open science

# Explaining gender segregation in higher education: longitudinal evidence on the French case 

Estelle Herbaut, Carlo Barone

## To cite this version:

Estelle Herbaut, Carlo Barone. Explaining gender segregation in higher education: longitudinal evidence on the French case. British Journal of Sociology of Education, 2021, 42 (2), pp.260-286. 10.1080/01425692.2021.1875199 . hal-03365195

HAL Id: hal-03365195
https://hal.science/hal-03365195
Submitted on 3 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Explaining Gender Segregation in Higher Education: Longitudinal Evidence on the French Case 

Estelle Herbaut ${ }^{\text {a }}$ * and Carlo Barone ${ }^{\text {b }}$<br>${ }^{a}$ Observatoire sociologique du changement (OSC), Sciences Po, Paris, France;<br>${ }^{b}$ Observatoire sociologique du changement (OSC), Sciences Po, Paris, France<br>*Contact details: estelle.herbaut@sciencespo.fr; 27, rue Saint-Guillaume, 75337 Paris Cedex 07, France.


#### Abstract

This article examines how girls and boys choose their tertiary field of study and how the different factors driving their choices contribute to Gender Segregation in Higher Education (GSHE) in France. We present seven theoretical explanations for GSHE, review the relevant literature and discuss their applicability to the French context. Using rich longitudinal data combining administrative sources, students, parents and school head questionnaires, we assess the heuristic value of these explanations. We employ multinomial logit models as well as decomposition techniques for categorical outcome variables. Our results refute explanations arguing that GSHE reflects gender differences in skills or girls' lower career ambitions. We conclude that curricular differentiation and occupational plans are key drivers of GSHE, even when controlling for ability selection into secondary curricula and for parental and teachers' genderstereotypical influences.


Keywords: gender segregation; higher education; field of study; STEM;

## 1.Introduction

In recent cohorts, women display significantly higher tertiary attainment rates than men in virtually all OECD countries (DiPrete, Buchmann 2013). Thanks to these educational performances, women have narrowed the gender gaps in the labour market and have gained increasing access to graduate jobs (OECD 2015, 2016). However, this equalisation process is still hindered by several factors, among which persisting Gender

Segregation in Higher Education (GSHE). Women are still strongly underrepresented in some of the most rewarding fields of study, such as engineering and information \& communications technology, while they are overrepresented in the humanities, human sciences and in other fields of study which are less economically rewarding (Buchmann et al. 2008; Charles, Bradley 2002). Hence, GSHE fosters gender segregation across occupations and gender inequalities in labour market outcomes (Smyth, Steinmetz 2008).

GSHE is highly resistant to change. Its stability in recent cohorts is striking (Mann, DiPrete 2013; Barone 2011; Charles, Bradley 2009), when compared with the dramatic increases in educational and labour market participation of women. Moreover, the qualitative pattern and the overall level of GSHE display remarkable similarities across countries, testifying to the low efficacy of gender mainstreaming policies promoted in some of them (Barone 2011). The level of GSHE appears to be even higher in countries with the highest levels of gender equality in overall educational attainment (Stoet, Geary 2018). While these descriptive patterns are well-documented, much less is known about the factors driving GSHE. This is unfortunate, because understanding how and when GSHE is produced is of utmost importance to develop effective policies to erase it.

To be sure, an extensive literature sheds light on GSHE by tracing it back to the gender stereotypes interiorised and performatively enacted by teachers, parents and students, as well as to the gender biases affecting educational contents, curricula and textbooks (Buchmann et al. 2008; Xie, Shauman 2005). Moreover, as discussed below, extensive theoretical work has suggested several detailed mechanisms that could link gender stereotypes to GSHE. However, less attention has been paid to the empirical assessment of these theoretical accounts and, in particular, to the relative contributions
of different layers of GSHE. We have a long list of potential explanatory candidates, but we know little about their relative weight. Hence, we do not know which are the most important targets for developing effective policies to desegregate higher education.

Moreover, the scant empirical research assessing the contribution of different factors has three limitations. First, these studies mainly report negative findings, suggesting for instance that gender differentials in mathematics or in the choice of elective subjects in high school fail to account for GSHE (Ellison and Swanson 2010; Frank et al. 2008). While these results are important, they leave our research puzzle without solution. Second, since these studies are largely confined to the US, their conclusions may not be generalizable to European educational systems. In particular, several European countries display a more rigid system of curricular tracking in secondary education than the US and access to these tracks is often constrained by previous performance (Blossfeld et al. 2016). Hence, the results concerning the limited role of school performance and curricular choices in the US may not apply to the European context, as reported by a recent study on the Italian case (Barone, Assirelli 2019). Third, these studies typically assess only one or two potential factors of GSHE, while it is often suggested that GSHE results from the convergence of multiple influences (Buchmann et al. 2008). Besides, if the omitted determinants of GSHE correlate with those included in the analyses, omitted-variable bias is a serious concern.

In this study, we provide an empirical assessment of seven potential explanations for GSHE: gender gaps in math, the comparative advantage of girls in humanistic subjects, teachers' gender biases in the perceptions of student skills, parents' informal pressures towards stereotypical field of study choices, curricular specialisation in high school, the gendered career orientations of students, and their preferences for different types of jobs. For this purpose, we exploit a rich, large-scale, longitudinal
survey carried out in France which followed a cohort of students from entry into secondary education to higher education completion, collecting data from students, teachers, parents and administrative sources.

## 2. Theoretical explanations for gender segregation in Higher Education

Sociological explanations for GSHE stress the pervasive character of normative expectations concerning gender-appropriate attitudes and behaviour. Even in societies formally committed to gender equality, gender stereotypes persistently influence the skills, preferences and expectations developed by girls and boys (Charles, Bradley 2002, 2009). However, different sociological explanations trace different pathways between gender stereotypes and GSHE. Broadly speaking, we can identify three main theoretical approaches: some explanations build on the Rational Choice Theory (RCT), some highlight social influence processes, while a last set of mechanisms refers to path dependencies and the linkages between the educational system and the labour market.

Within the RCT framework, gendered skill gaps, especially in math and science, are perhaps the most recurrent explanation for GSHE. They involve a choice mechanism driven by the anticipated chances of success across tertiary fields. Gender stereotypes embedded in socialisation practices are supposed to promote a faster development of numeric and spatial skills among boys (Ceci and Williams 2010). International surveys such as PISA and TIMSS report that boys outperform girls in math and science skills (OECD 2015). Since several male-dominated fields are mathintensive, these skill gaps could foster GSHE. However, in most OECD countries gender gaps in math are small (and those in science are even smaller). Boys are overrepresented in the upper tail of the distribution of math skills, but this tail is too small to account for a sizeable share of GSHE (ibidem). In France, recent evidence
suggests that the advantage of boys in math has declined over time (Chabanon and Steinmetz 2018).

The comparative advantage hypothesis is similarly concerned with gender differences in student skills and related chances of success across fields (Jonsson 1999). However, this second explanation stresses the strengths of girls, rather than their weaknesses, focusing on the female advantage in humanistic subjects. These skill gaps favouring girls are stronger than those favouring boys in math and science. Hence, the former may be more consequential for field of study choices, if girls maximise their chances of success in higher education by choosing those fields of study that value these skills more. However, the few empirical tests of this explanation reported negative evidence, because performance in humanistic subjects was a poor predictor of field of study choice (Barone, Assirelli 2019; Vaarmets 2018; Jonsson 1999; Van de Werfhorst et al. 2003).

Focusing on social influence dynamics, the social control hypothesis (Jacobs 1995) stresses instead the importance of parental pressures to conform to cultural norms about gender-stereotypical fields, and of the tacit stigmatisation of gender-atypical choices, perceived as a form of deviant behaviour. Indeed, parental curricular preferences foster gender segregation across secondary curricula (Gabay-Egozi et al. 2015).

