ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 160/26

Characterizing the Returns to STEM: Marginal and Policy-Relevant Treatment Effects

Aderonke Osikominu, Gregor Pfeifer, Tim Ruberg

www.rfberlin.com JUNE 2026

Characterizing the Returns to STEM: Marginal and Policy-Relevant Treatment Effects

Authors

Aderonke Osikominu, Gregor Pfeifer, Tim Ruberg

Reference

JEL Codes: C26; I26; I28; J16; J24

Keywords: Returns to Education; STEM; MTE; PRTE; Gender

Recommended Citation: Aderonke Osikominu, Gregor Pfeifer, Tim Ruberg (2026): Characterizing the Returns to STEM: Marginal and Policy-Relevant Treatment Effects. RFBerlin Discussion Paper No. 160/26

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Characterizing the Returns to STEM: Marginal and Policy-Relevant Treatment Effects*

Aderonke Osikominu †　　Gregor Pfeifer ‡　　Tim Ruberg §

June 4, 2026

Abstract

We estimate heterogeneous returns to STEM education by leveraging relative distances to technical versus general universities in Switzerland. While individuals who choose a STEM education gain on average, a declining marginal treatment effect curve indicates positive selection on gains, suggesting that low-resistance individuals benefit the most. Through policy simulations aimed at increasing STEM enrollment and estimating corresponding policy-relevant treatment effects, we demonstrate that these policies' effectiveness critically depends on both observable and unobservable characteristics of affected individuals. Furthermore, we highlight how policies should be designed to both increase STEM enrollment and generate positive returns for targeted groups, particularly women.

JEL Classification: C26; I26; I28; J16; J24.

Keywords: Returns to Education; STEM; Gender; MTE; PRTE.

*The authors would like to thank Martin Biewen, Thomas Cornelissen, Albrecht Glitz, Daniel Kamhöfer, Hikaru Kawarazaki, Philipp Kugler, Shinpei Sano, Gregory Veramendi, Rainer Winkelmann, and seminar participants at the Institute for Employment Research, Institute for Fiscal Studies, Hitotsubashi University, Keio University, University of Freiburg, University of Hohenheim, University of Tübingen, Zeppelin University in Friedrichshafen, and participants at the Annual Congress of the Swiss Society of Economics and Statistics, Hermes Workshop, AASLE Conference, meeting of the Population Economics Committee of the Verein für Socialpolitik, and ASSA Annual Meeting for their helpful comments. We gratefully acknowledge financial support from the Baden-Württemberg Stiftung and the JST ERATO Grant Number JPMJER2301, Japan. The views expressed in this paper are those of the authors and do not necessarily reflect the positions or policies of the organizations mentioned above. Any remaining errors are our own.

†Aderonke Osikominu, University of Hohenheim, Department of Economics, Stuttgart, Germany; CEPR; CESifo; IZA; Email: a.osikominu@uni-hohenheim.de.

‡Gregor Pfeifer, The University of Sydney, School of Economics, Social Sciences Building, NSW 2006, Australia; CESifo; IZA; RFBerlin; Email: gregor.pfeifer@sydney.edu.au.

§Corresponding author: Tim Ruberg, Market Design Center, University of Tokyo, Graduate School of Economics, 7-3-1 Hongo, Bunkyo-ku, 113-0033 Tokyo, Japan; Email: ruberg@g.ecc.u-tokyo.ac.jp.

1 Introduction

Because of their importance for innovation and growth (Hunt and Gauthier-Loiselle, 2010; Peri et al., 2015), many countries are attempting to increase the number of STEM graduates through initiatives aimed at expanding access to STEM education in high schools and universities, promoting interest in STEM, and providing financial support. As women are particularly underrepresented in STEM majors at universities and in STEM occupations (Mourifie et al., 2020),[1] such initiatives often target women.[2] Although individuals with STEM degrees tend to have relatively high wages on average (Altonji et al., 2012; Kirkebøen et al., 2016), less is known about what drives heterogeneity in the returns to STEM. Understanding who benefits most from a STEM education is essential to formulating effective policies that not only increase STEM enrollment, but also improve the earnings prospects of those affected.

This paper uses individual-level data from Switzerland to estimate the causal effect of earning a STEM university degree on wages within a generalized Roy (1951) model (Heckman and Vytlacil, 2007a). This framework accounts for self-selection into university majors based not only on potential earnings but also on non-earnings-related factors such as costs and tastes. Switzerland's unique institutional setting, where prospective students can freely choose their universities and majors, enables us to exploit regional differences in the proximity of technical universities (which specialize in STEM), relative to the proximity of general universities, for identification. We allow individual returns to vary along both observed and unobserved characteristics and adopt the marginal treatment effect (MTE) framework of Björklund and Moffitt (1987) and Heckman and Vytlacil (1999, 2001a, 2005, 2007b). Finally, we use the estimated MTEs to compute policy-relevant treatment effects (PRTEs, Heckman and Vytlacil, 2001b) for various hypothetical policies. This allows us to assess which initiatives are most effective in increasing STEM enrollment while ensuring positive average returns for those affected by the policy change.

Our identification strategy relies on instrumental variables that are conditionally indepen-

[1]In 2021, men account for more than 75% of all enrollments in STEM fields in OECD countries (OECD, 2023, Fig. B4.2). This female underrepresentation can be attributed to factors such as the influence of role models (Carrell et al., 2010; Canaan and Mouganie, 2023), lower self-efficacy in math among women (Saltiel, 2023), and differing work-related preferences (Wiswall and Zafar, 2018).

[2]In 2022, Germany invested €45 million in a program to combine STEM education with extracurricular activities, including family involvement, to stimulate interest in STEM at an early age (Federal Ministry of Education and Research, 2021). Similarly, the United Kingdom supports various programs such as the *STEM Learning* initiative, which provides professional development support for teachers, brings STEM role models into schools, and supports STEM activities in schools (STEM Learning, 2025). An initiative in Switzerland, called *Swiss Tecladies* by the Swiss Academy of Engineering Sciences (SATW), aims to help young women discover and develop their quantitative skills and encourages them to pursue a career in a STEM field (Swiss Academy of Engineering Sciences, 2025).

dent of treatment and potential outcomes and that satisfy the exclusion restriction—meaning they influence the relative cost of pursuing a STEM education but not its return. Our main instrument is based on the geographic distance to the nearest technical university relative to the nearest general university. While Switzerland's nine general universities offer mostly non-STEM programs, the two technical universities have larger STEM departments with a wider range of disciplines, programs, and areas of specialization. Because attending a nearby general university for a non-STEM major is less costly than traveling to a distant technical university for a STEM major, this relative distance serves as a cost factor influencing major choice but not wages. We account for abilities and tastes through indicators for a math and science specialization in high school and parental educational attainment. In addition, we control for characteristics of the residence municipality at the end of high school like city size and local employment and industry structure, as well as fixed effects at the level of local labor markets. To improve precision, we use a second instrument based on age. Individuals who are older at college entry tend to make more mature decisions, placing greater emphasis on long-term labor market benefits over immediate amenities. As STEM fields offer more stable career prospects (Wiswall and Zafar, 2015), we expect older students to be more inclined to choose them. Since we control flexibly for current age in the wage equation, age at college entry—when majors are chosen—should be unrelated to wages after graduation.

We provide several pieces of evidence to support the plausibility of the conditional independence and exclusion assumptions for our proposed instruments. In particular, we show that relative distance is only weakly associated with individual and parental characteristics reflecting abilities and tastes for STEM, and partial correlations vanish once regional control variables are included. This suggests that relative distance is unlikely to be confounded by unobserved regional heterogeneity arising from geographic sorting of families or nonrandom university locations (see, e.g., Mountjoy, 2022 for similar reasoning). Additionally, we conduct a placebo analysis using the progressivism index of Osikominu et al. (2020) as an alternative instrument. This enables us to empirically test whether our baseline instruments have a direct effect on wages in the second stage of a two-stage least squares regression—which is not the case. As a further robustness check, we exclude observations from the most densely populated and the most rural municipalities to ensure that our results are not driven by variation in relative distance associated with spatial heterogeneity within local labor markets.

The estimation of the return to a STEM education is implemented in two steps. In the first step, we estimate the reduced-form choice model of STEM major choice, where the set of explanatory variables includes on the one hand the earnings determinants and on the other hand the instrumental variables. Our results for the reduced-form choice model suggest that, as the relative distance to the nearest technical university increases and with it the relative cost of a STEM education, the probability of choosing a STEM major decreases significantly. Further, older individuals at college entry are significantly more likely to prefer a STEM major over a non-STEM major. When examining instrument strength in the first stage, we find effective F-statistics (Olea and Pflueger, 2013) well above the rule-of-thumb threshold and generalized critical values.

In the second step, we estimate the conditional expectation functions (CEFs) of the potential outcomes and the MTEs, which capture the average effect of a STEM education for individuals who are indifferent between STEM and non-STEM fields, and plot them against the percentiles of the unobserved resistance to STEM. The CEF for potential non-STEM wages is significantly increasing in the percentiles of the unobserved resistance to a STEM education, while it is significantly decreasing for potential STEM wages. Individuals with low (high) resistance to choosing a STEM major tend to have low (high) potential wages after graduating from a non-STEM field but high (low) potential wages after graduating from a STEM field. This heterogeneous pattern is likely driven by differences in skills: low-resistance individuals excel in areas where high-resistance individuals struggle, and vice versa (e.g., logical vs. social skills), with these skills being valued differently in the typical jobs of STEM and non-STEM graduates. This suggests that individuals select their major based on their comparative advantage, consistent with findings from, e.g., Kirkebøen et al. (2016) and Arcidiacono et al. (2020).[3]

This pattern of positive selection on returns translates into a downward-sloping MTE curve, obtained as the difference between the CEFs of potential STEM and non-STEM wages. Our MTE curve suggests marginal treatment effects between more than 57% (45 log points) for low-resistance individuals and less than −42% (−54 log points) for high-resistance individuals. Summarizing this heterogeneity in common global treatment parameters by computing weighted averages of the MTE curve, we find an average treatment effect on the treated (ATT) of 36% (31 log points) and an average treatment effect on the untreated (TUT) of −15% (−16 log points). The average treatment effect (ATE) is essentially zero.[4] We further show that returns are not only heterogeneous with respect to unobserved characteristics, but also across observed earnings determinants. In particular, the return from graduating from a STEM field is 8 log points lower for women than men (see also Daymont

[3]Mourifie et al. (2020) reject self-selection based on potential earnings for some parts of the population in Germany and Canada. However, their classification of STEM fields differs from ours, which follows that of the OECD. Institutional differences across countries, such as a lower intensity of gender profiling in Switzerland, could also contribute to explaining their differing results.

[4]For comparison, we also perform 2SLS regressions using our set of instruments and the propensity score from the reduced-form choice regression as instruments, respectively. The 2SLS estimates suggest a local average treatment effect (LATE) of 14 to 16 log points.

and Andrisani, 1984; Buffington et al., 2016; Imberman et al., 2026). Further, individuals who specialized in math and science during high school have a 24 log points higher return of graduating from a STEM field (see also Joensen and Nielsen, 2016; Saltiel, 2023). These results are robust to various sensitivity checks: our estimates remain unchanged when we allow for more flexible specifications of unobserved heterogeneity, apply further sample restrictions, or use alternative instruments.

Finally, we evaluate the effectiveness of counterfactual policies aiming at promoting STEM education by calculating PRTEs from our MTE curve. Our results suggest that policies are only successful (i.e., increase STEM enrollment and have a positive average effect on the earnings of those affected) if they are able to target individuals with low predicted probabilities to major in STEM based on their observed characteristics, i.e., those with a high relative monetary cost of studying STEM. The reason is that only in this group are there untreated individuals with low unobserved resistance and high pecuniary returns. Policies that are able to reduce the subjective costs of these individuals result in high pecuniary returns among those affected.

Our paper contributes to several strands of the literature. A recent literature has proposed and tested some policies promoting STEM that could increase women's enrollment in traditionally male-dominated fields (i.e., economics and STEM). As often suggested by policymakers, Canaan and Mouganie (2023) show that the gender gap in STEM fields can be reduced by direct contact with a female science advisor. Similarly, Porter and Serra (2020) conduct a randomized controlled trial and find that female role models in economics have a strongly positive effect on female students majoring in economics. Finally, Saltiel (2023) shows that policies which increase high-math-ability women's self-efficacy boost STEM enrollment and completion, leading to positive returns. We contribute to this literature by explicitly showing that successful support policies need to target a combination of observed characteristics and subjective tastes that imply high returns. We also provide examples of policies that might *not* be effective given the pattern of returns we find (e.g., generally reduced distance or subsidized public transport for everyone).

Our paper also relates to the literature studying factors explaining the gender gap in STEM. It is well known that women are strongly underrepresented in STEM majors at universities and in STEM occupations (Benbow and Stanley, 1980; Turner and Bowen, 1999; Goldin et al., 2006; Guiso et al., 2008; Kahn and Ginther, 2017). In 2019, across all OECD countries, only 20% of new entrants in IT and 26% of new entrants in engineering were women (OECD, 2021, Tab. B4.3).[5] An extensive literature has identified potential reasons for the STEM gender gap, including tastes and preferences (Zafar, 2013; Wiswall and Zafar,

[5] For Switzerland, these shares are only 13 and 19%, respectively.

2015; Ngo and Dustan, 2024; Campos et al., 2026), teachers' gender bias (Alan et al., 2018; Lavy and Sand, 2018; Carlana, 2019), gender gap in competitiveness (Buser et al., 2014, 2017), female role models (Carrell et al., 2010), cultural beliefs (Nollenberger et al., 2016), peer effects (Fischer, 2017; Bostwick and Weinberg, 2022), and false beliefs about own ability (Owen, 2023), among others. We contribute to this literature by showing that well-designed support policies explicitly targeting these factors can indeed reduce the STEM gender gap.

Finally, we contribute to the literature studying heterogeneous returns to college majors (Altonji et al., 2012, 2016; Lovenheim and Smith, 2023). While many papers document a wage advantage of STEM fields (James et al., 1989; Arcidiacono, 2004; Kinsler and Pavan, 2015; Choi et al., 2023), little attention has been paid to heterogeneity in returns with respect to unobserved characteristics.[6] One exception is Mourifie et al. (2020), who consider a Roy (1951) model for self-selection into STEM education. Their results show that not all individuals self-select into college majors based on potential earnings, but that self-selection behavior depends on gender, race, and region. The absence of self-selection behavior for some groups implies that policies aimed directly at increasing women's STEM enrollment may have negative effects on the gender wage gap. Unlike Mourifie et al. (2020), we consider a generalized Roy model (Heckman and Vytlacil, 2007a). Apart from documenting self-selection into university majors, we additionally show how the returns to a STEM education depend on observed earnings determinants and unobserved tastes. Our results on the heterogeneity of returns imply that only individuals with favorable observed and unobserved characteristics benefit on average from a STEM education, and support policies need to identify and target such individuals to be effective.

The remainder of this paper proceeds as follows. In Section 2, we describe the framework of marginal treatment effects (MTEs) and present our estimation strategy. In Section 3, we introduce the data and institutional background. In Section 4, we discuss the plausibility of the conditional independence and exclusion assumptions. In Section 5, we present our main results. In Section 6, we evaluate hypothetical policies promoting STEM education and evaluate their effectiveness by calculating policy-relevant treatment effects (PRTEs). Finally, Section 7 concludes.

[6] Seminal contributions to the literature on the returns to different college majors are Altonji (1993), Grogger and Eide (1995), Arcidiacono (2004), Hamermesh and Donald (2008), Beffy et al. (2012), Hastings et al. (2013), Kinsler and Pavan (2015), Kirkebøen et al. (2016), Bleemer and Mehta (2022), and Andrews et al. (2024), among others. Studies with a specific focus on the returns to a STEM education are Deming and Noray (2020) and Ng and Riehl (2024), among others.

2 Empirical Framework

2.1 The Marginal Treatment Effect

We adopt a generalized Roy (1951) model (Heckman and Vytlacil, 2007a) to study the returns to majoring in STEM at university. Let Y_1 and Y_0 be potential outcomes in the state with treatment and in the state without (Neyman, 1923; Fisher, 1935; Rubin, 1978). Here, Y_1 is the log hourly wage five years after graduation from a STEM field, while Y_0 is the corresponding outcome after graduation from a non-STEM field. We model the potential outcomes as follows:

$$Y_1 = \mu_1(X) + U_1, \quad \text{with } \mathbb{E}(Y_1 \mid X = x) = \mu_1(X) \tag{1}$$

$$Y_0 = \mu_0(X) + U_0, \quad \text{with } \mathbb{E}(Y_0 \mid X = x) = \mu_0(X), \tag{2}$$

where X denotes a vector of random variables that are observable to the researcher, and U_1 and U_0 are unobservable random variables. The function $\mu_d(X)$, $d = 0, 1$ corresponds to the orthogonal projection of the potential outcome Y_d on X, which implies that $\mathbb{E}(U_d \mid X = x) = 0$.[7] Let D denote a dummy equal to one if a person majors in STEM, i.e., is treated, and zero else. The observed outcome for each individual equals

$$Y = (1 - D)Y_0 + DY_1 = Y_0 + D(Y_1 - Y_0). \tag{3}$$

A student chooses to major in STEM according to the following rule:

$$D = 1 \text{ if } \mu_D(Z) - U_D > 0. \tag{4}$$

Here, $\mu_D(Z)$ is a function of Z, which denotes a vector of observed (from the perspective of the researcher) random variables determining major choice. U_D captures unobserved determinants including subjective costs and tastes for STEM relative to non-STEM, i.e., $U_D = U_C - (U_1 - U_0)$, where U_C is an unobservable subjective cost determinant of majoring in STEM vs. non-STEM and $(U_1 - U_0)$ is the unobserved idiosyncratic gain from choosing STEM over non-STEM. A low (negative) value of U_D is associated with a high unobserved gain from graduating in STEM relative to the unobserved cost. In the literature, U_D is often referred to as the unobserved resistance or distaste for treatment (see, e.g., Cornelissen et al., 2016).

