The Labor Market Determinants of the Payoffs to University Field of Study

Massimo Anelli (Bocconi University, CESifo, IZA) *

February 15, 2018

[PRELIMINARY: PLEASE DO NOT CIRCULATE]

Abstract

Recent studies have estimated a high degree of heterogeneity in payoffs to different university fields of study, marking the a shift of education literature towards them. I contribute to this literature by studying the labor market mechanisms through which the payoffs to university fields of study are so heterogenous, whether this heterogeneity persists over cohorts and by exploring the reasons why these differences persist. To do so, I match 20 cohorts of individual level administrative data from high schools and universities to social security records that allow me to follow them up to 25 years after university graduation. I exploit multiple institutional features to design instrumental variable strategies to estimate causal returns to field enrollment. I find large differences in long-run returns to choice of field with "Economics & Business" delivering around 100% higher returns with respect to "Humanities" (the lowest paying field). Income trajectories differ substantially across fields and returns are heterogenous by gender and family background. A higher probability to reach the top 1% of the income distribution for high-payoff fields, a substantially higher probability of having a discontinuous career plus higher mismatch rates for the low-payoff fields are the most relevant determinants of differences in payoffs across fields. Contrary to what theory would predict, I do not observe a decrease in enrollment in low-payoff fields over time and differences in payoffs tend to diverge. Volatility-adjusted returns and risk considerations do not justify the persistent choice of fields, leaving a lack of information and an increasing importance on non-monetary payoffs in the choice as most plausible candidates for the observed persistence.

*Address: Bocconi University, Department of Policy Analysis and Public Management, Via Roentgen 1, Milan, 20136, Italy. massimo.anelli@unibocconi.it

JEL Codes: I23, I26, J24
Key Words: Returns to education, choice of major, human capital

1 Introduction

While the college premium has been rising in the past decades[1], a growing number of recent studies have shown a high degree of heterogeneity in payoffs to different college fields of study in several countries (Hastings et al., 2014; Altonji et al., 2016; Kirkebøen et al., 2016). The returns to choosing the most rewarding fields is estimated to be comparable in magnitude to the college wage premium (Altonji et al., 2012; Kirkebøen et al., 2016). This evidence, together with the growing share and relevance of university graduates in the labor market, justifies a shift of the academic focus towards the returns of the choice of university field as one of the main determinants of labor market outcomes.

With this work, I contribute to this literature by studying the mechanisms through which the payoffs to university fields of study are so heterogenous, whether this heterogeneity persists over cohorts and explore the reasons why these differences persist. To do so, I match, at the individual level, Italian administrative data from high schools and universities to social security records that allow me to track every single labor market episode of university graduates up to 25 years after university graduation.

Estimating causal estimates of returns to choice field is very challenging because of multiple dimensions of selection on both ability and comparative advantages. For instance, while for returns to schooling ability is clearly correlated with higher schooling and higher outcomes, in the context of the choice of field of study, selection into field based on ability is not correlated with the payoffs of fields (e.g. students enrolling in humanities, the lowest-payoff field in my data are among the best in terms of high school performance).

[1] See Bound & Johnson (1992); Katz & Murphy (1992); Juhn, Murphy & Pierce (1993); Goldin & Katz (2007); Autor, Katz & Kearney (2008)

In the literature, causal identification has generally been sought by exploiting strategies to control for selection of observables (Hamermesh and Donald, 2008; Del Rossi and Hersch ,2008; Chavalier, 2011; Webber, 2014) or by designing discrete choice models (Arcidiacono, 2004; Beffy, Fougere and Maurel, 2012; Kinsler and Pavan, 2014; Altonji et al., 2016). More recently, Hastings et al. (2014) and Kirkebøen et al. (2016) exploited admission score cutoffs in countries where there exists a centralized admission procedure to estimate returns to field in a fuzzy regression discontinuity framework.

To get causal estimates of the returns to field, in my paper I first control for selection on relevant observable characteristics (i.e. high school performance, parental background, school and teacher fixed effects). The Italian institutional setting, as it is for the majority of countries, does not have a centralized college admission procedure that allows me to exploit a Fuzzy Regression Discontinuity approach such as in Chile (Hastings et al., 2014) or Norway (Kirkebøen et al., 2016). Nonetheless, I can exploit other institutional features (the field-of-study recommendation announced by the instructors at the end of high school to the student) and the richness of my high school data (the persistence of certain choices across cohort and within same group of instructors) in multiple IV strategies to estimate causal returns to choice of field.

The most relevant and original contributions of this paper to the literature leverage the richness of the social security data available for my analysis. The possibility of following 20 cohorts and the large number of years of labor market outcomes that I can track, allow me to study several research questions that have remained so far unexplored. Which are the labor market outcomes/events that explain these stark differences in returns across graduates of different fields? Do career trajectories of different fields have heterogenous patterns? Are lower returns to some fields driven by an excess supply of graduates from those fields who end up working in occupations that do not require a college degree or in industries that do not require the skills acquired in college?

In a second section of the paper, I explore whether the large differences in payoffs across fields persist over cohorts and I try to identify possible dimensions of the choice of major that might be relevant for justifying such persistence. Are differences in returns across fields decreasing over cohorts? If returns to certain fields are particularly low due to an excess supply of graduates, later cohorts might take this into account when making their field choice. This should reduce differences in payoffs across fields. If differences across fields persist, people might still be choosing fields optimally. Non-monetary returns could thus be driving it or there might be a trade-off between expected returns and their volatility. I thus explore in the data whether enrollment responds over time to differential payoffs and the extent of the high returns-high volatility trade off to understand whether these factors can justify the differences in the choice.

My analysis show large differences in long-run returns to choice of field with "Economics & Business" delivering around 100% higher returns with respect to "Humanities" (the lowest paying field). Income trajectories also differ substantially across fields with Engineering starting among the highest paying right after college, but having a way flatter trajectory than fields such as law or medicine, which start lower and surpass engineering after about ten years from graduation. While "Economics & Business" remain the highest trajectory , humanities is the lowest in every year after graduation. Among labor market determinants of the differential payoffs to field, the most relevant are a higher probability to reach the top 1% of the income distribution for high paying fields and a substantially higher probability of having a discontinuous career (fixed term contracts and unemployment spells) for the low-paying fields. I also link sector/occupation cells of destination with PIAAC survey data to estimate the degree of "over-education" for graduates of different fields and find that a much larger fraction of employment spells of fields such as humanities

(33%) , architecture and design (40%) are in sector/occupation cells that do not require a university degree. I interpret this as a clear sign of excess supply of graduates from these fields. I also explore differential returns to fields across gender and parental income. While returns for women appear to be lower for all fields, engineering, economics and business have much smaller penalties. Returns to field appear to depend also on family background with returns to Law, Architecture and Design being substantially lower for graduates with fathers having an income below median.

When I explore how returns vary over cohorts, I find that even though the Humanities have the lowest payoffs already for the earliest cohorts, its enrollment in later cohorts increased substantially, and its payoffs kept dropping over cohorts. Instead of converging, payoffs appear to diverge over time. This clearly excludes the possibility of an efficiency-improving adjustment in the choice of field over time and can only be explained by a lack of information on returns to field and by a possible increase over time of non-monetary returns of humanities. When I check for the possibility that the persistence in choice of major is driven by a high returns-high volatility trade-off, I do not find strong empirical evidence to support this hypothesis. It is true that some of the high-payoff fields have higher volatility (i.e. Law, Business & Economics), while low-payoff fields have low volatility (i.e. Humanities), however confidence intervals around payoff trajectories, calculated taking sample volatility into account, never overlap. Differentials across fields of risk-adjusted returns are more compressed, although differences remain across fields. This evidence appears to justify only a limited role of the high risk-high return trade-off in the optimal choice of field. Other risk-related considerations in the choice of field might be driven by relative insulation from the business cycle. I have explored this hypothesis and I found that graduates of high paying fields do not suffer long-run penalization for entering the labor market in a year of economic recession, contrary to graduates of social sciences (a low-payoff field) who appear to be the most penalized by graduating during an economic downturn.

While these findings are specific to the institutional and empirical context under study, they contribute to the literature by offering a first detailed analysis of the labor market mechanisms driving the strong heterogeneity in the payoffs to different university fields. Moreover, given that most recent causal evidence on returns to field has been drawn by specific countries and institutional contexts (e.g. Chile and Norway), I contribute by offering evidence on returns to field choice for a rather different country. The evidence presented in this paper has important insights also for policy. The clear presence of an excess supply of graduates of the lowest-payoff fields and an increase over time of enrollment in those same fields, that can not be explained by an increase in non-monetary payoff or risk-related considerations, suggests the presence of large inefficiencies in the choice of fields of Italian students. Lack of public information on the payoffs and labor market prospects of each field is the first candidate to explain such inefficient choice process. Absent information on payoffs by field, in an institutional setting in which university tuitions are highly subsidized, students might choose fields based solely on individual preferences for specific disciplines. Given that the Italian university system is almost exclusively state-run, it should be fairly inexpensive to aggregate information on payoffs and labor market prospects by fields and provide it to students and families at enrollment at university.

2 Empirical and Institutional Framework

For this paper, I use administrative data collected in the city of Milan (a large service-oriented metropolitan area in the wealthiest part of Italy), for about 30000 individuals graduating from college-preparatory high schools ("licei") between 1985 and 2005. Individual records were digitized in the archives of these schools from official hard-copy registries compiled for the high school exit exam.[2] These registries have a public nature and are safely kept by each school as the only formal record of high school graduation and of the final graduation score. From this registries, it is possible to retrieve final graduation score, class composition, assignment to a specific team of instructors, basic demographic characteristics and, importantly for this work, the field of study that is formally suggested by the group of instructors at the end of high school. Based on the address where students used to live at time of school, I geo-coded the position of the parental house and matched administrative data on house value by block. I exploit this measure as a control for parental wealth. After digitization of high school records, data have been merged with records from the five universities of Milan. Given the wide offer of university degrees in Milan and the high quality and reputation of its higher education, the large majority of high school graduates from Milan attend university in one of these five universities. University records contain individual information on the degree enrolled, completion, change of degree, time to graduation and final GPA. Finally, I linked the academic records to individual social security records from several data sources (data were made available and matched by the national social security institute, INPS). The main sources are the registry of all employment records from 1980 to 2016, the registry of self-employment records, professional occupation records from the regulated professional bars (e.g. lawyers, engineers, architects, business consultants, physicians, etc.) and welfare assistance records (e.g. unemployment, sickness, maternity leaves, etc.). All together, the different sources provide a complete overview of all possible post graduation records and events. These data contain an extremely rich set of labor market information. Thanks to the civil statistic office of the city of Milan, I have also identified parents of the students in my analytical sample and thus linked parental social security records.

[2]The final high school exit exam is centralized and run by the Ministry of Education and plays a very formal role in the education career of individuals.

2.1 The Italian Higher Education System

The Italian higher education context is prevalently State-run and characterized by some peculiarities with respect to other countries. In 2016 only 26.2% of Italians between 30 and 34 had a university degree, the lowest share in the EU (with the UE average being 39.2 %). Although the supply of university graduates remains low, their labor market prospects are on average rather disappointing. The unemployment rate of university graduates between 25 and 39 is 17.1% (only slightly lower than the 21.1% rate for Italian high school graduates). For those employed, the college premium is on average lower than in the United States. Colonna (2007) estimates an average college wage premium in Italy of 24% vs 55% for the US. However, Colonna (2007) finds a substantially lower degree of selectivity into university (based on ability) in Italy than in the US and show that after accounting for ability, college premium in Italy is substantially higher than in the US.

In the Italian context, the choice of university field of study takes place towards the end of the last year of high school when students enroll in a specific degree of a chosen university. This is different from the US context, but very similar to what happens in all European countries. In the period under analysis universities offered 4-year degrees and virtually no field of study had a selective admission test, except Medicine. In one of the 5 universities (private and the smallest in Milan), also Economics and Business degrees had a selective admission test. However, the overall admission by degree across universities was virtually unconstrained in Milan in the period under analysis.

When interpreting the results of my analysis, we need to consider that Milan higher education and labor market offers potentially higher quality education and better prospects for university graduates. Moreover, differently from the rest of the country two out of the five universities in Milan are private, selective schools. As a consequence, the results of this analysis might be representative for a rather high portion of the Italian student distribution.

3 Identification strategy

Estimating causal estimates of returns to choice field is very challenging because of multiple dimensions of selection on ability, comparative advantages and parental background. While in the context of returns to schooling, ability is positively correlated with higher schooling and better outcomes, in the context of the choice of field of study, selection into field based on ability is not clearly positively correlated with the payoffs of fields (e.g. students enrolling in humanities, the lowest-payoff field in my data are among the most positively selected in terms of high school performance). In the literature, causal identification has generally been sought by exploiting strategies to control for selection of observables (Hamermesh and Donald, 2008; Del Rossi and Hersch ,2008; Chavalier, 2011; Webber, 2014) or by designing discrete choice models (Arcidiacono, 2004; Beffy, Fougere and Maurel, 2012; Kinsler and Pavan, 2014; Altonji et al., 2016). More recently, Hastings et al. (2014) and Kirkebøen et al. (2016) exploited admission score cutoffs in countries where there exists a centralized admission procedure to estimate returns to field in a fuzzy regression discontinuity framework. The Italian institutional setting, as it is for the majority of countries, does not have a centralized college admission procedure that allows me to exploit a Fuzzy Regression Discontinuity approach such as in Chile (Hastings et al., 2014) or Norway (Kirkebøen et al., 2016). Nonetheless, I can exploit other institutional features (the field-of-study recommendation announced by the instructors at the end of high school to the student) and the richness of my high school data (the persistence of certain choices across cohort and within same group of instructors) in multiple IV strategies to estimate causal returns to choice of field.

