *Gakushikai Kaiin Shimeiroku* which were used to identify such cases. We also used *Imperial University Calendars*, High-school Registries (*High-school Calendars*), and online sources to track name changes.

assignees' names and addresses. We then hand-matched the information in those patent data with our database of engineering graduates. The matching process was initially conducted by a text-analysis script, followed by the disambiguation process and manual check by two members of the research team conducted independently to remove wrong matches. Different combinations and old forms of characters in which names were written were considered and name changes of university graduates were tracked by extensively investigating various sources, including Calendars of Imperial Universities, public high schools, Japanese Personnel Inquiry Records (*Jinji Koushinroku*), and online searches. To avoid false matches of same-name but different individuals, we manually checked the consistency of employer and address information between the graduates' records and the patent records for all individual matches and dropped those where we were not sure that it was the same person (the fraction of cases dropped through this manual check was about 10 percent of the total number of initial matches).[29] As in Akcigit et al. (2017), we also eliminated from the sample a small number of individuals where we could not definitively ascertain that we had a positive match.

As mentioned in the main text, it is important to measure not just the number but also the quality (impact) of patents (Lanjouw and Schankerman, 2004). Patent citations data are not available for our time periods, so following the spirit of Nicholas (2011), we matched information about Japanese inventors' patents and their names with the names of inventors/assignees and their patents in the Google patent data until 1940 to identify those who were also granted patents outside of Japan (U.S., Britain, and other countries). This matching process consisted of hand-typing English transliterations of the Japanese inventors' names (trying various possible versions of such transliteration) into the Google patent search engine and manually examining the returns to identify proper matches based on names, addresses, and patent characteristics.

To identify patents that were closely related to the fields of study of the graduates in our sample, we constructed a bridge between patent technology classes in the J-PlatPat database and Imperial University College of Engineering divisions. The process was rather straightforward. Details are available upon request.

D.5. Employer Assignment to Industry Codes and Ownership

The process of employer identification proceeded as follows. We started with simply detecting organization names found in the employer name columns extracted from *Gakushikai* surveys, based on a manually constructed organization name list. This was followed by multiple iterations on the manual check of every similar-sounding and differently spelled organization names. Old-style Japanese Kanji characters (*Kyujitai* or *Itaiji*) often resulted in varying representations of the same company. Also, some companies were reported in alumni surveys in abbreviated forms, as also happens in English company names (e.g., IBM or International Business Machines). For instance, The South Manchuria Railway Company (or *Mimami Manshu Tetsudo*, in Japanese), a semi-governmental railroad company in south Manchuria established after the Russo-Japanese War,

[29] In determining if the patent record belonged to a given graduate or not we chose to err on the side of caution and rejected a match if we could not definitively ascertain that we had a positive match.

could be represented by its full name but often would be abbreviated as *Mantetsu*. Another challenge came from some commonalities or patterns in company names. Japanese companies, particularly in the industrialization period, were quite often named as the combination of founder family names and business domains/industries, or city/country names and business domains/industries. For instance, we needed to be careful to distinguish between *Tanaka Mining*, founded by Heihachi Tanaka, and *Tanaka Mine*, founded by Chobei Tanaka, where Tanaka is a common family name. Finally, there were also numerous cases of misspells, omissions and reporting lags of employer names in *Gakushikai*. The self-reported nature of the *Gakushikai* records made all these cases happen frequently, which forced us to conduct multiple rounds of manual checks and triangulations, referring to various archival sources.

To examine employer characteristics, such as industry and ownership, we manually collected information on firm histories for all the firms disambiguated as explained in the previous paragraph, consulting with various archival sources. The primary data source was the restricted-use Business Archives Online (BAO) provided by the Japan Digital Archives Center (J-DAC) (https://j-dac.jp/infolib/meta_pub/G0000004kigyo), which offers scanned images of shareholders' reports for over 14,000 companies. If a firm's financial report for any given year was available in this database, the search results provided information on its primary industry and major shareholders and executives. Another source used to identify company ownership was Japan's Major Company Lineage Charts (*Honpo Shuyo Kigyo Keifuzu-shu*) published by Kobe University, which documented the historical lineage and restructuring events of major firms. We also searched for firm history books (*shashi*), company websites (if still active), and other online sources such as biographies of university graduates and academic papers in Japanese history. In cases where company records were unavailable in any of these sources, we employed alternative methods such as name-based classifications. For example, if a firm's name was "X Mining," we inferred that it was a mining company. However, some observations remained missing if uncertainty persisted. From the above sources we also extracted information on ownership by business groups, particularly ownership by the largest *zaibatsu* (*Mitsui, Mitsubishi, Sumitomo,* and *Yasuda*).

Many important employers in our data operated in more than one industry. In particular, this was true of most *zaibatsu*-owned firms. We manually assigned up to eight industry (SIC) codes to such companies.

D.6. Matching Industry Codes with Field of Study

We used industry codes assigned to employers as described above to create a bridge between the industry code and the field of study of the graduate, using College of Engineering division names in the process similar to matching between patent technology classes and field of study. The matching process is *m:m*, as multiple divisions were matched to the same industry code (e.g., civil, mechanical and electrical engineering divisions would all be matched to railroad transportation industry), while multiple industries were matched to the same division (e.g., both electric services and electric equipment industries would be matched to the electrical engineering division). Details are available upon request.