The variables in Z influence the monetary return and the relative monetary cost of

[7] With this specification, the partial effect of an element of X on $\mu_d(\cdot)$ does not have a causal interpretation (see Brinch et al., 2017; Cornelissen et al., 2018, for a similar approach).

majoring in STEM. We assume that some variables in Z are excluded from X, meaning $Z \supset X$ (*exclusion assumption*). Candidate variables included in Z but excluded from X are those that specifically affect the relative monetary cost of studying STEM. Moreover, we assume that Z is independent of (U_t, U_D), $t = 0, 1$, conditional on X (*conditional independence assumption*). In terms of the generalized Roy model (Roy, 1951; Heckman and Vytlacil, 2007a; French and Taber, 2011), Eq. (4) corresponds to the reduced-form choice equation.

Following Heckman and Vytlacil (2001a, 2005), we rewrite Eq. (4) as

$$D = 1 \text{ if } P > V \,, \tag{5}$$

where $V = F_{U_D}(U_D)$ is a uniform random variable corresponding to the percentiles of U_D and $P = F_{U_D}[\mu_D(Z)]$ is the propensity score. When $P = V$ an individual is indifferent between a STEM and a non-STEM major.

Our goal is to evaluate the marginal treatment effect (MTE), defined as the average effect of graduating in STEM for individuals at the margin of indifference between a STEM and a non-STEM major, evaluated at $V = P = p$ and $X = x$:

$$\text{MTE}(X = x, V = p) = \mathbb{E}(Y_1 - Y_0 \mid X = x, V = p). \tag{6}$$

For low values of p, the MTE is the average effect of treatment for individuals who become indifferent between majoring in STEM and not at low percentiles of the unobserved resistance to STEM (accordingly, at low values of the propensity score). Hence, these are the individuals most likely to choose a STEM major.

Following Brinch et al. (2017), we impose that

$$\mathbb{E}(Y_t \,|\, X, V) = \mu_t(X) + \mathbb{E}(U_t \,|\, V)\,,\; t = 0, 1\,, \tag{7}$$

meaning that the conditional expectation functions (CEFs) of the potential outcomes (and, consequently, the MTE) are additively separable in X and V. In theory, identification of the components of the MTE requires only the exclusion restriction and independence of Z and (U_t, U_D), $t = 0, 1$, conditional on X, but not additive separability as specified in Eq. (7) (Heckman and Vytlacil, 2007b; Carneiro and Lee, 2009; Brinch et al., 2017). However, empirically it is usually not feasible to trace out the distribution of (U_t, V) given X, $t = 0, 1$, when the number of observations for each value of X is limited and/or the variation of Z at each given value of X is limited (Carneiro et al., 2011). Therefore, we rely in addition on separability of the potential outcomes, which is common in empirical applications (see, e.g., Cornelissen et al., 2018).

To obtain the MTE we separately recover the CEFs of the potential outcomes and then take their difference. According to Carneiro and Lee (2009), these are identified by

$$\mathbb{E}(Y_1 \mid X=x, V=p) = \mathbb{E}(Y \mid X=x, P=p, D=1) + p\frac{\partial\,\mathbb{E}(Y \mid X=x, P=p, D=1)}{\partial p} \tag{8}$$

and

$$\mathbb{E}(Y_0 \mid X=x, V=p) = \mathbb{E}(Y \mid X=x, P=p, D=0) - (1-p)\frac{\partial\,\mathbb{E}(Y \mid X=x, P=p, D=0)}{\partial p}, \tag{9}$$

where, for $t = 0, 1$,

$$\mathbb{E}(Y \mid X=x, P=p, D=t) = \mu_t(X=x) + \mathbb{E}(U_t \mid P=p, D=t) \tag{10}$$

and

$$\frac{\partial\,\mathbb{E}(Y \mid X=x, P=p, D=t)}{\partial p} = \frac{\partial\,\mathbb{E}(U_t \mid P=p, D=t)}{\partial p}, \tag{11}$$

because we have specified the potential outcomes as additively separable (see Eq. (7)).

By appropriately integrating the MTE over X and V, we can then recover global treatment parameters such as the average treatment effect on the treated (ATT), the average treatment effect on the untreated (TUT), and the average treatment effect (ATE) (Heckman and Vytlacil, 2001a, 2005; Carneiro and Lee, 2009).[8] For instance, to obtain the ATT, we average across MTEs giving a higher weight to MTEs at lower percentiles of the unobserved resistance to STEM (Heckman and Vytlacil, 2005). Accordingly, to obtain the TUT, we average across MTEs giving a higher weight to MTEs at higher percentiles.

2.2 Estimation Strategy

In a first step, we estimate the reduced-form choice model, see Eq. (4), with a flexible Logit regression and predict the propensity score, P, for every observation. We trim the sample to ensure that we have at least five treated and non-treated observations, respectively, for each 0.01-interval of the propensity score. Although this reduces the range over which we can evaluate the conditional expectation functions of Y_0 and Y_1, it ensures that they can be estimated precisely at every evaluation point (especially at the boundaries of the support). After trimming, we estimate the CEF of the potential treatment outcome over the empirical

[8] See a detailed description in the Appendix A.

support of $\hat{P}$ in the treated sample ($D = 1$) and, analogously, the CEF of the potential non-treatment outcome over the empirical support of $\hat{P}$ in the non-treated sample ($D = 0$).

Assuming additive separability of the potential outcomes as in Eq. (7) the CEFs of the potential outcomes can be expressed as (Carneiro and Lee, 2009):

$$\mathbb{E}(Y_t \mid X = x, V = p) = \mu_t(x) + \lambda_t(p)\,, \tag{12}$$

for $t = 0, 1$, where the functions $\lambda_t(\cdot)$ equal

$$\begin{aligned} \lambda_1(p) &= \mathbb{E}(U_1 \mid P = p, D = 1) + p\frac{\partial\,\mathbb{E}(U_1 \mid P = p, D = 1)}{\partial p} \\ \lambda_0(p) &= \mathbb{E}(U_0 \mid P = p, D = 0) - (1-p)\frac{\partial\,\mathbb{E}(U_0 \mid P = p, D = 0)}{\partial p}\,. \end{aligned}$$

We model $\mu_t(X)$ as linear in a vector of parameters β_t, i.e., $\mu_t(X) = X\beta_t$, and specify $\lambda_t(\cdot)$ as a mean-zero polynomial in the propensity score, $t = 0, 1$. We use a second order polynomial for our main specification and verify robustness to other choices (see Section 5.4).[9] Then we estimate the CEFs with separate OLS regressions of Y on X and the mean-zero polynomial of $\hat{P}$ in the trimmed treated and untreated subsamples.[10]

The functions $\lambda_0(\cdot)$ and $\lambda_1(\cdot)$—polynomials of the propensity score—correspond to control functions. Nonzero coefficients on any of the terms involving the propensity score imply that U_t and V are not statistically independent and that individuals self-select into university majors based on U_t, $t = 0, 1$.

Taking the difference between the estimated CEFs of the potential outcomes at given values of X and P, for values of P where the support among the treated overlaps with that of the untreated (i.e., common support), we obtain estimates of the MTE as a function of X and P.[11] In particular, an estimate of the MTE is obtained as

$$\begin{aligned} \widehat{\mathrm{MTE}}(X = x, V = p) &= \hat{\mathbb{E}}(Y_1 \mid X = x, V = p) - \hat{\mathbb{E}}(Y_0 \mid X = x, V = p) \\ &= x(\hat{\beta}_1 - \hat{\beta}_0) + \hat{\lambda}_1(p) - \hat{\lambda}_0(p)\,. \end{aligned} \tag{13}$$

For statistical inference, we use a weighted bootstrap at the municipality level with 499 replications and reestimate each step of the procedure in each iteration to obtain standard errors that take estimation uncertainty in the estimated propensity score into account (Barbe

[9] We also estimated $\lambda_t(p)$, $t = 0, 1$, non-parametrically. The results are very similar and available on request.

[10] For estimation, we rely on the Stata command `mtefe` provided by Andresen (2018).

[11] Note that the supports differ between treated and non-treated observations: separate supports are used to estimate the conditional expectation functions, whereas only the common support is used to estimate the MTE.

and Bertail, 1995).

2.3 Goodness of Fit

Carneiro and Lee (2009) propose to evaluate goodness of fit by comparing estimates of $\mathbb{E}(Y_0 \mid D = 0)$ and $\mathbb{E}(Y_1 \mid D = 1)$ with sample-based estimates obtained as the means of observed non-STEM and STEM wages. The former can be obtained from the conditional expectation functions by integrating out the estimates of $\mathbb{E}(Y_t \mid X, V)$, $t = 0, 1$, with some suitable weights as described in Appendix A. We will present results from this goodness-of-fit check in Table 8 within our main results section (see Section 5).

3 Background and Data

3.1 Institutional Background

Secondary education in Switzerland is organized in different tracks. Depending on their academic performance, students are either grouped into secondary schools that prepare for university studies (academic track, ending after grade 12 or 13), or schools that are more vocationally oriented (vocational track) and prepare for an apprenticeship (after grade 9) or prepare for studies of a particular field at a University of Applied Sciences (after grade 12 or 13). Only students graduating from an academic track school with a general high school degree (*Gymnasiale Matura*) can directly access a Swiss university.[12] Around a fifth of a cohort holds such a degree (Swiss Federal Statistical Office, 2015).

At the time of our study, young people with a *Gymnasiale Matura* degree could choose freely among the available programs at all universities in Switzerland. There were no university entrance exams and no restrictions in terms of grade point average or specializations chosen at high school.[13] Tuition at Swiss universities is moderate in international comparison, both in absolute terms and relative to housing costs. It varies little across universities and is constant within university, i.e., it is not major-specific.

In the 1990s and early 2000s, around 75% of people with a Swiss general high school degree enrolled in a university within a year after high school graduation (Swiss Federal Statistical Office, 2013). Graduating from university meant completing a long first degree program,

[12]High school graduates with a *Gymnasiale Matura* degree but no professional experience are not entitled to enroll in a University of Applied Sciences (*Fachhochschule*), a second type of tertiary education institution established in the mid-1990s in Switzerland. The latter offers shorter, more practically oriented and occupation-specific programs compared with those at universities.

[13]Since 1998, medical schools select students according to their grade point average at high school and using an entrance exam. Such restrictions do not apply for the cohorts in our data set.

comparable to a Master's degree in Switzerland and other developed countries nowadays. In 2002, the share of graduates from tertiary-type A programs (i.e., those programs that are typically offered at universities) to the population of the relevant age group was 17.9% of a cohort (OECD, 2004, Tab. A3.1). This figure is comparable to Germany (19.2%), Austria (18%), and France (24.2%). Drop-out rates are similar across STEM and non-STEM programs and relatively low in international comparison, which may be explained by the fact that only a rather small, well-prepared fraction of a cohort is entitled to enroll at university. 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).

In the period under study, there were eleven universities in Switzerland, all of which are publicly funded and managed. Figure 1 provides a map on the municipality level showing corresponding locations. The two technical universities, called *Eidgenössische Technische Hochschulen* (ETH), are the only federal universities in Switzerland. They are based in Lausanne and Zurich. During our study period 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 general universities are governed at the cantonal (i.e., state) level. The nine general universities are based in Basel, Berne, Fribourg, Geneva, Lausanne, Lugano, Neuchâtel, St. Gallen, and Zurich. All general universities offer degree programs in non-STEM fields, while only a few offer programs in STEM fields. However, the technical universities have larger STEM departments with a wider range of disciplines, programs, and areas of specialization than the general universities. Thus, individuals who live relatively closer to a technical university are much more likely to enroll in a STEM field.[14] In our data set, 63.9% of university graduates in STEM attended one of the two technical universities. Table B1 in the Appendix provides further information on Swiss universities.

3.2 Data

Our main data source is the 'Swiss Graduate Survey' of the Federal Statistical Office, a comprehensive survey of the full population of graduates from tertiary academic education in Switzerland (Swiss Federal Statistical Office, 2008, 2009, 2012).[15] We consider all Swiss respondents who lived in Switzerland when graduating from high school and who graduated

[14]However, doctoral degrees can be obtained and are equally common at both types of universities.

[15]These data were also used by Osikominu et al. (2020) (see Section 3 in their work). We slightly deviate from their work in harmonizing the municipality codes according to the classification valid on January 1, 2014 (instead of December 31, 2010), to calculate distance measures.

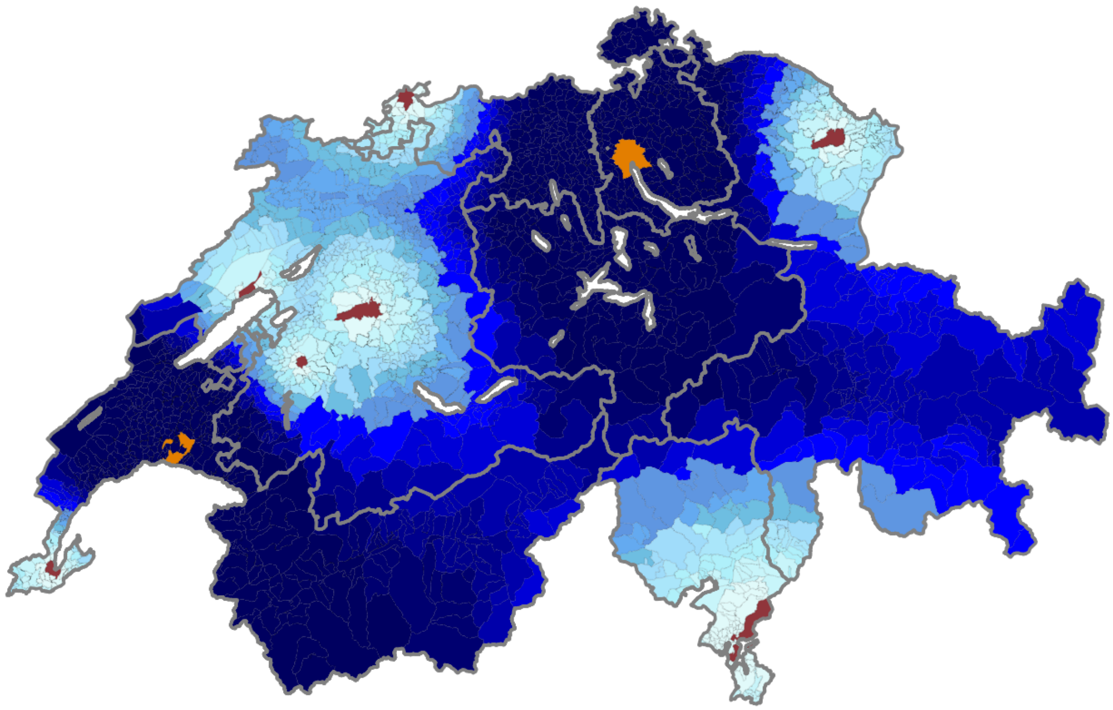


Figure 1: Swiss Municipalities and Relative Distance to Universities

Notes: The figure shows a map of Swiss municipalities. Those with a general university are colored in dark red. The municipalities of Lausanne and Zurich, marked in orange, host both a general and a technical university. Darker blue color indicates a smaller relative distance to the next technical university while brighter blue color indicates a larger relative distance. Grey color indicates NUTS-2 region boundaries.

from one of the nine general or the two technical universities in 2000 and 2002. 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. In all analyses, we use the probability weights provided by the Federal Statistical Office to account for potentially selective non-response and to ensure that the sample is fully representative.

The survey contains detailed information on respondents' demographic characteristics, study experience, and employment experience. In particular, we observe their age, sex, regular earnings and other pay components, hours worked, university major, specializations in general high school, and the level of education of mother and father. Another important piece of information in the 'Swiss Graduate Survey' is a graduate's home municipality at the end of high school as well as at the time of the survey. There are about 2,600 municipalities in Switzerland and most municipalities are small in terms of population size and area. They have an average (median) size of 17 (7.8) km^2 and an average (median) population of 2,522 (877) inhabitants in 2000. 95% of the municipalities are smaller than 59.1 km^2 and have less than 11,000 inhabitants.

We use the information on the home municipality at the end of high school to construct the Euclidean distance from the home municipality to the nearest technical university (i.e., ETH) and to the nearest general university. Further, we use the municipality of residence to match the survey data with municipality-level data from the 'Federal Population Census' (Eidgenössische Volkszählung) of 1990 and 2000. We use the 1990 Census to construct variables characterizing a respondent's economic environment at the end of high school, while the 2000 Census allows us to construct variables capturing local economic conditions at the time of labor market entry. To capture further potential heterogeneity in the economic environment, in all analyses, we also control for the broader Swiss region a student resided in when completing 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.

We follow Osikominu et al. (2020) and restrict the analysis sample to typical careers and drop observations with missing or implausible values in relevant variables. Specifically, we focus on employed graduates five years after graduation, without extreme or missing values of earnings, study duration, or age. Because the unemployment rate for tertiary graduates in Switzerland was only around 3% during our study period, we believe that our sample restriction is unlikely to affect the generalizability of our findings. In particular, there are only 9.4% nonemployed individuals in the original sample with no difference between STEM and non-STEM graduates.[16] In total, our sample consists of 4,766 individuals.

Outcome and Treatment Variable

Our outcome variable of interest is the log hourly gross wage (in CHF) five years after graduation. We compute it as total annual earnings from the main job, i.e., in either 2005 or 2007, including regular pay as well as overtime compensation and bonus payments, divided by the actual number of hours worked in that year.

