Sociocultural Background and Choice of STEM Majors at University*

Aderonke Osikominu,† Volker Grossmann‡ and Marius Osterfeld§

February 9, 2019

Abstract

This paper proposes a generalized Roy model to examine the role of students' sociocultural background for choosing a STEM major at university. We combine survey data on Swiss university graduates with rich municipality level information. We use a principal component analysis to construct an indicator capturing progressive attitudes in a student's home environment. Our structural approach allows directly comparing the importance of sociocultural background with that of pecuniary returns and costs in the choice of college major. Identification exploits individual differences in the relative cost of studying STEM that are unrelated to the local economic environment. Male students from conservative backgrounds are more likely to study STEM, whereas women are unaffected by sociocultural background besides majority language. The effect of the progressivism indicator for males is about half of the effect of the earnings return to STEM and twice as large as the effect of the relative monetary cost.

Key words: Choice of field of study; Generalized Roy model; Sociocultural environment; STEM fields.

JEL classification: I20, C31.

*We are grateful to the Associate Editor Ken Mayhew, two anonymous reviewers, Thomas Bauer, Bernd Fitzenberger, and seminar participants at the University of Fribourg, the University of St. Gallen, the Annual Congress of the International Institute of Public Finance (IIPF), the European Economic Association (EEA), the German Economic Association (Verein für Socialpolitik), the CESifo Area Conference in Economics of Education, the 4th SOLE/EALE World Conference, the annual conference of the population economics group of the German Economic Association, and the Asian Meeting of the Econometric Society for valuable comments and suggestions.

†University of Hohenheim, Germany; Institute for Employment Research (IAB), Nuremberg; Centre for Economic Policy Research (CEPR), London; CESifo, Munich; Institute of Labor Economics (IZA), Bonn. E-mail: a.osikominu@uni-hohenheim.de.

‡Corresponding author. University of Fribourg, Switzerland; CESifo, Munich; Centre for Research and Analysis of Migration (CReAM), University College London; Institute of Labor Economics (IZA), Bonn. E-mail: volker.grossmann@unifr.ch.

§swissstaffing. E-mail: marius.osterfeld@swissstaffing.ch.

1 Introduction

Technological progress relies on innovations that are created by scientists and engineers. There is consequently a widespread consensus that the so-called *STEM* (science, technology, engineering, and mathematics) skills are of major importance to sustain innovation and growth.[1] Yet, the average share of new university entrants into STEM fields in OECD countries in the year 2011 was only about 25%, where 39% of male students and 14% of female students choose a STEM field (OECD, 2013, Tab. C3.3b). These figures cause concern to policymakers and employers who perceive a lack of qualified candidates for jobs requiring STEM skills.

Akerlof and Kranton (2000, 2002) point out that the social setting may be an important driver of educational choices. They propose a human capital model that incorporates social incentives in addition to the usual economic returns and constraints. Macroeconomic evidence highlights the role of cultural factors for economic growth (e.g., Guiso et al., 2006; Tabellini, 2008; Becker and Woessmann, 2009). However, the relevance of the sociocultural setting for microeconomic outcomes has proven difficult to establish empirically.

This paper sets out to explore to what extent the decision to study a STEM field is influenced by the sociocultural background of a student at the micro level, combining unique survey data on the population of university graduates in Switzerland with rich municipality level information. We exploit the important differences in social and cultural norms and attitudes in the about 2,600 municipalities, while economically Switzerland is highly integrated. We construct an indicator capturing progressive attitudes in a student's home environment from a principal component analysis that is based on referenda results about gender- and science-related progressive issues, parliamentary election results, and the share of Catholics in the municipality a high school graduate resided in before entering university.[2]

At first glance, one may think that a progressive environment contributes to developing a taste for science. Alternatively, however, STEM fields may be preferred in conservative

[1] See Hunt and Gauthier-Loiselle (2010) and Winters (2014), who present evidence for large human capital externalities of STEM graduates, e.g., in the form of patents.

[2] In Switzerland, Catholics typically are associated with more conservative religious values compared to Protestants, which has been attributed to the Reformation process in Switzerland itself (e.g., Gordon, 2002). The share of Catholics indeed enters negatively in our progressivism indicator.

environments vis-à-vis, for instance, social sciences which may be viewed as being oriented to liberal political attitudes. One may also hypothesize that conservatism is related to a low fraction of females choosing a STEM field by creating an environment in which certain fields and occupations are perceived as being better suited for either males or females.

We derive a structural estimation framework from a generalized Roy model (Roy, 1951; Heckman and Vytlacil, 2007) that accounts for differences in both pecuniary and subjective benefits and costs of studying a STEM major rather than other university majors. Our approach allows us to directly compare the importance of sociocultural background with the pecuniary returns and costs in the choice of college major. When students make their choice they evaluate the return of a STEM major relative to a non-STEM major in terms of earnings as well as the cost differential between the two. Since we observe earnings associated with a particular major only for the group of people who self-select into the major, we have to estimate the conditional expectation functions of earnings associated with STEM and non-STEM majors by correcting for selection bias in order to construct the pecuniary return to STEM majors.

At the time of our study all Swiss inhabitants graduating from general high school could freely choose which field and at which Swiss university to study irrespective of prior specializations. Study fees are rather low in international comparison. They are similar across universities and the same for all majors within a university. The main driver of pecuniary costs is the geographic proximity of a student's home municipality to the next technical university (exclusively specializing in STEM fields and enjoying high reputation) relative to the next standard university. Thus, we exploit that the *relative* geographic proximity of a technical university affects the selection into majors. We argue that, controlling for local economic conditions in the region a university graduate resided in before entering university, relative distance to the next technical university is unrelated to their earnings capability, thus serving as an exclusion restriction. In extensive sensitivity analyses, we verify that our identification strategy proves robust to modifications of our benchmark model.

The empirical analysis focusses on Swiss university graduates who finished their studies in the early 2000s and for which we observe earnings five years after graduation, parental education, age, gender, and the home municipality in which they lived at the end of high

school. To measure a respondent's sociocultural environment at the municipality level in a way that is not contaminated by social desirability considerations, we take advantage of the unique direct democratic system in Switzerland. In addition to the fraction of votes that accrued to left-wing parties at the national parliamentary elections in 1995 and religious denomination, our indicator of progressivism in a municipality is constructed from the results of four national referenda on introducing equal rights of men and women in the constitution (referendum held in 1981), on providing addicts with medical prescriptions of heroin (1999), on allowing stem cell research (2004), and on the civil union of homosexual couples (2005), providing similar rights than to married couples. Further, we capture the sociocultural environment by the majority language in a municipality.

Our main results are as follows. First, students from more progressive municipalities are significantly less likely to study a STEM field at university. Regressions with gender interactions suggest further, that the sociocultural environment affects especially the behavior of young men. An increase in the progressivism indicator by one standard deviation reduces the probability to choose a STEM major at university by 3.9 percentage points (around 13%), according to our benchmark model. In absolute terms, its magnitude is about half of the effect of the earnings differential between STEM and non-STEM graduates and twice as large as the effect of relative distance to the next technical university. In contrast, a progressive environment does not affect the probability to study a STEM major for women. In the majority language dimension, sociocultural background matters for women as well. Moreover, while the pecuniary return to STEM majors plays a large role for men, our evidence suggests that it does not affect the probability to study STEM among women.

The main innovation of our approach is to capture social incentives as reflected in the sociocultural background of a student. Previous studies on college major choice focussed on the role of quantitative abilities (e.g., Arcidiacono, 2004; Wang, 2013; Stinebrickner and Stinebrickner, 2014), specialization in high school (Card and Payne, 2017), parental background (e.g., Boudarbat and Montmarquette, 2009; Sonnert, 2009), expected future labour force participation (e.g., Polachek, 1978; Blakemore and Low, 1984), expected lifetime earnings (e.g., Berger, 1988; Eidea and Waehrer, 1998; Boudarbat, 2008; Arcidiacono et al., 2012; Wiswall and Zafar, 2015) as well as tastes and social orientation of an individual (Humlum et

al., 2012; Wiswall and Zafar, 2015). Our paper contributes to this literature by employing a unique data set to examine the role of objective measures for the sociocultural environment for individual tastes based on a structural model. Addressing the gender gap in STEM field choice, Carrell et al. (2010) find that the gender of the professor plays a role for both females' performance in basic math and science classes and the choice of women with high quantitative skills to graduate from a STEM field. Card and Payne (2017) document that the course orientation in high school is important for the gender gap. Finally, Montmarquette et al. (2002) and Zafar (2013) that earnings are a more important determinant of major choice for men than for women. Our results are consistent with the latter finding and generally evaluate the impact of the relative return and costs to study STEM for men and women separately.

The paper is organized as follows. In Section 2, we sketch the theoretical background and the empirical strategy. Section 3 describes the institutional setup and data. Section 4 discusses the empirical implementation. Section 5 presents the estimation results. The last section concludes. The Online Appendix contains further details on the theoretical and empirical framework as well as a large body of additional empirical evidence.

2 Framework

We now sketch the generalized Roy model that we estimate.[3] We focus on the binary decision to study a STEM field (alternative 1) or a non-STEM field (alternative 0). According to the model individuals evaluate the benefits and costs over the life cycle associated with the choice of a particular major at university. Individuals select the major that yields the higher net benefit for them. In addition to monetary benefits and costs we also introduce net subjective costs that stem from a perceived subjective distance between the own sociocultural background and the values typically associated with a university major.

The monetary return to STEM-fields is not observable to an econometrician because earnings are only observed for the field actually chosen. The econometrician therefore has to take into account the self-selection of individuals into study fields according to their indi-

[3]For details of the model that gives rise to a structural binary choice model and the empirical identification strategy, see Online Appendix.

vidual comparative advantage. To identify the conditional expectation functions of earnings associated with STEM and non-STEM fields free of selection bias we rely on a multi-stage estimation procedure and exclusion restrictions.[4]

Specifically, the estimation proceeds as follows. Let $\mathbf{x}_i$ be a row vector of variables that influence earnings capability of an individual i and denote by $\mathbf{z}_i$ a row vector of individual characteristics that influence tastes for a university major, reflecting monetary and net subjective costs of the majors. Also define $\mathbf{w}_i \equiv (1, \mathbf{x}_i, \mathbf{z}_i)$. At stage 1, we estimate

$$\Pr\{s_i = 1 \,|\, \mathbf{w}_i\} = \Pr\{\mathbf{w}_i \pi \geq \varepsilon_i \,|\, \mathbf{w}_i\} = \Phi\left(\mathbf{w}_i \pi\right), \tag{1}$$

where s_i is an indicator function equal to one when utility from choosing a STEM field is at least as high as utility when choosing a non-STEM field, π is a column vector of coefficients, and ε_i is the error term. For practical purposes, we assume the error term to be standard normally distributed, i.e., $\varepsilon_i \,|\, \mathbf{w}_i \sim \mathcal{N}(0,1)$.[5] Φ thus denotes the standard normal c.d.f.; its p.d.f. is denoted by φ.

Suppose that earnings y_{ij} (observed at some point of the career) of individual i when choosing major $j \in \{0,1\}$ are in the same way for both majors proportional to the present discounted value of the earnings stream. y_{ij} is given by some major-specific function f_j, which depends on the observable characteristics $\mathbf{x}_i$ and an individual- and major-specific ability component, u_{ij}, that is unobservable for an econometrician; that is, $y_{ij} = f_j(\mathbf{x}_i, u_{ij})$. We assume a familiar linear form for log earnings of an individual such that

$$\ln y_{ij} = \beta_{0j} + \mathbf{x}_i \beta_j + u_{ij}\,,\ j \in \{0,1\}\,. \tag{2}$$

where β_{0j} is an intercept and β_j a vector of slope parameters. Further, express the relation-

[4]See, e.g., the seminal work by Heckman (1976, 1979) and Willis and Rosen (1979) in the context of college attendance. Heckman et al. (2006) provide an illuminating survey.

[5]Normality of ε_i and the linear index structure in (1) as well as the linearity assumptions in (2) and (3) are not needed for econometric identification since we also rely on exclusion restrictions. From a practical perspective, a flexible parametric specification that exploits exclusion restrictions seems a good compromise. The Online Appendix shows that we can alternatively relax normality and obtain similar estimation results.

ship between the error terms u_{ij} in wage equation (2) and ε_i as

$$u_{ij} = \gamma_j \varepsilon_i + \zeta_{ij} \,, \tag{3}$$

with expectation $\mathrm{E}\left[u_{ij} \,|\, \mathbf{w}_i, \varepsilon_i\right] = \gamma_j \varepsilon_i$, which implies that $\mathrm{E}\left[\zeta_{ij} \,|\, \mathbf{w}_i, \varepsilon_i\right] = 0$ and $\gamma_j = \mathrm{Cov}(u_{ij}, \varepsilon_i)$ is the covariance of u_{ij} and ε_i. To account for selection bias, at stage 2 we then estimate wage regressions

$$\ln y_{i0} = \beta_{00} + \mathbf{x}_i \beta_0 + \gamma_0 \lambda_{i0} + \eta_{i0} \,, \tag{4}$$

$$\ln y_{i1} = \beta_{01} + \mathbf{x}_i \beta_1 + \gamma_1 \lambda_{i1} + \eta_{i1} \,, \tag{5}$$

where

$$\lambda_{i0} \equiv \frac{\varphi(\mathbf{w}_i \pi)}{1 - \Phi(\mathbf{w}_i \pi)} \,, \;\; \lambda_{i1} \equiv -\frac{\varphi(\mathbf{w}_i \pi)}{\Phi(\mathbf{w}_i \pi)} \,, \tag{6}$$

denote inverse Mills ratios and we define error term $\eta_{ij} \equiv \gamma_j(\varepsilon_i - \lambda_{ij}) + \zeta_{ij}$, $j \in \{0, 1\}$, in the subsamples for which wages associated with alternatives 1 and 0, respectively, are observed. The greater $-\varepsilon_i$ the greater the unobserved relative advantage for STEM fields. Hence, if the estimate of $\gamma_1 = \mathrm{Cov}(u_{i1}, \varepsilon_i)$ is negative, the unobserved relative advantage for STEM, $-\varepsilon_i$, and the unobserved earnings capability u_{i1} are positively related, such that that mean earnings associated with a STEM field are higher in the subgroup of people who have chosen that major than in the total population of college graduates. The opposite holds if the estimate of γ_1 is positive. If $\gamma_0 = \mathrm{Cov}(u_{i0}, \varepsilon_i)$ is positive (negative), there is positive (negative) selection of non-STEM graduates.

In practice, λ_{i0} and λ_{i1} are unknown. We obtain estimates $\hat{\lambda}_{i1}$ and $\hat{\lambda}_{i0}$ by evaluating the right-hand sides of (6) using the estimated coefficients $\hat{\pi}$ from the first stage probit regression (1). As shown by Heckman (1976, 1979), the two-step estimation procedure yields consistent estimates $(\hat{\beta}_{00}, \hat{\beta}_0)$ and $(\hat{\beta}_{01}, \hat{\beta}_1)$ for coefficient vectors (β_{00}, β_0) and (β_{01}, β_1). The conventional OLS standard errors for the estimated coefficients in (4) and (5) are incorrect, however, when $\gamma_j \neq 0$, $j \in \{0, 1\}$, because the conditional variances of the error terms, $\mathrm{Var}(\eta_{i0} \,|\, \mathbf{w}_i, s_i = 0)$, $\mathrm{Var}(\eta_{i1} \,|\, \mathbf{w}_i, s_i = 1)$, are nonconstant and $\hat{\lambda}_{i0}$, $\hat{\lambda}_{i1}$ are generated regressors. Therefore, we bootstrap the full three-step estimation procedure using 499 bootstrap

replications. Specifically, we apply the weighted bootstrap suggested by Barbe and Bertail (1995). For each person in our data set we generate 499 weights based on random draws from a gamma distribution with shape and scale parameters equal to one. Thus, the bootstrap weights are non-integer and the probability that a weight exactly equals zero is zero. With a binary dependent variable and a number of discrete regressors, this bootstrap procedure has the advantage that we avoid having to repeat the sampling if, in a given resample, the maximum likelihood estimation fails to converge or certain covariate settings perfectly predict the dependent variable (see also Fitzenberger and Muehler, 2015, for a similar argument).

In the final stage 3 (structural major choice), we estimate the structural choice equation

$$\Pr\{s_i = 1 \,|\, \mathbf{x}_i, \mathbf{z}_i\} = \Pr\left\{\delta_0 + \alpha[\hat{\beta}_{01} - \hat{\beta}_{00} + \mathbf{x}_i(\hat{\beta}_1 - \hat{\beta}_0)] + \mathbf{z}_i\delta \geq \varepsilon_i \,|\, \mathbf{x}_i, \mathbf{z}_i\right\}, \tag{7}$$

where α is the coefficient on the estimated expected log wage differential between STEM and non-STEM fields, $\hat{\beta}_{01} - \hat{\beta}_{00} + \mathbf{x}_i(\hat{\beta}_1 - \hat{\beta}_0)$, δ is the vector of coefficients on taste characteristics, and δ_0 is a constant. Again, we rely on the bootstrap procedure described above to obtain standard errors that are valid when the generated wage difference is included as regressor.

As discussed in detail in the coming section, we think of the taste characteristics in $\mathbf{z}_i$ that are not included in $\mathbf{x}_i$ as the distance of the students' home municipality to the next technical university relative to the next cantonal university, and the sociocultural environment (progressivism and majority language) in the home municipality at the time the major is chosen. Both $\mathbf{x}_i$ and $\mathbf{z}_i$ include dummies for the broader region a student lives in before entering university as well as, in sensitivity analyses, other variables capturing the economic background that may be correlated with unobserved earnings capability of students. Technical universities specialize in STEM fields whereas the cantonal universities offer only a limited range of less prestigious STEM programs. Thus, relative distance to the next technical university is a prime candidate for an excluded variable in the earnings regressions to capture the variation of relative costs to study a STEM field within regions. We will show that it is unrelated to the earnings differential across the two alternatives and, on average, different for STEM and non-STEM graduates.