Teachers are not immune from gender stereotypes. Their beliefs on the 'natural' inclinations of girls and boys and their most appropriate educational pathways may thus contribute to GSHE (Gunderson 2012). Teachers' track recommendations represent the most visible source of influence on student choices and can operate as a second mechanism of social control reinforcing parental pressures. However, teachers’ influences need not to operate only via formal recommendations. Indeed, their
perceptions of student skills and inclinations, vehiculated in daily interactions in the classroom and in regular assessments, may shape students' self-images and career plans even more deeply, with lasting consequences for GSHE (Gabay-Egozi et al. 2015; Frank et al. 2008; Correll 2004). Previous research reported that, among students with comparable performance as measured in standardised tests, girls display lower selfconfidence in math, and in France the gender gap in mathematics anxiety is one of the largest among the participating countries in PISA 2012 (OECD 2013). These gender biases originate from a variety of sources, but it seems likely that teachers, who regularly assess students, play a major role in this respect, thus representing a fourth source of GSHE.

Finally, path dependencies and anticipated linkages between school outcomes and the labour market may also contribute to GSHE. In all western countries, secondary education involves some form of curricular differentiation, be it a formal school tracking model, a more informal model of elective subject choice, or a combination of the two (Blossfeld et al. 2016). The gender-typing of secondary curricula represents a fifth potential explanation for GSHE. Gender stereotypes affect the extent to which the subjects taught in different curricula, the skills and attitudes that they are supposed to instil, and the career pathways for which they train, are perceived as appropriate for female and male students. These normative expectations are interiorised and enacted by girls and boys as 'spontaneous' inclinations for these gender-typed curricula (Buchmann et al. 2008). In fact the subject and career preferences of students are often quite vague and fluid in secondary education (Barone et al. 2017), but curricular tracking structures them into a limited set of predefined, gender-typical pathways constraining their future educational choices. Secondary curricula are thus strong predictors of tertiary field choices, thus mediating GSHE. In countries like the US,
where curricular differentiation arrives later in the educational career, is less intensive and more reversible, its contribution to GSHE may not be very important, as reported in some studies (Frank et al. 2008; Morgan et al. 2013). However, in several European countries, curricular differentiation arrives earlier, and it is more intensive and rigid. Two comparative analyses indicate that earlier curricular choice is associated with higher GSHE (Imdorf et al. 2015; Smyth, Steinmetz, 2008). In Italy, where track choices are made at the age of 14 and are hardly reversible, the secondary curricular track was found to account for about two thirds of GSHE (Barone, Assirelli 2019).

The last two explanations focus on the connections between tertiary fields and anticipated labour market outcomes. Boys may be more attracted to lucrative, prestigious, career-oriented jobs, while girls may be more oriented toward less ambitious but family-friendly jobs, that is, occupations displaying flexible work schedules and facilitated access to part-time contracts (Zafar 2013; Ceci and Williams 2010). Then, fields of study associated with welfare occupations, teaching and whitecollar jobs in the public sector may be more attractive for girls. However, the evidence for recent cohorts casts some doubts on the main premise of this explanation, namely that girls are less ambitious than boys (Bobbitt-Zeher 2007; Konrad et al. 2000; Mann and DiPrete 2013).

An alternative explanation is that girls and boys care for future instrumental rewards to a similar extent, but differ in their expressive preferences for specific occupations (Barone 2011; Barone, Assirelli 2019; Morgan et al. 2013). In particular, girls may prefer care-oriented jobs involving direct interaction with people oriented towards their personal development and welfare (teachers, psychologists, social workers, doctors), while boys are more interested in occupations that involve technical
tasks entailing the formal manipulation of objects. Our seventh explanation thus focuses on expressive preferences for detailed occupations.

These seven explanations for GSHE may be complementary and mutually reinforcing. For instance, student performances across subjects may affect curricular choices, which could in turn amplify these initial skill gaps. Similarly, occupational preferences can be both a cause and a consequence of curricular choices. Due to data limitations, we cannot incorporate all these patterns of reciprocal causation in our analyses, but it seems important to acknowledge these complexities.

## 3. The French educational system

In France, education provides a common curriculum until the end of lower secondary education ( $9^{\text {th }}$ grade), when students make a first curricular choice at the age of 15 . They select the type of high school they wish to attend, either vocational school (lycée professionnel) or general \& technological school (lycée général et technologique). The former offers several occupation-specific curricula, most being strongly gender-typed (e.g. early child care, fashion, construction). The latter involves one year of common instruction in $10^{\text {th }}$ grade (cycle de determination), when students have to choose two compulsory options, followed by a second bifurcation between technological and academic track in grade 11. The former offers eight streams, characterised by a prevalence of technical and economic curricula, while the latter has three streams (science, literature, social sciences). In grade 12, students in the academic track further choose a subject speciality within their stream (for example, students in the scientific stream can choose physics, life sciences, etc.), which entails more instruction time and a larger weight in the final high school examination taking place in this year. A recent reform has modified to some extent these curricular paths, but we do not discuss it in this work because it has not affected the cohort under study. Overall, France is an
intermediate case between rigid tracking systems, such as Germany or Switzerland, and comprehensive models prevalent in Scandinavian and Anglo-Saxon countries.

At the end of lower secondary school (grade 9), families formulate a preference for a school type (vocational or academic) and teachers respond with a written recommendation. In case of disagreement, the final decision is taken by the school head and is binding. Students demanding a general \& technological high school do not make any curricular choice nor do they get any teacher recommendation for a specific curriculum; this happens instead in the first high school year when they are still attending a common core instruction (grade 10).

All three tracks afford an upper secondary certificate (baccalauréat) ensuring eligibility to higher education. Using a longitudinal survey on students who entered secondary education in 1995 (see below), it is possible to estimate that the majority of students ( $62 \%$ ) entering an academic and technological track in 10th grade obtained an academic upper secondary certificate, while $27 \%$ of them obtained a technical upper secondary certificate. In contrast, among students entering a vocational track only $21 \%$ obtained a vocational upper secondary certificate. This partly reflects high dropout risks in this track, but also the possibility to stop after two or three years of instruction and obtain a short vocational qualification that does not afford eligibility to higher education.

French higher education includes two-year vocational programmes which award the "Brevet de Technicien Supérieur" (BTS) or the "Diplôme Universitaire de Technologie"-(DUT), and academic instruction offered in universities and in the small sector of elite institutions (Grandes écoles). Bachelor's programmes in universities are accessible to any high school graduate, without additional requirements nor selection.

Grandes écoles grant only five-year master's degrees to students selected via competitive examinations.

## 4. Data, variables and models

The study draws on the survey "Panel d'élèves du second degré, recrutement 1995", a nationally representative, longitudinal survey which followed a cohort of students who entered lower secondary education ( $6^{\text {th }}$ grade) in 1995 throughout secondary and tertiary education. It combines administrative data on school careers, written information reported for each student by school principals in 1995, a parental questionnaire administered in 1998 (in the summer between $8^{\text {th }}$ and $9^{\text {th }}$ grade), and a student questionnaire administered in 2002 ( $12^{\text {th }}$ grade, the last high school year). Students were also surveyed about their higher education studies every following year until they left education for two consecutive years. The parents and student questionnaires were administered via either post mail or phone interviews (in case of non-response to the mail questionnaire).

Our main dependent variable refers to the first field of study of enrolment in higher education, either in academic or in vocational programmes (BTS, DUT). The original data differentiate between 40 programmes. We first classified them into 10 categories by adapting the ISCED-F13 classification for fields of study ${ }^{1}$ to the French case and to the information available in the survey. The ten categories are: humanities (arts, literature, foreign languages, etc.); human sciences (psychology, sociology, teacher education, etc.); law \& political science; business \& economics; natural sciences (mathematics, biology, chemistry, etc.); technical fields (engineering, information \&

[^0]communications technology, construction); sports; health \& welfare (medicine, nursery, social work); sales \& services (personal services, sales, tourism); a residual missing field category which refers to programmes with no information on the field of study in the original dataset.

Taking into account sample size and the need to have a manageable number of outcomes, we further grouped these categories into a 5-category classification: Humanities and human sciences, business and administration (business \& economics, law \& political sciences), STEM programmes (natural sciences and technical fields, sport), care-oriented fields (health \& welfare; sales \& services); and a missing category. The dissimilarity index by gender ${ }^{2}$ amounts to 0.38 when using the original information (40 programmes), 0.36 for the 10 -category classification and 0.32 for the final classification, which thus captures $85 \%$ of the GSHE observed in the original data. Finally, we add a category "no postsecondary education" to account for selection into higher education: as transition rates into higher education vary by gender, we model postsecondary educational choices for the whole sample of students eligible to it.

