Was Your Degree Worth It? Heterogenous Returns to College Majors and Skill Mismatch[*]

Lorenzo Navarini[†] (University of Vienna)

Dieter Verhaest[‡] (KU Leuven)

March 2025

Abstract

On average, returns to college are positive and substantial, but they hide substantial heterogeneity across individuals and college majors. This paper estimates the heterogeneous returns to college majors and the impact of skill mismatch using a dynamic model of educational and labor market choices, controlling for unobserved heterogeneity. Identification leverages exclusion restrictions, including local labor market conditions, distance to college, and graduation timing. While most majors yield positive average returns, a sizable share of individuals face low or negative returns, largely due to college major choices and substantial penalties from skill mismatch.

[*]This paper benefited from helpful comments at various stages from Olivier De Groote, Christophe Bruneel-Zupanc, Edwin Leuven, Jo Van Biesebroeck, Koen Declercq, and several audiences at KU Leuven and the University of Vienna. It also benefited from discussions during and after presentations at the 2023 Transition in Youth, Edinburgh, and the LEER Conference 2023, the Vienna Applied Micro Workshop and the UCL-IFS Human Capital and Labor Markets Seminar. Funding for this project was generously provided by the Research Foundation Flanders (FWO) - G079420N.

[†]University of Vienna; KU Leuven; Department of Economics, KU Leuven; Leuven Economics of Education (LEER); Email: lorenzo.navarini@univie.ac.at

[‡]KU Leuven; Department of Economics, KU Leuven; Leuven Economics of Education (LEER); Email: dieter.verhaest@kuleuven.be

1 Introduction

Was your degree worth it? This question is increasingly being asked in the news and the academic literature (Webber, 2016; Hastings et al., 2013; Cappelli, 2015).[1] On average, returns to college are positive and substantial (Altonji et al., 2016a; Oreopoulos and Salvanes, 2011), but returns are also heterogeneous. Some individuals benefit substantially, while others see limited gains (Arcidiacono, 2004; Rodríguez et al., 2016; Altonji et al., 2012). This heterogeneity exists both across and within fields of study. Indeed, wage gaps between majors are large and growing (Altonji et al., 2014), with STEM and Health degrees consistently outperforming the Humanities (Webber, 2014; Kirkebøen et al., 2016; Hastings et al., 2013; Beffy et al., 2012), despite attracting fewer students (Oosterbeek and Webbink, 1997; Altonji et al., 2016b). However, even within the same major, outcomes differ based on college quality (Loury et al., 1995; Andrews et al., 2016), student ability (Arcidiacono, 2004; Rodríguez et al., 2016), early investments in human capital (Humphries et al., 2023), and non-monetary returns (Arcidiacono et al., 2020). When low wage returns are not compensated by such non-pecuniary benefits, some students may have been better off choosing a different major, or forgoing college altogether.

A key determinant of heterogeneity within and across college majors might be attributed to skill mismatch, when a graduate's job does not align with their qualifications (Kinsler and Pavan, 2015; Lemieux, 2014; Robst, 2007). Mismatch leads to lower wage premiums, particularly when the job does not require a college degree (Leuven et al., 2013) or does not utilize the graduate's field of study (Somers et al., 2019). Mismatch rates also differ by major: Humanities graduates, for instance, face higher rates of both vertical and horizontal mismatch (Frenette, 2004; Ghignoni and Verashchagina, 2014).

As noted by Leuven and Oosterbeek (2011), both college major choice and skill mismatch are subject to non-random selection. Ideally, we would observe two independent sources of quasi-experimental variation: one influencing tertiary educational choices, and another affecting occupational choices and, therefore, skill mismatch. There are a few studies exploiting quasi-experimental variation in college admission cut-offs to estimate causal returns to college majors (Hastings et al., 2013; Kirkebøen et al., 2016). Moreover, even with a source of exogenous variation in educational choices, individ-

[1] Recent coverage includes The Economist (Was your degree worth it? and How to make your degree worth the investment) and The Wall Street Journal (Employers Rethink Need for College Degrees and Half of College Grads Are Working Jobs That Dont Use Their Degrees).

uals may differ in preferences or occupation-specific abilities, leading to non-random selection into skill mismatched jobs. To the best of our knowledge, no study combines such variation with a second source that identifies the effect of skill mismatch on wage returns. Indeed, it requires a very specific and rare setting.

In this paper, we estimate heterogeneous causal returns to college majors and investigate the importance of skill mismatches as an underlying mechanism. Rather than relying on a (quasi-)experimental design, we address non-random selection using a dynamic model of joint educational choices and labor market outcomes, accounting for dynamic selection and unobserved heterogeneity (Ashworth et al., 2021; Heckman and Navarro, 2007; Heckman et al., 2018a,b, 2016). We identify the latter using the panel nature of the data, initial conditions, local labor market conditions, and a set of exclusion restrictions, including relative distance to college and graduation timing (cf. Humphries et al., 2023). Conditional on observed state variables and unobserved heterogeneity, relative distance to college shifts the costs to college major enrollment and graduation timing shifts the probability to end up in a skill mismatched occupation at the entry of the labor market. We show that these variables are likely to be excluded from subsequent choices and outcomes, conditional on observed and unobserved characteristics, and are key to identify unobserved types and recover causal estimates of returns to college majors and skill mismatch impact. Moreover, we control for several observed characteristics and endogenous academic achievements, including retention, grades and year of completion, and secondary education outcomes. We include skill mismatches in the first job, and estimate their associated wage penalties at later ages. Moreover, besides investigating the role of skill mismatch, we also account for other indirect channels that may contribute to heterogenous wage returns between college majors such as differences in the probability of graduating or obtaining a higher grade.

We estimate this model using detailed data from Belgium. Our initial sample consists of three samples of individuals born in 1976, 1978, and 1980, with information on educational choices and labor market outcomes up to 29 years old. This dataset follows individuals choices from primary education until they enter the labor market, with detailed calendar data on enrollment, track choices, grades, and college majors.

Unlike when relying on OLS estimation, we find evidence of significant and positive wage returns to each college major based on our dynamic model. The highest returns are for obtaining a Bachelor's and a Master's in Health and Business and Law, respectively. There are important differentials across college majors: a BA in STEM pays almost 4.9 percentage points more than a BA in Business and Law. The same applies to an MA in Business and Law, which pays almost 10.4 percentage points more than a Social sciences MA. Moreover, also returns within college majors are heterogeneous: 34.4% (42.1%) of individuals receive a negative return to college when obtaining a bachelors (masters) in Social sciences. However, when individuals are adequately qualified for their first job, this percentage reduces substantially, and, for some college majors, no individuals experience negative returns to college when adequately matched. This happens because skill mismatch generates a substantial penalty, albeit only when a mismatch in term of college major is combined with overeducation: this penalty goes up to 8% (12.1%) for a BA (MA) in Health. This penalty is relatively lower for a BA (MA) in Business and law: 3.6% (7.7%). However, a degree in Health substantially reduces the probability of both horizontal and vertical mismatches. Humanities and Arts and Social Sciences, meanwhile, are associated with substantially larger probabilities of being mismatched in the first job. Interestingly, while STEM degrees have a similar average probability of vertical mismatch as Humanities and Arts degrees, they are associated with lower probabilities of horizontal mismatch.

Related Literature

Our paper contributes to three related branches of the literature. First, we contribute to the literature on the average returns to the field of study or college major choice (Altonji et al., 2012, 2016b). While most previous studies have relied on a selection-on-observables approach, several more recent papers have estimated causal returns to college majors by exploiting discontinuities in admission cut-offs (Hastings et al., 2013; Kirkebøen et al., 2016; Andrews et al., 2016; Bleemer and Mehta, 2022). In general, these studies confirm that STEM degrees yield larger returns relative to other college degrees. Similarly, by exploiting a lottery for admission to medical school in the Netherlands, Ketel et al. (2016) found evidence of substantial returns to majoring in health. However, as these admission rules are often confined to a restricted set of programs only, these approaches usually do not allow one to deliver a more fine-grained comparison based on a wider set of majors (Andrews et al., 2022). Other studies have therefore relied on fixed-effects models to exploit the variation in earnings between prior to and post enrollment. Relying on this approach to evaluate different programs at US community colleges, Jepsen et al. (2014) confirmed the large earning gains that are associated with health programs. However, students with earnings prior to enrollment is a specific group that is not necessarily representative to the full population and results may still be biased due to problems of dynamic selection. Henceforth, we rely on a dynamic discrete choice model to deal with these problems. Arcidiacono (2004) and Beffy et al. (2012) have embraced a comparable methodology, noting notably higher returns among science and business majors. Different from these studies, though, we consider a wider set of majors and, in particular, differentiate between health and (other) STEM fields. Moreover, we use a set of exclusion restrictions at the main nodes of interest, college major choices and skill mismatch treatment, to identify and distinguish between the unobserved persistent heterogeneity from random shocks.

Our paper also relates to the growing literature on heterogeneous returns to college in general and to college majors in particular. Using data from Chile, Rodríguez et al. (2016) documented substantial heterogeneity in the returns to post-secondary degrees and large fractions of individuals earning negative returns to college. Similarly, Andrews et al. (2016) have found substantial heterogeneity in the returns to college quality in Texas. More recently, using similar data and relying on selection-on-observables approach, Andrews et al. (2022) showed that there is also substantial variation in the returns to college majors both across individuals and within individuals over time. In particular, for 4-year programs, they found the ex-ante risk of science, health and business majors to be much higher relative to liberal arts, with in particular those at the top of the earnings distributions realizing substantial gains. Based on our dynamic discrete choice model, our study provides further evidence on these heterogeneous returns to college majors. Our methodological approach offers several advantages relative to the earlier evidence in this respect. First, while Andrews et al. (2022) controlled for a rich set of observable characteristics, we account for both observable and unobservable determinants of college major choice and earnings. Second, we model high school attainment and college enrollment next to college major choices. Henceforth, rather than merely simulating wage differentials between college majors, we are able to assess exact returns to a college degree and show how college major choice contributes to the existence of negative wage returns to college for a large part of the population. Third, we also account for differences in graduation risk between and within majors. Suppose ones likelihood to drop out is relatively higher in STEM fields of study. In that case, it may be rational to prefer a non-STEM field despite being associated with a lower return conditional on graduation. Our results do, indeed, show that differences in drop-out rates contribute to part of the heterogeneity in returns between and within college majors. Fourth and most importantly, our modeling approach also allows us to demonstrate the importance of skill mismatch in explaining heterogeneous returns between and within college majors.

Indeed, our paper also contributes to the literature on skill mismatch and college major choice. Starting from Robst (2007), the literature has demonstrated that college graduates receive a larger earnings premium when their occupation is a good match for their college majors or is typical for their major (Nordin et al., 2010; Lemieux, 2014; Lindley and McIntosh, 2015). Moreover, college major choice is also an important predictor of the quality of the match between a workers education and job. Indeed, many studies also found both overeducation and field-of-study mismatch to be related to the college major, particularly with graduates from humanities facing problems finding matching jobs (e.g., Frenette, 2004; Ghignoni and Verashchagina, 2014). For instance, relying on data for European graduates, Verhaest et al. (2017) found graduates in Health to be successful in avoiding overeducation and field of-study mismatch; alternatively, graduates in Sciences, Mathematics, and Computing performed averagely in both respects, while those in Engineering, Manufacturing and Construction seemed to combine low incidences of full mismatches (i.e., combining overeducation with horizontal mismatch) with high incidences of mere overeducation. Similar results were found by Chevalier (2017) when relying on data for the UK, with more than half of the graduates in STEM fields other than Health Sciences being employed outside their domain and with the high wage returns to STEM fields being confined to those working in STEM jobs. Further, based on a review of the evidence, also Cappelli (2015) concluded that a substantial part of the STEM graduates is employed outside their domain and that the facts do not warrant complaints about substantial STEM skill shortages.

The most closely connected to our research is the paper by Kinsler and Pavan (2015), who developed a dynamic model for the US graduate labor market to investigate how field-of-study mismatch affects the return to college majors. While they find graduates in Business and Science to realize higher average returns to college than graduates from other fields of study, this advantage for Science majors is found to be absent for those working in non-science jobs. They attribute this finding to science majors being relatively more intensive in specific skills. We add several ways to the literature compared to Kinsler and Pavan (2015). First, our data include multiple measures of mismatches, allowing us to accommodate measurement error problems partly. Second, we identify unobserved heterogeneity and the model using a set of strategies, including exclusion restrictions. Third, like Arcidiacono (2004), Beffy et al. (2012) and Kinsler and Pavan (2015) have not differentiated between Health and (other) STEM fields of study either. Indeed, consistent with findings from non-dynamic studies (Chevalier, 2017; Verhaest et al., 2017), we find marked differences between these two subfields with health degrees being the most effective in avoiding mismatches. Fourth, besides looking into how mismatches explain wage differentials across and within majors, we also show that skill mismatch contributes to differences between college majors in negative wage returns to college.

The rest of our paper is structured as follows. Section 2 introduces the data. Next, in Section 3, we describe the model. The subsequent section presents the results. We end the paper in section 5 with a discussion and conclusion.

2 Institutional setting and data

Institutional setting Belgium is an important setting for studying the relationship between tertiary education choices, skill mismatch and returns to college. The country combines an open-access higher education system with a high share of college graduates. Tertiary education in Belgium is considered universally accessible due to the absence of formal entry barriers, no ex ante selection and the relatively low tuition fees (Declercq and Verboven, 2018).[2] All students who obtain a high school diploma are eligible to enroll in most higher education programs, regardless of their specific secondary school track.

As of 2022, 51.36% of the population aged 25-34 in Belgium held a tertiary degree, a figure well above the OECD average. This represents a substantial increase from 40.61% in 2005 and 32.94% in 1995, both of which were already significantly higher than the OECD average at the time.[3] The rapid expansion of higher education attainment has generated growing concern over a potential overproduction of graduates and the associated risk of labor market mismatches - both vertical and horizontal - attracting increasing attention from researchers and policymakers (Verhaest et al., 2017; Chevalier, 2017; Kinsler and Pavan, 2015).

[2]For a detailed overview of the tertiary education system in Belgium, see Declercq and Verboven (2018). Notably, higher education institutions are not permitted to establish their own admission standards. In addition, tuition fees remain relatively low, currently capped at 890 in Flanders.

[3]See OECD data from OECD: Population with tertiary education

SONAR dataset We analyze this setting using the SONAR dataset. This data provides representative samples from three cohorts (born in 1976, 1978, and 1980) and includes around 3,000 individuals from Flanders, the northern Dutch-speaking region of Belgium. The survey was conducted when participants were 23 years old, and follow-up surveys were completed at ages 26 and 29 for different cohorts (1976 and 1978 at age 26, and 1976 and 1980 at age 29), with response rates ranging from 60% to 70%. The dataset includes detailed information on education and labor market outcomes, including educational choices made from age six onwards and core labor market history every month. It also includes a range of indicators related to family background, skill mismatch status and wages at the start of the first job and at the time of the various surveys.[4]

In this paper, we model tertiary education choices and outcomes, skill mismatch in the first job at the entry in the labor market, and labor market outcomes at subsequent ages (23, 26 and 29). Choices and outcomes are determined by a set of observed characteristics and initial conditions.

Observed characteristics and initial conditions We consider the following observed characteristics: gender (as a dummy variable), number of siblings, foreign origin (as a dummy variable), years of education of both the mother and the father (beyond primary education), day of birth within the calendar year, and cohort fixed effects.[5] We show the descriptive statistics in Table 1.

For tractability, we start modeling choices of individuals at the beginning of tertiary education. However, we account for differentials in initial conditions by including information on grade repetition in primary and secondary education, track choices, and self-reported grades in secondary education and we condition the unobserved heterogeneity on these variables. This is equivalent to fully model unobserved heterogeneity influencing choices starting from primary education. Indeed, unobserved heterogeneity is orthogonal to exogenous variables, while being correlated with initial conditions and choices of the model.

[4]To ensure the model is tractable, we exclude individuals with (i) a delay of more than one year in starting primary education (76 individuals), (ii) special needs requiring care in schools (124 individuals), and (iii) inconsistent, erroneous, or incomplete data on the exogenous variables and educational career (638 individuals). After these exclusions, we estimate the equations related to educational outcomes using a final sample of 8,162 individuals.

[5]These variables are standard background characteristics that are typically included in dynamic discrete choice models on educational careers. See: Cameron and Heckman, 1998, 2001; Heckman et al., 2016, 2018a,b; Neyt et al., 2022; Navarini and Verhaest, 2023.

Table 1: Descriptive statistics: observables and initial conditions

	Mean	SD	Min	Max
Female	0.493	0.500	0.000	1.000
Number of siblings	1.672	1.426	0.000	18.000
Foreign origin	0.057	0.232	0.000	1.000
Father years of education	5.734	3.436	1.000	13.000
Mother years of education	6.214	3.675	1.000	13.000
Day of birth	171.756	100.195	1.000	365.000
Cohort 1978	0.338	0.473	0.000	1.000
Cohort 1980	0.344	0.475	0.000	1.000
Delay in primary education	0.015	0.123	0.000	1.000
Delay in secondary education	0.103	0.303	0.000	1.000
HS grades - 1st Q	0.052	0.223	0.000	1.000
HS grades - 2nd Q	0.117	0.321	0.000	1.000
HS grades - 3rd Q	0.391	0.488	0.000	1.000
HS grades - 4th Q	0.440	0.496	0.000	1.000
HS track - dropout	0.115	0.319	0.000	1.000
HS track - vocational	0.168	0.374	0.000	1.000
HS track - technical	0.294	0.456	0.000	1.000
HS track - general	0.423	0.494	0.000	1.000

Notes: Summary statistics of 8,189 observations from the SONAR data (cohorts: 1976, 1978, 1980). Delay in primary and secondary education includes information about grade retention based on the age of individuals. Higher secondary education grades are self-reported grades, which gives us information about the position of individuals relative to their peers in higher secondary education, expressed in quarters (where 1st Q is the best quarter and 4th Q the worst).

Tertiary education Based on observed characteristics and initial conditions, individuals choose to enroll in a specific program in tertiary education. We include choices for 6 academic years during college education, as this includes individuals who complete tertiary education within 6 years (by the age of 23) and start working before the age of 23. We define 11 tertiary programs, as included in Table 2. These programs are combinations of 6 majors offered at 2 types of tertiary education institutions. We distinguish between an academic track in university (Univ) with a non-academic track at vocationally oriented colleges (Col).

We choose to differentiate between STEM (STEM) and Health (HEA) degrees because previous research suggests that HEA degrees perform much better than STEM degrees regarding skill mismatch and graduation rates (Verhaest et al., 2017). We also include Education (EDU) as a distinct program because of the institutional setting of the Belgian system.

Table 2: Descriptive statistics: college major and 1st year success rate

1st year success rate	Failed	Passed(s)	Passed(d)	Passed(h)	Total
HUMA(Col)	50.0	35.7	12.6	1.7	100.0
HUMA(Univ)	52.5	31.3	12.7	3.5	100.0
BULA(Col)	50.5	38.3	10.1	1.1	100.0
BULA(Univ)	51.4	35.4	10.5	2.7	100.0
SSOC(Col)	52.9	37.6	9.2	0.3	100.0
SSOC(Univ)	49.5	35.4	12.2	2.9	100.0
HEA(Col)	41.7	35.8	18.9	3.6	100.0
HEA(Univ)	44.0	33.9	12.9	9.3	100.0
STEM(Col)	51.9	31.3	13.4	3.5	100.0
STEM(Univ)	51.4	30.0	14.3	4.3	100.0
EDU	48.3	39.9	10.7	1.1	100.0
Total	49.7	35.3	12.4	2.7	100.0

Notes: The college majors are the following: Humanities and arts (HUMA), Business and Law (BULA), Social sciences (SSOC), Health and biomedical sciences (HEA), Science, technology, engineering and mathematics (STEM), Education (EDU). (Univ) stands for academic degree in university, (Col) stands for non-academic track at vocationally oriented collages. The columns represent: (i) Passed (s) - Passed satisfactorily, (ii) Passed (d) - Passed with distinction, (iii) Passed (h) - Passed with highest distinction.