3 Background and Data

3.1 Institutional Background

Like Germany, Switzerland is well known for a strong vocational training system. In 2000, around the time we observe the university graduates in our data, 90% of the Swiss population aged 25-34 held at least an upper secondary degree, a figure that has remained stable thereafter (OECD, 2017, Tab. A1.2). Attending a Swiss university requires an upper secondary degree from a general high school (*Gymnasiale Matura*), which is obtained by about 20% of a cohort (Federal Statistical Office, 2015). In 2002, the share of graduates from tertiary-type A programmes to the population of the typical age group was 17.9% of a cohort (OECD, 2004, Tab. A3.1). The figure is comparable to Germany (19.2%), Austria (18%) and France (24.2%). At the time of our study, individuals with a *Gymnasiale Matura* degree – irrespective of the chosen specialization in secondary school – could choose freely among the available programs at all universities in Switzerland, without university entrance exams, restrictions in terms of minimum high school grade point average, or other selection procedures.[6] Tuition fees in Swiss universities are moderate in international comparison, both in absolute terms and relative to housing costs. They are similar across universities and the same for all majors within a university.

In the 1990s and early 2000s, around 75% of the young people with a Swiss general high school degree enrolled at a university within a year after high school graduation (Federal Statistical Office, 2013). There was no distinction between undergraduate and graduate university degrees. Graduating from university meant completing a curriculum comparable to a Master's degree in Switzerland and other advanced countries nowadays.[7]

In the period under study, there were eleven universities in Switzerland, all of which are publicly funded and managed. The two technical universities, called *Eidgenössische Technis-*

[6] Only medical schools began, in the year 1998, to select students according to their grade point average at high school and in an entrance exam. These restrictions are not relevant for the cohorts in our data set.

[7] In the mid-1990s, a second type of academic tertiary education institution, the Universities of Applied Sciences (*Fachhochschulen*), was established. They offered shorter, more practically oriented and occupation-specific programs compared with those at universities. In 1999, 7% of a cohort obtained an upper secondary degree from a vocational high school (*Berufsmatura*) that prepares for study of a particular field at a University of Applied Sciences. High school graduates with a *Gymnasiale Matura* degree but no job experience are not entitled to enrol.

che Hochschulen (ETH), are the only federal universities in Switzerland. They are located in the cities of Lausanne and Zurich. At the time of our data they offered programs in the STEM fields physical sciences, chemistry, biology, geography, geology, mathematics, computing, and technical sciences (primarily engineering). No other fields were offered. All other universities are governed at the cantonal (i.e., state) level. The nine cantonal universities are located in the cities of Basel, Berne, Fribourg, Geneva, Lausanne, Lugano, Neuchâtel, St. Gallen, and Zurich. The cantonal universities offer degree programs in both STEM and non-STEM fields. For STEM fields, however, the technical universities are better endowed, have much bigger departments and offer a wider variety of programs and specializations than the cantonal universities. Thus, a STEM degree from a technical university is considered as somewhat more prestigious than one from a cantonal university.[8] In our data set, 62.8% of STEM university graduates attended one of the two technical universities.

Nationwide drop-out rates are similar across STEM and non-STEM programs and relatively low in international comparison, suggesting a comparatively well informed selection of university students on average. Of those university students enrolled in a particular program in the year 2001, for instance, 30.2% did not graduate from that program within 10 years and the vast majority of them had definitely dropped out of it (Wolter et al., 2014).[9]

There is a pronounced gender gap in choosing STEM fields in Switzerland, that is comparable to Germany, Austria, and France, for instance. According to OECD (2013, web Tab. C3.3b), in 2011, 35.1% of males in tertiary education are in STEM programs (excluding life sciences), but only 8.1% of females. Respective figures are 44.6 vs. 12.2% in Germany, 38.1 vs. 10% in Austria, and 29.4 vs. 13.8% in France.

[8]For instance, Nobel Prize winner Albert Einstein studied between 1896 and 1900 at the ETH Zurich where he later also served as professor. Until today, 21 Nobel prize winners have studied or worked at the ETH Zurich. However, doctoral degrees can be obtained and are equally common at both types of universities.

[9]The drop-out patterns are very different from those reported by Arcidiacono (2004) and Stinebrickner and Stinebrickner (2014) for the US. Their evidence suggests that an extraordinarily high fraction of students, who intend to or actually start a science major, graduate in another major or drop out of university.

3.2 Graduate Survey

Our main data source is the 'Swiss Graduate Survey' of the Federal Statistical Office, a unique survey of the full population of graduates from tertiary academic education in Switzerland. We consider all Swiss respondents who lived in Switzerland when graduating from high school and graduated in 2000 and 2002 from one of the nine cantonal or the two federal universities. All graduates of these two cohorts received a questionnaire one year after graduation. All respondents in the first wave received a follow-up questionnaire five years after graduation. Participation in the survey was voluntary. The response rate was about 60% in the first wave. 65% of those who responded in the first wave responded in the second wave. We use the probability weights provided by the Federal Statistical Office to account for potentially selective nonresponse.

The survey contains a large array of individual characteristics including earnings, hours worked, major at university, gender, specializations in general high school, the level of education of mother and father, as well as the home municipality before entering university. To construct our main dependent variable, we categorize graduates into two groups according to their field of study. We define STEM majors as those offered by the two federal technical universities in Switzerland. The remaining majors are classified as non-STEM or Humanities.

Since expected earnings differences between fields of study may have an important impact on study major choice, the availability of individual earnings several years after graduation is crucial for our estimation strategy. We therefore restrict the analysis sample to those who participate also in the second wave that includes information on earnings five years after graduation. We consider earnings in the main job including also overtime compensation and bonus payments. We divide total earnings by total hours worked (contractual hours plus overtime) to obtain the hourly wage rate we use in the estimations.

To focus on a typical career and mitigate potential measurement error, we focus on employed graduates without extreme or missing values of earnings, study duration, or age. There is little difference between those excluded because of extreme earning values according to gender or field. Graduates with extreme study duration or high graduation age are excluded since we want to focus on first-time graduates. Moreover, some individuals take advantage of the unrestricted university access and enroll without intention to base their

career on the acquired education. Our data reveals that this is the case especially in social sciences, history & culture and literature, but much less so in STEM fields. In total, we drop 25.7% of the original sample. Our final sample includes 4,767 individuals.[10]

3.3 Geographical and Sociocultural Data

For our analysis, an important piece of information in the 'Swiss Graduate Survey' is the home municipality of each graduate at the end of high school. We draw on this variable to characterize a student's sociocultural background at the time of major choice. There are about 2,600 municipalities in Switzerland, which allows us to reconstruct the sociocultural environment at a very detailed regional level.

Most municipalities are small in terms of population size and area. 95% of the municipalities are smaller than 59.1 km^2 (22.8 mi^2) or have less than 11,000 inhabitants. In the 53 municipalities that are larger than 100 km^2, the average fraction of populated land is 1.5%, as most people live downtown.

First, we construct the driving distance from downtown of the home municipality to the next technical university (ETH) relative to the driving distance to the next cantonal university with the help of Google Maps. Second, we determine the majority language and the religious environment (share of Catholics) of a graduate's home municipality using information from the 'Federal Population Census' in 1990. Third, we calculate the total vote share which accrued to left-wing parties in the Swiss national election in 1995 based on municipality level election data. Fourth, we use the results from four nationwide referenda that were particularly salient in the public debate to capture how progressive views were on gender equality and science-related issues: on introducing equal rights of men and women in the constitution held in 1981, on providing drug addicts with medical prescriptions of heroin held in 1999, on allowing stem cell research held in 2004, and on a civil union of homosexual couples held in 2005. We consider the latter referendum to be science-related

[10] In the Online Appendix, we show that the main results are robust to lifting the sample restrictions except that we still require an individual to be employed and have valid information on hours worked, and not to have extreme study duration. Keeping these two restrictions is necessary for our structural approach that requires estimating earnings regressions and focussing on persons who enroll at university with the intention to base their career on the acquired education. These are the ones we hypothesize to respond to the modelled incentives.

because a large body of publicly well-received research has severely questioned the argument typically put forward by the religiously conservative that homosexuality is "unnatural". Also the referendum on novel ways to cope with criminal activity of heroin addicts is an example of science-based changes in political attitudes. Details of these referenda and results are taken from from Année Politique Suisse, a data set described in Linder et al. (2010),[11]. We list and describe this source and the other employed administrative data sources ('Swiss Graduate Survey', 'Swiss Historical Municipality Register', 'Federal Population Census', 'Federal Elections') in the Online Appendix.

To extract the common information contained in the share of yes-votes of the four referenda on gender equality and science-related issues described above, the vote share of left-wing parties, and the share of Catholics, we do a principal component analysis with these six variables (see the Online Appendix). They all load particularly well on a single principal component, with positive scoring coefficients for the vote shares and a negative one for the share of Catholics. Notably, the latter is, in the Swiss context, associated with more conservative attitudes (Altermatt, 1979; Gordon, 2002). We thus interpret the first principal factor as an indicator for 'progressivism' of a municipality.[12]

3.4 Economic Environment

To capture the economic environment, we control for the broader Swiss region a student has resided in when leaving high school. We follow the Federal Statistical Office that distinguishes seven medium sized (so-called *NUTS*-2) regions based on both geographic and economic criteria (see Online Appendix).

In sensitivity analyses we also use information from the 'Federal Population Census' in 1990 on municipality size as well as, for the total population and for females, the employment rate, the sectoral structure (employment shares in manufacturing, business services, agriculture, and construction), or, alternatively, the occupational structure among the high-skilled (employment shares in management occupations, technical occupations, and in a category

[11]Downloadable at `www.swissvotes.ch`, last retrieved on December 19, 2013.

[12]To see whether sociocultural attitudes are constant over time we run a fixed effects regression of the share of Catholics and the five vote shares on municipality and year fixed effects. Reassuringly, the year fixed effects are insignificant, see the Online Appendix.

summarizing health, education, research, and culture occupations). As Switzerland is highly economically integrated within *NUTS*-2 regions, we do not expect these indicators to play a role for the outcomes of interest.

4 Implementation

We now discuss how we implement the three stage estimation procedure, particularly how we distinguish between variables entering stage 2 (earnings) and stage 3 (structural major choice). Tab. 1 provides an overview of the different (sets of) variables used in the empirical analysis and shows for the benchmark model in which stages of the estimation they enter.

— Table 1 about here —

4.1 Variables Affecting Pecuniary Gains

Recall that variables in $\mathbf{x}_i$ are those which affect the earnings capability $y_{ij} = f_j(\mathbf{x}_i, u_{ij})$ for major j of an individual i at some point of the professional career (observed five years after graduating from university). They enter the earnings equations at estimation stage 2. We control for the age (in logs) of an individual five years after graduating from university to capture work experience (variable 'log age'). Some older graduates may have gained work experience prior or during attending university, the latter possibly prolonging their study duration, to the benefit of higher earnings early in the career. We also account for the fact whether an individual has participated in a post-graduate education program for a period of at least six months (variable 'postgraduate education'). We expect individuals who have participated in such a program to earn significantly less early in the career than those who have not because, for a given age, they tend to have shorter work experience at the time of observation. Finally, we include mutually exclusive measures of parental education indicating whether father or mother have tertiary academic education, vocational education, or no such education (those with missing information are allocated to the latter category). In this way we account for a possible intergenerational transmission of cognitive ability. We expect the education of parents to be much less important for success in the labour market within the group of university graduates as compared to the whole population.

4.2 Variables Affecting Monetary and Net Subjective Costs

Identification requires that we find convincing exclusion restrictions. That is, we need variables $\mathbf{z}_i$ that reflect monetary and net subjective costs of major choice but are distinct from variables in $\mathbf{x}_i$ that affect the earnings capability of a student i.

4.2.1 Relative Distance to Next Technical University

With nine university locations in a small country like Switzerland, but only two technical universities that exclusively offer STEM programmes and are located in the high-cost areas Zurich and Lausanne, we hypothesize that the distance from the home municipality to the next technical university relative to the distance to the next cantonal university is a major determinant of the economic cost to study a STEM major.[13] For instance, a high school graduate living at their parents' home in the municipality of Rorschach faces the trade-off between studying, say, economics at the nearby cantonal university in St. Gallen (16.2 km driving distance) or studying a STEM field at ETH Zurich which is six times farther away. Our main identifying assumption is that the *relative* distance to the next technical university does not, at the same time, affect earnings associated with a given major, conditional on the economic environment (as captured by the broader region in which the home municipality is located). We take the log of relative distance to the next technical university to capture that the marginal impact of an additional kilometer on major choice is decreasing with distance.

The unrestricted access of those holding a *Gymnasiale Matura* degree to university programs together with the geographical distribution of technical and cantonal universities in Switzerland provide a unique source of variation in the cost determinants of university majors.

4.2.2 Sociocultural Characteristics

Our preferred specification also holds that variables capturing the sociocultural background of students affect the decision whether to study a STEM field but do not affect the earnings

[13] We assume that all individuals within a home municipality live downtown. This is innocuous since 98% of municipalities have a radius of about 5.7 km or less (see Online Appendix for further municipality characteristics).

gain associated with studying a STEM field, given the economic environment. Most importantly for our main research question, we include in $\mathbf{z}_i$ the indicator for progressive attitudes in the home municipality of a college graduate before going to university derived from principal component analysis involving the share of yes-votes of the four referenda on gender equality and science-related issues described above, the vote share of left-wing parties, and the share of Catholics in a municipality. Although political attitudes and religious denomination may have been related to cognitive skills in the 19th century (Becker and Woessmann, 2009; Boppart et al., 2013, 2014), they are unlikely to affect the contemporaneous earnings differential between studying a STEM and non-STEM field.

Moreover, in Switzerland, the motivation to study a STEM field may depend on the majority language (German, French, Italian) of an individual's home environment. For instance, the two technical universities in Switzerland offer programs in German, French, and English but not in Italian. Further, language may also be perceived as a cultural characteristic that affects tastes. We include dummy variables for French and Italian as majority language, i.e., German as majority language is the left-out category.[14]

4.3 Variables Affecting Both Pecuniary Gains and Costs

All stages of the estimation contain a dummy for females (variable 'female'), a dummy indicating whether the individual first responded in the 2002 survey rather than in 2000, and region dummies indicating the broader Swiss region a student has resided in when leaving high school. Including region fixed effects that characterize the economic environment strengthens our exclusion restriction that the distance to the next technical university relative to the next cantonal university is not correlated with unobserved earnings ability of a student. In sensitivity analyses, we also control for municipality size and further economic characteristics like the employment rate of females and of the total population, the sectoral structure or, alternatively, the occupational structure of a municipality at all stages of the estimation.

[14]We subsume the very few observations from Rhaeto-Romansh speaking municipalities to the category Italian speaking. Inhabitants of these municipalities speak Italian at least as their second language, typically being bilingual. Also notably, graduates from a general high school with mother tongue Italian are typically fluent in German or French.

4.4 Gender Differences

In order to examine whether the sociocultural characteristics contribute to explaining gender differences with respect to choosing a STEM field in university, in an extension of our benchmark estimations in Section 5.2, Section 5.3 allows for interaction effects with gender (except for the variables entering all stages, i.e., region fixed effects and the cohort dummy). In particular, we investigate whether there are gender differences in the role of sociocultural characteristics for major choice. Moreover, we use this specification to check the well-known hypothesis that female students are less motivated by earnings than males by including an interaction effect at stage 3 between the (expected) earnings differential across fields and gender.

5 Results

5.1 Summary Statistics

Tab. 2 provides summary statistics of the variables employed in the baseline estimations. Five years after graduation, the average age of the respondents is 31 years. About (1,438/4,767=) 30% of the graduates studied a STEM field. Among STEM field graduates, 23.7% were woman. About 53% of the respondents participated in post-graduate education (slightly more among non-STEM graduates). About three quarters of the parents have at least vocational education. Gross hourly wages in the main job are somewhat lower for STEM graduates.[15] *Inter alia*, this reflects the extraordinarily high wages in the Swiss financial industry.[16]

The average distance of the home municipality of graduates before entering university to the next technical university (ETH) is 71 km, whereas the average distance to the next cantonal university is about 28 km. While average distances to the next ETH do not differ

[15] Hourly wages are CHF 34.52 (≈ US$ 28.77) vs. CHF 36.66 (≈ US$ 30.55). (Exchange rates are for 2007, the year of earnings information of the second cohort of graduates; see https://data.oecd.org/conversion/exchange-rates.htm).

[16] One may also keep in mind that, under positive externalities of STEM graduates, market earnings in STEM occupations are too low from a social point of view. Switzerland does not provide tax support for business R&D, for instance, that could internalize positive externalities from innovative activity.

between STEM and non-STEM graduates, STEM graduates live on average about 4 km farther away from the next cantonal university. This is an indication that the distance to the next ETH relative to the next cantonal university is an important cost component affecting university major choice. The average relative distance to the next ETH is significantly smaller for STEM than for non-STEM graduates (11.7 vs. 15.8).

We also see significant differences between STEM and non-STEM graduates in their sociocultural environment. For instance, in home municipalities of STEM graduates, the share of yes votes in the referenda on gender equality and stem cell engineering are lower whereas the share of votes in favor of left-wing parties and the share of Catholics are higher, on average.