Regarding the role of math skills for GSHE, we use performance measures at three points in time: a standardised test taken in grade 6 ; this is a low-stakes assessment administered for diagnostic purposes, whose results may be communicated to the parents. As a measure of performance before track choices, we use teacher grades in math at the end of lower secondary school. These grades are an important criterion for families and teachers to choose among school tracks. Finally, we use the grades obtained in the national examinations in high school (grade 12) to capture math skills

[^1]with a standardised measure. These three performance measures are available also for French. Following the seminal paper by Jonsson (1999), we test the comparative advantage hypothesis by taking the ratio between the scores in French and in math at each assessment.

To assess the hypothesis that teacher perceptions of student skills are genderbiased, we further include the evaluations reported for each student by the school principals on student skills in math and in French in grade 6, when students were administered the standardised test. To report this information, school principals have either directly spoken to the teachers or used their evaluations formulated in the school reports.

There are two measures of parental influences in our data. First, the parental questionnaire had a question on the preferred track (vocational, technical academic) and, for the academic track, the specific stream (scientific, social sciences, humanistic). Unfortunately, for the two other tracks we do not have this detailed information on curricular preferences. This is less of a problem for the vocational track, since only a low share of its students enrol in higher education anyway, while the loss of information is more relevant for the technological track. Our second measure concerns structured out-of-school activities in lower secondary education, assuming that parents either influence the choice of these activities directly at this young age or more indirectly shape their children's preferences for gender stereotypical activities. We picked up from a larger battery ${ }^{3}$ of items three leisure activities that are both frequent and gendered: sports, artistic activities and being registered in a library (see Table 1).

[^2]Curricular specialisation in high school is first measured in grade 10. For vocational tracks, we used the speciality of the first programme attended after lower secondary education. As regards students enrolled in academic and technological tracks following in grade 10 the common curriculum, they must choose two compulsory options. For more than $90 \%$ of them, the first option is the second foreign language, while there is more diversity in the second option. Hence, in this study, we use the subject chosen as second option, which identifies a first curricular specialisation at the beginning of high school. Specialties and options are classified to match our categories of fields of study in higher education (see Table 1). As an additional measure of curricular specialisation at the end of high school (grade 12), we use the detailed curriculum of the high school certificate, based on the three tracks (academic, technological, vocational), the stream and subject speciality chosen by the student. These two measures are based on administrative sources.

Finally, we use two sets of indicators of occupational plans based on the student questionnaire administered in the senior high school year. The first one assesses the hypothesis that girls are less career-oriented than boys: students were asked to choose two priorities for their future job among several possibilities. We focus on two attributes: earnings and job security. The second set of indicators focuses on detailed occupational preferences, collected through an open question about students' aspired occupations, which was coded into 22 categories in the original data, which we recoded to match the categories that we use for fields of study in higher education.

We additionally include the following control variables: the highest level of education and social class (EGP schema) among the caregivers, city size in 1995, and parents' country of birth. Among the 10,226 students who obtained the high school diploma granting access to higher education, $10 \%$ have missing information on the
transition into higher education and are thus excluded. Among the remaining 9,183 students with data on the dependent variable, $40 \%$ have missing data on at least one independent variable. Listwise deletion (or complete case analysis) can lead to biased estimates when the mechanism producing missingness is not at random (Rubin, 1987). To address this issue, we opted for multiple imputation. Incomplete variables were imputed in 100 datasets using a fully conditional specification approach (Van Buuren et al., 2006), where imputations are done sequentially for each variable using the other variables in the study (including the outcome variable), along with auxiliary variables to reduce bias.

The association between gender and field of study is estimated with multinomial logit models. The analyses are performed for each imputed dataset separately, with estimates and standard errors combined using Rubin's rules (Rubin, 1987). We report the average marginal effects (AME), which are easily interpretable. We first estimate a baseline model with only gender and the control variables, then add sequentially each of the variables presented above, following a temporal order.

In order to estimate the mediation effects of different variables, we further rely on the KHB method (Breen, Karlson, \& Holm, 2013). This method allows comparing coefficients between two nested non-linear models, taking into account the rescaling of logit coefficients in logit models; it is implemented with the Stata routine khb (Kohler, Karlson, and Holm 2011). This method has not been developed for imputed datasets, yet. We thus replicated our analyses on the original dataset with a missing category for each categorical variable and set to the mean missings for continuous variables. As shown in the appendix, estimated average marginal effects with this method are almost identical to the effects estimated with multiple imputation (Tables B1 - B6). Hence, applying the KHB method on the original dataset with missing categories is not
expected to bias the results to any significant extent. We identify the contribution to GSHE of each mediator by comparing logit coefficients in the reduced-form model (which includes only gender and the control variables) and in the subsequent, nested models that progressively incorporate mediating variables. We use the category "missing field of study" as base outcome, since this is the category with the smallest gender differential (2 p.p.) and which has less substantive interest.

Before presenting the results in the next section, we wish to stress that they cannot be interpreted in causal terms. As explained above, this study improves substantially over previous research by incorporating a rich set of determinants of GSHE, but ultimately we cannot exclude that results reflect omitted-variable bias and reverse causality bias.

### 5.1 Descriptive results

As illustrated by Figure 1, our data on enrolments revealed clear patterns of GSHE. The largest gender gap involves technical programmes, attended by $4.3 \%$ of women compared to $26.1 \%$ of men. Men also enrolled in natural sciences (12.6\%) more often than women $(6.8 \%)$. Conversely, women more often chose the humanities $(13.6 \%$ vs. $4.5 \%$ ), the human sciences ( $9.4 \%$ vs. $4 \%$ ), health and welfare programmes $(12.5 \%$ vs. $3.9 \%$ ), sales and services programmes ( $19.5 \%$ vs $11.8 \%$ ). We detect only marginal gender differences for business \& economics and for law and political sciences. It is reassuring that enrolments in programmes which could not be classified vary only slightly by gender. The share of students not enrolling in higher education was higher for men.
[Figure 1]

Our five-fold classification of tertiary fields reproduces these patterns: $23 \%$ of women chose the humanities and human sciences compared with $8.5 \%$ of men; $12.6 \%$ of women chose STEM fields compared with $42.2 \%$ of men and $31.9 \%$ of women selected care-oriented programmes compared with $15.6 \%$ of men (see figure A1 in appendix). Enrolments in business and administration were gender balanced.

Table 1 shows the distribution by gender of the potential mediators of GSHE. Performance in math and French, assessed by standardised tests, differed by gender. In grade 6, boys outperformed girls in math on average, while girls enjoyed a comparative advantage in French; both differences were statistically significant, but small. Although the same pattern was detected in the concomitant school heads' assessments, the male advantage in math looked even smaller, and the competitive advantage of girls in French was not even statistically significant. Interestingly, in grade 9 , that is, just before the first branching point, we did not find any significant difference in GPA in math, while girls displayed a moderate competitive advantage in French.

Despite these rather similar performances of girls and boys, parents revealed gender-stereotypical patterns of curricular preferences. They more often considered the scientific stream of the academic track for boys ( $22.9 \%$ vs $15.9 \%$ for girls) and the literature stream for girls ( $5.5 \%$ vs $1.3 \%$ for boys). However, several parents of female ( $27.6 \%$ ) and of male students ( $21 \%$ ) declared only a generic preference for the academic track (but not for a specific stream). The technological track was preferred for boys ( $5.7 \%$ vs $1.7 \%$ for girls). One out of five families did not express any expectation about the track or stream of their children. Overall, the gendered patterns were of moderate intensity and indecision was high among parents. Our second indicator referred to child participation in structured out-of-school leisure activities in lower
secondary education. Gender differences were large for practising sport activities, while artistic activities were somehow more common among girls.
[Table 1]

Regarding curricular tracking, we found already large differences at the beginning of high school (grade 10): $44.3 \%$ of boys chose a specialisation in scientific and technical fields, but only $10.3 \%$ of girls. Girls chose more often an option in the field of economics and administration (55\%) or human sciences (26.6\%).