Our treatment dummy takes the value one if an individual has graduated from a STEM major and zero if an individual has graduated from a non-STEM major. We define as STEM majors those offered by the two technical universities in Switzerland, i.e., physical sciences, chemistry, biology, geography, mathematics, computing, and technical sciences (primarily engineering). The remaining majors are classified as non-STEM, i.e., theology, language

[16]For further details on the sample construction, see the Online Appendix of Osikominu et al. (2020).

and literature, history and culture, other humanities, economics and business, law, human medicine, dentistry, veterinary medicine, and pharmacy.

Instrumental Variables

Our empirical strategy relies on instrumental variables that affect the decision to major in a STEM field but have no direct effect on hourly wages. Our main instrument is based on the geographic distance to the nearest technical university (i.e., ETH) relative to the geographic distance to the nearest general university. In particular, we calculate these geographic distances as Euclidean distances from the municipality centroids.[17]

The instrument is based on the ratio of the two distances. For individuals from Zurich or Lausanne, which host both a general and a technical university, this ratio equals one, the lowest possible value. This also applies to individuals for whom the nearest universities of both types are located in Zurich or Lausanne. A relative distance above one indicates that the nearest technical university is farther away than the nearest general university. For example, a relative distance of two implies that the nearest technical university is twice as far from an individual's home municipality as the nearest general university. When both distances are equal, the relative distance does not affect the relative cost of studying STEM versus non-STEM, whereas values above one indicate a higher monetary cost of studying STEM.

Importantly, a large relative distance does not necessarily correspond to a rural municipality. Rather, it often reflects large cities that host a general university but no technical university (e.g., Basel, Bern, Geneva, and Lugano). At the same time, high relative distance values are not confined to large cities: they are observed across all municipality size classes, including very small municipalities. Relative distance is low either when both types of universities are far away, which is typical of small rural municipalities but also applies to larger cities such as Lucerne, or when both are located nearby (e.g., Zurich, Lausanne, and Winterthur). Consequently, relative distance provides little information about municipality size or population density.

In the empirical analysis, we use the logarithm of relative distance to allow for nonlinear effects, recognizing that a proportional change in relative distance may have different implications depending on the initial level. For example, an increase from two to four is likely to have a different impact than an increase from four to eight. In addition, we control for various sources of regional heterogeneity, including municipality size, local employment and

[17]Results do not change if we use travel distance instead (see Osikominu et al., 2020). Because travel distance is constantly changing, which complicates replication of our results, we focus on the stable Euclidean distances instead.

industry structure, and local labor market fixed effects.

Although our results do not change if we rely on the log relative distance as sole instrument (see Section 5.4), we use a second instrument, age at college entry, to improve precision. After graduating from general high school, Swiss students have many options before they begin their studies. Some travel, do an internship, or complete voluntary military service; others take time to focus on personal interests and gather information about possible courses of study. Here, we exploit that people who are older when they start college tend to make more mature decisions, in the sense that they give higher weight to future (labor market) benefits than to immediate amenities such as a relaxed student lifestyle. In addition, they are more risk averse (Morin and Suarez, 1983). According to these considerations, older people should be more likely to major in a STEM field, which is known to lead to more stable careers (Wiswall and Zafar, 2015).

Control Variables

We consider a number of variables that appear as controls in all estimation steps. In particular, we control for sex, age, age squared, and a dummy for having specialized in math and science at high school.[18] We account for parental education using dummies indicating a college degree and vocational training, respectively. We further control for structural differences between major cities and less-populated areas using a dummy variable for municipalities with a population above 100,000.[19] In order to control for local economic conditions in an individual's home municipality, we control for the local (female) employment rate and the local employment shares of various industries in 1990, including manufacturing and construction. Finally, all estimations include region or local labor market fixed effects at the NUTS-2 level to capture potential heterogeneities across regions that may affect the decision to study a STEM field, hourly wages as well as the instrument based on distance.[20] The NUTS-2 regions can be understood as local labor markets that aggregate the granular municipalities into larger geographic units based on their economic integration and commuter flows. Table 1 provides an overview of the variables used in the empirical analysis.

[18] Card and Payne (2021) highlight the role that STEM readiness, proxied by high school math and science courses, can play a role in choosing a STEM major.

[19] Major cities in our context are Zurich and Lausanne with both a general and a technical university and Basel, Bern, and Geneva.

[20] There are seven NUTS-2 regions: Lake Geneva, Espace Mittelland, North-Western Switzerland, Zurich, Eastern Switzerland, Central Switzerland, and Ticino.

Table 1: Variable Description

Outcome	Log hourly gross wage
Treatment	= 1 if graduated from a STEM major, = 0 if graduated from a non-STEM major
Instruments	Log distance to nearest technical university relative to nearest general university; age at college entry
Control variables	
Individual characteristics	Sex; age; age squared; high school math & science specialization; cohort
Parental characteristics	University degree; vocational training degree
Regional characteristics	Major city; Employment rate; female employment rate; employment share manufacturing; employment share construction; local labor market FE (NUTS-2 regions)

Notes: This table summarizes all variables used in the empirical analysis. Major city, employment rates, and employment shares are measured in 1990 in an individual's home municipality.

Descriptive Statistics

Table 2 presents descriptive statistics of our dependent variable and the treatment dummy, the instruments, and the main control variables for the subgroups of treated and untreated individuals, respectively. In addition, results from t-tests of equality of means are indicated by p-values in column (7).

Our sample consists of 4,766 individuals of which 3,328 (70%) have a non-STEM and 1,438 (30%) have a STEM degree. Individuals with a non-STEM degree earn on average CHF 37.15 (around US$ 46) per hour five years after graduation, significantly more than individuals with a STEM degree (CHF 34.43, around US$ 42). Thus, there is a raw wage gap of CHF −2.71 (−9.83 log points) in our sample.[21] The distance from the home municipality to the nearest technical university (i.e., ETH) relative to the distance to the nearest general university is larger for non-STEM graduates than for STEM graduates. This difference is due to a significant difference in the distance to the nearest general university. Table 2 also suggests that STEM graduates are slightly older when starting their university education. For most control variables, the means differ significantly between treatment and control group, suggesting that the raw wage gap between STEM graduates and non-STEM graduates is likely contaminated by selection bias.

[21] It is not surprising that non-STEM graduates (e.g., business administration, economics, or law) earn a higher income on average given the strong banking and finance sector in Switzerland.

Table 2: Summary Statistics

	Non-STEM major			STEM major			
	(1) Obs.	(2) Mean	(3) SD	(4) Obs.	(5) Mean	(6) SD	(7) $p > \mid t \mid$
Panel A: Outcome variable							
Gross hourly wage (in CHF)	3328	37.148	13.53	1438	34.434	14.47	0.000
Log gross hourly wage (in CHF)	3328	3.557	0.35	1438	3.458	0.41	0.000
Panel B: Instrumental variables							
Distance to technical university (km)	3328	54.240	36.46	1438	52.714	36.20	0.180
Distance to general university (km)	3328	20.184	19.12	1438	22.621	19.24	0.000
Log relative distance	3328	1.228	1.44	1438	0.991	1.32	0.000
Age at college entry	3328	20.340	1.47	1438	20.480	1.26	0.000
Panel C: Control variables							
= 1 if female	3328	0.546	0.50	1438	0.261	0.44	0.000
Age	3328	31.182	1.97	1438	30.894	1.57	0.000
= 1 if math & science spec.	3328	0.164	0.37	1438	0.571	0.50	0.000
= 1 if second cohort	3328	0.562	0.50	1438	0.566	0.50	0.780
= 1 if father university	3328	0.362	0.48	1438	0.341	0.47	0.180
= 1 if father voc. training	3328	0.513	0.50	1438	0.568	0.50	0.000
= 1 if mother university	3328	0.145	0.35	1438	0.114	0.32	0.000
= 1 if mother voc. training	3328	0.613	0.49	1438	0.675	0.47	0.000

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: Columns (1) and (4) show the number of observations, Columns (2) and (5) show the mean, and Columns (3) and (6) show the standard deviation for non-STEM graduates and STEM graduates, respectively. Column (7) shows the p-value of a two-sample t-test of no difference in means between non-STEM and STEM graduates. The log difference between non-STEM and STEM wages is -0.098.

4 Conditional Independence & Exclusion Assumptions

Instruments based on geographic variation in the proximity or supply of certain facilities have a long tradition in economics and are frequently used to estimate the returns to college education.[22] While they tend to deliver strong first-stage relationships, some precautions are necessary to ensure that they do not pick up heterogeneity across regions that is also correlated with the outcome variable. In the context of estimating the returns to college education, researchers have argued that the geographic distribution of colleges in a country

[22]See, e.g., Card (1995), Kane and Rouse (1995), Kling (2001), Currie and Moretti (2003), Cameron and Taber (2004), Dee (2004), Carneiro et al. (2011), Nybom (2017), and Mountjoy (2022).

and/or location choices of parents may not be exogenous to the earnings of their children (Cameron and Taber, 2004). For instance, the location of colleges could be related to the existence or development of industrial clusters in the surrounding region, or college-educated parents could choose to locate in regions where the return to a college education is highest.

Analogous arguments in our case could be that the industry structure in the regions around the two technical universities differs from that in regions with only a general university, or that parents who pass on a taste or talent for STEM to their children prefer to locate in regions that are relatively closer to one of the two technical universities. Further, one could argue that our relative distance instrument reflects an urban-rural divide, with comparisons effectively being drawn between residents of major cities and those in surrounding or more rural areas. Such arguments would cast doubt on the validity of the conditional independence assumption, as they suggest that a distance-based instrument is potentially correlated with regional heterogeneity that also affects the outcome. To address such concerns related to the nonrandom location of universities and geographic sorting of families, we employ the following diagnostics and strategies to safeguard our estimates against violation of conditional independence.

First, we control in all estimation steps for fixed effects at the level of local labor markets. Hence, we argue that relative distance to the nearest technical university has no direct effect on wages within a given local labor market (see, e.g., Mountjoy, 2022 for a similar reasoning). Further, we control for municipality size as well as local employment and industry structure to account for potential sources of spatial heterogeneity within local labor markets, including differences in urbanization and industrial composition, that may be correlated with both the instrument and the outcome. In addition, we control for a math and science specialization in high school and for parental educational attainment. With these variables, we intend to account for abilities and tastes for STEM that could be related to our outcome variable through inter-generational transmission or geographic sorting of parents.

Second, we evaluate the balancing of our distance-based instrument across individual and parental characteristics that potentially predict wages and the return to STEM.[23] Specifically, we regress log relative distance on the individual and parental characteristics, adding regional control variables step by step. The results, presented in Table 3, allow us to assess to what extent the mean of log relative distance varies as a function of individual and parental characteristics, before and after accounting for regional heterogeneity. If log relative distance is not or only weakly associated with key variables that reflect abilities and tastes for STEM once regional heterogeneity is accounted for, there is little scope for the instrument to pick up neglected regional heterogeneity that could bias our estimates of the pecuniary return

[23] For more information on balancing tests see, e.g., Pei et al. (2019) and Mountjoy (2022).

Table 3: Balancing of Log Relative Distance Instrument

	(1)	(2)	(3)	(4)	(5)
= 1 if math & science spec.	0.081	0.072	0.078*	−0.023	0.007
	(0.052)	(0.046)	(0.047)	(0.035)	(0.034)
= 1 if father university	−0.041	−0.050	−0.059	0.073	0.009
	(0.072)	(0.071)	(0.067)	(0.051)	(0.047)
= 1 if father voc. training	−0.142*	−0.075	−0.066	−0.047	−0.023
	(0.086)	(0.075)	(0.072)	(0.054)	(0.050)
= 1 if mother university	0.070	0.055	0.037	0.133*	0.067
	(0.080)	(0.078)	(0.076)	(0.069)	(0.062)
= 1 if mother voc. training	−0.098	−0.066	−0.075	−0.059	−0.067
	(0.061)	(0.058)	(0.056)	(0.044)	(0.041)
Employment shares		✓	✓		✓
Industry shares			✓		✓
Region FE				✓	✓
Observations	4766	4766	4766	4766	4766

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows results from OLS regressions of the log distance to the nearest technical university relative to the nearest general university on a set of controls. All regressions control in addition for age and its square, a female and a cohort dummy. Standard errors robust to clustering at the municipality-level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

to STEM.[24] Column (1) of Table 3, which does not condition on any regional variables, suggests that basically no variable correlates with our instrument. This pattern remains stable with point estimates shrinking even more when controls for local employment, industry structure, and region fixed effects are added. Interestingly, the results in Table 3 suggest no association between the indicator for a math and science specialization in high school and log relative distance, which is remarkable because the high school specialization dummy is an important predictor of major choice, wages, and the return to STEM, see Tables B2 and B3 in the Appendix. Overall, these findings suggest that key individual characteristics reflecting abilities and tastes for STEM vary only weakly with relative distance and that our controls for regional heterogeneity effectively address concerns that distance-based instruments may capture confounding regional heterogeneity.

To ensure that our second instrument, age at college entry, is not associated with field-

[24] While conditional balancing of log relative distance across individual and parental characteristics supports our identification strategy, any remaining imbalances do not imply that the instrument is not valid since the individual and parental characteristics examined in the balancing exercise are included as controls when we estimate the return to STEM.

specific wage premia, we control in all estimation steps for age and age squared at the survey date, i.e., the time at which the wage is measured. In doing so, we are able to account for the potential direct effect of later college entry on earnings. However, in a sensitivity analysis presented in Section 5.4, we show that our baseline estimation results of the return to STEM are very similar if we instead rely on the relative distance as the sole instrument.

Finally, to test the validity of conditional independence and exclusion for both of our instruments, we rely on an alternative instrument originally proposed by Osikominu et al. (2020). Specifically, they constructed a progressivism index that captures attitudes towards science and gender issues present in an individual's home municipality at the end of high school. Their research shows that an individual's sociocultural background, captured by the progressivism index, affects their preferences for university majors, but is unrelated to wages. We use this index in a placebo IV analysis, where it serves as instrument that enters only in the first stage, while our primary instruments—log relative distance and age at college entry—are included in both estimation stages. This setup allows us to test whether our primary instruments have a direct effect on wages. Results are presented in Table 4.[25] While log relative distance and age at college entry remain highly relevant for the STEM major choice after the inclusion of the progressivism index as instrument, we can show that neither variable has any direct effect on wages in the second stage.[26]

5 Evaluating the Returns to STEM

5.1 Preliminary Evidence

In a first step, we estimate a Logit regression of STEM major choice on the excluded instruments and the set of control variables. Table 5 presents the average partial effects (APE) for the excluded instruments in Column (1). Our instruments strongly explain STEM major choice. In particular, an increase in the relative distance to the nearest technical university and, thus, rising costs for a STEM major by 10% decreases the probability of choosing a STEM major by 0.25 percentage points on average. This effect is not only highly statistically significant but also economically relevant. A 50% increase in the relative distance

[25]The slightly smaller sample size is due to missing observations in the progressivism index for five municipalities.

[26]We further estimate the MTE schedule using this progressivism index as excluded instrument, and present the results in Figure B6 in the Appendix. The MTE curve also slopes downwards, much like our preferred specification, but is slightly flatter due to less heterogeneity in the conditional expectation function for Y_1. The corresponding IV estimate (i.e., local average treatment effect (LATE)) differs somewhat from our main specification in Table 5, because compliers to changes in the progressivism index are different from compliers to changes in the other instruments: they have a higher unobserved resistance on average and, thus, lower returns to STEM (IV weights are not reported).

Table 4: Placebo Test of Exclusion Restriction

	First stage	Second stage
	(1) APE	(2) IV ($\hat{P}$)
= 1 if STEM major		0.147
		(0.093)
Progressivism index	-0.027^{***}	
	(0.009)	
Log relative distance	-0.022^{***}	0.007
	(0.005)	(0.006)
Age at college entry	0.073^{***}	0.001
	(0.008)	(0.009)
Control variables	✓	✓
Region FE	✓	✓
Effective F-stat for test for weak instruments		327.67
Observations	4761	4736

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: Column (1) shows average partial effects (APE) from a Logit regression of graduating from a STEM major against a non-STEM major on the progressivism index of Osikominu et al. (2020) and the instruments Z defined in Table 1. Column (2) shows results from a 2SLS regression of the log hourly gross wage on graduating from a STEM major using the estimated propensity score $\hat{P}$ from the Logit regression in the first stage as instrument. Effective F-statistics are based on the approach by Olea and Pflueger (2013). The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, in Column (2). Standard errors are robust to clustering at the municipality-level and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

corresponds, for instance, to a 50 km increase in the distance to the nearest technical university when this distance was 100 km, holding the distance to the nearest general university constant. Such an increase would lead to a reduction in the probability of choosing a STEM major by 1.25 percentage points. The coefficient on age at university start suggests that older individuals are more likely to choose a STEM major over a non-STEM major. This holds true even after controlling for the actual age at the survey date. In particular, the probability of choosing a STEM major increases by 7.4 percentage points on average when being one year older at the time of college major choice.

