Online Appendix

Vocational versus General Upper Secondary Education and Earnings

Eskil Heinesen and Elise Stenholt Lange

Contents:

Appendix A. Supplementary tables and figures

Table A1. Sample selection

	Criteria	N
1.	Birth year 1986-1989, living in Denmark in at least one year before age 14	253,923
2.	Was enrolled in at least one of the grades 8, 9 or 10 in Denmark	243,135
3.	Did not enroll in upper secondary school and then re-enter lower secondary	242,197
4.	Only Danish upper secondary education programs are included [a]	241,014
5.	Was living in Denmark Jan. 1st in the year leaving lower secondary school	240,840
6.	Did not complete 9th grade earlier than prescribed if born in 1986 [b]	238,629
7.	Was living in Denmark at age 25 (when we measure basic education status)	228,328
8.	Distances to educational institutions (the instruments) are observed [c]	221,812
9.	Outcomes at age 28 are observed (lived in Denmark at age 28)	214,871
10.	Was living in Denmark in 2015-2018 [d]	210,339
11.	Enrolled in an upper secondary education program by age 25	198,048
12.	Completed an upper secondary education program by age 25	161,432

[a] Was enrolled in an upper secondary program if had completed such a program.

[b] Persons born in 1986 (the first cohort) who completed 9th grade earlier than 2002 are excluded, because we do not have data on 9th grade test scores before 2002.

[c] We exclude persons living on small islands without a bridge to the mainland, because distances are based on the road network. Residential address is registered on January 1st in the year of leaving lower-secondary school.

[d] The last year in our dataset is 2018, and data for 2015-2018 are used in the estimation of predicted earnings at age 40.

Table A2. Main types of general and vocational upper secondary education by gender (percent)

General programs	Males	Females
General, 3-year (stx)	55.39	66.82
General, 2-year (hf)	5.84	11.01
Mercantile (hhx)	22.34	19.36
Technical (htx)	16.43	2.81
Total	100.00	100.00

Vocational programs	Males	Females
Commercial field, office	17.36	48.94
Technology, communication	18.49	3.45
Construction	25.59	3.98
Craftsmanship	8.40	1.29
Food production, catering	10.86	18.98
Mechanics, transport, logistics	18.07	1.40
Service, care, health	1.24	21.96
Total	100.00	100.00

Note. Students are categorized by the first program in which they enroll.

Table A3. Educational attainment by age 28-31: Highest completed education by level and field of study (percent)

	Males	Females
Upper secondary education		
General upper secondary	9.0	5.8
Vocational: Commercial field, office	6.9	12.1
Vocational: Technology, communication	9.4	0.9
Vocational: Construction and craftsmanship	11.9	1.1
Vocational: Food production, catering	5.1	4.8
Vocational: Mechanics, transport, logistics	6.6	0.2
Vocational: Service, care, health	0.8	8.0
2-year post-secondary education		
Social science	2.8	3.4
Technical educations	2.4	0.7
Other 2-year post-secondary programs	1.9	1.7
4-year post-secondary education		
Teacher training programs	4.8	11.6
Social science	2.8	4.5
Technical educations	5.8	0.7
Health	1.4	10.7
Other 4-year post-secondary programs	1.9	2.5
University programs (bachelor, master, PhD)		
Humanities and arts	5.4	10.4
Science	3.5	2.7
Social science	10.4	11.5
Technical sciences	5.2	2.9
Health	1.9	3.8
Total	100.0	100.0
N	79,070	82,362

Note. If an individual has completed both a general and a vocational upper secondary education, but no post-secondary education, their highest completed education is defined (by Statistics Denmark) as the vocational education. In the lower part of the table, we have merged the three levels of university programs (bachelors, masters and PhD) because the number having completed a bachelors but not a masters or PhD by age 28-31 is small, as the number having completed a PhD.

Table A4. Level of education attained by age 28-31 (percent)

	Males	Females
General upper secondary	9.0	5.8
Vocational upper secondary	40.6	27.1
2-year post-secondary	7.1	5.7
4-year post-secondary	16.9	30.0
Bachelor	4.4	4.3
Master	21.3	26.7
PhD	0.7	0.3
Total	100.0	100.0
N	79,070	82,362

Note. This table shows highest completed education by level, and distinguishes between the bachelor, master and PhD levels. If an individual has completed both a general and a vocational upper secondary education, but no post-secondary education, their highest completed education is defined (by Statistics Denmark) as the vocational education.