My baseline model take advantage of the richness of my pre-university data to control for as much selection into field as possible. A large set of school and teacher fixed effects allows me to identify returns to field within school, cohort and teachers and at the same time controlling for high school performance, family background, high school class and peer characteristics. To exploit all the information available for the labor market, I pool all individual observations and define y_{itcsg} as the labor market outcome (e.g. income, unemployment, type of contract, etc.) observed t years after expected graduation for individual i who completed high school with graduating cohort c, attended school s with group of instructors g. Given that I observe multiple social security records per year, I had to collapse the relevant information at the yearly level. For instance, I have summed yearly income from different working spells and I have chosen employment status in a given year to be the status that had the longest duration during the year (e.g. if the individual spent 6 months and 1 day in unemployment I define her to be unemployed for that yearly observation). In order to include the multiple observations for each individual in my sample, I cluster standard errors at the individual level. Since time to graduation is per se an outcome of the choice of field, I prefer to measure years on the labor market starting from a theoretical year of university of graduation, defined as the 5^{th} year after high school graduation (and university enrollment). While in the period under analysis, university students attended 4-year degrees, I leave an extra year for late graduates and define the 6^{th} year after university enrollment as the first year of labor market experience. The full model is specified as follows:

$$y_{itcsg} = \alpha + \pi_i^f + \beta X_i + \gamma Z_{-i} + \phi_{sc} + \psi_g + \epsilon_{itcsg} \tag{1}$$

where π_i^f is a set of dummies for field of study choice (9 fields or 4 broad fields in some specifications). Humanities is the omitted field category. I define the choice of field of study as the one taken at the very first university enrollment (i.e. the year of high school graduation). While I can effectively control for selection into and instrument the first choice of field, field of graduation might already be the result of a process endogenous with respect to the initial choice. X_i is a set of individual characteristics such as gender, high school exit score, parents' wealth and a dummy for whether the student was commuting daily from outside Milan. Z_{-i} is a set of average class peer characteristics such as gender share, share of low SES classmates, share of high SES classmates. In the Italian secondary school context, students are assigned to a class (i.e. a group of classmates), which remain fixed for 5 years, attend the same courses and shares the same exact instructors.[3] As shown in Anelli and Peri (2017) using the same high school data, high school peers play a relevant role in influencing the choice of field substantially. Controlling for peer characteristics is thus crucial in this context as well. The baseline model include also a set of fixed effects that allow to identify returns to choice of field using variation from a very restricted set of homogenous individuals. ϕ_{sc} is a set of school/cohort fixed effects while ψ_g is a set of instructor team fixed effects. In the Italian high school context each class is assigned to a specific team of instructors who is collectively responsible for class performance. These teams are persistent across cohorts and it is thus possible to control for fixed effects. Overall the identifying variation is thus coming within school across cohorts, holding constant the effect of instructor teams and controlling for individual and classroom characteristics. Finally, standard errors clustered at the individual level.

A consequence of pooling all observations available together is that the earliest cohort (i.e. those graduating from high school in 1985) can be followed on the labor market for up to 26 years after expected graduation, while the latest cohort for 11 post-education labor market years. Hence, I observe the first 11 years of labor market outcomes post-graduation for all individuals in my dataset, while I loose one year of labor market outcomes after the 11^{th} for each cohort after the earliest. In Figure 1 I present a simple timeline diagram of cohorts and number of observations on the labor market per individual. As a consequence,

[3] There are no elective courses in the Italian secondary school

the pool dataset will tend to mechanically overweight the earlier labor market outcomes after graduation. To avoid this issue, I weight my regressions by the inverse of the total number of observations per labor market year after graduation (i.e. I weight labor market observations past the 11^{th} post-graduation labor market years relatively more).[4]

3.1 IV strategy - Teacher field suggestion

To reinforce the identification of returns to choice of field, I exploit a unique feature of the Italian institutional context. At the end of high school the team of teachers who is responsible for each class formally suggest students a specific university degree to continue their studies. This formal suggestion is recorded in the registry of the high school exit exam ("Maturitá). In Figure 2 I report an example of what the suggestion looks like in the registries digitized for this paper. The example in Figure 2 reports the following: *"We suggest the student to continue her/his studies with an Economics and Business Degree"* This registry entry is complete and digitized for all cohorts graduating until the year 1999, which was the last year in which this formal suggestion was recorded in the high school exit exam registries. It is likely that the suggestion was communicated to students also in following years, however I do not observe it.

In a first simple instrumental variable strategy, I define $F+1$[5] dummies $S_i^f = 1$ if the student was suggested to enroll in field f and 0 otherwise and then use the F dummies S_i^f in model 1 as instrumental variable for the F dummies π_i^f of field choice. This strategy has the advantage of exploiting a shifter of field choice which is in theory exogenous to the choice of student. However, teachers are likely to suggest students to enroll in a field if

[4] I could alternatively

[5] Although the analysis is conditional on enrolling in university and almost all students in the high school sample enroll, a sizeable number of students received the formal suggestion to "not continue with university studies". I thus use these students as reference category for the set of suggestion dummies. The model is thus overidentified with F endogenous variables and $F+1$ instrumental variables

they observe them having a comparative advantage in that same field. This suggests that the local average treatment effect of this instrumental variable strategy might be capturing the returns to choice of field for those students who learn from their instructors about their own comparative advantage and choose a field following it. While the choice of field of the compliers of this instrumental variable will be potentially exogenous, we need to keep in mind that they are likely to perform better in all fields since they enroll in fields for which they have a comparative advantage.

A second alternative instrumental variable consists in estimating teacher teams' suggestion fixed effects leaving student i out and using the suggestions given by each team to student i's classmates and students in the 4 cohorts adjacent to student i's cohort (two earlier and two later cohorts) with same background characteristics as student i. In practice, Estimate probability of field f suggestion by teacher group g to students other than you and to students of the 2 cohorts ahead and behind you with similar background characteristics. In practice, I estimate the model $S^f_{j \neq i, c \pm 2, s, g} = \alpha + \beta X_i + \psi_g + \epsilon_{jcsg}$ separately for each i and f and predict $\hat{S}^f_i = \hat{\alpha} + \hat{\beta} X_i + \hat{\psi}_g$. Finally, I use $\hat{S}^f_i$ to instrument π^f_i. This strategy ideally exploits the fact that certain instructor teams have consistent preferences towards suggesting certain field across cohorts and similar students. Being assigned in the first year of high school to a teacher team that is consistently more likely to suggest enrollment in a law degree to top-performance male students exogenously increase a student probability to choose a law degree 5 years later. This strategy requires students assignment to teacher teams at the beginning of high school to be exogenous with respect to teacher preferences. While institutional rules for the high schools in the period under analysis prescribed school principals to randomize student assignment to classes in the very first year of high school (at around age 13), I do not have explicit tracking of the randomization mechanisms nor their details. Anelli and Peri (2017) present a series of checks for randomness that show no systematic correlation of individual observable characteristics to classrooms using the same exact data. Moreover, it is highly unlikely that students were assigned to classrooms and hence to teacher teams based on their future preferences for specific fields of study at university. I am thus confident that first-year-of-high-school assignment to classrooms and teacher teams is exogenous to teacher preferences for specific fields. While this is my ideal identification strategy, it is also the most demanding. When estimating returns to nine separate fields of study, the first stage is too weak. However, I obtain enough first stage power to estimate returns to four broader fields of study (Humanities, Stem, Natural Sciences, Economics & Business + Law). For certain outcomes I show estimates also from this more demanding specification.

A slightly different instrumental variable strategy that however provides enough first stage power to estimate returns to the nine fields of study, consists in estimating a field choice fixed effects, excluding one's own choice, within instructor team and use it to instrument one's choice. In this case, I estimate the model $\pi^f_{j\neq i,c\pm 2,s,g} = \alpha + \beta X_i + \psi_g + \epsilon_{jcsg}$ separately for each i and f and predict $\hat{\pi}^f_i = \hat{\alpha} + \hat{\beta} X_i + \hat{\psi}_g$. Finally, I use $\hat{\pi}^f_i$ to instrument π^f_i. Intuitively, this instrument captures the fact that within teacher team g students, other than oneself, are persistently more likely to go to a specific field (potentially because of the persistence in field suggestion of the team). The assumptions for this instrumental variable strategy to be valid remain the same as for $\hat{S}^f_i$.

Finally, in some specifications I take advantage of the teacher field suggestion variable in an alternative way. If the suggestion is truly capturing comparative advantages for certain fields that are otherwise unobserved to us, it can be included in the regression as a separate control. Estimates obtained including teacher suggestion as control should thus reflect returns to choice of field net of individual comparative advantages.

4 Main Results

In this section I first present estimates for the long-run payoffs to choice of field and then explore the labor market determinants that explain payoff differentials across fields.

Table 1 presents the main estimates for long-run payoffs to choice of field for the first 26 years after university graduation. Income has been adjusted for inflation using an official deflation index and should be interpreted in 2016 euros and included in the left hand side of the regression in logarithms. The estimates in this table and all following ones can thus be interpreted as percent premium with respect to the average income of the baseline field of study, Humanities. Column 1 reports estimates from my baseline model 1 with school/cohort fixed effects and teacher team fixed effects, controlling for a large set of individual and high school class peers' pre-university characteristics. All fields yield significantly higher payoffs than Humanities with premia ranging between 21% (Architecture & Design) and 82% (Economics & Business). All other fields yield premia similar to these two thus identifying two wider groups of field in terms of payoffs. Law (71%), Medicine (62%) and Engineering (61%) yield payoffs slightly lower then Economics & Business, while Natural Sciences (29%), Social Sciences (31%), Math & Physics (36%) yield payoffs closer to Architecture & Design. I will refer to the former group of fields as the "high-payoff" fields, and to the latter group as the "low-payoff" fields, although their returns are higher than Humanities, which is clearly the lowest-payoff field. In columns 2 to 6 of Table 1 I present estimates for the same premia obtained with different specifications and the instrumental variable strategies described in 3. In column 2 I add to the specification of column 1 teachers' field suggestions to obtain returns to choice of field that should be net of comparative advantages of students choosing them. In column 3 I use teachers' field suggestion S_i^f as instrument for the actual choice, while in column 4 I use the predicted choice $\hat{\pi}_i^f$ as instrument. The IV estimates appear to be generally slightly higher than the baseline model, but the overall picture presented by the baseline model remains true. Two exceptions are Math & Physics and Social Sciences, with the former one yielding substantially higher premia (70-80%) and the latter yielding smaller and no more significant premia in column 4 and 5. It is thus possible that although the baseline model seem to do quite well in capturing selection on observables, for these two fields in particular there seem to be stronger unobserved dimensions of selection. For all the IV estimates in columns 3-5 the first stage Cragg-Donald F statistic is high (ranging from 148 to 811) confirming a strong enough first stage power for valid IV estimation. Finally, in column 6 I present results for my most demanding identification strategy using the instrument $\hat{S}_i^f$, presented in section 3, however the first stage for this instrument is extremely low (F is 3.8) when instrumenting the nine different fields. For this reason, in Table 2 I replicate my baseline and the IV models, by focusing only on 4 aggregated categories of study fields to obtain enough first-stage power for the most demanding IV specification. Column 6 in Table 2 has indeed a first stage F statistic of 95 which allows for a valid IV estimate also using $\hat{S}_i^f$ as instrument. When comparing estimates of column 6 to the baseline model estimates, $\hat{S}_i^f$ appears to deliver slightly higher returns for STEM fields and Natural Sciences with respect to column 1 and very similar returns for Economics & Business + Law broad category. With respect to the instrument $\hat{\pi}_i^f$, $\hat{S}_i^f$ seem to deliver similar, but slightly higher returns. I interpret this as evidence that estimates obtained with $\hat{\pi}_i^f$ (the one instrument that allows identification for the nine fields) should be the more conservative. In presenting the remaining results I will thus focus on the IV estimates obtained using the instrument $\hat{\pi}_i^f$. Figure 3 provides a visual representation of the four main payoff estimates from Table 1 to allow for a more direct comparison across fields. In the rest of the paper, I present similar complement figures for most analysis.

Since estimates on income returns are estimated using the logarithm of income as dependent variable, they are by definition conditional to observing positive income. On average long-run employment rate for the Milan university graduates in my sample is very high (above 90% for this sample). However, in the first years right after graduation the there might still be substantial differences across fields in the time it takes to find a stable job. The possibility of tracking the labor market status of graduates over time, allows me to estimate whether graduates of different fields have differentials probabilities of having zero income yearly spells between university graduation and the 26^th year on the labor market. Figure 4 and Table 3 show differences in the probability of observing non-zero income spells in the first 26 years in the labor market. With respect to students choosing Humanities (who have on average 64% non-zero income spells in the first 25 years after graduation), high-payoff fields (Law, Medicine, Engineering, Economics and Business) have between 8 and 25 percentage points higher probability to have non-zero income spells, with Medicine delivering the highest probability of having non-zero income spells (with an average total of 90% of all spells in the first 25 years). Other fields do not guarantee a significantly higher probability of employment. Since the low-payoff fields of Figure 3 and Table 1 are also those with a higher probability of having zero-income spells, payoff estimates that included zero income spells would certainly amplify the income differentials across fields.

The nice feature of the social security data available for this analysis is the possibility of following labor market outcomes for many years after university graduation. This allows me to study career trajectories by field. To do so I estimate my baseline specification 1 income returns to field choice separately for every post-graduation labor market year (e.g. in the 10^{th} year on labor market).

$$\forall t Log(y_{icsg}^{t=1,\dots,15}) = \alpha + \pi_i^f + \beta X_i + \gamma Z_{-i} + \phi_{sc} + \psi_g + \epsilon_{icsg} \tag{2}$$

Since conditioning on single labor market year reduces the sample substantially, in order to avoid that idiosyncratic yearly shocks to income of individuals choosing certain fields, I use the logarithm of the three-year moving average of income as dependent variable. Moreover, given that for years $t > 15$ I have only a limited number of cohorts available by construction of the dataset, I focus only on the trajectories in the first 15 years after university graduation. Figure 5 presents the trajectories obtained applying the premium of each field (estimated with model 2) to the baseline income of humanities in every year after university graduation. The figure reveals very interesting heterogeneities in the income trajectories of different fields. We can identify three main patterns. Payoffs to Economics & Business remain always above those of other fields, while payoffs to Humanities are always the lowest. The trajectory of Math & Physics is somewhat in the middle between this two extremes. An interesting fact emerging from Figure 5 is the very different trajectories of some high-payoff fields. For instance, Engineering which does very well right after university graduation (and is for this fact notoriously considered the highest payoff field) have a quite flat trajectory over years on the labor market. To the contrary, Law and Medicine start off substantially lower than Engineering and have initial payoffs similar to Social Sciences have very steep trajectories (almost linear) in the first 15 years on the labor market and cross the Engineering trajectory already around the 10^{th} year after university graduation. All the remaining fields have some degree of heterogeneity in the first years after graduation, but seem to converge to the Math & Physics payoffs after the 10^{th} year of labor market. While the trajectories in Figure 5 are estimated using the baseline fixed effect model 1, it can also be constructed using IV estimates. However, the smaller sample size makes the first stage estimates for some years t weak and estimates are a bit less reliable.