Looking at the detailed track, stream and speciality chosen at the end of high school, gender differentiation was no less strong. While $33.9 \%$ of all boys graduated from the scientific stream of the academic track, this was the case for only $24.3 \%$ of the girls. Furthermore, students differed also by the subject speciality chosen for the final examinations: math and engineering were more often chosen by boys, while girls opted more often for the life sciences. In contrast, only $3.8 \%$ of boys graduated from the humanistic stream of the academic track, but $16.1 \%$ of girls did so. Gender differences were particularly striking in technological and vocational tracks, where $28.8 \%$ of boys graduated in a technical speciality, but only $2.3 \%$ of girls.

Regarding performance in math in the high school final examinations, we found a small advantage of girls. However, the difficulty of examinations depends on the chosen track, so the advantage of girls in math may reflect their lower enrolment in scientific streams. The competitive advantage of girls in French was small and nonsignificant. Once more, gender differences in performance across subjects looked weak.

Finally, gender affects the preferred job attributes and detailed occupational plans. Boys more often prioritised "earning a good living" ( $53.2 \%$ vs $45.4 \%$ for girls), while girls more often chose 'having job security' ( $13.8 \%$ vs $7.8 \%$ for boys). Finally,
gender differences in occupational plans were large: for instance, $40.7 \%$ of boys but only $10.6 \%$ of girls mentioned a STEM job. Overall, the most striking gender differentials involved the choice of different curricula, streams and subjects, as well as the detailed occupational plans of high school students. In the next section, we explore the role of these different mediators for the explanation of GSHE.

### 5.2 Multivariate results

We now turn to the multivariate models for field of study choices (Tables 2a to 2c). Since we detected a weak gender gap for "business \& administration" (only 2 p.p.), the results for this category are included in the Appendix, together with results for missing field of study and no postsecondary education (Tables A1-A3). The average marginal effects for gender in model 0 confirmed the above descriptive patterns: women were 15 p.p. more likely to enrol in the humanities and human sciences; 29 p.p. less likely to take STEM programmes and 16 p.p. more likely to choose care-oriented programmes.

Early performance in math (model 1) had a negligible effect on enrolments in humanistic and care fields, while each additional point (out of 20) increased by 3 p.p. enrolments in STEM programmes. Despite this large impact, gender gaps in STEM fields decreased by only 1 p.p. in model 1 after the inclusion of this mediating variable, since gender differences in math were small. Our measure of comparative advantage, the ratio between French and math scores in grade 6, influenced enrolments across the three outcomes, but the effects were small. Hence, also the inclusion of this mediator in model 2 changed only marginally the gender coefficients. Interestingly, when we added school principals' assessments in these two subjects in model 3, both scores in standardised tests and school heads assessments had independent effects, especially for STEM fields. However, the results did not confirm the hypothesis of a teachers' bias: the gender coefficient changed only marginally between model 2 and 3 . Our last set of
performance measures before track choice is the GPA in math and the ratio between French and math grades in grade 9 (model 4). GPA in math had a statistically significant coefficient only for STEM programmes, while a comparative advantage in French mattered as well for enrolling in the humanities and human sciences. Both performance measures were irrelevant for access to care-oriented fields.

Once controlling for all indicators of student performance in lower secondary education, the coefficient for gender decreased from 15 p.p. to 13 p.p. for the humanities and human sciences (model 0 to 4 ), from -29 p.p. to -26 p.p. for STEM programmes and stayed unchanged for care-oriented programmes, thus pointing to the limited explanatory power of these mediators.

Parents' curricular preferences (model 5) displayed large effects. For example, a preference for the literature stream was associated with an increase by 17 p.p. in the probability that the child enrolled in the humanities and human sciences (compared to having no preference), while considering the scientific stream of the academic track increased by 6 p.p. child's enrolments in STEM programmes. In contrast, the coefficients for spare time activities were small and showed no association with field of study choices. As can be seen, overall, these two indicators only marginally mediated GSHE.

Curricular specialisation in grade 10 (first year of high school, model 6) had large independent effects controlling for all the above-mentioned predictors, pointing to systematic affinities with later fields of study choices. Moreover, the introduction of this variable substantially reduced the gender coefficients across all outcomes, for instance they shrank by 6 p.p. for STEM programmes. The coefficients for the detailed track and subject speciality taken in grade 12 were even larger. These later curricular choices explained out the coefficients for early curricular choices, pointing to a strong
continuity between the two. For instance, compared to students in the scientific stream of the academic track who took a mathematic speciality, those graduating from the academic literature track were 52 p.p. more likely to enrol in the humanities and human sciences. The comparisons of gender coefficients across models 5 and 6 suggest that curricular choices substantially contributed to GSHE.

## [Table 2a]

## [Table 2b]

## [Table 2c]

Regarding performance in the high school examinations in French and math, we detected only small effects, which left the gender coefficients unchanged between models 7 and 8 for all outcomes.

Preferences for earnings and job security for the future job were negatively associated with enrolment in the humanities and human sciences, positively associated with care-oriented programmes, and unrelated to enrolment in STEM programmes. We did not find any support for the hypothesis that GSHE reflects a lack of ambition on the side of girls, as the inclusion of this variable leaves the gender coefficient unchanged (model 9) ${ }^{4}$. Finally, we found large coefficients for students' occupation plans: students who named an occupation in one field (compared to those who did not know) were always much more likely to enrol in the corresponding field category. Including this variable resulted in marked reductions of the gender coefficients, especially for STEM and care-oriented programmes.

[^3]We now turn to the results of the KHB decomposition. Since we found no support for the hypothesis of a teachers' bias, we grouped together the contributions of all performance measures before track choice in figure 2 for better readability (the contribution of each variable is presented in table C 1 in the Appendix). A negative value indicates that the variable acts as a suppressor of the gender effect on fields of study choice.
[Figure 2]

First of all, overall these mediators accounted for most of the association between gender and fields of study. The unexplained proportion of GSHE was only $33 \%$ for care-oriented programmes, $16 \%$ for STEM and $13 \%$ for the humanities and human sciences. Curricular specialisation in high school was the key driver of GSHE: curricular choices in grade 10 and 12 altogether accounted for $57.3 \%$ of the gender gap in STEM fields, $47,3 \%$ in the humanities and human sciences, and $34 \%$ in care-oriented fields. Occupational plans mediated $28 \%$ of the gender gap in care-oriented and in STEM programmes, and $21 \%$ of the gender gap in the humanities and human sciences, while the priority assigned to earnings and job security made a negligible contribution. Math performances played a minor role across fields and the comparative advantage of girls in French mediated 8\% of their over-representation in the humanities and human sciences, which was also slightly driven by parental influences (9.3\%).

### 5.3 Robustness checks

Our results are robust to several robustness checks (available upon request). First, we reran the models for the smaller sample $(\mathrm{N}=5,474)$ of students with data on all independent variables: the overall gender gaps across fields and the patterns of
mediation remained unchanged. Second, we tested a different coding of fields by separating health and welfare programmes from sales and services: the only different mediation pattern was that curricular specialisation was more relevant for GSHE in sales \& services programmes, while detailed occupational plans were more relevant for health and welfare programmes.

Third, we reran the models excluding students enrolled in vocational higher education: in academic higher education we found a larger gender gap in the humanities and human sciences (21 p.p.), a smaller one in STEM (-18 p.p.) and in care-oriented programmes (4 p.p.) than in the whole sector of higher education. The mediation patterns suggest that the comparative advantage of girls in French was more relevant for GSHE in academic humanities and human sciences fields, while curricular specialisation explained a smaller share of GSHE in academic STEM fields. The small over-representation of women in care programmes was explained out by occupational plans. Again, this redefinition of the analytical sample did not affect our key substantive conclusions

Fourth, we used the field of study of graduation, instead of the field of enrolment. We lost a large number of individuals, due to attrition and missing information. In this selected sample, GSHE was smaller than for enrolments, but the patterns of mediation remained largely unchanged.

Fifth, we used information on curricular specialisation (track and stream) in grade 11 instead of the detailed high school degree, which further includes subject specialisation. We found that the mediation power of curricular specialisation is reduced, especially for STEM programmes, suggesting that subject specialisation in grade 12 was most relevant for GSHE in these fields.