Table A5. Means of covariates by type of completed upper secondary program

	General	Vocational	Total
Male	0.423	0.637	0.490
Western immigrant	0.004	0.004	0.004
Western 2nd generation immigrant	0.004	0.002	0.003
Non-Western immigrant	0.029	0.029	0.029
Non-Western 2nd generation immigrant	0.030	0.023	0.028
Only child	0.160	0.191	0.170
1 full sibling	0.498	0.481	0.492
2 full siblings	0.252	0.232	0.246
First-born	0.389	0.346	0.375
Any older half siblings, mother	0.071	0.101	0.080
Any older half siblings, father	0.099	0.112	0.103
Parents separated at age 14	0.253	0.333	0.279
Not with parent(s) at age 14	0.005	0.013	0.008
Child not in population, age 14	0.004	0.004	0.004
Mother not in population, age 14	0.012	0.014	0.013
Father not in population, age 14	0.042	0.044	0.043
Mother vocational education	0.334	0.429	0.364
Mother short further education	0.100	0.058	0.086
Mother medium further education	0.294	0.116	0.238
Mother higher education	0.066	0.008	0.048
Father vocational education	0.393	0.510	0.430
Father short further education	0.107	0.059	0.092
Father medium further education	0.155	0.048	0.122
Father higher education	0.124	0.016	0.090
Mother self-employed	0.047	0.042	0.046
Mother unemployed	0.023	0.033	0.026
Mother disability pension	0.024	0.045	0.031
Mother not in work force	0.077	0.101	0.085
Father self-employed	0.113	0.114	0.114
Father unemployed	0.018	0.022	0.019
Father disability pension	0.023	0.036	0.027
Father not in work force	0.045	0.057	0.049
Log income mother	3.861	3.752	3.827
Log income father	4.094	3.925	4.041
Mother conviction: Child age 0-13	0.026	0.044	0.031
Father conviction: Child age 0-13	0.088	0.145	0.106
Father imprisonment/probation: Child age 0-13	0.026	0.056	0.036
Father violence (conviction): Child age 0-13	0.015	0.030	0.020
Test score math	8.435	7.038	8.020
Test score Danish	8.685	7.209	8.245
Test score English	8.999	7.230	8.509
Test score science	8.418	7.224	8.096
Math score by teacher	8.665	7.320	8.265
Danish score by teacher	8.722	7.262	8.289

	General	Vocational	Total
English score by teacher	8.667	7.083	8.202
Science score by teacher	8.435	7.251	8.096
Test score project report	9.243	7.746	8.819
Age 17+ grade 9 test scores	0.006	0.041	0.017
No 9th grade test scores	0.021	0.072	0.037
N	110,707	50,725	161,432

Note. If a student has completed both a general and a vocational upper secondary education program before age 25, they are categorized by their first completed program. Dummies for cohorts (3), municipalities (92) and missing information on specific variables (14) are not shown.

Table A6. OLS estimation results. Effect on earnings of completing a vocational instead of a general program ($1,000).

	Males		Females	
	Earnings age 28	Earnings age 40	Earnings age 28	Earnings age 40
	Controls: Cohort and municipality dummies			
Vocational	2.681***	-19.084***	-8.917***	-22.057***
	(0.187)	(0.341)	(0.171)	(0.243)
	Controls: Cohort and municipality dummies, and family background			
Vocational	2.619***	-15.681***	-7.931***	-18.406***
	(0.202)	(0.353)	(0.180)	(0.246)
	Controls: Cohort and municipality dummies, family background, and test scores			
Vocational	7.161***	-5.573***	-2.713***	-8.789***
	(0.242)	(0.367)	(0.214)	(0.260)
N	79,070	79,070	82,362	82,362

Note. The table shows OLS estimates of effects of completing a vocational instead of a general upper secondary education. If a student completed both a general and a vocational upper secondary education program, they are categorized by their first completed program. The lower panel of the table shows results when the full set of controls are included: 92 municipality dummies, 3 cohort dummies, and 62 family background and 9th grade test score variables. These estimates are identical to those reported in Table 4. The two upper panels show estimates with reduced sets of controls, as indicated by their subtitles. Appendix Table A5 shows means for the family background and test score controls included. Heteroskedasticity robust standard errors are in parentheses. For predicted earnings at age 40, standard errors are bootstrapped.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7. OLS results. Effects of completing a vocational instead of a general upper-secondary program on labor market status and educational attainment at age 28.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Males						Females					
	Employed	Wage earner	Student	YOE	Post-sec. degree	University degree	Employed	Wage earner	Student	YOE	Post-sec. degree	University degree
Voc.	0.080***	0.060***	-0.055***	-0.689***	-0.419***	-0.194***	-0.009*	-0.021***	-0.021***	-1.441***	-0.540***	-0.088***
	(0.004)	(0.004)	(0.003)	(0.014)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.014)	(0.004)	(0.003)
N	79,070	79,070	79,070	79,070	79,070	79,070	82,362	82,362	82,362	82,362	82,362	82,362

Note. The table shows OLS estimates of effects of completing a vocational instead of a general upper secondary education. If a student completed both a general and a vocational upper secondary education program, they are categorized by their first completed program. The full set of controls are included: 92 municipality dummies, 3 cohort dummies, and 62 family background and 9th grade test score variables. Appendix Table A5 shows means for the family background and test score controls included. Labor market status at age 28 is represented by the indicator variables: Employed (that is, wage earner or self-employed), wage earner and student (that is, enrolled in education). Educational attainment by age 28 is represented by years of education (YOE), and indicator variables for having completed any post-secondary education and a university degree (a bachelors, masters or PhD degree). Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8. OLS results. Effects of completing a vocational instead of a general program on working hours per month and the hourly wage rate at age 28 conditional on being a wage earner.