— Table 2 about here —

5.2 Benchmark Model Estimations

Tab. 3 presents the estimation results of all three stages for the benchmark model. Results from the reduced-form (stage 1) and structural choice estimations (stage 3) are shown in column (1) and (4), respectively. They correspond to average partial effects on the probability to choose a STEM field for one-standard-deviation changes of continuous regressors (the progressivism indicator, log relative distance to the next technical university, log age five years after graduation, and the predicted log wage differential between STEM and non-STEM fields) and to discrete changes in the response probability for dummy regressors. First, an increase in the progressivism indicator by one standard deviation lowers the likelihood to study a STEM field by 2.3 percentage points, according to the structural choice estimation. The effect is similar to the reduced-form estimation and statistically different from zero at the 1%-level. Second, as expected from the summary statistics, women are significantly less likely to study a STEM field. Our structural estimation predicts that the probability for females to graduate in a STEM field is 9.8 percentage points lower than for men. This mirrors the widely-discussed substantial gender differences in university major choice we see in most OECD countries, including Switzerland. Third, and important for our identification strategy, an increase in the log relative distance to the next technical university significantly

reduces the probability of choosing a STEM major. The effect size of 2.2 percentage points (column (4)) is comparable to that of an increase in the progressivism indicator. Fourth, students with a Francophone background are less likely to choose a STEM field than students with a home municipality where German or Italian is the majority language. This result does not only underline the importance to control for language effects in the Swiss context, but also suggests that cultural background may matter for important choices.[17]

— Table 3 about here —

Results of the earnings regressions (stage 2) for STEM and non-STEM fields are given in column (2) and (3) of Tab. 3, respectively. In columns (2) and (3), the partial effect of log age is also scaled to reflect a change by one standard deviation. For the other variables in the earnings regressions, we show the partial effects of unit changes (dummy variables and estimates of correction terms (6)). We find that the estimated coefficients for the correction terms $\hat{\lambda}_{i1}$ and $\hat{\lambda}_{i0}$ are both positive albeit only for non-STEM graduates significantly different from zero when tested individually (at the 1%-level). A joint Wald test suggests that they are significantly different from zero (p-value is 0.006). The positive sign of the coefficient for non-STEM graduates suggests that they are positively selected. Individuals with high unobserved earnings capability for non-STEM fields may select into fields like economics, management or law because of extraordinarily high wages in the (quantitatively important) Swiss financial industry. This demonstrates the importance of accounting for potential selection bias for our general research question using Swiss data.

The coefficients on the female dummy in both earnings regressions are insignificant after conditioning on observed characteristics and the correction terms for self-selection.[18] Moreover, older and more experienced non-STEM graduates have higher earnings. An increase in the age five years after graduation from 32 to 34 would raise earnings by 3.3%.[19] The wage effect of an increase in log age is insignificant, however, for STEM graduates.

[17] Eugster et al. (2017) and Steinhauer (2018) also make use of the cultural diversity in Switzerland to study how majority language affects unemployment and labor force participation of mothers, respectively.

[18] Without correcting for self-selection, we find that female STEM (non-STEM) graduates earn 10% (8.1%) less than their male counterparts (see Online Appendix).

[19] According to Tab. 2, one standard deviation of log age for non-STEM graduates is 6.2%. Evaluated at age 32, these are two years. The average partial effect of log age in column (3) of Tab. 3 is 0.033.

As expected, individuals with post-graduate education earn significantly less five years after graduation, particularly STEM graduates. Interestingly, earnings do not systematically increase with the education level of parents. This is not implausible. The education of parents would certainly matter for earnings in a sample with both graduates and non-graduates but there is not much reason for such an effect when restricting focus on those who attend university. Furthermore, a Wald test rejects the hypothesis that the coefficients on the earnings determinants $\mathbf{x}_i$ are equal between STEM and non-STEM graduates (p-value is 0). Thus, the expected return to STEM fields is statistically different from zero.

The key difference between third stage estimates in column (4) and those in column (1) is that at stage 3 we employ the estimated expected log wage differential across fields (see eq. (7)) rather than the variables that affect earnings capability. A change in the log wage differential by one standard deviation increases, on average, the fraction of STEM major graduates by 4.8 percentage points. The average partial effect is highly significant. Comparing the effect size (of a one-standard-deviation change) to that of the progressivism indicator and relative distance to the next technical university suggests that the role of the sociocultural background for university major choice is substantial compared with standard economic determinants. For instance, the absolute magnitude of the effect of progressivism is about half the effect of earnings and somewhat larger than the effect of relative distance to the next technical university.

To check empirical relevance of our exclusion restrictions, we conduct Wald tests on the coefficient estimates in the reduced form model (see the Online Appendix). Specifically, we test individual and joint significance of the regressors grouped according to the categories in Tab. 1. Supporting our identification strategy, the coefficient on the monetary cost indicator (log relative distance to the next technical university, excluded from stage 2) is significantly different from zero at the 5%-level. The coefficients on the sociocultural variables (progressivism and majority language, excluded from stage 2) are jointly significant at the 1%-level. The same is true for the variables affecting earnings capability (log age, dummies for post-graduate training and parental education, excluded from stage 3). Coefficients on region fixed effects have a p-value of 12% in a joint test.

5.3 Interactions Across Gender

We now examine gender-specific effects of sociocultural characteristics and pecuniary incentives for university major choice. Tab. 4 presents the results analogous to Tab. 3, with effects separated for men and women. They are based on an estimation with gender interactions with coefficients shown in the Online Appendix.

— Table 4 about here —

The quantitative effects of relative distance to the next technical university and majority language are similar to those presented in Tab. 3 for both genders. The coefficients on the corresponding gender interactions are not significantly different from zero. Most interestingly, we see that the effects of progressivism largely differ across gender both in magnitude and statistical significance, by contrast. For men (first panel of Tab. 4), an increase of the progressivism indicator by one standard deviation lowers the probability to study a STEM major by 3.9 percentage points in the structural estimation (column (4)). The average partial effect is significantly different from zero at the 1%-level and similar to the reduced-form estimation (column (1)). For women (second panel), by contrast, not only is the average partial effect of the progressivism indicator statistically not different from zero but also literally zero in magnitude in both the reduced-form and the structural estimation, according to column (1) and (4), respectively.

Average partial effects of pecuniary incentives to study a STEM field are different for men and women, too. For men, the increase in the fraction of STEM major graduates from an increase in the log wage differential by one standard deviation is 7.4 percentage points and highly significant (column (4)). By contrast, the average partial effect of the log wage differential (STEM versus non-STEM) for women is very small and insignificant. It is thus safe to conclude that pecuniary returns to graduating in a STEM field matter considerably less for women than for men. As for the progressivism indicator, the gender interaction effect is highly significant.

With respect to earnings regressions, we see that postgraduate education reduces earnings for both genders, according to columns (2) and (3). Moreover, interestingly, the positive estimates for the coefficient on the selection correction term in the non-STEM equation

($\hat{\gamma}_0 > 0$) suggest that both men and women are positively selected into non-STEM fields. Albeit the coefficient is significantly different from zero for men only, the gender difference is insignificant. Both genders do not seem being particularly selected into STEM fields.

5.4 Sensitivity Analyses

We summarize the results of sensitivity analyses we implemented to examine support for our exclusion restrictions and robustness of size and significance of the effects of interest. All detailed results are relegated to the Online Appendix.

Our first set of sensitivity analyses addresses the concern that the effect of sociocultural characteristics in a municipality on individual major choice could be contaminated by unmeasured heterogeneity in local economic conditions that may shape tastes as well as earnings associated with different study fields. Controlling for the broader (i.e., *NUTS*-2) region a student resided in before entering university may not suffice if economic integration within regions is low. Moreover, it could be the case that municipalities with more progressive attitudes also have higher female employment rates. We thus estimate a richer model that includes at all stages additional variables capturing the economic environment in the home municipality at the time of major choice. We start by including the total employment rate, the female employment rate, and the municipality size. Coefficients of these additional variables are not significantly different from zero in a joint test. Average partial effects on the progressivism indicator, relative distance to the next technical university, and the log wage differential for university major choice at stage 3 remain at significance levels of the benchmark model. An increase by one standard deviation in the progressivism indicator and in relative distance reduces the probability to study STEM by 2.6 and 2.3 percentage points, respectively, comparable to the benchmark model. A similar picture arises when also adding the industry structure (employment shares in agriculture, manufacturing, construction, and business services) or the occupational structure among the high-skilled (employment shares in management occupations, technical occupations, and health, education, research, culture occupations). Results are remarkably robust and the additional economic indicators do not play any role, reflecting high economic integration of *NUTS*-2 regions and Switzerland as a whole.

Second, motivated by Card and Payne (2017), we extend the benchmark model to allow for high school choices of students, thereby controlling for major-specific ability. We include two dummy variables at all stages of the estimations that capture specializations in general high school: in math & sciences and in Latin language. Both specializations are typically also associated with high analytical ability. 57.2% of STEM graduates specialized in math & science in secondary school, compared to only 17.5% of the non-STEM graduates. 38.8% of non-STEM graduates have chosen a specialization containing Latin language, compared to only 23.3% of STEM graduates. Specializing in math & sciences raises the probability to choose a STEM field by 11.9 percentage points, whereas specializing in Latin language raises it by 2.6 percentage points. Importantly, according to the structural equation (stage 3), effect sizes and significance levels of the progressivism indicator, relative distance to the next technical university, and the earnings differential are similar to the benchmark model.

Third, we relax the restriction that the sociocultural characteristics do not affect earnings and allow them to enter the estimated equations in all stages. The progressivism indicator enters insignificantly in the earnings regressions. The point estimates in the structural equation are very similar but less precise compared with the specification that excludes the sociocultural characteristics at stage 2. Overall, this evidence lends support to our assumption that progressivism affects net subjective costs rather than pecuniary gains.

Fourth, we allow the relative distance to the next technical university to enter stage 2. We find that the coefficients on relative distance to the next technical university in the two earnings equations are not significantly different from zero in a joint test. Moreover, the earnings differential between STEM and non-STEM ("return to STEM") is not significantly related to relative distance, suggesting that potentially differential labour demand for graduates by municipality is not important for the return to STEM.[20]

[20]To interpret the results of our third or fourth robustness check as a test of overidentification restrictions it would be necessary to maintain the hypothesis that the exclusion restrictions not under test are valid (i.e. log relative distance in the third robustness check and sociocultural characteristics in the fourth). We leave it to the reader to choose the maintained exclusion restriction.

6 Conclusion

We have examined the role of the sociocultural background of students for choosing a STEM field in university. The motivation to focus on the formation of STEM skills is rooted in their salient role for innovation-led economic growth. We have exploited the large regional variation of sociocultural attitudes combined with a high degree of economic integration in Switzerland. We employ rich survey data on university graduates and complement them with municipality level information from the census as well as nationwide referenda and parliamentary election results. In particular, we characterize a student's home environment with respect to progressive attitudes. The unique opportunity for our research mainly comes from two institutional features in Switzerland, (i) the frequently held national referenda in the Swiss direct democratic system and (ii) the fact that at the time of our study all inhabitants with a general upper secondary education degree were free which field and at which university to study at very moderate tuition fees. We based the empirical identification on a generalized Roy model (Roy, 1951; Heckman and Vytlacil, 2007) which accounts for differences monetary and subjective costs and earnings across majors as well as for selection bias.

Our structural approach allowed us to directly compare the effect of the sociocultural background on major choice with the effects of pecuniary returns (the STEM wage differential) and costs (relative distance to the next technical university). Our findings suggest that sociocultural background is an important driver of major choice of men. In particular, male students from more progressive municipalities are less likely to study a STEM field. The effect of the sociocultural background of males is about half of the effect of the earnings return to STEM and twice as large as the effect of relative cost of studying a STEM field. For females, a progressive background plays no role for the decision to study STEM. Consistent with previous studies, we also find that female students are considerably less motivated by earnings than men.

From a policy point of view, at least for males, our results suggest important challenges for promoting the formation in STEM fields. Motivating students for those fields may require addressing culturally rooted biases against STEM education in progressive environments.

Future research may attempt to differentiate among the non-STEM fields. As this would probably require modelling major choice among more than two alternatives, identification will be an important challenge beyond the scope of the current paper. The differential impact of the sociocultural environment on males and females certainly deserves further attention in future research as well.

References

[1] Akerlof, G.A. and Kranton R.E. (2000). Economics and identity, *Quarterly Journal of Economics* 115, 715-753.

[2] Akerlof, G.A. and Kranton R.E. (2002). Identity and schooling: some lessons for the economics of education, *Journal of Economic Literature* 40, 1167-1201.

[3] Altermatt, U. (1979). Conservatism in Switzerland: a study in antimodernism, *Journal of Contemporary History* 14, 581-610.

[4] Arcidiacono, P. (2004). Ability sorting and returns to college major, *Journal of Econometrics* 121, 343-375.

[5] Arcidiacono, P., Hotz J., and Kang S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals, *Journal of Econometrics* 166, 3-16.

[6] Barbe, P. and Bertail P. (1995). *The Weighted Bootstrap*, Springer, New York.

[7] Becker, S.O. and Woessmann, L. (2009). Was Weber wrong? A human capital theory of protestant economic history, *Quarterly Journal of Economics* 124, 531-596.

[8] Berger, M.C. (1988). Predicted future earnings and choice of college major, *Industrial and Labour Relations Review* 41, 418-29.

[9] Blakemore, A.E. and Low, S.A. (1984). Sex differences in occupational selection: the case of college majors, *Review of Economics and Statistics* 66, 157-63.

[10] Boppart, T., Falkinger, J., Grossmann, V., Woitek U., and Wüthrich, G. (2013). Under which conditions does religion affect educational outcomes?, *Explorations in Economic History* 50, 242-266.

[11] Boppart, T., Falkinger, J., and Grossmann, V. (2014). Protestantism and education: reading (the Bible) and other skills, *Economic Inquiry* 52, 874-895.

[12] Boudarbat, B. (2008). Field of study choice by community college students in Canada, *Economics of Education Review* 27, 79-93.

[13] Boudarbat, B. and Montmarquette C. (2009). Choice of fields of study of university Canadian graduates: The role of gender and their parents' education, *Education Economics* 17, 185-213.

[14] Carrell, S.E., Page, M., and West, J.E. (2010). Sex and science: how professor gender perpetuates the gender gap, *Quarterly Journal of Economics* 125, 1101-1144.

[15] Card, D. and Payne, A.A. (2017). High School choices and the gender gap in STEM, NBER Working Paper No. 23769, Cambridge, MA.

[16] Eidea, E. and Waehrer, G. (1998). The role of the option value of college attendance in college major choice, *Economics of Education Review* 17, 73-82.

[17] Eugster, B., Lalive, R., Steinhauer, A., and Zweimüller, J. (2017). Culture, work attitudes and job search: evidence from the Swiss language border, *Journal of the European Economic Association* 15, 1056-1100.

[18] Federal Statistical Office (2013). Bildungssystem Schweiz – Indikatoren: Abschlüsse und Kompetenzen – Übertrittsquote Maturität-HS, Neuchâtel.

[19] Federal Statistical Office (2015). Bildungssystem Schweiz – Indikatoren: Abschlüsse und Kompetenzen – Maturitätsquote, Neuchâtel.

[20] Fitzenberger, B. and Muehler, G. (2015). Dips and floors in workplace training: using personnel records to estimate gender differences, *Scottish Journal of Political Economy* 62, 400–429.

[21] French, E. and Taber, C. (2011). Identification of models of the labor market, in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 4A, ch. 6, Elsevier Science, Amsterdam, 537-617.

[22] Gordon, B. (2002). *The Swiss Reformation*, Manchester University Press, Manchester, New York.

[23] Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes, *Journal of Economic Perspectives* 20, 23-48.

[24] Heckman, J.J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *Annals of Economic and Social Measurement* 5, 475-492.

[25] Heckman, J.J. (1979). Sample selection bias as a specification error, *Econometrica* 47, 153-161.

[26] Heckman, J.J., Lochner, L.J., and Todd, P.E. (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond, in: E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Vol. 1, ch. 7, Elsevier Science, Amsterdam, 307-458.

[27] Heckman, J.J. and Vytlacil, E.J. (2007). Econometric evaluation of social programs, part I: causal models, structural models and econometric policy evaluation, in: J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, Vol. 6B, ch. 70, Elsevier Science, Amsterdam, 4779-4874.

[28] Humlum, M.K., Kleinjans, K.J., and Nielsen, H.S. (2012). An economic analysis of identity and career choice, *Economic Inquiry* 50, 39-61.

[29] Hunt, J. and Gauthier-Loiselle, M. (2010). How much does immigration boost innovation, *American Economic Journal: Macroeconomics* 2, 31-56.

[30] Linder, W., Bolliger, C., and Rielle, Y. (2010). Handbuch der eidgenössischen Volksabstimmungen 1848-2007, Haupt Verlag, Bern.

[31] Montmarquette, C., Cannings, K., and Mahseredjian, S. (2002). How do young people choose college majors?, *Economics of Education Review* 21, 543-556.

[32] OECD (2004). Education at a glance, Paris.

[33] OECD (2012). Public and private schools: how management and funding relate to their socio-economic profile, Paris.

[34] OECD (2013). Education at a glance, Paris.

[35] OECD (2017). Education at a glance, Paris.

[36] Polachek, S.W. (1978). Sex differences in college major, *Industrial and Labor Relations Review* 31, 498-508.

[37] Roy, A.D. (1951). Some thoughts on the distribution of earnings, *Oxford Economic Papers* 3, 135-146.

[38] Sonnert, G. (2009). Parents who influence their children to become scientists: effects of gender and parental education, *Social Studies of Science* 39, 927-941.

[39] Steinhauer, A. (2013). Working moms, childlessness, and female identity, CEPR Discussion Paper No. 12929, London.

[40] Stinebrickner, R. and Stinebrickner, T.R. (2014). A major in science? Initial beliefs and final outcomes for college major and dropout, *Review of Economic Studies* 81, 426-472.

[41] Tabellini, G. (2008). Institutions and culture, *Journal of the European Economic Association* 6, 255-294.

[42] Wang, X. (2013). Why students choose STEM majors: motivation, high school learning, and postsecondary context of support, *American Educational Research Journal* 50, 1081-1121.

[43] Willis, R.J. and Rosen, S. (1979). Education and self-selection, *Journal of Political Economy* 87, S7-S36.

[44] Winters, J.V. (2014). STEM graduates, human capital externalities, and wages in the U.S., *Regional Science and Urban Economics* 48, 190-198.

[45] Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: identification using an information experiment, *Review of Economic Studies* 82, 791-824.