In Table 2, we highlight a key feature of the Belgian higher education system. Since universities and colleges lack the autonomy to screen students prior to admission, selection occurs ex post through the awarding of credits based on academic performance during the early years of higher education. As a result, student success rates are particularly low after the first year. As shown in Table 2, only about 50% of students successfully complete the required coursework in their first year. This period is marked by high rates of dropout and program reorientation. Our data track student performance at the end of each academic year, categorizing outcomes as follows: failed, passed satisfactorily (Passed (s)), passed with distinction (Passed (d)), and passed with high or highest distinction (Passed (h)). These classifications apply consistently throughout all years of higher education, including both bachelor's and master's programs completion. This allows us to observe the full academic trajectory of each student, including annual grades and the final grade obtained upon completion of a degree. Moreover, it allows us to track students who fail or switch programs.

Table 3 shows the choices of individuals who fail the first year. A large fraction (47%) drops out of college after an unsuccessful first year. This differs by college ma-

Table 3: Descriptive statistics: individuals who fail the first year and Bachelor's degree

1st year program	Drop	Switch	Stay	Total
HUMA(Col)	45.6	34.0	20.4	100.0
HUMA(Univ)	33.6	36.9	29.5	100.0
BULA(Col)	60.8	9.9	29.3	100.0
BULA(Univ)	31.2	44.5	24.3	100.0
SSOC(Col)	55.6	17.3	27.2	100.0
SSOC(Univ)	27.9	51.9	20.1	100.0
HEA(Col)	51.8	18.1	30.2	100.0
HEA(Univ)	22.9	56.0	21.1	100.0
STEM(Col)	57.3	10.9	31.8	100.0
STEM(Univ)	23.9	46.9	29.1	100.0
EDU	59.6	15.6	24.8	100.0
Total	47.0	26.0	27.0	100.0

Notes: The college majors are the following: Humanities and arts (HUMA), Business and Law (BULA), Social sciences (SSOC), Health and biomedical sciences (HEA), Science, technology, engineering and mathematics (STEM), Education (EDU). (Univ) stands for academic degree in university, (Col) stands for non-academic track at vocationally oriented collages.

jor and it is as high as 60.8% for college programs in Business and Law in colleges, while substantially lower for university programs in STEM or Health (around 23%). Drop-out rates are substantially larger for individuals who do not pass the first year in college programs. However, despite a failed first year, other individuals attain a Bachelor's degree in a different major (26%) or in the same program (27%). Switching rates are substantially larger for individuals who do not pass the first year in university programs. This ability sorting patterns reflects a cascade effect (as called in Declercq and Verboven, 2018): students start in the more difficult majors (in university) and update their choices depending on their grades.

For those who pass the first year, in Appendix Table 13, the large majority (87.4%) attain a Bachelor's degree in the same program. Only a smaller percentage choose to drop out later on (9.5%). However, this average covers a substantial heterogeneity across programs: drop out rates range from 4.4% in a university program in Humanities and arts to a 19.4% in a university program in Health. A residual proportion of students choose to earn a Bachelor's degree in a different program despite a successful first year.

On average, Appendix Table 14 shows the reorientation patterns and the completion rate at the Bachelor's level for different choices in the first year of college. Relative to Declercq and Verboven (2018), we find similar rates of college enrollment and reorientation. For individuals starting in a universtity (college) program, the probability to get a Bachelorś degree are approximatley 60% (50%). Individuals may acquire credits (i.e. one year of completed study) and we observe the results each year and the potential changes at both field of study or institution.

Skill mismatch Regarding skill mismatch, we rely on a large set of measures. For vertical mismatch, based on Navarini and Verhaest (2023), we measure it using a composite measure based on a latent factor approach, which includes a job-analysis (JA) measure, a direct self-assessment (DSA) measure, and an indirect self-assessment (ISA) measure. The JA measure is determined by comparing job requirements with the level of education. Jobs are classified based on the Standard Occupation Classification of Statistics Netherlands.[6] The classification groups jobs based on five educational levels. The DSA measure is derived from the survey question: "According to your opinion, do you have a level of education that is too high, too low or appropriate for your job?". The ISA measure is based on the survey question: "What is (was), in your opinion, the most appropriate educational level to execute your first job?" As this question was not included in the survey for the 1976 cohort, we implemented a modified procedure following Baert et al. (2013).[7] Relying on these measures, we construct a latent factor measure, which controls for measurement error and assumes each of the three usual measures to capture one dimension of vertical mismatch (see Navarini and Verhaest, 2023).

Furthermore, we rely on two measures for horizontal mismatch: job analysis (JA) and direct self-assessment (DSA). Each horizontal mismatch measure includes three possible outcomes: complete mismatch, somewhat match, or complete match.

In our approach, we assume that these 5 measures capture a latent variable of skill

[6] Link to the dataset: Dutch Standard Classification of Occupations (SBC) 1992 (Last accessed: 15.02.2023)

[7] First, we calculate each occupation's mean self-assessed required level based on the available information. To do this, we rely on the aforementioned five categories of education levels. Second, we extrapolate this mean to all jobs in each occupation. Third, we classify an individual as being overeducated if their attained level of education exceeds the mean required level within their occupation.

mismatch, measured with an error:

$$\text{mismatch}_i^j = \beta^j \mu_i + \varepsilon_i^j \text{ for } j \in J, \tag{1}$$

where J includes three measures of vertical mismatch (JA, DSA, ISA) and two measures of horizontal mismatch (JA, DSA). Using this approach, we minimize the measurement error regarding skill mismatch and we provide a clear way of interpreting all these different measures. We classify an individual as skill mismatched with a dummy variable if the latent variable μ_i is greater than 0. We also conduct robustness checks to confirm that our results hold when using indicators such as JA or DSA. In this context, a higher value of the latent mismatch factor increases the likelihood of both vertical and horizontal mismatch. Therefore, throughout the rest of the paper, I will also refer to skill mismatch as “full mismatch", that is, a higher probability of experiencing both types of mismatch.

Labor market outcomes Our main interest is the impact of the first job match quality on subsequent labor market outcomes up until age 29. This includes the probability of persistent skill mismatch after the first job, but, potentially, also the scarring effect on subsequent wages even if the individuals find an adequately matched job. We follow Heckman et al. (2006, 2018a) and focus our analysis on individuals before the age of 30.

In Table 4, we include the endogenous labor market outcomes. First, we include potential experience at ages 23, 26 and 29. We follow Adda and Dustmann (2023) and construct potential experience by considering the age of entry in the labor market of each individual. This captures the late entry into the labor market of individuals with a degree, but also of individuals who dropped out of college. On average, at age 23, individuals have 1.79 years of potential experience. This means that the average individual enter the labor market around at the age of 21. This increases to 4.38 years of potential experience at 26 and 7.15 at 29.

Second, we account for unemployment rate at ages 23, 26 and 29. For those with employment information, unemployment rate is as high as 42.8% at age 23 and it decreases to 21.5% at 26 and 10.4% at 29. This accounts for differences between matched and mismatched individuals in their career and in the probability of finding a well-matched job. At last, we include wage selection and log-hourly wages. Of those with employment information at age 23, 26 and 29, we observe 74.7% of the wages at age 23,

Table 4: Descriptive statistics: labor market outcomes

	Mean	SD	Min	Max
Potential experience at age 23	2.117	1.680	0.000	6.000
Unemployment at age 23	0.428	0.495	0.000	1.000
Wage observed at age 23	0.747	0.435	0.000	1.000
Hourly wage at age 23	7.354	1.590	2.791	19.963
Potential experience at age 26	4.367	2.115	0.000	9.000
Unemployment at age 26	0.215	0.411	0.000	1.000
Wage observed at age 26	0.894	0.308	0.000	1.000
Hourly wage at age 26	8.126	1.859	2.927	19.493
Potential experience at age 29	7.462	2.163	0.000	12.000
Unemployment at age 29	0.104	0.305	0.000	1.000
Wage observed at age 29	0.982	0.134	0.000	1.000
Hourly wage at age 29	8.563	1.855	2.853	20.007

Notes: Summary statistics of 8,189 observations from the SONAR data (cohorts: 1976, 1978, 1980). Potential experience is calculated from the age of entry in the labor market, as in Adda and Dustmann (2023). At each age, we observe unemployment and, for those employed, we have a probability of not observing the wage: we take this into account by modelling it.

89.4% at age 26 and 98.2% at age 29.

3 Model

We study the causal returns to college majors, jointly with the causal impact of skill mismatch on subsequent wages, using a dynamic model. We estimate dynamic treatment effects using a dynamic model of human capital accumulation and labor market outcomes, identified using a set of instruments to address the non-random selection into educational choices and skill mismatch (Heckman and Navarro, 2007; Aakvik et al., 2005; Humphries et al., 2023).

First, we specify the model starting from a potential outcome framework. Second, we present the full specification of our model. Third, we discuss how to credibly estimate causal effects, while recovering unobserved heterogeneity and controlling for non-random (dynamic) selection through a dynamic model. Moreover, we present initial evidence supporting our choice of instruments, relative distance to higher education and the timing of labor market entry. Fourth, we introduce the treatment effects of interest of this paper.

3.1 Potential outcomes framework

Given the attainment of an educational program $E_i = j$, let $Y_{ijt,1}$ be the log hourly wage for individual i with attainment j at age t when first entering the labor market in a mismatched job ($D_i = 1$), and $Y_{ijt,0}$ be the log hourly wage when entering in an adequately matched job ($D_i = 0$).[8] For each individual i, we only observe the realized wage Y_{it} at age t:

$$Y_{it} = \sum_{j}^{J} \mathbf{1}(E_i = j)(D_i Y_{1ijt} + (1 - D_i) Y_{0ijt}), \tag{2}$$

where E_i represents the educational attainment of individual i from a set of $j \in J$ educational programs available. Moreover, D_i and Y_{ijt} are a function of observables (respectively, Z and X), and persistent unobservables (θ).

Wages Y_{1ijt} and Y_{0ijt} capture the potential scarring effect of starting a career in an occupation requiring a different educational level or field of study. To address endogeneity issues, we model the impact of the first occupations on subsequent wages at subsequent ages (t might be interpreted as any age after the first job). Therefore, individuals who enter the labor market with a mismatched occupation might experience lower wage growth even if they later switch to adequately matched occupation later.

It is possible to think of treatment D_i as a discrete variable (i.e., full mismatch, only vertical or only horizontal mismatch) or as a binary treatment, depending on the measure at the disposal of the researcher. Regarding educational choices, this framework allows for substantial flexibility: E_i might be considered as years of schooling, educational attainment level, specific educational programs or combinations of college majors and level of attainment.

As noted in Leuven and Oosterbeek (2011), there is a non-random selection of individuals into the completed educational attainment, E_i, and into the skill mismatch treatment, D_i. Addressing this endogeneity problem is far from trivial: endogeneity arises because there exists a common unobserved component, θ_i, which jointly generate educational choices E_i, non-random sorting into skill mismatch in the first job D_i, and subsequent potential wages, $Y_{ijt,1}$ and $Y_{ijt,0}$. To identify causal effects of educational choices and skill mismatch on wages, we need to identify and control for this unobserved factor (see Section 3.3). As in Heckman et al. (2018a), conditional on θ, X, Z,

[8] Educational attainment j can be interpreted both as years of schooling or any possible combination between educational attainment and specific educational programs, such as higher secondary education tracks or college majors.

choices and outcomes are statistically independent: controlling for this set of variables eliminates selection effects.

Dynamic model Based on observables X and unobservables θ, individuals make a sequence of educational choices, enter the labor market in a specific occupation and realize a sequence of wages up until age t. We present this process into three main stages.

In the first stage, individuals makes a sequence of choices (C^{ht}) regarding their higher educational attainment. This includes choices about college majors, institutions, major switching behavior and academic results. In a second stage, individuals enter the labor market in a mismatched job ($D_i = 1$) or in an adequately matched job ($D_i = 0$). In the spirit of Aakvik et al. (2005), this can be thought of as a selection-into-treatment equations, based on observables and unobservable characteristics. Moreover, it includes a random shocks, which may be indicative of matching frictions or non-persistent labor market shocks. In the third and final stage, based on their educational choices and the skill mismatch treatment, they realize wages up until age t which are specific to a career trajectory if started in a mismatched occupation or not. These three main stages can be represented by a system of core equations:

$$\begin{cases} C^{ht} &= \Psi^{ht}\Big(\{\phi^{ht}(X,\theta,Z^{ht})\}_{l^{ht}\in D^{ht}}, \varepsilon^{ht}_{l^{ht}}\Big), \\ D &= \Psi^{d}\Big(\{\phi^{e}(X,\theta,C^{ht},Z^{d})\}_{l^{d}\in D}, \varepsilon^{d}_{l^{d}}\Big), \\ Y_{1it} &= f^{1w}(X,\theta,D^{ht},Z^{1w}) + \varepsilon^{1w}, \\ Y_{0it} &= f^{0w}(X,\theta,D^{ht},Z^{0w}) + \varepsilon^{0w}, \end{cases} \tag{3}$$

where θ represents a set of unobserved characteristics which jointly determines educational choices in secondary and tertiary education, the skill mismatch treatment and the relative potential wages up until age 29, given the initial matching quality at the labor market entry. Without accounting for θ, the estimates of college majors or skill mismatch are biased. Decisions function Ψ allow for different choice models (binary, discrete or ordered).

When exploiting quasi-experimental variation from admission cutoffs (e.g., Kirkebøen et al., 2016), we can identify reduced-form estimates of the causal returns to college majors. This approach addresses endogeneity in major choice and recovers the local average treatment effect (LATE) of enrolling in a particular major. However, it does not control for the non-random selection into skill mismatch after graduation. Since individuals with different unobserved abilities or preferences may sort differently into occupations, even conditional on their major, reduced-form estimates based on admission cutoffs capture both the direct effect of major choice and the indirect effect through differential selection into matched or mismatched jobs.

3.2 Model stages

In this section, we present the full specification of three main key stages in reverse order: first, we introduce labor market outcomes up until age 29, including the specification for log-hourly wages, followed by the skill mismatch treatment equation, and finally the higher education stage. For the sake of clarity, we remove the subscript i.

Labor market outcomes: wages, employment selection, unemployment and potential experience We model Y_{st} as log-hourly wages earned by individual i at age t with college major diploma j, working in either an adequately ($s = 0$) or a mismatched occupation ($s = 1$):

$$
\begin{aligned}
Y_{st} = \log(\text{wage}_{st}) = {} & \alpha_0^s + \alpha^{s,X} X + \alpha^{s,L} L^w + \alpha^{s,\text{delay}^{se}} \text{delay}^{se} \\
& + \text{hs\_track}^{hs}(\alpha^{s,\text{hs\_track}} + \alpha^{s,\text{hs\_track}\times X} X) + \alpha^{s,\text{hs\_grade}} \text{hs\_grade}^{hs} \\
& + \text{major}_j(\alpha^{s,j} + \alpha^{s,j\times X} X) \\
& + \alpha^{s,\text{BA\_grade}} \text{BA\_grade} + \alpha^{s,\text{MA}} \text{MA} + \alpha^{s,\text{MA\_grade}} \text{MA\_grade} \\
& + \alpha^{s,\exp_t} \exp_t + \phi_t^s + \alpha^{\theta,s,m} \theta^{s,m} + \varepsilon_t^s \ \text{ for } s \in \{0,1\},
\end{aligned}
$$

with a set of secondary education variables, such as grade repetition in secondary education (delay^{se}), high school track (hs_track^{hs}) with heterogeneous effects based on observable characteristics (X) and the final high school grade (hs_grade^{hs}). Post-secondary educational attainment is captured through the completed college major (major_j), the corresponding Bachelor's grade (BA_grade), a Master's degree dummy (MA_grade) and the Master's grade (MA_grade), if applicable. We also control for potential experience in years at each age t ($\exp_t$), time fixed effects (ϕ_t^s), unobserved heterogeneity (θ^m) and an idiosyncratic error term (ε_t^s). Of course, each coefficient is mismatch specific and reflects the different wage setting of observationally identical individuals entering the labor market with a mismatched position.

This specification allows us to identify the returns to college majors by explicitly controlling for a rich set of pre-college and post-college characteristics, including secondary school track, grades, delays, and demographic factors. By including fixed effects for the college major and their interactions with individual characteristics, we can isolate the contribution of the major itself to wage outcomes, net of selection on observables. The inclusion of experience and fixed effects by time and unobserved type further helps disentangle major-specific returns from general labor market trends and persistent individual heterogeneity, which affects educational choices and other labor market outcomes.

Besides wages, we account for three additional labor market outcomes: unemployment, endogenous potential experience in years and a wage selection equation at ages 23, 26, and 29. The specification for these outcomes is similar:

$$\begin{aligned}
Y_{o,jt} = {} & \alpha_0^o + \alpha^{o,X} X + \alpha^{o,L} L^w + \alpha^{o,\text{delay}^{se}} \text{delay}^{se} \\
& + \alpha^{o,\text{hs_track}} \text{hs_track}^{hs} + \alpha^{o,\text{hs_grade}} \text{hs_grade}^{hs} \\
& + \alpha^{o,j} \text{major}_j + \alpha^{o,\text{BA_grade}} \text{BA_grade} \\
& + \alpha^{o,\text{MA}} \text{MA} + \alpha^{o,\text{MA_grade}} \text{MA_grade} \\
& + \alpha^{o,\text{mismatch}} \text{mismatch} \\
& + \phi_t^o + \alpha^{\theta,o,m} \theta^{o,m} + \varepsilon_{jt}^o \qquad \text{for } o \in \{\text{unem}, \text{exp}, \text{sele}\},
\end{aligned}$$

where these outcomes incorporate the effect of skill mismatch (mismatch), which, unlike wages, are not treated as potential outcomes but are essential for identifying unobserved heterogeneity (θ) and the risk of mismatch. These variables allow us to characterize heterogeneous employment trajectories from age 23 to 29 and help distinguish between observed and unobserved sources of variation in labor market outcomes. Controlling for unemployment, work experience accumulation, and selection into employment enables us to account for dynamic feedback effects and selection bias that would otherwise confound estimates of the causal impact of college major choice on wages. In particular, individuals with different unobserved abilities or preferences may self-select into majors and subsequently experience different employment dynamics. Modeling these additional outcomes improves identification and robustness in the estimation of heterogeneous returns to college.

Skill mismatch Skill mismatch is not a choice per se; rather, it may arise as an outcome of the assignment process in the labor market. Following the framework of Aakvik et al. (2005) and Heckman et al. (2018a), we model mismatch as a treatment assigned through a threshold-crossing index function. That is, individuals are selected into mismatch ($D_i = 1$) if their latent index falls below a threshold:

$$D_i = \begin{cases} 0 & \text{if } I_{jt} \geq 0 \\ 1 & \text{otherwise} \end{cases} \tag{4}$$

We approximate the latent index I_{jt} using a separable linear model:

$$\begin{aligned} I_{jt}^{mm} = \alpha_0^{\text{mm}} &+ \alpha^{\text{mm},X} X + \alpha^{\text{mm},L} L^w + \alpha^{\text{mm,delay}^{se}} \text{delay}^{se} \\ &+ \alpha^{\text{mm,hs\_track}} \text{hs\_track}^{hs} + \alpha^{\text{mm,hs\_grade}} \text{hs\_grade}^{hs} \\ &+ \alpha^{\text{mm},j} \text{major}_j + \alpha^{\text{mm,BA\_grade}} \text{BA\_grade} \\ &+ \alpha^{\text{mm,MA}} \text{MA} + \alpha^{\text{mm,MA\_grade}} \text{MA\_grade} \\ &+ \phi_t^{\text{mm}} + \alpha^{\theta,mm,m} \theta^{\text{mm},m} + \varepsilon_{jt}^{\text{mm}}. \end{aligned}$$

where, this specification captures how individual and educational characteristics jointly influence the likelihood of being mismatched in the labor market. Modeling mismatch selection is crucial for isolating the causal effects of college major on labor market outcomes, as individuals self-select into educational paths that influence not only wages but also the probability of mismatch. This helps correct for selection bias and allows for a cleaner interpretation of heterogeneous returns to college education.

higher education Following Declercq and Verboven (2018), we fully exploit the dynamic sequence of individuals choices throughout their higher education phase and their corresponding academic outcomes at the end of each year. Students may update their decisions annually based on academic performance and accumulated credits, choosing to continue in their current program, switch to a different study option, or drop out entirely. If they drop out, based on their potential experience, they enter the labor market and earn wage, as described in the previous paragraphs.