	(1)	(2)	(4)	(5)
	Males		Females	
	Log hours per month	Log hourly wage	Log hours per month	Log hourly wage
Vocational	0.117***	0.072***	0.008	-0.044***
	(0.006)	(0.003)	(0.008)	(0.003)
N	63,456	63,456	62,226	62,226

Note. The table shows OLS estimates of effects of completing a vocational instead of a general upper secondary education. If a student completed both a general and a vocational upper secondary education program, they are categorized by their first completed program. The full set of controls are included: 92 municipality dummies, 3 cohort dummies, and 62 family background and 9th grade test score variables. Appendix Table A5 shows means for the family background and test score controls included. Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A9. OLS results. Effect of completing a vocational instead of a general program on the probability of having children by age 28.

	(1)	(2)
	Males	Females
Vocational	0.128***	0.121***
	(0.004)	(0.005)
N	79,070	82,362

Note. The table shows OLS estimates of effects of completing a vocational instead of a general upper secondary education. If a student completed both a general and a vocational upper secondary education program, they are categorized by their first completed program. The full set of controls are included. Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10. OLS results. Effects of completing a vocational instead of a general program on earnings at age 28 when limiting the sample to those not enrolled in education at age 28, at age 27-28, and at age 26-28 ($1,000)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Males				Females			
	Baseline	Not enrolled age 28	Not enrolled age 27-28	Not enrolled age 26-28	Baseline	Not enrolled age 28	Not enrolled age 27-28	Not enrolled age 26-28
Vocational	7.161***	5.297***	4.235***	3.352***	-2.713***	-3.623***	-4.108***	-4.460***
	(0.242)	(0.237)	(0.242)	(0.250)	(0.214)	(0.212)	(0.215)	(0.220)
N	79,070	73,199	68,482	63,013	82,362	75,657	70,618	64,821

Columns 1 and 5 show the baseline OLS estimates of Table 4. In the other models the sample is restricted to those not enrolled in education at age 28, those not enrolled at ages 27 and 28, and those not enrolled at ages 26, 27 and 28. Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11. MTE results. Effects on labor market status and educational attainment at age 28

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Males						Females					
	Employed	Wage earner	Student	YOE	Post-sec. degree	University degree	Employed	Wage earner	Student	YOE	Post-sec. degree	University degree
ATE	0.097***	0.064**	-0.079***	-1.091***	-0.561***	-0.299***	-0.050	-0.054	0.012	-1.792***	-0.672***	-0.220***
	(0.020)	(0.021)	(0.014)	(0.122)	(0.020)	(0.013)	(0.045)	(0.047)	(0.028)	(0.253)	(0.034)	(0.032)
ATT	0.154***	0.157***	-0.117***	-1.100***	-0.557***	-0.191***	0.029	0.032	-0.052***	-2.047***	-0.728***	0.011
	(0.023)	(0.023)	(0.016)	(0.174)	(0.023)	(0.023)	(0.020)	(0.020)	(0.014)	(0.128)	(0.020)	(0.030)
ATUT	0.057*	0.000	-0.053**	-1.086***	-0.563***	-0.373***	-0.072	-0.079	0.031	-1.719***	-0.655***	-0.287***
	(0.029)	(0.031)	(0.019)	(0.168)	(0.030)	(0.014)	(0.058)	(0.059)	(0.035)	(0.323)	(0.043)	(0.040)
N	79,070	79,070	79,070	79,070	79,070	79,070	82,362	82,362	82,362	82,362	82,362	82,362
Polynomial	2	2	2	4	2	2	2	2	2	4	2	3
p obs. het.	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
p unobs. het.	0.0023	0.0004	0.0009	0.0128	0.1515	<0.0001	0.0031	0.0020	0.0699	<0.0001	<0.0001	<0.0001

Note. The table shows effects of completing a vocational instead of a general upper secondary education. The table presents estimation results for MTE models with polynomials in the propensity score as control functions. We report results for a polynomial of order *n* if the *n* order terms are significant (for either k_0 or k), and if higher order terms are not significant. Labor market status at age 28 is represented by the indicator variables: Employed (that is, wage earner or self-employed), wage earner and student (that is, enrolled in education). Educational attainment by age 28 is represented by years of education (YOE), and indicator variables for having completed any post-secondary education and a university degree (a bachelors, masters or PhD degree). The last two rows show p-values of tests for observed and unobserved heterogeneity: the test for observed heterogeneity is a test that $\beta_1 - \beta_0 = 0$; the test for unobserved heterogeneity is a test that the coefficients of the polynomial for k are equal to zero. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12. MTE results. Effects on working hours per month and the hourly wage rate at age 28, conditional on being a wage earner.

	(1)	(2)	(4)	(5)
	Males		Females	
	Log hours per month	Log hourly wage	Log hours per month	Log hourly wage
ATE	0.111***	0.059***	0.049	-0.072**
	(0.023)	(0.016)	(0.070)	(0.027)
ATT	0.167***	0.114***	0.107**	-0.002
	(0.029)	(0.017)	(0.038)	(0.010)
ATUT	0.069*	0.018	0.033	-0.091**
	(0.035)	(0.025)	(0.088)	(0.034)
N	63,456	63,456	62,226	62,226
Order of polynomials in p	1	2	2	1
p-value, observed heterogeneity	<0.0001	<0.0001	<0.0001	<0.0001
p-value, unobserved heterogeneity	0.4247	0.0553	0.0879	0.1683

Note. The table shows the effects of completing a vocational instead of a general upper secondary education. The table presents estimation results for MTE models with polynomials in the propensity score as control functions. We report results for a polynomial of order *n* if the *n* order terms are significant (for either k_0 or k), and if higher order terms are not significant. The test for observed heterogeneity is a test that $\beta_1 - \beta_0 = 0$. The test for unobserved heterogeneity is a test that the coefficients of the polynomial for *k* are equal to zero. Bootstrapped standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13. MTE results. Effects on the probability of having children by age 28.