In a second step, we perform two-stage least squares (2SLS) regressions using two different instrument specifications: first, the variables in Z, and second, the predicted propensity score $\hat{P}$ derived from the Logit regression for major choice. Summary results of these regressions can be found in Columns (2) and (3) of Table 5, while the full estimation results are presented

Table 5: Preliminary Estimates

	First stage	Second stage	
	(1) APE	(2) IV (Z)	(3) IV ($\hat{P}$)
Log relative distance	-0.025^{***}		
	(0.005)		
Age at college entry	0.074^{***}		
	(0.008)		
$= 1$ if STEM major		0.162^{**}	0.143^{**}
		(0.077)	(0.065)
Control variables	✓	✓	✓
Region FE	✓	✓	✓
Effective F-stat for test for weak instruments		183.40	451.16
Observations	4766	4631	4631

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: Column (1) shows average partial effects (APE) from a Logit regression of graduating from a STEM major against a non-STEM major. Columns (2) and (3) show results from 2SLS regressions of the log hourly gross wage on graduating from a STEM major. The specification in Column (2) uses the set of instruments Z defined in Table 1 in the first stage. The specification in Column (3) uses the estimated propensity score $\hat{P}$ from the Logit regression in the first stage as instrument. Full results are presented in Table B3 in the Appendix. Effective F-statistics are based on the approach by Olea and Pflueger (2013). The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support in Columns (2) and (3). Standard errors are robust to clustering at the municipality-level and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

in Table B3 in the Appendix. The IV estimates correspond to local average treatment effects (LATE) that capture the average effect of graduating from a STEM major for those individuals who are induced to major in STEM when the instruments change to values that make the choice of STEM more likely, the so-called compliers. Our findings suggest that these compliers gain on average in terms of wages. According to the estimates in Column (2), that use the log relative distance and age at college entry as instruments, STEM graduates earn on average 18% (16 log points) more. The effect is very similar (15% or 14 log points) if we instead use $\hat{P}$ as excluded instrument (see Column (3)). These LATEs may differ substantially from common treatment parameters such as the ATE or the ATT in the presence of heterogeneous treatment effects. Our analysis of the MTEs will allow us to systematically explore the heterogeneity of treatment effects as a function of unobserved resistance and observed characteristics.[27]

We follow Andrews et al. (2019) and compute the effective F-statistic of Olea and Pflueger (2013) to test for weak instruments in the context of overidentified and non-homoscedastic models with one endogenous variable.[28] These effective F-statistics are 183 and 451, respectively, well above the rule-of-thumb threshold of 10 (Staiger and Stock, 1997), but more importantly, also well above the generalized critical values provided by Olea and Pflueger (2013).

We also test the sensitivity of our IV regression results to different sample restrictions. Results are presented in Table B3 in the Appendix.[29] Although the effects increase slightly when dropping Zurich and Lausanne, when border municipalities are omitted, or when small or large municipalities are omitted, respectively, our estimates are very similar across specifications, providing strong support for the robustness of our results.[30]

5.2 Marginal Treatment Effects

In the next step, we estimate the conditional expectation functions (CEFs) of the potential outcomes, following the approach outlined in Section 2, and use them to construct the MTE curve for percentiles of the unobserved resistance within the range of common support. At this stage, we use the trimmed sample for estimation.[31] The interval of common support for which we can identify the MTE curve is $[0.04; 0.81]$. While this range is narrower than the full unit interval, our results still allow us to draw meaningful conclusions for a highly relevant segment of the population: individuals with low to medium-high resistance to STEM fields, who are typically the focus of initiatives promoting STEM education.

Figure 2 presents the estimated CEFs of potential wages associated with graduating from a non-STEM or STEM major. While the CEF of potential non-STEM wages (Panel (a)) is increasing almost linearly in the unobserved resistance V, the CEF of potential STEM wages (Panel (b)) is decreasing. This suggests that individuals with a low resistance to a STEM education have low potential wages when graduating from a non-STEM major, but high potential wages when graduating from a STEM major. Conversely, individuals with a

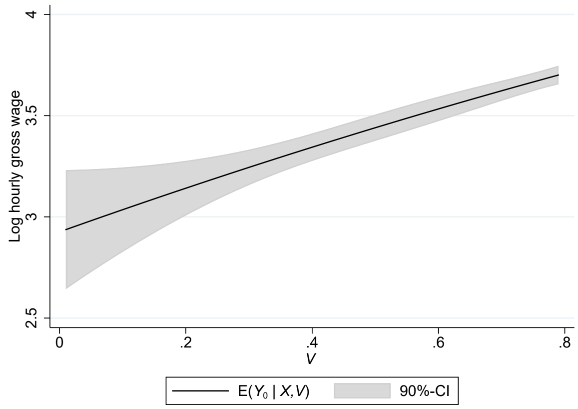
[27] Figure B2 in the Appendix shows the IV weights used to aggregate the MTE curve to the 2SLS estimate. These weights are highest for individuals with a low unobserved resistance to STEM and high returns (see Figure 3).

[28] In the just-identified case, the effective F-statistic is equivalent to the commonly used robust F-statistic. For implementation, we rely on the Stata command `weakivtest` provided by Pflueger and Wang (2015).

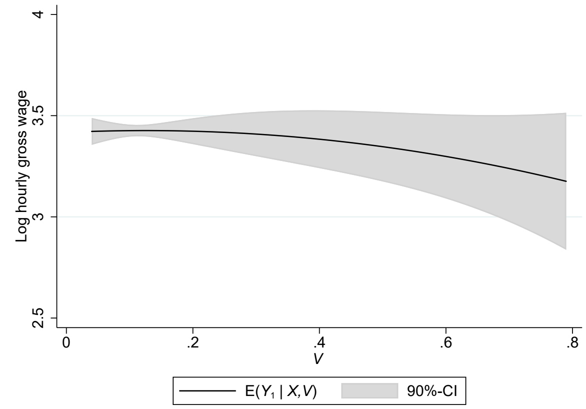
[29] We perform the same robustness checks for the first-stage Logit regressions. Results are presented in Table B2 in the Appendix.

[30] The intuition behind these robustness checks will be explained in more detail in Section 5.4.

[31] Figure B1 in the Appendix plots the distribution of the estimated propensity score. Trimming reduces the sample size by 11 treated and 3 nontreated observations. Imposing common support leads to a further reduction by 121 observations, resulting in a final sample size of 4,631.



(a) $\mathbb{E}(Y_0 \mid X, V)$



(b) $\mathbb{E}(Y_1 \mid X, V)$

Figure 2: Conditional Expectation Functions

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: These figures show the estimated CEFs of Y_0 in Panel (a) and of Y_1 in Panel (b) together with their 90% confidence intervals. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively. Confidence intervals are obtained from a municipality-level weighted bootstrap with 499 replications.

high resistance to a STEM education have high potential wages when graduating from a non-STEM major, but low potential wages when graduating from a STEM major. This pattern is consistent with the idea that low- and high-resistance individuals excel in different skills (e.g., logical skills vs. social skills) that in turn are rewarded differently in typical STEM and non-STEM jobs. Hence, individuals favoring a STEM field over a non-STEM field benefit from a comparative advantage in STEM and vice versa for those favoring a non-STEM field.

The visual impression of the two CEFs suggests nonconstant profiles for both, implying that people self-select into university majors based on their unobserved major-specific strengths. We can test this impression formally with a statistical test of the null hypothesis that the coefficients on the terms involving the propensity score in the $\lambda_t(\cdot)$-function, $t = 0, 1$ in Eq. (12) are jointly equal to zero. Our results suggest that we can reject the null hypothesis of a flat CEF for both potential non-STEM and potential STEM wages at any common significance level as the χ^2-statistics are 25.8 and 10.8, respectively, with corresponding p-values of close to 0.

Taking the difference between the CEF of potential STEM wages and that of potential non-STEM wages evaluated at the means of the observed characteristics and at various percentiles of the unobserved resistance we obtain the MTE curve displayed in Figure 3. It shows the average wage return to a STEM education for individuals with mean observable characteristics who are indifferent between a STEM and a non-STEM education at differ-

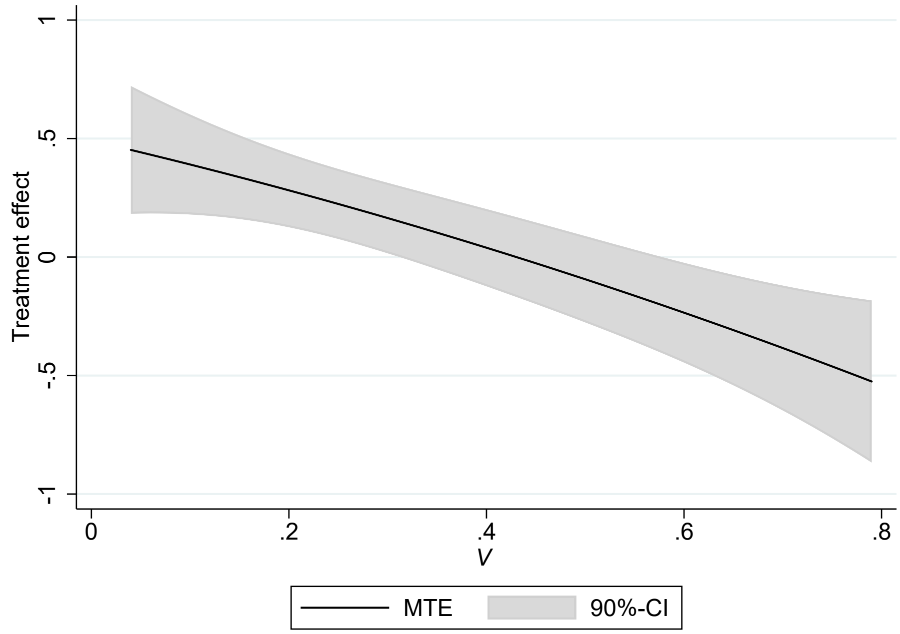


Figure 3: MTE of a STEM Education on Wages

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This figure shows the MTE curve obtained by the difference between the conditional expectations functions of Y_1 and Y_0 and its 90% confidence interval. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support. Confidence intervals are obtained from a municipality-level weighted bootstrap with 499 replications.

ent quantiles of unobserved resistance V. The MTE curve is downward-sloping, and the marginal treatment effects vary from more than 57% (45 log points) to less than -42% (-54 log points). Only individuals with a low resistance to a STEM education (i.e., a value of unobserved resistance less or equal than the 43rd quantile, $V \leq 0.43$) gain in terms of wages, while individuals with a medium to high resistance ($V > 0.43$) lose on average.

From the marginal treatment effects we can obtain three standard global treatment parameters: the average treatment effect on the treated (ATT), the average treatment effect on the untreated (TUT), and the average treatment effect (ATE). They are computed as weighted averages of the marginal treatment effects, where each parameter uses a different weighting scheme presented in Figure B2 in the Appendix.[32] It is important to emphasize that the average individual characteristics also differ somewhat between non-STEM graduates (TUT), STEM graduates (ATT), and the full sample (ATE); this is reflected in the vertically shifted MTE curves used for integration (see Panel (b) of Figure B2 in the Ap-

[32]The weights are rescaled such that they sum up to one over the interval of common support. For more details, see also Appendix A.

Table 6: Average Treatment Parameters

	(1) $\widetilde{\text{ATT}}$	(2) $\widetilde{\text{ATE}}$	(3) $\widetilde{\text{TUT}}$	(4) $\widetilde{\text{ATT}} - \widetilde{\text{TUT}}$
Treatment effect	0.308***	−0.013	−0.155	0.464***
	(0.092)	(0.086)	(0.112)	(0.130)

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows estimates of average treatment parameters obtained from the MTE curve (see Figure 3) in Columns (1) to (3). The weights used to estimate the ATT, ATE, and TUT are based on the description in Appendix A and are presented in Figure B2 in the Appendix. Column (4) shows the difference between the ATT and TUT. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support. Standard errors are obtained from a municipality-level weighted bootstrap with 499 replications and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

pendix). Since we do not have full common support, we label our estimated parameters $\widetilde{\text{ATT}}$, $\widetilde{\text{ATE}}$, and $\widetilde{\text{TUT}}$ to distinguish them from the non-identified population parameters. Our results in Table 6 show that an individual randomly selected from the population of STEM graduates earns on average 36% (31 log points) more after treatment (i.e., ATT), with the effect being highly statistically significant. In contrast, an individual randomly selected from the population of non-STEM graduates shows a negative return to a STEM education (i.e., TUT) of 15% (16 log points). The ATE is essentially zero. Column (4) shows the difference between ATT and TUT. It is positive and statistically significant at any common significance level. In line with the visual impression from Figure 3 and results from a test of essential heterogeneity discussed before, this finding suggests that individuals self-select into university majors based on their unobserved idiosyncratic gain (i.e., the difference between the unobservables affecting potential STEM and non-STEM wages).

Our MTE curve illustrates the heterogeneity of returns to a STEM education across the percentiles of unobserved resistance to a hypothetical individual with average values of the observed characteristics. Since no individual in our sample has exactly average observable characteristics, it is important to examine heterogeneity with respect to these characteristics. In this way, we gain a more complete understanding of who benefits from STEM education and who does not. In Table 7, we present separate ATT estimates for men and women, as well as for individuals who specialized in math and science in high school and those who did not. Our findings suggest that male STEM graduates earn on average 39% (33 log points) more than in the counterfactual without STEM, corresponding to a return roughly 8 log points higher than that for women. Our finding of a gender gap in returns for women in

Table 7: Heterogeneity Based on Observed Characteristics

	Sex		Math & science spec.	
	(1) Male	(2) Female	(3) No	(4) Yes
$\widetilde{\text{ATT}}$	0.327***	0.248***	0.169*	0.414***
	(0.094)	(0.092)	(0.093)	(0.100)

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows estimates of the ATT obtained from the MTE curve for different subgroups. The weights used to estimate the ATT are based on the description in Appendix A and are presented in Figure B2 in the Appendix. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support. Standard errors are obtained from a municipality-level weighted bootstrap with 499 replications and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

STEM is in line with previous research for the United States, documenting that women earn significantly less than men in STEM (Daymont and Andrisani, 1984; Buffington et al., 2016; Imberman et al., 2026). Another interesting pattern emerges for math and science specialization during high school. While the ATT for people with such a specialization in high school is 51% (41 log points), it is only 19% (17 log points) for STEM graduates without a prior math and science specialization. Overall, our findings suggest that the heterogeneity in returns across observed earnings determinants is economically important, but somewhat weaker than the heterogeneity across unobserved earnings determinants.

To complete the analysis of selection into STEM based on gains, it is useful to contrast the effects of sex and math and science skills on the return to STEM with those on the probability to major in STEM. The estimation results of the reduced-form choice model in Table B2 in the Appendix suggest that women are 12 percentage points less likely than men to major in STEM and that people with a math and science specialization are 32 percentage points more likely to major in STEM than those without. Given that women have lower returns to STEM and people with a math and science specialization have higher returns, this suggests positive selection on gains not only on unobserved but also on these observed earnings determinants.

5.3 Goodness of Fit

In Table 8, we present results from a goodness-of-fit check of our MTE curve. In particular, we follow Carneiro and Lee (2009) and compare the sample means of non-STEM and STEM wages (Column (1)), respectively, with the weighted averages of our estimated CEFs

Table 8: Goodness of Fit

	(1) Sample	(2) Model (MTE)
Estimate of $\mathbb{E}(Y_0 \mid D = 0)$	3.556	3.553
	(0.006)	(0.010)
Estimate of $\mathbb{E}(Y_1 \mid D = 1)$	3.458	3.430
	(0.011)	(0.016)

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows estimates that are used for a goodness-of-fit evaluation. Column (1) is based on sample estimates. Column (2) is based on the MTE and the CEFs for Y_0 weighted by TUT weights and Y_1 weighted by ATT weights (see Figure B2 in the Appendix). The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support. Standard errors are given in parentheses. In Column (2), standard errors are obtained from a municipality-level weighted bootstrap with 499 replications.

(Column (2)). In particular, we weight the CEF for non-STEM graduates with the TUT weights and the CEF for STEM graduates with the ATT weights presented in Figure B2 in the Appendix. The resulting estimates are essentially identical to the sample means, providing supporting evidence for the correct specification of our MTE curve.

5.4 Sensitivity Analysis

In order to evaluate sensitivity of our main results, we provide three sets of robustness checks in Figures B3, B4, and B6 in the Appendix.

Different Functional Form

As a first check, we evaluate the robustness of the estimated CEFs of potential outcomes and the corresponding MTE curve to the way we approximate the conditional expectation function of U_t given $V = p$ and its derivative. Our main specification models the control function $\lambda_t(\cdot)$, $t = 0, 1$, as a second degree polynomial in the estimated propensity score. We reestimate the CEFs of potential wages using a linear and a cubic specification instead and present the resulting curves in Figure B3 in the Appendix. Whether we estimate a linear, quadratic, or cubic specification in the estimated propensity score does not change our results. The CEFs of potential non-STEM wages differ slightly for low values of V, but are very close for intermediate and high values. The CEFs of potential STEM wages are virtually the same. As a result, the MTE curves are essentially identical, thereby suggesting robustness of our results.

Different Sample Restrictions

Next, we provide evidence that our results are not driven by specific types of municipalities. First, we address the possibility that the municipalities of Zurich and Lausanne, where both a technical and a general university are located, have a disproportionate influence on the shape of the estimated CEFs and MTE curve. Second, since Switzerland is surrounded by several countries with open borders, one may argue that individuals from border municipalities might be more attracted to universities in surrounding countries. If this emigration pattern is selective, and stronger or weaker prospective students in terms of STEM ability are more likely to remain in Switzerland, our estimates could be biased. To address both concerns, we reestimate the CEFs after excluding Zurich and Lausanne and, separately, after excluding border municipalities; the resulting curves are shown in Figure B4. The CEFs for non-STEM graduates are very similar across samples, although they differ slightly for low values of V. The CEF for STEM graduates is less steep when border municipalities are excluded, leading to a small deviation for $V > 0.4$. Nevertheless, the MTE curves for the full and restricted samples are very similar and show the same downward-sloping pattern. Thus, we can conclude that our results are neither driven by Zurich and Lausanne nor sensitive to the inclusion of individuals who grew up in border municipalities.

We further examine whether our findings reflect differences between highly populated municipalities and less populated municipalities. This is relevant because the relative distance instrument could, in principle, partly capture spatial heterogeneity within local labor markets. We therefore reestimate the CEFs and MTE curve after excluding small municipalities with fewer than 2,000 inhabitants and, separately, after excluding large municipalities with more than 50,000 inhabitants. Figure B5 shows that the resulting CEFs and MTE curves are very similar to the baseline estimates. Excluding large municipalities leaves both the CEFs and the MTE curve almost unchanged. Excluding small municipalities leads to somewhat stronger deviations in the CEFs, especially at lower and higher values of V, but the resulting MTE curve continues to exhibit the same overall shape and a clear downward slope. These findings indicate that the main pattern of positive selection on gains is not driven by either very small or highly populated municipalities.