As in Arcidiacono (2004) and Declercq and Verboven (2018), in each academic year a, a student selects an option $j \in \{0, 1, \ldots, J\}$, where $j = 0$ denotes the decision to drop out and $j > 0$ represents the various study programs available. The decision in year a

is captured by a vector $d_a = (d_a^0, d_a^1, \ldots, d_a^J)$, where d_a^j is a binary indicator equal to 1 if option j is chosen and 0 otherwise. The flow utility associated with each study option $j > 0$ in year a is composed of current consumption benefits and associated costs. It is modeled as:

$$\begin{aligned} u_a^j(X, Z_a^{ht}) = \alpha_{0j}^{\text{ht}} &+ \alpha_j^{\text{ht},X} X + \alpha_j^{\text{ht},L} L^{ht} + \alpha_j^{\text{ht,delay}^{se}} \cdot \text{delay}^{se} \\ &+ \alpha_j^{\text{ht,hs\_track}} \cdot \text{hs\_track}^{hs} + \alpha_j^{\text{ht,hs\_grade}} \cdot \text{hs\_grade}^{hs} \\ &+ \alpha^{ht,C} C^j + \alpha^{ht,E} E_a + \alpha_j^{ht,d} d_{j,a-1} + \theta_j^m + \psi_{ja} + \varepsilon_a^j, \end{aligned}$$

where X includes individual background characteristics and Z_a^{ht} includes outcome- and age-specific variables, such as L^{ht} local labor market conditions, delay^{se} delays in secondary education, together with tracks and grades in higher secondary education. The variable C^j represents travel or access costs to program j, following Declercq and Verboven (2018). The term E_a measures the number of accumulated academic credits at the beginning of year a, which reflects progress and ability, as emphasized in Arcidiacono (2004). Consistent with Altonji et al. (2016a), the inclusion of $d_{j,a-1}$ captures switching costs associated with continuing or changing ones major.

Each year, students also receive a grade reflecting their academic performance, categorized as not passed, passed with satisfactory, intermediate, or highest marks. These outcomes serve as proxies for individual college major specific ability (as major programs directly affects grades) and influence subsequent decisions. Moreover, we also include a final treatment, which is if a student earn a degree at a BA or a MA level in the second period of graduation. If they pass a year of coursework with any grades, they collect an additional credit (E_a).

Modeling these sequential decisions and outcomes allows us to capture persistent unobserved heterogeneity among students. This helps identify latent types that reflect differences in academic preferences and abilities. As a result, the model can generate diverse trajectories and outcomes even among observationally identical individuals, enabling us to better estimate the heterogeneous returns to college majors. Relative to Arcidiacono (2004) and Declercq and Verboven (2018), we do not fully model the dynamic expectation of individuals, which enters the model only through unobserved heterogeneity, as in Heckman et al. (2018a). This also allow for a flexible specification with the advantage of being agnostic relative to agent rationality and expectations formation (Heckman et al., 2018a, 2016).

Likelihood function We estimate the full model by combining all its components into a single likelihood function and maximizing it with respect to the parameter vector $\hat{\Theta}$:

$$
\begin{aligned}
\hat{\Theta} &= \arg\max_{\Theta} \mathcal{L}(\Theta) \\
&= \arg\max_{\Theta} \prod_{i=1}^{I} \sum_{m=1}^{M} \pi_m(Z_i^0) \Bigg[\prod_{a=1}^{A_i} \prod_{j=0}^{J} \Big(P(d_{ia}^j = 1 | Z_{ia}^{ht}, \theta_m) \Big)^{d_{ia}^j} \\
&\quad \times \prod_{a=1}^{A_i} f(E_{ia} | Z_{ia}^{ht}, \theta_m) \times \prod_{a=1}^{A_i} f(\text{grade}_{ia} | Z_{ia}^{ht}, \theta_m) \\
&\quad \times f(\text{graduation}_i | Z_i^{grad}, \theta_m) \\
&\quad \times P(D_i = 1 | Z_i^{mm}, \theta_m)^{D_i} \times P(D_i = 0 | Z_i^{mm}, \theta_m)^{1-D_i} \\
&\quad \times \prod_{t=23}^{29} \prod_{s=0}^{1} \Big[f(Y_{ist} | Z_{it}^s, \theta_m)^{s \cdot D_i + (1-s)\cdot(1-D_i)} \Big]^{I_{it}} \\
&\quad \times \prod_{t=23,26,29} \prod_{o \in \{\text{unem,exp,sele}\}} f(Y_{iot} | Z_{it}^o, \theta_m) \qquad \Bigg],
\end{aligned}
$$

where we assume there are a number M of different heterogeneity types, which are conditioned on a set of secondary education variables (Z_i^0). This includes delay in primary and secondary education, higher secondary track diploma and higher secondary grade. This is equivalent to start modeling unobserved heterogeneity from the end of primary education.

We estimate this model by using an Expectation-Maximization (EM) algorithm, iterating between two main steps. In the expectation step, we compute the probability of each individual to be assigned to heterogeneity type m, based on the likelihood value for each $m \in M$. Given current parameter estimates, we compute the probability that each individual i belongs to type m using Bayes rule and the individual likelihood contribution:

$$
\hat{\tau}_{im} = P(\text{type} = m | \text{data}_i, \hat{\Theta}) = \frac{\pi_m(Z_i^0)\mathcal{L}_i^m(\hat{\Theta}, \hat{\theta}_m)}{\sum_{m=1}^{M} \pi_m(Z_i^0)\mathcal{L}_i^m(\hat{\Theta}, \hat{\theta}_m)} \tag{5}
$$

As shown by Arcidiacono and Jones (2003), this step can be implemented in stages. Once the heterogeneity probabilities $\hat{\tau}_{im}$ are treated as known and given, the likelihood becomes separable across model components, allowing for a modular estimation approach.

$$\hat{\Theta}^{(r+1)} = \arg\max_{\Theta} \sum_{i=1}^{N} \sum_{m=1}^{M} \hat{\tau}_{im}^{(r)} \log \mathcal{L}_i^m(\Theta, \theta_m) \tag{6}$$

After each maximization step, we update the posterior probabilities and iterate until convergence. To identify the optimal number of heterogeneity types m, we re-estimate the model by gradually adding up to four types to the model. Since the likelihood function may have multiple local maxima, we perform multiple estimations with different starting values and retain the model based on AIC and BIC.

3.3 Identification

For identification, we need to address at least two types of biases. First, a selection bias arises because treated individuals could differ from the control group in various respects not included by observables and educational choices for instance, unobserved preferences for a college major with higher or lower mathematics content. Second, the estimates may be biased due to dynamic selection bias. This happens through the increasing negative correlation between the treatment and the unobservable characteristics as students progress in their educational careers (Cameron and Heckman, 1998, 2001). This is especially true for the Belgian higher education system, where there is ex-post selection and those who complete a program are selected from the first year onward.

Identification: unobserved heterogeneity In this model, unobserved heterogeneity through θ induces correlation across different choices, addressing the issue of dynamic selection. The literature calls this matching on unobservables (Heckman and Navarro, 2007). Indeed, choices and outcomes of the model are correlated and this rationalizes differences in outcomes between observationally identical individuals (Aakvik et al., 2005). Given the model specification, unobserved heterogeneity enters as discrete-type random-effects, where θ is a random effect, independent of ε, and independent across individuals. This random effect captures unobserved determinants and is assumed independent of the observed exogenous individual characteristics.

Following the literature on dynamic discrete choice models, we use a finite mixture distribution to model the unobserved random variable θ_m (cf. Heckman and Singer, 1984; Arcidiacono, 2004).[9] We assume this distribution to be characterized by an a priori unknown number of M different heterogeneity types with type-specific heterogeneity parameters for each outcome. This avoids relying on strong distributional assumptions and, therefore, also minimizes any bias resulting from misspecification in this respect (Heckman and Singer, 1984; Hotz et al., 2002).

We use strategies to identify unobserved heterogeneity and correctly identify the model. First, the panel dimension of the data, specifically the autocorrelation of educational choices, occupational choices and wages given observed covariates, plays a crucial role in identifying the returns associated with skills while accounting for unobserved heterogeneity and dynamic selection. Our model starts from the first-year college major choice and, therefore, this would be the initial condition for estimating our unobserved heterogeneity. Therefore, we condition unobserved heterogeneity on a set of primary and secondary education variables, including grade repetition, grades and high school tracks. Therefore, unobserved types could be thought of having a strong effect on educational choices in both secondary and tertiary education.

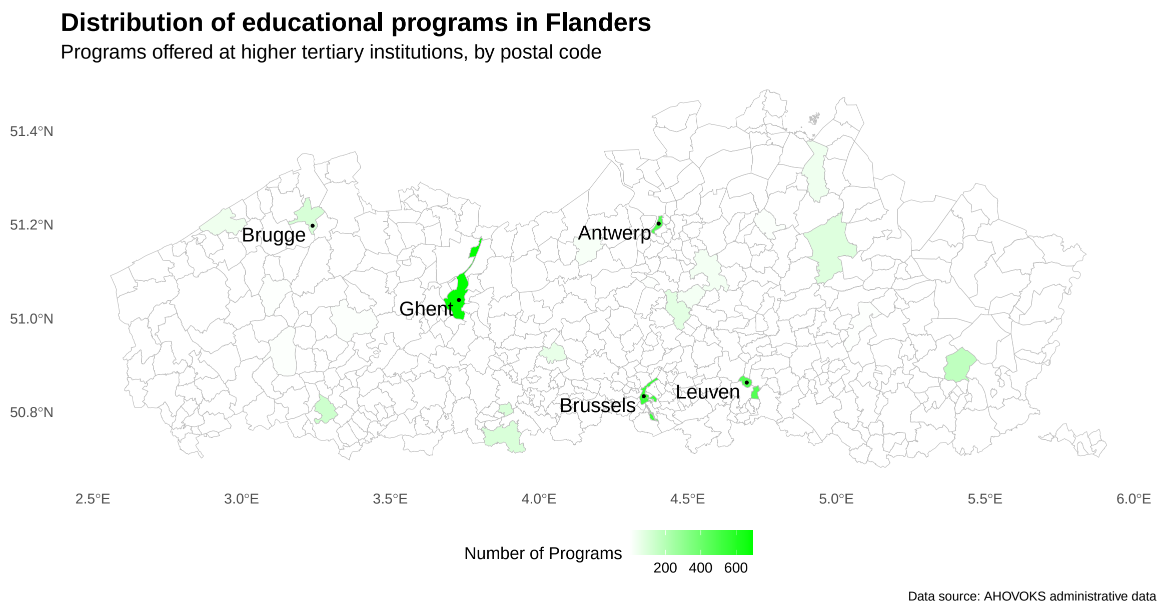
Second, we impose exclusion restrictions using variables that affect educational choices but not later outcomes, following Heckman and Navarro (2007); Heckman et al. (2016, 2018a,b) and Ashworth et al. (2021). These restrictions support identification of preference parameters and selection effects. We include relative distance to higher education that affects college enrollment choice, while the timing in the graduation affects skill mismatch. The combined use of these instruments and unobserved heterogeneity help us in estimating the causal effect of college majors and skill mismatch on wages.

Relative distance to higher education Following the literature on educational choices, we use the relative distance to higher education institutions as an instrument for higher enrollment choice. Distance affects program choice via enrollment costs, but is uncorrelated with unobserved ability or preferences conditional on background characteristics and unobserved characteristics.

Similar to previous studies (Declercq and Verboven, 2018; De Groote and Declercq, 2021), we define the relative distance as the distance to the closest university institutions, subtracted by the distance to the closest college institution (in kilometers). Indeed, as pointed out in Declercq and Verboven (2018), relative distance is the most important factor for students in deciding at which higher institution to enroll, because

[9] It enters each likelihood contribution as a constant parameter, but, given the probability weight for each observation, it becomes a dummy capturing type-specific shocks.

Figure 1: Distribution of educational programs in Flanders



of public funding on the basis of enrolled students, no capacity constraints and uniform standards. Figure 1 includes the number of programs offered by postal code in Flanders from AHOVOKS administrative data on the entire history of officially recognized higher programs. For example, a student living in a postal code close to college campuses, but no university campus, will be more likely to choose for a college program than a student living in an area with access to both college and university programs (Declercq and Verboven, 2018).

We check the minimal set of assumptions of the instrument, such that it is relevant and independent from the unobservable. The first condition implies that distance should have a strong impact on higher education enrollment choice. Table 5 shows the results of the first stage regression, with different specifications, estimating the impact of relative distance on first-year college major choices. In odd columns, the model estimate a single coefficient for each higher education program, while, in even columns, the coefficient is allowed to be program specific. This estimate does not change substantially when controlling for observed student characteristics (which is what we would expect from a valid instrument).

Travel distance significantly affects both the decision to pursue higher education and the type of institution chosen. For example, students in areas with only colleges

Table 5: First stage estimates: distance and college major choice

	(1)	(2)	(3)	(4)	(5)	(6)
Humanities and Arts (Col)						
Relative distance to closest tertiary education	-0.007***	-0.019***	-0.006***	-0.018***	-0.005***	-0.006
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Humanities and Arts (Univ)						
Relative distance to closest tertiary education	-0.007***	-0.008***	-0.006***	-0.007**	-0.005***	-0.011**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.004)
Business and Law (Col)						
Relative distance to closest tertiary education	-0.007***	-0.006***	-0.006***	-0.005***	-0.005***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Business and Law (Univ)						
Relative distance to closest tertiary education	-0.007***	-0.014***	-0.006***	-0.012***	-0.005***	-0.006*
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Social sciences (Col)						
Relative distance to closest tertiary education	-0.007***	-0.004*	-0.006***	-0.004	-0.005***	-0.006
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Social sciences (Univ)						
Relative distance to closest tertiary education	-0.007***	-0.011***	-0.006***	-0.009***	-0.005***	-0.012***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Health (Col)						
Relative distance to closest tertiary education	-0.007***	-0.001	-0.006***	-0.001	-0.005***	-0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Health (Univ)						
Relative distance to closest tertiary education	-0.007***	-0.000	-0.006***	0.004	-0.005***	0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
STEM (Col)						
Relative distance to closest tertiary education	-0.007***	-0.004**	-0.006***	-0.004**	-0.005***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
STEM (Univ)						
Relative distance to closest tertiary education	-0.007***	-0.013***	-0.006***	-0.011***	-0.005***	-0.015***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Education						
Relative distance to closest tertiary education	-0.007***	-0.004**	-0.006***	-0.004*	-0.005***	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
Exogenous variables	No	No	Yes	Yes	Yes	Yes
Province FE	No	No	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	57323	57323	57323	57323	57029	57029

Notes: Exogenous variables include gender, foreign background, number of siblings, parental background, day of birth. Province FE includes provicinal fixed effects. Relative distance to closest tertiary education is measure in kilometers. Each college major is either (Col), vocational university, or (Univ), academic university. Education is a single program.

are more likely to enroll in college programs than those with access to both universities and colleges (Declercq and Verboven, 2018). To validate the instrument, we verify both its relevance and exogeneity. Appendix Table 20 shows that it is reasonable to exclude relative distance to closest tertiary education from labor market outcomes, as this is not the relevant variable explaining these outcomes and the estimated coefficients are all around zero. The second condition implies that differences in distances must be independent of unobservables in the outcome equation. As in De Groote and Declercq (2021), we check this by regressing relative distance on exogenous variables, showing that it is not correlated with observed student characteristics, as shown in Appendix Table 21.

Graduation timing As emphasized by Leuven and Oosterbeek (2011), selection into skill mismatch (D_i) is not random, even after conditioning on rich educational and demographic covariates. To identify the causal effect of skill mismatch in the first job on early-career wages, we exploit quasi-random variation in labor market entry timing induced by the institutional structure of Belgian universities. Specifically, students who fail one or more courses during the regular academic session can retake them in a second sit session held in August-September. Although this allows them to graduate within the same academic year, it delays labor market entry by several months compared to peers who pass all exams in June.

In a standard instrumental variables framework, second sit graduation would not qualify as a valid instrument, given its correlation with unobserved ability. However, in our setting, we explicitly model this endogeneity and use the residual variation in labor market entry timing to jointly identify both unobserved ability and selection into skill mismatch.

To formalize our approach, consider the following system of equations, which is a simplified version of our full dynamic model:

$$
\begin{aligned}
\text{Graduation timing:} \quad & Z_i = \mathbf{1}[X_i'\gamma + \delta\theta_i + u_i > 0] \\
\text{Skill mismatch:} \quad & D_i = \mathbf{1}[\alpha Z_i + X_i'\beta + \psi\theta_i + \varepsilon_i > 0] \\
\text{Wage function:} \quad & Y_{it} = \begin{cases} X_{it}'\phi^0 + \lambda^0\theta_i + \eta_{it}^0 & \text{if } D_i = 0 \\ X_{it}'\phi^1 + \lambda^1\theta_i + \eta_{it}^1 & \text{if } D_i = 1, \end{cases}
\end{aligned}
$$

where, θ_i is a latent factor that captures unobserved ability, preferences, or motivation, which affects graduation timing (Z_i), skill mismatch (D_i), and wages (Y_i).[10] The error term u_i introduces idiosyncratic shocks, such as illness or temporary family events, that influence graduation timing but are uncorrelated with persistent unobserved and

[10] In our full benchmark model, this latent factor also drives choices in secondary and tertiary education, together with labor market outcomes, such as unemployment and potential work experience.

observed characteristics, θ_i and X_i. Conditional on θ_i and X_i, this random shock will make identical individuals sort differently in mismatched occupations (through α) and it will isolate the causal impact of mismatch from the non-random selection into late graduation and skill mismatch. By comparing wage trajectories of individuals with the same observed characteristics but differing in graduation timing and mismatch status, we isolate the effect of mismatch on wages net of selection. Because graduation timing does not perfectly predict mismatch and does not directly affect wages, conditional on unobservable and observables, it provides variation in mismatch orthogonal to latent type, enabling the model to attribute the remaining correlation between mismatch and wages to unobserved heterogeneity.

To illustrate this identification argument intuitively, consider two students, A and B, who both study economics, have the same grades over different years, with the same set of observed characteristics (X_i). A passes all her exams in June and enters the labor market during the main hiring season. B, who fails one exam, retakes it in the August-September second sit, graduating later and entering the market off-cycle. Of course, this might be rationalized by individuals having different unobserved ability (θ_i), where higher ability individuals might have different secondary education outcomes, choose different majors, graduate faster, find better jobs, with higher potential wages. However, if A and B are individuals who shares the same observed (X_i) and unobserved characteristics (θ_i), the only remaining variation which made them sort into late graduation is u_i, which is a random shock uncorrelated with persistent unobserved heterogeneity, mismatch and wages.

The model captures this variation and it disentangles the persistent variation from the random shocks: θ_i is identified by using individuals with similar state variables, but different outcomes. But, as in Heckman et al. (2016), conditional on θ, we identify the causal effect of mismatch on wages, by using the as-good-as-random allocation into graduation timing, given by factors which are uncorrelated with observed and unobserved characteristics (i.e. temporary family shocks or illnesses). Then, the difference in wages is attributable to mismatch caused by graduation timing rather than to unobserved ability.

Table 6 presents first-stage evidence that late graduation significantly increases the likelihood of mismatch, conditional on a rich set of educational and demographic controls. Graduating later may disadvantage students in securing a well-matched job relative to peers graduating earlier in the hiring cycle, because high-quality positions may be filled earlier in the hiring cycle.

Table 6: First stage estimates: graduation timing and skill mismatch

	(1) Skill mismatch (first job)	(2) Skill mismatch (first job)	(3) Skill mismatch (first job)
Graduated in the second period	0.149***	0.143***	0.098***
	(0.021)	(0.021)	(0.021)
Exogenous variables	No	Yes	Yes
Educational attainment	No	No	Yes
Observations	3738	3738	3738
R^2	0.013	0.015	0.071

Notes: Exogenous variables include gender, foreign background, number of siblings, parental background, day of birth. Educational attainment includes college majors program in BA and MA, grades in BA and MA, and secondary education variables. Each college major is either (Col), vocational university, or (Univ), academic university. Education is a single program.