	(1)	(2)
	Males	Females
ATE	0.136***	0.065
	(0.028)	(0.047)
ATT	0.234***	0.172***
	(0.024)	(0.024)
ATUT	0.068	0.035
	(0.044)	(0.060)
N	79,070	82,362
Order of polynomials in p	2	2
p-value, observed heterogeneity	<0.0001	<0.0001
p-value, unobserved heterogeneity	<0.0001	0.0055

Note. The table shows effects of completing a vocational instead of a general upper secondary education on the probability of having at least one child by age 28. The table presents estimation results for MTE models with polynomials in the propensity score as control functions. We report results for a polynomial of order *n* if the *n* order terms are significant (for either k_0 or k), and if higher order terms are not significant. The test for observed heterogeneity is a test that $\beta_1 - \beta_0 = 0$. The test for unobserved heterogeneity is a test that the coefficients of the polynomial for *k* are equal to zero. Bootstrapped standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

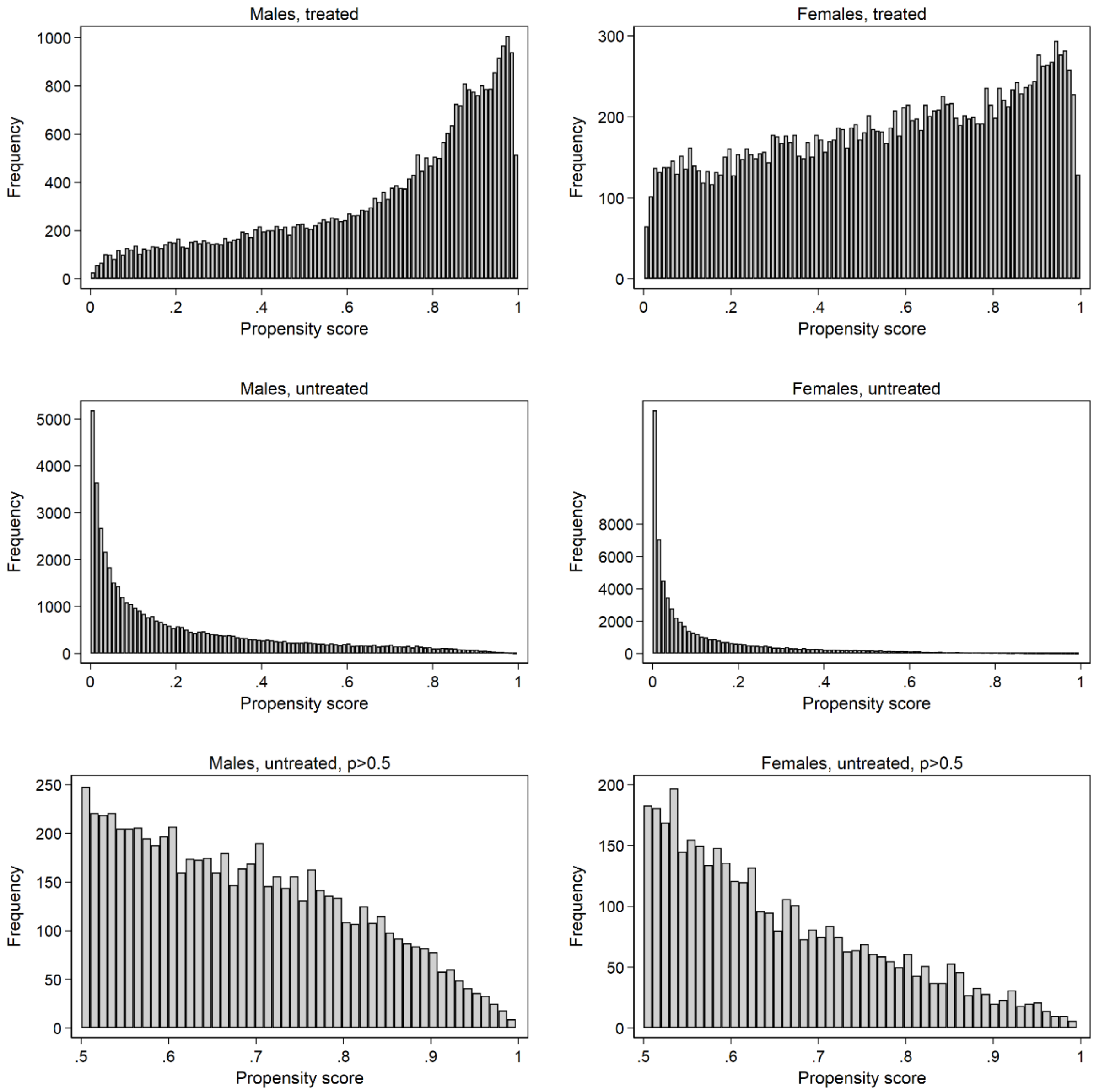
Table A14. MTE results. Effects of completing a vocational instead of a general program on earnings at age 28 when limiting the sample to those not enrolled in education at age 28, at age 27-28, and at age 26-28 ($1,000)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Males				Females			
	Baseline	Not enrolled age 28	Not enrolled age 27-28	Not enrolled age 26-28	Baseline	Not enrolled age 28	Not enrolled age 27-28	Not enrolled age 26-28
ATE	8.036***	5.333***	4.262**	3.051*	-6.739**	-6.520***	-6.845***	-6.937***
	(1.393)	(1.380)	(1.360)	(1.371)	(2.066)	(1.837)	(1.907)	(1.838)
ATT	14.946***	11.053***	8.480***	6.578***	5.131***	3.337**	2.748*	2.848*
	(1.581)	(1.549)	(1.598)	(1.650)	(1.187)	(1.179)	(1.226)	(1.223)
ATUT	3.254	1.132	0.960	0.020	-10.189***	-9.453***	-9.815***	-10.188***
	(2.137)	(2.087)	(1.994)	(2.066)	(2.632)	(2.367)	(2.478)	(2.412)
N	79,070	73,199	68,482	63,013	82,362	75,657	70,618	64,821
Order of polynomial in p	2	2	2	2	2	2	2	2