Finally, we changed the order of introduction of the independent variables. In the above analyses, we followed a temporal order and included occupational preferences in the last model because they were measured at the end of high school. However, these preferences may have developed earlier, and they may have even driven curricular choices. We have therefore included the corresponding variables before the measures of curricular specialisation, and their explanatory contribution was enhanced. Compared to the model controlling for performances before tracking and parental expectations, adding occupational plans reduced the gender coefficient by 14 p.p. for STEM programmes, by 7 p.p. for care-oriented programmes, while for the humanities and human sciences the reduction was smaller (4 p.p.). The effect of occupational plans on STEM programmes was largely reduced when controlling for curricular choices, while it was more robust for care-oriented programmes. This confirms the patterns found in the KHB decomposition (which is insensitive to the order of introduction of variables): curricular specialisation is a more important mediator for STEM programmes, while the weight of occupational plans is larger for care-oriented programmes. Additional research, using repeated measures of occupational plans before and after tracking, is necessary to clarify the interplay between curricular differentiation and occupational preferences.

## 6. Conclusions

In this work we used a rich longitudinal dataset on educational careers in France to carry out a comprehensive assessment of seven explanations for GSHE. Overall, in line with studies on Anglo-Saxon countries, we conclude that skill-based explanations display limited heuristic value. School performance across subjects is a significant predictor of fields of study choices, but it significantly contributes to GSHE only in the humanities and human sciences. Importantly, this reflects the stronger performance of
girls in French, in line with the comparative advantage hypothesis, rather than math gaps across genders. This is unsurprising, since the latter are quite small, while girls clearly outperform boys in French in junior school.

Teachers' gender biases in the assessment of student skills play a minor role for GSHE, too. The reports of school principals based on the subjective evaluations of teachers point to limited gender differences in math, which are even smaller than those recorded by standardised tests passed in the same school year. These results confirm previous studies reporting that teachers do not downgrade girls in math in secondary education (OECD 2015). However, we cannot exclude other teacher influences on GSHE operating, in particular, via counselling activities at the end of lower secondary education. More generally, we could measure teacher attitudes only indirectly with our data.

Even though girls and boys do not systematically differ much in their math performances at the end of junior school, nor in their perceptions by teachers, parents display a clear preference toward scientific and technical streams for boys and toward humanistic streams for girls. The type of out-of-school activities chosen by children in junior school are also suggestive of gender biases in primary socialisation. Overall, parental biases are more evident than teacher biases in our data. However, even the indicator of parental pressures fails to significantly contribute to GSHE. An important consideration here is that about half of the parents do not declare any specific preference concerning the curricular choices of their children, and even among those who express a preference, gender-stereotypical patterns are of moderate intensity. This suggests that the kind of direct pressures to choose a secondary and tertiary field of study described by the social control hypothesis may not be very important. However, parental influences may operate in more indirect and tacit forms that are difficult to
capture with survey data. The only indirect measure available in the data, based on spare time activities of children, is possibly too crude in this respect. Moreover, with our data we cannot observe any parental influence operating later in the educational career.

Curricular differentiation in France combines the model based on tracking (grade 10) and streaming (grade 11) often found in continental Europe with the model of elective subject choice (in grade 12) prevalent in Anglo-Saxon countries. This gradual, soft process of curricular differentiation is by far the single most important mediator of GSHE across all fields, although its role is larger for humanistic and STEM fields than for care-oriented fields. Importantly, its contribution to GSHE looks much larger than what was found in previous research, largely based on subject choice in the US. Gender differences across school types, streams, specialties and subjects are strong and display marked associations with fields of study choices. In the tracked systems found in continental Europe, curricular differentiation is possibly a much more important driver of GSHE than in Anglo-Saxon countries.

On one side, curricular differentiation brings girls and boys to specialise in different disciplines in high school, thus diverging in the disciplines and fields of study in tertiary education that attract them and that they perceive as realistic options. On the other side, curricular differentiation can both confirm and reinforce gender differences in occupational plans. We found that preferences for specific occupations are strong predictors of fields of study choice that significantly mediate GSHE, particularly for STEM and care-oriented fields. Hence, the interplay between curricular choices and aspired occupations captures the core processes driving GSHE in France.

On the contrary, gender differences in the importance attached to earnings and job security seem of limited importance. The hypothesis that GSHE is driven by a lower
career orientation of girls is not supported by our analyses. More generally, our study lends limited empirical support to rational choice models of GSHE, considering that the three explanations derived from this theory (skill gaps in math, comparative advantage in French and career orientations) found limited support. This suggests that mechanisms mobilising cost-benefit calculations do not shed much light on GSHE.

This does not mean that student choices are irrational, though. Instead students seem to choose their fields of study following two simple heuristics that seem reasonable, given the limited information available to them: 'what I like do' and 'what I can do'. Curricular choices are relevant in both respects: they affect their preferences for different disciplines and occupations, as well as their chances to succeed in different fields of study. Girls and boys do not differ much in their academic performances before being tracked, nor in their occupational priorities even after tracking. However, they are progressively channelled into different academic pathways through a sequence of curricular decisions (track choices, streaming, elective subjects), whose cumulative impact is a powerful differentiation of their opportunity sets and preferences for specific disciplines and occupations.

Overall, our mediators accounted for most gender differences across fields of study. However, their relative weight varied significantly across fields of study. Occupational plans matter less for gender gaps in the humanities and human sciences, which are mainly driven by girls' preference for school streams and subjects displaying clear affinities with these disciplines in their educational contents and, to a minor extent, by girls' stronger performance in French; the 'vocational logic' of preferences for specific occupations is less relevant here. To the contrary, this vocational logic is more important for girls choosing care-oriented fields like medicine or social work. Previous academic results did not contribute to GSHE in these fields and curricular
differentiation is less predictive of enrolment than for other fields. The strong overrepresentation of boys in STEM fields results from both dynamics: their specialisation in scientific curricula and subjects fostering their chances to succeed in these fields, as well as their preferences for the related occupations (e.g. engineer). While explanations for GSHE are typically formulated in general terms, our results suggest that their heuristic value differs across fields.

Most previous research on GSHE was either descriptive or involved testing only one or two mechanisms. This study presented a comprehensive assessment of seven explanations using rich, longitudinal data. Our contribution to the literature was twofold. On one side, we could refute explanations arguing that GSHE reflects gender differences in skills or girls' lower ambitions. On the other side, we concluded that curricular differentiation and detailed occupational plans are the key drivers of GSHE, even when controlling for ability selection into secondary curricula and for parental and teacher gender-stereotypical influences. To the best of our knowledge, this is the first study arriving at this important conclusion.

Let us also stress some limitations of our study and directions for future research. First, we could not test some potential explanations for GSHE relating to students' motivation in different subjects, peer effects and to anticipated discrimination in gender-atypical fields. Given that we explain large part of GSHE, this does not seem a major limitation, but of course omitted-variable bias is a potential concern. Second, reverse causality is also possible: for instance, students may somehow rationalise their anticipated fields of study choices by declaring gender-typical occupational plans. Hence, it must be reiterated that the correlational results of this study cannot be interpreted in causal terms. The third, and perhaps most important, limitation of this work is that our data do not allow any detailed assessment of the mechanisms driving
the allocation of girls and boys to different curricula. Given that our study documented the importance of curricular differentiation for GSHE, identifying these mechanisms will advance our theoretical understanding of GSHE and will provide important indications to policy-makers.

## References

Barone, Carlo and Giulia Assirelli. 2020. "Gender segregation in higher education: an empirical test of seven explanations." Higher Education 79(1):55 78.

Barone, Carlo. 2011. "Some Things Never Change: Gender Segregation in Higher Education across Eight Nations and Three Decades." Sociology of Education 84(2):157-76.

Barone, Carlo, Antonio Schizzerotto, Giovanni Abbiati, and Gianluca Argentin. 2017. "Information Barriers, Social Inequality, and Plans for Higher Education: Evidence from a Field Experiment." European Sociological Review 33(1):8496.

Blossfeld, Hans-Peter, Sandra Buchholz, Jan Skopek, and Moris Triventi. 2016. Models of Secondary Education and Social Inequality: An International Comparison. Edward Elgar Publishing.

Bobbitt-Zeher, Donna. 2007. "The Gender Income Gap and the Role of Education." Sociology of Education 80(1):1-22.