A potential concern is that some individuals may delay graduation because they are already employed and do not require a degree to access a specific occupation. However, this scenario appears unlikely in our context. During our observation window (graduation years 1997-2004, prior to the Bologna reform), 95% of individuals graduating with a bachelor's degree and 97% with a master's degree report no prior work experience. Furthermore, 98% and 99%, respectively, report no prior full-time employment contracts. This pattern suggests that most students enter the labor market only after completing their studies, and that late graduation is not systematically related to pre-existing employment.

At last, in the spirit of Heckman et al. (2018a) we include unemployment rate at the district level for each choice and outcome in the model. Using unemployment rate shocks at the district level, we aim at isolating possible local labor market shocks at the selection-into-treatment equation, without affecting future wages. The combined use of these potentially non-excluded exclusion restrictions (quasi-IV) with the unemployment rate at the district level is necessary to identify unobserved ability, as explained by Bruneel-Zupanc and Beyhum (2024); Bruneel-Zupanc (2023).

4 Results

In this section, we first present the benchmark model estimates. Second, we report the average returns to each college program, after controlling for unobserved heterogeneity and dynamic selection. Then, we examine differences within and across college majors in the distribution of returns. Using counterfactual simulations, we estimate the distribution of individual returns and the fraction of negative returns, that is, individuals who would have benefited more from a high school degree than from a specific college program in terms of hourly wages. Third, we relate these findings to skill mismatch, focusing on the probability of mismatch by college major and how college major influence the probability of receiving a negative return. Fourth, we explore further heterogeneity in returns, including analyses of individuals at the margin of choice and differences by unobserved types. At last, we present a series of robustness checks to validate our main results.

4.1 Model estimates

Initial conditions We estimate our benchmark model including three unobserved types. Unobserved types are conditioned on a set of initial conditions and they are, therefore, correlated with variables from primary and secondary education. As mentioned already, this is equivalent to estimate unobserved heterogeneity starting from primary education. Correlations from Appendix Table 22 indicate that individuals in Types 2 and 3 are more likely to have repeated a year in either primary or secondary education, to have dropped out of higher secondary education, and to have a lower probability of obtaining an academic higher secondary diploma (i.e. low achievers). They are also more likely to report grades in the 3rd and 4th quartiles relative to their class. In contrast, individuals in Type 1 have a lower probability of grade repetition, are more likely to complete an academic secondary diploma, and tend to achieve better grades overall. These individuals may be viewed as having higher secondary education ability (i.e. high achievers).

Higher education Table 7 reports the estimates from the higher education choice model stage. Despite substantial differences in secondary education achievement, Type 1 individuals exhibit significantly lower probabilities of enrolling in most tertiary education programs relative to Type 2 and 3. Type 3 individuals have the highest probability of enrolling in all college programs.

We find that higher relative travel distance (in km) has a significant and negative effect on enrolling in higher education. Moreover, Table 7 includes the switching cost coefficients of going choosing a university program in a different field compared to staying in Humanities. These estimates are consistently large and negative across all

Table 7: Higher tertiary education regression

	HUM (Col)	HUM (Univ)	BUL (Col)	BUL (Univ)	SSOC (Col)	SSOC (Univ)	HEA (Col)	HEA (Univ)	STEM (Col)	STEM (Univ)	EDU
Credits	10.446***										
	(0.323)										
Travel distance	-0.121***										
	(0.010)										
Unemployment rate	-0.014***										
	(0.002)										
HUM (a-1)			-3.343***		-3.552***		-4.508***		-3.895***		-3.339***
			(0.112)		(0.169)		(0.256)		(0.157)		(0.140)
BUL (a-1)	-4.105***				-3.235***		-4.326***		-4.030***		-3.737***
	(0.180)				(0.128)		(0.205)		(0.145)		(0.149)
SSOC (a-1)	-3.956***		-3.303***				-4.083***		-5.119***		-3.459***
	(0.194)		(0.112)				(0.208)		(0.295)		(0.149)
HEA (a-1)	-4.859***		-4.124***		-4.019***				-4.222***		-4.031***
	(0.283)		(0.151)		(0.197)				(0.184)		(0.183)
STEM (a-1)	-4.076***		-3.895***		-4.338***		-3.815***				-3.789***
	(0.184)		(0.129)		(0.225)		(0.174)				(0.161)
EDU (a-1)	-3.815***		-3.549***		-2.737***		-3.724***		-3.769***		
	(0.245)		(0.175)		(0.152)		(0.246)		(0.220)		
Col (a-1)		-3.312***									
		(0.079)									
Univ (a-1)	-1.818***										
	(0.053)										
Constant	-9.795***	-11.384***	-8.408***	-11.443***	-9.481***	-11.298***	-9.395***	-15.603***	-7.766***	-12.118***	-9.435***
	(0.299)	(0.340)	(0.240)	(0.319)	(0.301)	(0.327)	(0.335)	(1.073)	(0.235)	(0.354)	(0.288)
Type 2	0.246***	0.189*	0.275***	0.239***	0.286***	0.179*	0.227***	0.199*	0.211***	0.162	0.214***
	(0.089)	(0.107)	(0.066)	(0.087)	(0.090)	(0.102)	(0.086)	(0.114)	(0.070)	(0.100)	(0.075)
Type 3	0.744***	0.776***	0.618***	0.600***	0.590***	0.722***	0.461***	0.613***	0.610***	0.515***	0.666***
	(0.134)	(0.158)	(0.106)	(0.139)	(0.149)	(0.157)	(0.143)	(0.180)	(0.109)	(0.159)	(0.117)

alternatives, ranging from -2.737 (EDU to SSOC) to -5.119 (SSOC to STEM), and statistically significant at the 1% level. Consistent with the descriptives, individuals face lower switching cost from a university program to a college program, relative to the opposite.

Based on the model, we can simulate educational trajectories for individuals, after controlling for unobserved heterogeneity and dynamic selection. Overall, Figure 4 in Appendix shows that, among those who start a college major, only a small fraction (lower than 50%) stays in the same major before earning enough years of coursework to graduate. As shown in Figure 5 in Appendix, only a fraction included between 20% and 40% earn enough credits by year 6 to obtain a BA degree in the same starting major.

Unobserved types At last, Table 8 includes the estimated parameters of unobserved heterogeneity types for each step of the model, with Type 1 serving as the reference category. The estimates show that the model successfully recovers meaningful unobserved types that are strongly linked to key educational and labor market outcomes.

For example, individuals of Type 2 and 3 are much more likely to graduate late. Despite this, Type 3 are also substantially more likely to experience skill mismatch in the labor market compared to Type 1 and 2. At last, Type 2 have substantial less work experience relative to Type 1.

Table 8: Unobserved heterogeneity types parameters

	Type 2	Type 3
Results higher tertiary	-0.049	-0.074
	(0.031)	(0.048)
Graduation (second sit)	5.715***	1.584***
	(0.302)	(0.387)
Skill mismatch	0.084	3.963***
	(0.070)	(0.169)
Potential experience (ages 23, 26, 29)	-0.098***	-0.057
	(0.021)	(0.037)
Unemployment (ages 23, 26, 29)	-0.023	-0.119
	(0.051)	(0.086)
Wage selection (ages 23, 26, 29)	-0.002	0.003
	(0.033)	(0.059)
Log-hourly wage matched (ages 23, 26, 29)	0.001	0.004
	(0.005)	(0.020)
Log-hourly wage mismatched (ages 23, 26, 29)	-0.007	0.004
	(0.008)	(0.009)

The model captures a persistent heterogeneity that connects college major choice, late graduation and skill mismatch, with observationally identical individuals performing very differently based on their estimated unobserved type. These patterns validate the models identification strategy: the latent types capture persistent traits related to both college major choices, where all type coefficients are significant (Table 7), and post-graduation outcomes such as mismatch and timing of graduation. This suggests the types reflect underlying differences in ability, preferences, or constraints that drive educational pathways and their associated returns.

4.2 Returns to college

Using our benchmark model, we can estimate individual i counterfactual wages and obtain returns Δ_{ijd} to field of study program j, at BA or MA level ($d \in \{BA, MA\}$):

$$\Delta_{ijd} = \log(\text{wage})_{ijd} - \log(\text{wage})_{iHS}, \tag{7}$$

where $\log(\text{wage})_{iHS}$ represents the counterfactual wage when an individual starts working immediately after high school without enrolling in college. For this reason, in this counterfactual, we force high school graduates at age 23 (26, 29) to have 5 (8, 11) years of potential experience. Moreover, we account for age and cohort fixed effects.

As a first step, we obtain average returns, rather than the full distribution, using $\Delta_{jd} = \frac{1}{n}\sum_i^n \Delta_{ijd}$. Table 9 includes Δ_{jd} for each college major, after controlling for dynamic selection and unobserved heterogeneity.

Table 9: Returns to college majors

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts	0.045***	0.098***	0.036*	0.096***
	(0.015)	(0.015)	(0.020)	(0.017)
Business and Law	0.035***	0.096***	0.152***	0.211***
	(0.011)	(0.014)	(0.018)	(0.017)
Social Sciences	0.017	0.078***	0.039**	0.091***
	(0.016)	(0.018)	(0.020)	(0.018)
Health	0.099***	0.160***	0.128***	0.190***
	(0.014)	(0.016)	(0.023)	(0.023)
STEM	0.079***	0.137***	0.110***	0.169***
	(0.012)	(0.013)	(0.021)	(0.018)
Education	0.012			
	(0.010)			

Notes: Log-hourly wage at age 23, 26, 29, with year and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

We find significant and important returns to most college major programs after controlling for observed characteristics, endogenous tertiary education choices, unobserved heterogeneity, and dynamic selection.

The highest paying Bachelor's (Master's) programs are university degrees in Business and Law, Health, and STEM. Respectively, these programs pay, on average, 16.1% (22.0%), 13.2% (19.4%), and 11.3% (17.3%) more than wages for high school graduates, considering earnings between ages 23 and 29. University programs in Humanities and Arts and Social Sciences yield substantially lower returns, with wages in Education programs not significantly different from high school wages. At the college level, a program in Social Sciences does not provide a statistically significant return. On average, college programs yield lower returns than university programs, though this difference is not statistically significant except for Business and Law programs.

In most cases, obtaining a degree in each college major already positively affects your wages, without considering the two risks arising from dropping out and not finding an adequately matched job.

Difference in returns to college majors The differences in returns to college majors may be driven by higher demand for specific jobs relative to others, especially those requiring Health or STEM skills, together with Business and law. This happens because of various factors, such as technical change, aging societies, and globalization (Acemoglu and Autor, 2011). The higher demand for these graduates drives up their wages and relative returns. However, large differences may also be attributed to the different supply of graduates in the labor market. STEM and Health majors are more challenging; therefore, those with these degrees might face less competition in the labor market and higher wages.

As shown in Table 9, there are important differences in returns across college majors. We can compute the college major premium by looking at the difference between college major j and college major k:

$$\begin{aligned} \Delta^{cp}_{jkd} = \frac{1}{I}\sum_{i}^{I}\Big(\log(wage)_{ijd} - \log(wage)_{ikd}\Big) \\ \text{for } j \in \{H, B, S, HE, ST, E\} \text{ and} \\ d \in \{BA, MA\}, \end{aligned} \tag{8}$$

We estimate the differences across returns to college majors and the relative college premium (Δ^{cp}_{jd}). Moreover, we document the distribution of individual returns to each college major j and, therefore, the difference within a college major j.

Difference across college majors In Appendix Table 23 we report college premia (Δ^{cp}_{jkd}) for k =Humanities and Arts (at the relative attainment level) with respect to each college major. STEM and Health degrees show consistently positive and significant premiums across all degree types. University degrees in Business and Law also show high premiums, while Social Sciences and Education generally have no significant differences with programs in Humanities and arts. At last, we find that Education programs pay significantly less than Humanities and arts programs.

Difference within college majors Appendix Table 23 shows significant differences across college majors. However, there are substantial differences within college majors, where some individuals could earn substantially less than others for various reasons, such as observed or unobserved characteristics. As shown in Carneiro et al. (2003), Aakvik et al. (2005) and Abbring and Heckman (2007), we are able to identify the joint distribution of treated and untreated potential outcomes and generate counterfactual distributions using factor models assumptions. The counterfactual distribution are key for estimating the heterogeneity in returns to college majors. In this case, some individuals would earn significantly higher returns than the average, while others could earn a negative return. In the latter case, these individuals would have benefited from the counterfactual while not enrolling in a BA degree and earning a wage right after the HS degree without considering non-pecuniary returns.

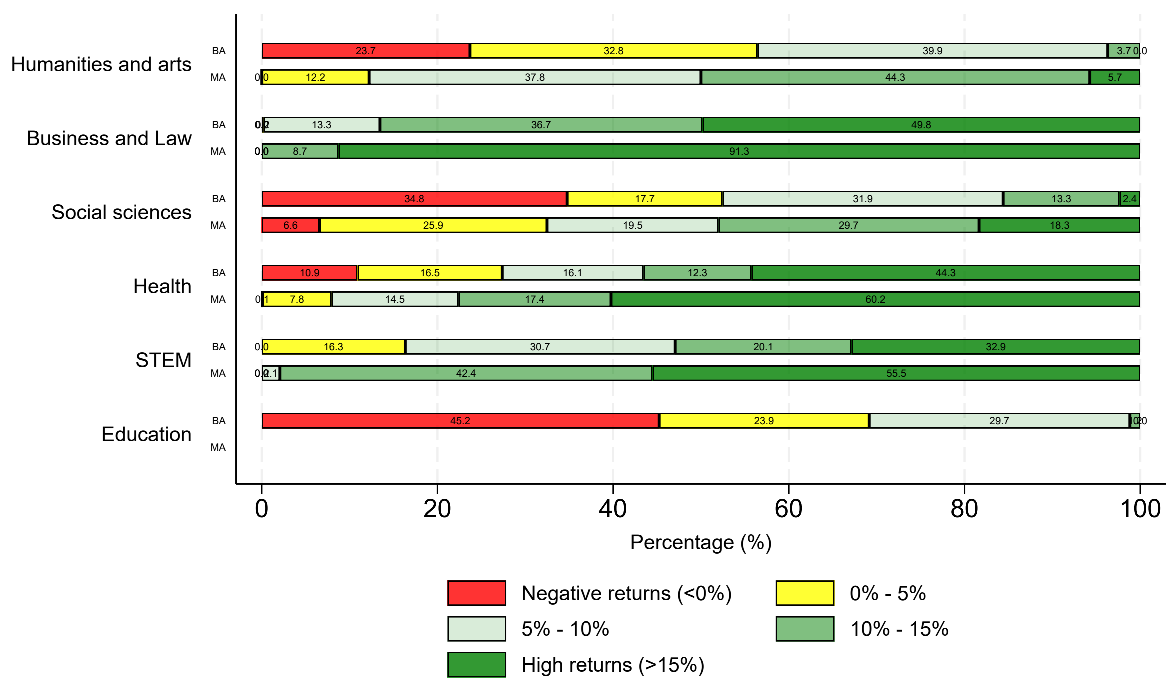


Figure 2: Heterogeneous returns to university programs (Δ_{ijd}, BA and MA, Univ)

Notes: BA and MA denote Bachelor's and Master's degrees. In Belgium, an MA Degree in Education does not exist. The figure includes Δ_{ijd}, which is the individual return to enrolling and obtaining a BA or an MA in college major j. Δ_{ijd} is computed as in Equation 7.

In Figure 2, we document the heterogeneity in these returns by showing the fraction of individuals earning: (a) negative returns (< 0%), (b) returns between 0 and 5%, (c) between 5 and 10%, (d) between 10 and 15% and, (e) high returns (>15%). We include the results on college programs in Appendix Figure 6.

Negative and low returns As documented in previous studies (Rodríguez et al., 2016), a substantial fraction of individuals experience negative returns to college education: they would have earned higher wages had they entered the labor market directly after high school.

Table 10: Negative returns by level of education and college majors (in %)

Level of Education	College Major	(1) Univ Average Mismatch (in %)	(2) Univ Without Mismatch (in %)	(3) Univ Full Mismatch (in %)	(4) Col Average Mismatch (in %)	(5) Col Without Mismatch (in %)	(6) Col Full Mismatch (in %)
BA Degree	Humanities and Arts	24.2	25.0	22.2	23.9	17.4	38.3
		(9.2)	(11.1)	(10.5)	(5.1)	(4.8)	(12.7)
	Business and Law	1.1	0.3	5.7	21.2	22.0	17.7
		(1.1)	(0.8)	(4.3)	(6.3)	(7.1)	(8.8)
	Social Sciences	25.7	17.2	43.1	32.1	31.3	34.9
		(7.3)	(9.1)	(10.8)	(6.7)	(7.7)	(8.4)
	Health	11.7	11.2	15.8	4.9	2.4	45.0
		(6.7)	(7.4)	(8.9)	(2.7)	(2.7)	(10.6)
	STEM	4.5	3.3	12.6	4.1	1.9	14.0
		(4.3)	(4.7)	(11.2)	(2.5)	(2.1)	(7.9)
	Education	33.9	33.5	40.2	33.9	33.5	40.2
		(4.4)	(4.6)	(11.2)	(4.4)	(4.6)	(11.2)
MA Degree	Humanities and Arts	6.8	5.8	9.3	10.8	5.6	19.1
		(5.1)	(7.1)	(5.7)	(4.3)	(3.0)	(9.9)
	Business and Law	0.3	0.0	1.6	2.3	1.4	5.2
		(0.5)	(0.1)	(2.1)	(2.3)	(2.4)	(3.8)
	Social Sciences	12.7	3.8	26.0	13.0	10.6	22.0
		(4.3)	(4.5)	(9.1)	(6.9)	(8.1)	(10.2)
	Health	4.2	3.1	10.0	2.6	0.1	33.1
		(3.7)	(3.9)	(7.1)	(1.1)	(0.3)	(11.1)
	STEM	0.7	0.1	3.5	0.8	0.0	3.3
		(0.8)	(0.3)	(4.4)	(0.8)	(0.0)	(3.2)

Notes: Each cell reports the simulated fraction (percentage) of individuals with negative returns in the distribution of individual returns by education and mismatch type. Standard errors in parentheses.

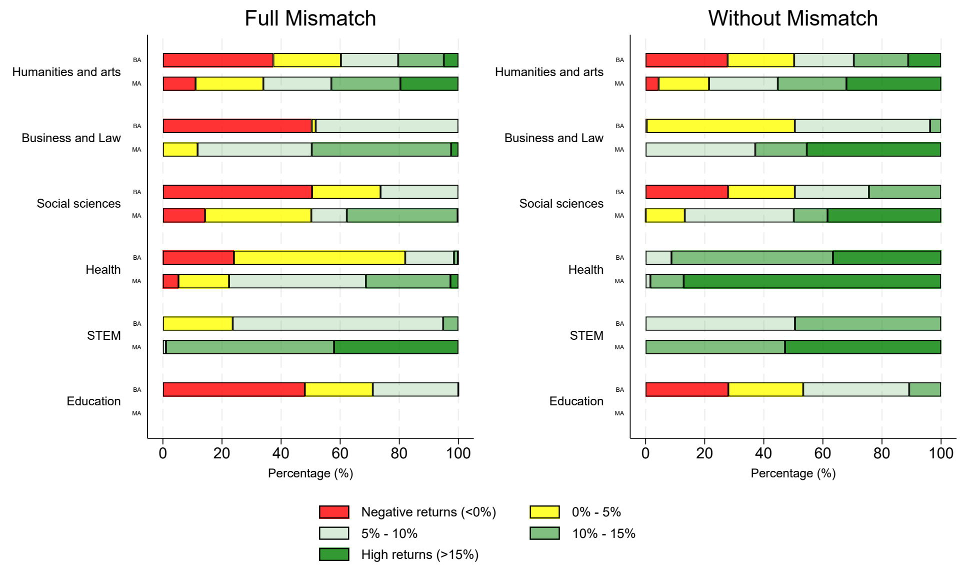
Table 10 reports the simulated share of individuals with negative returns, disaggregated by education level (BA vs. MA), institution type (university vs. college), and skill mismatch (average, none, or full mismatch). Average mismatch is defined as the average mismatch rates at the field of study level, including a weighted average between individuals who are mismatched with those who are not. Full mismatch and none are defined as the counterfactual simulations where each individual is either mismatched or not. The simulations are based on the full distribution of individual-level returns.