Note. The table shows MTE-based estimates of the effects of completing vocational instead of general education on earnings at age 28. Columns 1 and 5 show the baseline estimates of Table 6. In the other models the sample is restricted to those not enrolled in education at age 28, those not enrolled at ages 27 and 28, and those not enrolled at ages 26, 27 and 28. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A1. Common support: Frequency distribution of propensity score by treatment status and gender



Note. The two upper graphs show the frequency distribution of the propensity score for the treatment group by gender. The next two graphs show the distribution for the control group by gender. Here the distribution is highly skewed, making it difficult to assess whether there is common support for high values of the propensity score. In consequence, the last two graphs show frequencies for propensity scores above 0.5.

Appendix B. Standard IV estimates

Standard IV (2SLS) estimates are difficult to interpret because our instruments are continuous, and also because we have more than one instrument - which implies that the monotonicity assumption of Imbens and Angrist (1994), needed for a LATE interpretation of the estimates, is very restrictive; see Mogstad, Torgovitsky, and Walters (2021). Use of the estimated propensity score as instrument in a linear IV model ensures that the monotonicity assumption is satisfied. It is common to report such IV estimates for comparison when applying MTE methods; see for instance Carneiro, Heckman and Vytlacil (2011) and Nybom (2017). Table B1 shows standard IV estimates and linear IV estimates with the estimated propensity score as instrument. Upon using the estimated propensity score as instrument, the estimates are rather close to the ATT effects in Table 6. The standard IV estimates are very different from the estimates using the propensity score as instrument, both in terms of the size of the effect and in some cases also the sign. This is not unusual; see for instance Heckman (2010), and Brinch, Mogstad, and Wiswall (2017). We do not discuss these estimates further because, as stated above, they do not have a clear interpretation in this application.

Table B1. Linear IV estimates with $P(Z)$ and Z^e as instruments. Effect of completing vocational instead of general upper secondary education on earnings ($1,000).

	(1)	(2)	(3)	(4)
	Males		Females	
	Earnings age 28	Earnings age 40	Earnings age 28	Earnings age 40
$P(Z)$ as instrument	11.830***	4.261***	3.007***	8.740***
	(0.886)	(1.196)	(0.661)	(0.895)
Z^e as instruments	31.239***	10.986**	-1.870	-10.049*
	(3.951)	(3.861)	(5.461)	(4.593)
N	79,070	79,070	82,362	82,362

Note. The excluded instruments (Z^e) are distances to vocational and general education and their interactions with the math test score and a dummy for missing math test score. The first row shows linear IV estimates with *P(Z)* as instrument, where *P(Z)* is constructed from the parameter estimates of the first-stage logit models for which we report average marginal derivates in Columns 1 and 3 of Table 5. The second row shows standard linear IV estimates with Z^e as instruments. All models control for socio-demographics, test scores and cohort and municipality dummies. Appendix Table A5 shows the family background and test score controls included. Bootstrapped standard errors are in parentheses (except when Z^e are the instruments and earnings at age 28 is the outcome, for which we report analytical heteroskedasticity robust standard errors).

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C. Robustness and specification checks

Appendix C.1. Alternative specifications of the distance instruments

In the logit models for choice of education we include, in the main specification, cross terms between the distance instruments and math test scores. This tends to reduce the estimated standard errors of the treatment effect parameters, although not very much and not always. Table C1 shows treatment effect estimates when we omit these cross terms from the choice models (see Columns 2 and 4 in Table 5 for the first stage results). The treatment effect estimates are very similar to the corresponding baseline estimates in Table 6 and there are no significant differences.

In the main specification we use as instruments the weighted distances to the nearest vocational and general schools, in order to allow for the fact that different tracks are offered by different institutions within each of these two main types of upper secondary programs; see Section IV.C. Our results do not differ significantly if we instead use the distance to the nearest vocational school and the distance to the nearest general school as instruments. These simpler instruments are also very strong in the first stage; see Table C2.

Table C1. Robustness check: Distance instruments not interacted with math test score in first stage. Effects of completing vocational instead of general education on earnings ($1,000).