[46] Wolter, S.C., Diem, A., and Messer, D. (2014). Drop-outs from Swiss universities: an empirical analysis of data on all students between 1975 and 2008, *European Journal of Education* 49, 471-483.

[47] Zafar, B. (2013). College major choice and the gender gap, *Journal of Human Resources* 48, 545-595.

Tables 1-4

Table 1: Benchmark Model and Exclusion Restrictions

	Estimation Stage			
	Red. form	Earnings	Structural	
Variable	Choice		Choice	Interpretation
Progressivism indicator	x		x	Net subjective cost
Female	x	x	x	
Log rel. dist. to techn. univ.	x		x	Monetary cost
Majority language	x		x	Net subjective cost
Log age	x	x		Earnings capability
Post-graduate education	x	x		Earnings capability
Parental education	x	x		Earnings capability
NUTS-2 region fixed effects	x	x	x	Economic environment

Note: The model includes in addition a cohort dummy at all stages.

Table 2: Descriptive Statistics of Core Variables

	STEM (1)	Non-STEM (2)	t-Statistic (3)
Individual level variables			
Female (1=yes, 0=no)	0.237 (0.425)	0.516 (0.500)	-19.709
Second cohort (1=yes, 0=no)	0.451 (0.498)	0.433 (0.496)	1.152
Age	30.89 (1.557)	31.11 (1.959)	-4.209
Log age	3.429 (0.050)	3.436 (0.062)	-3.850
Postgraduate education (1=yes, 0=no)	0.493 (0.500)	0.550 (0.498)	-3.601
Both parents with no/missing educ.	0.062 (0.241)	0.082 (0.274)	-2.532
F: no/missing, M: vocational educ.	0.033 (0.180)	0.042 (0.200)	-1.403
F: no/missing, M: university educ.	0.002 (0.045)	0.004 (0.065)	-1.331
F: vocational, M: no/missing educ.	0.107 (0.309)	0.116 (0.320)	-0.925
Both parents with vocational educ.	0.450 (0.498)	0.365 (0.481)	5.502
F: vocational, M: university educ.	0.015 (0.123)	0.025 (0.155)	-2.176
F: university, M: no/missing educ.	0.049 (0.216)	0.051 (0.219)	-0.262
F: university, M: vocational educ.	0.190 (0.392)	0.195 (0.396)	-0.415
Both parents with university educ.	0.092 (0.289)	0.122 (0.327)	-3.145
Gross hourly wage	34.52 (14.82)	36.66 (13.35)	-4.731
Log gross hourly wage	3.459 (0.414)	3.544 (0.343)	-6.833
Municipality level variables			
Distance to next university (km)	30.82 (30.25)	26.70 (30.37)	4.319
Distance to next ETH (km)	70.49 (49.92)	70.57 (49.48)	-0.050
Relative distance to ETH	11.65 (27.70)	15.81 (30.81)	-4.598
Log relative distance to ETH	1.058 (1.428)	1.342 (1.584)	-6.101
Share in favor of gender equality	0.610 (0.126)	0.645 (0.129)	-8.693
Share in favor of heroin program	0.547 (0.101)	0.545 (0.104)	0.715
Share in favor of stem cell engineering	0.676 (0.097)	0.702 (0.102)	-8.594
Share in favor of gay marriage	0.591 (0.086)	0.596 (0.085)	-2.027
Share of left-wing parties	0.307 (0.125)	0.335 (0.133)	-6.873
Share of Catholics	0.502 (0.242)	0.490 (0.227)	1.609
Majority French (1=yes, 0=no)	0.257 (0.437)	0.374 (0.484)	-8.223
Majority Italian (1=yes, 0=no)	0.061 (0.240)	0.058 (0.234)	0.426
Observations	1,438	3,329	

Source: Author's calculations using Federal Statistical Office of Switzerland, Année Politique Suisse. *Note:* The first two columns show the means and standard deviations (in parentheses) of the variables, the last column the t-statistics of a test of equality of means.

Table 3: Benchmark Estimations

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Average partial effects and standard errors				
Progressivism indicator	-0.024			-0.023
	(0.007)***			(0.007)***
Female	-0.124	-0.029	0.022	-0.098
	(0.008)***	(0.032)	(0.021)	(0.018)***
Log relative distance	-0.018			-0.022
	(0.009)**			(0.009)**
Majority French	-0.038			-0.027
	(0.012)***			(0.013)**
Majority Italian	0.012			0.010
	(0.028)			(0.029)
Log age	-0.058	-0.002	0.033	
	(0.007)***	(0.020)	(0.011)***	
Postgraduate education	-0.030	-0.111	-0.027	
	(0.007)***	(0.013)***	(0.009)***	
F: no/missing, M: vocational educ.	0.005	-0.022	-0.000	
	(0.008)	(0.015)	(0.007)	
F: no/missing, M: university educ.	-0.001	-0.031	0.010	
	(0.007)	(0.007)***	(0.005)**	
F: vocational, M: no/missing educ.	0.013	-0.002	0.005	
	(0.010)	(0.016)	(0.010)	
Both parents with vocational educ.	0.043	-0.009	-0.008	
	(0.012)***	(0.025)	(0.015)	
F: vocational, M: university educ.	-0.004	0.005	-0.006	
	(0.008)	(0.015)	(0.008)	
F: university, M: no/missing educ.	0.014	-0.012	0.002	
	(0.008)*	(0.013)	(0.009)	
F: university, M: vocational educ.	0.021	-0.017	-0.004	
	(0.011)*	(0.019)	(0.012)	
Both parents with university educ.	0.011	-0.044	-0.025	
	(0.010)	(0.017)**	(0.011)**	
Correction term λ		0.074	0.384	
		(0.109)	(0.122)***	
Log wage differential				0.048
				(0.014)***

<continued on next page>

Table 3: Benchmark Estimations <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Model statistics				
Observations	4,767	1,438	3,329	4,767
(Pseudo) R^2	0.09	0.14	0.07	0.08

Source: Author's calculations using Federal Statistical Office of Switzerland, Année Politique Suisse. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressors). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

Table 4: Estimations with Gender Interactions

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Average partial effects for men				
Progressivism indicator	-0.042			-0.039
	(0.011)***			(0.011)***
Log relative distance	-0.021			-0.023
	(0.012)*			(0.013)*
Majority French	-0.036			-0.024
	(0.016)**			(0.016)
Majority Italian	0.014			0.011
	(0.034)			(0.035)
Log age	-0.081	-0.001	0.033	
	(0.011)***	(0.022)	(0.016)**	
Postgraduate education	-0.048	-0.108	-0.030	
	(0.011)***	(0.015)***	(0.013)**	
F: no/missing, M: vocational educ.	0.001	-0.017	-0.000	
	(0.011)	(0.015)	(0.010)	

<continued on next page>

Table 4: Estimations with Gender Interactions <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
F: no/missing, M: university educ.	0.008	-0.038	0.007	
	(0.013)	(0.005)***	(0.009)	
F: vocational, M: no/missing educ.	0.021	0.004	0.006	
	(0.014)	(0.019)	(0.013)	
Both parents with vocational educ.	0.065	-0.009	-0.006	
	(0.017)***	(0.028)	(0.020)	
F: vocational, M: university educ.	0.003	0.001	-0.008	
	(0.011)	(0.016)	(0.010)	
F: university, M: no/missing educ.	0.009	-0.012	0.024	
	(0.012)	(0.015)	(0.010)**	
F: university, M: vocational educ.	0.037	-0.022	0 (0.016)	
	(0.016)**	(0.022)		
Both parents with university educ.	0.018	-0.037	-0.027	
	(0.014)	(0.020)*	(0.014)*	
Correction term λ		0.052	0.291	
		(0.106)	(0.120)**	
Log wage differential				0.074
				(0.023)***
Average partial effects for women				
Progressivism indicator	-0.000			0.000
	(0.009)			(0.009)
Log relative distance	-0.015			-0.022
	(0.009)			(0.009)**
Majority French	-0.035			-0.025
	(0.012)***			(0.013)**
Majority Italian	0.014			0.013
	(0.023)			(0.024)
Log age	-0.028	-0.010	0.019	
	(0.008)***	(0.030)	(0.010)*	
Postgraduate education	-0.009	-0.124	-0.030	
	(0.008)	(0.025)***	(0.009)***	
F: no/missing, M: vocational educ.	0.007	-0.037	0.002	
	(0.009)	(0.032)	(0.010)	

<continued on next page>

Table 4: Estimations with Gender Interactions <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
F: no/missing, M: university educ.	-0.008	-0.010	0.012	
	(0.007)	(0.010)	(0.006)**	
F: vocational, M: no/missing educ.	-0.000	-0.026	0.008	
	(0.012)	(0.033)	(0.015)	
Both parents with vocational educ.	0.014	-0.011	0.003	
	(0.017)	(0.050)	(0.021)	
F: vocational, M: university educ.	-0.015	0.027	-0.005	
	(0.009)	(0.016)*	(0.012)	
F: university, M: no/missing educ.	0.018	-0.018	-0.024	
	(0.010)*	(0.028)	(0.014)*	
F: university, M: vocational educ.	0.001	-0.007	-0.002	
	(0.015)	(0.039)	(0.017)	
Both parents with university educ.	0.002	-0.063	-0.022	
	(0.013)	(0.036)*	(0.015)	
Correction term λ		0.171	0.289	
		(0.163)	(0.184)	
Log wage differential				0.009
				(0.011)

Source: Authors' calculations using Federal Statistical Office of Switzerland, Année Politique Suisse. *Note:* See Tab. 3.

Online Appendix

Sociocultural Background and Choice of STEM Majors at University

Aderonke Osikominu,[♮] Volker Grossmann,[♭] and Marius Osterfeld[♯]

[♮]University of Hohenheim and IAB, CEPR, CESifo, IZA, a.osikominu@uni-hohenheim.de

[♭]University of Fribourg, CESifo, CReAM, IZA, volker.grossmann@unifr.ch

[♯]swissstaffing, marius.osterfeld@swissstaffing.ch

A Full Structural Model

We present the full foundation of our generalized Roy model (Roy, 1951, Heckman and Vytlacil, 2007) for university major choice. We first develop a simple theoretical life-cycle model that features individual university major choice and then derive from it our structural estimation approach outlined in Section 2. This allows us to employ an identification strategy which is inspired by the literature on the (binary) college participation decision. Although we assume linear specifications when estimating the model, nothing in our identification strategy would prevent us to estimate the model semi-parametrically, as will become apparent.

A.1 Life-Cycle Model of University Major Choice

Consider an individual i with time horizon T who chooses university major $j \in \{0, 1\}$ in period 0. For simplicity and following standard arguments, suppose individuals cannot borrow against future income while attending university in period 0. In periods $t \geq 1$, individuals can freely borrow and lend at the time-invariant (world market) interest rate r. We assume that, in period 0, individual i possesses and uses resources a_i for consumption.[1]

Earnings of individual i in period $t \geq 1$ when choosing major j are denoted by Y_{ijt}. That is, the present discounted value of earnings of such an individual reads as $W_{ij} \equiv \sum_{t=1}^{T} Y_{ijt}(1+r)^{-t}$. We assume that earnings observed at some point of the career, y_{ij}, are proportional to W_{ij}, where the proportionality factor is independent of j. The log of earnings

[1]Sources of income in period 0 could be income from a sideline job during study, scholarships, transfers by parents or prior savings. Individuals may save in period 0 but provided that future income is sufficiently higher than available resources in period 0, which we implicitly assume, they optimally choose zero savings.

y_{ij} is given by

$$\ln y_{ij} = m_j^y(\mathbf{x}_i) + u_{ij}\,,\ j \in \{0,1\}\,, \tag{A.1}$$

where m_j^y is a major specific function of observable characteristics $\mathbf{x}_i$ of individual i and u_{ij} is an individual- and major-specific ability component that is unobservable for an econometrician. $m_j^y(\mathbf{x}_i)$ corresponds to the conditional expectation function of log earnings, $\mathrm{E}\left[\ln y_{ij} \,|\, \mathbf{x}_i\right]$. Anticipating the next paragraph, we introduce another set of observed individual characteristics distinct from $\mathbf{x}_i$, $\mathbf{z}_i$, and strengthen the argument to $m_j^y(\mathbf{x}_i) = \mathrm{E}\left[\ln y_{ij} \,|\, \mathbf{x}_i, \mathbf{z}_i\right]$, which means that $\mathbf{z}_i$ has no influence on the conditional expectation of log earnings.

Individuals draw utility from their consumption stream. Importantly, each individual i also draws utility h_{ij} from studying (and graduating from) major j. Total intertemporal utility of individual i when choosing field j is given by

$$U_{ij} = \sum_{t=0}^{T} \rho^t \ln c_{ijt} + h_{ij}\,, \tag{A.2}$$

where $\rho \in (0,1)$ is the discount rate and c_{ijt} denotes the level of material consumption in t of an individual i that has chosen major j. We assume that h_{ij} depends on observable characteristics $\mathbf{z}_i$ of individual i and on an unobservable individual- and major-specific taste component τ_{ij}. More specifically, we specify h_{ij} as

$$h_{ij} = m_j^h(\mathbf{z}_i) + \tau_{ij}\,, \tag{A.3}$$

where $m_j^h(\mathbf{z}_i)$ is the conditional expectation function of the immediate utility associated with major j, $\mathrm{E}\left[h_{ij} \,|\, \mathbf{x}_i, \mathbf{z}_i\right]$. Thus, individual characteristics in $\mathbf{z}_i$ influence tastes for a major, whereas variables in $\mathbf{x}_i$ influence earnings capability. Eq. (A.1) and (A.3) may also depend on common observable characteristics, but we keep the conditioning on them implicit to simplify the notation.

After choosing a major j, individual i solves the following maximization problem that smooths consumption over time:

$$\max_{\{c_{ijt}\}_{t=1}^{T}} \sum_{t=1}^{T} \rho^t \ln c_{ijt} \text{ s.t. } \sum_{t=1}^{T} \frac{c_{ijt}}{(1+r)^t} = W_{ij}. \tag{A.4}$$

This leads to the well-known Euler equation

$$c_{ijt+1} = (1+r)\rho c_{ijt}\,,\ t = 1, 2, \ldots, T-1\,. \tag{A.5}$$

Using (A.5) in (A.2) and observing that $c_{ij0} = a_i$ and $\sum_{t=1}^{T} \rho^t = \frac{\rho(1-\rho^T)}{1-\rho} \equiv \alpha$, indirect life-time utility of individual i conditional on major choice j reads as

$$U_{ij} = \alpha \ln c_{ij1} + h_{ij} + \ln a_i + const. \tag{A.6}$$

Combining (A.5) with the intertemporal budget constraint in (A.4), we obtain the (optimal) level of consumption of individual i under college major choice j in the first working period ($t = 1$):

$$c_{ij1} = \frac{(1-\rho)(1+r)}{(1-\rho^T)} W_{ij} \,. \tag{A.7}$$

Substituting both (A.3) and (A.7) for $j \in \{0, 1\}$ into (A.6) and using that y_{ij} is proportional to W_{ij} in the same way for $j \in \{0, 1\}$, we find that the difference in utility between studying a STEM major (alternative 1) and a non-STEM major (alternative 0) is given by

$$\Delta U_i = \alpha[\ln y_{i1} - \ln y_{i0}] + m_1^h(\mathbf{z}_i) - m_0^h(\mathbf{z}_i) + \nu_i \,, \tag{A.8}$$

where $\nu_i \equiv \tau_{i1} - \tau_{i0}$. As an implication of the logarithmic form of instantaneous utility from material consumption, [2] the expression for the utility difference ΔU_i does not depend on the interest rate, r. Moreover, because income (and consumption) while attending university, a_i, is independent of the major choice, it cancels out.

Substituting the expression for $\ln y_{ij}$ in (A.1) for $j \in \{0, 1\}$ into (A.8) and defining

$$m(\mathbf{x}_i, \mathbf{z}_i) \equiv \alpha[m_1^y(\mathbf{x}_i) - m_0^y(\mathbf{x}_i)] + m_1^h(\mathbf{z}_i) - m_0^h(\mathbf{z}_i) \tag{A.9}$$

we can thus write

$$\Delta U_i = m(\mathbf{x}_i, \mathbf{z}_i) - \varepsilon_i \,, \tag{A.10}$$

where $-\varepsilon_i \equiv \alpha(u_{i1} - u_{i0}) + \nu_i$. Clearly, individual i prefers alternative 1 (STEM major) over alternative 0 if $\Delta U_i \geq 0$, i.e., if $m(\mathbf{x}_i, \mathbf{z}_i) \geq \varepsilon_i$.

Willis and Rosen (1979) derived expressions for ΔU_i which are similar to (A.8) and (A.10), albeit in a rather different set up. Whereas they analyzed the choice of college participation, our framework captures the choice of the university major given that an individual goes to university.[3]

[2] The assumption means that the coefficient of relative risk aversion is unity, consistent with empirical evidence (e.g., Chetty, 2006).

[3] Willis and Rosen (1979) did not consider consumption smoothing and utility from attending university. The (log) linear forms in their model are implied by a first-order Taylor-approximation.