The results show that Social Sciences and Education stand out, especially among BA degree holders, where over 30% experience negative returns under average mismatch conditions. Under full mismatch, almost half of graduates experience a negative returns and could have earned a higher wage by not enrolling in college. Indeed, the percentage of individuals experience negative returns increases further, reaching 43.1% and 40.2% for university graduates in Social Sciences and Education, respectively, and 45.0% and 40.2% for college graduates. Humanities and Arts also show high shares of negative returns at the BA level (e.g., 24.2% under average mismatch), while STEM and Health majors consistently display the lowest incidence of negative returns, often below 10%, even under full mismatch.

For MA graduates, the incidence of negative returns is notably lower across most fields compared to BA holders. Social Sciences remain the most affected, with negative return rates ranging from 12.7% (average mismatch) to 26.1% (full mismatch) for university graduates. Humanities and Arts and Health also show moderate vulnerability under mismatch, while STEM and Business and Law consistently exhibit minimal negative return rates, often below 5%. Notably, Business and Law shows almost zero negative returns across all mismatch types, with some estimates as low as 0.3%. This suggests that at the master's level, educational investments are generally safer, particularly in applied and technical fields. However, mismatch still matters: full mismatch can substantially increase the fraction of graduates with negative returns, especially in vulnerable majors.

These findings indicate that field of study and job match matter substantially. The Full Mismatch columns show consistently higher rates of negative returns than the Without Mismatch columns, highlighting the economic importance of reducing education-employment mismatch. Furthermore, MA degrees are associated with lower negative return rates than BA degrees across nearly all fields, suggesting that promoting graduate education, especially in high-risk fields like Humanities and Social Sciences, could help reduce the incidence of low or negative returns.

Figure 3: The Role of Mismatch in Defining Negative (Low) Returns



Notes: BA and MA denote Bachelor's and Master's degrees. In Belgium, an MA Degree in Education does not exist. The figure includes Δ_{ijd}, which is the individual return to enrolling and obtaining a BA or an MA in college major j. Δ_{ijd} is computed as in Equation 7. This is by including full mismatch or without mismatch.

4.3 Skill mismatch

Skill mismatch is a key driver of low or negative returns to college. In essence, this suggests that, in a counterfactual scenario, an individual might have been better off entering the labor market directly after high school or choosing a different college major. This is intuitive, as individuals may end up employed in the same occupation regardless of whether they hold a high school diploma, a bachelor's degree (BA), or a master's degree (MA) in major j. Therefore, in Figure 3, we show the fraction of individuals earning heterogeneous returns with and without a full mismatch. As it is clear from Figure 3, in most college majors, if substantial fractions of individuals earn negative or low returns with a complete skill mismatch, this is not the case when looking at the case without mismatch.

Sorting into skill mismatch At first, this may be driven by different mismatch rates across college majors.

Table 11 reports the difference in skill mismatch rates between individuals with a tertiary education degree and those with only a high school degree, controlling for ex-

Table 11: Skill mismatch sorting

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts			-0.032	-0.038
			(0.035)	(0.043)
Business and Law	-0.089***	-0.103***	-0.110***	-0.132***
	(0.029)	(0.032)	(0.031)	(0.035)
Social Sciences	-0.110***	-0.131***	0.018	0.020
	(0.032)	(0.034)	(0.035)	(0.039)
Health	-0.195***	-0.239***	-0.141***	-0.171***
	(0.032)	(0.033)	(0.033)	(0.039)
STEM	-0.068**	-0.080**	-0.119***	-0.143***
	(0.028)	(0.032)	(0.030)	(0.036)
Education	-0.190***			
	(0.029)			

Notes: Skill mismatch rates difference with skill mismatch rates when holding a high-school degree. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

perience, unemployment rates, and dynamic selection. The estimates are presented separately by type and level of degree (BA or MA, and vocational (Col) or academic (Univ) institutions) across six fields of study. Negative coefficients indicate that tertiary education is associated with lower mismatch rates compared to graduates in Humanities and arts college degrees (at BA or MA level). The results show that college degrees (Col) are consistently associated with significantly lower mismatch rates across all fields, especially in Health (up to -24 percentage points), Education (-19 pp), and Social Sciences (-13 pp). Academic university degrees (Univ) also reduce mismatch in some fields, though to a lesser extent, particularly in Health (-14 to -17 pp) and STEM (-12 to -14 pp). In contrast, Humanities and Arts show no significant improvement, suggesting that these fields do not shield graduates from skill mismatch. These findings highlight substantial heterogeneity in how different types and levels of tertiary education influence the likelihood of being mismatched in the labor market.

Skill mismatch and returns to college majors Table 12 presents the estimated returns to different college majors and the penalty associated with skill mismatch in the labor market. The analysis is conducted separately for BA and MA degree holders across six major categories: Humanities and Arts, Business and Law, Social Sciences, Health, STEM, and Education. Columns (1) and (4) show returns without skill mismatch for university and college graduates respectively, while columns (2) and (5) present returns with full skill mismatch. The skill mismatch penalty, calculated as the difference between returns with and without mismatch, is reported in columns (3) and (6).

Table 12: Returns to college majors and skill mismatch

Level of Education	College Major	(1) Univ Without Mismatch Returns	(2) Univ Full Mismatch Returns	(3) Univ Full Mismatch Penalty	(4) Col Without Mismatch Returns	(5) Col Full Mismatch Returns	(6) Col Full Mismatch Penalty
BA Degree	Humanities and Arts	0.035	0.039	0.005	0.069***	-0.004	-0.073**
		(0.024)	(0.030)	(0.037)	(0.018)	(0.026)	(0.032)
	Business and Law	0.160***	0.115***	-0.045	0.036***	0.035*	-0.000
		(0.021)	(0.027)	(0.032)	(0.011)	(0.019)	(0.019)
	Social Sciences	0.068***	-0.015	-0.083**	0.017	0.012	-0.005
		(0.025)	(0.030)	(0.036)	(0.017)	(0.029)	(0.031)
	Health	0.127***	0.129**	0.002	0.108***	-0.033	-0.141***
		(0.027)	(0.055)	(0.064)	(0.014)	(0.040)	(0.042)
	STEM	0.118***	0.069*	-0.049	0.086***	0.053**	-0.032
		(0.024)	(0.036)	(0.043)	(0.012)	(0.022)	(0.022)
	Education	0.013	-0.007	-0.020	0.013	-0.007	-0.020
		(0.011)	(0.029)	(0.031)	(0.011)	(0.029)	(0.031)
MA Degree	Humanities and Arts	0.100***	0.089***	-0.011	0.134***	0.045*	-0.089***
		(0.023)	(0.024)	(0.032)	(0.018)	(0.025)	(0.032)
	Business and Law	0.225***	0.164***	-0.061**	0.101***	0.085***	-0.016
		(0.020)	(0.024)	(0.028)	(0.015)	(0.022)	(0.024)
	Social Sciences	0.132***	0.034	-0.099***	0.082***	0.061*	-0.021
		(0.024)	(0.025)	(0.033)	(0.020)	(0.034)	(0.037)
	Health	0.192***	0.178***	-0.013	0.172***	0.016	-0.156***
		(0.026)	(0.051)	(0.058)	(0.018)	(0.041)	(0.045)
	STEM	0.182***	0.118***	-0.064*	0.151***	0.103***	-0.048**
		(0.021)	(0.029)	(0.036)	(0.015)	(0.021)	(0.023)

Notes: Log-hourly wage at age 23, 26, 29, with time and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. Standard errors in parentheses.

Table 12 shows that, when fully mismatched, on average individuals do not benefit from a BA degree, except for Business and law, Health (Univ), and STEM. This is associated to high skill mismatch penalty, which substantially reduces the average returns to college.

The estimates reveal several important patterns regarding returns to education and skill mismatch effects. First, there are substantial wage premiums associated with most college majors, with Business and Law, Health, and STEM fields generally yielding the highest returns across both BA and MA levels. Second, skill mismatches significantly reduce these wage premiums, with some notable variation across fields. For example, Health majors face particularly severe mismatch penalties (up to 15.4% for MA graduates), while Business and Law graduates experience relatively smaller penalties. Third, the magnitude of mismatch penalties tends to be larger for college graduates (column 6) compared to university graduates (column 3), suggesting that university education may provide more transferable skills. Finally, advanced degrees (MA) generally offer higher returns than bachelor's degrees (BA) in matched employment, but this advantage diminishes or disappears under conditions of skill mismatch. These findings highlight the economic importance of both field of study and job-skill alignment in determining labor market outcomes for college graduates.

Table 12 presents the estimated returns to college majors and the role of skill mismatch, using log-hourly wages at ages 23, 26, and 29 as the outcome. The table distinguishes between Bachelors (BA) and Masters (MA) degree holders and compares university and college pathways. Columns (1) and (4) report returns without accounting for mismatch, while columns (2) and (5) include full returns accounting for mismatch. Columns (3) and (6) isolate the mismatch penalty. Across both education levels, majors in Business and Law, Health, and STEM tend to yield the highest wage returns, particularly in university settings. However, incorporating mismatch reveals significant penalties for certain majors, notably in Health and Humanities and Arts at the MA level, indicating that skill mismatch can substantially reduce returns. Returns are consistently higher for MA graduates compared to BA graduates. The table also highlights the importance of accounting for mismatch and selection dynamics when assessing the value of different college majors.

4.4 Heterogeneity analysis

Unobserved heterogeneity At first, we present results by estimating heterogeneous returns by unobserved type. Unobserved types generate different returns to college, different sorting patterns into skill mismatch occupations and, therefore, different skill mismatch penalty. Table 24 in Appendix includes the estimated returns for each unobserved type. Type 1 and 2 have similar returns to most college major programs. They do no benefit only from programs in Education and university programs in Humanities and Arts. On the other side, individuals in Type 3 do not benefit as much from college, especially they get not significant results in college programs in Humanities and arts, Social Sciences, and Education and university programs in Social Sciences. Moreover, they receive lower returns to university programs in STEM (both BA and MA degrees) and Business and law.

Table 25 in Appendix includes the results relative to sorting into skill mismatch. Even if they are the ones benefiting the less from college, Type 3 individuals demonstrate the best outcomes in terms of mismatch, with the most negative coefficients indicating the lowest mismatch rates, particularly excelling in Health (-0.255 to -0.201) and Education (-0.222). These are individuals with lower returns to college, but with better match quality. On the other side, Type 1 individuals show intermediate performance, with reductions in mismatch rates in Health (-0.174 to -0.224) and Education (-0.170), but not as pronounced as Type 3. Type 2 individuals face the highest risk of skill mismatch, showing the smallest negative coefficients (closest to zero) across most fields, with Health (-0.207 to -0.257) and Education (-0.202) still providing some protection but less than the other types. Notably, all types benefit from avoiding Humanities and Arts majors (the reference category), but Type 2 individuals see the smallest improvements when switching to more specialized fields.

Table 26 in Appendix shows the returns to college majors and the skill mismatch penalty for each unobserved type. Regarding the BA degree, Type 3 generally experiences higher returns in the first job, regardless of the first occupation's matching quality, relative to Types 1 and 2. Moreover, they experience a significant mismatch penalty only relative to college programs in Health (-14.3). Type 3 individuals are those with a higher risk of entering the labor market in a mismatched occupation. Type 1 experiences penalties in college programs in Humanities and arts (-7.9) and in Health (-14.4). In contrast, Type 2 experiences significant penalties in college programs in STEM (-4.3) Health (-15.3), Humanities and arts (-8.8), and university programs in Business and Law (-6.0) and Social Sciences (-8.0). Relative to the BA degree, Type 2 individuals are at the highest risk of a mismatch penalty. Regarding the MA degree, we find significant mismatch penalties for Types 1 and 2 in college programs in Humanities and Arts (-9.5 and -10.3), Health (-16.7 and -17.4), STEM (-4.7 and -5.5), and in university programs in Business and Law (-6.0 and -7.0), Social Sciences (-8.8 and -9.6), and STEM (-8.6 and -9.3). For Type 3 individuals, we find significant penalties for a smaller number of MA degree programs. Moreover, the size of penalties is generally lower.

4.5 Robustness checks

We perform different robustness checks on our model's skill mismatch measures and the inclusion of unobserved heterogeneity. As a first analysis, we focus on the measures of skill mismatch. We do so by re-estimating the model using each skill mismatch alternative measure. Then, we reproduce the following tables using the parameters from each model for returns to college majors (Table 9), low and negative returns (Table 10), skill mismatch sorting (Table 11) and skill mismatch penalty (Table 12).

At first, on average, returns to college majors from our benchmark model (Table 9) are robust to use an alternative measure of skill mismatch, as in done in Appendix Table 27 and Appendix Table 28. Indeed, average returns to college major shows that there are no significant returns to programs in Social Sciences and Education, regardless of the skill mismatch measure used. Overall, the highest returns are university programs in Business and law (MA degree) in all models. Our estimates of returns to college majors are not sensitive to the skill mismatch measure used in the model.

Low and negative returns The results from our benchmark model (Table 10) is robust when compared to the alternative skill mismatch measures in Table 33 and Table 34 in the Appendix. The patterns remain consistent across all specifications: Education majors consistently show the highest rates of negative returns (around 33-40% for BA degrees), followed by Social Sciences and Humanities and Arts, while STEM fields exhibit the lowest negative return rates (typically under 5% for BA degrees). Business and Law consistently performs well with very low share of negative returns across all measures. The relative ranking of majors by negative return vulnerability remains virtually unchanged, and the general pattern that university graduates experience higher negative return rates than college graduates persists across all mismatch definitions. Despite some magnitude differences, the fundamental economic story about which fields of study carry greater risk of negative returns remains remarkably stable across different methodological approaches.

Skill mismatch sorting In Appendix, Table 31 and Table 32, shows the main findings relative to skill mismatch rates when using a different skill mismatch measure. First, Health and Education shows the lowest mismatch rates across JA, DSA and our benchmark model (Table 11). In all three models, STEM has generally lower mismatch rates relative to Humanities and arts, but in line with Business and Law and Social Sciences.

At last, university programs in Humanities and arts and Social Sciences show no significant differences from college programs in Humanities and arts (reference category). When using the JA model, university programs in Social Sciences have larger skill mismatch rates. Despite these magnitude differences, the overall pattern of negative skill mismatch sorting remains consistent across all three specifications, reinforcing the robustness of our core results.

Skill mismatch penalty In the Appendix, we include Table 29 and Table 30, which present new estimates from the model using the JA and DSA measures, without relying on our latent factor approach.

Regarding skill mismatch penalties, we find significant effects for university programs in Social Sciences (at both the BA and MA levels), which are robust and consistent across models using different skill mismatch measures. We also observe robust and consistent penalties for Health programs when using our benchmark model and the DSA model. However, when using skill mismatch measures based on JA or DSA alone, fewer programs exhibit statistically significant penalties. Our latent factor approach addresses measurement issues and helps to uncover these penalties more clearly.

The Job Analysis (JA) measure generally yields higher returns to education in the absence of mismatch. For example, Business and Law MA degrees show returns of 0.207 using the JA measure, compared to 0.234 in the benchmark. JA also produces more pronounced mismatch penalties, particularly for Social Sciences BA degrees (-0.121 vs. -0.093). In contrast, the Developed Skills Assessment (DSA) measure produces more conservative estimates: for instance, Health MA degrees yield returns of 0.179 without mismatch versus 0.198 in the benchmark, and generally show smaller penalty coefficients. Notably, the DSA measure indicates less severe penalties for college graduates, with Health BA penalties of -0.152 compared to -0.146 in the benchmark.

Despite differences in magnitude, all three measures consistently indicate positive returns to higher education in the absence of mismatch, substantial penalties for skill mismatch (particularly in Health and Social Sciences fields), and systematically better outcomes for university graduates relative to college graduates. This consistency across different skill measurement approaches reinforces the robustness of our core findings regarding the heterogeneous effects of educational mismatch.

5 Conclusions

This paper estimates heterogeneous causal returns to college majors and investigates the importance of skill mismatch risk as underlying mechanisms generating heterogeneous returns to college. We use a dynamic model of joint educational choices and labor market outcomes, accounting for dynamic selection and unobserved heterogeneity (Ashworth et al., 2021; Heckman and Navarro, 2007; Heckman et al., 2018a,b, 2016). We identify the latter using the panel nature of the data, initial conditions, local labor market conditions, and exclusion restrictions, including relative distance to higher education and graduation timing (cf. Humphries et al., 2023). Moreover, we control for several observed characteristics and endogenous academic achievements, including secondary education outcomes, year retention, and grades in secondary and tertiary education. We estimate the impact of educational choices and outcomes on skill mismatch probability and returns to college majors. We include skill mismatch in the first job and estimate its associated wage penalties at later ages, up until age 29. Moreover, besides investigating the role of skill mismatch, we also account for other indirect channels that may contribute to heterogeneous wage returns between college majors, such as differences in the probability of graduating or obtaining a higher grade.

This paper contributes to the literature in three main ways. First, we contribute to the literature on the average returns to the college major choice by considering a broader set of majors and, in particular, differentiating between health and (other) STEM fields. We estimate direct and indirect returns to college majors, controlling for unobserved heterogeneity, a large set of observed characteristics, and endogenous educational choices. Moreover, we estimate risk components associated with skill mismatch. Second, we contribute to the growing literature on heterogeneous returns to college in general and college majors in particular. Using a dynamic model, we can assess individual returns to a college degree and show how college major choice contributes to negative wage returns to college for a fraction of the population. Third, our paper contributes to the literature on skill mismatch and college major choice. We account for vertical and horizontal mismatches, including multiple measures of mismatches, allowing us to accommodate measurement error problems.

Our results suggest that returns to college majors are substantial and positive after controlling for observed and unobserved characteristics, endogenous educational choices, and other factors, in line with other studies (Kirkebøen et al., 2016; Beffy et al., 2012; Arcidiacono, 2004; Altonji et al., 2016b). This is different from the OLS estimates provided in our preliminary result section. However, when considering the distribution of individual returns, we show that a substantial fraction of individuals earn a negative wage return: this is almost 33.6% of individuals enrolling and obtaining a BA degree in social sciences. When removing skill mismatch, these fractions reduce substantially, suggesting that part of the issue is the probability of individuals ending up in adequately matched jobs.

These results have key policy implications. Indeed, a college degree always has a positive average return, so individuals should be encouraged to pursue higher education. However, they also suggest that individuals may incur negative returns to college even in higher-paying college majors, such as Business and law. Therefore, re-directing individuals to these college majors may not always be the correct policy answer. Moreover, skill mismatch risk are not as important in canceling the associated positive expected returns of most degrees. Therefore, most degrees have still a strong expected return, which make them a valuable investment. The only degree with a zero expected return is an MA degree in social sciences, where the skill mismatch risk component is so high that it reduces the expected returns substantially. Doing this analysis could help in understanding which are the true causal expected returns and which could be potentially the best college majors to enroll.

References

Aakvik, Arild, James J. Heckman, and Edward J. Vytlacil (2005) "Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs," *Journal of Econometrics*, 125 (1), 15–51, https://doi.org/10.1016/j.jeconom.2004.04.002.

Abbring, Jaap H. and James J. Heckman (2007) "Chapter 72 Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation," in Heckman, James J. and Edward E. Leamer eds. *Handbook of Econometrics*, 6 of Handbook of Econometrics, 5145–5303: Elsevier.

Acemoglu, Daron and David Autor (2011) "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Ashenfelter, O. and D. Card eds. *Handbook of Labor Economics*, 4 of Handbook of Labor Economics, Chap. 12, 1043–1171: Elsevier.

Adda, Jérôme and Christian Dustmann (2023) "Sources of Wage Growth," *Journal of Political Economy*, 131 (2), 456–503, 10.1086/721657.

Altonji, J.G., P. Arcidiacono, and A. Maurel (2016a) "Chapter 7 - The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," 5 of Handbook of the Economics of Education, 305–396: Elsevier, https://doi.org/10.1016/B978-0-444-63459-7.00007-5.