Breen, Richard, Kristian Bernt Karlson, and Anders Holm. 2013. "Total, Direct, and Indirect Effects in Logit and Probit Models." Sociological Methods \& Research 42(2):164-91.

Buchmann, Claudia, Thomas A. DiPrete, and Anne McDaniel. 2008. "Gender Inequalities in Education." Annual Review of Sociology 34(1):319-37.

Ceci, Stephen J. and Wendy M. Williams. 2010. "Sex Differences in Math-Intensive Fields." Current Directions in Psychological Science 19(5):275-79.

Chabanon, Léa and Claire Steinmetz. 2018. "Écarts de performance des élèves selon le sexe." Éducation \& formations (96):20.

Charles, Maria and Karen Bradley. 2002. "Equal but Separate? A Cross-National Study of Sex Segregation in Higher Education." American Sociological Review 67(4):573-99.

Charles, Maria and Karen Bradley. 2009. "Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries." American Journal of Sociology 114(4):924-76.

Correll, Shelley J. 2004. "Constraints into Preferences: Gender, Status, and Emerging Career Aspirations." American Sociological Review 69(1):93-113.

DiPrete, Thomas A. and Claudia Buchmann. 2013. The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools. Russell Sage Foundation.

Ellison, Glenn and Ashley Swanson. 2010. "The Gender Gap in Secondary School Mathematics at High Achievement Levels" Journal of Economic Perspectives 24(2):109-28.

Frank, Kenneth A., Chandra Muller, Kathryn S. Schiller, Catherine Riegle- Crumb, Anna Strassmann Mueller, Robert Crosnoe, and Jennifer Pearson. 2008. "The Social Dynamics of Mathematics Coursetaking in High School." American Journal of Sociology 113(6):1645-96.

Gabay-Egozi, Limor, Yossi Shavit, and Meir Yaish. 2015. "Gender Differences in Fields of Study: The Role of Significant Others and Rational Choice Motivations." European Sociological Review 31(3):284-97.

Gunderson, Elizabeth A., Gerardo Ramirez, Susan C. Levine, and Sian L. Beilock. 2012. "The Role of Parents and Teachers in the Development of Gender-Related Math Attitudes." Sex Roles 66(3):153-66.

Imdorf, Christian, Kristinn Hegna, Verena Eberhard, and Pierre Doray. 2015. "Educational Systems and Gender Segregation in Education" Pp. 83-122 in Gender Segregation in Vocational Education. Vol. 31, Comparative Social Research. Emerald Group Publishing.

Jacobs, Jerry A. 1995. "Gender and Academic Specialties: Trends among Recipients of College Degrees in the 1980s." Sociology of Education 68(2):81-98.

Jonsson, Jan O. 1999. "Explaining Sex Differences in Educational Choice: An Empirical Assessment of a Rational Choice Model." European Sociological Review 15(4):391-404.

Kohler, Ulrich, K. B. Karlson, and A. Holm. 2011. "Comparing Coefficients of Nested Nonlinear Probability Models Using Khb." Stata Journal 11(3):420-438.

Mann, Allison and Thomas A. DiPrete. 2013. "Trends in Gender Segregation in the Choice of Science and Engineering Majors." Social Science Research 42(6):1519-41.

Morgan, Stephen L., Dafna Gelbgiser, and Kim A. Weeden. 2013. "Feeding the Pipeline: Gender, Occupational Plans, and College Major Selection." Social Science Research 42(4):989-1005.

OECD. 2013. PISA 2012 Results: Ready to Learn (Volume III). Paris: Éditions OCDE.
OECD. 2015. The ABC of Gender Equality in Education. Paris: Éditions OCDE.

OECD. 2016. Education at a Glance 2016. Paris: Organisation for Economic Cooperation and Development.

Rubin, D. B. (1987) Multiple Imputation for Nonresponse in Surveys, Chichester: John Wiley \& Sons.

Smyth, Emer and Stephanie Steinmetz. 2008. "Field of Study and Gender Segregation in European Labour Markets." International Journal of Comparative Sociology 49(4-5):257-81.

Stoet, Gijsbert and David C. Geary. 2018. «The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education ». Psychological Science 29(4):581-93.

Vaarmets, Tarvo. 2018. "Gender, Academic Abilities and Postsecondary Educational Choices." Journal of Applied Research in Higher Education 10(3):380-98.

Van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., Rubin, D. B. (2006) 'Fully conditional specification in multivariate imputation', Journal of Statistical Computation and Simulation, 76(12): 1049-64.

Van De Werfhorst, Herman, Alice Sullivan, and Sin Yi Cheung. 2003. "Social Class, Ability and Choice of Subject in Secondary and Tertiary Education in Britain." British Educational Research Journal 29(1):41-62.

Xie, Yu and A. Shauman. 2005. Women in Science: Career Processes and Outcomes. Cambridge, Mass: Harvard University Press.

Zafar, Basit. 2013. "College Major Choice and the Gender Gap." Journal of Human Resources 48(3):545-95.

Table 1: Bivariate associations between gender and independent variables

|  | Men | Women | $P$-value ${ }^{\text {a }}$ |
| :---: | :---: | :---: | :---: |
| Score in math at entrance in $6^{\text {th }}$ grade (1995), out of 20 | 14.68 | 14.16 | 0.00 |
| Ratio between French and math scores at entrance in 6th grade | 1.00 | 1.09 | 0.00 |
| Score given by the head of school in math at entrance in $6^{\text {th }}$ grade (1995), out of 10 | 7.22 | 7.06 | 0.00 |
| Ratio between French and math scores given by head of school | 1.11 | 1.08 | 0.68 |
| GPA in math in 9th grade | 11.98 | 11.92 | 0.31 |
| Ratio between GPA in French and math in 9th grade | 0.97 | 1.07 | 0.00 |
| Track Do not know or no answer <br> considered by Apprenticeship/CAP/BEP <br> parents in Vocational track <br> 1998 for high Technological track <br> school Academic track, any stream <br>  Academic track-literature <br>  Academic track-economics <br>  Academic track-scientific <br>  Missing | 19.42 3.53 6.14 5.70 21.04 1.33 1.79 22.85 18.21 | $\begin{gathered} \hline 20.96 \\ 3.03 \\ 3.97 \\ 1.67 \\ 27.60 \\ 5.53 \\ 3.73 \\ 15.90 \\ 17.61 \end{gathered}$ | 0.00 |
| Out-of-school None <br> activities in Sport only <br> 1998 Sport and library <br>  Sport and arts <br>  Library only <br>  Arts only <br>  Library and arts <br>  Sport, library and arts <br>  Missing | $\begin{gathered} \hline 15.07 \\ 31.16 \\ 19.57 \\ 5.34 \\ 8.94 \\ 3.26 \\ 2.95 \\ 5.63 \\ 8.09 \end{gathered}$ | $\begin{gathered} \hline 19.19 \\ 12.02 \\ 11.38 \\ 7.97 \\ 15.21 \\ 7.97 \\ 9.42 \\ 9.00 \\ 7.83 \\ \hline \end{gathered}$ | 0.00 |
| Option/ Language, arts, human sciences <br> specialty in  <br> grade 10 Economics, business, administration <br>  Sciences, engineering, production, sport, agriculture <br>  <br>  <br>  <br> Health, welfare, services <br> Missing | $\begin{gathered} \hline 12.25 \\ 40.07 \\ 44.28 \\ 2.80 \\ 0.60 \end{gathered}$ | $\begin{gathered} 26.59 \\ 55.03 \\ 10.27 \\ 7.63 \\ 0.48 \end{gathered}$ | 0.00 |
| Track, stream Academic, scientific, math <br> and subject Academic, scientific, physics <br> specialty at Academic, scientific, life sciences <br> upper Academic, scientific, engineering <br> secondary Academic, economics, economics <br> graduation Academic, economics, math <br>  Academic, economics, other specialties <br>  Academic, literature, all specialties <br>  Technological, industrial technologies <br>  Technological, business and administration <br>  Technological, laboratory technologies and agronomy <br>  Technological, health, welfare and services <br>  Vocational, manufacturing, production, IT, agriculture <br>  Vocational, social sciences, communication <br>  Vocational, administration, business <br>  Vocational, health, welfare \& services | $\begin{gathered} \hline 10.97 \\ 9.69 \\ 8.00 \\ 5.22 \\ 7.51 \\ 3.53 \\ 1.79 \\ 3.79 \\ 15.70 \\ 11.91 \\ 3.00 \\ 0.87 \\ 13.07 \\ 0.31 \\ 1.40 \\ 3.26 \\ \hline \end{gathered}$ | $\begin{gathered} \hline 6.40 \\ 6.68 \\ 10.63 \\ 0.56 \\ 10.87 \\ 5.93 \\ 4.74 \\ 16.14 \\ 0.89 \\ 17.49 \\ 1.92 \\ 6.50 \\ 1.39 \\ 3.43 \\ 2.30 \\ 4.12 \\ \hline \end{gathered}$ | 0.00 |
| Score in math at high school degree examinations, out of 20 | 10.55 | 10.78 | 0.01 |