A.2 Identification

Ideally, we would like to estimate the probability that a STEM field is chosen as a function of the monetary return, $(\ln y_{i1} - \ln y_{i0})$, and taste characteristics, $\mathbf{z}_i$, of individual i. However, an econometrician observes earnings of an individual only for the major which has actually been chosen. The individual pecuniary gain from studying a STEM field, $(\ln y_{i1} - \ln y_{i0})$, is not observable for any i. Recalling (A.9), a closer look at (A.10) suggests that we may identify the structural choice model without knowledge of the individual pecuniary gain, if we can compute the expected pecuniary gain associated with characteristics $\mathbf{x}_i$, $[m_1^y(\mathbf{x}_i) - m_0^y(\mathbf{x}_i)]$, for every individual i. Define $s_i \equiv \mathbb{1}(\Delta U_i \geq 0)$ as a dummy that takes on value one if choosing STEM educations yields higher utility for individual i. Using (A.8), (A.9), and (A.10), we get

$$\begin{aligned} &\Pr\{s_i = 1 \,|\, \mathbf{x}_i, \mathbf{z}_i\} = \\ &\quad \Pr\{\alpha[\ln y_{i1} - \ln y_{i0}] + m_1^h(\mathbf{z}_i) - m_0^h(\mathbf{z}_i) + \nu_i \geq 0 \,|\, \mathbf{x}_i, \mathbf{z}_i\} = \\ &\quad \Pr\{\alpha[m_1^y(\mathbf{x}_i) - m_0^y(\mathbf{x}_i)] + m_1^h(\mathbf{z}_i) - m_0^h(\mathbf{z}_i) \geq \varepsilon_i \,|\, \mathbf{x}_i, \mathbf{z}_i\}\,. \end{aligned} \tag{A.11}$$

Thus, a key ingredient for semi-parametric identification of the structural choice model (A.11) is nonparametric identification of $m_j^y(\mathbf{x}_i) = \mathrm{E}[\ln y_{ij} \,|\, \mathbf{x}_i]$, $j \in \{0, 1\}$, free of selection bias.

The identification analysis proceeds in three steps[4]. In the first step, we use (A.10) to set up the reduced form for the probability of choosing a STEM major. Defining $\mathbf{w}_i \equiv (1, \mathbf{x}_i, \mathbf{z}_i)$ we express the reduced form choice probability conditional on the combined set of regressors $\mathbf{w}_i$ as

$$\Pr(s_i = 1 \,|\, \mathbf{w}_i) = \Pr\big(m(\mathbf{w}_i) \geq \varepsilon_i \,|\, \mathbf{w}_i\big) = F_\varepsilon\big(m(\mathbf{w}_i)\big)\,, \tag{A.12}$$

where F_ε is the c.d.f of ε_i. The index function and the distribution of ε_i are identified (up to a monotonic transformation) based on standard arguments for nonparametric identification of discrete choice models (Matzkin, 1992). In particular, this would require a full independence assumption such as

$$\mathbf{x}_i, \mathbf{z}_i \perp\!\!\!\perp u_{ij}, \nu_i\,,\ j \in \{0, 1\}\,, \tag{A.13}$$

where $\perp\!\!\!\perp$ denotes statistical independence. Importantly, the independence assumption is *not* necessary if identification of the distributions of ε_i and u_{ij} is not of interest and the reduced form choice (A.12) is modeled as a nonparametric linear probability model (Das

[4]See Heckman and Vytlacil (2007, Appendix B) and French and Taber (2011, Sections 3-4), for a detailed formal account of nonparametric identification of the generalized Roy model. Das, Newey, and Vella (2003) consider nonparametric identification of the closely related sample selection model under a weaker mean independence assumption.

et al., 2003).[5] In the second step, we can identify the components of $\mathrm{E}[\ln y_{i1} \,|\, \mathbf{w}_i, s_i = 1]$ and $\mathrm{E}[\ln y_{i0} \,|\, \mathbf{w}_i, s_i = 0]$ exploiting knowledge of $m(\mathbf{w}_i)$ from the first step. To see how consider the observable conditional expectation of log earnings for STEM fields for those who graduated from a STEM field ($s_i = 1$),

$$\mathrm{E}[\ln y_{i1} \,|\, \mathbf{w}_i, s_i = 1] = m_1^y(\mathbf{x}_i) + \mathrm{E}[u_{i1} \,|\, \mathbf{w}_i, s_i = 1]\,. \tag{A.14}$$

Eq. (A.1) implies that $\mathrm{E}[u_{i1} \,|\, \mathbf{w}_i] = \mathrm{E}[u_{i1} \,|\, \mathbf{x}_i] = 0$. However, $\mathrm{E}[u_{i1} \,|\, \mathbf{w}_i, s_i = 1]$ is in general different from zero, because the expectation is taken only in the part of the population for which $s_i = 1$. Under (A.13) or the weaker mean independence condition

$$\mathrm{E}[u_{ij} \,|\, \mathbf{w}_i, \varepsilon_i] = \mathrm{E}[u_{ij} \,|\, \varepsilon_i]\,,\ j \in \{0, 1\}\,, \tag{A.15}$$

the conditional expectation of u_{i1} in the subpopulation with $s_i = 1$ is a function of $m(\mathbf{w}_i)$ only:

$$\begin{aligned} \mathrm{E}[u_{i1} \,|\, \mathbf{w}_i, s_i = 1] &= \mathrm{E}[u_{i1} \,|\, \mathbf{w}_i, m(\mathbf{w}_i) \geq \varepsilon_i] \\ &= \mathrm{E}\big[\mathrm{E}\left[u_{i1} \,|\, \mathbf{w}_i, \varepsilon_i\right] \,|\, \mathbf{w}_i, m(\mathbf{w}_i) \geq \varepsilon_i\big] \\ &= \mathrm{E}\big[\mathrm{E}\left[u_{i1} \,|\, \varepsilon_i\right] \,|\, m(\mathbf{w}_i) \geq \varepsilon_i\big] \equiv q_1\big(m(\mathbf{w}_i)\big)\,, \end{aligned} \tag{A.16}$$

where the third equality follows either from (A.13) or (A.15) together with the fact that ε_i is mean independent of $\mathbf{w}_i$[6]. Thus, we have the following additive structure for the conditional expectation of log earnings for STEM fields under self-selection.[7]

$$\mathrm{E}[\ln y_{i1} \,|\, \mathbf{w}_i, s_i = 1] = m_1^y(\mathbf{x}_i) + q_1\big(m(\mathbf{w}_i)\big)\,. \tag{A.17}$$

Its components are identified nonparametrically, if $\mathbf{x}_i$ and $m(\mathbf{w}_i)$ can be varied independently from each other. This is the case if $m(\mathbf{w}_i)$ is a nontrivial function of $\mathbf{z}_i$. Put formally, it is required that

$$\mathrm{Supp}\big(m(\mathbf{w}_i), \mathbf{x}_i\big) = \mathrm{Supp}\big(m(\mathbf{w}_i)\big) \times \mathrm{Supp}(\mathbf{x}_i)\,, \tag{A.18}$$

where $\mathrm{Supp}(\cdot)$ denotes the support. Moreover, the joint distribution of (u_{ij}, ε_i) can be identified nonparametrically, if we rely on the full independence assumption (A.13). Finally, identification of $[m_1^h(\mathbf{z}_i) - m_0^h(\mathbf{z}_i)]$ in the structural choice model (A.10) follows immediately

[5] It is always possible to express s_i as $s_i = \mathrm{E}\left[s_i \,|\, \mathbf{w}_i\right] + \varepsilon_i$ and $\mathrm{E}\left[\varepsilon_i \,|\, \mathbf{w}_i\right] = 0$ by definition. Further, define $\mathrm{E}\left[s_i \,|\, \mathbf{w}_i\right] \equiv m(\mathbf{w}_i)$ and we obtain (A.12) when F_ε is equal to the identity function.

[6] See also Subsection A.3 and Das et al. (2003) for further details. Heckman (1976) developed a similar expression in the case of joint normality of (u_{ij}, ε_i).

[7] An analogous expression holds for the conditional expectation of log earnings for non-STEM fields.

given identification of the expected pecuniary gain $[m_1^y(\mathbf{x}_i) - m_0^y(\mathbf{x}_i)]$ in step 2.

From the identification analysis it is clear that robust estimation requires two sets of exogenous regressors, $\mathbf{x}_i$ and $\mathbf{z}_i$, that contain distinct variables.[8] Here the vector $\mathbf{x}_i$ captures regressors that reflect abilities affecting earnings capability, see eq. (A.1), while $\mathbf{z}_i$ subsumes regressors that measure sociocultural and economic factors reflecting monetary and net subjective costs of the majors, see eq. (A.3). Importantly, we measure the sociocultural indicators at the municipality level rather than at the individual level. These sociocultural indicators correspond to the mean of the underlying individual sociocultural attitudes evaluated at the municipality level. Therefore, they are exogenous to the individual level unobserved components (u_{ij}, ν_i) by construction.[9]

Having multiple variables in $\mathbf{z}_i$ distinct from $\mathbf{x}_i$ is useful to extend the range of variation in the reduced form choice index $m(\mathbf{w}_i)$ given $\mathbf{x}_i$ and to enhance precision of the estimates in stage 2 and 3.[10]

A.3 Estimation

Since we observe earnings of an individual only for the major actually chosen we have to estimate the conditional expectation functions of earnings associated with STEM and non-STEM fields by correcting for selection bias and then plug the estimates for $[m_1^y(\mathbf{x}_i) - m_0^y(\mathbf{x}_i)]$ into the structural choice equation, i.e. eq. (A.11). For this purpose, we use flexible parametric specifications. From a practical perspective, a flexible parametric specification that exploits also exclusion restrictions seems a good compromise between the conceptual goal of nonparametric identification and empirical tractability.

In a first step, we estimate the reduced form choice equation (A.12). As outlined in Section 2, we implement eq. (A.12) as a probit model with a linear index function. Moreover, for applying the strategy at stage 2 outlined in Section 2, we set $m_j^y(\mathbf{x}_i) = \beta_{0j} + \mathbf{x}_i\boldsymbol{\beta}_j$ to obtain (2). From (2), we find by using (1) and (3) that the expected log earnings associated

[8] The conditional expectation functions, $m_j^y(\cdot)$ and $m_j^h(\cdot)$, $j \in \{0, 1\}$, may also depend on common variables. To simplify the notation we keep the conditioning on these variables implicit.

[9] In a regression sense, the municipality level sociocultural indicators correspond to the fit from the nonparametric population regression of the underlying individual level attitudes on municipality dummies. The municipality level sociocultural indicators are therefore orthogonal to individual level variation in unobserved tastes and abilities that is part of the nonparametric regression residual.

[10] In this respect, the structural approach adopted here differs from reduced form instrumental variable estimations that aim at recovering the average effect of a treatment choice for those who change their choice in response to the change in the instrument, i.e. the local average treatment effect (LATE). Interpretability of the LATE hinges on the possibility to link the instrumental variation to an interesting policy experiment, see Heckman (2010) for a comparison of the different approaches. Thus, unlike in structural approaches, it is often preferable to use just a single instrument in reduced form instrumental variable estimations.

with alternative 1 (STEM field) conditional on self-selecting to that field are given as:

$$\begin{aligned}
\mathrm{E}(\ln y_{i1} \,|\, \mathbf{w}_i, s_i = 1) &= \mathrm{E}\left[\,\mathrm{E}(\ln y_{i1} \,|\, \mathbf{w}_i, \varepsilon_i) \,|\, \mathbf{w}_i, s_i = 1\right] \\
&= \beta_{01} + \mathbf{x}_i\boldsymbol{\beta}_1 + \gamma_1 \mathrm{E}(\varepsilon_i \,|\, \mathbf{w}_i, s_i = 1) \\
&= \beta_{01} + \mathbf{x}_i\boldsymbol{\beta}_1 + \gamma_1 \mathrm{E}(\varepsilon_i \,|\, \mathbf{w}_i\boldsymbol{\pi} \geq \varepsilon_i) \\
&= \beta_{01} + \mathbf{x}_i\boldsymbol{\beta}_1 + \gamma_1 \lambda_{i1}\,, \qquad \text{(A.19)}
\end{aligned}$$

and, analogously,

$$\mathrm{E}(\ln y_{i0} \,|\, \mathbf{w}_i, s_i = 0) = \beta_{00} + \mathbf{x}_i\boldsymbol{\beta}_0 + \gamma_0 \lambda_{i0}, \qquad \text{(A.20)}$$

with definitions of λ_{i0} and λ_{i1} as in (6) in the main text. These expressions confirm that we can estimate wage regressions (4) and (5) in the main text, respectively.

In the final stage 3, we estimate the structural choice model (A.11) with a probit regression as

$$\Pr\left\{\delta_0 + \alpha[\hat{\beta}_{01} - \hat{\beta}_{00} + \mathbf{x}_i(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_0)] + \mathbf{z}_i\boldsymbol{\delta} \geq \varepsilon_i \,|\, \mathbf{x}_i, \mathbf{z}_i\right\}, \qquad \text{(A.21)}$$

where we replace $m_1^y(\mathbf{x}_i) - m_0^y(\mathbf{x}_i)$ with the estimated value from a linear regression and also specify $m_1^h(\mathbf{z}_i) - m_0^h(\mathbf{z}_i)$ linearly.

B Details on Data Sources

- The 'Swiss Graduate Survey' (*Absolventenstudie*) of the Swiss Federal Statistical Office (*Bundesamt für Statistik*) is a biennial survey of the population of students who graduate from tertiary academic education in Switzerland (Federal Statistical Office, 2008, 2009, 2012a). We use data on the cohorts who graduated in 2000 and 2002 and completed the questionnaires one year and five years after graduation. [11]

- The 'Swiss Historical Municipality Register' (*Historisiertes Gemeindeverzeichnis der Schweiz*) records all municipality changes since 1960 (Federal Statistical Office, 2014). We use it to harmonize the municipality codes across data sources to the classification valid on December 31, 2010. We obtain additional information on geographical classifications and region types from Federal Statistical Office (2010, 2011).

- The 'Federal Population Census' (*Eidgenössische Volkszählung*) is a mandatory survey of the entire resident population that takes place every 10 years (Federal Statistical

[11] We employ probability weights provided by the Federal Statistical Office to account for potentially selective nonresponse. The weights are constructed from the 'Swiss Information System of Universities' that collects information on the distribution of graduates across different universities, majors and cantons of origin (see Federal Statistical Office, 2012a).

Office, 1990, 1996). Every household in Switzerland receives a household specific questionnaire and individual specific questionnaires for each person living in the household. The census collects information on the demographic, economic, social and cultural structure of Switzerland and its development over time. We use the census of the year 1990 to construct the share of Catholics, indicators for the majority language, female and total employment rates, as well as occupation and industry shares of a municipality.

- We use municipality level information on the results of the 'Federal Elections' (*Nationalratswahlen*) of the year 1995 (Federal Statistical Office, 1995) to calculate the share of left-wing parties in each municipality. We classify the following parties as left wing: SP, PdA, Sol., FGA, and GPS.

- The municipality level data on content and results of the nationwide referenda is provided by Année Politique Suisse ('Datensatz der eidgenössischen Volksabstimmungen ab 1848, Institut für Politikwissenschaft, Bern') on the web page `www.swissvotes.ch`; see Linder et al. (2010) for a description. We use it to compute the share of yes-votes in the following four referenda that were particularly salient in the public debate:[12] on introducing equal rights of men and women in the constitution (referendum held in 1981), on providing addicts with medical prescriptions of heroin (1999), on the regulation of stem cell research (2004), and on the civil union of homosexual couples (2005). Details are provided in Section E.

- With the help of Google Maps, we compiled a data base recording the distances between the downtown of the municipalities that existed in 2010 and the nine locations of the cantonal and federal universities. The distances correspond to the driving distances when travelling by car.

C Sample Restrictions on Individual-Level Variables

This section first addresses the concern that our goal to focus on typical careers and to omit observations with implausible values in key variables may lead to selective sample restrictions according to field of study or gender. Second, we show that the results remain fairly robust when we lift most of the sample restrictions.

[12] In the period 1980-2005, there were additional referenda on similar topics than the ones included. In 1985 and in 2000 there were two further referenda on gender equality. In the late 1990s, there were four additional referenda on topics related to science issues (genetic engineering, transplantation medicine) and drug policy. Statistically, these additional referenda capture similar variation than the ones we have retained.

Tab. C.1 reveals that we exclude those who are nonemployed (9.4 percent of the original sample) or work less than 20 percent of the fulltime amount despite being employed (0.6 percent). With only about two percent of tertiary educated classified as unemployed in OECD statistics, we assume that those who are nonemployed are in general out of the labor force and voluntarily not working. According to Tab. C.2, these are somewhat more females than males (11.5 vs. 8.7 percent), while the difference between STEM and non-STEM graduates is not statistically significant (Tab. C.1).

We also exclude those who did not report hours worked, report extreme earnings relative to the reported hours worked, took less than eight or more than 18 semesters to complete their degree, or were older than 32 years when graduating. Extreme earnings values mean that we exclude those reporting monthly earnings below CHF 1,000, above CHF 50,000 (one observation), and those who report monthly earnings higher than CHF 12,546 (about twice the median) despite working less than 50 percent of the fulltime hours. These restrictions seem reasonable by noting that the first percentile of reported monthly earnings is CHF 1,692 and the 99th percentile CHF 17,692. Tab. C.1 and C.2 suggest that there is little difference between STEM and non-STEM graduates and between males and females of those excluded because of extreme earning values, respectively.

Note that the percentages of individuals that fall under single selection criteria do not add up to the total 25.7 percent of individuals dropped from the original sample as selection criteria may overlap. For instance, some of those with earnings less than CHF 1,000 might also report working less than 20 percent of the fulltime amount. In total, the sample selection is somewhat asymmetric for STEM vs. non-STEM graduates (21.7 vs. 27.1 percent of observations are dropped). However, this difference is primarily due to exclusion of those who were 33 years or older when graduating (3.5 percent STEM graduates vs. 9.3 percent non-STEM graduates, according to Tab. C.1).

Table C.1: Selectivity of Sample Restrictions with Respect to Study Major

	Total (1)	STEM (2)	Non-STEM (3)	t-Stat. (4)
Initial number of observations	6,427	1,809	4,618	
Nonemployed	0.094 (0.291)	0.085 (0.280)	0.097 (0.295)	-1.418
Employed < 20% of full time	0.006 (0.079)	0.006 (0.075)	0.006 (0.080)	-0.326
Hours worked missing	0.066 (0.249)	0.071 (0.257)	0.065 (0.246)	0.863
Extreme earnings	0.064 (0.245)	0.063 (0.244)	0.065 (0.246)	-0.167
Part-time emp. with ext. earn.	0.010 (0.099)	0.009 (0.095)	0.010 (0.101)	-0.417
Extreme study duration	0.064 (0.245)	0.044 (0.206)	0.071 (0.258)	-4.402
Extreme age	0.078 (0.268)	0.035 (0.185)	0.093 (0.291)	-9.508
Total share of obs. dropped	0.257 (0.437)	0.217 (0.412)	0.271 (0.445)	-4.668
Final number of observations	4,767	1,438	3,329	

Source: Graduate Survey, own calculations. *Note:* The first three columns show the means and standard deviations (in parentheses) of the variables, the last column the t-statistics of a test of equality of means.