4.1 Labor market determinants of the payoffs to the choice of fields

After estimating returns to choice of field, I take advantage of the richness of social security records available, to explore which are the labor market events that can explain such large differentials in payoffs to fields. In Figure 6 and Table 4, I study whether some fields grant some kind of "lottery-ticket effect" by yielding higher probability to reach the top 1% of the Italian income distribution already in the first 26 years on the labor market (the time window I can observe). While only 1% of individuals choosing Humanities reach the top 1%, this probability is substantially higher for Economics & Business and Law (around 10 percentage points higher). For all other fields the probability is at most 2 percentage points higher than Humanities, but often not significant. This evidence is consistent with the exceptional performance of Law, Economics & Business especially in the long-run and with the rather flat trajectory of Engineering that appears not to deliver a "lottery-ticket effect". Strictly linked to the probability of reaching the top 1% of the income distribution, is the probability of becoming a top manager. Table 5and Figure 7 show that in this case it is Medicine, Engineering and Economics & Business that guarantee substantially higher probabilities of becoming top managers (lawyers typically have self-employment careers and thus do not really cover top management jobs).

On the opposite side of the income distribution, individuals of certain fields might face higher probabilities to face very negative labor market outcomes. Tables 6,7 and 8 and Figures 8,9 and 10 explore the probabilities by field of observing labor market spells with income below the poverty line[6], unemployment spells and the probability of having a fixed-term contract. Overall, this analysis depicts a rather negative situation for individuals choosing low-payoff fields with Humanities yielding on average 15% of all spells in the first 26 years of labor market experience with income below the poverty line and 41%

[6]I refer to the absolute poverty line calculated by ISTAT for urban areas of northern Italy

of all spells with fixed-term contracts. The probability of receiving an income below the poverty line is virtually driven down to zero for all high-paying fields and the probability of having a fixed-term contract goes down a lot for Engineering and Medicine, but not as much for Economics & Business. Table 8 focuses instead on the probability of receiving unemployment benefits, however this outcome should be interpreted with caution, because eligibility for unemployment benefits is conditional to having a permanent contract. Since a large part of the careers in low-payoff fields are in fixed-term contracts, their probability of observing receiving unemployment benefits is rather low. Despite this, it is clear that all high-payoff fields deliver a significantly lower probability of receiving unemployment benefits. Overall, this focus on negative labor market outcomes characterizes low-payoff field careers as penalized by highly discontinuous careers that drive down income (even below the poverty line in certain cases). This evidence is also consistent with the very flat career trajectories of these fields shown in Figure 5

4.2 Over Education as measure of excess supply of some fields

The strong evidence of negative labor market perspectives for low-payoff fields just presented can already be interpreted as a sign of excess supply of graduates of certain fields. In this section, I construct a more robust measure of excess supply to confirm the presence of inefficiencies in the market for university graduates. I take advantage of the OECD PIAAC survey data for Italy and calculate for each "sector/occupation cell", the percentage of respondents who answers that no college degree is required for her/his job. I then matched each labor market spell in my data to these cells and classified as over-educated those individuals who are in a "sector/occupation cell" that does not require a college degree (e.g. all respondents in that cell answers that no college degree was required). Table 9 and Figure 11 presents the results of this analysis. I find that a much larger fraction of employment spells in the first 26 years on the labor market of low-payoff fields such as humanities (33%) , architecture and design (40%) are in sector/occupation cells that do not require a university degree. High-payoff fields such as Law, Medicine and Economics & Business have substantially lower probability to be over-educated on the job. Interestingly, also fields that are relatively higher payoff than Humanities such as Math & Physics and Natural Sciences do have significantly lower probability (around 7-8 percentage points) than Humanities to be over-educated. I interpret this as a clear sign of excess supply of graduates from Humanities, Architecture & Design. Moreover, the patterns of this overeducation outcomes appear to strictly map into the payoffs differentials estimated across fields. This reinforces the hypothesis of a mismatch between demand and supply of university graduates from different fields.

4.3 Individual heterogeneity of payoffs to field choice

Finally, I explore possible heterogeneity in the payoffs to field choice along two individual dimensions such as gender and parental background. For instance, I explore whether some fields have higher penalties for women and whether high-payoff fields are as rewarding for individuals with a low SES background or whether instead do not guarantee an equally high probability to reach the top 1% for individuals with a low SES background. Figure 12 plots the value of an field-women interaction dummy. Results show that although all fields seem to have an income penalty for women, Law and Medicine (and to a lesser extent Engineering, Math & Physics) appear to penalize women substantially more. Figures 13 and 14 instead plot the estimates of a field-low-SES interaction (after linking social security records for fathers, I split the parental sample in below- and above-median income groups). Interestingly, children of low-income fathers appear to do substantially worse than their high income background peers in the same field if they choose Architecture & Design or

Law, while payoffs for all other fields seem to be pretty much independent from parental background. This evidence is consistent with the one in Figure 14 showing a similar pattern for the probability of reaching the top 1% of the income distribution. Architecture & Design or Law again show strong intergenerational income persistence.

5 Do students learn from the differentials across fields?

If we believe that the large differentials across fields estimated are the results of suboptimal individual choices and students were to have some information about these differentials, we should expect them to adjust their choices consequently, reducing the excess of supply in the low-paying fields. This would ultimately drive down differentials across fields. In this second main section of the analysis I explore this hypothesis, by estimating my main specification 1 separately for three different cohort windows (cohorts graduating from high school between 1985 and 1989, 1990-94 and 1995-2000) and considering only the first 10 years on labor market for each cohort window.[7] Table 10 shows that fields that were already high-payoff in the earliest cohorts have even higher premia with respect to Humanities in the latest cohorts. On the other hand, relatively low-paying fields that yielded significantly more than Humanities in earlier cohorts (e.g. Social Sciences, Natural Sciences, Architecture & Design) appear to have converged towards Humanities over time. Overall, Table 10 certificates a divergence between high- and low-payoff fields over this period. This appears to exclude the possibility of an optimal adjustment of students' choice of field over time. To complement this evidence, Figure 15 plots the share of students in my sample enrolling each field over time. Surprisingly, although payoffs to Humanities were already the lowest for the earliest cohorts, the share of students enrolling in Humanities in my sample have

[7] Income has been adjusted for inflation using an official deflation index provided by the Italian Statistical Office ISTAT and should be interpreted in 2016 euros.

doubled in the same period. This can explain the further drop in payoffs to Humanities. Taken altogether, the evidence of Table 10 and 15 describe a situation that is reversed with respect to the hypothesised learning story. Payoffs across fields diverge over time and students seem not to respond to the excess supply of graduates from low-payoff fields. This can only be explained by a lack of information on returns to field or by a possible increase over time of non-monetary returns of low-paying fields. Since there is no evidence of this latter hypothesis, and no change in non-monetary returns to Humanities can justify a doubling of enrollment in this field, I interpret these results as compelling evidence of a lack of information in the field choice process of students.

5.1 Volatility and Risk in Payoffs to Field Choice

One reason that could justify persistence in choice of field and in payoffs is risk and volatility. Risk averse students might be optimally trading off higher payoffs for lower volatility. I explore this hypothesis by incorporating payoff volatility in my analysis. To do so, I estimate sample payoff volatility after excluding volatility attributable to background characteristics. In practice, I estimate the following model, including dummies θ_t for each different year on labor market after university graduation, school and teacher group dummies:

$$y_{itcsg} = \alpha + \beta X_i + \gamma Z_{-i} + \phi_s + \psi_g + \theta_t + \epsilon_{ictsg} \tag{3}$$

I interpret the residuals $\epsilon_{\hat{ictsg}}$ of this model as deviation from mean sample income in each post-graduation year, adjusted for pre-university background characteristics. I then calculate payoff volatility as the standard deviation σ_ϵ of $\epsilon_{\hat{ictsg}}$ within field and post-graduation year t. In Figure 16 I plot the average of $\epsilon_{\hat{ictsg}}$ by field and post-graduation year together with a confidence interval computed using σ_ϵ. Some of the high-payoff fields clearly have higher volatility (i.e. Law, Business & Economics), while low-payoff fields have substantially lower volatility (i.e. Humanities). However, 95% confidence intervals of high- and low-payoff fields around payoff trajectories, which should be interpreted as deviations from sample mean income in labor market year t after university graduation, never overlap. For instance, the lower confidence interval for Law, Economics & Business in the long-run remains more €15,000 above the average sample income, while the upper confidence interval of Humanities is more than €20,000 below average sample income.

An alternative way to incorporate risk considerations in the payoffs estimation is to compute sharp-ratios, which I construct from the previous figure as the average deviation from mean income by year t divided by its standard deviation, $\epsilon_{i\bar{\hat{c}}tsg}/\sigma_\epsilon$ Figure 17 plots volatility-adjusted payoffs $\epsilon_{i\bar{\hat{c}}tsg}/\sigma_\epsilon$ by choice of field. Differentials of risk-adjusted payoffs across fields appear more compressed (most of the difference between high-payoff fields disappear), however large differences remain across low-payoff and high-payoff fields, with Humanities still performing poorly, despite of the very low volatility. Overall, Figure 17 and 16 appear to justify only a limited role of risk in the optimal choice of field. Payoffs to Economics & Business or Law indeed appear to be an optimal choice also for risk averse individuals.

An alternative risk-related consideration in the choice of field might be driven by relative insulation from the business cycle. I have explored this hypothesis, by exploring whether returns of individuals choosing low-payoff fields are penalized less by graduating in a year of economic recession. Figure 18 shows a rather high degree of heterogeneity in response to economic recession at time of graduation. Individuals choosing Humanities appear to be indeed insulated from the risks of entering the labor market during an economic downturn, however other low-payoff fields (such as Social Sciences, Architecture and Design) do not appear to guarantee the same degree of insulation, with individuals from Social Sciences earning on average around €6000 less if they graduate during an economic downturn. Among high-payoff fields, Engineering, Economics & Business only have a small not significant penalization for entrance in the labor market , while payoffs of Law and Medicine interestingly appear to be counter-cyclical (higher payoffs if graduation happens during an economic downturn)

6 Conclusions

I exploit Italian administrative data from high schools and universities matched to social security records and track every single labor market episode of university graduates up to 25 years after university graduation. I estimate long-run payoffs to choosing nine fields of study, analyze the labor market determinants of the heterogeneity in payoffs across fields, whether this heterogeneity persists over cohorts and explore the reasons why these differences persist. I find large differences in long-run returns to choice of field with Economics & Business delivering almost 100% higher returns with respect to Humanities (the lowest paying field). Income trajectories also differ substantially across fields with Engineering starting among the highest paying right after college, but having a flatter trajectory than fields such as law or medicine, which start lower and surpass engineering after about ten years from graduation. While Economics & Business remain the highest trajectory , humanities is the lowest in every year after graduation. The most relevant labor market determinants of the differential payoffs to field are a higher probability to reach the top 1% of the income distribution for high-paying fields and a substantially higher probability of having a discontinuous career (fixed term contracts and unemployment spells) for the low-paying fields. I also link sector/occupation cells of destination with OECD PIAAC survey data to estimate the degree of "over-education" for graduates of different fields and find that a much larger fraction of employment spells in the first 26 years on the labor market of fields such as humanities (33%) , architecture and design (40%) are in sector/occupation cells that do not require a university degree. I interpret this as a clear sign of excess supply of graduates from these fields. I also explore differential returns to fields across gender and parental income. While returns for women appear to be lower for all fields, Social Sciences and Engineering have smaller penalties. Returns to field appear to depend also on family background with returns to Law, Architecture and Design being substantially lower for graduates with fathers having an income below median. When I explore how returns vary over cohorts, I find that even though the Humanities had the lowest payoffs already for the earliest cohorts, its enrollment in later cohorts increased substantially, and its payoffs kept dropping over cohorts. Instead of converging, payoffs diverged over time. This clearly excludes the possibility of an efficiency-improving adjustment in the choice of field over time and can only be explained by a lack of information on returns to field or by a possible increase over time of non-monetary returns of humanities. When I check for the possibility that the persistence in choice of major is driven by a high returns-high volatility tradeoff, I do not find strong empirical evidence to support this hypothesis. It is true that some of the high-payoff fields have higher volatility (i.e. Law, Business & Economics), while low-payoff fields have low volatility (i.e. Humanities), however confidence intervals around payoff trajectories, calculated taking sample volatility into account, never overlap. Field differentials of risk-adjusted returns are more compressed (most of the difference between high payoff fields disappear), however differences remain across low-payoff and high-payoff fields. This evidence appears to justify only a limited role of the high risk-high return trade-off in the optimal choice of field. Other risk-related considerations in the choice of field might be driven by relative insulation from the business cycle. I have explored this hypothesis and I found that graduates of high paying fields do not suffer long-run penalization for entering the labor market in a year of economic recession, contrary to graduates of social sciences (a low-payoff field) who appear to be the most penalized by graduating during an economic downturn.

While these findings are specific to the institutional and empirical context under study, they contribute to the literature by offering a first detailed analysis of the labor market mechanisms driving the strong heterogeneity in the payoffs to different university fields and by exploring whether and why field differentials persist over time. Moreover, given that most recent causal evidence on returns to field has been drawn by specific countries and institutional contexts (e.g. Chile and Norway), I contribute by offering evidence on returns to field choice for a large European country with a university admission system that is more representative for continental Europe. The evidence presented in this paper has important insights also for policy. The clear presence of an excess supply of graduates of the lowest-payoff fields and an increase over time of enrollment in those same fields, that can not be explained by an increase in non-monetary payoff or risk-related considerations, suggests the presence of potentially large inefficiencies in the choice of fields of Italian students. Lack of public information on the payoffs and labor market prospects of each field is the first candidate to explain such inefficient choice process. Absent information on payoffs by field, in an institutional setting in which university tuition fees are highly subsidized, students might choose fields based solely on individual preferences for specific disciplines. Given that the Italian university system is by a large majority State-run, it should be fairly inexpensive to aggregate information on payoffs and labor market prospects by fields and provide it to students, families and high school teachers at the time of enrollment at university.

References

[1] Altonji, Joseph, Erica Blom and Costas Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," Annual Review of Economics, vol. 4(1): 185-223, 07.

[2] J.G. Altonji, P. Arcidiacono, A. Maurel, Chapter 7 - The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects, Editor(s): Eric A. Hanushek, Stephen Machin, Ludger Woessmann, Handbook of the Economics of Education, Elsevier, Volume 5, 2016, Pages 305-396.

[3] Andrews, R.J., Li, J. & Lovenheim, M.F. 2012, Quantile Treatment Effects of College Quality on Earnings: Evidence from Administrative Data in Texas, National Bureau of Economic Research, Inc, NBER Working Papers: 18068.

[4] Arcidiacono, Peter. 2004. Ability sorting and the returns to college major. Journal of Econometrics, 121: 343-375

[5] Autor, David H., Lawrence F. Katz and Melissa S. Kearney. 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. Review of Economics and Statistics, 90(2): 300-323.

[6] Bey, M., Fougere, D. & Maurel, A. (2012), 'Choosing the field of studies in postsecondary education: Do expected earnings matter?', Review of Economics and Statistics 94, 334-347.