Alternative Instruments

Finally, we test sensitivity of our main results with respect to the choice of instruments. In particular, we reestimate the CEFs and the MTE curve for specifications in which we use the progressivism index of Osikominu et al. (2020) and the relative distance as the sole instrument, respectively. We present results in Figure B6 in the Appendix. Although the

CEFs for non-STEM graduates are very similar, the CEFs for STEM graduates are slightly increasing, resulting in less steep MTE curves. Nevertheless, the main finding of a downward-sloping MTE curve and heterogeneity based on unobserved characteristics remains true for these specifications as well.

6 Exploring Alternative Policies to Promote STEM

Many countries are attempting to increase the number of STEM graduates by implementing initiatives aimed at expanding access to STEM education, promoting interest in STEM, and providing financial support, because people trained in STEM significantly contribute to innovation and economic growth (Hunt and Gauthier-Loiselle, 2010; Peri et al., 2015). Although several studies have documented that individuals with STEM degrees tend to earn relatively high wages on average (Altonji et al., 2012; Kirkebøen et al., 2016), our results highlight that the return to a STEM education varies strongly along observed and unobserved individual-specific factors. Therefore, it is likely that the effectiveness of an initiative will depend on who it attracts to a STEM education. If an initiative targets people with unfavorable characteristics that are associated with low or negative returns, the impact of the policy on those affected may be low or negative. In contrast, if an initiative targets people with favorable characteristics, it may have important positive effects on these individuals. Standard global treatment parameters such as the ATE, ATT, and TUT offer limited insights into which of a set of potential policies would be most effective and why (Heckman and Vytlacil, 2001b). In contrast, the policy-relevant treatment effect (PRTE), which can be calculated from the MTE, allows one to track who responds to a policy change of interest and what the impact of that policy change is on those responding.

In this section, we use the PRTE framework to evaluate counterfactual policies, such as expanding local STEM capacity, lowering students' monetary costs, providing information and mentoring, or offering targeted encouragement to certain groups, that may shift students into STEM. In particular, we consider PRTEs for policy changes that perturb the distribution of the propensity score, while leaving the potential outcomes, and in particular their unobserved components (U_0, U_1), unaffected.[33] Specifically, let D' denote the treatment indicator under the alternative policy, where $D' = 1$ if $P' > V$ and P' is the propensity score under the alternative policy. $Y' \equiv D'Y_1 + (1 - D')Y_0$ is the observed outcome under the

[33] Although most of our proposed policies only affect a small number of individuals, some envision changes to the propensity score that potentially induce general equilibrium effects. For simplicity, however, we focus on the direct, microeconomic effects and assume that the MTE schedule is policy invariant (Heckman and Vytlacil, 2007b).

alternative policy. The PRTE is defined as

$$\text{PRTE}(x) = \mathbb{E}(Y' \mid X = x) - \mathbb{E}(Y \mid X = x) = \int_0^1 \text{MTE}(x, v)\omega_{\text{PRTE}}(v)dv, \tag{14}$$

where the weight $\omega_{\text{PRTE}}(v)$ reflects the change in the fraction of individuals choosing a STEM education in response to the policy change at a given v.[34] We evaluate the PRTE at the means of the observed characteristics of the compliers to the alternative policy which may differ from those across the total population (used to evaluate the MTE schedule under the baseline policy), potentially leading to a vertical shift of the MTE schedule under the alternative policy (see Figure 4).

We use the PRTE to estimate the average effect of counterfactual policies that aim to increase STEM enrollment. Therefore, we represent these policy margins as shifts in the propensity to choose STEM, i.e., $\Pr(P' > v) \geq \Pr(P > v)$ for $v \in (0, 1)$, and then evaluate the average pecuniary return to STEM for the additional individuals shifted into STEM. This exercise allows us to gain insights into the effectiveness of different policies and determine which are most successful in increasing enrollment while implying positive pecuniary returns for those affected at the same time. We consider two sets of policies: three standard policies (A, B, C) that improve access to and interest in STEM education in an untargeted way; and three policies (D, E, F) specifically geared toward individuals with potentially high returns to STEM, and women, who are often underrepresented in STEM. It is important to emphasize that these counterfactuals should be interpreted as policy margins induced by classes of interventions, not as reduced-form estimates of a specific program.

Policy A captures a place-based expansion of local STEM capacity that reduces the relative distance to the closest technical university for some individuals and is, thus, supposed to increase STEM enrollment by reducing the relative cost for these individuals. This policy experiment assumes that a third technical university was founded in Lucerne, a city in central Switzerland. It therefore modifies a real policy reform: in 2000, the small, specialized institute of higher education in Lucerne became a full general university.[35] Instead, Policy A assumes that the university had been founded as a technical university, thereby reducing the relative distance to the nearest technical university for individuals living in the central and southern parts of Switzerland.[36] This supply-side policy allows us to test whether an

[34] See Appendix A for a description of the weights.

[35] Starting as a theological institute and later expanding to include an institute of philosophy and a history department, it now comprises six faculties. Individuals in our sample graduated from a university in 2000 or 2002 and were, therefore, unable to enroll at the University of Lucerne at the time they started their university education.

[36] Figure B7 in the Appendix shows which municipalities are affected by this policy.

additional technical university, and thus an improvement in the accessibility of a STEM education, benefits those individuals who are induced to change their university major in response to the policy.

Policy B captures a broad-based, untargeted encouragement policy, that increases STEM enrollment to achieve an overall STEM share of 39%, corresponding to Germany's share in 2019, the OECD country with the highest STEM share in tertiary education (OECD, 2021, Tab. B4.3). Specifically, to achieve an overall STEM share of 39%, we add the constant 0.615 to the index function of the propensity score ($\mu_D(Z)$ in Eq. (4)). This policy change can be interpreted as a location shift of the distribution of the subjective cost component of unobserved resistance (the unobservable U_C in Section 2). A general encouragement policy that uniformly increases young people's interest in STEM-related topics, such as a national STEM awareness campaign, could achieve such a location shift in STEM distaste and, hence, a higher likelihood of enrolling into STEM education at universities.

Finally, *Policy C* captures uniform reductions in the monetary cost of accessing STEM programs (i.e., the cost of living relatively far away) by halving the estimated coefficient on the relative distance measure in the model for major choice and recalculating the propensity score accordingly. Policies that could mitigate the negative impact of distance include subsidies for public transport or affordable housing for students in Zurich and Lausanne (i.e., the cities with a technical university).

Panel (a) of Table 9 presents the results of the untargeted policy experiments. Since the identification of the PRTE requires full common support (see, e.g., Heckman et al., 2010), we will refer to our parameters estimated over the empirical common support as $\widetilde{\text{PRTE}}$, to distinguish them from parameter estimates obtained under full common support.

The establishment of the University of Lucerne as a technical university (Policy A) reduces our relative distance measure for 36% of all municipalities and 29% of all individuals in our sample, respectively (see Figure B7 in the Appendix). However, enrollment in STEM in response to the policy change hardly increases, as the mean propensity score only increases from 0.284 to 0.288 (1.5% in relative terms). The estimated PRTE for individuals who change their university major in response to this policy change is small and statistically insignificant. When the overall enrollment rate is increased to 0.39 (that is, by 37.2%) by adding a constant to each individual's index function (Policy B), the estimated PRTE is again small and statistically insignificant. Policy C, that uniformly reduces the relative pecuniary cost, has very similar effects as Policy A: it leads to a small increase in STEM enrollment by 1.5 percentage points (5.1% in relative terms), but the estimated PRTE is slightly larger and marginally statistically significant.

Why do policies that improve access to and reduce the relative cost of a STEM education

Table 9: PRTEs of Policies Promoting STEM

(a) Untargeted policies			
	(1) **Policy A:** Expanding local STEM capacity	(2) **Policy B:** General STEM encour.	(3) **Policy C:** Travel cost reduction
Δ STEM	0.004	0.106	0.015
Δ STEM (%)	1.5%	37.2%	5.1%
$\widetilde{\text{PRTE}}$	0.093	0.082	0.130*
	(0.087)	(0.078)	(0.078)

(b) Targeted policies			
	(4) **Policy D:** Targeted encour., all	(5) **Policy E:** General STEM encour., women	(6) **Policy F:** Targeted encour. women
Δ STEM	0.028	0.102	0.015
Δ STEM (%)	9.9%	35.9%	5.4%
$\widetilde{\text{PRTE}}$	0.226**	0.083	0.232**
	(0.092)	(0.092)	(0.105)

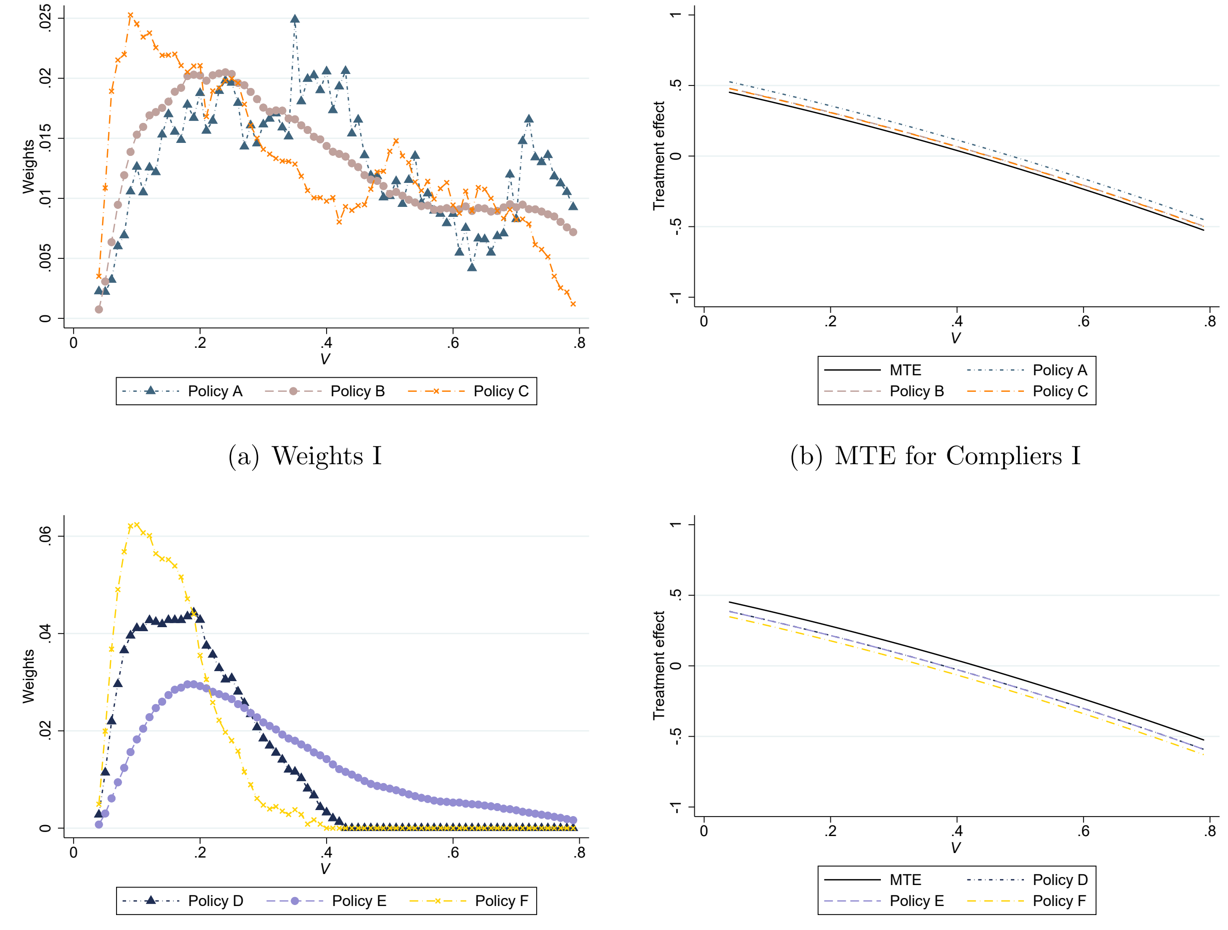
Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows PRTEs for different policies together with mean propensity scores before and after the respective policy. Policy A corresponds to a reduction in the relative distance to the closest technical university caused by the establishment of the University of Lucerne as a technical university. Policy B adds 0.615 to the index of the propensity score, i.e., $\hat{\mu}'_D(Z_i) = 0.615 + \hat{\mu}_D(Z_i)$, leading to a mean propensity score of 0.39. Policy C halves the estimated coefficient on relative distance in the model for major choice. Policy D increases enrollment of low-resistance individuals (i.e., those with positive pecuniary returns) overproportionally such that $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i) + 0.5 \times \max((0.43 - \hat{P}_i)/0.43,\ 0)$. Policy E raises the mean propensity score of women to that of men by adding 1.36 to the propensity score index of women. Policy F increases enrollment of low-resistance women (i.e., those with positive pecuniary returns) overproportionally by setting $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i) + 0.5 \times \max((0.40 - \hat{P}_i)/0.40,\ 0)$ if person i is female. "Δ STEM" refers to the change in the STEM enrollment rate, which is 0.284 at baseline. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support. Standard errors are obtained from a municipality-level weighted bootstrap with 499 replications and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

fail to significantly improve the earnings prospects of those affected? The answer can be found by tracing the observed and unobserved characteristics of those responding to the policy changes. Panel (a) in Figure 4 presents the weights used to aggregate the MTE schedule into the PRTEs of Policies A, B, and C. At a given value of V, the weight corresponds to the fraction of persons shifted into STEM at that value of V out of all persons who are shifted into STEM by the hypothetical policy change. Thus, the weights show the representation of the different percentiles of unobserved resistance among those responding to the policy change. From the analysis of the MTE schedule in Section 5.2, we know that the pecuniary return to STEM decreases in the percentiles of unobserved resistance. Policy A does not only assign a high weight to individuals with low to intermediate values of unobserved resistance, but also to individuals with high unobserved resistance, i.e., $V > 0.7$. The latter have a highly negative MTE, resulting in a low PRTE overall. Similarly, the small PRTEs of Policies B and C can be explained by the weight attributed to individuals with a high resistance to a STEM education, who show a negative MTE.

Not only the unobserved but also the observed characteristics of those responding to the policy change affect the PRTE. When evaluating the MTE curve at the observed characteristics of those who change their treatment status in response to the new policy, we see in Panel (b) of Figure 4 that people reacting to the change to Policy A have more favorable observed characteristics than those who respond to the changes to Policies B and C, which are very similar to those of the total population. The positive contribution of more favorable observed characteristics to the PRTE of Policy A, however, is outweighed by the negative contribution of worse unobserved characteristics.

We next illustrate the potential of more targeted policies aiming to promote STEM. *Policy D* simulates an increase in the propensity scores limited to individuals with positive MTE (i.e., $V \leq 0.43$). Specifically, we set $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i) + 0.5 \times \max((0.43 - \hat{P}_i)/0.43,\ 0)$. This policy should not be interpreted as a literal intervention that targets individuals on the basis of their unobserved MTE. Rather, it represents the best-case policy margin attainable through interventions that exploit information on the distribution of unobserved characteristics to target students who are underrepresented in STEM despite having strong potential to benefit from it. We discuss below how such students can be identified based on their observed characteristics. Panel (c) in Figure 4 shows that this policy successfully targets only low-resistance individuals, increasing the STEM share by 2.8 percentage points (9.9% in relative terms) and resulting in a large and statistically significant PRTE of 23 log wage points (see Table 9). Therefore, well-designed encouragement policies that target the "right" people can not only achieve their goal of increasing STEM enrollment, but also benefit those affected by the policy with respect to wages.

Since women are particularly underrepresented in STEM (e.g., Mourifie et al., 2020), they are often the focus of initiatives promoting STEM. Our results have shown that women are 12 percentage points less likely to enroll in a STEM major (see Table B2 in the Appendix) and earn 8% (8.0 log points) less after graduating from a STEM major compared to men



(a) Weights I

(b) MTE for Compliers I

(c) Weights II

(d) MTE for Compliers II

Figure 4: PRTE Weights and MTE Evaluated at Mean Characteristics of Compliers

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: Panels (a) and (c) show the weighting applied to the MTE schedule to calculate the PRTEs in Table 9, and Panels (b) and (d) the MTE schedule evaluated at the mean observed characteristics of those affected by the respective policy change. Policy A corresponds to a reduction in the relative distance to the closest technical university caused by the establishment of the University of Lucerne as a technical university. Policy B adds 0.615 to the index of the propensity score, i.e., $\hat{\mu}'_D(Z_i) = 0.615 + \hat{\mu}_D(Z_i)$, leading to a mean propensity score of 0.39. Policy C halves the estimated coefficient on relative distance in the model for major choice. Policy D increases enrollment of low-resistance individuals (i.e., those with positive pecuniary returns) overproportionally and sets $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i) + 0.5 \times \max((0.43 - \hat{P}_i)/0.43,\ 0)$. Policy E raises the mean propensity score of women to that of men by adding 1.36 to the propensity score index of women. Policy F increases enrollment of low-resistance women (i.e., those with positive pecuniary returns) overproportionally by setting $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i) + 0.5 \times \max((0.40 - \hat{P}_i)/0.40,\ 0)$ if person i is female.