D.7. Study Abroad

In the mediation analyses in the next section of this appendix we add experience studying abroad after graduation as a potentially important mediator between academic achievement and future invention. This is motivated by the fact that higher-ranked graduates were more likely that lower-ranked graduate to go for a study abroad and that such studies, which lasted for two-three years and took the graduates to top universities in Europe and the U.S. (as well to important industrial firms in some cases for internships), allowed them to absorb even higher-level engineering knowledge.

Gakushikai survey reports contain annually updated information on addresses and workplaces. In particular, if the individual resided abroad in a given year, this is noted in the surveys. The transliteration is not uniform, however, ranging from explicit "study abroad" ("*gaikoku ryugaku*") or "Ministry of Education-sponsored student abroad" ("*Monbusho ryugakusei*") to just listing the individual as "overseas" or "abroad" ("*yoko*" or "*zaigai*" or "*gaikoku yuki*", etc.). Being in specific countries (the United States, Britain, Germany, France) is also often part of the entries.

For the purpose of this paper, we adopted the following convention to classify graduates as on a study abroad. First, if the graduate was explicitly listed as conducting a study abroad he was assigned a "study abroad" indicator in those years. The same indicator was assigned for years where the graduate was listed as staying in one of the frontier countries such as above, with some exceptions noted below.

With respect to listings such as "overseas" or "abroad" ("*yoko*" or "*zaigai*" or "*gaikoku yuki*", etc.), we checked those listing against other sources where those were available. These checks revealed that these listings were also generally associated with travel abroad for study or internship purposes. Thus, we also provisionally assigned the study abroad indicator to those listings. However, in some cases, job or address entries showed that individuals were working in branches of Japan's trading firms abroad or in Japanese firms in Korea, Taiwan, Manchuria, or China while they were listed "abroad." In such cases we excluded those individual-year observations from our "study abroad" definition as those were just jobs associated with foreign travel but not necessarily educational training or internship. We did the same for a few cases where staying in frontier countries was also clearly associated with working for a Japanese trading company branch.

Finally, we checked our study abroad variable constructed from the *Gakushikai* panel against the list of all Japanese nationals who went to study abroad during Meiji era kindly provided by Katsuya Takii and found a strong correlation between that list and our constructed variable. We use our constructed variable in this paper because the other data only cover the period until 1912.

E. Panel Data Analyses and Mediation Estimation

While there is a robust association between academic performance in higher technical education and graduates' inventive outcomes, the propensity to become inventors may also be shaped by the jobs the graduates sorted to after graduation. For example, those who acquire jobs in public administration or military service may have fewer opportunities or incentives to engage in research leading to invention, while those who work for research institutions almost by definition will engage in activities more likely to lead to inventions and patents. Even within the same sector of employment, different organizations (firms) operate in different industries and have different environments. These would also expectedly impact the likelihood of individuals becoming up inventors. Graduates also sorted (or were selected by the organizations they worked for) to go study at top universities, research institutes, or major innovative firms in the U.S., Britain, Germany, France, and so on. They would spend at least a year (often two or three years) on such study abroad stints before returning to Japan having absorbed even higher-level engineering knowledge than what they learned at Imperial University. In this appendix, we utilize the panel nature of our data to measure how such sorting mediates the effect of academic performance on the probability of becoming an inventor.

E.1 Preliminary Examination of Sorting to the First Job

We begin by examining job sorting based on the first job after graduation because the first job choice would be most likely to be influenced by the university education. In contrast, later job sorting may share a common cause with the graduation rankings and an already previously taken patent, potentially leading to the collider bias (Frake, Hagemann, and Uribe, 2024). Almost nobody in our data has patents at the time of their first job selection, however, so this bias would not be an issue if we use the first job characteristics. We also examine sorting to studying abroad as another potential mechanism for the impact of graduation rankings on inventor outcome. As mentioned, obtaining further advanced knowledge through such study was an important factor given that Japan was catching up to the technological frontier.

In Table E1 we present summary statistics on some salient characteristics of the first organizations to which graduates sorted based on whether they became inventors within 20 years after graduation, as well as based upon whether they were in the top decile in terms of their graduation rankings or in the bottom nine deciles (the findings are robust to using other cutoffs). We also show the differences across the same categories in terms of going for a study abroad and whether the graduates sorted into a high-tech industry, or an industry closely matched to their field of study.

[Table E1 around here]

Table E1, Panel A shows that inventors were 10 percentage points (20 percent compared to non-inventors) more likely to sort into industry for their first jobs, and more likely to be first employed in academia or research sectors. Inventors were 18 percentage points (42 percent compared to non-inventors) less likely to be employed in the public sector (excluding public universities and research institutes), however. Comparing this with Panel

B, we can see that graduates in the top decile of graduation rankings were also more likely to be first employed in academia or research compared to graduates in the bottom nine deciles but were actually 11 percentage points *less* likely to be first employed in industry, with no difference in the likelihood of having their first job in the non-research public sector. Thus, in terms of a cross-sector first job sorting, the correlation between future invention and graduation rankings is only partial—while graduation rankings and the propensity to become inventors are both higher for those who choose academia and research, those who choose industry for their first job have relatively lower graduation rankings compared to those who choose public or military service, but they are nevertheless much more likely to become inventors in the future.