	(1)	(2)	(3)	(4)
	Males		Females	
	Earnings age 28	Earnings age 40	Earnings age 28	Earnings age 40
		Baseline estimates		
ATE	8.036^{***}	-8.682^{***}	-6.739^{**}	-12.195^{***}
	(1.393)	(1.053)	(2.066)	(2.656)
ATT	14.946^{***}	3.701^{*}	5.131^{***}	8.138^{***}
	(1.581)	(1.494)	(1.187)	(1.891)
ATUT	3.254	-17.251^{***}	-10.189^{***}	-18.108^{***}
	(2.137)	(1.409)	(2.632)	(3.384)
		No cross terms in 1st stage		
ATE	8.416^{***}	-8.370^{***}	-6.253^{**}	-9.561^{***}
	(1.466)	(1.057)	(2.149)	(2.870)
ATT	14.908^{***}	3.784^{*}	5.184^{***}	8.490^{***}
	(1.580)	(1.501)	(1.219)	(1.907)
ATUT	3.922^{+}	-16.780^{***}	-9.576^{***}	-14.812^{***}
	(2.189)	(1.413)	(2.729)	(3.660)
N	79,070	79,070	82,362	82,362
Order of polynomials in p	2	1	2	3

Note. The table shows the main results from Table 6 (upper panel) on treatment effects of completing vocational instead of general upper secondary education, and the results when we omit cross terms between the distance instruments and math test scores in the first stage (lower panel). Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C2. Robustness check: Instruments are distance to nearest vocational and general school (instead of weighted average distance to nearest sub-categories of vocational and general schools). Effects of completing vocational instead of general education on earnings ($1,000).

	(1)	(2)	(4)	(5)
	Males		Females	
	Earnings age 28	Earnings age 40	Earnings age 28	Earnings age 40
	Baseline estimates			
ATE	8.036***	-8.682***	-6.739**	-12.195***
	(1.393)	(1.053)	(2.066)	(2.656)
ATT	14.946***	3.701*	5.131***	8.138***
	(1.581)	(1.494)	(1.187)	(1.891)
ATUT	3.254	-17.251***	-10.189***	-18.108***
	(2.137)	(1.409)	(2.632)	(3.384)
	Distance to nearest vocational and general school			
ATE	8.063***	-8.478***	-6.313**	-11.962***
	(1.447)	(1.058)	(2.124)	(2.636)
ATT	14.855***	3.643*	4.978***	8.103***
	(1.569)	(1.489)	(1.206)	(1.883)
ATUT	3.362	-16.865***	-9.595***	-17.798***
	(2.160)	(1.418)	(2.703)	(3.355)
N	79,070	79,070	82,362	82,362
Order of polynomials in p	2	1	2	3
Chi squared, 1st stage exclusion restrictions	289.7	289.7	98.1	98.1
p-value, 1st stage exclusion restrictions	1.333e-59	1.333e-59	6.249e-19	6.249e-19
Degrees of freedom, exclusion restrictions	6	6	6	6

Note. This table presents estimation results for models similar to the baseline in Table 6, except that the instruments are the distance to the nearest vocational school and the distance to the nearest general school, instead of the weighted average distances to nearest sub-categories of vocational and general schools. As in the baseline specification, cross terms between distance and math test score variables are also used as instruments. The models include polynomials in the propensity score. We report results for a polynomial of order n if the n order terms are significant (for either k_0 or k), and if higher order terms are not significant. The first stage statistics in the last rows can be compared to Table 5 (Columns 1 and 3) with the baseline first stage results. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C.2. Total income

Some females are on maternity leave at age 28, and maternity pay in the form of public transfers is not included in our measure of earnings whereas maternity pay provided by the employer is included. We therefore conduct a robustness check using total income at age 28 as the outcome. In addition to earnings and maternity pay, total income includes student grants, unemployment benefits, social security benefits, and interest and dividend income. The overall pattern of results is not changed very much by using total income instead of earnings as the outcome at age 28; see Table C3. The effect of vocational education on total income is more positive for females, which may be explained by the fact that it is more common for females with a post-secondary education to have a job that entitles them to maternity leave on full pay (provided by the employer), whereas females with vocational education rely more on maternity pay in the form of public transfers. For males, the ATE, ATT and ATUT effects on total income are slightly (although not significantly) smaller than the effects on earnings in Table 6. This difference is consistent with more males who have a general upper secondary education receiving student grants or unemployment benefits at age 28; see Table 1.

Table C3. Robustness check: Results using total income at age 28 as the outcome instead of earnings ($1,000).

	(1)	(2)	(3)	(4)
	Males		Females	
	Earnings age 28 (baseline)	Total income age 28	Earnings age 28 (baseline)	Total income age 28
ATE	8.036***	6.890***	-6.739**	-4.139**
	(1.393)	(1.236)	(2.066)	(1.349)
ATT	14.946***	12.932***	5.131***	6.816***
	(1.581)	(1.404)	(1.187)	(0.916)
ATUT	3.254	2.704	-10.189***	-7.325***
	(2.137)	(1.844)	(2.632)	(1.719)
N	79,070	78,987	82,362	82,332
Mean of outcome with vocational education	50.023	53.401	31.248	41.986
Mean of outcome with general education	47.117	51.242	40.251	47.711