(see Table 7). This suggests that support policies for women need to be even more nuanced, focusing on women with very low unobserved resistance in order to overcome the disadvantage in terms of monetary returns they face. We design and evaluate two potential policies specifically for women, one that affects all women and another that targets women with low unobserved resistance. *Policy E* captures a broad female-focused encouragement policy, such as mentoring, exposure to female STEM role models, information campaigns aimed at young women, or interventions that raise confidence in quantitative subjects. We model these interventions as a general upward shift in women's propensity to choose STEM. In particular, it adds a constant to the index function of the propensity score ($\hat{\mu}_D(Z_i)$ in Eq. (4)) of women such that the female enrollment rate increases from 0.161 to 0.386 and the overall enrollment rate from 0.284 to 0.386 (by 35.9%), which is the male enrollment rate in our sample. Specifically, we set $\hat{\mu}'_D(Z_i) = 1.36 + \mu_D(Z_i)$ if person i is female, and $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i)$ else. Panel (c) in Figure 4 shows that, while this policy change gives greater weight to women with a low resistance and positive MTE, it still gives some weight to women with higher resistance and negative MTE. Consequently, the PRTE is small and statistically insignificant (see Panel (b) of Table 9). In contrast, *Policy F* captures a targeted female STEM support policy, such as scholarships, mentoring, or advising directed at women with a strong intrinsic interest in STEM. It therefore represents a practically relevant contrast to broad female-oriented encouragement policies that do not differentiate by likely fit or expected return. In particular, it increases the propensity scores of low-resistance women only (similar to Policy D), by exclusively targeting women with positive MTE (i.e., $V \leq 0.40$).[37] Specifically, we set $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i) + 0.5 \times \max((0.40 - \hat{P}_i)/0.40,\ 0)$ if person i is female, and $\hat{\mu}'_D(Z_i) = \hat{\mu}_D(Z_i)$ else. This policy change increases the female enrollment rate from 0.161 at baseline to 0.195 (by 21%) and the overall enrollment rate from 0.284 to 0.299 (by 5.4%). This type of encouragement policy produces a large and statistically significant PRTE of 23 log wage points (see Panel (b) of Table 9), because it places particularly high weight on women with low resistance and large MTE.[38]

In sum, this exercise shows that well-designed policies can be successful—in the sense that they both increase STEM enrollment and have a positive average effect on the earnings of those affected—if they target individuals with positive pecuniary returns.[39] Policies that target individuals with low or even negative pecuniary returns are ineffective and do not benefit those affected. It is, therefore, important that initiatives aiming to promote STEM identify those individuals who likely benefit most from a STEM education in the first place and then encourage these particular individuals to pursue a career in STEM. Specifically,

[37] Note that the MTE curve for women is shifted downward due to their lower average returns (see Panel (d) in Figure 4), meaning that only women with $V \leq 0.40$ achieve positive returns.

[38] Since there are many women who, despite having a low resistance, do not enroll in a STEM education, this presents an opportunity for potential support policies (see Figure B8 in the Appendix).

[39] For this exercise, we ignore the potential costs of each policy. It is clear that some policies may be more expensive than others (e.g., Policy A) and, thus, potentially infeasible.

our analysis suggests that such untreated individuals can be found among those who have low predicted probabilities to major in STEM based on their observed characteristics, i.e., those with high relative costs of studying STEM. The reason is that only in this group are there individuals with unobserved characteristics implying high pecuniary returns to STEM.

Turning to the question of what policy initiatives could effectively attract more of those with potentially high pecuniary returns into STEM, there are two possibilities: on the one hand, policies aiming to promote STEM could try to mitigate unfavorable, observed characteristics, i.e., high monetary costs, that reduce the propensity to enroll in STEM among targeted people. On the other hand, policy initiatives could focus on reducing the non-pecuniary costs of STEM, i.e., target the component of unobserved resistance not related to the pecuniary return (the unobservable U_C in Section 2). Our approach, which focuses on modeling the "hard" pecuniary constraints and incentives, and subsumes "soft" factors (i.e., non-pecuniary, subjective costs and benefits) as unobserved resistance, allows us to identify the relationship between the pecuniary returns to STEM and these "soft" factors as a whole. Our findings suggest that support policies are most effective when they reduce subjective rather than monetary costs (Policies D and F vs. Policies B and E). To get a more differentiated picture of these "soft" factors, we can draw on research that focuses on shedding more light on them. For instance, Zafar (2013) suggests that one dimension of subjective benefits that matters particularly to women is workplace attributes. Unlike men, whose choices are mostly guided by earnings prospects, women place more value on the possibility of reconciling work and family commitments and on the enjoyment of the work itself (see also Campos et al., 2026).[40] Similarly, Wiswall and Zafar (2015) show that, for women, enjoyment of specific college majors plays a bigger role than expected earnings (see also Berger, 1988; Eide and Waehrer, 1998; Boudarbat, 2008; Arcidiacono et al., 2012). Finally, Carrell et al. (2010) show how female role models can help overcome the STEM enrollment gap. If promotion policies successfully identify and change these "soft" factors—e.g., by improving the work environment in STEM jobs or by providing encouragement through role models—they may be able to inspire more high-potential individuals, and especially women, to choose a STEM education.

7 Conclusion

This paper exploits the unique institutional setting of Switzerland, where prospective students can freely choose universities and majors, to estimate the causal effects of a STEM

[40] Arcidiacono (2004) also shows that subjective tastes for major-specific jobs can explain parts of the sorting into different college majors in general.

education on individuals' wages. We evaluate these effects using the framework of marginal treatment effects (MTEs) (Björklund and Moffitt, 1987; Heckman and Vytlacil, 1999, 2001a, 2005, 2007b) that allows for selection into university majors based on potential earnings as well as non-earnings-related considerations (i.e., costs and tastes) that are partially unobserved to the researcher. Identification relies on instrumental variables that affect the relative cost of pursuing a STEM education but not the return. Our main instrument is based on the geographic distance to the nearest technical university relative to the nearest general university. To improve precision, we also use age at college entry as an instrument.

We provide supporting evidence for the plausibility of the conditional independence and exclusion assumptions for our proposed instruments. In particular, we show that the partial correlations between individual and parental characteristics and the relative distance to the nearest technical university are at most weak and disappear when regional control variables are included. This suggests that relative distance is unlikely to be confounded by unobserved regional heterogeneity arising from geographic sorting of families or nonrandom university locations (see, e.g., Mountjoy, 2022 for similar reasoning). Further, we conduct a placebo analysis using an alternative instrument proposed by Osikominu et al. (2020). This enables us to confirm empirically that our baseline instruments have no direct effect on wages in the second stage of a 2SLS regression.

Estimating reduced-form Logit regressions of major choice, we find that our instruments strongly predict STEM major choice. Individuals who live closer to a technical university, relative to a general one, are more likely to enroll in a STEM field. The same is true for those who are relatively older at the time of college major choice. In the next step, we estimate the CEFs of potential wages for non-STEM and STEM graduates using the predicted propensity score from the reduced-form choice regression to account for self-selection into university majors based on unobservables. Our results suggest that the CEFs of both potential non-STEM and STEM wages vary strongly as a function of unobserved resistance to a STEM education. Low-resistance individuals have low potential non-STEM wages but high potential STEM wages, while the opposite is true for high-resistance individuals. Evaluating the MTEs, i.e., the difference between the CEFs of potential STEM and non-STEM wages, over increasing percentiles of unobserved resistance to STEM produces a downward-sloping curve, revealing a clear pattern of positive selection on unobserved idiosyncratic returns. Individuals with a low resistance gain up to 57% from a STEM education, while individuals with a high resistance lose up to 42%. This finding of positive selection on returns suggests that low- and high-resistance individuals excel in different skills (e.g., logical skills vs. social skills), which in turn are rewarded differently in typical STEM and non-STEM jobs. Hence, individuals favoring a STEM field over a non-STEM field benefit from a comparative advantage in STEM and vice versa for those favoring a non-STEM field. Results are robust to various sensitivity checks.

Given the importance of STEM graduates for innovation and economic growth (Hunt and Gauthier-Loiselle, 2010; Peri et al., 2015), understanding how different policies aiming to increase the number of STEM graduates work is important to develop strategies that effectively reach their goals. To better understand which policies are most effective in attracting individuals into a STEM education while also implying positive average pecuniary returns for those affected, we estimate policy-relevant treatment effects (PRTEs) (Heckman and Vytlacil, 2001b) for various counterfactual policies. We consider two sets of policies: policies that more or less uniformly improve access to STEM education for a broad range of people; and targeted policies that focus on groups with high pecuniary returns, as well as on women who are often underrepresented in STEM. Our results suggest that policies are only successful (i.e., increase STEM enrollment and have a positive average effect on the earnings of those affected) if they are able to target individuals with low predicted probabilities to major in STEM based on their observed characteristics, i.e., those with a high relative monetary cost of studying STEM. The reason is that only in this group are there untreated individuals with low unobserved resistance and high pecuniary returns. Policies that are able to reduce the subjective costs of these individuals result in high pecuniary returns among those affected.

References

Alan, S., Ertac, S., and Mumcu, I. (2018). Gender stereotypes in the classroom and effects on achievement. *Review of Economics and Statistics*, 100(5):876–890.

Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 11(1, Part 1):48–83.

Altonji, J. G., Arcidiacono, P., and Maurel, A. (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 5, chapter 7, pages 305–396. Elsevier, Amsterdam.

Altonji, J. G., Blom, E., and Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, 4(1):185–223.

Andresen, M. E. (2018). Exploring marginal treatment effects: Flexible estimation using Stata. *Stata Journal*, 18(1):118–158.

Andrews, I., Stock, J. H., and Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11:727–753.

Andrews, R. J., Imberman, S. A., Lovenheim, M. F., and Stange, K. (2024). The returns to college major choice: Average and distributional effects, career trajectories, and earnings variability. *Review of Economics and Statistics*, pages 1–45.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2020). Ex ante returns and occupational choice. *Journal of Political Economy*, 128(12):4475–4522.

Barbe, P. and Bertail, P. (1995). *The weighted bootstrap*. Springer, New York.

Beffy, M., Fougere, D., and Maurel, A. (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? *Review of Economics and Statistics*, 94(1):334–347.

Benbow, C. P. and Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 210(4475):1262–1264.

Berger, M. C. (1988). Predicted future earnings and choice of college major. *ILR Review*, 41(3):418–429.

Björklund, A. and Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *Review of Economics and Statistics*, 69(1):42–49.

Bleemer, Z. and Mehta, A. (2022). Will studying economics make you rich? A regression discontinuity analysis of the returns to college major. *American Economic Journal: Applied Economics*, 14(2):1–22.

Bostwick, V. K. and Weinberg, B. A. (2022). Nevertheless she persisted? Gender peer effects in doctoral STEM programs. *Journal of Labor Economics*, 40(2):397–436.

Boudarbat, B. (2008). Field of study choice by community college students in Canada. *Economics of Education Review*, 27(1):79–93.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). Beyond LATE with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

Buffington, C., Cerf, B., Jones, C., and Weinberg, B. A. (2016). STEM training and early career outcomes of female and male graduate students: Evidence from UMETRICS data linked to the 2010 census. *American Economic Review*, 106(5):333–338.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *Quarterly Journal of Economics*, 129(3):1409–1447.

Buser, T., Peter, N., and Wolter, S. C. (2017). Gender, competitiveness, and study choices in high school: Evidence from Switzerland. *American Economic Review*, 107(5):125–130.

Cameron, S. V. and Taber, C. (2004). Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112(1):132–182.

Campos, C., Muñoz, P., Bucarey, A., and Contreras, D. (2026). College major choice, payoffs, and gender gaps. NBER Working Paper No. 34736.

Canaan, S. and Mouganie, P. (2023). The impact of advisor gender on female students' STEM enrollment and persistence. *Journal of Human Resources*, 58(2):593–632.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. In Christofides, L. N., Grant, K. E., and Swidinsky, R., editors, *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, pages 201–222. University of Toronto Press, Toronto.

Card, D. and Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, 59(1):9–28.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *Quarterly Journal of Economics*, 134(3):1163–1224.

Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6):2754–2781.

Carneiro, P. and Lee, S. (2009). Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality. *Journal of Econometrics*, 149(2):191–208.

Carrell, S. E., Page, M. E., and West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics*, 125(3):1101–1144.

Choi, W., Kinsler, J., Orellana, A., and Pavan, R. (2023). College majors and earnings growth. Unpublished manuscript.

Cornelissen, T., Dustmann, C., Raute, A., and Schönberg, U. (2016). From LATE to MTE: Alternative methods for the evaluation of policy interventions. *Labour Economics*, 41:47–60.

Cornelissen, T., Dustmann, C., Raute, A., and Schönberg, U. (2018). Who benefits from universal child care? Estimating marginal returns to early child care attendance. *Journal of Political Economy*, 126(6):2356–2409.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics*, 118(4):1495–1532.

Daymont, T. N. and Andrisani, P. J. (1984). Job preferences, college major, and the gender gap in earnings. *Journal of Human Resources*, 19(3):408–428.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, 88(9-10):1697–1720.

Deming, D. J. and Noray, K. (2020). Earnings dynamics, changing job skills, and STEM careers. *Quarterly Journal of Economics*, 135(4):1965–2005.

Eide, E. and Waehrer, G. (1998). The role of the option value of college attendance in college major choice. *Economics of Education Review*, 17(1):73–82.

Federal Ministry of Education and Research (2021). MINT-Aktionsplan 2.0. https://www.bildung-forschung.digital/digitalezukunft/de/unsere-ueberzeugungen/digitalstrategie-des-bmbf/mint-aktionsplan/mint-aktionsplan_node.html, retrieved January 27, 2025.

Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men's and women's STEM persistence. *Labour Economics*, 46:211–226.

Fisher, R. A. (1935). *Design of experiments*. Oliver and Boyd, London.

French, E. and Taber, C. (2011). Identification of models of the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, chapter 6, pages 537–617. Elsevier, Amsterdam.

Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4):133–156.

Grogger, J. and Eide, E. (1995). Changes in college skills and the rise in the college wage premium. *Journal of Human Resources*, 30(2):280–310.

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science*, 320(5880):1164–1165.

Hamermesh, D. S. and Donald, S. G. (2008). The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias. *Journal of Econometrics*, 144(2):479–491.

Hastings, J. S., Neilson, C. A., and Zimmerman, S. D. (2013). Are some degrees worth more than others? Evidence from college admission cutoffs in Chile. NBER Working Paper No. 19241.

Heckman, J. J., Schmierer, D., and Urzua, S. (2010). Testing the correlated random coefficient model. *Journal of Econometrics*, 158(2):177–203.

Heckman, J. J. and Vytlacil, E. J. (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences*, 96(8):4730–4734.

Heckman, J. J. and Vytlacil, E. J. (2001a). Local instrumental variables. In Hsiao, C., Morimune, K., and Powell, J. L., editors, *Nonlinear statistical modeling: Essays in honor of Takeshi Amemiya*, chapter 1. Cambridge University Press, New York.

Heckman, J. J. and Vytlacil, E. J. (2001b). Policy-relevant treatment effects. *American Economic Review*, 91(2):107–111.

Heckman, J. J. and Vytlacil, E. J. (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, 73(3):669–738.

Heckman, J. J. and Vytlacil, E. J. (2007a). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6B, chapter 70, pages 4779–4874. Elsevier, Amsterdam.

Heckman, J. J. and Vytlacil, E. J. (2007b). Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6B, chapter 71, pages 4875–5143. Elsevier, Amsterdam.

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

Imberman, S. A., Lovenheim, M. F., Massey, P., Stange, K., and Andrews, R. J. (2026). The contribution of college majors to gender and racial earnings differences. NBER Working Paper No. 34726.

James, E., Alsalam, N., Conaty, J. C., and To, D. (1989). College quality and future earnings: Where should you send your child to college? *American Economic Review*, 79(2):247–252.

Joensen, J. S. and Nielsen, H. S. (2016). Mathematics and gender: Heterogeneity in causes and consequences. *Economic Journal*, 126(593):1129–1163.

Kahn, S. and Ginther, D. (2017). Women and STEM. NBER Working Paper No. 23525.

Kane, T. J. and Rouse, C. E. (1995). Labor market returns to two- and four-year colleges: Is a credit a credit and do degrees matter? *American Economic Review*, 85(3):600–614.

Kinsler, J. and Pavan, R. (2015). The specificity of general human capital: Evidence from college major choice. *Journal of Labor Economics*, 33(4):933–972.

Kirkebøen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection. *Quarterly Journal of Economics*, 131(3):1057–1111.

Kling, J. R. (2001). Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business & Economic Statistics*, 19(3):358–364.

Lavy, V. and Sand, E. (2018). On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases. *Journal of Public Economics*, 167:263–279.

Lovenheim, M. and Smith, J. (2023). Returns to different postsecondary investments: Institution type, academic programs, and credentials. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 6, chapter 4, pages 187–318. Elsevier, Amsterdam.

Morin, R.-A. and Suarez, A. F. (1983). Risk aversion revisited. *Journal of Finance*, 38(4):1201–1216.

Mountjoy, J. (2022). Community colleges and upward mobility. *American Economic Review*, 112(8):2580–2630.

Mourifie, I., Henry, M., and Meango, R. (2020). Sharp bounds and testability of a Roy model of STEM major choices. *Journal of Political Economy*, 128(8):3220–3283.

Neyman, J. (1923). Statistical problems in agricultural experiments. *Journal of the Royal Statistical Society II*, Suppl. (2):107–180.

Ng, K. and Riehl, E. (2024). The returns to STEM programs for less-prepared students. *American Economic Journal: Economic Policy*, 16(2):37–77.

Ngo, D. and Dustan, A. (2024). Preferences, access, and the STEM gender gap in centralized high school assignment. *American Economic Journal: Applied Economics*, 16(4):257–287.

Nollenberger, N., Rodríguez-Planas, N., and Sevilla, A. (2016). The math gender gap: The role of culture. *American Economic Review*, 106(5):257–261.

Nybom, M. (2017). The distribution of lifetime earnings returns to college. *Journal of Labor Economics*, 35(4):903–952.