Table E1 also reveals that both graduates with high academic rankings as well as inventors sorted to studies abroad and to first jobs in more innovation-oriented organizations although, once again, there are some subtle differences in sorting patterns between top-ranked graduates and future inventors. For instance, both top-decile graduates and future inventors were more likely to secure their first jobs in organizations that already employed more graduates-inventors compared to lower-ranked graduates and future non-inventors. However, future inventors sorted even more compared to top-ranked graduates to organizations that had more peer inventors who took the same field of study in the university. We also see from Table E1 that inventors were more likely to secure first jobs in high-tech industries and industries closely related to the divisions they graduated from, but not necessarily in firms owned by the largest business groups (*zaibatsu*, i.e., *Mitsui, Mitsubishi, Sumitomo*, and *Yasuda*), while there was no difference along these dimensions between top-decile and all other graduates.[30]

All in all, the raw evidence in Table E1 suggests that various aspects of job sorting and eventual invention outcome were often correlated, even conditional on academic achievement. Given these correlations, we examine how these mediate the impact of university graduation rankings on future inventions through the mediating factors using a formal empirical model.

E.2 Employer Sorting and Invention

We now investigate, in a framework similar to Abowd et al. (1999) (AKM; see, e.g., Bhaskarabhatla et al., 2021), how much academic performance continues to predict invention outcomes after controlling for various post-graduation employer characteristics, job experience, and travel abroad. We follow a two-stage empirical strategy. In the first stage, we use our panel data to estimate the contribution of individual fixed-effects to the probability of a post-graduation patent application event in any given year, controlling for having studied abroad as well as various job characteristics (where eventually we include all employer and year fixed effects). In the second stage, we use the individual fixed-effects estimated in the first stage as the dependent variable in what is otherwise the same cross-section regression as in equation (1) above. This allows us to evaluate the role of academic

[30] These results are driven by lower average graduation rankings of those who sort into industry noted above. If we limit the sample to those who selected to be employed in industry (private sector) for their first jobs (not shown), we see that both high-tech and *zaibatsu*-owned firms attracted a significantly higher share of top-decile graduates than other firms.

performance at each stage of higher technical education in predicting post-graduation invention after controlling for employer and year fixed effects as well as the experience of studying abroad.

Specifically, we estimate the following linear probability model in the first stage:

$$y_{it} = X_{it}^{'}\gamma + f_i + f_t + f_{e(i,t)} + \eta_{it}, \tag{2}$$

where y_{it} is an indicator equal to one if graduate i applied for a patent in year t and zero otherwise; X_{it} is a vector which includes job characteristics in year t and a dummy equal to one if the graduate had been exposed to a study abroad by year t and zero otherwise; f_i and f_t are individual and year fixed-effects; respectively, and $f_{e(i,t)}$ are year t employer fixed effects which, when included, absorb the time-invariant employer-characteristics components of vector X_{it}.

Although the main purpose of estimating equation (2) is to obtain estimated individual fixed-effects f_i for the second stage, the estimated coefficients on various components of the vector X_{it} may be of independent interest. We present the estimation results in Table E2.

[Table E2 around here]

The first four columns include the following four indicator variables one by one: one if the individual worked in a research-related job (whether a public research institute or a research lab in the private sector) in year t, one if he worked for a firm affiliated with one of the four largest *zaibatsu* in year t, one if he worked in a firm belonging to a high-tech industry[31] in year t, and one if he already had experience of study abroad by year t. The specifications in columns (5) and (6) include all the indicators, with the specification in column (6) absorbing employer fixed effects, so that the coefficients on employer characteristics are identified through within-employer time variation, in addition to within-individual variation.

The estimated coefficients in the first four columns show working in a research-related job or in a high-tech industry—but not in a firm affiliated with a large *zaibatsu*—was economically and statistically significantly associated with the probability of applying for a patent. A job in research more than doubles the probability in any given year, while working in a high-tech industry increases this probability by about 45 percent. The estimated coefficient on the indicator for having completed a study abroad implies that such experience increased the patenting probability by 3.5 times the mean, underscoring the critical importance of absorbing frontier technologies by Japanese engineers at the time.[32]

[31] Given the historical context, we define high-tech industries based on whether the firm produced output in at least one of the following two-digit SIC-code industry: 28 (Chemical and Allied Products), 35 (Industrial and Commercial Machinery and Computer Equipment), 36 (Electronic and other Electrical Equipment and Components), or 38 (Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods, Watches and Clocks). The findings are not sensitive to expanding this list to include 37 (Transportation Equipment).

[32] Note that year fixed effects which are included in all regressions are perfectly collinear with individuals' ages, so the length of post-graduation experience is also controlled for.

The estimated coefficients in column (5), where we include all the dummies at the same time, are similar to the coefficients in each of the previous columns, indicating all these factors affect invention probability independently of each other. Once employer fixed effects are also absorbed in column (6), however, the coefficient on a research-related job dummy becomes smaller in magnitude and statistically marginally significant, while the coefficient on working in a high-tech industry, in contrast, is almost twice as large as in column (3). This suggests employers' time-varying research facilities had a smaller than average impact, while their time-varying diversification into a high-tech industry mattered more for patenting than the average effect of being employed a high-tech industry. On the other hand, the influence of study abroad on patenting was not much affected by employer changes, as the magnitude of the estimated coefficient on that indicator does not substantially fall between columns (4) and (6).