Altonji, Joseph G., Erica Blom, and Costas Meghir (2012) "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 4 (1), 185–223, 10.1146/annurev-economics-080511-110908.

Altonji, Joseph, Lisa Kahn, and Jamin Speer (2014) "Trends in Earnings Differentials across College Majors and the Changing Task Composition of Jobs," *American Economic Review*, 104 (5), 387–93.

——— (2016b) "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success," *Journal of Labor Economics*, 34 (S1), S361 – S401.

Andrews, Rodney J, Scott A Imberman, Michael F Lovenheim, and Kevin M Stange (2022) "The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability," 10.3386/W30331.

Andrews, Rodney J., Jing Li, and Michael F. Lovenheim (2016) "Quantile Treatment Effects of College Quality on Earnings," *Journal of Human Resources*, 51 (1), 200–238, 10.3368/jhr.51.1.200.

Arcidiacono, Peter (2004) "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 121, 343–375, https://EconPapers.repec.org/RePEc:eee:econom:v:121:y:2004:i:1-2:p:343-375.

Arcidiacono, Peter, V. Joseph Hotz, Arnaud Maurel, and Teresa Romano (2020) "Ex ante returns and occupational choice," *Journal of Political Economy*, 128, 4475–4522, 10.1086/710559.

Arcidiacono, Peter and John Bailey Jones (2003) "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71, 933–946, 10.1111/1468-0262.00431.

Ashworth, Jared, V. Joseph Hotz, and Arnaud Maurel (2021) "Changes Across Cohorts in Wage Returns to Schooling and Early Work Experiences," *Journal of Labor Economics*, 39, 931–964, 10.1086/711851/SUPPL_FILE/18195DATA.ZIP.

Baert, Stijn, Bart Cockx, and Dieter Verhaest (2013) "Overeducation at the Start of the Career: Stepping Stone or Trap?," *Labour Economics*, 25, 123–140, 10.1016/J.LABECO.2013.04.013.

Beffy, Magali, Denis Fougere, and Arnaud Maurel (2012) "Choosing the Field of Study in Post-secondary Education: Do Expected Earnings Matter?" *The Review of Economics and Statistics*, 94 (1), 334–347.

Bleemer, Zachary and Aashish Mehta (2022) "Will Studying Economics Make You Rich? A Regression Discontinuity Analysis of the Returns to College Major," *American Economic Journal: Applied Economics*, 14 (2), 1–22, 10.1257/app.20200447.

Blom, Erica, Brian C. Cadena, and Benjamin J. Keys (2021) "Investment over the Business Cycle: Insights from College Major Choice," *Journal of Labor Economics*, 39 (4), 1043–1082, 10.1086/712611.

Bruneel-Zupanc, Christophe (2023) "Don't (Fully) Exclude Me, it's Not Necessary! Identification with Semi-IVs," *Working Paper*.

Bruneel-Zupanc, Christophe and Jad Beyhum (2024) "Identification with Possibly Invalid IVs," *Working Paper*.

Cameron, Stephen V. and James J. Heckman (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106, 262–333, 10.1086/250010.

——— (2001) "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 109, 455–499, 10.1086/321014.

Cappelli, Peter H. (2015) "SKILL GAPS, SKILL SHORTAGES, AND SKILL MISMATCHES: EVIDENCE AND ARGUMENTS FOR THE UNITED STATES," *ILR Review*, 68 (2), 251–290.

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman (2003) "2001 Lawrence R. Klein Lecture: Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice," *International Economic Review*, 44 (2), 361–422, http://www.jstor.org/stable/3663472.

Chevalier, Arnaud (2017) "To Be or Not to Be a Scientist?," in *Skill Mismatch in Labor Markets*, 45, 1–39: Emerald Group Publishing Limited.

De Groote, Olivier and Koen Declercq (2021) "Tracking and specialization of high schools: Heterogeneous effects of school choice," *Journal of Applied Econometrics*, 36 (7), 898–916, https://doi.org/10.1002/jae.2856.

Declercq, Koen and Frank Verboven (2018) "Enrollment and degree completion in higher education without admission standards," *Economics of Education Review*, 66 (C), 223–244, https://EconPapers.repec.org/RePEc:eee:ecoedu:v:66:y:2018:i:c:p:223-244.

Frenette, Marc (2004) "Access to College and University: Does Distance to School Matter?" *Canadian Public Policy*, 30, 427–443, https://EconPapers.repec.org/RePEc:cpp:issued:v:30:y:2004:i:4:p:427-443.

Ghignoni, Emanuela and Alina Verashchagina (2014) "Educational qualifications mismatch in europe. is it demand or supply driven?" *Journal of Comparative Economics*, 42, 670–692, 10.1016/j.jce.2013.06.006.

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman (2013) "Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile," Working Paper 19241, National Bureau of Economic Research, 10.3386/w19241.

Heckman, J and B Singer (1984) "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," 52, 271–320.

Heckman, James J., John Eric Humphries, and Gregory Veramendi (2016) "Dynamic Treatment Effects," *Journal of Econometrics*, 191 (2), 276–292, https://doi.org/10.1016/j.jeconom.2015.12.001.

——— (2018a) “Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking,” *The Journal of Political Economy*, 126, S197, 10.1086/698760.

——— (2018b) “The Nonmarket Benefits of Education and Ability,” *Journal of Human Capital*, 12, 282, 10.1086/697535.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006) “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behaviour,” *Journal of Labor Economics*, 24, http://jenni.uchicago.edu/noncog/.

Heckman, James and Salvador Navarro (2007) “Dynamic Discrete Choice and Dynamic Treatment Effects,” *Journal of Econometrics*, 136, 341–396, https://EconPapers.repec.org/RePEc:eee:econom:v:136:y:2007:i:2:p:341-396.

Hotz, V. Joseph, Lixin Xu, Marta Tienda, and Avner Ahituv (2002) “Are There Returns To The Wages Of Young Men From Working While In School?,” *The Review of Economics and Statistics*, 84, 221–236, 10.1162/003465302317411497.

Humphries, John Eric, Juanna Schrøter Joensen, and Gregory F Veramendi (2023) “Complementarities in High School and College Investments,” *Working Paper*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4560152.

Jepsen, Christopher, Kenneth Troske, and Paul Coomes (2014) “The Labor-Market Returns to Community College Degrees, Diplomas, and Certificates,” *Journal of Labor Economics*, 32 (1), 95–121.

Ketel, Nadine, Edwin Leuven, Hessel Oosterbeek, and Bas van der Klaauw (2016) “The Returns to Medical School: Evidence from Admission Lotteries,” *American Economic Journal: Applied Economics*, 8 (2), 225–54.

Kinsler, Josh and Ronni Pavan (2015) “The specificity of general human capital: Evidence from college major choice,” *Journal of Labor Economics*, 33, 933–972, 10.1086/681206.

Kirkebøen, Lars, Edwin Leuven, Magne Mogstad, Lars Kirkebøen, Edwin Leuven, and Magne Mogstad (2016) “Editor’s Choice Field of Study, Earnings, and Self-Selection,” *The Quarterly Journal of Economics*, 131, 1057–1111, 10.1093/QJE/QJW019.

Leighton, Margaret and Jamin D Speer (2020) “Labor market returns to college major specificity,” *European Economic Review*, 128, 103489, 10.1016/j.euroecorev.2020.103489.

Lemieux, Thomas (2014) “Occupations, fields of study and returns to education,” *Canadian Journal of Economics/Revue canadienne d’économique*, 47, 1047–1077, 10.1111/CAJE.12116.

Leuven, Edwin and Hessel Oosterbeek (2011) "Overeducation and Mismatch in the Labor Market," IZA Discussion Papers 5523, Institute of Labor Economics (IZA), https://EconPapers.repec.org/RePEc:iza:izadps:dp5523.

Leuven, Edwin, Hessel Oosterbeek, and Inge de Wolf (2013) "The Effects Of Medical School On Health Outcomes: Evidence From Admission Lotteries," *Journal of Health Economics*, 32 (4), 698–707.

Lindley, Joanne and Steven McIntosh (2015) "Growth In Within Graduate Wage Inequality: The Role Of Subjects, Cognitive Skill Dispersion And Occupational Concentration," *Labour Economics*, 37 (C), 101–111.

Loury, Linda, David Garman, Linda Loury, and David Garman (1995) "College Selectivity and Earnings," *Journal of Labor Economics*, 13, 289–308, 10.1086/298375.

Navarini, Lorenzo and Dieter Verhaest (2023) "Educational Attainment, Overeducation, and Wages: Evidence from a Dynamic Model," *Working Paper*.

Neyt, Brecht, Dieter Verhaest, Lorenzo Navarini, and Stijn Baert (2022) "The Impact of Internship Experience on Schooling and Labour Market Outcomes," *CESifo Economic Studies*, 68, 127–154, 10.1093/CESIFO/IFAC001.

Nordin, Martin, Inga Persson, and Dan-Olof Rooth (2010) "Education-occupation mismatch: Is there an income penalty?" *Economics of Education Review*, 29 (6), 1047–1059.

Oosterbeek, Hessel and Dinand Webbink (1997) "Is there a hidden technical potential?" *Economist*, 145, 159–177, 10.1023/A:1002938120342/METRICS.

Oreopoulos, Philip and Kjell G. Salvanes (2011) "Priceless: The Nonpecuniary Benefits of Schooling," *Journal of Economic Perspectives*, 25, 159–84, 10.1257/JEP.25.1.159.

Robst, John (2007) "Education and job match: The relatedness of college major and work," *Economics of Education Review*, 26 (4), 397–407, https://doi.org/10.1016/j.econedurev.2006.08.003.

Rodríguez, Jorge, Sergio Urzúa, and Loreto Reyes (2016) "Heterogeneous Economic Returns to Post-Secondary Degrees: Evidence from Chile," *Journal of Human Resources*, 51, 416–460, 10.3368/JHR.51.2.0213-5474R1.

Sellami, Sana, Dieter Verhaest, and Walter Van Trier (2018) "How to Measure Field-of-Study Mismatch? A Comparative Analysis of the Different Methods," *LABOUR*, 32 (4), 141–173, https://doi.org/10.1111/labr.12129.

Somers, Melline A., Sofie J. Cabus, Wim Groot, and Henriëtte Maassen van den Brink (2019) "Horizontal Mismatch Between Employment And Field Of Education: Evidence From A Systematic Literature Review," *Journal of Economic Surveys*, 33 (2), 567–603.

Verhaest, Dieter, Sana Sellami, and Rolf Van der Velden (2017) "Differences in horizontal and vertical mismatches across countries and fields of study," *International Labour Review*, 156 (1), 1–23, https://doi.org/10.1111/j.1564-913X.2015.00031.x.

Webber, Douglas A. (2014) "The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors," *Labour Economics*, 28, 14–23, 10.1016/J.LABECO.2014.03.009.

——— (2016) "Are college costs worth it? How ability, major, and debt affect the returns to schooling," *Economics of Education Review*, 53, 296–310, 10.1016/J.ECONEDUREV.2016.04.007.

A Data

A.1 higher education in Belgium

A.2 Preliminary evidence

A.3 Skill Mismatch, College Majors and Occupations

Our mismatch measures are based on educational variables, such as educational attainment and college majors. However, skill mismatches should not necessarily coincide with educational mismatches, with the first being defined as working in a job where the required skills do not match the acquired skills (Sellami et al., 2018).

Using an occupation-based measure, we check the correlation between skill mismatch and educational mismatch. Similarly to Blom et al. (2021), we characterize occupations based on the concentration of different college majors (a Herndahl-Hirschman index). The higher the concentration, the more specialized an occupation is: for instance, to be a doctor, you need a Health degree, and it is highly probable that people around you have a Health degree. This is not true for low-skilled occupations, such as waiters, or non-specialized occupations with lower entry barriers, such as salespersons or management analysts. Of course, it is more likely to find higher skill (and educational) mismatch rates for non-specialized occupations. And, while it is intuitive that

Table 13: Descriptive statistics: individuals who fail the first year and Bachelor's degree

1st year program	Drop	Switch	Stay	Total
HUMA(Col)	18.0	6.3	75.7	100.0
HUMA(Univ)	4.4	1.5	94.1	100.0
BULA(Col)	9.0	2.2	88.8	100.0
BULA(Univ)	9.6	2.8	87.6	100.0
SSOC(Col)	11.8	5.6	82.6	100.0
SSOC(Univ)	9.6	3.8	86.6	100.0
HEA(Col)	7.6	5.0	87.4	100.0
HEA(Univ)	19.4	2.9	77.7	100.0
STEM(Col)	7.0	2.0	91.0	100.0
STEM(Univ)	5.5	1.0	93.5	100.0
EDU	9.1	2.4	88.4	100.0
Total	9.5	3.1	87.4	100.0

Notes: The college majors are the following: Humanities and arts (HUMA), Business and Law (BULA), Social sciences (SSOC), Health and biomedical sciences (HEA), Science, technology, engineering and mathematics (STEM), Education (EDU). (Univ) stands for academic degree in university, (Col) stands for non-academic track at vocationally oriented collages. Columns include the information on: (i) Drop - if drop out and never attained a Bachelor's degree, (ii) Switch - attained a Bachelor's degree but in a different major, (iii) Stay - attained a Bachelor's degree in the same major.

Table 14: Descriptive statistics: reorientation and completion in higher education

Bachelor's Degree	HUMA(Col)	HUMA(Univ)	BULA(Col)	BULA(Univ)	SSOC(Col)	SSOC(Univ)	HEA(Col)	HEA(Univ)	STEM(Col)	STEM(Univ)	EDU	Total
No college	31.8	19.7	35.2	20.7	35.0	18.6	26.0	21.0	33.1	15.0	33.5	28.1
HUMA(Col)	48.1	2.1	0.0	0.6	0.3	0.6	0.2	0.8	0.6	1.2	0.3	4.2
HUMA(Univ)	0.0	60.2	0.0	0.2	1.0	0.0	0.2	0.0	0.0	0.5	0.3	3.4
BULA(Col)	10.4	6.3	58.7	15.6	2.0	9.0	2.5	4.4	1.9	2.4	1.1	14.2
BULA(Univ)	0.2	0.4	0.2	55.1	0.3	1.0	0.0	0.8	0.1	1.0	0.0	5.6
SSOC(Col)	1.7	1.4	0.7	2.7	53.3	7.7	1.0	1.6	0.2	0.0	3.6	4.7
SSOC(Univ)	0.2	0.7	0.6	0.8	1.6	53.7	0.0	0.8	0.1	1.2	0.2	3.6
HEA(Col)	1.0	1.1	0.7	0.8	1.6	1.9	63.5	8.5	1.1	0.7	1.6	7.0
HEA(Univ)	0.0	0.4	0.0	0.4	0.3	0.3	2.1	52.8	0.0	0.2	0.0	2.8
STEM(Col)	2.2	0.4	1.8	2.0	0.3	1.0	1.0	4.0	60.3	14.0	1.7	11.7
STEM(Univ)	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.5	60.4	0.0	4.8
EDU	4.1	7.4	2.2	1.2	4.2	6.1	3.4	4.4	2.0	3.4	57.7	9.8
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: The college majors are the following: Humanities and arts (HUMA), Business and Law (BULA), Social sciences (SSOC), Health and biomedical sciences (HEA), Science, technology, engineering and mathematics (STEM), Education (EDU). (Univ) stands for academic degree in university, (Col) stands for non-academic track at vocationally oriented collages. Columns include the information on: (i) Drop - if drop out and never attained a Bachelor's degree, (ii) Switch - attained a Bachelor's degree but in a different major, (iii) Stay - attained a Bachelor's degree in the same major.

specialized workers can be damagingly mismatched, it is less clear that a similar correspondence can be defined for generally skilled workers (Leighton and Speer, 2020).

Table 15: Preliminary evidence: labor market outcomes and tertiary education programs

	potential_experience_	unemployment_	Endo_SM_factor	Log_Endo_Hwage	Log_Endo_Hwage	Log_Endo_Hwage
Humanities and Arts (Col)	-1.727***	0.009	-0.020	0.041**	0.025	0.041**
	(0.080)	(0.027)	(0.048)	(0.015)	(0.014)	(0.015)
Humanities and Arts (Univ)	-1.461***	-0.050	-0.176**	0.080***	0.066***	0.078***
	(0.092)	(0.031)	(0.057)	(0.017)	(0.017)	(0.017)
Business and Law (Col)	-1.110***	-0.088***	-0.181***	0.056***	0.044***	0.051***
	(0.045)	(0.015)	(0.025)	(0.008)	(0.008)	(0.008)
Business and Law (Univ)	-1.524***	-0.077**	-0.224***	0.159***	0.145***	0.154***
	(0.080)	(0.028)	(0.050)	(0.015)	(0.015)	(0.015)
Social sciences (Col)	-1.259***	-0.084***	-0.288***	0.052***	0.038***	0.043***
	(0.069)	(0.022)	(0.039)	(0.012)	(0.011)	(0.012)
Social sciences (Univ)	-1.842***	-0.001	-0.066	0.076***	0.057***	0.076***
	(0.088)	(0.030)	(0.057)	(0.017)	(0.017)	(0.017)
Health (Col)	-0.956***	-0.172***	-0.360***	0.116***	0.104***	0.105***
	(0.058)	(0.019)	(0.033)	(0.010)	(0.009)	(0.010)
Health (Univ)	-1.850***	-0.211***	-0.259***	0.180***	0.161***	0.171***
	(0.097)	(0.036)	(0.074)	(0.019)	(0.018)	(0.019)
STEM (Col)	-1.448***	-0.089***	-0.202***	0.099***	0.086***	0.094***
	(0.050)	(0.017)	(0.028)	(0.009)	(0.008)	(0.009)
STEM (Univ)	-1.580***	-0.037	-0.298***	0.148***	0.133***	0.140***
	(0.084)	(0.029)	(0.053)	(0.016)	(0.016)	(0.016)
Education	-0.787***	-0.038*	-0.350***	0.044***	0.033***	0.034***
	(0.050)	(0.016)	(0.026)	(0.008)	(0.008)	(0.008)
Endo_MA_=1	-0.196***	0.067***	0.028	0.054***	0.048***	0.055***
	(0.050)	(0.017)	(0.032)	(0.009)	(0.009)	(0.009)
potential_experience_		-0.064***		0.008***		0.009***
		(0.003)		(0.001)		(0.001)
Endo_SM_factor						-0.031***
						(0.004)
Exogenous variables	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24441	12842	5879	10901	10901	10901

Notes:

Moreover, by construction, if peers in your occupation have different educational histories or skills, you might find yourself not adequately matched. On the other side, if individuals around you are all similar, it would be hard to be mismatched because of entry barriers.