| Ratio between French and math scores at high school degree <br> examinations | 1.51 | 1.70 | 0.68 |  |
| :--- | :--- | :---: | :---: | :---: |
| Students | Earnings \& other attributes | 53.19 | 45.41 |  |
| preferred job | Job security \& other attributes | 7.78 | 13.84 |  |
| attributes in | Earnings \& job security | 9.71 | 10.59 |  |
| 2002 | Other attributes | 19.15 | 23.26 | 0.90 |
|  | Missing | 10.17 | 16.79 | 12.00 |
| Student | Do not know, no answer or unclear | 7.90 | 18.36 |  |
| occupational | Humanistic and human science professions | 12.00 | 23.10 |  |
| plans in 2002 | Law, Economics, business, administration | 40.72 | 10.59 |  |
|  | Sciences, engineering, production, sport, agriculture | 11.43 | 28.20 | 0.00 |
|  | Health, welfare, sales and services | 11.16 | 7.75 |  |
|  | Missing | $\mathbf{N = 9 , 1 8 3}$ | $\mathbf{4 , 1 4 0}$ | $\mathbf{5 , 0 4 3}$ |

${ }^{\text {a }} \mathrm{P}$-values from t-test (mean comparisons) for continuous variables and from Pearson's chi-squared for categorical variables.
Source: Panel des élèves 1995

Table 2a: Multinomial logistic model of the field of study in higher education (imputed data, $\mathrm{N}=9$ 183); AME for the outcome "Humanities and human sciences"


Table 2b: Multinomial logistic model of the field of study in higher education (imputed data, $\mathrm{N}=9$ 183); AME for the outcome "STEM"

| Variables |  | Model 0 | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 | Mode | el 9 | Mode | el 10 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Gender | Men (reference category) Women | -0.29*** (0.01) | -0.28*** (0.01) | -0.27*** (0.01) | -0.26*** (0.01) | -0.26*** (0.01) | -0.23*** (0.01) | -0.17*** (0.01) | -0.08*** (0.01) | -0.08***(0.01) | -0.08*** | *(0.01) | -0.02*** | *(0.01) |
| Score in math at entrance in 6th grade, out of 20 |  |  | $0.03{ }^{* * *}(0.00)$ | $0.02{ }^{* * *}(0.00)$ | 0.01 *** (0.00) | $0.00 \quad(0.00)$ | $0.00 \quad(0.00)$ | -0.00 (0.00) | -0.00 (0.00) | -0.00 (0.00) | -0.00 | (0.00) | -0.00 | (0.00) |
| Ratio between French and math scores at entrance in 6th grade |  |  |  | $-0.15 * * *(0.03)$ | $-0.15^{* * *}(0.03)$ | $-0.12{ }^{* * *}(0.03)$ | $-0.11^{* * *}(0.03)$ | $-0.10^{* * *}(0.03)$ | -0.04* (0.03) | -0.05* (0.03) | -0.05* | (0.03) | -0.04 | (0.02) |
| Score given by head of school in math in 6th grade, out of 10 |  |  |  |  | $0.02^{* * *}(0.00)$ | $0.00 \quad(0.00)$ | $0.00 \quad(0.00)$ | $0.00 \quad(0.00)$ | $0.00 \quad(0.00)$ | $0.00 \quad(0.00)$ | 0.00 | (0.00) | 0.00 | (0.00) |
| Ratio between French and math scores given by head of school |  |  |  |  | -0.05* (0.02) | -0.05* (0.02) | -0.04* (0.02) | -0.03 (0.02) | -0.01 (0.02) | -0.01 (0.02) | -0.01 | (0.02) | -0.01 | (0.02) |
| GPA in math in 9th grade |  |  |  |  |  | 0.02*** (0.00) | 0.02 *** (0.00) | 0.02*** (0.00) | $0.00 \quad(0.00)$ | $0.00 \quad(0.00)$ | 0.00 | (0.00) | -0.00 | (0.00) |
| Ratio between GPA in French and math in 9th grade |  |  |  |  |  | -0.16*** (0.03) | -0.13*** (0.03) | $-0.08^{* * *}(0.03)$ | -0.03 (0.02) | -0.03 (0.02) | -0.03 | (0.02) | -0.02 | (0.02) |
| Track considered by parents in 1998 for high school | Do not know or no answer (ref.) |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Vocational high school degree |  |  |  |  |  | 0.02 (0.02) |  | 0.00 | 0.00 0 | 0.00 | (0.02) | -0.00 |  |
|  | Technological high school degree Academic high school, any stream |  |  |  |  |  | $\left\|\begin{array}{\|c\|c\|c\|} 0.12^{* * *} \\ -0.03^{* * *} & (0.02) \end{array}\right\|$ | $0.07 * * *$ <br> -0.02 <br> 0 | 0.03 ${ }^{0}(0.02)$ | 0.04* (0.02) | $0.04^{*}$ | (0.02) | 0.02 -0.01 | $(0.02)$ |
|  | Academic high school-literature |  |  |  |  |  | -0.11*** (0.03) | -0.08*** (0.03) | 0.02 (0.03) | 0.03 $\quad 10.03)$ | 0.03 | (0.03) | 0.04 | $\begin{aligned} & (0.01) \\ & (0.03) \end{aligned}$ |
|  | Academic high school-economics |  |  |  |  |  | $-0.09 * * *(0.02)$ | -0.06** (0.03) | $0.01 \quad(0.03)$ | $0.02 \quad(0.03)$ | 0.01 | (0.03) | 0.04 | $\begin{gathered} (0.03) \\ (0.03) \end{gathered}$ |
|  | Academic high school-scientific |  |  |  |  |  | 0.06*** (0.01) | 0.05*** (0.01) | -0.00 $\quad(0.01)$ | -0.00 $\quad 10.01)$ | -0.00 | (0.01) | 0.00 |  |
| Out-of-school activities in 1998 | None (ref.) |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Sport only |  |  |  |  |  | 0.03** (0.01) | 0.03** (0.01) | 0.03** (0.01) | 0.03*** (0.01) | 0.03*** | (0.01) | 0.03** | (0.01) |
|  | Sport and library |  |  |  |  |  | 0.04** (0.02) | 0.04*** (0.01) | 0.03* (0.01) | 0.03** (0.01) | 0.03** | (0.01) | 0.03*** | (0.01) |
|  | Sport and arts |  |  |  |  |  | -0.00 (0.02) | -0.00 (0.02) | -0.01 (0.02) | -0.01 (0.02) | -0.01 | (0.02) | 0.00 | (0.01) |
|  | Library only |  |  |  |  |  | 0.00 (0.02) | $0.01 \quad(0.02)$ | -0.00 (0.01) | -0.00 (0.01) | -0.00 | (0.01) | 0.01 | (0.01) |
|  | Arts only |  |  |  |  |  | 0.00 (0.02) | 0.01 (0.02) | 0.02 (0.02) | 0.01 (0.02) | 0.01 | (0.02) | 0.01 | (0.02) |
|  | Library and arts |  |  |  |  |  | 0.00 (0.02) | $0.00 \quad(0.02)$ | -0.01 | -0.01 (0.02) | -0.01 | (0.02) | -0.00 | (0.02) |
|  | Sport, library and arts |  |  |  |  |  | 0.02 (0.02) | 0.02 (0.02) | 0.01 (0.02) | $0.01 \quad(0.02)$ | 0.01 | (0.02) | 0.01 | (0.01) |
| Option or specialty in 10th grade | Language, arts, social sciences (ref.) |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Economics, business, administration |  |  |  |  |  |  | -0.05*** (0.01) | 0.01 (0.01) | $0.01 \quad(0.01)$ | 0.01 | (0.01) | 0.01 | (0.01) |
|  | ences, production, sport, agriculture |  |  |  |  |  |  | 0.17*** (0.01) | 0.03*** (0.01) | 0.03*** (0.01) | 0.03*** | (0.01) | 0.02* | (0.01) |
|  | Health, welfare and services |  |  |  |  |  |  | -0.15*** (0.02) | 0.01 (0.04) | 0.02 (0.04) | 0.02 | (0.04) | 0.03 | (0.03) |
| Track and specialty of high school degree | Academic-sciences-math (ref.) Academic-sciences-physics |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Academic-sciences-physics Academic-sciences-life sciences |  |  |  |  |  |  |  | $\begin{aligned} & -0.04^{*}(0.02) \\ & -0.16^{* *}(0.02) \end{aligned}$ | $\begin{array}{\|c} -0.03 \\ -0.14^{* * *}(0.02) \end{array}$ |  | $\begin{gathered} (0.02) \\ *(0.02) \end{gathered}$ |  |  |
|  | Academic-sciences-life sciences Academic-sciences-engineering |  |  |  |  |  |  |  | $\begin{gathered} -0.16^{* * *}(0.02) \\ 0.12^{* * *} \\ (0.04) \end{gathered}$ | $\left\|\begin{array}{c} -0.14^{* * *}(0.02) \\ 0.13^{* * *}(0.04) \end{array}\right\|$ | $\begin{aligned} & -0.14^{* * *} \\ & 0.13^{* * *} \end{aligned}$ | ${ }^{*}(0.02)$ | $\begin{aligned} & -0.08^{* * *} \\ & 0.04 \end{aligned}$ | $\begin{array}{r} *(0.02) \\ (0.04) \end{array}$ |
|  | Academic-economics-economics |  |  |  |  |  |  |  | -0.49*** (0.02) | -0.50*** (0.02) | -0.50*** | *(0.02) | $-0.37^{* *}$ | * (0.03) |
|  | Academic-economics-math |  |  |  |  |  |  |  | -0.47*** (0.03) | -0.50*** (0.03) | $-0.50 * * *$ | (0.03) | $-0.35 * * *$ | * (0.03) |
|  | Academic-economics-other specialties |  |  |  |  |  |  |  | $-0.50 * * *(0.03)$ | -0.52*** (0.03) | -0.52*** | (0.03) | -0.39*** | * (0.03) |
|  | Academic-literature |  |  |  |  |  |  |  | -0.52*** (0.02) | -0.54*** (0.02) | -0.54*** | *(0.02) | -0.42** | * (0.03) |
|  | Technological-industrial technologies |  |  |  |  |  |  |  | 0.13*** (0.03) | 0.12*** (0.03) | 0.12*** | (0.03) | 0.08*** | (0.03) |
|  | Technological-business\&administration |  |  |  |  |  |  |  | -0.51*** (0.02) | -0.54*** (0.02) | -0.54*** | (0.02) | -0.43*** | *(0.02) |
|  | Technological-laboratory technologies, agronomy |  |  |  |  |  |  |  | 0.11*** (0.04) | $0.05 \quad(0.04)$ | 0.05 | (0.04) | -0.04 | (0.04) |
|  | Technological-health \& welfare or hotel |  |  |  |  |  |  |  | $-0.52^{* * *}(0.03)$ | -0.54*** (0.03) | -0.54*** | (0.03) | $-0.42^{* *}$ | (0.03) |
|  | Vocational-manufacturing, IT, agriculture |  |  |  |  |  |  |  | -0.28*** (0.04) | -0.31*** (0.04) | -0.31*** | (0.04) | $-0.30 * *$ | (0.03) |
|  | Vocational-social sciences, communication |  |  |  |  |  |  |  | -0.55*** (0.02) | -0.57*** (0.02) | -0.57** | (0.02) | $-0.47^{* *}$ | (0.02) |
|  | Vocational-administration \& business |  |  |  |  |  |  |  | -0.52*** (0.03) | -0.54***(0.03) | -0.54** | (0.03) | $-0.42^{* *}$ | (0.04) |
|  | Vocational-health, welfare \& services |  |  |  |  |  |  |  | -0.50*** (0.03) | -0.53*** (0.03) | -0.53*** | (0.03) | -0.40*** | (0.03) |
| Score in math in high school degree exam, out of 20 |  |  |  |  |  |  |  |  |  | 0.01 *** (0.00) | 0.01*** | (0.00) | 0.01*** | (0.00) |
| Ratio between French and math scores in high school degree exams |  |  |  |  |  |  |  |  |  | -0.00 (0.01) | -0.00 | (0.01) | -0.00 | (0.01) |
| Preferred job attributes in 2002 | Other attributes (ref.) |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Earning \& other |  |  |  |  |  |  |  |  |  | -0.01 | (0.01) | -0.01 | (0.01) |
|  | Job security \& |  |  |  |  |  |  |  |  |  |  | 0.0 | 0.00 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Occupational plans in 2002 | w, no answer or unclear (ref.) Human and social sciences |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Law, Economics, business, adminisistration |  |  |  |  |  |  |  |  |  |  |  | -0.18*** | (0.01) |
|  | Sciences, engineering, production, sport, agriculture Health, welfare, sales and services |  |  |  |  |  |  |  |  |  |  |  | 0.15*** | (0.01) |
|  |  |  |  |  |  |  |  |  |  |  |  |  | -0.15*** | * (0.01) |