Our main focus in this exercise is to utilize the estimated individual fixed effects from the first stage to see how much they are affected by educational achievement (knowledge absorption at Imperial University). For this purpose, as mentioned, we estimate the following model that mirrors equation (1) but replaces the dependent variable with the estimated individual fixed-effects, f_i, from equation (2).

$$f_i = \sum_k \beta_k Acad_Perf_{k(i)i} + d_{c(i)d(i)} + d_{c(i)h(i)} + d_{g(i)} + \varepsilon_i \qquad (3)$$

The results are in Table E3. The dependent variable in each of the six specifications corresponds to the estimations in Table E2. That is, the individual fixed effect that serves as the dependent variable in column (1) is obtained from the version of equation (2) that controls for working in a research-related job (plus all fixed effects noted in Table E2). This continues up to column (6), where the dependent variable are the individual fixed effects estimated from the version of equation (2) that includes all four dummies in Table 5 and employer fixed effects.

[Table E3 around here]

The baseline coefficient on university graduation rankings (not shown) obtained from equation (3) while not including any components of vector X_{it} is 0.0189. The magnitudes of the estimated coefficients on graduation rankings in columns (1)-(3) are 15-20 percent smaller than this, but remain statistically significant. This suggests that, at least under the identification assumptions of the AKM model, individuals sorting into research jobs, high-tech industries, and *zaibatsu*-affiliated firms did not much diminish the impact on patenting of educational achievement at the university. However, when we strip the individual fixed-effects of the influence of studying abroad, the magnitude of the estimated coefficient on the university graduation rankings is reduced by about one-third and is no longer statistically significant at conventional levels. The same is true if all four dummies are included at the same time, as shown in column (5).

Notably, the estimated coefficient on university graduation rankings in column (6), where the individual fixed effects in (2) are estimated including all employer fixed-effects as well (making it the closest counterpart

to the estimations implemented by Bharaskabhatla et al., 2021), is once again statistically significant and even somewhat larger in magnitude than the baseline estimate of 0.0189. This may indicate negative assortative matching between individuals' and firms' capabilities, in line with Bharaskabhatla et al. (2021) (see also Metcaffe et al., 2023). Investigating such matching in our historical context is a task for future research.

E.3 Empirical Model of Mediation

While estimation results above show that the impact of academic performance overall remains strong even after controlling for post-graduation job characteristics and study abroad, they do not address the issue of how much of this impact is mediated by sorting to different employers and into study abroad. To decompose the effect of graduation rankings on the probability of becoming an inventor into direct effects (or unmeasured indirect effects) and measured indirect effects of the mediators such as those above, we employ the mediation analysis framework (Heckman and Pinto, 2015, Heckman et al., 2013; see also Fagereng, Mogstad, and Rønning, 2021).

Let y_{ir} denote the potential outcome of whether graduate i became an inventor within 20 years after graduation if his measured academic performance in the $k(i)$-th year in the program (first-year, second-year, university graduation, and high-school graduation rankings) was r ($Acad_Perf_{k(i)i} = r$, for a given k). The vector of mediators when academic performance is *fixed* at r is given by $\theta_{ir} = \left\{\theta_{ir}^{j} : j \in \mathcal{J}\right\}$, where $\mathcal{J}$ is an index set for mediators. Let $\mathcal{J}_p \subseteq \mathcal{J}$ be the index set of mediators that we can measure. Then, our analysis is based on the following linear model:

$$\begin{aligned} y_{ir} &= \kappa_0 + \kappa r + \underbrace{\sum\nolimits_{j\in\mathcal{J}_p} \alpha_r^j\, \theta_{ir}^j}_{\text{measured mediators}} + \underbrace{\sum\nolimits_{j\in\mathcal{J}\setminus\mathcal{J}_p} \alpha_r^j\, \theta_{ir}^j}_{\text{unmeasured mediators}} + d_{c(i)d(i)} + d_{c(i)h(i)} + d_{g(i)} + \tilde{\varepsilon}_{ir} \\ &= \tau_r + \sum\nolimits_{j\in\mathcal{J}_p} \alpha_r^j\, \theta_{ir}^j + d_{c(i)d(i)} + d_{c(i)h(i)} + d_{g(i)} + \varepsilon_{ir}, \end{aligned}$$

where κ represents the direct effect of academic performance and $\tilde{\varepsilon}_{ir}$ is an error term assumed to be uncorrelated with the vector of mediators θ_{ir}. As in Section IV, since academic rankings are cohort $c(i)$- and division $d(i)$-specific, we include cohort-division fixed effects $d_{c(i)d(i)}$ in all specifications and cohort-high-school fixed effects $d_{c(i)h(i)}$ whenever we include high-school graduation rankings in the specifications. We also include the dummies for graduates' birthplace prefecture $g(i)$. All these variables, along with academic performance rankings r, are fixed and not affected by the mediators.