To check this, we follow the framework of Blom et al. (2021) and Altonji et al. (2012). The first paper computes major-specific measures of occupational concentration using a Hirschman-Hirfindahl Index (HHI). Altonji et al. (2012) calculate the share of graduates from each major employed in the three most popular occupations for that major. Using the same intuition, we propose a similar measure to assess an occupation's entry barriers and degree of specialization. We compute the college major concentration using an HHI by occupation. We weight this index by the rates of college graduates in each occupation:

$$HHI_o = s_{co}\left(\sum_{m=1}^{M} s_{mo}^2\right), \tag{9}$$

where m denotes the major, o the occupation, and s_{mo} denotes the share of graduates

Table 16: Hourly wages by college major and skill mismatch

	Wage First Job	Wage Age 23	Wage Age 26	Wage Age 29
High-school graduate (Adequate match)	7.111	7.405	7.997	8.450
High-school graduate (Skill mismatch)	6.949	7.220	7.983	8.388
Humanities and Arts (Col) (Adequate match)	6.800	7.297	8.076	8.410
Humanities and Arts (Col) (Skill mismatch)	7.690	8.253	9.369	9.221
Humanities and Arts (Univ) (Adequate match)	8.873	8.412	8.579	9.335
Humanities and Arts (Univ) (Skill mismatch)	8.210	8.144	8.612	9.270
Business and Law (Col) (Adequate match)	7.376	7.505	8.243	8.916
Business and Law (Col) (Skill mismatch)	7.049	7.474	8.667	8.824
Business and Law (Univ) (Adequate match)	8.225	8.661	10.232	10.009
Business and Law (Univ) (Skill mismatch)	8.693	8.873	10.594	10.647
Social sciences (Col) (Adequate match)	7.517	7.414	7.857	8.204
Social sciences (Col) (Skill mismatch)	8.265	7.906	7.617	8.009
Social sciences (Univ) (Adequate match)	8.511	8.984	9.465	9.760
Social sciences (Univ) (Skill mismatch)	8.505	8.120	9.102	8.872
Health (Col) (Adequate match)	7.661	7.934	8.546	9.248
Health (Col) (Skill mismatch)	8.142	8.219	8.107	8.812
Health (Univ) (Adequate match)	9.214	8.913	9.921	10.124
Health (Univ) (Skill mismatch)	8.581	9.433	9.071	9.895
STEM (Col) (Adequate match)	8.392	8.506	9.335	9.735
STEM (Col) (Skill mismatch)	7.419	7.603	9.014	9.930
STEM (Univ) (Adequate match)	8.715	8.991	9.881	10.906
STEM (Univ) (Skill mismatch)	8.364	8.423	9.602	10.765
Education (Adequate match)	7.436	7.263	7.691	8.346
Education (Skill mismatch)	7.224	6.853	7.906	7.892
Total	7.130	7.354	8.126	8.563

in m that work in occupation o. This is adjusted by the share of graduates s_{co} in each occupation o. In this case, values close to 1 denote a high concentration of graduates in a single college major (e.g., doctors with a Health degree). Conversely, values close to 0 denote a low concentration of graduates (e.g., kitchen helpers as in Appendix Figure **??**) or a higher number of graduates in different college majors (for instance, contact center salesperson as in Appendix Figure **??**). For example, Appendix Table 18 shows that lawyers have one of the highest index (0.885), while elementary workers have one of the lowest index (0.141). Moreover, Appendix Table **??** shows that more than 50% of Health and Education graduates sort into the top 5 common occupations for each major, as computed by Altonji et al. (2012). Similarly, in Appendix Figure **??**, these are the majors with the highest average adjusted HHI by occupation and the lowest rates of full mismatch.

Figure **??** shows a clear negative relationship between college major concentration and different skill mismatch rates. Indeed, occupations with higher levels of concentration present lower rates of mismatch. Based on Figure **??**, it is clear that occupations with higher concentrations of college majors have lower skill mismatch rates. For occupations with an index higher than 0.75, the reported skill mismatch rate is close to 10%. For occupations with low concentration, this is not true: almost 65% of the individuals are classified as a mismatch (full, horizontal or vertical). There is a clear relationship between the concentration of an occupation and the likelihood of a mismatch. In Table 18, we show that among the top 10 occupations based on the rates of full mismatch are specialized and highly skilled professions, such as dentists, lawyers, accounting professionals, and nurses. These are also occupations with a higher concentration of majors, e.g., individuals holding different majors in that specific occupation are less likely to be observed.

Sorting into occupations with a higher concentration index may be beneficial in terms of wage returns. Appendix Table **??** shows that occupations with higher concentration are associated with higher wages.

B Model

B.1 higher education model

B.2 Instruments

Table 17: Horizontal Mismatch Measure Definition

	JA				
	Complete mismatch	Somewhat match	Complete match	Missing	Total
DA					
Complete mismatch	1,765	221	394	90	2,470
Somewhat match	619	465	2,215	67	3,366
Complete match	557	274	716	45	1,592
Missing	26	4	31	673	734
Total	2,967	964	3,356	875	8,162

Table 18: Occupations and Skill Mismatch

ISCO-08 Occupations	Vertical Mismatch	Horizontal Mismatch	Full Mismatch	HHI Index
Building Architects	0.136	0.068	0.000	0.421
Dentists	0.000	0.000	0.000	0.692
Optometrists and Ophthalmic Opticians	0.347	0.067	0.000	0.677
Special Needs Teachers	0.064	0.029	0.000	0.885
Lawyers	0.000	0.000	0.000	0.849
Nursing Associate Professionals	0.042	0.056	0.000	0.280
Accounting Associate Professionals	0.170	0.085	0.000	0.561
Nursing Professionals	0.037	0.025	0.006	0.641
Social Work and Counselling Professionals	0.208	0.110	0.013	0.503
Payroll Clerks	0.143	0.181	0.019	0.270
[…]				
Shop Sales Assistants	0.806	0.609	0.497	0.158
Waiters	0.782	0.683	0.542	0.151
Crane, hoist and related plant operators	0.680	0.900	0.620	0.147
Cashiers and Ticket Clerks	0.789	0.737	0.632	0.158
Hand Launderers and Pressers	0.828	0.776	0.672	0.137
Freight Handlers	0.798	0.854	0.692	0.141
Dairy Products Makers	0.818	0.849	0.698	0.141
Mail Carriers and Sorting Clerks	0.842	0.877	0.754	0.146
Hand Packers	0.841	0.885	0.758	0.141
Elementary Workers Not Elsewhere Classified	0.946	0.838	0.784	0.141

Figure 4: Stayers in college major programs

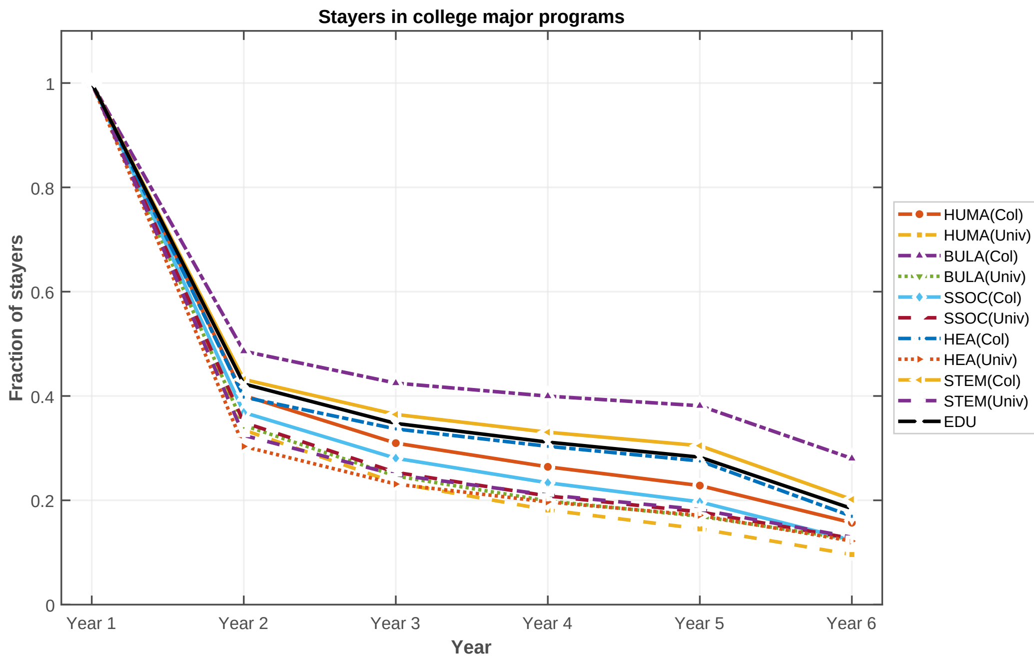


Table 19: Exogenous variables parameters

	Female	Foreign	Number of siblings	Father education	Mother education	Day of birth	Cohort 1978	Cohort 1980
Results higher tertiary	0.334***	-0.570***	-0.022*	-0.004	0.017***	0.019	-0.074**	-0.179***
	(0.029)	(0.088)	(0.012)	(0.004)	(0.005)	(0.013)	(0.034)	(0.038)
Graduation (second sit)	-0.763***	1.155***	0.165***	-0.053**	0.013	0.054	0.764***	0.592***
	(0.147)	(0.434)	(0.060)	(0.022)	(0.024)	(0.068)	(0.181)	(0.177)
Skill mismatch	0.029	0.197	0.052**	-0.038***	0.009	0.001	-0.036	0.146*
	(0.063)	(0.159)	(0.023)	(0.011)	(0.011)	(0.031)	(0.077)	(0.079)
Potential experience (ages 23, 26, 29)	-0.027	-0.575***	-0.019**	-0.039***	-0.029***	-0.022**	0.118***	0.214***
	(0.020)	(0.052)	(0.008)	(0.003)	(0.003)	(0.009)	(0.024)	(0.024)
Unemployment (ages 23, 26, 29)	0.324***	0.416***	0.085***	0.009	-0.009	-0.048**	-0.283***	0.029
	(0.048)	(0.113)	(0.017)	(0.008)	(0.008)	(0.023)	(0.059)	(0.057)
Wage selection (ages 23, 26, 29)	-0.032	-0.236***	0.004	-0.010**	-0.014**	-0.005	-0.601***	-0.510***
	(0.032)	(0.084)	(0.012)	(0.005)	(0.005)	(0.015)	(0.038)	(0.038)
Log-hourly wage matched (ages 23, 26, 29)	-0.078***	0.026*	0.000	0.001	0.004***	-0.002	0.009	0.019***
	(0.007)	(0.015)	(0.002)	(0.001)	(0.001)	(0.002)	(0.006)	(0.006)
Log-hourly wage mismatched (ages 23, 26, 29)	-0.131***	0.001	-0.008***	0.001	-0.001	-0.004	-0.000	0.019**
	(0.010)	(0.018)	(0.003)	(0.002)	(0.002)	(0.004)	(0.009)	(0.009)

Table 20: Relative distance and labor market outcomes

	(1)	(2)	(3)	(4)
	Skill mismatch (first job)	Log-hourly wage (age 23)	Log-hourly wage (age 26)	Log-hourly wage (age 29)
Relative distance to closest tertiary education	0.000	-0.001	-0.004	0.004
	(0.001)	(0.004)	(0.002)	(0.002)
Observations	4391	4391	3391	3119
R^2	0.002	0.002	0.004	0.004

Figure 5: Cumulative BA degree diploma staying in the same major

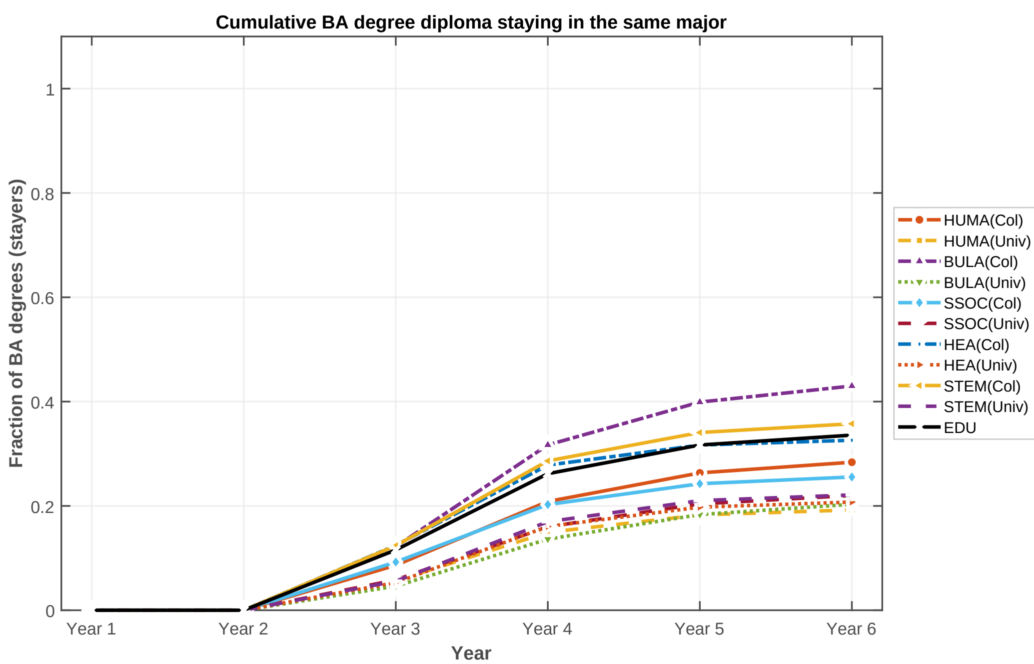


Table 21: Regression of relative distance on individual characteristics

	(1)
Female	0.101
	(0.118)
Number of siblings	-0.085
	(0.141)
Foreign origin	-1.653
	(1.975)
Father years of education	-0.017
	(0.047)
Mother years of education	-0.106
	(0.043)
Day of birth	0.000
	(0.001)
Cohort 1978	1.023
	(0.934)
Cohort 1980	1.376
	(1.327)
Province FE	Yes
Observations	8147
R^2	0.589

C Results

C.1 Model estimates

C.2 Difference across college majors

C.3 Difference within college majors

C.4 Heterogeneity

C.5 Robustness checks

Table 22: Unobserved types and higher secondary education

	Type 1	Type 2	Type 3
Repeated (Primary)	-0.085	0.022	0.111
Repeated (Secondary)	-0.325	0.181	0.277
HS Dropout	-0.203	0.041	0.283
Vocational HS	0.067	-0.061	-0.020
Technical HS	-0.072	0.074	0.009
General HS	0.147	-0.049	-0.175
1st Quarter	0.246	-0.218	-0.085
2nd Quarter	-0.014	0.025	-0.014
3rd Quarter	-0.323	0.316	0.066
4th Quarter	-0.052	-0.023	0.124

Table 23: Difference in returns across college majors

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Business and Law	-0.010	-0.002	0.116***	0.115***
	(0.015)	(0.015)	(0.024)	(0.023)
Social Sciences	-0.029	-0.021	0.003	-0.005
	(0.020)	(0.021)	(0.022)	(0.022)
Health	0.053***	0.061***	0.092***	0.094***
	(0.017)	(0.019)	(0.026)	(0.026)
STEM	0.033**	0.038**	0.074***	0.073***
	(0.015)	(0.016)	(0.024)	(0.022)
Education	-0.034**			
	(0.016)			

Notes: Log-hourly wage at age 23, 26, 29, with year and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

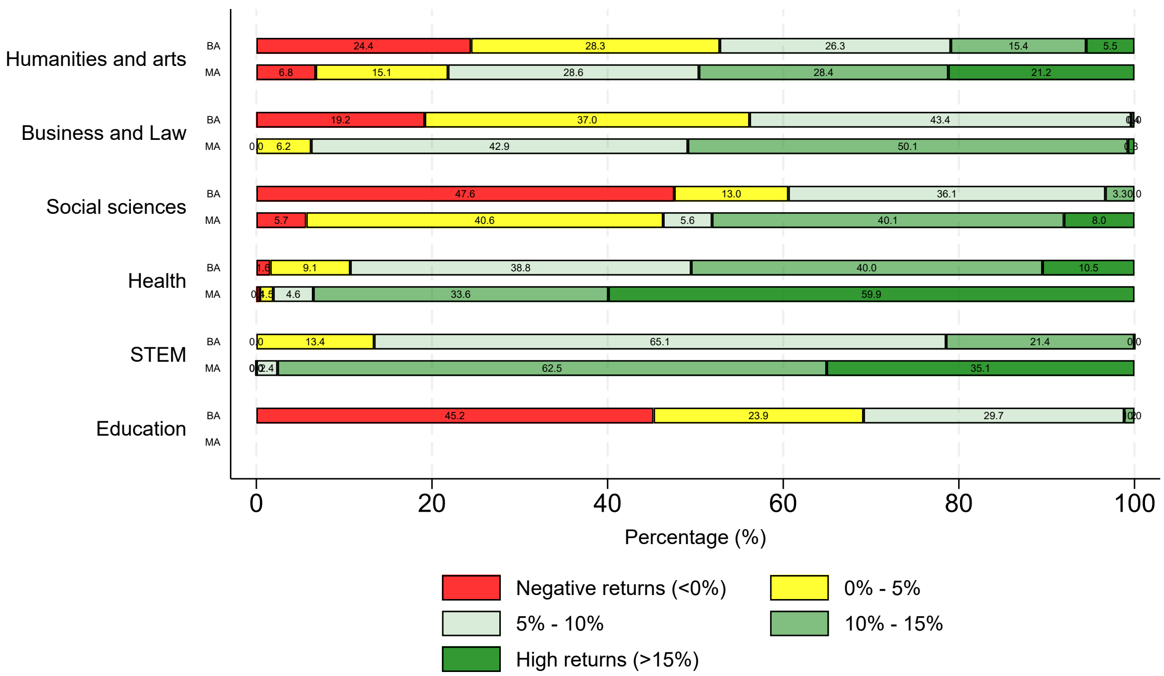


Figure 6: Heterogeneous Returns to college programs (Δ_{ijd}, BA and MA, Col)

Notes: BA and MA denote Bachelor's and Master's degrees. In Belgium, an MA Degree in Education does not exist. The figure includes Δ_{ijd}, which is the individual return to enrolling and obtaining a BA or an MA in college major j. Δ_{ijd} is computed as in Equation 7.

Table 24: Returns to college majors

	Unobserved types	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts	(Type 1)	0.049**	0.099***	0.033	0.090***
		(0.018)	(0.014)	(0.026)	(0.020)
Business and Law		0.032***	0.090***	0.157***	0.213***
		(0.010)	(0.013)	(0.020)	(0.017)
Social Sciences		0.017	0.076***	0.045**	0.095***
		(0.014)	(0.011)	(0.019)	(0.018)
Health		0.097***	0.155***	0.125***	0.184***
		(0.015)	(0.011)	(0.024)	(0.026)
STEM		0.076***	0.131***	0.116***	0.172***
		(0.011)	(0.012)	(0.018)	(0.016)
Education		0.010			
		(0.011)			
Humanities and Arts	(Type 2)	0.049**	0.097***	0.038	0.093***
		(0.018)	(0.014)	(0.025)	(0.020)
Business and Law		0.038***	0.094***	0.161***	0.215***
		(0.009)	(0.014)	(0.020)	(0.017)
Social Sciences		0.023*	0.080***	0.046**	0.093***
		(0.013)	(0.012)	(0.018)	(0.019)
Health		0.103***	0.160***	0.132***	0.190***
		(0.015)	(0.013)	(0.022)	(0.025)
STEM		0.080***	0.133***	0.120***	0.174***
		(0.010)	(0.012)	(0.018)	(0.016)
Education		0.017			
		(0.010)			
Humanities and Arts	(Type 3)	0.023	0.068***	0.066*	0.113***
		(0.027)	(0.020)	(0.038)	(0.030)
Business and Law		0.057***	0.105***	0.146***	0.192***
		(0.017)	(0.026)	(0.026)	(0.024)
Social Sciences		0.041	0.089**	0.025	0.070*
		(0.026)	(0.041)	(0.037)	(0.034)
Health		0.040*	0.078**	0.153***	0.202***
		(0.023)	(0.034)	(0.043)	(0.045)
STEM		0.078***	0.124***	0.092**	0.138***
		(0.019)	(0.022)	(0.036)	(0.023)
Education		0.024			
		(0.016)			

Notes: Log-hourly wage at age 23, 26, 29, with year and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

Table 25: Skill mismatch sorting (Unobserved types)

	Unobserved type	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts	(Type 1)			-0.027	-0.032
				(0.035)	(0.042)
Business and Law		-0.081**	-0.098**	-0.102***	-0.127***
		(0.030)	(0.038)	(0.027)	(0.035)
Social Sciences		-0.104***	-0.129***	0.014	0.015
		(0.028)	(0.031)	(0.031)	(0.037)
Health		-0.174***	-0.224***	-0.129***	-0.164***
		(0.032)	(0.037)	(0.036)	(0.046)
STEM		-0.059*	-0.073*	-0.115***	-0.143***
		(0.033)	(0.039)	(0.026)	(0.035)
Education		-0.170***			
		(0.031)			
Humanities and Arts	(Type 2)			-0.029	-0.034
				(0.040)	(0.044)
Business and Law		-0.092**	-0.106**	-0.116***	-0.139***
		(0.034)	(0.040)	(0.034)	(0.038)
Social Sciences		-0.120***	-0.142***	0.015	0.015
		(0.032)	(0.033)	(0.034)	(0.039)
Health		-0.207***	-0.257***	-0.150***	-0.182***
		(0.039)	(0.042)	(0.043)	(0.051)
STEM		-0.066*	-0.078*	-0.132***	-0.157***
		(0.038)	(0.041)	(0.030)	(0.036)
Education		-0.202***			
		(0.036)			
Humanities and Arts	(Type 3)			-0.010	-0.007
				(0.013)	(0.009)
Business and Law		-0.039**	-0.028*	-0.058***	-0.043**
		(0.016)	(0.013)	(0.016)	(0.016)
Social Sciences		-0.061***	-0.044***	0.004	0.003
		(0.015)	(0.015)	(0.008)	(0.008)
Health		-0.255***	-0.201***	-0.097***	-0.072**
		(0.055)	(0.057)	(0.034)	(0.028)
STEM		-0.024*	-0.018	-0.075***	-0.055***
		(0.012)	(0.011)	(0.023)	(0.017)
Education		-0.222***			
		(0.029)			