[7] Bound, John and George Johnson. 1992. Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations. American Economic Review, 82(3): 371-392.

[8] Chevalier, A. (2011), 'Subject choice and earnings of uk graduates', Economics of Education Review 30, 1187-1201.

[9] Colonna, Fabrizio. "2007 Labor market and Schooling Choice: Italy vs US". Mimeo

[10] Del Rossi, Alison F. and Hersch, Joni, Double Your Major, Double Your Return?. Economics of Education Review, Forthcoming; Vanderbilt Law and Economics Research Paper No. 07-09.

[11] Goldin, Claudia and Lawrence F. Katz. 2007. Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing. Brookings Papers on Economic Activity, 38(2): 135-168.

[12] Hamermesh, D. & Donald, S. (2008), 'The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias', Journal of Econometrics 144(2), 479-491.

[13] Hastings, J.S., Neilson, C.A. & Zimmerman, S.D. 2013, Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile, National Bureau of Economic Research, Inc, NBER Working Papers: 19241.

[14] Katz, Lawrence F. and Kevin M. Murphy. 1992. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. Quarterly Journal of Economics, 107(1): 35-78.

[15] Kinsler, J. & Pavan, R. (forthcoming), 'The specicity of general human capital: Evidence from college major choice', Journal of Labor Economics .

[16] Kirkeboen, L, Leuven, E, & Mogstad, M 2014, 'Field of Study, Earnings, and Self-Selection'. NBER working paper series, wp. 20816.

[17] Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. 1993. Wage Inequality and the Rise in Returns to Skill. Journal of Political Economy, 101(3): 410-442.

[18] Oreopoulos, P. & Petronijevic, U. 2013, Making College Worth It: A Review of Research on the Returns to Higher Education, National Bureau of Economic Research, Inc, NBER Working Papers: 19053.

[19] Webber, D. (2014), 'The lifetime earnings premia of different majors: Correcting for selection based on cognitive, non-cognitive, and unobserved factors', Labour Economics 28, 14-23.

Figure 1: Data Timeline

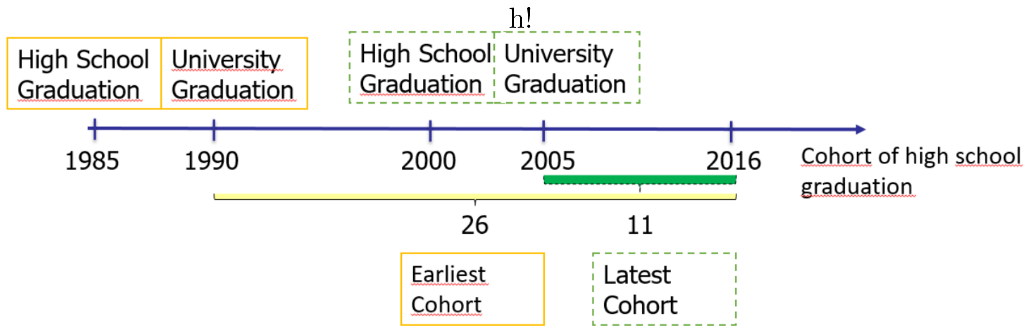
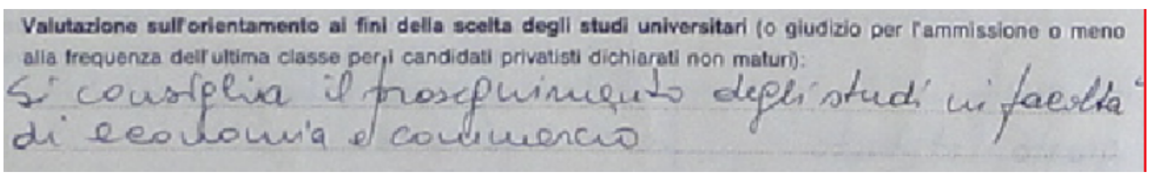
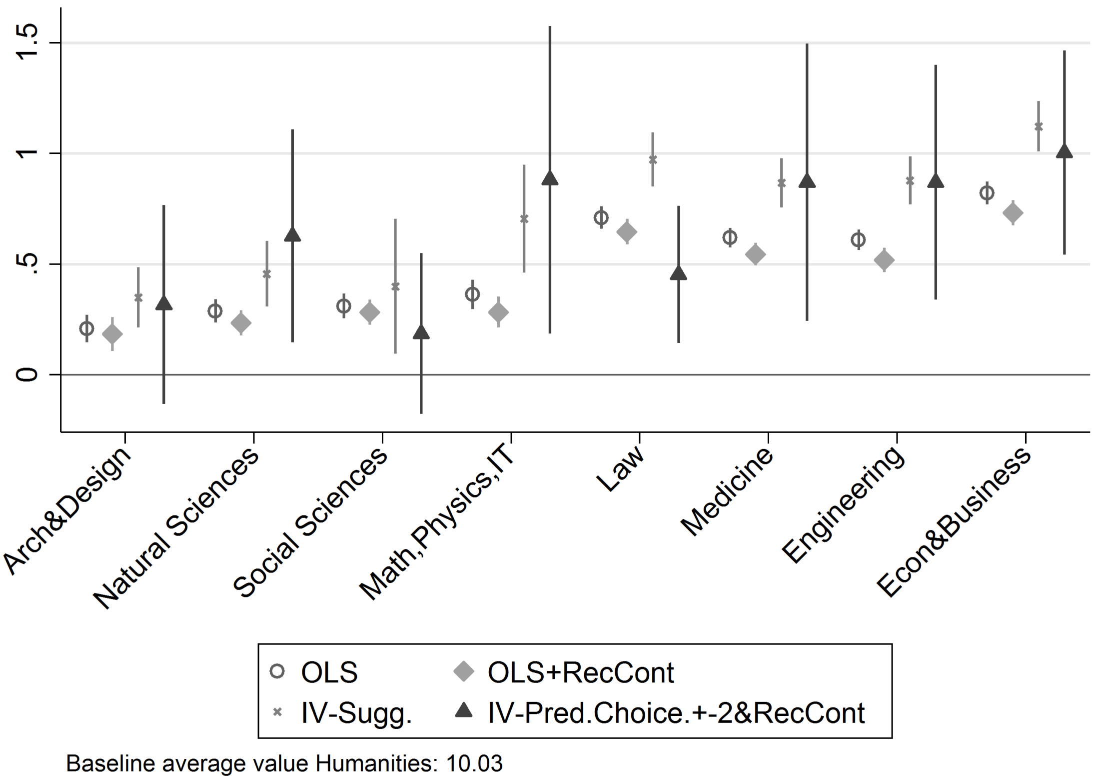


Figure 2: Example of suggestion from digitized hard copy registry


Valutazione sull'orientamento ai fini della scelta degli studi universitari (o giudizio per l'ammissione o meno alla frequenza dell'ultima classe per i candidati privatisti dichiarati non maturi):
Si consiglia il proseguimento degli studi in facoltà di economia e commercio

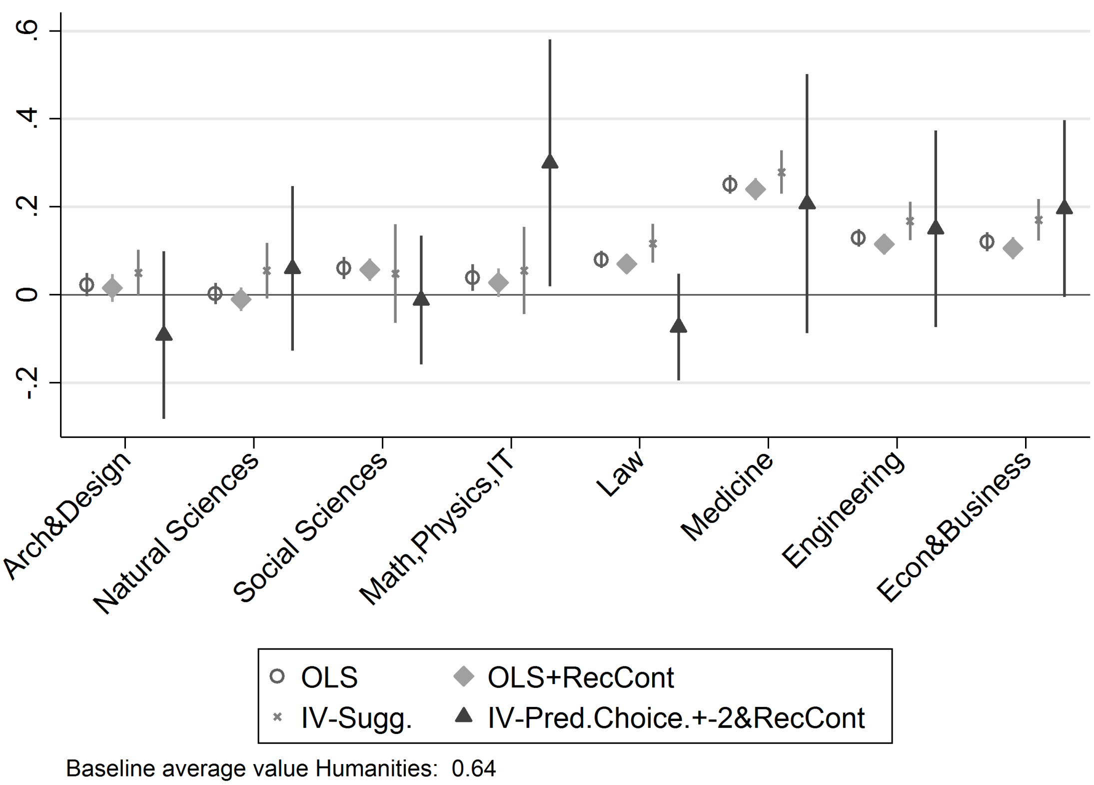
Figures

Figure 3: Payoffs to choice of field 1-26 years on labor mkt - Log of income



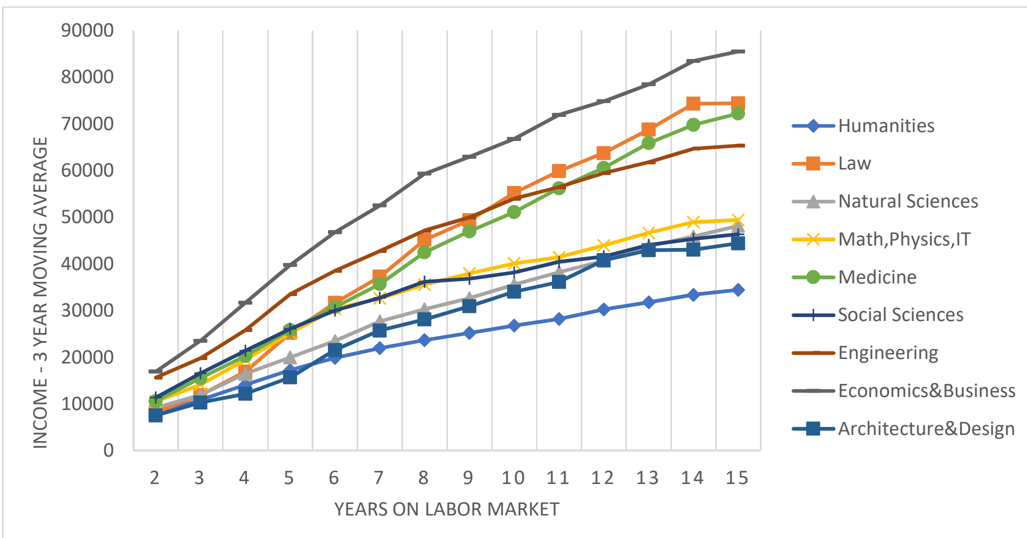
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 4: Probability of observing non-zero income 1-26 years after university graduation



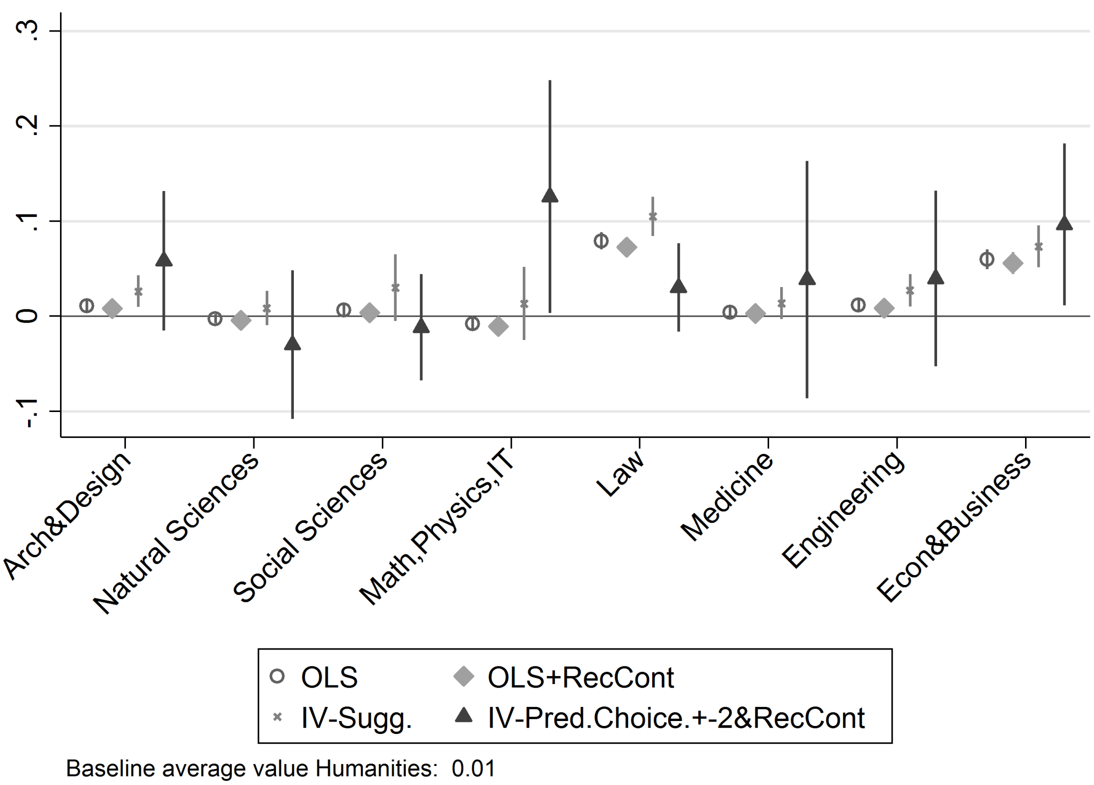
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 5: Income trajectories by field of study 0-15 years on labor market (OLS)