The main challenge to identifying the sources of mediating effects is that we do not measure all the possible mediators. This is highlighted in the first line of the equation above where the set of mediators $\mathcal{J}$ is split in two subsets, one comprised of the mediators we do measure ($\mathcal{J}_p$) and the other, of those we do not measure ($\mathcal{J}\setminus\mathcal{J}_p$). The second line rewrites this by absorbing the unmeasured components into the intercept, $\tau_r = \kappa_0 + \kappa r +$

$\sum_{j\in\mathcal{J}\setminus\mathcal{J}_p}\alpha_r^j\,E\left[\theta_{ir}^j\right]$ and a mean-zero error term, $\varepsilon_{ir}=\tilde{\varepsilon}_{ir}+\sum_{j\in\mathcal{J}\setminus\mathcal{J}_p}\alpha_r^j\left(\theta_{ir}^j-E\left[\theta_{ir}^j\right]\right)$. The τ_r term captures the sum of the direct rank-specific effect and the average effect of the unmeasured mediators. Note that ε_{ir}, the error term in the second line, can be correlated with $\left\{\theta_{ir}^j:j\in\mathcal{J}_p\right\}$ if the measured and unmeasured mediators are correlated.

Next, for each mediator, we use the following linear model specification:

$$\theta_{ir}^j=\mu_0^j+\mu_1^j r+d_{c(i)d(i)}+d_{c(i)h(i)}+d_{g(i)}+\eta_j \tag{E1}$$

where η_j is a mean-zero independent error term. Equation (E1) implies that mediators linearly depend on academic rankings, conditional on division-cohort, high-school-cohort and birthplace fixed effects. Furthermore, we assume that while academic rankings would affect sorting, they would not affect the impact of mediating variables on the outcome, i.e., $\alpha_r^j=\alpha^j$ for all r. Under this uncorrelatedness assumption (Fagereng et al., 2021), the model can be simplified to yield the following observed outcome equation:

$$y_i=\tau_0+\tau r+\sum\nolimits_{j\in\mathcal{J}_p}\alpha^j\,\theta_r^j+d_{c(i)d(i)}+d_{c(i)h(i)}+d_{g(i)}+\varepsilon_{ir}, \tag{E2}$$

where $\tau=\kappa+\sum_{j\in\mathcal{J}\setminus\mathcal{J}_p}\alpha^j\,\mu_1^j$, the sum of the direct academic rankings' effects and the effects of unmeasured mediators. Based on the outcome equation (E2), we can decompose the average effect of academic performance associated with being ranked at r instead of being ranked at r' as in:

$$\begin{aligned}E[y_{ir}-y_{ir'}]&=\tau(r-r')+\sum\nolimits_{j\in\mathcal{J}_p}\alpha^j\,E\left[\theta_r^j-\theta_{r'}^j\right]\\&=\underbrace{\tau(r-r')}_{\text{DE \& UIE}}+\sum\nolimits_{j\in\mathcal{J}_p}\underbrace{\alpha^j\mu_1^j(r-r')}_{\text{MIE}},\end{aligned} \tag{E3}$$

The second term in equation (E3) represents the measured indirect effect (MIE) of academic performance (the effect that operates through measured mediators), while the first term represents the sum of the direct effect (DE) and unmeasured indirect effect (UIE). Decomposition of these effects is straightforward to implement if the measured mediators are independent of the unmeasured ones. In such a case, α^j can be consistently estimated through an ordinary least square regression.

However, as mentioned, any correlation between measured and unmeasured mediators would bias our estimates of the coefficients on the measured mediators. This is likely to be the case in our sample, given that we cannot observe a comprehensive set of firm-level characteristics in the data. In this case, the measured indirect effect would also include any indirect effect of unmeasured components that are correlated with the

measured mediators, so that the first term of the decomposition becomes the sum of the direct effect and indirect effect of those unmeasured components that are orthogonal to the measured mediator variables.

We have a pathway to mitigate this concern for the estimation bias, however, since we observe the identity of the graduates' employers in our data. We can thus include employer-year fixed effects when estimating the outcome equation (E2). The estimated values of these fixed effects allow us to create a unidimensional measure of the employer-year effect. We then employ this measure as another mediator to capture *all* time-variant employer characteristics, including those we do not directly observe in our data. We then include the study abroad dummy as an individual-specific mediator.

Accordingly, our estimation proceeds in two steps. In the first step, we estimate the outcome equation (E2) including employer-year fixed effects. Corralling the estimated values of these fixed effects generates the aforementioned unidimensional measure of time-variant employer effects. The second step involves estimating the mediator equation (E1) for each observed mediator, including also an additional mediator comprised of the estimated employer-year fixed effects, which, by construction, is orthogonal to the remaining variations of all observed mediators. Corralling the coefficients from those regressions, we obtain the estimates of the parameters in equation (E3).