Notes: Skill mismatch rates difference with skill mismatch rates when holding a high-school degree. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

Table 26: Returns to college majors and skill mismatch

	Unobserved types	(1) BA degree (mismatch)	(2) BA degree (match)	(3) BA degree (penalty)	(4) MA degree (mismatch)	(5) MA degree (match)	(6) MA degree (penalty)
Humanities and Arts (Col)	(Type 1)	-0.008	0.071***	-0.079**	0.037*	0.132***	-0.095***
		(0.029)	(0.019)	(0.030)	(0.021)	(0.016)	(0.029)
Humanities and Arts (Univ)		0.040	0.031	0.009	0.090**	0.092***	-0.003
		(0.040)	(0.030)	(0.049)	(0.032)	(0.026)	(0.041)
Business and Law (Col)		0.031*	0.033***	-0.002	0.075***	0.095***	-0.020
		(0.016)	(0.011)	(0.018)	(0.026)	(0.012)	(0.026)
Business and Law (Univ)		0.113***	0.164***	-0.051	0.165***	0.225***	-0.060*
		(0.028)	(0.022)	(0.031)	(0.028)	(0.019)	(0.030)
Social Sciences (Col)		0.014	0.018	-0.004	0.058	0.080***	-0.022
		(0.026)	(0.016)	(0.031)	(0.042)	(0.013)	(0.047)
Social Sciences (Univ)		-0.006	0.065***	-0.071	0.039	0.127***	-0.088*
		(0.038)	(0.021)	(0.042)	(0.037)	(0.022)	(0.043)
Health (Col)		-0.042	0.102***	-0.144***	-0.004	0.163***	-0.167***
		(0.031)	(0.015)	(0.034)	(0.050)	(0.011)	(0.048)
Health (Univ)		0.127**	0.125***	0.002	0.176***	0.186***	-0.010
		(0.050)	(0.030)	(0.069)	(0.055)	(0.032)	(0.073)
STEM (Col)		0.048**	0.083***	-0.034	0.096***	0.144***	-0.047*
		(0.019)	(0.012)	(0.020)	(0.021)	(0.014)	(0.026)
STEM (Univ)		0.056	0.124***	-0.067	0.099***	0.185***	-0.086**
		(0.044)	(0.022)	(0.051)	(0.031)	(0.020)	(0.040)
Education		-0.008	0.011	-0.019			
		(0.019)	(0.011)	(0.025)			
Humanities and Arts (Col)	(Type 2)	-0.009	0.078***	-0.088***	0.036*	0.140***	-0.103***
		(0.029)	(0.018)	(0.031)	(0.021)	(0.016)	(0.028)
Humanities and Arts (Univ)		0.039	0.039	0.000	0.088**	0.100***	-0.012
		(0.040)	(0.029)	(0.049)	(0.032)	(0.025)	(0.039)
Business and Law (Col)		0.030*	0.041***	-0.011	0.074***	0.102***	-0.028
		(0.016)	(0.010)	(0.017)	(0.024)	(0.013)	(0.023)
Business and Law (Univ)		0.112***	0.172***	-0.060*	0.163***	0.233***	-0.070**
		(0.029)	(0.022)	(0.033)	(0.027)	(0.020)	(0.031)
Social Sciences (Col)		0.013	0.026*	-0.013	0.057	0.087***	-0.030
		(0.027)	(0.015)	(0.030)	(0.042)	(0.013)	(0.046)
Social Sciences (Univ)		-0.007	0.073***	-0.080*	0.039	0.134***	-0.096**
		(0.040)	(0.019)	(0.042)	(0.037)	(0.021)	(0.043)
Health (Col)		-0.043	0.110***	-0.153***	-0.003	0.171***	-0.174***
		(0.030)	(0.015)	(0.032)	(0.048)	(0.013)	(0.045)
Health (Univ)		0.125**	0.132***	-0.007	0.172***	0.194***	-0.021
		(0.051)	(0.029)	(0.069)	(0.055)	(0.032)	(0.073)
STEM (Col)		0.047**	0.090***	-0.043**	0.096***	0.151***	-0.055**
		(0.020)	(0.010)	(0.019)	(0.020)	(0.014)	(0.024)
STEM (Univ)		0.055	0.132***	-0.076	0.100***	0.193***	-0.093**
		(0.044)	(0.022)	(0.053)	(0.029)	(0.022)	(0.041)
Education		-0.009	0.019	-0.028			
		(0.020)	(0.011)	(0.026)			
Humanities and Arts (Col)	(Type 3)	0.019	0.097***	-0.078	0.065***	0.158***	-0.093**
		(0.028)	(0.033)	(0.046)	(0.021)	(0.028)	(0.037)
Humanities and Arts (Univ)		0.067	0.057	0.010	0.114***	0.118***	-0.004
		(0.041)	(0.040)	(0.063)	(0.033)	(0.034)	(0.050)
Business and Law (Col)		0.058***	0.059**	-0.001	0.104***	0.120***	-0.016
		(0.019)	(0.027)	(0.035)	(0.027)	(0.023)	(0.030)
Business and Law (Univ)		0.140***	0.190***	-0.050	0.188***	0.251***	-0.063
		(0.030)	(0.038)	(0.052)	(0.027)	(0.033)	(0.044)
Social Sciences (Col)		0.041	0.044	-0.003	0.087*	0.105***	-0.019
		(0.030)	(0.033)	(0.044)	(0.045)	(0.028)	(0.052)
Social Sciences (Univ)		0.021	0.091***	-0.070	0.067*	0.152***	-0.085*
		(0.039)	(0.031)	(0.051)	(0.036)	(0.028)	(0.044)
Health (Col)		-0.015	0.128***	-0.143***	0.030	0.189***	-0.159***
		(0.034)	(0.032)	(0.050)	(0.049)	(0.026)	(0.052)
Health (Univ)		0.154***	0.151***	0.003	0.200***	0.212***	-0.012
		(0.053)	(0.036)	(0.074)	(0.053)	(0.035)	(0.069)
STEM (Col)		0.075***	0.108***	-0.033	0.122***	0.170***	-0.047
		(0.021)	(0.025)	(0.036)	(0.023)	(0.022)	(0.031)
STEM (Univ)		0.083*	0.150***	-0.066	0.130***	0.211***	-0.081*
		(0.042)	(0.033)	(0.059)	(0.026)	(0.029)	(0.041)
Education		0.019	0.037	-0.018			
		(0.021)	(0.026)	(0.034)			

Notes: Log-hourly wage at age 23, 26, 29, with time and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. Mismatch and match explanation:

Table 27: Returns to college majors (JA)

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts	0.041***	0.102***	0.041**	0.103***
	(0.015)	(0.014)	(0.020)	(0.017)
Business and Law	0.036***	0.095***	0.152***	0.208***
	(0.011)	(0.013)	(0.017)	(0.016)
Social Sciences	0.014	0.072***	0.027	0.089***
	(0.016)	(0.018)	(0.020)	(0.018)
Health	0.095***	0.150***	0.142***	0.197***
	(0.014)	(0.017)	(0.025)	(0.026)
STEM	0.077***	0.134***	0.105***	0.162***
	(0.012)	(0.013)	(0.020)	(0.017)
Education	0.013			
	(0.010)			

Notes: Log-hourly wage at age 23, 26, 29, with year and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

Table 28: Returns to college majors (DSA)

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts	0.043***	0.103***	0.038*	0.097***
	(0.014)	(0.014)	(0.020)	(0.018)
Business and Law	0.037***	0.098***	0.145***	0.206***
	(0.011)	(0.014)	(0.018)	(0.017)
Social Sciences	0.015	0.077***	0.035*	0.096***
	(0.016)	(0.018)	(0.019)	(0.019)
Health	0.097***	0.160***	0.121***	0.182***
	(0.014)	(0.017)	(0.023)	(0.024)
STEM	0.079***	0.140***	0.098***	0.159***
	(0.011)	(0.013)	(0.020)	(0.018)
Education	0.013			
	(0.010)			

Notes: Log-hourly wage at age 23, 26, 29, with year and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

Table 29: Returns to college majors and skill mismatch (JA)

Level of Education	College Major	(1) Univ Without Mismatch Returns	(2) Univ Full Mismatch Returns	(3) Univ Full Mismatch Penalty	(4) Col Without Mismatch Returns	(5) Col Full Mismatch Returns	(6) Col Full Mismatch Penalty
BA Degree	Humanities and Arts	0.057**	0.021	-0.036	0.053***	0.028	-0.025
		(0.025)	(0.030)	(0.037)	(0.019)	(0.025)	(0.032)
	Business and Law	0.154***	0.136***	-0.018	0.034***	0.042**	0.007
		(0.020)	(0.030)	(0.033)	(0.012)	(0.017)	(0.017)
	Social Sciences	0.090***	-0.031	-0.121***	0.008	0.028	0.019
		(0.027)	(0.029)	(0.038)	(0.018)	(0.026)	(0.029)
	Health	0.140***	0.074	-0.066	0.098***	0.078*	-0.020
		(0.024)	(0.113)	(0.119)	(0.015)	(0.041)	(0.044)
	STEM	0.116***	0.069*	-0.047	0.090***	0.039*	-0.050**
		(0.024)	(0.036)	(0.042)	(0.012)	(0.022)	(0.022)
	Education	0.015	-0.005	-0.020	0.015	-0.005	-0.020
		(0.011)	(0.027)	(0.028)	(0.011)	(0.027)	(0.028)
MA Degree	Humanities and Arts	0.110***	0.093***	-0.017	0.106***	0.100***	-0.006
		(0.023)	(0.025)	(0.032)	(0.018)	(0.026)	(0.032)
	Business and Law	0.207***	0.208***	0.001	0.087***	0.114***	0.026
		(0.019)	(0.026)	(0.029)	(0.015)	(0.023)	(0.025)
	Social Sciences	0.143***	0.041*	-0.102***	0.062***	0.100***	0.038
		(0.027)	(0.024)	(0.035)	(0.021)	(0.031)	(0.035)
	Health	0.193***	0.146	-0.047	0.151***	0.150***	-0.001
		(0.024)	(0.110)	(0.113)	(0.018)	(0.042)	(0.046)
	STEM	0.169***	0.141***	-0.028	0.143***	0.111***	-0.032
		(0.020)	(0.030)	(0.036)	(0.015)	(0.022)	(0.024)

Notes: Log-hourly wage at age 23, 26, 29, with time and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. Standard errors in parentheses.

Table 30: Returns to college majors and skill mismatch (DSA)

Level of Education	College Major	(1) Univ Without Mismatch Returns	(2) Univ Full Mismatch Returns	(3) Univ Full Mismatch Penalty	(4) Col Without Mismatch Returns	(5) Col Full Mismatch Returns	(6) Col Full Mismatch Penalty
BA Degree	Humanities and Arts	0.034	0.057	0.023	0.050***	0.014	-0.036
		(0.022)	(0.038)	(0.042)	(0.016)	(0.034)	(0.039)
	Business and Law	0.144***	0.149***	0.005	0.037***	0.038*	0.001
		(0.018)	(0.043)	(0.045)	(0.011)	(0.022)	(0.022)
	Social Sciences	0.054**	-0.031	-0.084**	0.015	0.021	0.006
		(0.022)	(0.037)	(0.040)	(0.016)	(0.033)	(0.035)
	Health	0.116***	0.190***	0.074	0.105***	-0.047	-0.152***
		(0.026)	(0.069)	(0.076)	(0.014)	(0.050)	(0.052)
	STEM	0.101***	0.073	-0.028	0.080***	0.072***	-0.008
		(0.022)	(0.047)	(0.051)	(0.012)	(0.027)	(0.027)
	Education	0.014	0.000	-0.014	0.014	0.000	-0.014
		(0.011)	(0.038)	(0.040)	(0.011)	(0.038)	(0.040)
MA Degree	Humanities and Arts	0.098***	0.095***	-0.003	0.114***	0.052	-0.062*
		(0.021)	(0.030)	(0.034)	(0.016)	(0.033)	(0.037)
	Business and Law	0.208***	0.187***	-0.021	0.101***	0.076***	-0.025
		(0.018)	(0.039)	(0.041)	(0.014)	(0.029)	(0.030)
	Social Sciences	0.117***	0.007	-0.110***	0.079***	0.059	-0.020
		(0.021)	(0.032)	(0.037)	(0.019)	(0.041)	(0.043)
	Health	0.179***	0.228***	0.049	0.169***	-0.009	-0.178***
		(0.025)	(0.064)	(0.069)	(0.017)	(0.051)	(0.055)
	STEM	0.165***	0.111***	-0.054	0.144***	0.110***	-0.034
		(0.019)	(0.038)	(0.042)	(0.014)	(0.029)	(0.030)

Notes: Log-hourly wage at age 23, 26, 29, with time and cohort fixed effects. We control for potential experience, unemployment rates and dynamic selection. Standard errors in parentheses.

Table 31: Skill mismatch sorting (JA)

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts			0.039	0.039
			(0.036)	(0.036)
Business and Law	-0.098***	-0.098***	-0.166***	-0.169***
	(0.025)	(0.025)	(0.027)	(0.027)
Social Sciences	-0.115***	-0.116***	0.081**	0.080***
	(0.028)	(0.028)	(0.031)	(0.030)
Health	-0.250***	-0.253***	-0.236***	-0.240***
	(0.027)	(0.026)	(0.029)	(0.030)
STEM	-0.138***	-0.140***	-0.139***	-0.140***
	(0.024)	(0.024)	(0.027)	(0.028)
Education	-0.241***			
	(0.024)			

Notes: Skill mismatch rates difference with skill mismatch rates when holding a high-school degree. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

Table 32: Skill mismatch sorting (DSA)

	(1) BA degree (Col)	(2) MA degree (Col)	(3) BA degree (Univ)	(4) MA degree (Univ)
Humanities and Arts			-0.004	-0.005
			(0.032)	(0.030)
Business and Law	-0.068***	-0.061***	-0.083***	-0.076***
	(0.024)	(0.020)	(0.026)	(0.023)
Social Sciences	-0.073***	-0.067***	0.016	0.013
	(0.027)	(0.022)	(0.030)	(0.026)
Health	-0.112***	-0.100***	-0.101***	-0.091***
	(0.027)	(0.022)	(0.028)	(0.025)
STEM	-0.063***	-0.058***	-0.062**	-0.058**
	(0.023)	(0.020)	(0.026)	(0.024)
Education	-0.114***			
	(0.024)			

Notes: Skill mismatch rates difference with skill mismatch rates when holding a high-school degree. We control for potential experience, unemployment rates and dynamic selection. BA degree are defined when attaining 3 years of higher tertiary education, while MA degree are defined over attaining more than 3 years of higher tertiary education. Col includes vocational higher tertiary education institutions, while Univ includes accademic higher tertiary education institutions. In the Belgian system, there is not a MA degree for Education degrees.

Table 33: Negative returns by level of education and college majors (JA, in %)

Level of Education	College Major	(1) Univ Average Mismatch (in %)	(2) Univ Without Mismatch (in %)	(3) Univ Full Mismatch (in %)	(4) Col Average Mismatch (in %)	(5) Col Without Mismatch (in %)	(6) Col Full Mismatch (in %)
BA Degree	Humanities and Arts	22.2	19.7	28.6	23.3	24.4	24.7
		(8.8)	(10.5)	(12.2)	(5.1)	(5.4)	(10.2)
	Business and Law	1.0	0.2	8.6	22.1	25.5	15.3
		(1.1)	(0.8)	(4.7)	(5.1)	(6.4)	(7.4)
	Social Sciences	28.5	11.2	46.9	33.7	35.5	28.8
		(5.8)	(7.4)	(8.0)	(6.6)	(7.5)	(10.5)
	Health	11.5	9.1	34.0	5.6	4.8	12.7
		(5.9)	(6.5)	(6.2)	(3.8)	(3.8)	(12.7)
	STEM	4.3	3.2	12.8	3.7	1.5	18.1
		(4.2)	(4.6)	(11.4)	(2.7)	(2.1)	(10.7)
	Education	34.4	33.8	40.3	34.4	33.8	40.3
		(5.1)	(5.6)	(10.7)	(5.1)	(5.6)	(10.7)
MA Degree	Humanities and Arts	5.9	5.8	8.2	9.6	11.5	6.9
		(4.9)	(7.5)	(5.4)	(2.7)	(3.7)	(4.4)
	Business and Law	0.2	0.0	2.3	3.2	3.9	1.4
		(0.3)	(0.1)	(2.4)	(4.2)	(5.2)	(2.1)
	Social Sciences	14.6	3.7	26.8	16.7	18.9	10.5
		(4.8)	(3.9)	(8.2)	(8.1)	(10.3)	(7.8)
	Health	5.5	3.0	29.7	0.8	0.5	3.6
		(3.3)	(3.3)	(5.0)	(1.2)	(1.1)	(6.8)
	STEM	0.3	0.2	1.8	0.2	0.0	1.4
		(0.5)	(0.5)	(3.0)	(0.3)	(0.1)	(2.1)

Notes: Each cell reports the simulated fraction (percentage) of individuals with negative returns in the distribution of individual returns by education and mismatch type. Standard errors in parentheses.

Table 34: Negative returns by level of education and college majors (DSA, in %)

Level of Education	College Major	(1) Univ Average Mismatch (in %)	(2) Univ Without Mismatch (in %)	(3) Univ Full Mismatch (in %)	(4) Col Average Mismatch (in %)	(5) Col Without Mismatch (in %)	(6) Col Full Mismatch (in %)
BA Degree	Humanities and Arts	20.4	20.1	20.9	24.0	22.6	30.8
		(10.9)	(12.8)	(10.8)	(5.2)	(5.4)	(17.0)
	Business and Law	0.6	0.2	4.8	17.1	16.7	19.1
		(0.8)	(0.6)	(4.9)	(9.4)	(10.3)	(10.2)
	Social Sciences	25.1	18.9	47.4	32.5	32.3	32.3
		(7.8)	(9.2)	(12.1)	(6.8)	(7.4)	(9.3)
	Health	14.0	14.0	14.0	3.8	1.5	47.2
		(8.6)	(9.2)	(8.4)	(2.3)	(2.3)	(10.8)
	STEM	5.6	3.9	17.9	1.5	0.6	10.1
		(4.5)	(4.7)	(11.9)	(1.2)	(0.9)	(8.4)
	Education	33.7	33.5	36.6	33.7	33.5	36.6
		(3.7)	(3.9)	(11.9)	(3.7)	(3.9)	(11.9)
MA Degree	Humanities and Arts	4.0	2.4	12.2	9.1	7.9	16.0
		(4.4)	(5.0)	(7.6)	(3.5)	(3.3)	(12.7)
	Business and Law	0.2	0.0	2.2	1.1	0.4	8.2
		(0.2)	(0.0)	(2.8)	(1.0)	(0.6)	(7.5)
	Social Sciences	9.9	3.8	35.8	11.4	10.0	23.7
		(3.9)	(4.5)	(10.6)	(8.7)	(9.4)	(10.2)
	Health	3.8	3.3	10.7	1.8	0.0	39.3
		(5.1)	(5.3)	(7.3)	(0.8)	(0.3)	(11.4)
	STEM	1.1	0.1	10.3	0.3	0.0	3.3
		(1.0)	(0.4)	(8.4)	(0.4)	(0.0)	(4.1)

Notes: Each cell reports the simulated fraction (percentage) of individuals with negative returns in the distribution of individual returns by education and mismatch type. Standard errors in parentheses.