Note. The table shows the effects of completing vocational instead of general education on outcomes at age 28: Earnings as in Table 6, and total income including public transfers and interest and dividend income. The order of the polynomial in the propensity score is 2 in all analyses. The number of observations is a little smaller in the models for total income. This is because we trim observations with income below zero and above DKK 1m. The last two rows show descriptive means of earnings and total income for individuals who completed vocational and general education respectively. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C.3. Robustness and specification checks for predicted earnings at age 40

The main results in Columns 2 and 4 of Table 6 are based on predicting earnings at age 40 using a sample of older cohorts born in 1974-77, and using covariates measured in 2018 (the last year in the dataset), as well as 3 years of lagged variables for earnings, labor market status, educational level and their interactions. Table C4 shows results for two alternative specifications of the model for predicting earnings at age 40. First, we include 3 extra older cohorts when estimating the model; specifically, we use data for cohorts 1971-77 instead of cohorts 1974-77. Second, we omit the 3 years of lagged variables in the model. The results for these two alternative specifications are very similar to the baseline results in Table 6. The robustness of the results with respect to the inclusion of extra cohorts provides some confidence that the assumption of comparability of samples is plausible. The robustness concerning the number of lags included reflects the fact that many highly relevant short-term outcome variables are included; see Section IV.B.

Table C4. Robustness checks: Results for alternative specifications of predicted earnings at age 40 ($1,000).

	(1)	(2)	(3)	(4)	(5)	(6)
		Males			Females	
	Baseline	Cohorts 71-77	No lags	Baseline	Cohorts 71-77	No lags
ATE	-8.682***	-9.053***	-9.601***	-12.195***	-11.970***	-12.700***
	(1.053)	(1.101)	(0.989)	(2.656)	(2.600)	(2.625)
ATT	3.701*	3.232*	2.991*	8.138***	8.833***	7.538***
	(1.494)	(1.414)	(1.435)	(1.891)	(1.749)	(1.800)
ATUT	-17.251***	-17.554***	-18.314***	-18.108***	-18.020***	-18.586***
	(1.409)	(1.473)	(1.312)	(3.384)	(3.297)	(3.326)
N	79,070	79,070	79,070	82,362	82,362	82,362
Order of polynomials in p	1	1	1	3	3	3

Note. Columns 1 and 4 in this table show the main results for predicted earnings at age 40 from Table 6 on the treatment effects of completing vocational instead of general upper secondary education. Columns 2 and 5 show results when we base predicted earnings on 3 extra older cohorts. Columns 3 and 6 show results when we omit lagged variables of earnings, labor market status and educational level in the model for predicting earnings at age 40. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C5. Results when only educational status at age 28-31 is used to predict earnings at age 40. Effects of completing a vocational instead of a general program on earnings at age 40 ($1,000).

	(1)	(2)	(3)	(4)
	Males		Females	
	Baseline	Education	Baseline	Education
ATE	-8.682***	-14.336***	-12.195***	-13.763***
	(1.053)	(0.816)	(2.656)	(1.974)
ATT	3.701*	-8.977***	8.138***	0.282
	(1.494)	(1.157)	(1.891)	(1.392)
ATUT	-17.251***	-18.047***	-18.108***	-17.851***
	(1.409)	(1.082)	(3.384)	(2.490)
N	79,070	79,070	82,362	82,362
Order of polynomials in p	1	2	3	3

Note. Columns 1 and 3 in this table show the main results for predicted earnings at age 40 from Table 6 on the treatment effects of completing vocational instead of general upper secondary education. Columns 2 and 4 show results when we base predicted earnings on only educational attainment (290 categories of education in terms of both level and field of study) by age 28-31. The models include polynomials in the propensity score. We report results for a polynomial of order *n* if the *n* order terms are significant (for either k_0 or k), and if higher order terms are not significant. Bootstrapped standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The inclusion of a broad spectrum of short-term outcomes by age 28-31 is indeed important for our results. This is illustrated in Table C5, which compares our baseline estimates of treatment effects on predicted earnings at age 40 with estimates when only educational attainment at age 28-31 and gender are used to construct predicted earnings.[1] Even though we measure educational attainment very precisely in terms of both educational level and field of study (we have about 290 specific categories of the highest completed education), the two sets of estimates are substantially different. In particular, the differences are marked for the ATT, which for males is positive with a value of $3,700 in the baseline specification, but negative with a value of -$9,000 in the alternative specification; for females, the baseline estimate is $8,100, whereas the alternative estimate is close

[1] Predicted long-term earnings based on (only) the type of education attained correspond to what is sometimes called earnings-weighted education; see for instance Björklund and Sundström (2006).

to zero and is not statistically significant. These differences reflect the fact that, even with very detailed information on education, important earnings heterogeneity exists within education groups which is related to treatment; the additional short-term outcomes, including earnings and labor market status, capture such heterogeneity. This indicates that education and gender alone do not capture the causal chain from treatment to long-term earnings.

We now assess the validity of the surrogacy and comparability assumptions. Following the idea of Athey et al. (2019), we treat an observed short-term outcome, in this case earnings at age 28, as if it were unobserved, and predict it using lagged values of earnings and other short-term outcomes (at age 25-27), including educational attainment and labor market status, as well as pre-treatment variables. Thus, we estimate the model for earnings at age 28 on the sample of older cohorts, and use the estimates to predict earnings at age 28 in the main sample of cohorts 1986-89. Finally, we estimate the ATE, ATT and ATUT parameters in the MTE model using the predicted earnings at age 28. For all three treatment effect parameters and for both genders, the estimates using predicted and observed earnings at age 28 are very similar; there are no significant differences (see Columns 1-3 in Table C6). This strengthens our confidence in the validity of the surrogacy and comparability assumptions.