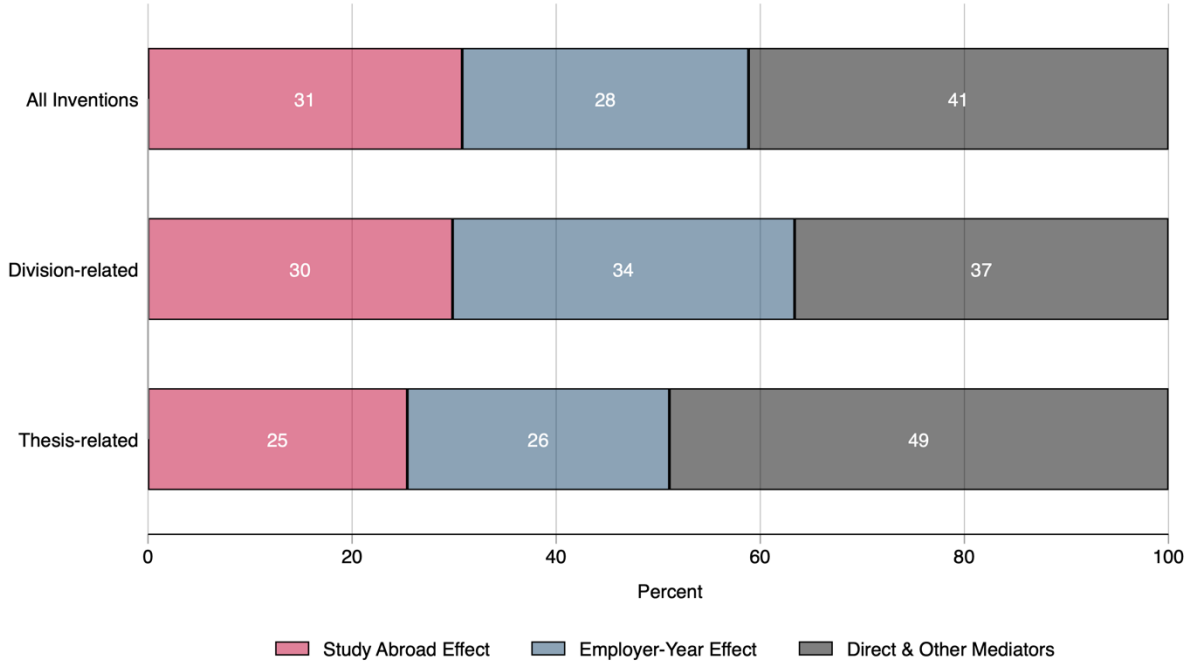
E.3 Estimation Results from Mediation Model

Estimation results from the mediation model are presented in Figure E1. The Figure shows the estimated decomposition of the average effect of academic performance into two parts as per equation (E3): the direct effect of university graduation rankings, which, as noted, may also include the unmeasured indirect effects (the effects of unobserved mediators), and the (measured) indirect effects of having completed a study abroad and employer-year fixed-effects which subsume job characteristics considered in Section V of the main text, such as research-related jobs, jobs in *zaibatsu*-affiliated firms, and firms in high-tech industries. These employer-year fixed effects also absorb time-varying characteristics even at the same employer, such as the number of other inventors in the organization, etc. in Table E1.

The first thing to note is that around 40 percent of the effect of university graduation rankings on becoming an inventor is estimated to come from the direct (including unmeasured indirect) effect for all inventions and also for inventions in technological classes closely related to graduates' divisions. The share of the direct effect of graduation rankings, however, is estimated to be above 50 percent in the sample of inventions in technological classes related to graduates' theses themes. We also conducted the same decompositions for academic rankings at various earlier stages of higher education (not shown) and confirmed that the direct effect was still important, although the magnitude was smaller compared to the direct effect of rankings at graduation. This corroborates the conclusion in Section IV that integrating all knowledge obtained through higher technical education was of utmost importance for graduates to become inventors.

Second, employer-year and study abroad mediating effects are similar in magnitude across all three categories and range from 24 percent each in the sample of inventions in technological classes related to graduates' theses themes to 27-32 percent in the other two samples. While employer-year effects remain something like a "black box" at this stage, summary statistics in Table E1 above suggest that the number of other inventors the organization (firm) employed, especially in the same field of study as the focal graduate, and the number of inventors with global patents may have played a role. The mediating role of peer inventors in the impact of knowledge absorption by the focal graduate on his invention probability may stem from the increased exposure to other inventors (i.e., the spillover effect) and/or from human capital complementarities where talented people in teams work together to produce disproportionately higher output compared to "going it alone" (Stiglitz, 1975).

Figure E1. Decomposition of the Effect of Graduation Rankings on Invention Probability



Note: The figure shows the decomposition of the average effect of graduation rankings: the direct effect of university graduation rankings (including unmeasured indirect effects), and the (measured) indirect effects of the study abroad and employer-year fixed effects.

Table E1. Summary Statistics on First Employment Organizations

Panel A. Future Inventors versus Non-Inventors

	(1) Inventors	(2) Non-inventors	(3) Diff.
1(first employed in non-research private sector)	0.58	0.49	0.10***
	(0.49)	(0.50)	(4.35)
1(first employed in non-research public sector)	0.25	0.43	-0.18***
	(0.43)	(0.49)	(-8.94)
1(first employed in academia)	0.11	0.05	0.06***
	(0.32)	(0.21)	(4.62)
1(first employed in public or private research)	0.04	0.01	0.03**
	(0.18)	(0.09)	(3.29)
Number of peer inventors in the first organization	1.79	1.57	0.22
	(3.34)	(2.53)	(1.50)
Number of peer inventors from the same division in the first organization	0.52	0.37	0.15**
	(1.10)	(0.95)	(2.99)
Number of peer global inventors in the first organization	0.31	0.11	0.20***
	(0.92)	(0.44)	(4.89)
Number of peer global inventors from the same division in the first organization	0.05	0.00	0.04***
	(0.27)	(0.07)	(3.32)
Number of patents by other graduates per year in the first organization	0.75	0.48	0.27***
	(1.65)	(1.20)	(3.71)
1(first employed in a firm owned by *zaibatsu*)	0.16	0.14	0.02
	(0.37)	(0.35)	(1.44)
1(first employed in high-tech industry)	0.37	0.22	0.15***
	(0.48)	(0.41)	(7.08)
1(first employed in division-related industry)	0.69	0.61	0.07**
	(0.46)	(0.49)	(3.29)
1(study abroad)	0.39	0.23	0.16***
	(0.49)	(0.42)	(7.37
Observations	571	3092	3663