Table C.2: Selectivity of Sample Restrictions with Respect to Gender

	Total (1)	Male (2)	Female (3)	t-Stat. (4)
Initial number of observations	6,427	3,359	3,068	
Nonemployed	0.094 (0.291)	0.082 (0.275)	0.107 (0.310)	-3.598
Employed < 20% of full time	0.006 (0.079)	0.005 (0.067)	0.008 (0.090)	-2.140
Hours worked missing	0.066 (0.249)	0.064 (0.244)	0.070 (0.255)	-1.032
Extreme earnings	0.064 (0.245)	0.060 (0.237)	0.069 (0.254)	-1.588
Part-time emp. with ext. earn.	0.010 (0.099)	0.006 (0.079)	0.014 (0.118)	-3.882
Extreme study duration	0.064 (0.245)	0.062 (0.242)	0.067 (0.249)	-0.676
Extreme age	0.078 (0.268)	0.063 (0.243)	0.096 (0.295)	-5.301
Total share of obs. dropped	0.257 (0.437)	0.235 (0.424)	0.282 (0.450)	-4.333
Final number of observations	4,767	2,572	2,195	

Source: Graduate Survey, own calculations. *Note:* The first three columns show the means and standard deviations (in parentheses) of the variables, the last column the t-statistics of a test of equality of means.

Finally, Tab. C.3 shows the results obtained when we lift all sample restrictions except two. First, we maintain the restriction that individuals are employed and have nonmissing information on hours worked because in our earnings regressions the hourly wage is the dependent variable. Second, we continue to exclude persons with extreme study duration who are likely to just exploit the free university access without studying seriously. The sample size increases from 4,767 in the baseline estimations (Tab. 3 of the main text) to 5,068 observations. Comparing the results with Tab. 3 of the main text, we see that, in the structural STEM choice regression (stage 3), the average partial effects of the progressivism indicator and the log wage differential remain nearly unchanged whereas that of the log relative distance drops slightly both in size and significance.

Table C.3: Benchmark Specification with Less Restrictive Sample

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Average partial effects and standard errors				
Progressivism indicator	-0.024			-0.023
	(0.009)***			(0.009)***
Female	-0.120	-0.022	0.016	-0.102
	(0.007)***	(0.034)	(0.020)	(0.023)***
Log relative distance	-0.013			-0.019
	(0.010)			(0.010)*
Majority French	-0.038			-0.026
	(0.014)***			(0.014)*
Majority Italian	0.011			0.012
	(0.028)			(0.030)
Log age	-0.072	0.029	0.055	
	(0.009)***	(0.025)	(0.011)***	
Postgraduate education	-0.029	-0.101	-0.028	
	(0.006)***	(0.014)***	(0.008)***	
F: no/missing, M: vocational educ.	0.001	-0.020	0.004	
	(0.008)	(0.016)	(0.008)	
F: no/missing, M: university educ.	-0.003	-0.031	0.011	
	(0.007)	(0.007)***	(0.006)*	
F: vocational, M: no/missing educ.	0.009	0.001	0.008	
	(0.010)	(0.016)	(0.011)	

<continued on next page>

Table C.3: Benchmark Estimations <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Both parents with vocational educ.	0.034	-0.007	-0.004	
	(0.013)***	(0.024)	(0.016)	
F: vocational, M: university educ.	-0.006	0.007	-0.005	
	(0.008)	(0.016)	(0.008)	
F: university, M: no/missing educ.	0.010	-0.009	0.004	
	(0.009)	(0.013)	(0.010)	
F: university, M: vocational educ.	0.014	-0.014	-0.001	
	(0.011)	(0.019)	(0.012)	
Both parents with university educ.	0.007	-0.041	-0.022	
	(0.011)	(0.018)**	(0.010)**	
Correction term λ		0.096	0.350	
		(0.113)	(0.120)***	
Log wage differential				0.044
				(0.019)**
Model statistics				
Observations	5,068	1,481	3,587	5,068
(Pseudo) R^2	0.09	0.13	0.08	0.08

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressors). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

D Characteristics of Geographical Units in the Data Set

D.1 Region Fixed Effects

As outlined in Section 3.4, to support our identification strategy, we include region fixed effects at all stages of the estimation, distinguishing the seven so-called *NUTS*-2 regions as defined by the Federal Statistical Office:

- Region Lake Geneva (cantons of Geneva, Vaud, Valais)
- Espace Mittelland (cantons of Berne, Fribourg, Jura, Neuchâtel, Solothurn)
- North-Western Switzerland (cantons of Aargau, Basel-Stadt, Basel-Land)
- Canton of Zurich
- Eastern Switzerland (cantons of Appenzell Ausserrhoden, Appenzell Innerrhoden, Glarus, Grisons, St. Gallen, Schaffhausen, Thurgau)
- Central Switzerland (cantons of Lucerne, Nidwalden, Obwalden, Schwyz, Uri, Zug)
- Canton of Ticino.

The five larger *NUTS*-2 regions have between one and 1.7 million inhabitants whereas Central Switzerland and Ticino have only 750 thousand and 330 thousand inhabitants, respectively. Total population size in Switzerland in the year 2010 is 7.87 million inhabitants (Federal Statistical Office, 2012b)

D.2 Descriptive Statistics on Municipalities

We next provide summary statistics on the geographical characteristics of municipalities.[13] Tab. D.1 shows that the median population share of municipalities is merely 0.015 percent (1,218 inhabitants), the figure for the 95th percentile is 0.137 percent (10,803 inhabitants). Even the by far most populous municipality, Zurich, is inhabited by less than 5 percent of the total population in Switzerland. The median area is only 7.7 km^2 (3.0 mi^2), that of the 95th percentile is 59.1 km^2 (22.8 mi^2).

Table D.1: Characteristics of Municipalities

	Mean	St. Dev.	Min.	P05	Median	P95	Max.
Population share (%)	0.040	0.138	0.0002	0.002	0.015	0.137	4.738
Area (km^2)	16.0	27.5	0.3	1.8	7.7	59.1	430.2

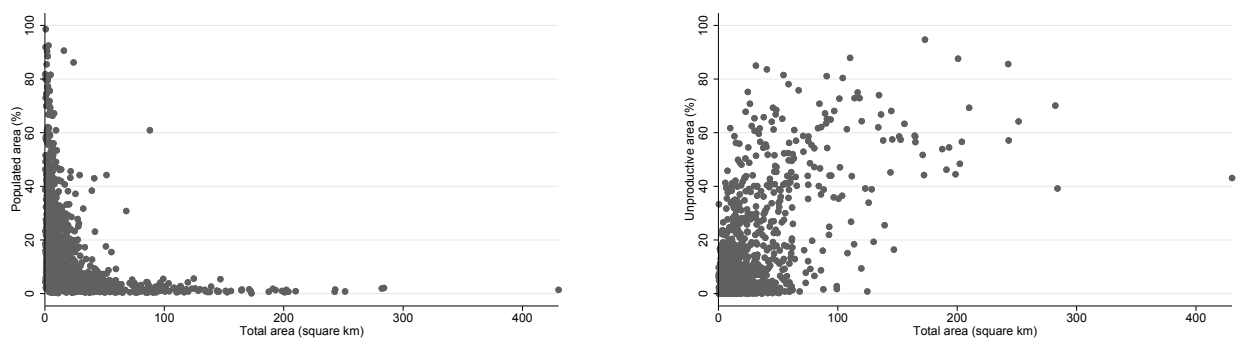
Source: Federal Statistical Office (2012b); own calculations. *Note:* Calculations are based on 2,495 municipalities.

Further support for measuring distances from downtowns is provided in Fig. D.1. Panel (a) shows the relation between the size of a municipality and the fraction of the area populated within a municipality. In those municipalities larger than Zurich (with an area of 88 km^2, i.e. a radius of only 5.3 km = 3.3 mi) only five percent or less of the land area is inhabited (less than two percent on average). Panel (b) shows that there is a positive relation between land size of a municipality and the fraction of unproductive land (i.e. mountainous

[13] Not all information is available for all municipalities. Thus, the number of municipalities in the following tables varies somewhat.

areas that are not used for industry, agriculture or forestry). Together with panel (a), we can thus conclude that most people in municipalities with land area larger than Zurich live downtown. The other municipalities are small enough to reasonably approximate distances of an individual to the next technical or cantonal university by using distances from downtowns.

Figure D.1: Land Use of Municipalities (in %)



(a) Populated

(b) Unproductive

Source: Federal Statistical Office (2012b); own calculations. *Note:* Calculations are based on 2,495 municipalities.

E Background Evidence on Sociocultural Indicators

This section first provides summary statistics of the sociocultural indicators employed to construct our 'progressivism' indicator as first principal component: the share of yes-votes of four referenda on gender equality and science-related issues, the vote share of left-wing parties, and share of Catholics in a graduate's home municipality as well as background information on the employed referenda. We then show detailed evidence from the principal component analysis. We finally demonstrate that these sociocultural indicators are stable over time.

E.1 Summary Statistics and Background Information

Tab. E.1 illustrates the large variation across municipalities of the variables from which we construct the 'progressivism' indicator.

From Linder et al. (2010) and Année Politique Suisse (`www.swissvotes.ch`) we extract the following background information on the employed referenda.

- The referendum on gender equality in 1981 was intended to give equal rights to women with respect to professional life, family life, and education. A particular focus of the initiative was to ensure by law equal pay for equal work. Opponents feared that the proposal would interfere in wage negotiations and endanger private autonomy of families. The referendum passed in 17 of the 26 cantons. The overall support for the

Table E.1: Minimum, Quartiles, and Maximum of Sociocultural Indicators

	Min. (1)	P25 (2)	Median (3)	P75 (4)	Max. (5)
Share in favor of gender equality	0.179	0.560	0.652	0.716	0.914
Share in favor of heroin program	0.164	0.488	0.557	0.614	0.766
Share in favor of stem cell engineering	0.314	0.634	0.674	0.765	0.970
Share in favor of gay marriage	0.182	0.543	0.615	0.655	0.766
Share of left-wing parties	0.000	0.239	0.320	0.410	0.868
Share of Catholics	0.000	0.215	0.395	0.686	0.986

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* The number of observations is 4,767 survey respondents.

proposal was 60.3 percent. A nationwide representative poll revealed that opponents of the initiative disapproved equal rights for men and women in general.

- After the isolation of embryonic stem cells in 1998, advances in stem cell engineering prompted the debate on how to deal with the use of embryonic stem cells legally and ethically. The Swiss government proposed a law allowing scientists to use stem cells, albeit only from embryos left over from in vitro fertilization procedures. In 2004, the electorate accepted the law with a majority of 66.4 percent. All cantons voted in favor of the proposal. Polls showed that 40 percent of the opponents said that ethical concerns were the principal reason for voting against the proposal. About 50 percent of adversaries expressed that doubts about the merits of scientific research in general and fears of unwelcome consequences were decisive factors to oppose the law.

- In a scientific pilot project set up by the Swiss Government in 1994, drug addicts were entitled to receive heroin from a physician for free. Evaluation of this program suggested positive effects on both the health and the social situation of drug addicts. Subsequently, the Swiss government was seeking to enlarge the set of therapies by a state-controlled distribution of heroin all over Switzerland. Adversaries argued, in particular, that the state would financially and morally support drug addicts by this law. In 1999, the people approved the proposal by a majority of 54.4 percent. Ten cantons refused the law.

- Finally, in 2005, a referendum was held on the introduction of equal rights for homosexual couples in civil law. Registered homosexual partnerships were supposed to get the same rights as married couples except for the right to adopt children and access to in vitro fertilization. 58 percent of voters approved the law. In seven of the 26 cantons the majority refused it. A nationwide poll found that voters based their decision on their fundamental conviction whether homosexual partnerships should be legally and socially recognized.

E.2 Principal Component Analysis

Tab. E.2 shows that the shares of yes votes in the four referenda described in Section E.1, the vote share of left-wing parties, and the share of Catholics load particularly high on a single principal component.

Table E.2: Principal Component Analysis of the Six Variables for the Cultural Environment

Component	Eigenvalue	Difference	Proportion	Cumulative
1	2.598	1.573	0.433	0.433
2	1.025	0.017	0.171	0.604
3	1.008	0.373	0.168	0.772
4	0.635	0.154	0.106	0.878
5	0.480	0.226	0.080	0.958
6	0.254		0.042	1.000

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* The six variables are the share of yes votes in four national referenda on gender equality and science related issues, the share of left-wing parties in the federal elections, the share of Catholics. The number of observations is 2,575 municipalities.

Tab. E.3 displays the scoring coefficients. Consistent with the interpretation of the first principal component as 'progressivism' indicator, they are positive for the vote shares in the referenda and the national elections and negative for the share of Catholics in a municipality.

Table E.3: Loadings on the First Principal Component

Variable	Value
Share in favor of gender equality	0.413
Share in favor of heroin program	0.358
Share in favor of stem cell engineering	0.418
Share in favor of gay marriage	0.524
Share of left-wing parties	0.430
Share of Catholics	-0.257

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* The number of observations is 2,575 municipalities.

E.3 Stability of Sociocultural Variables Over Time

Tab. E.4 shows the results from regressing these sociocultural variables on municipality fixed effects and time fixed effects. Time fixed effects are not significantly different from zero, neither separately nor jointly. Thus, we conclude that sociocultural variables are relatively constant over time for the relevant period.

Table E.4: Fixed Effects Regression of Standardized Sociocultural Variables on Year and Municipality Dummies

	Coefficient	Std. Error	t-Statistic	p-Value
Year = 1990	-0.0044	0.0271	-0.16	0.87
Year = 1995	-0.0001	0.0224	0.00	1.00
Year = 1999	0.0002	0.0243	0.01	0.99
Year = 2004	0.0002	0.0197	0.01	0.99
Year = 2005	0.0002	0.0219	0.01	0.99
Constant	0.0007	0.0150	0.04	0.97

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* Fixed Effects regression is based on 15,525 observations from 2,575 municipalities. Municipality dummies are included. Left-out time dummy is for year 1981 (year of referendum on gender equality). Standard errors are adjusted for clustering at the municipality level. P-value for Wald test of joint significance of year dummies is 1.00.

F Gender Wage Gap When Not Accounting for Self-Selection into Field of Study

According to Tab. 3 in Section 5.2 of the main text (Benchmark Model Estimations), females do not earn significantly less than men after conditioning on observed characteristics and the correction term for self-selection.

Tab. F.1 shows the earnings regressions without the correction term for self-selection, i.e. without using first stage results from the reduced-form estimation.We see that, in this case, female STEM graduates earn 10 percent less and female non-STEM graduates earn 8.1 percent less than their male counterparts. The earnings differences between males and females are highly significant.

Table F.1: Gender Wage Gap When Not Accounting for Self-Selection into Field of Study

	Coefficient	Std. Error	t-Statistic	p-Value
Dependent Variable: Log Hourly Wage of STEM Graduates (1,438 Obs.)				
Female	-0.100	0.026	-3.89	0.000
Log age	-0.011	0.015	-0.76	0.445
Postgraduate education	-0.234	0.022	-10.73	0.000
F: no/missing, M: vocational educ.	-0.108	0.077	-1.41	0.160
F: no/missing, M: university educ.	-0.526	0.099	-5.34	0.000
F: vocational, M: no/missing educ.	0.000	0.051	0.00	0.996
Both parents with vocational educ.	-0.002	0.044	-0.05	0.962
F: vocational, M: university educ.	0.028	0.104	0.27	0.785
F: university, M: no/missing educ.	-0.044	0.057	-0.78	0.438
F: university, M: vocational educ.	-0.034	0.047	-0.74	0.460
Both parents with university educ.	-0.135	0.056	-2.44	0.015
Constant	4.242	0.843	5.03	0.000
Dependent Variable: Log Hourly Wage of Non-STEM Graduates (3,329 Obs.)				
Female	-0.081	0.012	-6.60	0.000
Log age	0.005	0.006	0.85	0.395
Postgraduate education	-0.085	0.012	-6.97	0.000
F: no/missing, M: vocational educ.	0.015	0.033	0.44	0.658
F: no/missing, M: university educ.	0.158	0.075	2.12	0.034
F: vocational, M: no/missing educ.	0.040	0.029	1.39	0.165
Both parents with vocational educ.	0.036	0.025	1.44	0.150
F: vocational, M: university educ.	-0.057	0.049	-1.17	0.244
F: university, M: no/missing educ.	0.039	0.036	1.08	0.278
F: university, M: vocational educ.	0.017	0.027	0.62	0.534
Both parents with university educ.	-0.065	0.029	-2.25	0.025
Constant	3.256	0.345	9.45	0.000

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* Log age is normalized to have a standard deviation of one. The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy. Heteroskedasticity robust standard errors.

G Estimations with Gender Interactions – Coefficients

In Tab. 4 in Section 5.3 of the main text we presented average partial effects from the estimations with gender interactions, allowing us to separately assess results for men and women. Tab. G.1 shows the coefficients (rather than the average partial effects) of the same estimations, including t-statistics for the null hypothesis that coefficients on gender interactions are equal to zero (i.e. statistical significance is marked with stars).

We see from column (4) of Tab. G.1 that the gender interactions of the progressivism indicator and the log wage differential ("return to STEM") in stage 3 estimations are statistically significant, as claimed in the main text. The main text also states the coefficients on the interaction of gender with the selection correction terms in the earnings regressions (stage 2) are not significantly different from zero, according to columns (2) and (3) of Tab. G.1.