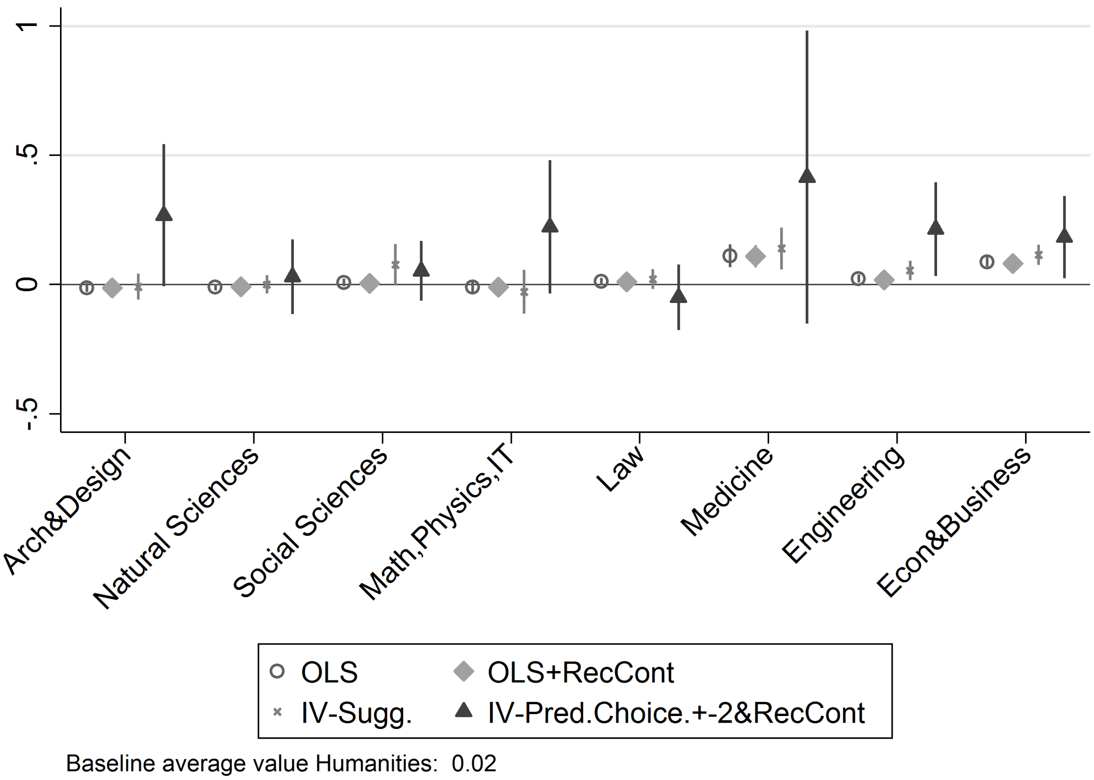
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 6: Lottery ticket effect - Probability of reaching top 1%



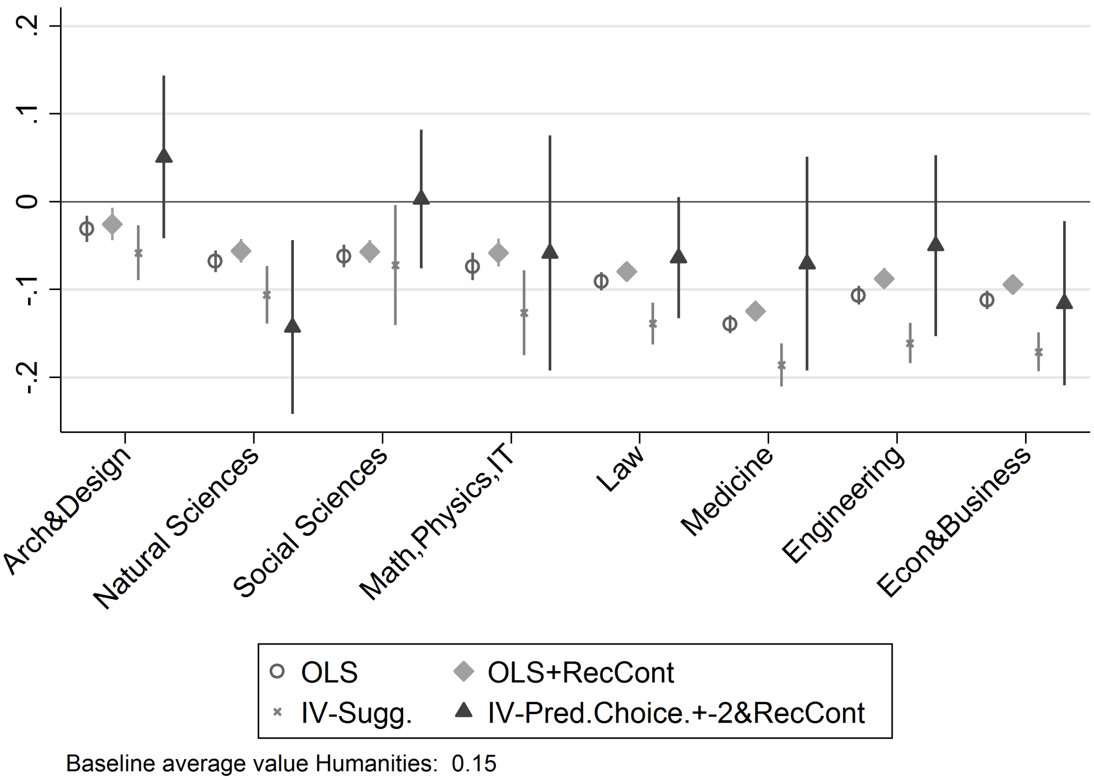
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 7: Probability of becoming a top manager



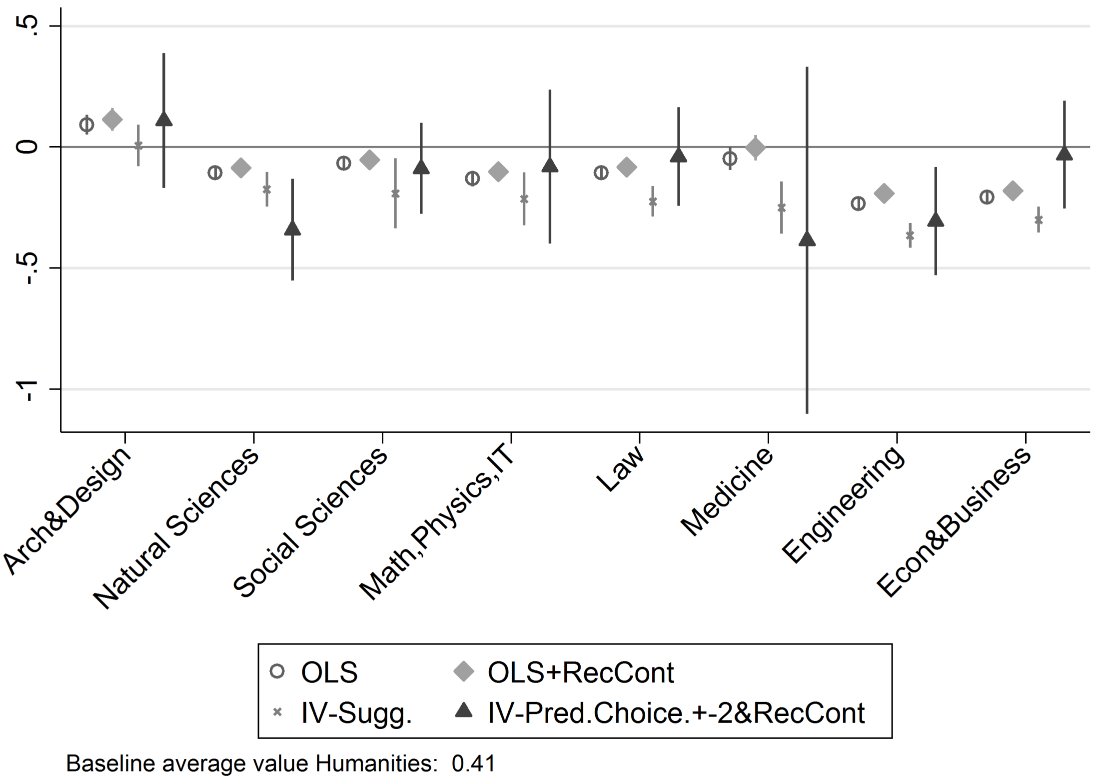
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 8: Probability of having income below the poverty line - 0-25 years on labor market



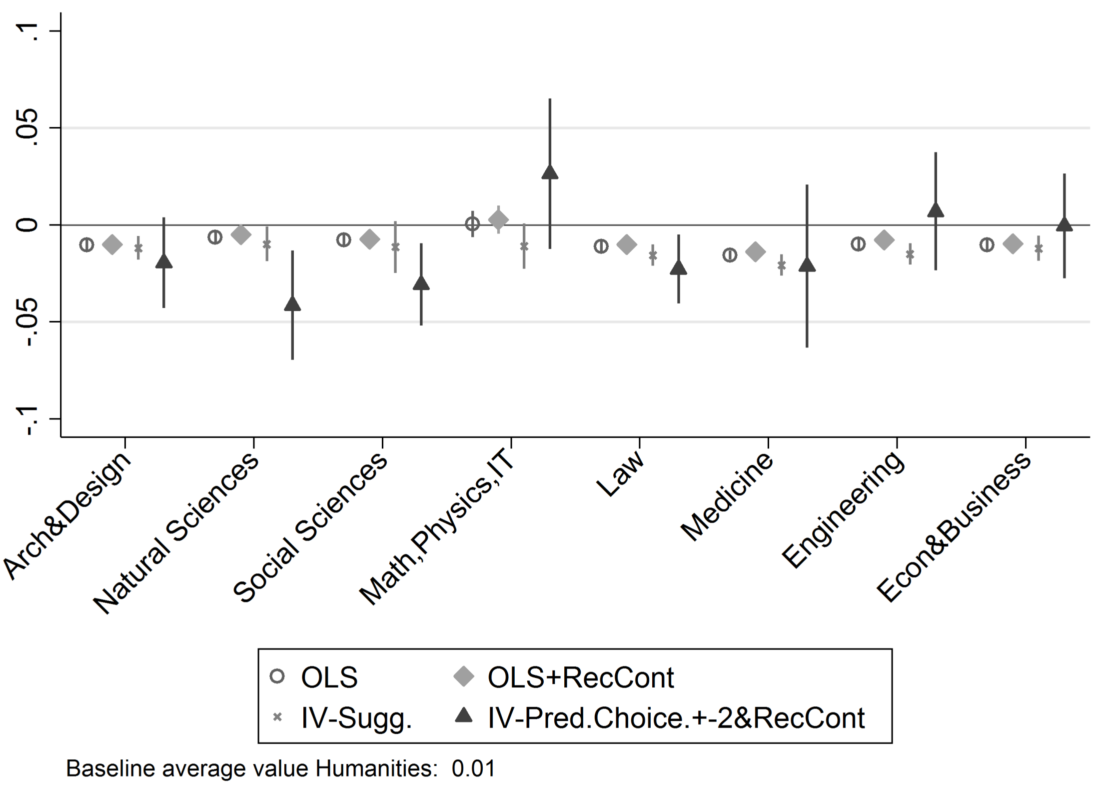
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 9: Probability of having a fixed-term contract



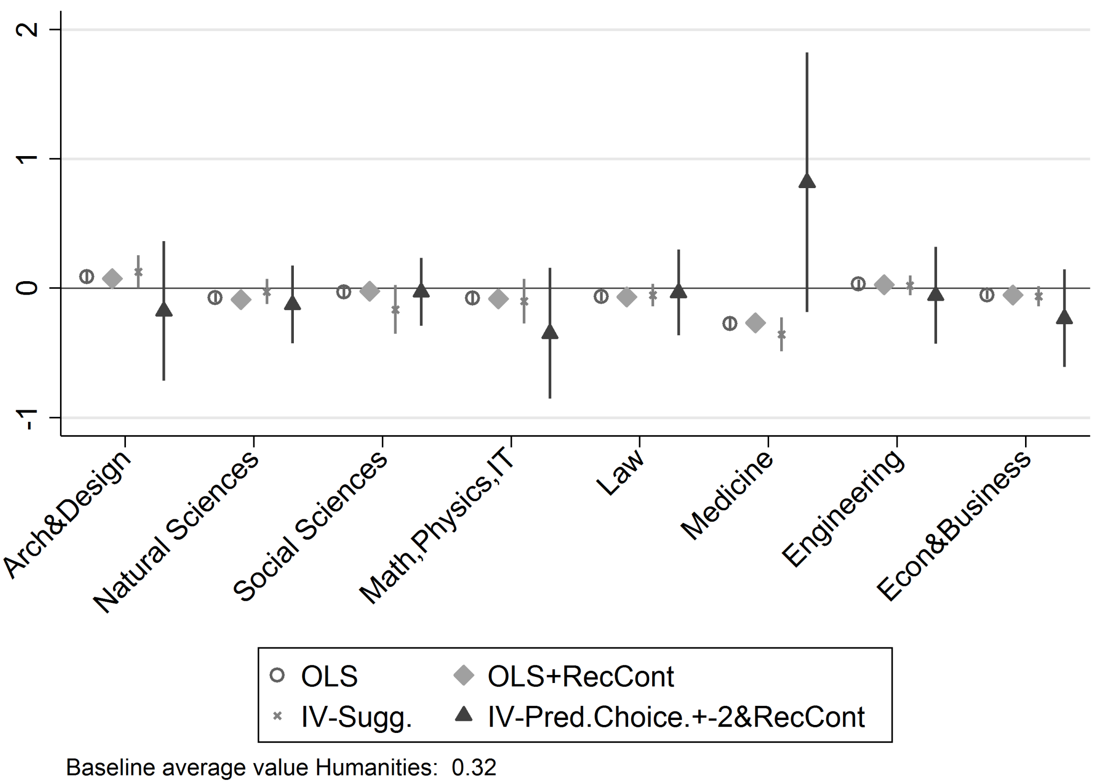
Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 10: Probability of receiving unemployment benefits



Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 11: Probability of working in a sector/occupation cell that does not require a college degree - 0-25 years on labor market



Note: Estimate 1 controls for individual and peer characteristics, school/cohort and teacher group fixed effects, Estimate 2 like 1 but controlling for teacher suggestion, 3 **IV1** - Own suggestion S_i^f as instrument for own choice π_i^f, 4 **IV2** Own choice instrumented with $\hat{\pi}_i^f$ controlling for teacher suggestion S_i^f