OECD (2004). Education at a glance 2004. OECD indicators, OECD publishing, Paris.

OECD (2021). Education at a glance 2021. OECD indicators, OECD publishing, Paris.

OECD (2023). Education at a glance 2023. OECD indicators, OECD publishing, Paris.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

Osikominu, A., Grossmann, V., and Osterfeld, M. (2020). Sociocultural background and choice of STEM majors at university. *Oxford Economic Papers*, 72(2):347–369.

Owen, S. (2023). College major choice and beliefs about relative performance: An experimental intervention to understand gender gaps in STEM. *Economics of Education Review*, 97:102479.

Pei, Z., Pischke, J.-S., and Schwandt, H. (2019). Poorly measured confounders are more useful on the left than on the right. *Journal of Business & Economic Statistics*, 37(2):205–216.

Peri, G., Shih, K., and Sparber, C. (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics*, 33(S1):S225–S255.

Pflueger, C. E. and Wang, S. (2015). A robust test for weak instruments in Stata. *Stata Journal*, 15(1):216–225.

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–254.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, 6(1):34–58.

Saltiel, F. (2023). Multi-dimensional skills and gender differences in STEM majors. *Economic Journal*, 133(651):1217–1247.

Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586.

STEM Learning (2025). STEM Learning. https://www.stem.org.uk/, retrieved January 27, 2025.

Swiss Academy of Engineering Sciences (2025). Swiss TecLadies. https://tecladies.ch/, retrieved January 27, 2025.

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

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

Swiss Federal Statistical Office (2012). Erhebungen, Quellen – Absolventenstudien Hochschulen: Steckbrief, Neuchâtel.

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

Swiss Federal Statistical Office (2015). Bildungssystem Schweiz – Indikatoren: Abschlüsse und Kompetenzen – Maturitäsquote, Neuchâtel.

Turner, S. E. and Bowen, W. G. (1999). Choice of major: The changing (unchanging) gender gap. *ILR Review*, 52(2):289–313.

Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies*, 82(2):791–824.

Wiswall, M. and Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *Quarterly Journal of Economics*, 133(1):457–507.

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(4):471–483.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, 48(3):545–595.

A Appendix – MTE Weights

Aggregate treatment parameters such as the ATT, ATE, TUT and PRTE can be obtained as weighted averages over the MTE curve (Heckman and Vytlacil, 2007b). In particular, we can write them as

$$\text{Treatment parameter } (j) = \int_0^1 \text{MTE}(x, v)\omega_j(v)dv, \tag{A1}$$

where $\omega_j(v)$ is the weight for treatment parameter j. Table A1 summarizes all relevant treatment parameters and their estimated weights assuming a discrete uniform distribution for V with s support points (Andresen, 2018).

Table A1: Treatment Effect Parameters and MTE Weights

Panel A: Treatment effect parameters
$\text{ATE}(x) = \mathbb{E}(Y_1 - Y_0 \mid X = x) = \int_0^1 \text{MTE}(x, v)\omega_{\text{ATE}}(v)dv$
$\text{ATT}(x) = \mathbb{E}(Y_1 - Y_0 \mid X = x, D = 1) = \int_0^1 \text{MTE}(x, v)\omega_{\text{ATT}}(v)dv$
$\text{TUT}(x) = \mathbb{E}(Y_1 - Y_0 \mid X = x, D = 0) = \int_0^1 \text{MTE}(x, v)\omega_{\text{TUT}}(v)dv$
$\text{PRTE}(x) = \mathbb{E}(Y' \mid X = x) - \mathbb{E}(Y \mid X = x) = \int_0^1 \text{MTE}(x, v)\omega_{\text{PRTE}}(v)dv$
Panel B: Definition of weights
$\omega_{\text{ATE}}(v) = \dfrac{1}{s}$
$\omega_{\text{ATT}}(v) = \dfrac{\Pr(P > v)}{s\,\mathbb{E}(P)}$
$\omega_{\text{TUT}}(v) = \dfrac{1 - \Pr(P > v)}{s\,[1 - \mathbb{E}(P)]}$
$\omega_{\text{PRTE}}(v) = \dfrac{\Pr(P' > v) - \Pr(P > v)}{s\,[\mathbb{E}(P') - \mathbb{E}(P)]}$

Source: Heckman and Vytlacil (2007b) and Andresen (2018). *Notes*: This table describes common average treatment parameters in Panel A and the weights used to calculate them in Panel B assuming a discrete uniform distribution with s support points for unobserved resistance, V.

B Appendix – Additional Tables and Figures

Table B1: Stylized Facts about Swiss Universities

University	Founding date	Students	STEM faculties	Non-STEM faculties
ETH Lausanne	1853	12,700	6	1
ETH Zurich	1855	25,022	14	2
University of Basel	1460	13,039	1	6
University of Bern	1834	19,297	0	8
University of Fribourg	1889	10,724	1	4
University of Geneva	1559	19,078	1	8
University of Lausanne	1537	16,908	1	6
University of Lucerne	2000	3,346	0	6
University of Lugano	1995	4,190	3	3
University of Neuchâtel	1838	4,409	1	3
University of St. Gallen	1898	9,590	1	5
University of Zurich	1833	27,895	1	6

Notes: This table presents stylized facts about Swiss universities. Information is retrieved from the universities' homepages, last visited on September 7, 2023. Student numbers are for 2021 or 2022. Note that the University of Lucerne is not taken into account in the main analysis because it was founded after the individuals in our sample have already chosen their college major.

Table B2: Robustness Check – First Stage Regressions for Major Choice

	Baseline	W/o Zurich & Lausanne	W/o border munic.	W/o small munic.	W/o large munic.
	(1)	(2)	(3)	(4)	(5)
Log relative distance	-0.025^{***}	-0.028^{***}	-0.027^{***}	-0.026^{***}	-0.033^{***}
	(0.005)	(0.008)	(0.005)	(0.006)	(0.009)
Age at college entry	0.074^{***}	0.072^{***}	0.075^{***}	0.076^{***}	0.075^{***}
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
= 1 if female	-0.122^{***}	-0.123^{***}	-0.128^{***}	-0.110^{***}	-0.124^{***}
	(0.012)	(0.013)	(0.013)	(0.012)	(0.014)
Age	0.353^{***}	0.319^{***}	0.352^{***}	0.277^{***}	0.388^{***}
	(0.091)	(0.092)	(0.098)	(0.102)	(0.094)
Age squared	-0.007^{***}	-0.006^{***}	-0.007^{***}	-0.005^{***}	-0.007^{***}
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
= 1 if math & science spec.	0.322^{***}	0.322^{***}	0.330^{***}	0.310^{***}	0.268^{***}
	(0.015)	(0.016)	(0.016)	(0.016)	(0.011)
= 1 if father university	0.027	0.023	0.022	0.031	0.037
	(0.022)	(0.023)	(0.024)	(0.024)	(0.024)
= 1 if father voc. training	0.030	0.025	0.026	0.019	0.037
	(0.021)	(0.021)	(0.023)	(0.023)	(0.023)
= 1 if mother university	0.005	0.004	−0.002	0.010	0.006
	(0.022)	(0.023)	(0.023)	(0.023)	(0.026)
= 1 if mother voc. training	0.022	0.023	0.025	0.035^{**}	0.024
	(0.015)	(0.016)	(0.016)	(0.017)	(0.017)
Regional Characteristics	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓
Observations	4766	4506	4120	4019	3966

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows average partial effects (APE) from Logit regressions of major choice (STEM versus non-STEM). Column (2) drops the municipalities of Zurich and Lausanne. Column (3) drops municipalities on the border with surrounding countries (i.e., Germany, France, Italy, Austria, and Liechtenstein). Column (4) drops small municipalities with a population of less than 2,000. Column (5) drops large municipalities with a population of more than 50,000. Standard errors are robust to clustering at the municipality-level and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table B3: Robustness Check – IV Regressions of Returns to STEM

	Baseline		W/o Zurich & Lausanne		W/o border munic.		W/o small munic.		W/o large munic.	
	(1) IV (Z)	(2) IV ($\hat{P}$)	(3) IV (Z)	(4) IV ($\hat{P}$)	(5) IV (Z)	(6) IV ($\hat{P}$)	(7) IV (Z)	(8) IV ($\hat{P}$)	(9) IV (Z)	(10) IV ($\hat{P}$)
= 1 if STEM major	0.162**	0.143**	0.219**	0.181**	0.240**	0.205**	0.251***	0.213***	0.245***	0.221***
	(0.077)	(0.065)	(0.089)	(0.075)	(0.096)	(0.086)	(0.096)	(0.081)	(0.091)	(0.079)
= 1 if female	−0.048***	−0.051***	−0.047***	−0.052***	−0.043**	−0.048***	−0.055***	−0.059***	−0.039**	−0.043**
	(0.015)	(0.014)	(0.017)	(0.015)	(0.018)	(0.017)	(0.017)	(0.016)	(0.018)	(0.017)
Age	0.016	0.019	0.021	0.028	0.012	0.018	0.062	0.066	−0.043	−0.037
	(0.070)	(0.070)	(0.087)	(0.085)	(0.087)	(0.086)	(0.082)	(0.080)	(0.090)	(0.087)
Age squared	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000	−0.001	−0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
= 1 if math & science spec.	−0.088***	−0.082***	−0.104***	−0.091***	−0.120***	−0.109***	−0.108***	−0.097***	−0.105***	−0.097***
	(0.031)	(0.027)	(0.034)	(0.029)	(0.037)	(0.034)	(0.033)	(0.029)	(0.034)	(0.030)
= 1 if second cohort	0.011	0.011	0.010	0.010	0.007	0.007	0.007	0.008	0.003	0.004
	(0.011)	(0.011)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)
= 1 if father university	−0.007	−0.007	−0.003	−0.003	0.015	0.015	−0.006	−0.006	0.003	0.004
	(0.022)	(0.022)	(0.023)	(0.023)	(0.026)	(0.025)	(0.025)	(0.025)	(0.024)	(0.024)
= 1 if father voc. training	0.019	0.020	0.024	0.025	0.031	0.032	0.032	0.032	0.023	0.024
	(0.024)	(0.025)	(0.023)	(0.023)	(0.027)	(0.027)	(0.028)	(0.027)	(0.022)	(0.021)
= 1 if mother university	−0.089***	−0.089***	−0.099***	−0.099***	−0.099***	−0.100***	−0.098***	−0.098***	−0.108***	−0.108***
	(0.020)	(0.020)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)	(0.026)	(0.026)
= 1 if mother voc. training	−0.022	−0.022	−0.024	−0.022	−0.034**	−0.032**	−0.034*	−0.033*	−0.032*	−0.031*
	(0.015)	(0.015)	(0.017)	(0.016)	(0.017)	(0.016)	(0.019)	(0.018)	(0.017)	(0.016)
Regional Characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Effective F-stat	183.40	451.16	141.58	352.41	141.20	273.49	124.28	276.04	136.58	325.07
Observations	4631	4631	4101	4101	3769	3769	3574	3574	3616	3616

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This table shows full results from 2SLS regressions of the log hourly gross wage on graduating from a STEM major. The specification in Columns (1), (3), (5), (7), and (9) uses the set of instruments Z defined in Table 1 in the first stage. The specification in Columns (2), (4), (6), (8), and (10) uses the propensity score P obtained from a Logit regression in the first stage. Columns (3) and (4) drop the municipalities of Zurich and Lausanne. Columns (5) and (6) drop municipalities on the border with surrounding countries (i.e., Germany, France, Italy, Austria, and Liechtenstein). Columns (7) and (8) drop small municipalities with a population of less than 2,000. Columns (9) and (10) drop large municipalities with a population of more than 50,000. Effective F-statistics are based on the approach by Olea and Pflueger (2013). The samples are trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support. The propensity score used in the second stage is reestimated for each sample. Standard errors are robust to clustering at the municipality-level and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

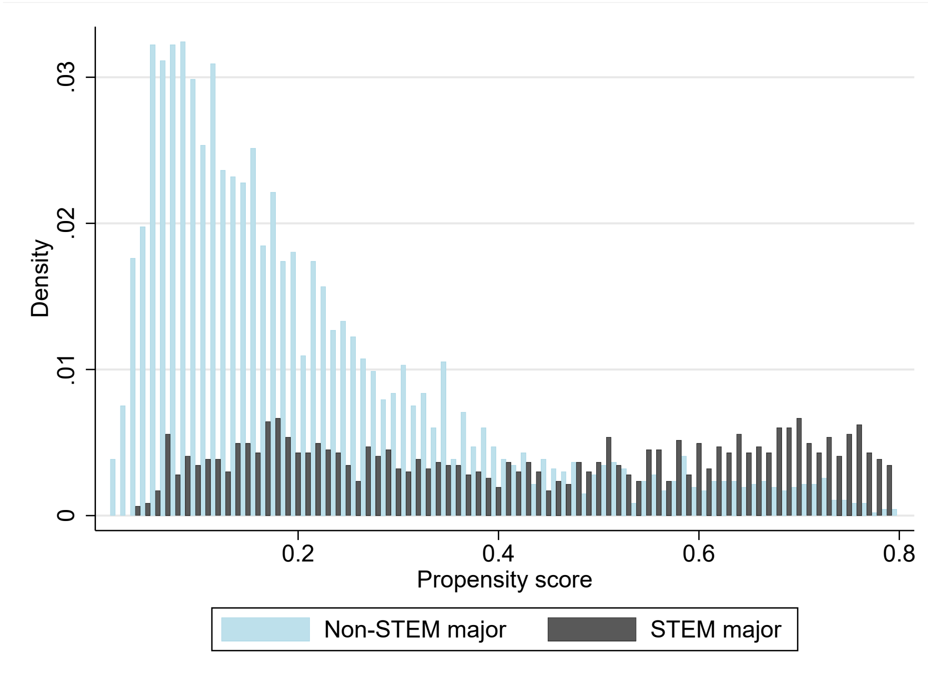
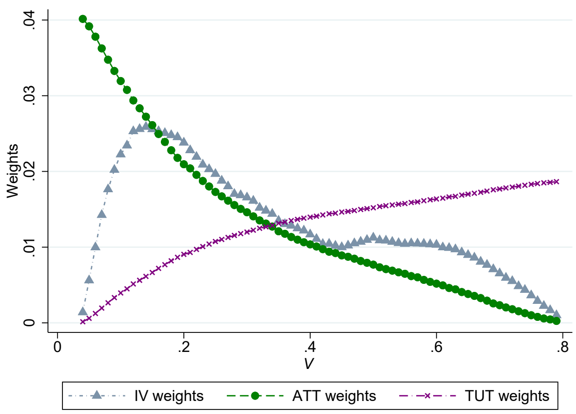
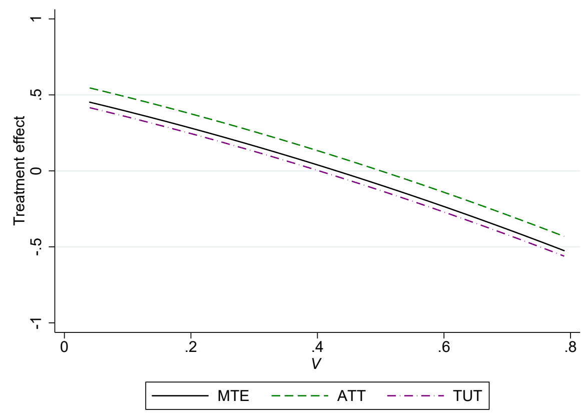


Figure B1: Common Support

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This figure shows the share of individuals who have graduated from a non-STEM major or STEM major, respectively, summarized in 0.01-intervals of the propensity score. The propensity score is based on the Logit regression presented in Table 5. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively.



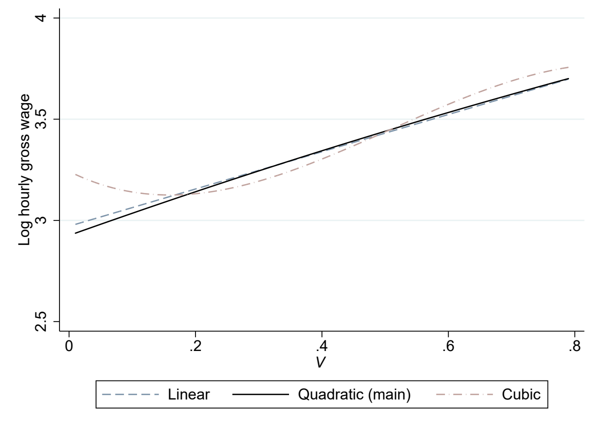
(a) MTE weights



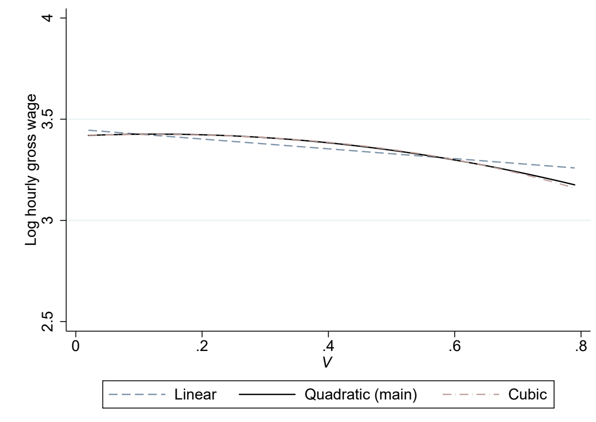
(b) MTE curves

Figure B2: MTE Weights and MTE Curves

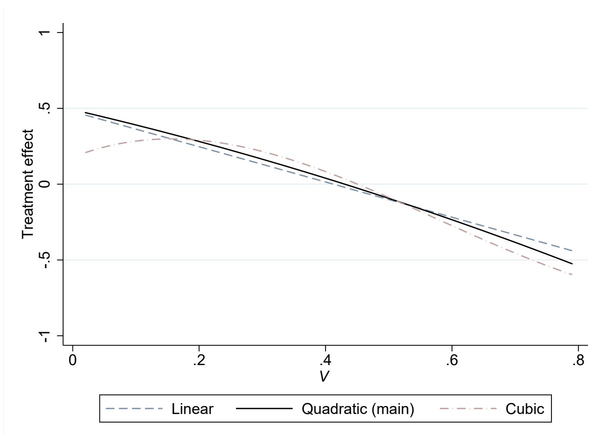
Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: These figures show the weights used to calculate the average treatment parameters in Tables 6 and 8 in Panel (a) and MTE curves evaluated at average observable characteristics of the full sample (MTE), STEM graduates (ATT), and non-STEM graduates (TUT) in Panel (b).