Table G.1: Estimations with Gender Interactions (Coefficients)

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.116			-0.106
	(0.030)***			(0.030)***
Log relative distance	-0.058			-0.061
	(0.035)*			(0.035)*
Majority French	-0.100			-0.065
	(0.044)**			(0.044)
Majority Italian	0.039			0.031
	(0.096)			(0.095)
Log age	-0.223	-0.001	0.033	
	(0.031)***	(0.022)	(0.016)**	
Postgraduate education	-0.129	-0.108	-0.030	
	(0.029)***	(0.015)***	(0.013)**	
F: no/missing, M: vocational educ.	0.003	-0.017	-0.000	
	(0.031)	(0.015)	(0.010)	
F: no/missing, M: university educ.	0.022	-0.038	0.007	
	(0.036)	(0.005)***	(0.009)	
F: vocational, M: no/missing educ.	0.059	0.004	0.006	
	(0.040)	(0.019)	(0.013)	
Both parents with vocational educ.	0.183	-0.009	-0.006	
	(0.049)***	(0.028)	(0.020)	

<continued on next page>

Table G.1: Estimations with Gender Interactions (Coefficients) <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
F: vocational, M: university educ.	0.009	0.001	-0.008	
	(0.031)	(0.016)	(0.010)	
F: university, M: no/missing educ.	0.024	-0.012	0.024	
	(0.033)	(0.015)	(0.010)**	
F: university, M: vocational educ.	0.101	-0.022	0.000	
	(0.044)**	(0.022)	(0.016)	
Both parents with university educ.	0.049	-0.037	-0.027	
	(0.040)	(0.020)*	(0.014)*	
Correction term λ		0.052	0.291	
		(0.106)	(0.120)**	
Log wage differential				0.200
				(0.064)***
Female × progressivism indicator	0.114			0.107
	(0.047)**			(0.048)**
Female × log relative distance	-0.008			-0.033
	(0.046)			(0.047)
Female × majority French	-0.049			-0.040
	(0.048)			(0.047)
Female × majority Italian	0.021			0.023
	(0.047)			(0.048)
Female × log age	0.100	-0.010	-0.014	
	(0.049)**	(0.033)	(0.015)	
Female × postgraduate education	0.090	-0.016	0.000	
	(0.044)**	(0.028)	(0.015)	
Female × F: no/missing, M: vocational educ.	0.026	-0.020	0.002	
	(0.051)	(0.035)	(0.015)	
Female × F: no/missing, M: university educ.	-0.056	0.028	0.005	
	(0.050)	(0.011)**	(0.011)	
Female × F: vocational, M: no/missing educ.	-0.059	-0.030	0.002	
	(0.066)	(0.038)	(0.020)	
Female × both parents with vocational educ.	-0.119	-0.001	0.009	
	(0.089)	(0.055)	(0.027)	
Female × F: vocational, M: university educ.	-0.075	0.027	0.003	
	(0.054)	(0.023)	(0.015)	

<continued on next page>

Table G.1: Estimations with Gender Interactions (Coefficients) <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Female × F: university, M: no/missing educ.	0.052	-0.005	-0.048	
	(0.054)	(0.032)	(0.017)***	
Female × F: university, M: vocational educ.	-0.096	0.015	-0.003	
	(0.080)	(0.044)	(0.023)	
Female × both parents with university educ.	-0.040	-0.026	0.005	
	(0.069)	(0.042)	(0.020)	
Female × correction term λ		0.118	-0.002	
		(0.164)	(0.152)	
Female × log wage differential				-0.161
				(0.068)**
Female	-6.534	0.700	0.847	-0.650
	(2.872)**	(1.931)	(0.897)	(0.211)***
Constant	12.614	3.710	1.516	-0.488
	(1.802)***	(1.218)***	(0.974)	(0.202)**

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the coefficients. The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

H Sensitivity Analyses

This section presents all detailed results referred to in Section 5.4 of the main paper in the order of appearance. We start with summary statistics of the additional control variables and an overview of results from joint significance tests of variable categories of the different first stage estimates in the extended models.

H.1 Additional Control Variables

For the sensitivity analysis, we also control for total and female employment rates (total and for females), municipality size, the sectoral / occupational structure of municipalities, and the specialization of graduates in general high school. The (female) employment rate is defined as those (females) employed divided by the (female) population aged 15+. The sectoral and occupational structure is measured by employment shares (those who work in

some sector or occupation divided by all workers), in total or just considering females living in a municipality).

Tab. H.1 presents the summary statistics by distinguishing STEM and non-STEM graduates. We see that 57.2 percent of STEM graduates specialized in general high school in mathematics & science, compared to only 17.5 percent of the non-STEM graduates – a highly significant difference. 38.8 percent of non-STEM graduates have chosen a specialization containing Latin language (helpful to study non-STEM fields like law or languages), compared to 23.3 percent of STEM graduates.

With respect to the sectoral and occupational structure of municipalities, backgrounds of STEM and non-STEM graduates sometimes differ statistically, although quantitatively differences are quite small. For instance, the municipality employment rate of females and in total does not seem to be associated with university major choice. Even the employment share in technical occupations is unimportant. For instance, the average share of females working in a technical occupation in the female working population is 1.7 percent in the home municipality of a graduate for both STEM and non-STEM graduates.

H.2 Overview on Extended Models

Tab. H.2 presents for the benchmark model and its extensions (discussed in the following) the results from a Wald test for the joint significance of different variable categories in the reduced-form estimation (stage 1): the relative distance (capturing the monetary cost to study at university, panel (a)), sociocultural background (progressivism indicator and majority language dummies, panel (b)), earnings capability measures (panel (c)), region fixed effects (panel (e)), additional controls in the extensions (panel (d)), and the remaining variables (dummy for female and second cohort, panel (f))). Recall that the relative distance measure and sociocultural factors are excluded at stage 2 whereas the earnings capability measures are excluded at stage 3.

Column (1) of Tab. H.2 shows that for the benchmark model we can reject for each excluded instrument category (panels (a)-(c)) the hypothesis that coefficients are zero at stage 1 in a joint test. According to columns (2)-(5), the same is true for the extended models. By contrast, coefficients on region fixed effects (panel (e)) are not significantly different from zero in any of the specifications. Also coefficients on additional controls that capture the economic environment (employment rates, municipality size, sectoral / occupational structure) are insignificant in a joint test, according to columns (2)-(4), panel (d). By contrast, coefficients on the two dummies capturing specialization of graduates in general high school in mathematics & sciences and Latin language are significantly different from zero at the one percent level in a joint test, according to column (5), panel (d).

We will next examine whether the effects of interest remain robust to extending our benchmark model in the directions indicated in columns (2)-(5).

Table H.1: Descriptive Statistics of Additional Control Variables

	STEM (1)	Non-STEM (2)	t-Statistic (3)
Spec. in math & sciences at high school (1=yes, 0=no)	0.572 (0.495)	0.175 (0.380)	27.154
Specialized in Latin at high school (1=yes, 0=no)	0.233 (0.423)	0.388 (0.487)	-11.065
Employment rate	0.626 (0.046)	0.623 (0.045)	2.534
Female employment rate	0.480 (0.053)	0.480 (0.050)	-0.112
Share of agriculture	0.037 (0.052)	0.032 (0.049)	2.729
Share of manufacturing	0.203 (0.086)	0.189 (0.086)	5.463
Share of construction	0.086 (0.027)	0.084 (0.029)	2.760
Share of business services	0.143 (0.053)	0.151 (0.053)	-4.809
Share females in agriculture	0.029 (0.046)	0.024 (0.042)	3.455
Share females in manufacturing	0.147 (0.080)	0.138 (0.077)	3.663
Share females in construction	0.018 (0.009)	0.017 (0.009)	3.301
Share females in business services	0.152 (0.049)	0.161 (0.050)	-5.714
Share of technical occupations	0.077 (0.020)	0.075 (0.019)	2.929
Share of managerial occupations	0.211 (0.052)	0.220 (0.055)	-5.518
Share of health, education, research, culture occ.	0.131 (0.038)	0.137 (0.039)	-4.638
Share females in technical occupations	0.017 (0.007)	0.017 (0.007)	-1.055
Share females in managerial occupations	0.269 (0.057)	0.279 (0.059)	-5.184
Sh. females in health, education, research, culture occ.	0.202 (0.051)	0.208 (0.051)	-3.367
100,000 and more inhabitants (1=yes, 0=no)	0.111 (0.314)	0.154 (0.361)	-4.126
Between 50,000 and 99,999 (1=yes, 0=no)	0.042 (0.201)	0.036 (0.186)	0.997
Between 20,000 and 49,999 (1=yes, 0=no)	0.086 (0.280)	0.100 (0.300)	-1.616
Between 10,000 and 19,999 (1=yes, 0=no)	0.184 (0.388)	0.177 (0.382)	0.603
Between 5,000 and 9,999 (1=yes, 0=no)	0.185 (0.389)	0.170 (0.376)	1.238
Between 2,000 and 4,999 (1=yes, 0=no)	0.215 (0.411)	0.214 (0.410)	0.119
Between 1,000 and 1,999 (1=yes, 0=no)	0.104 (0.306)	0.078 (0.268)	2.807
Less than 1,000 inhabitants (1=yes, 0=no)	0.072 (0.258)	0.070 (0.256)	0.151
Observations	1,438	3,329	

Source: Federal Statistical Office of Switzerland, own calculations. *Note:* The first two columns show the means and standard deviations (in parentheses) of the variables, the last column the t-statistics of a test of equality of means.

Table H.2: Wald Tests of Joint Significance for Reduced Form Choice Models with Additional Controls

	Benchmark	Benchmark+ Munic. sizes+ Empl. rates	Benchmark+ Munic. sizes+ Empl. rates+ Industries	Benchmark+ Munic. sizes+ Empl. rates+ Occupations	Benchmark+ High school specialization
	(1)	(2)	(3)	(4)	(5)
			(a) Monetary cost		
χ^2-Statistic	4.116	4.430	2.497	2.921	2.896
Degr. of freedom	1	1	1	1	1
p-Value	0.042	0.035	0.114	0.087	0.089
			(b) Sociocultural background		
χ^2-Statistic	19.481	19.763	9.582	16.778	22.587
Degr. of freedom	3	3	3	3	3
p-Value	0.000	0.000	0.022	0.001	0.000
			(c) Earnings capability		
χ^2-Statistic	115.108	115.794	116.617	116.043	67.323
Degr. of freedom	10	10	10	10	10
p-Value	0.000	0.000	0.000	0.000	0.000
			(d) Additional controls		
χ^2-Statistic	–	8.127	19.153	11.463	448.443
Degr. of freedom	–	9	17	15	2
p-Value	–	0.521	0.320	0.719	0.000
			(e) Region fixed effects		
χ^2-Statistic	10.155	9.850	10.725	9.072	12.704
Degr. of freedom	6	6	6	6	6
p-Value	0.118	0.131	0.097	0.170	0.048
			(f) Remaining variables		
χ^2-Statistic	282.284	278.411	276.162	279.952	106.444
Degr. of freedom	2	2	2	2	2
p-Value	0.000	0.000	0.000	0.000	0.000

Source: Swiss Federal Statistical Office, Année Politique Suisse, own calculations. *Note:* The panel labeled 'Monetary cost' refers to a Wald test of the log relative distance to the next technical university (ETH). The panel 'Sociocultural background' includes the progressivism variable and dummies for the majority language. The panel 'Earnings capability' includes the variables log age, dummies for parental education, and postgraduate training. The panel 'Additional controls' includes the group of variables indicated in the column header. The panel 'Region fixed effects' contains dummies for the *NUTS*-2 regions. The remaining variables in the last panel are dummies for female and second cohort. Variance matrices for test statistics are bootstrapped.

H.3 Economic Environment

Tab. H.3 shows the results when adding at all stages the total employment rate, the female employment rate and dummies for the municipality size brackets to the benchmark model, corresponding to column (2) of Tab. H.2. Results confirm that average partial effects on the progressivism indicator, the relative distance measure and the log wage differential for university major choice at stage 3 remain at the significance levels of the benchmark model and have similar magnitudes. Employment rates do not significantly affect any of the outcomes.

Table H.3: Extended Model with Employment Rates and Municipality Size

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.030			-0.026
	(0.009)***			(0.009)***
Female	-0.222	-0.000	0.041	-0.206
	(0.013)***	(0.079)	(0.045)	(0.029)***
Employment rate	-0.021	-0.014	0.025	-0.005
	(0.019)	(0.029)	(0.020)	(0.021)
Female employment rate	0.021	0.031	-0.018	-0.000
	(0.019)	(0.029)	(0.018)	(0.021)
50,000 to 99,999 inhabitants	0.008	-0.076	0.052	0.060
	(0.039)	(0.068)	(0.046)	(0.052)
20,000 to 49,999 inhabitants	-0.011	-0.059	0.012	0.014
	(0.030)	(0.047)	(0.029)	(0.037)
10,000 to 19,999 inhabitants	0.005	-0.036	0.009	0.021
	(0.028)	(0.045)	(0.030)	(0.033)
5,000 to 9,999 inhabitants	0.011	-0.086	0.009	0.045
	(0.030)	(0.047)*	(0.031)	(0.038)
2,000 to 4,999 inhabitants	-0.016	0.010	-0.013	-0.020
	(0.030)	(0.051)	(0.031)	(0.034)
1,000 to 1,999 inhabitants	0.049	-0.067	-0.054	0.060
	(0.038)	(0.064)	(0.042)	(0.044)
Less than 1,000 inhabitants	0.009	-0.044	-0.038	0.020
	(0.038)	(0.058)	(0.043)	(0.042)
Log relative distance	-0.021			-0.024
	(0.010)**			(0.010)**
Majority French	-0.081			-0.055
	(0.024)***			(0.027)**

<continued on next page>

Table H.3: Extended Model with Employment Rates and Municipality Size <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Majority Italian	0.056			0.042
	(0.120)			(0.122)
Log age	-0.058	0.012	0.032	
	(0.007)***	(0.022)	(0.012)***	
Postgraduate education	-0.059	-0.204	-0.054	
	(0.013)***	(0.029)***	(0.018)***	
F: no/missing, M: vocational educ.	0.023	-0.122	0.000	
	(0.041)	(0.077)	(0.037)	
F: no/missing, M: university educ.	-0.030	-0.520	0.187	
	(0.097)	(0.138)***	(0.077)**	
F: vocational, M: no/missing educ.	0.040	-0.019	0.015	
	(0.033)	(0.054)	(0.032)	
Both parents with vocational educ.	0.090	-0.047	-0.017	
	(0.027)***	(0.055)	(0.033)	
F: vocational, M: university educ.	-0.026	0.043	-0.041	
	(0.048)	(0.103)	(0.053)	
F: university, M: no/missing educ.	0.066	-0.075	0.009	
	(0.041)	(0.063)	(0.044)	
F: university, M: vocational educ.	0.053	-0.062	-0.010	
	(0.029)*	(0.051)	(0.032)	
Both parents with university educ.	0.035	-0.154	-0.081	
	(0.033)	(0.057)***	(0.034)**	
Correction term λ		0.192	0.379	
		(0.136)	(0.131)***	
Log wage differential				0.038
				(0.018)**

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressors). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

Tab. H.4 shows the results when also adding at all stages the industry structure to the

benchmark model, corresponding to column (3) of Tab. H.2. Again, the average partial effects on the progressivism indicator, the relative distance measure and the log wage differential for university major choice at stage 3 remain at similar magnitudes albeit significance levels decrease.

Table H.4: Extended Model with Employment Rates, Municipality Size, and Industry Structure

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.025			-0.021
	(0.010)**			(0.010)**
Female	-0.223	0.031	0.071	-0.207
	(0.013)***	(0.109)	(0.053)	(0.035)***
Employment rate	-0.021	-0.028	0.046	0.006
	(0.021)	(0.033)	(0.023)**	(0.026)
Female employment rate	0.027	0.042	-0.044	-0.005
	(0.021)	(0.034)	(0.022)**	(0.028)
Share of agriculture	-0.015	0.023	-0.032	-0.035
	(0.021)	(0.036)	(0.021)	(0.024)
Share of manufacturing	0.056	-0.008	-0.057	0.043
	(0.024)**	(0.051)	(0.030)*	(0.030)
Share of construction	0.004	-0.002	-0.013	0.004
	(0.010)	(0.018)	(0.011)	(0.012)
Share of business services	0.012	0.026	-0.015	0.007
	(0.023)	(0.042)	(0.024)	(0.027)
Share females in agriculture	0.011	0.006	0.029	0.022
	(0.019)	(0.032)	(0.018)	(0.020)
Share females in manufacturing	-0.058	0.035	0.045	-0.055
	(0.020)***	(0.048)	(0.027)*	(0.024)**
Share females in construction	-0.005	0.024	0.001	-0.013
	(0.008)	(0.014)*	(0.009)	(0.010)
Share females in business services	-0.025	-0.002	0.025	-0.023
	(0.020)	(0.038)	(0.022)	(0.023)
50,000 to 99,999 inhabitants	-0.007	-0.088	0.074	0.050
	(0.039)	(0.072)	(0.049)	(0.058)
20,000 to 49,999 inhabitants	-0.017	-0.077	0.028	0.015
	(0.031)	(0.051)	(0.031)	(0.041)

<continued on next page>

Table H.4: Extended Model with Employment Rates, Municipality Size, and Industry Structure) <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
10,000 to 19,999 inhabitants	0.001	-0.068	0.023	0.028
	(0.030)	(0.049)	(0.032)	(0.039)
5,000 to 9,999 inhabitants	0.007	-0.129	0.021	0.054
	(0.033)	(0.053)**	(0.035)	(0.049)
2,000 to 4,999 inhabitants	-0.018	-0.037	0.003	-0.004
	(0.033)	(0.057)	(0.036)	(0.038)
1,000 to 1,999 inhabitants	0.045	-0.129	-0.046	0.077
	(0.042)	(0.071)*	(0.048)	(0.051)
Less than 1,000 inhabitants	0.012	-0.118	-0.021	0.053
	(0.044)	(0.068)*	(0.052)	(0.052)
Log relative distance	-0.016			-0.019
	(0.010)			(0.010)*
Majority French	-0.055			-0.028
	(0.027)**			(0.030)
Majority Italian	0.067			0.058
	(0.122)			(0.126)
Log age	-0.058	0.021	0.039	
	(0.007)***	(0.028)	(0.014)***	
Postgraduate education	-0.060	-0.192	-0.046	
	(0.013)***	(0.036)***	(0.020)**	
F: no/missing, M: vocational educ.	0.022	-0.128	0.001	
	(0.041)	(0.078)	(0.038)	
F: no/missing, M: university educ.	-0.028	-0.539	0.176	
	(0.096)	(0.150)***	(0.080)**	
F: vocational, M: no/missing educ.	0.038	-0.028	0.013	
	(0.033)	(0.056)	(0.033)	
Both parents with vocational educ.	0.088	-0.064	-0.027	
	(0.027)***	(0.062)	(0.034)	
F: vocational, M: university educ.	-0.026	0.044	-0.038	
	(0.048)	(0.107)	(0.055)	
F: university, M: no/missing educ.	0.064	-0.089	-0.002	
	(0.041)	(0.068)	(0.046)	
F: university, M: vocational educ.	0.051	-0.077	-0.016	
	(0.029)*	(0.055)	(0.032)	

<continued on next page>

Table H.4: Extended Model with Employment Rates, Municipality Size, and Industry Structure) <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Both parents with university educ.	0.032	-0.163	-0.086	
	(0.033)	(0.059)***	(0.035)**	
Correction term λ		0.254	0.473	
		(0.193)	(0.158)***	
Log wage differential				0.039
				(0.024)

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressors). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

Tab. H.5 shows the results when adding at all stages the occupational structure instead of the industry structure to the benchmark model (again also controlling for the employment rate, female employment rate, and municipality size dummies), corresponding to column (4) of Tab. H.2.