Panel B. Top Decile versus Bottom Nine Deciles by University Graduation Rankings

	(1) Top Decile	(2) Bottom Nine Deciles	(3) Diff.
1(first employed in non-research private sector)	0.41	0.52	-0.11**
	(0.49)	(0.50)	(-4.73)
1(first employed in non-research public sector)	0.40	0.40	0.00
	(0.49)	(0.49)	(0.13)
1(first employed in academia)	0.15	0.04	0.11***
	(0.36)	(0.20)	(6.76)
1(first employed in public or private research)	0.03	0.01	0.01
	(0.16)	(0.10)	(1.96)
Number of peer inventors in the first organization	2.19	1.50	0.69***
	(3.51)	(2.49)	(4.23)
Number of peer inventors from the same division in the first organization	0.53	0.37	0.16**
	(1.16)	(0.94)	(2.87)
Number of peer global inventors in the first organization	0.27	0.12	0.15***
	(0.88)	(0.47)	(3.78)
Number of peer global inventors from the same division in the first organization	0.03	0.01	0.02*
	(0.20)	(0.10)	(2.13)
Number of patents by other graduates per year in the first organization	0.69	0.49	0.20**
	(1.51)	(1.24)	(2.86)
1(first employed in a firm owned by *zaibatsu*)	0.14	0.14	-0.01
	(0.34)	(0.35)	(-0.52)
1(first employed in high-tech industry)	0.25	0.24	0.01
	(0.43)	(0.43)	(0.26)
1(first employed in division-related industry)	0.60	0.63	-0.03
	(0.49)	(0.48)	(-1.22)
1(study abroad)	0.49	0.22	0.27***
	(0.50)	(0.41)	(11.73)
Observations	518	3145	3663

Note: ***, **, and * refer to statistical significance using two-sided t-test at the 1%, 5%, and 10% levels, respectively. "Peer inventors" identified from the universe of university graduates.

Table E2. Panel-Data Estimation of the Probability of Invention

	(1)	(2)	(3)	(4)	(5)	(6)
1(research-related job)	0.0237***				0.0233***	0.0165*
	(0.00846)				(0.00819)	(0.00846)
1(firm affiliated with a *zaibatsu*)		0.00437			0.000548	0.0134
		(0.00594)			(0.00589)	(0.0309)
1(firm in a high-tech industry)			0.0112***		0.0104***	0.0213***
			(0.00338)		(0.00333)	(0.00597)
1(after study abroad)				0.0781***	0.0779***	0.0717***
				(0.0105)	(0.0104)	(0.00956)
Mean of DV	0.0225	0.0225	0.0225	0.0225	0.0225	0.0226
Individual FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Employer FEs	No	No	No	No	No	Yes
Observations	56995	56995	56995	56995	56995	56793
Adj R-Squared	0.181	0.181	0.181	0.186	0.186	0.204

Note: The dependent variable is a dummy equal to one if the individual applied for a patent in a given year and zero otherwise (see equation (2) in the main text). Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table E3 Academic Performance and Individual Fixed-Effects from Table E2

	(1)	(2)	(3)	(4)	(5)	(6)
1st year class ranking	0.00239	0.00246	0.00154	0.00152	0.000616	-0.000102
	(0.00704)	(0.00705)	(0.00701)	(0.00687)	(0.00677)	(0.00734)
2nd year class ranking	0.00564	0.00586	0.00549	0.00271	0.00231	-0.00528
	(0.00740)	(0.00738)	(0.00735)	(0.00708)	(0.00705)	(0.00842)
University graduation ranking	0.0151*	0.0155**	0.0161**	0.0121	0.0117	0.0192**
	(0.00768)	(0.00764)	(0.00751)	(0.00743)	(0.00730)	(0.00917)
High-school graduation ranking	0.00436	0.00437	0.00482	0.00307	0.00328	0.00282
	(0.00613)	(0.00614)	(0.00613)	(0.00606)	(0.00596)	(0.00653)
Mean of DV	0.0201	0.0202	0.0182	0.0170	0.0136	0.00910
Division X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
High-school X cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Birthplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3087	3087	3087	3087	3087	3087
Adj R-Squared	0.0558	0.0603	0.0580	0.0266	0.0199	0.00173

Note: The dependent variables in each column are individual fixed-effects estimated in the corresponding column in Table 5 (see equation (2) in the main text). Standard errors are clustered at the cohort-division level in the columns (1) to (3) at the cohort-division level and cohort-high-school level in the columns (4) to (6) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

F. Details of Structural Equation Parameters Estimation

In this section, we demonstrate the exact algebraic identification of the structural parameters in our human capital accumulation model. Assume that observed test scores $\tilde{s}_{0i}$, $\tilde{s}_{1i}$, $\tilde{s}_{2i}$ and $\tilde{s}_{Gi}$ are all standardized to have a mean of zero and a variance of one. The weights w_1, w_2, and w_3 are known constants. Then, we know that $\tilde{s}_{Gi}$ can be written as:

$$\tilde{s}_{Gi} = \frac{w_1\tilde{s}_{1i} + w_2\tilde{s}_{2i} + w_3\tilde{s}_{3i}}{\sigma_G} \tag{11}$$

where σ_G is the unknown scaling parameter to make the variance of $\tilde{s}_G$ equal to one. The problem is that we do not observe $\tilde{s}_{3i}$ and its correlation with $\tilde{s}_{1i}$ and $\tilde{s}_{2i}$, so we cannot estimate σ_G immediately. Our goal is to recover the structural parameters solely from the observable variables.