Figure 12: Log of annual income - interaction of field choice x dummy female

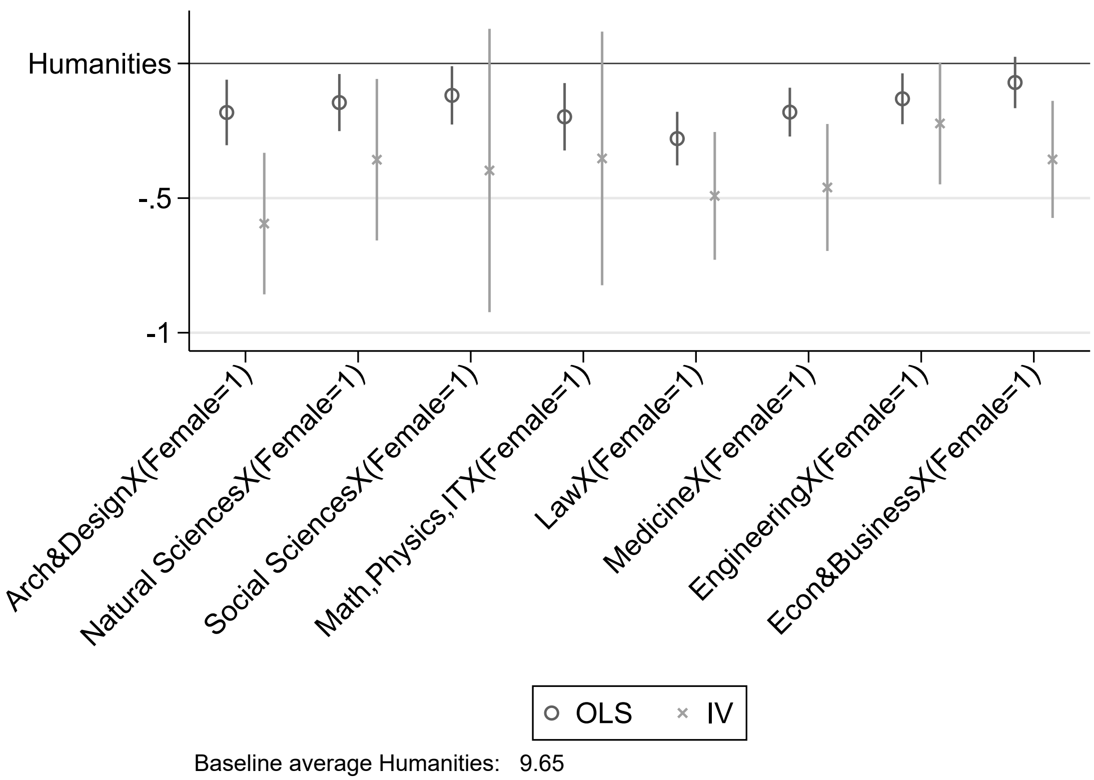


Figure 13: Log of annual income - interaction of field choice x dummy father has income below median

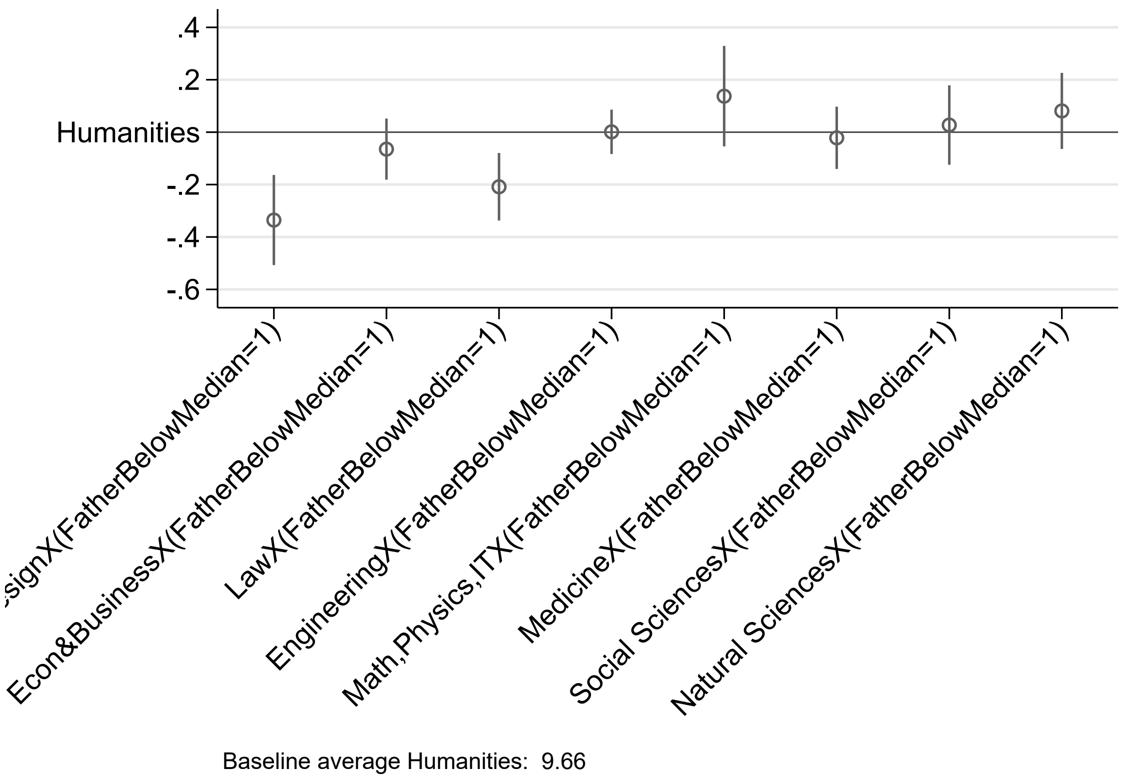


Figure 14: Probability of reaching top 1% - interaction of field choice x dummy father has income below median