Average partial effects on the progressivism indicator, the relative distance measure and the log wage differential for university major choice at stage 3 have the same significance levels as the benchmark model. Results also suggest that progressivism plays an even larger role quantitatively than in the benchmark model. An increase of the progressivism indicator by one standard deviation reduces the probability to choose a STEM field by 3.1 percentage points.

Table H.5: Extended Model with Employment Rates, Municipality Size, and Occupational Structure

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.034			-0.031
	(0.011)***			(0.011)***
Female	-0.222	-0.015	0.048	-0.198
	(0.013)***	(0.093)	(0.049)	(0.033)***
Employment rate	-0.023	-0.012	0.035	-0.005
	(0.024)	(0.039)	(0.025)	(0.028)
Female employment rate	0.024	0.028	-0.028	0.001
	(0.023)	(0.036)	(0.022)	(0.026)
Share of technical occupations	0.014	-0.007	0.003	0.016
	(0.011)	(0.018)	(0.013)	(0.012)
Share of managerial occupations	-0.008	0.034	0.032	-0.005
	(0.019)	(0.033)	(0.020)	(0.022)
Share of health, education, research, culture occupations	0.011	-0.013	-0	0.006
	(0.030)	(0.045)	(0.030)	(0.031)
Share females in technical occupations	-0.011	0.021	0.005	-0.015
	(0.008)	(0.014)	(0.008)	(0.009)
Share females in managerial occupations	0.005	-0.038	-0.025	0.008
	(0.017)	(0.032)	(0.019)	(0.021)
Share females in health, education, research, culture occupations	-0.004	0.002	-0.001	0.003
	(0.024)	(0.037)	(0.025)	(0.026)
50,000 to 99,999 inhabitants	-0.004	-0.076	0.045	0.045
	(0.040)	(0.068)	(0.047)	(0.053)
20,000 to 49,999 inhabitants	-0.020	-0.061	0.007	0.004
	(0.033)	(0.050)	(0.030)	(0.040)
10,000 to 19,999 inhabitants	-0.003	-0.037	-0.001	0.010
	(0.031)	(0.046)	(0.030)	(0.035)
5,000 to 9,999 inhabitants	0.000	-0.086	-0.006	0.030
	(0.033)	(0.049)*	(0.031)	(0.041)
2,000 to 4,999 inhabitants	-0.025	0.009	-0.025	-0.034
	(0.033)	(0.051)	(0.032)	(0.036)
1,000 to 1,999 inhabitants	0.040	-0.063	-0.067	0.043
	(0.041)	(0.065)	(0.042)	(0.045)
Less than 1,000 inhabitants	0.004	-0.034	-0.043	0.008
	(0.040)	(0.058)	(0.043)	(0.043)

<continued on next page>

Table H.5: Extended Model with Employment Rates, Municipality Size, and Occupational Structure <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Log relative distance	-0.018			-0.022
	(0.011)*			(0.011)**
Majority French	-0.078			-0.053
	(0.026)***			(0.029)*
Majority Italian	0.070			0.057
	(0.121)			(0.124)
Log age	-0.058	0.010	0.033	
	(0.007)***	(0.024)	(0.013)***	
Postgraduate education	-0.059	-0.206	-0.053	
	(0.013)***	(0.031)***	(0.019)***	
F: no/missing, M: vocational educ.	0.021	-0.113	0.001	
	(0.041)	(0.076)	(0.037)	
F: no/missing, M: university educ.	-0.031	-0.528	0.180	
	(0.097)	(0.140)***	(0.078)**	
F: vocational, M: no/missing educ.	0.039	-0.016	0.014	
	(0.033)	(0.054)	(0.032)	
Both parents with vocational educ.	0.090	-0.042	-0.021	
	(0.027)***	(0.058)	(0.033)	
F: vocational, M: university educ.	-0.028	0.051	-0.043	
	(0.048)	(0.105)	(0.054)	
F: university, M: no/missing educ.	0.064	-0.068	0.002	
	(0.041)	(0.064)	(0.044)	
F: university, M: vocational educ.	0.051	-0.055	-0.015	
	(0.029)*	(0.053)	(0.032)	
Both parents with university educ.	0.033	-0.148	-0.086	
	(0.033)	(0.058)**	(0.034)**	
Correction term λ		0.170	0.400	
		(0.159)	(0.143)***	
Log wage differential				0.043
				(0.021)**

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressor). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

H.4 High School Specialization

Tab. H.6 shows the results when adding the high school specialization to the benchmark model, corresponding to column (5) of Tab. H.2. Average partial effects on both dummies, for specialization in mathematics & sciences and for specialization in Latin language are highly significant and have sizable effects. Importantly, as discussed in Section 5.4 of the main text, the inclusion of these variables does not change any of our main results.

Table H.6: Extended Model with High School Specialization

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.024			-0.022
	(0.007)***			(0.007)***
Female	-0.136	-0.066	-0.030	-0.121
	(0.013)***	(0.045)	(0.025)	(0.019)***
Spec. in math & sc. at high school	0.132	-0.020	-0.086	0.119
	(0.006)***	(0.036)	(0.026)***	(0.015)***
Spec. in Latin at high school	0.016	-0.050	-0.028	0.026
	(0.007)**	(0.017)***	(0.007)***	(0.008)***
Log relative distance	-0.015			-0.017
	(0.009)*			(0.009)**
Majority French	-0.081			-0.067
	(0.023)***			(0.024)***
Majority Italian	0.066			0.062
	(0.114)			(0.116)
Log age	-0.037	-0.008	0.014	
	(0.006)***	(0.017)	(0.008)*	
Postgraduate education	-0.051	-0.223	-0.064	
	(0.012)***	(0.025)***	(0.016)***	
F: no/missing, M: vocational educ.	0.048	-0.117	-0.007	
	(0.039)	(0.075)	(0.035)	
F: no/missing, M: university educ.	-0.079	-0.508	0.184	
	(0.087)	(0.121)***	(0.067)***	
F: vocational, M: no/missing educ.	0.023	-0.005	0.029	
	(0.030)	(0.051)	(0.029)	
Both parents with vocational educ.	0.067	-0.015	0.005	
	(0.025)***	(0.048)	(0.026)	

<continued on next page>

Table H.6: Extended Model with High School Specialization <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
F: vocational, M: university educ.	-0.030	0.026	-0.042	
	(0.045)	(0.099)	(0.050)	
F: university, M: no/missing educ.	0.070	-0.042	0.018	
	(0.038)*	(0.059)	(0.039)	
F: university, M: vocational educ.	0.037	-0.033	0.010	
	(0.027)	(0.048)	(0.028)	
Both parents with university educ.	0.039	-0.126	-0.070	
	(0.031)	(0.055)**	(0.031)**	
Correction term λ		0.060	0.278	
		(0.101)	(0.102)***	
Log wage differential				0.038
				(0.013)***

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressors). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

I Further Evidence on Exclusion Restrictions

We finally experiment with including sociocultural indicators (for progressivism and majority language) into stage 2 and relative distance to the next technical university into stage 3.

Tab. I.1 confirms that the progressivism indicator does not enter significantly at stage 2. Moreover, the point estimates of the progressivism indicator, the relative distance measure and the log earnings differential remain similar at stage 3 (structural major choice), compared to the benchmark model (Tab. 3 in the main text).

Table I.1: Model with Sociocultural Indicators in Earnings Equations

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.024	0.001	0.011	-0.019
	(0.007)***	(0.030)	(0.010)	(0.009)**
Female	-0.124	-0.027	0.014	-0.104
	(0.008)***	(0.105)	(0.030)	(0.030)***
Log relative distance	-0.018			-0.022
	(0.009)**			(0.009)**
Majority French	-0.076	0.001	-0.083	-0.083
	(0.024)***	(0.086)	(0.034)**	(0.042)**
Majority Italian	0.053	0.006	-0.049	0.024
	(0.120)	(0.129)	(0.248)	(0.133)
Log age	-0.058	-0.001	0.026	
	(0.007)***	(0.054)	(0.016)	
Postgraduate education	-0.030	-0.111	-0.030	
	(0.007)***	(0.028)***	(0.010)***	
F: no/missing, M: vocational educ.	0.005	-0.022	-0.000	
	(0.008)	(0.016)	(0.007)	
F: no/missing, M: university educ.	-0.001	-0.031	0.010	
	(0.007)	(0.009)***	(0.005)**	
F: vocational, M: no/missing educ.	0.013	-0.003	0.005	
	(0.010)	(0.023)	(0.010)	
Both parents with vocational educ.	0.043	-0.010	-0.008	
	(0.012)***	(0.048)	(0.017)	
F: vocational, M: university educ.	-0.004	0.005	-0.008	
	(0.008)	(0.017)	(0.008)	
F: university, M: no/missing educ.	0.014	-0.012	0.002	
	(0.008)*	(0.021)	(0.010)	
F: university, M: vocational educ.	0.021	-0.018	-0.006	
	(0.011)*	(0.031)	(0.013)	
Both parents with university educ.	0.011	-0.044	-0.027	
	(0.010)	(0.023)*	(0.011)**	
Correction term λ		0.082	0.336	
		(0.386)	(0.178)*	
Log wage differential				0.044
				(0.030)

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressors). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

Tab. I.2 shows the results when we include the relative distance measure also in the earnings regressions (stage 2). It confirms that the coefficient on relative distance is not significantly different from zero in the STEM equation. For non-STEM graduates it is marginally significant (at the 10 percent level). This raises the question whether the coefficients are zero when jointly tested, answered below.

Table I.2: Model with Log Relative Distance in Earnings Equations

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Progressivism indicator	-0.024			-0.023
	(0.007)***			(0.007)***
Female	-0.124	-0.019	0.048	-0.091
	(0.008)***	(0.042)	(0.027)*	(0.022)***
Log relative distance	-0.018	0.007	0.019	-0.016
	(0.009)**	(0.019)	(0.011)*	(0.011)
Majority French	-0.037			-0.027
	(0.012)***			(0.013)**
Majority Italian	0.012			0.010
	(0.028)			(0.029)
Log age	-0.058	0.003	0.045	
	(0.007)***	(0.023)	(0.013)***	
Postgraduate education	-0.030	-0.109	-0.021	
	(0.007)***	(0.015)***	(0.010)**	
F: no/missing, M: vocational educ.	0.005	-0.022	-0.001	
	(0.008)	(0.015)	(0.008)	
F: no/missing, M: university educ.	-0.001	-0.031	0.010	
	(0.007)	(0.007)***	(0.005)**	
F: vocational, M: no/missing educ.	0.013	-0.003	0.002	
	(0.010)	(0.017)	(0.011)	
Both parents with vocational educ.	0.043	-0.012	-0.018	
	(0.012)***	(0.026)	(0.017)	
F: vocational, M: university educ.	-0.004	0.005	-0.006	
	(0.008)	(0.015)	(0.009)	
F: university, M: no/missing educ.	0.014	-0.013	-0.001	
	(0.008)*	(0.013)	(0.010)	
F: university, M: vocational educ.	0.021	-0.019	-0.009	
	(0.011)*	(0.020)	(0.013)	

<continued on next page>

Table I.2: Model with Log Relative Distance in Earnings Equations <continued>

	Reduced Form STEM Choice (1)	Log Wage STEM (2)	Log Wage Non-STEM (3)	Structural STEM Choice (4)
Both parents with university educ.	0.011	-0.045	-0.028	
	(0.010)	(0.017)**	(0.012)**	
Correction term λ		0.114	0.543	
		(0.144)	(0.156)***	
Log wage differential				0.054
				(0.015)***

Source: Federal Statistical Office of Switzerland, Année Politique Suisse, own calculations. *Note:* In columns (1) and (4), the dependent variable is a dummy for graduation in a STEM field. In columns (2) and (3), the dependent variable is the log hourly wage of STEM and non-STEM graduates, respectively. Columns (1) to (4) show the average partial effect of the corresponding regressor for a regressor change by one standard deviation (continuous regressors, except correction terms) or from zero to one (dummy regressor). The model includes in addition *NUTS*-2 region fixed effects and a cohort dummy at all stages. Bootstrapped standard errors are shown in parentheses. *, ** and *** denote significance at the 10%-, 5%- and 1%-level, respectively.

Tab. I.3 refers to the estimates in Tab. I.2 and presents the results from two tests. First, it tests the hypothesis that the coefficients of the log relative distance to the next technical university in the two earnings equations are equal to zero. As claimed in Section 5.4 of the paper, the hypothesis cannot be rejected. Second, it tests the hypothesis that the coefficients on log relative distance in the STEM and non-STEM ("return to STEM") wage equation are equal to each other. The test statistic shows that the hypothesis cannot be rejected, suggesting that potentially differential labor demand for graduates by municipality is not important for the return to STEM.

Table I.3: Further Results on Log Relative Distance in Earnings Regressions

	χ^2-Statistic (1)	Degr. of Freedom (2)	p-Value (3)
H_0 : Coefficients on log relative distance = 0 in STEM and non-STEM wage equations	2.986	2	0.225
H_0 : Coefficients on log relative distance in STEM and non-STEM wage equations equal	0.378	1	0.539

References

[1] Federal Statistical Office (1990). Eidgenössische Volkszählung 1990, Ergebnisse nach Gemeindeebene, Neuchâtel.

[2] Federal Statistical Office (1995). Statistik der Nationalratswahlen, Ergebnisse nach Gemeinden, Neuchâtel.

[3] Federal Statistical Office (1996). Evaluation der Volkszählung 1990: Erhebungsmethoden und Alternativen, Neuchâtel.

[4] Federal Statistical Office (2008). Schlüsselkompetenzen der Schweizer Hochschulabsolvent/innen: Thematischer Sammelband mit empirischen Ergebnissen der Absolventenstudie, Neuchâtel.

[5] Federal Statistical Office (2009). Von der Hochschule ins Berufsleben: Erste Ergebnisse der Absolventenbefragung 2007, Neuchâtel.

[6] Federal Statistical Office (2010). Die Raumgliederungen der Schweiz 2010, Neuchâtel.

[7] Federal Statistical Office (2011). Die Raumgliederungen der Schweiz 2011, Neuchâtel.

[8] Federal Statistical Office (2012a). Erhebungen, Quellen – Absolventenstudien Hochschulen: Steckbrief, Neuchâtel.

[9] Federal Statistical Office (2012b), Statistisches Lexikon der Schweiz, Neuchâtel.

[10] Federal Statistical Office (2014). Historisiertes Gemeindeverzeichnis der Schweiz, Stand: 01. May 2014, Neuchâtel.

[11] Federal Statistical Office (2015). Bildungssystem Schweiz – Indikatoren: Abschlüsse und Kompetenzen – Maturitätsquote, Neuchâtel.

[12] Chetty, Raj (2006). A New Method of Estimating Risk Aversion, *American Economic Review* 96, 1821-1834.

[13] Das, Mitali, Whitney K. Newey, and Francis Vella (2003). Nonparametric Estimation of Sample Selection Models, *Review of Economic Studies* 70, 33-58.

[14] French, Eric and Christopher Taber (2011). Identification of Models of the Labor Market, in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 4A, ch. 6, Elsevier Science, Amsterdam, 537-617.

[15] Heckman, James J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement* 5, 475-492.

[16] Heckman, James J. (2010). Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy, *Journal of Economic Literature* 48, 356-398.

[17] Heckman, James J. and Edward J. Vytlacil (2007). Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation, in: J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, Vol. 6B, ch. 70, Elsevier Science, Amsterdam, 4779-4874.

[18] Linder, Wolf, Christian Bolliger, and Yvan Rielle (2010). Handbuch der eidgenössischen Volksabstimmungen 1848-2007, Haupt Verlag, Bern.

[19] Matzkin, Rosa L. (1992). Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models, *Econometrica* 60, 239-270.

[20] Roy, Andrew D. (1951). Some Thoughts on the Distribution of Earnings, *Oxford Economic Papers* 3, 135-146.

[21] Willis, Robert J. and Sherwin Rosen (1979). Education and Self-Selection, *Journal of Political Economy* 87, S7-S36.