Table 2c: Multinomial logistic model of the field of study in higher education (imputed data, $\mathrm{N}=9$ 183); AME for the outcome "care-oriented programmes"


Figure 1: Tertiary field of study of enrolment by gender


Source: Panel des élèves 1995.

Figure 2: Estimated contribution of each mediator to GSHE, using the KHB decomposition method

\begin{tabular}{|c|c|c|c|c|}
\hline \& Humanities \& human sciences \& STEM \& Care-oriented \& $\square$ Unexplained <br>
\hline \multirow[t]{2}{*}{100\%} \& \& \& \& $\square$ Occupational plans <br>
\hline \& \& \& \& 日Job attributes <br>
\hline $80 \%$

$60 \%$ \& \& \& \& 图 Ratio French/math in high school exam <br>
\hline \multirow[t]{2}{*}{60\%} \&  \& \& \& ©Performance in math in high school exam <br>
\hline \& \& \& \& - Detailed track of high school degree <br>
\hline \multirow[t]{2}{*}{40\%} \& \& \& \& $\square$ Option in grade 10 <br>
\hline \& \& \& \& ■Out-of-school activities <br>
\hline 20\% \&  \&  \& \& ■Parental expectations about track <br>
\hline 0 \&  \&  \&  \& Ratio French/math in lower secondary <br>
\hline -20\% \& \& \& \& Performance in math in lower secondary <br>
\hline
\end{tabular}

Results from the KHB decomposition method for logit coefficients, showing the percentage of the effect of being a woman mediated by each variable, with socio-demographic variables as concomitants. See table B1 in appendix.


[^0]:    ${ }^{1}$ The International Standard Classification of Education (ISCED) is the classification developed by UNESCO for organising and comparing education programmes. Adopted in 2013, ISCED-F categorises fields of study in secondary and tertiary education.

[^1]:    ${ }^{2}$ The dissimilarity index ranges from 0 to 1 and can be interpreted as the percentage of women (or men) who would have to change their field of study for the distribution of men and women to be equal. A value of 0 means that the share of women in every field of study is the same as women's share in higher education as a whole, while a value of 1 indicates complete gender segregation across fields of study.

[^2]:    ${ }^{3}$ The out-of-school activities in the questionnaire are: sport, computer club, scout movement, registration in the youth centre of the neighbourhood, artistic activities.

[^3]:    ${ }^{4}$ We carried out additional analyses with two other job attributes, which suggest that the higher propensity of girls to choose 'a job they are passionate about' and which 'allows to meet a lot of people' contributes to their over-representation in care-oriented programmes.