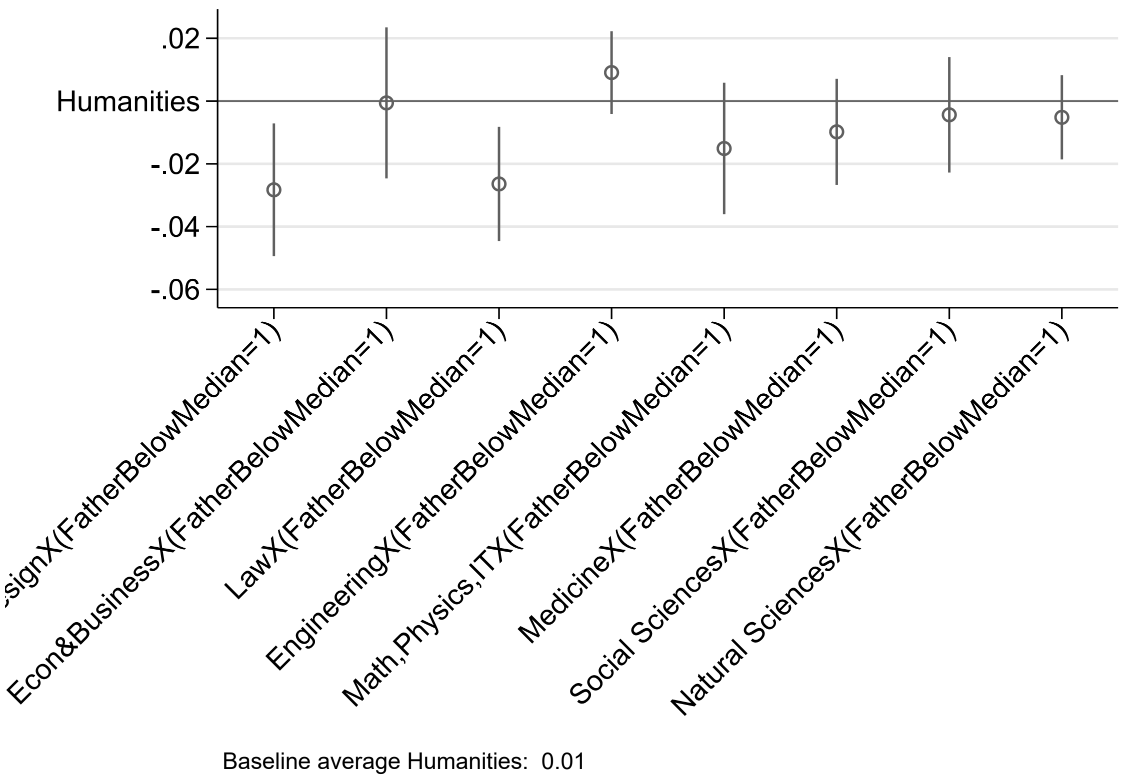


Figure 15: University Enrollment time series by Field

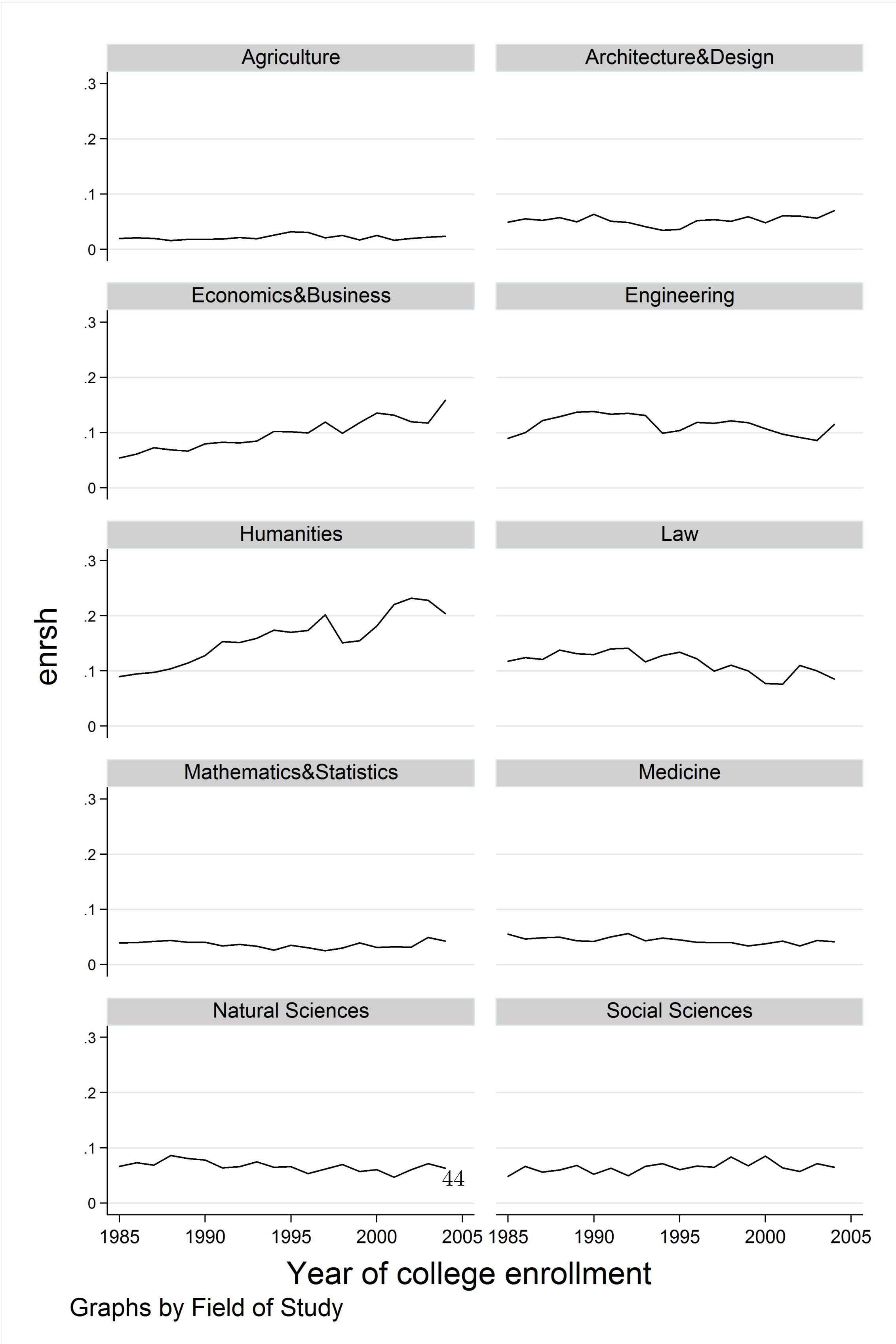


Figure 16: Payoff Volatility by Field and Experience

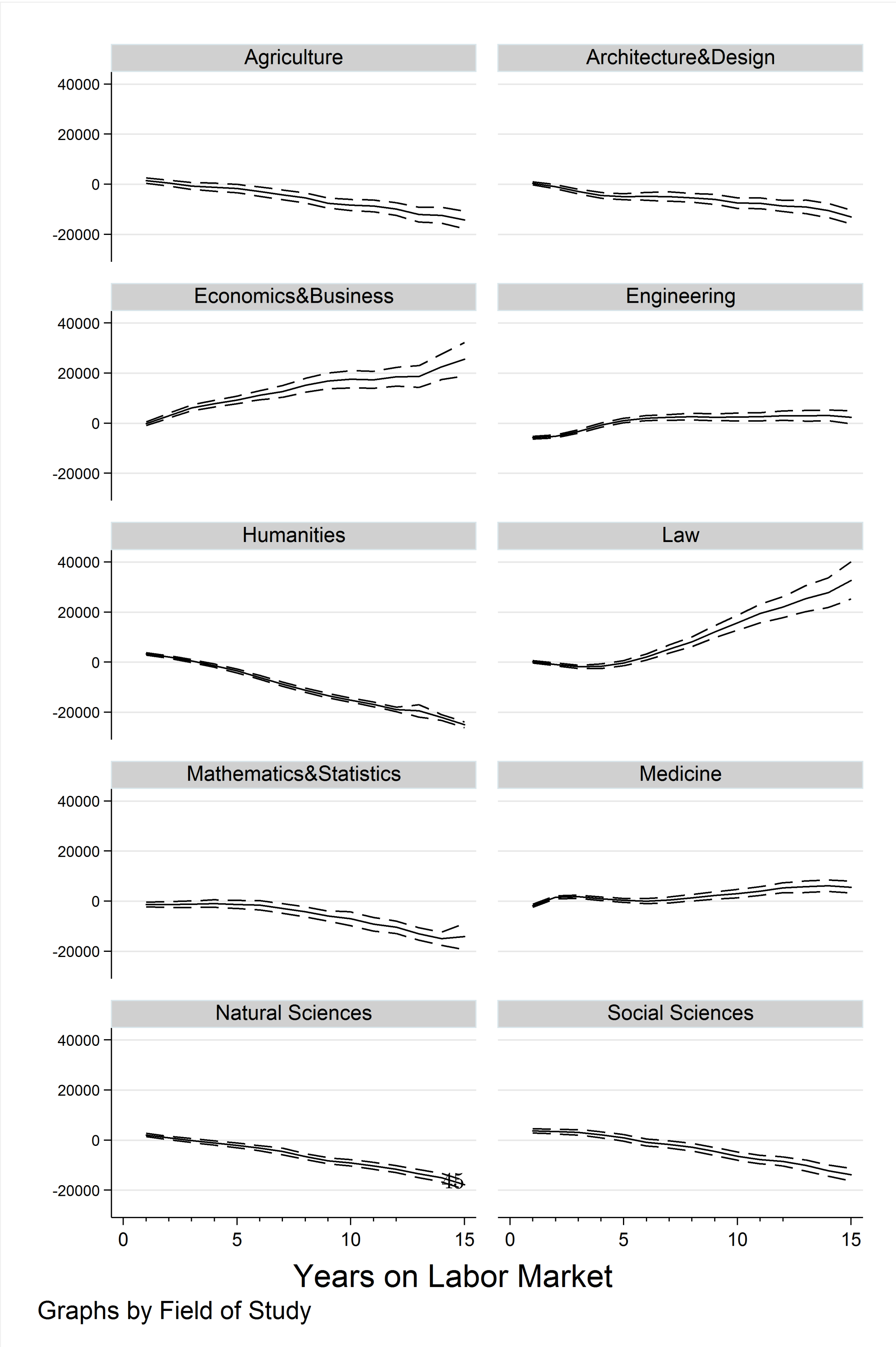


Figure 17: Risk-adjusted Payoffs (Sharp-ratios) by Field and Experience

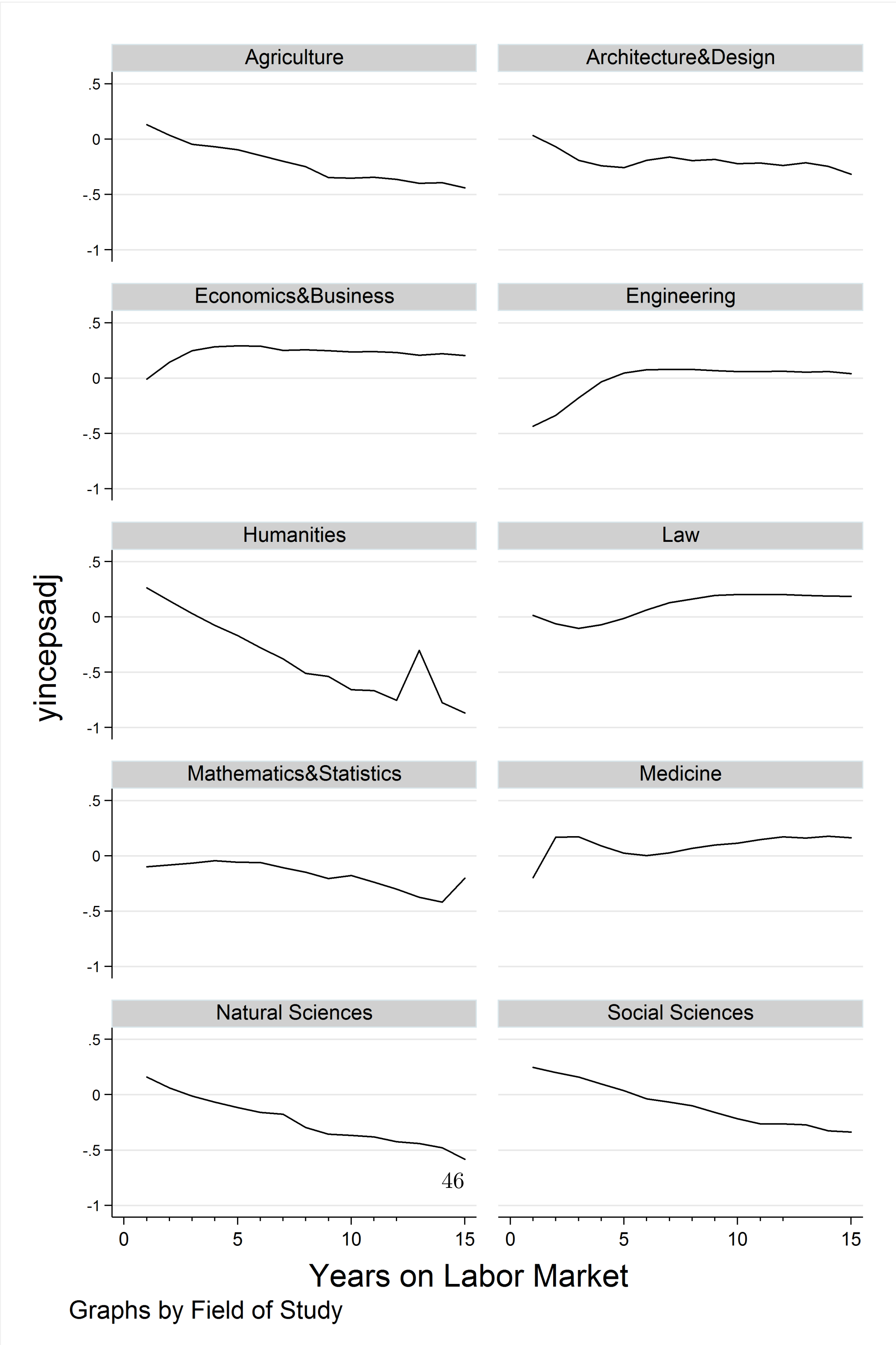
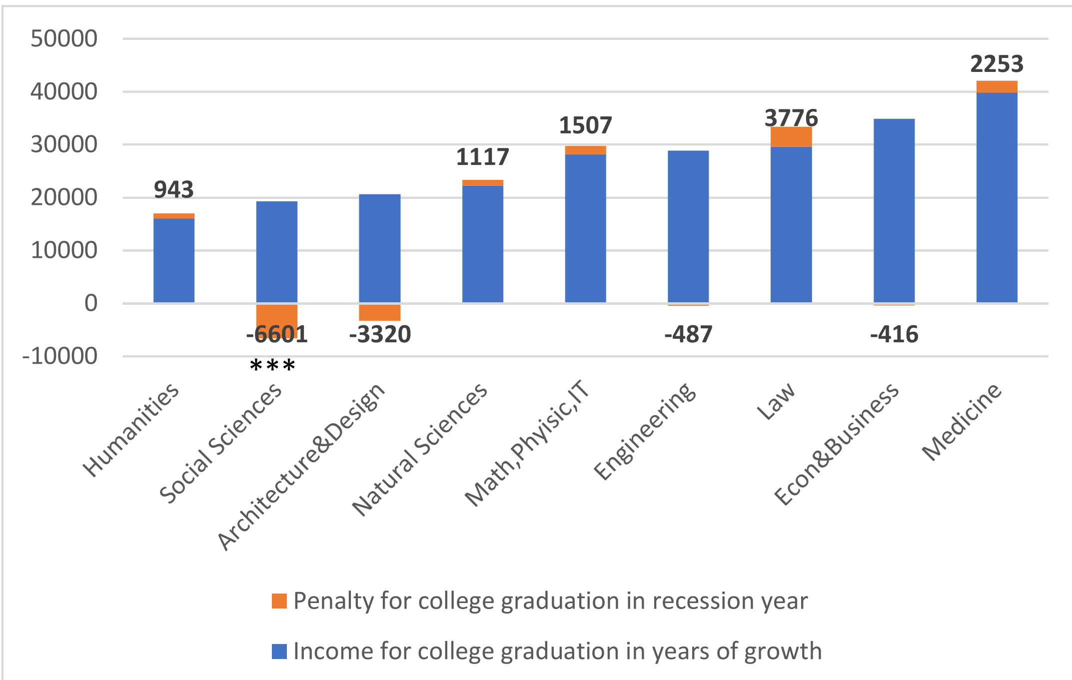


Figure 18: Penalty for entering labor market in recession years



Tables

Table 1: Payoffs on labor market 1-26 years after university graduation - Log of income

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion S_i^f	(4) IV Predicted Choice $\hat{\pi}_i^f$	(5) IV Predicted Choice $\hat{\pi}_i^f$	(6) IV Predicted Suggestion $\hat{S}_i^f$
Arch&Design	0.209***	0.184***	0.350***	0.296*	0.317	0.730
	(0.032)	(0.039)	(0.070)	(0.152)	(0.229)	(0.775)
Econ&Business	0.822***	0.732***	1.123***	1.027***	1.004***	-0.165
	(0.026)	(0.029)	(0.058)	(0.168)	(0.235)	(1.070)
Law	0.711***	0.647***	0.973***	0.551***	0.454***	0.423
	(0.026)	(0.029)	(0.062)	(0.131)	(0.158)	(0.980)
Engineering	0.610***	0.519***	0.878***	0.796***	0.870***	-0.129
	(0.024)	(0.028)	(0.056)	(0.177)	(0.271)	(1.014)
Math,Physics,IT	0.364***	0.284***	0.706***	0.795***	0.882**	1.774
	(0.034)	(0.036)	(0.124)	(0.266)	(0.354)	(2.064)
Medicine	0.620***	0.545***	0.867***	0.822***	0.870***	0.621
	(0.022)	(0.026)	(0.057)	(0.209)	(0.320)	(0.833)
Social Sciences	0.311***	0.283***	0.400***	0.233	0.187	-1.518
	(0.028)	(0.029)	(0.155)	(0.168)	(0.185)	(2.647)
Natural Sciences	0.289***	0.235***	0.457***	0.548***	0.628**	-0.272
	(0.027)	(0.029)	(0.076)	(0.196)	(0.246)	(1.946)
Observations	200,025	200,025	200,025	199,999	199,999	199,999
R-squared	0.153	0.156	0.144	0.135	0.130	-0.224
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	10.03	10.03	10.03	10.03	10.03	10.03
Cragg-Donald F			811	208.3	148.7	3.791
Underid pval			0	0	0	0.417

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2: Returns to broad field 0-25 on labor mkt - Log of income (cond. on inc¿0)

VARIABLES	(1) OLS	(2) OLS	(3) IV1 Own Suggestion	(4) IV2 Predicted Choice +-2	(5) IV3 Predicted Choice -3	(6) IV4 Predicted Suggestion +-2
Stem	0.570***	0.459***	0.723***	0.553***	0.209	0.697**
	(0.017)	(0.018)	(0.036)	(0.126)	(0.319)	(0.309)
Natural Sciences	0.176***	0.191***	0.307***	0.273**	0.395	0.326
	(0.023)	(0.022)	(0.054)	(0.139)	(0.363)	(0.429)
Econ&Bus+Law	0.714***	0.653***	0.938***	0.530***	0.889***	0.706***
	(0.020)	(0.020)	(0.042)	(0.094)	(0.187)	(0.252)
Observations	197,124	197,093	197,093	197,067	140,050	197,067
R-squared	0.127	0.147	0.136	0.141	0.096	0.141
SchoolXCohort FE	X	X	X	X	X	X
Teacher Group FE	X	X	X	X	X	X
Indiv. Cont.	-	X	X	X	X	X
Peer Controls	-	X	X	X	X	X
Baseline hr wage omitted cat.	10.03	10.03	10.03	10.03	9.956	10.03
First Stage Cragg-Donald F			8883	1030	84.61	95.16
Underid pval			0	0	5.16e-05	2.26e-05

Omitted category is “Humanities, Social Sciences, Architecture and Design”.
Standard Errors Cluster at individual level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Probability of observing non-zero income 1-26 years after university graduation

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion S_i^f	(4) IV Predicted Choice $\hat{\pi}_i^f$	(5) IV Predicted Choice $\hat{\pi}_i^f$	(6) IV Predicted Suggestion $\hat{S}_i^f$
Arch&Design	0.023*	0.015	0.050*	-0.066	-0.092	-0.164
	(0.013)	(0.016)	(0.027)	(0.066)	(0.097)	(0.294)
Econ&Business	0.120***	0.106***	0.170***	0.185**	0.196*	-0.445
	(0.011)	(0.013)	(0.024)	(0.072)	(0.102)	(0.444)
Law	0.080***	0.070***	0.117***	-0.038	-0.073	-0.316
	(0.010)	(0.011)	(0.023)	(0.050)	(0.062)	(0.388)
Engineering	0.129***	0.115***	0.168***	0.129*	0.150	-0.192
	(0.010)	(0.012)	(0.022)	(0.076)	(0.114)	(0.473)
Math,Physics,IT	0.039**	0.028*	0.055	0.257**	0.300**	-0.125
	(0.016)	(0.016)	(0.051)	(0.112)	(0.143)	(0.814)
Medicine	0.251***	0.240***	0.279***	0.210**	0.207	0.149
	(0.011)	(0.013)	(0.025)	(0.102)	(0.150)	(0.320)
Social Sciences	0.061***	0.057***	0.048	-0.001	-0.011	-1.323
	(0.013)	(0.013)	(0.057)	(0.069)	(0.075)	(0.957)
Natural Sciences	0.003	-0.010	0.055*	0.047	0.060	-0.462
	(0.012)	(0.014)	(0.032)	(0.075)	(0.095)	(0.949)
Observations	291,117	291,117	291,117	291,083	291,083	291,083
R-squared	0.044	0.045	0.042	0.018	0.011	-0.502
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	0.638	0.638	0.638	0.638	0.638	0.638
Cragg-Donald F			1407	298	211.6	5.095
Underid pval			0	0	0	0.392

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Lottery ticket effect - Probability of reaching top 1%

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion S_i^f	(4) IV Predicted Choice $\hat{\pi}_i^f$	(5) IV Predicted Choice $\hat{\pi}_i^f$	(6) IV Predicted Suggestion $\hat{S}_i^f$
Arch&Design	0.011***	0.008*	0.026***	0.034	0.058	-0.075
	(0.003)	(0.005)	(0.008)	(0.024)	(0.037)	(0.202)
Econ&Business	0.060***	0.056***	0.073***	0.078**	0.096**	-0.431
	(0.005)	(0.006)	(0.011)	(0.031)	(0.043)	(0.312)
Law	0.079***	0.073***	0.105***	0.041**	0.030	-0.170
	(0.005)	(0.005)	(0.011)	(0.019)	(0.024)	(0.272)
Engineering	0.012***	0.009**	0.027***	0.012	0.040	-0.354
	(0.004)	(0.004)	(0.009)	(0.031)	(0.047)	(0.341)
Math,Physics,IT	-0.008**	-0.011**	0.013	0.091*	0.126**	0.380
	(0.004)	(0.004)	(0.020)	(0.049)	(0.062)	(0.563)
Medicine	0.004	0.003	0.014	0.017	0.038	-0.170
	(0.003)	(0.004)	(0.009)	(0.042)	(0.064)	(0.227)
Social Sciences	0.006**	0.004	0.030*	-0.015	-0.012	-0.738
	(0.003)	(0.004)	(0.018)	(0.026)	(0.029)	(0.648)
Natural Sciences	-0.003	-0.004	0.008	-0.042	-0.030	-0.745
	(0.002)	(0.003)	(0.009)	(0.031)	(0.040)	(0.670)
Observations	291,117	291,117	291,117	291,083	291,083	291,083
R-squared	0.062	0.062	0.060	0.035	0.025	-2.229
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	0.0113	0.0113	0.0113	0.0113	0.0113	0.0113
Cragg-Donald F			1407	298	211.6	5.095
Underid pval			0	0	0	0.392