(a) $\mathbb{E}(Y_0 \mid X, V)$



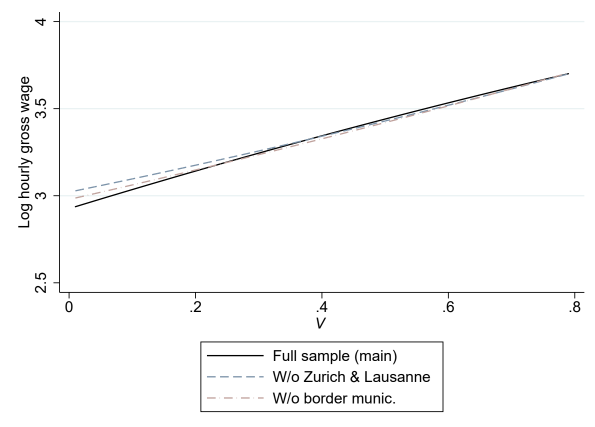
(b) $\mathbb{E}(Y_1 \mid X, V)$



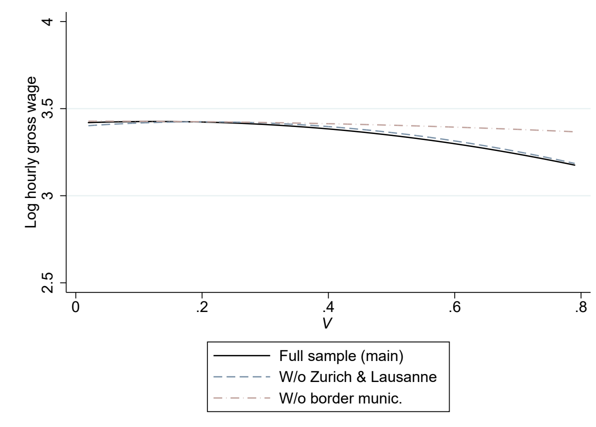
(c) $\text{MTE}(X, V)$

Figure B3: Robustness Check – Different Functional Forms

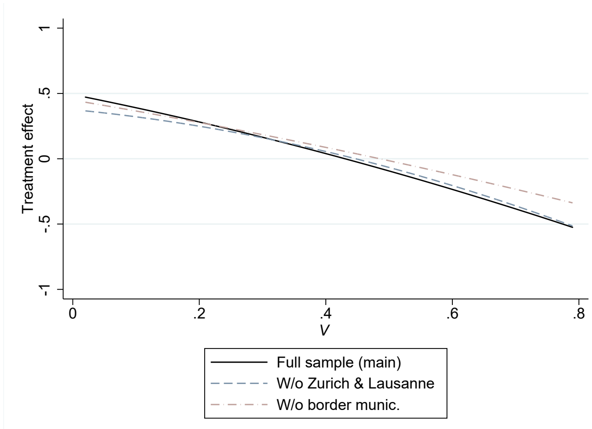
Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: These figures show the estimated CEFs of Y_0 in Panel (a) and of Y_1 in Panel (b), and the MTE curves obtained as their difference in Panel (c). Curves in blue use a linear function of $k_t(p)$, curves in black use a quadratic function of $k_t(p)$ and correspond to our main results, and curves in red use a cubic function of $k_t(p)$. The samples are trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and nontreated individuals in Panels (a) and (b), respectively, and are restricted to the common support in Panel (c).



(a) $\mathbb{E}(Y_0 \mid X, V)$



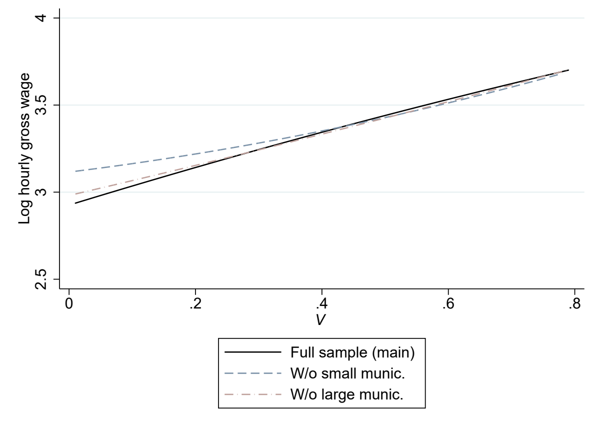
(b) $\mathbb{E}(Y_1 \mid X, V)$



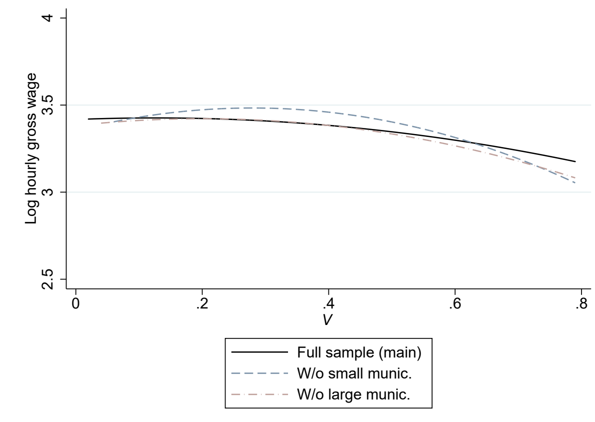
(c) $\mathrm{MTE}(X, V)$

Figure B4: Robustness Check – Different Sample Restrictions 1

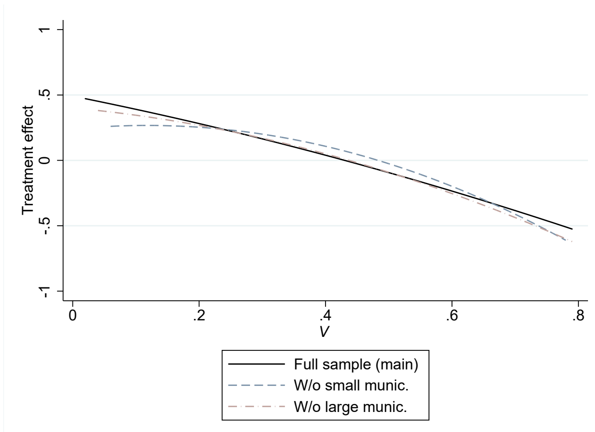
Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: These figures show the estimated CEFs of Y_0 in Panel (a) and of Y_1 in Panel (b), and the MTE curves obtained as their difference in Panel (c). Curves in black are for the full sample and thus correspond to our main results, curves in blue drop the municipalities of Zurich and Lausanne, and curves in red drop municipalities on the border with surrounding countries (i.e., Germany, France, Italy, Austria, and Liechtenstein). The samples are trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and nontreated individuals in Panels (a) and (b), respectively, and are restricted to the common support in Panel (c).



(a) $\mathbb{E}(Y_0 \mid X, V)$



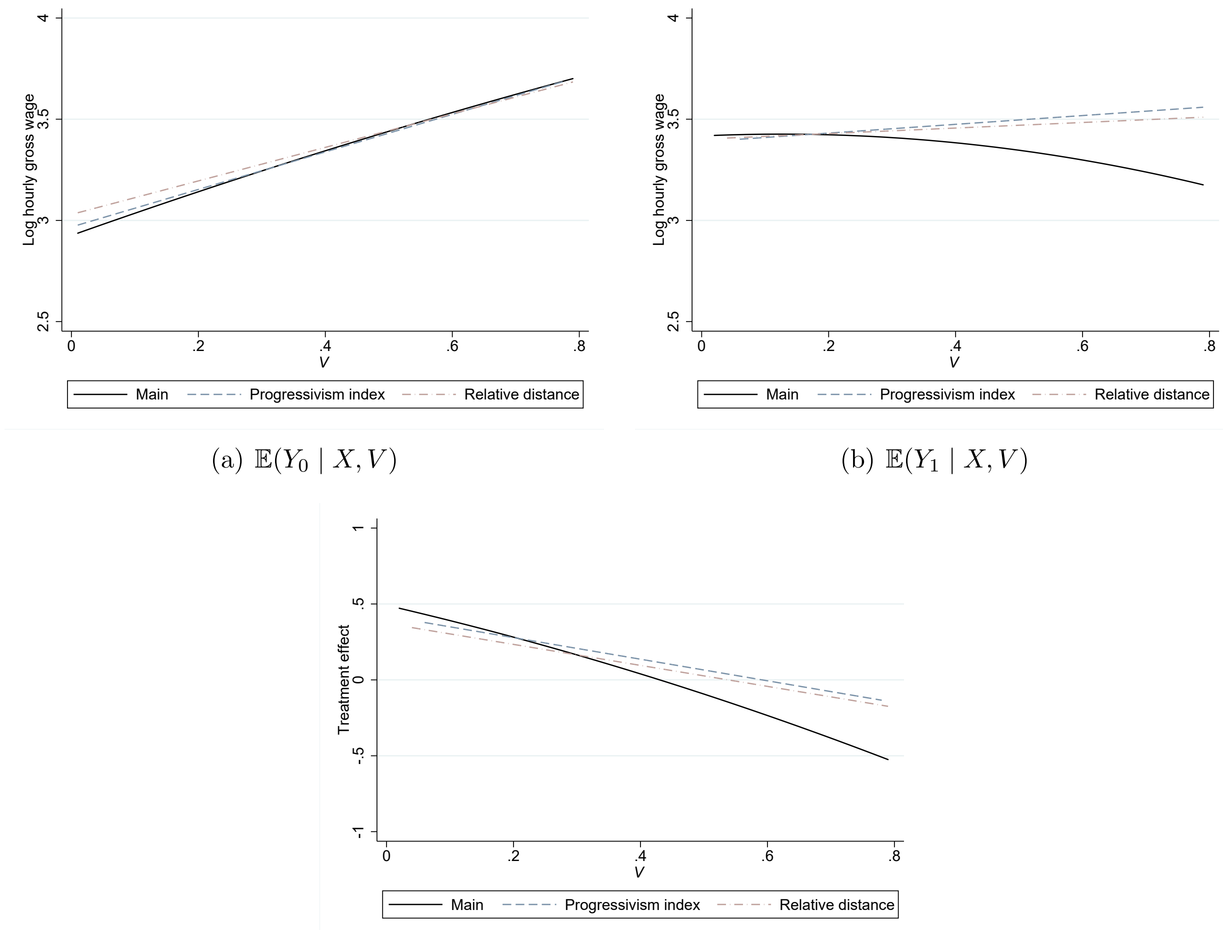
(b) $\mathbb{E}(Y_1 \mid X, V)$



(c) $\mathrm{MTE}(X, V)$

Figure B5: Robustness Check – Different Sample Restrictions 2

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: These figures show the estimated CEFs of Y_0 in Panel (a) and of Y_1 in Panel (b), and the MTE curves obtained as their difference in Panel (c). Curves in black are for the full sample and thus correspond to our main results, curves in blue drop small municipalities with a population of less than 2,000, and curves in red drop large municipalities with a population of more than 50,000. The samples are trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and nontreated individuals in Panels (a) and (b), respectively, and are restricted to the common support in Panel (c).



(a) $\mathbb{E}(Y_0 \mid X, V)$

(b) $\mathbb{E}(Y_1 \mid X, V)$

(c) $\mathrm{MTE}(X, V)$

Figure B6: Robustness Check – Alternative Choice of Instrument

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: These figures show the estimated CEFs of Y_0 in Panel (a) and of Y_1 in Panel (b), and the MTE curves obtained as their difference in Panel (c). Curves in black correspond to our main results, curves in blue use the progressivism index of Osikominu et al. (2020) as excluded instrument, and curves in red use only the relative distance as excluded instrument. The samples are trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and nontreated individuals in Panels (a) and (b), respectively, and are restricted to the common support in Panel (c).

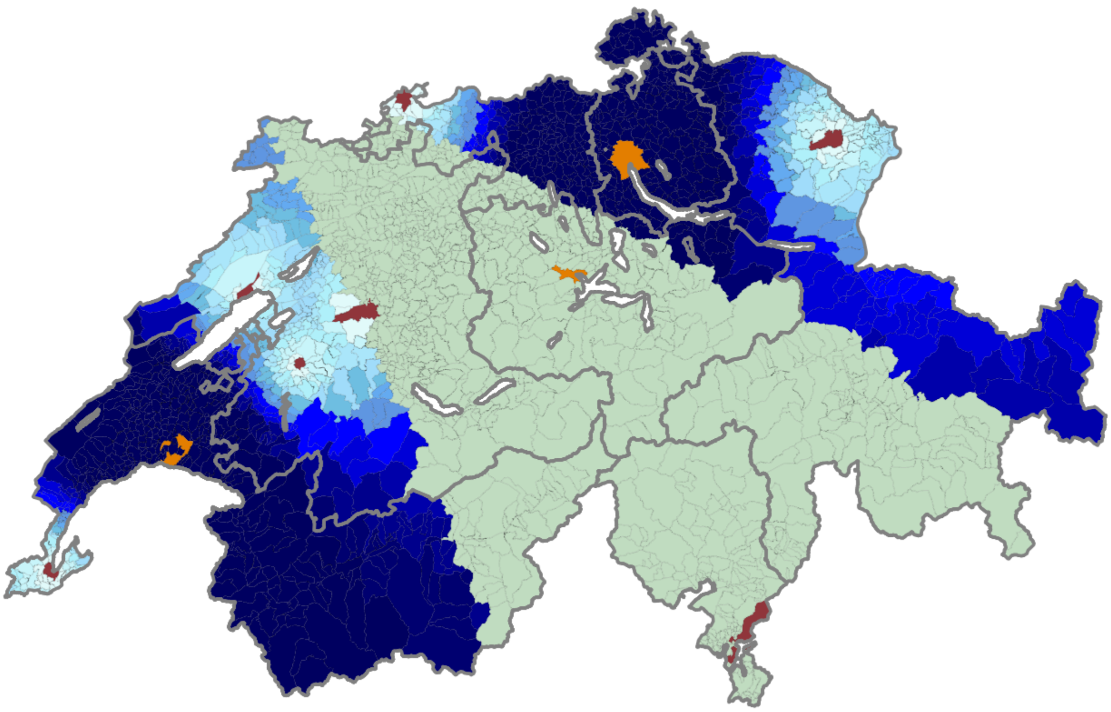


Figure B7: Municipalities Affected by the Establishment of the University of Lucerne

Notes: The figure shows a map of Swiss municipalities. Those whose closest distance changes with the establishment of the University of Lucerne as a technical university are marked in olive (Policy A). Municipalities with a general university are colored in dark red. The municipalities of Lausanne and Zurich, marked in orange, host both a general and a technical university. Lucerne is also marked in orange. Darker blue color indicates a smaller relative distance to the next technical university while brighter blue color indicates a larger relative distance. Grey color indicates NUTS-2 region boundaries.

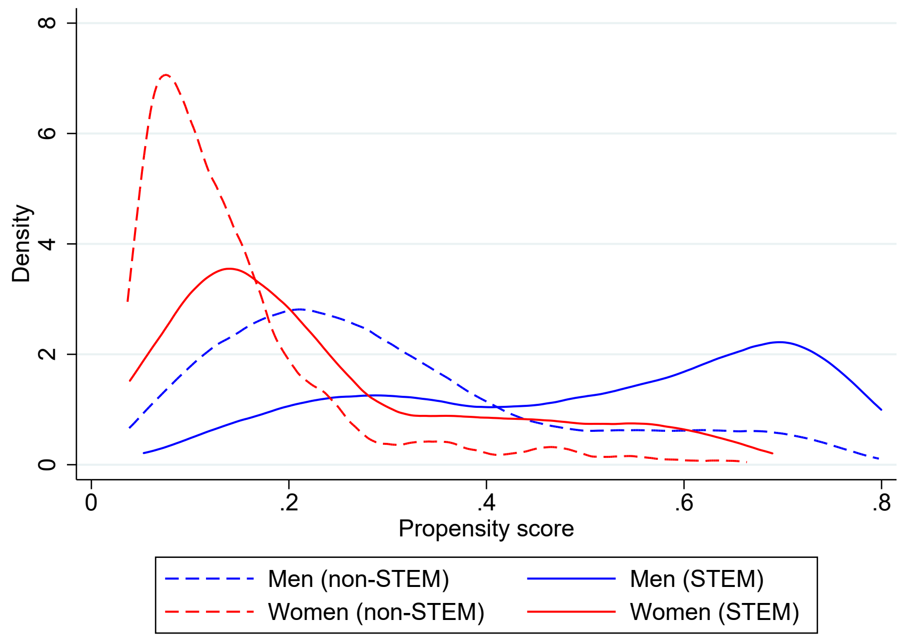


Figure B8: Distribution of Propensity Score by Sex

Source: Authors' calculations using data from Swiss Federal Statistical Office (2008, 2009, 2012). *Notes*: This figure shows the distribution of the propensity score for men and women, respectively, for non-STEM and STEM graduates separately. The propensity score is based on the Logit regression presented in Table 5. The sample is trimmed to contain at least 5 observations in each 0.01-interval of the propensity score for treated and non-treated individuals, respectively, and restricted to the common support.