Step 1. Because $\tilde{s}_{0i}$, $\tilde{s}_{1i}$, $\tilde{s}_{2i}$ are observed, the parameters ρ_{10}, ρ_{20}, and ρ_{21} are simply identified: regressing s_1 on s_0 yields, $\hat{\rho}_{10}$, and regressing s_2 on s_0 and s_1 yields $\hat{\rho}_{20}$ and $\hat{\rho}_{21}$.

Step 2. To identify the parameters in the third-year human capital accumulation, we first estimate the reduced-form regression of the observed $\tilde{s}_G$ on the observables $\tilde{s}_{0i}$, $\tilde{s}_{1i}$, $\tilde{s}_{2i}$:

$$\tilde{s}_G = \gamma_0\tilde{s}_{0i} + \gamma_1\tilde{s}_{1i} + \gamma_2\tilde{s}_{2i} + \tilde{\varphi}_{Gh} + \tilde{e}_{Gi} \tag{12}$$

By substituting the third-year human capital accumulation equation into (11), we obtain the theoretical composition of $\tilde{s}_G$:

$$\tilde{s}_G = \frac{1}{\sigma_G}[w_1\tilde{s}_{1i} + w_2\tilde{s}_{2i} + w_3(\rho_{30}\tilde{s}_{0i} + \rho_{31}\tilde{s}_{1i} + \rho_{32}\tilde{s}_{2i} + \tilde{\varphi}_{3h} + \tilde{e}_{3i})] \tag{13}$$

Rearranging the terms, we map the structural parameters to the reduced-form coefficients γ:

$$\gamma_0 = \frac{w_3\rho_{30}}{\sigma_G}, \gamma_1 = \frac{w_1 + w_3\rho_{31}}{\sigma_G}, \gamma_2 = \frac{w_2 + w_3\rho_{32}}{\sigma_G} \tag{14}$$

If the scaling factor σ_G is known, the structural parameters ρ_{30}, ρ_{31}, and ρ_{32} in the third-year human capital accumulation equation can be algebraically backed out.

Step 3. We can derive σ_G algebraically without observing $\tilde{s}_{3i}$. Let C_{xy} denote the correlation between variables x and y. Note that the variance of $w_1\tilde{s}_{1i} + w_2\tilde{s}_{2i} + w_3\tilde{s}_{3i}$ is equal to ${\sigma_G}^2$:

$$\sigma_G{}^2 = w_1{}^2 + w_2{}^2 + w_3{}^2 + 2w_1w_2C_{12} + 2w_1w_3C_{13} + 2w_2w_3C_{23} \tag{15}$$

where C_{km} are the correlations between $\tilde{s}_k$ and $\tilde{s}_m$. We can eliminate the unobserved correlations C_{13} and C_{23} by using the covariances between the observed standardized variables. By definition:

$$C_{G1} = \mathrm{Cov}(\tilde{s}_G, \tilde{s}_1) = \frac{w_1 + w_2cr_{12} + w_3C_{13}}{\sigma_G} \tag{16}$$

$$C_{G2} = \mathrm{Cov}(\tilde{s}_G, \tilde{s}_2) = \frac{w_1cr_{12} + w_2 + w_3C_{23}}{\sigma_G} \tag{17}$$

Solving for w_3C_{13} and w_3C_{23} from these equations and substituting into (15) yields:

$$\sigma_G{}^2 - 2(w_1C_{G1} + w_2C_{G2})\sigma_G + (w_1{}^2 + w_2{}^2 + 2w_1w_2C_{12} - w_3{}^2) = 0 \tag{18}$$

Letting $A \equiv w_1C_{G1} + w_2C_{G2}$ and $B \equiv w_1{}^2 + w_2{}^2 + 2w_1w_2C_{12} - w_3{}^2$ and solving for the positive root gives us:

$$\sigma_G = A + \sqrt{A^2 - B}$$

Note that C_{12}, C_{G1}, and C_{G2} can be fully estimated using the coefficients obtained from the regressions in Step 1 and Step 2.

Step 4. With $\hat{\sigma}_G$ recovered, we can identify the remaining structural parameters:

$$\hat{\rho}_{30} = \hat{\gamma}_0\left(\frac{\hat{\sigma}_G}{w_3}\right), \hat{\rho}_{31} = \hat{\gamma}_1\left(\frac{\hat{\sigma}_G}{w_3}\right) - \frac{w_1}{w_3}, \hat{\rho}_{32} = \hat{\gamma}_2\left(\frac{\hat{\sigma}_G}{w_3}\right) - \frac{w_2}{w_3}$$

Then, we can also use $\hat{\sigma}_G$ to recover $\tilde{s}_{3i}$. Using $\tilde{s}_{3i}$ directly, we can estimate the inventor production function and recover the remaining structural parameters β.