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Probability of becoming a top manager

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion	(4) IV Predicted Choice +-2	(5) IV Predicted Choice +-2	(6) IV Predicted Suggestion +-2
Arch&Design	-0.012	-0.012	-0.008	0.182*	0.268*	1.735
	(0.008)	(0.011)	(0.026)	(0.097)	(0.140)	(2.154)
Econ&Business	0.088***	0.082***	0.115***	0.126**	0.184**	-0.560
	(0.009)	(0.010)	(0.020)	(0.057)	(0.081)	(0.943)
Law	0.012*	0.010	0.021	-0.052	-0.050	0.118
	(0.007)	(0.008)	(0.020)	(0.050)	(0.065)	(1.325)
Engineering	0.023***	0.019**	0.054***	0.133**	0.215**	-0.069
	(0.008)	(0.009)	(0.019)	(0.059)	(0.092)	(0.454)
Math,Physics,IT	-0.010	-0.010	-0.028	0.119	0.223*	-0.076
	(0.010)	(0.011)	(0.043)	(0.095)	(0.132)	(1.069)
Medicine	0.111***	0.109***	0.139***	0.414*	0.415	0.779
	(0.022)	(0.022)	(0.041)	(0.223)	(0.289)	(1.598)
Social Sciences	0.008	0.005	0.077*	0.034	0.053	-1.702
	(0.007)	(0.008)	(0.041)	(0.050)	(0.059)	(2.843)
Natural Sciences	-0.010	-0.009	0.000	-0.032	0.031	0.082
	(0.006)	(0.007)	(0.018)	(0.056)	(0.074)	(1.722)
Observations	118,309	118,309	118,309	118,296	118,296	118,296
R-squared	0.099	0.100	0.092	0.014	-0.009	-6.392
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Sector FEs	-	-	-	-	-	-
Baseline hr wage omitted cat.	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188
Cragg-Donald F			411.2	77.92	58.32	0.953
Underid pval			0	6.42e-10	2.82e-08	0.793

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Probability of having income below the poverty line - 0-25 years on labor market

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion S_i^f	(4) IV Predicted Choice $\hat{\pi}_i^f$	(5) IV Predicted Choice $\hat{\pi}_i^f$	(6) IV Predicted Suggestion $\hat{S}_i^f$
Arch&Design	-0.031***	-0.026***	-0.058***	0.004	0.051	-0.118
	(0.008)	(0.009)	(0.016)	(0.032)	(0.047)	(0.148)
Econ&Business	-0.112***	-0.094***	-0.171***	-0.141***	-0.116**	-0.303
	(0.005)	(0.006)	(0.011)	(0.034)	(0.048)	(0.207)
Law	-0.091***	-0.079***	-0.139***	-0.085***	-0.064*	-0.225
	(0.005)	(0.006)	(0.012)	(0.029)	(0.035)	(0.191)
Engineering	-0.107***	-0.087***	-0.161***	-0.090**	-0.050	-0.198
	(0.005)	(0.006)	(0.012)	(0.036)	(0.053)	(0.196)
Math,Physics,IT	-0.074***	-0.058***	-0.126***	-0.101**	-0.058	-0.238
	(0.008)	(0.008)	(0.025)	(0.051)	(0.068)	(0.396)
Medicine	-0.139***	-0.125***	-0.186***	-0.114***	-0.071	-0.261
	(0.005)	(0.006)	(0.013)	(0.042)	(0.062)	(0.163)
Social Sciences	-0.062***	-0.057***	-0.072**	-0.022	0.003	-0.559
	(0.007)	(0.007)	(0.035)	(0.037)	(0.040)	(0.515)
Natural Sciences	-0.068***	-0.056***	-0.106***	-0.160***	-0.143***	-0.416
	(0.006)	(0.007)	(0.017)	(0.040)	(0.050)	(0.366)
Observations	200,025	200,025	200,025	199,999	199,999	199,999
R-squared	0.042	0.044	0.037	0.028	0.027	-0.174
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	0.147	0.147	0.147	0.147	0.147	0.147
Cragg-Donald F			811	208.3	148.7	3.791
Underid pval			0	0	0	0.417

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Probability of having a fixed-term contract

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion S_i^f	(4) IV Predicted Choice $\hat{\pi}_i^f$	(5) IV Predicted Choice $\hat{\pi}_i^f$	(6) IV Predicted Suggestion $\hat{S}_i^f$
Arch&Design	0.092***	0.114***	0.006	0.073	0.109	0.112
	(0.021)	(0.024)	(0.044)	(0.099)	(0.142)	(0.512)
Econ&Business	-0.207***	-0.180***	-0.301***	-0.099	-0.032	0.196
	(0.012)	(0.014)	(0.027)	(0.079)	(0.114)	(0.501)
Law	-0.107***	-0.083***	-0.225***	-0.075	-0.040	-0.049
	(0.013)	(0.015)	(0.032)	(0.085)	(0.104)	(0.675)
Engineering	-0.234***	-0.192***	-0.366***	-0.293***	-0.306***	0.047
	(0.012)	(0.014)	(0.026)	(0.077)	(0.114)	(0.459)
Math,Physics,IT	-0.131***	-0.101***	-0.214***	-0.097	-0.081	-0.092
	(0.017)	(0.018)	(0.056)	(0.121)	(0.162)	(0.719)
Medicine	-0.048**	-0.003	-0.250***	-0.328	-0.386	-0.126
	(0.024)	(0.027)	(0.055)	(0.280)	(0.366)	(0.766)
Social Sciences	-0.067***	-0.052***	-0.191***	-0.116	-0.089	1.126
	(0.015)	(0.016)	(0.074)	(0.083)	(0.096)	(1.236)
Natural Sciences	-0.106***	-0.086***	-0.174***	-0.293***	-0.342***	0.171
	(0.014)	(0.015)	(0.037)	(0.085)	(0.107)	(0.918)
Observations	145,504	145,504	145,504	145,491	145,491	145,491
R-squared	0.080	0.083	0.069	0.045	0.032	-0.394
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	0.410	0.410	0.410	0.410	0.410	0.410
Cragg-Donald F			609.8	109.5	77.70	3.256
Underid pval			0	0	0	0.430

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Probability of receiving unemployment benefits

	(1)	(2)	(3)	(4)	(5)	(6)
			IV	IV	IV	IV
			Own	Predicted	Predicted	Predicted
			Suggestion	Choice	Choice	Suggestion
VARIABLES	OLS	OLS	S_i^f	$\hat{\pi}_i^f$	$\hat{\pi}_i^f$	$\hat{S}_i^f$
Arch&Design	-0.010***	-0.010***	-0.012***	-0.023***	-0.020	-0.052
	(0.001)	(0.002)	(0.003)	(0.008)	(0.012)	(0.040)
Econ&Business	-0.010***	-0.010***	-0.012***	-0.006	-0.001	-0.010
	(0.001)	(0.002)	(0.003)	(0.010)	(0.014)	(0.073)
Law	-0.011***	-0.010***	-0.016***	-0.024***	-0.023**	-0.033
	(0.001)	(0.002)	(0.003)	(0.007)	(0.009)	(0.061)
Engineering	-0.010***	-0.008***	-0.015***	-0.003	0.007	0.024
	(0.001)	(0.002)	(0.003)	(0.011)	(0.016)	(0.080)
Math,Physics,IT	0.000	0.003	-0.011*	0.015	0.026	-0.138
	(0.003)	(0.004)	(0.006)	(0.016)	(0.020)	(0.116)
Medicine	-0.016***	-0.014***	-0.021***	-0.026*	-0.021	-0.027
	(0.001)	(0.001)	(0.003)	(0.015)	(0.021)	(0.048)
Social Sciences	-0.008***	-0.007***	-0.011*	-0.033***	-0.031***	-0.111
	(0.002)	(0.002)	(0.007)	(0.010)	(0.011)	(0.150)
Natural Sciences	-0.006***	-0.005***	-0.010**	-0.042***	-0.041***	-0.012
	(0.002)	(0.002)	(0.005)	(0.012)	(0.014)	(0.139)
Observations	291,117	291,117	291,117	291,083	291,083	291,083
R-squared	0.010	0.011	0.010	-0.011	-0.013	-0.220
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	0.0127	0.0127	0.0127	0.0127	0.0127	0.0127
Cragg-Donald F			1407	298	211.6	5.095
Underid pval			0	0	0	0.392

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Probability of working in a sector/occupation cell that does not require a college degree - 0-25 years on labor market

VARIABLES	(1) OLS	(2) OLS	(3) IV Own Suggestion S_i^f	(4) IV Predicted Choice $\hat{\pi}_i^f$	(5) IV Predicted Choice $\hat{\pi}_i^f$	(6) IV Predicted Suggestion $\hat{S}_i^f$
Arch&Design	0.090***	0.076**	0.128**	-0.120	-0.176	0.247
	(0.028)	(0.031)	(0.064)	(0.196)	(0.275)	(1.927)
Econ&Business	-0.051***	-0.051**	-0.062	-0.174	-0.233	0.822
	(0.017)	(0.020)	(0.039)	(0.132)	(0.192)	(0.867)
Law	-0.063***	-0.065***	-0.053	-0.045	-0.033	1.740
	(0.018)	(0.020)	(0.044)	(0.133)	(0.169)	(2.129)
Engineering	0.034**	0.028	0.020	-0.023	-0.055	0.753
	(0.017)	(0.020)	(0.039)	(0.125)	(0.190)	(1.010)
Math,Physics,IT	-0.075***	-0.082***	-0.101	-0.305	-0.348	-0.518
	(0.023)	(0.025)	(0.088)	(0.193)	(0.258)	(1.800)
Medicine	-0.273***	-0.265***	-0.357***	0.567	0.819	1.647
	(0.019)	(0.024)	(0.067)	(0.392)	(0.512)	(3.872)
Social Sciences	-0.027	-0.023	-0.164*	-0.059	-0.029	0.299
	(0.020)	(0.020)	(0.096)	(0.116)	(0.134)	(3.063)
Natural Sciences	-0.072***	-0.087***	-0.025	-0.100	-0.126	0.911
	(0.018)	(0.020)	(0.050)	(0.120)	(0.153)	(1.506)
Observations	94,549	94,549	94,549	94,536	94,536	94,536
R-squared	0.060	0.061	0.051	-0.041	-0.072	-1.592
SchoolXCohort FE	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-
Baseline hr wage omitted cat.	0.319	0.319	0.319	0.319	0.319	0.319
Cragg-Donald F			330.3	68.52	51.34	0.556
Underid pval			0	5.61e-10	5.93e-08	0.852

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 10: Returns to field 0-25 on labor mkt - Log of income by Cohort

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Cohort 85-89 OLS	Cohort 90-94 OLS	Cohort 95-99 OLS	Cohort 85-89 IV Own Suggestion	Cohort 90-94 IV Own Suggestion	Cohort 95-99 IV Own Suggestion	Cohort 85-89 IV Predicted Choice +-2	Cohort 90-94 IV Predicted Choice +-2	Cohort 95-99 IV Predicted Choice +-2
Arch&Design	0.214***	0.146***	0.141***	0.454***	0.183**	0.284***	0.212	0.016	0.057
	(0.055)	(0.044)	(0.051)	(0.122)	(0.092)	(0.101)	(0.182)	(0.259)	(0.154)
Econ&Business	0.840***	0.742***	0.724***	1.349***	0.836***	1.015***	0.564**	1.006***	0.783***
	(0.046)	(0.033)	(0.038)	(0.104)	(0.066)	(0.088)	(0.238)	(0.286)	(0.127)
Law	0.650***	0.495***	0.588***	0.958***	0.609***	0.826***	0.501***	0.442**	0.627***
	(0.043)	(0.033)	(0.042)	(0.106)	(0.073)	(0.086)	(0.173)	(0.204)	(0.144)
Engineering	0.660***	0.569***	0.549***	1.068***	0.642***	0.726***	0.535***	0.481*	0.664***
	(0.042)	(0.029)	(0.040)	(0.102)	(0.064)	(0.092)	(0.198)	(0.262)	(0.151)
Math,Physics,IT	0.429***	0.364***	0.378***	0.663***	0.727***	0.729**	0.460	0.297	0.342
	(0.052)	(0.049)	(0.058)	(0.184)	(0.169)	(0.298)	(0.280)	(0.359)	(0.234)
Medicine	0.450***	0.467***	0.515***	0.736***	0.594***	0.796***	0.590***	0.552*	0.780***
	(0.039)	(0.027)	(0.045)	(0.104)	(0.068)	(0.118)	(0.197)	(0.302)	(0.205)
Social Sciences	0.388***	0.303***	0.216***	0.447	0.357**	0.266	0.541***	0.030	0.042
	(0.050)	(0.041)	(0.045)	(0.372)	(0.158)	(0.194)	(0.172)	(0.278)	(0.221)
Natural Sciences	0.316***	0.239***	0.285***	0.593***	0.236***	0.693***	0.832***	0.524*	0.330*
	(0.048)	(0.036)	(0.046)	(0.138)	(0.092)	(0.136)	(0.299)	(0.303)	(0.191)
Observations	37,347	54,431	33,773	37,347	54,431	33,773	37,347	54,426	33,762
R-squared	0.103	0.085	0.096	0.086	0.080	0.082	0.071	0.065	0.089
SchoolXCohort FE	X	X	X	X	X	X	X	X	X
5yr Teacher FE	X	X	X	X	X	X	X	X	X
Indiv. Cont.	X	X	X	X	X	X	X	X	X
Peer Controls	X	X	X	X	X	X	X	X	X
Field suggestion Controls	-	X	-	-	X	-			
Baseline Humanities	9.840	9.798	9.692	9.840	9.798	9.692	9.840	9.798	9.692
Num. Individuals	4632	6201	3725	4632	6201	3725	4632	6200	3724
Cragg-Donald F				82.87	260.4	100.2	62.63	41.66	79.42
Underid pval				9.35e-06	0	2.00e-09	0	1.03e-07	1.08e-07

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1