As a further check, we predict earnings at age 28 based on an estimation using the main sample instead of the sample of older cohorts. Comparison of the result using this prediction (Column 4 in Table C6) with the result using observed earnings (Columns 1-2) is a more direct check of the surrogacy assumption, because the assumption of comparability of samples is irrelevant when predictions are based on the main sample only. Similarly, comparison of results using predictions based on the main sample (Column 4) with results using predictions based on the sample of older cohorts (Column 3) is a more direct check of the comparability assumption. These comparisons again indicate that the assumptions are plausible; there are no significant differences.

Table C6. Specification check. Effects on earnings at age 28 ($1,000): Observed, predicted from the sample of older cohorts, and predicted from the main sample. Predictions are based on observations at age 25-27

	(1) Earnings age 28 (baseline)	(2) Earnings age 28 (reduced sample)	(3) Fitted (older cohorts)	(4) Fitted (main sample)	(5) Fitted (main sample, all controls)
			Males		
ATE	8.036***	8.018***	7.157***	7.743***	7.803***
	(1.393)	(1.440)	(1.212)	(1.302)	(1.294)
ATT	14.946***	15.016***	13.657***	14.729***	14.625***
	(1.581)	(1.607)	(1.311)	(1.403)	(1.377)
ATUT	3.254	3.159	2.643	2.892	3.066
	(2.137)	(2.123)	(1.815)	(1.940)	(1.926)
N	79,070	78,563	78,563	78,563	78,563
			Females		
ATE	-6.739**	-6.622***	-6.104***	-5.580***	-5.302***
	(2.066)	(1.977)	(1.555)	(1.571)	(1.562)
ATT	5.131***	4.968***	4.275***	4.826***	4.746***
	(1.187)	(1.181)	(1.100)	(0.994)	(0.979)
ATUT	-10.189***	-9.998***	-9.128***	-8.612***	-8.229***
	(2.632)	(2.520)	(1.960)	(1.983)	(1.973)
N	82,362	81,969	81,969	81,969	81,969

Note. The table shows in Column 1 the main results for earnings at age 28 from Table 6 on treatment effects of completing vocational instead of general upper secondary education. Column 2 also shows results using observed earnings at age 28. It differs from Column 1 only because the sample is a little smaller and identical to the sample used in Columns 3-5 (see below for more details). Columns 3-5 show results using instead predicted earnings at age 28 conditional on outcomes by age 27, based on estimates for the sample of older cohorts (Column 3) and the main sample (Columns 4-5). The predicted earnings used in Columns 4 and 5 are different because the set of pre-treatment controls used for prediction differs; in Column 5, all control variables used in the MTE analysis are used, whereas in Column 4 a more limited set of pre-treatment controls is used. The number of observations is about 0.5% smaller when predicted earnings at age 28 are used as the outcome. Column 2 therefore shows the results for observed earnings estimated on this reduced sample. The reason for the small loss of observations when we use predicted earnings is that, to predict earnings at age 28, we use short-term outcomes measured at age 25-27, which are observed only for persons living in Denmark at these ages. When we predict earnings at age 40, the sample restriction is somewhat different. For instance, for cohort 1986, for which we can observe outcomes up to age 31, we condition on outcomes at age 28-31. The order of the polynomial in the propensity score is 2 in all models. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Finally, we check whether, conditional on the large set of intermediate outcomes used in predicting long-term earnings, the more limited set of pre-treatment controls used in the predictions is sufficient compared to the larger set of controls used in the main MTE analysis. This larger set of controls includes variables which are observed in the main sample but not in the sample of older cohorts, most importantly 9th grade test scores. When we use the main sample to predict earnings at age 28, we can include this larger set of controls. The resulting estimates of treatment effect parameters are very similar to the estimates using the more limited set of pre-treatment controls (see Columns 4 and 5 of Table C6). Again, there are no significant differences.

Appendix C.4. Enrollment as treatment

In our main analysis, the treatment is completion of a vocational instead of a general upper secondary education. We consider here an alternative definition of treatment, namely enrollment in vocational instead of general education; Carneiro, Lokshin, and Umapathi (2017) also consider enrollment in secondary education as the treatment. These two treatments are distinct, because dropout rates from especially vocational education are high. In our analysis, with enrollment as treatment, the sample size is considerably larger because we include all individuals who enroll in upper secondary education (vocational or general), whereas our main analysis is restricted to those who complete an upper secondary education. We consider enrollment up to age 25, and individuals who by age 25 had enrolled in both types of programs are categorized by their first program. About 90% of those enrolled in a general program complete such a program. In contrast, among those enrolled in a vocational program, the proportion completing such a program is only about 55%, and about 35% do not complete any upper secondary education program; see Table C7. The high dropout rates from vocational programs may partly be due to difficulties in finding apprenticeship positions and the much weaker academic skills of vocational education enrollees who do not complete an upper secondary education.

Table C7. Educational status at age 25 by type of upper secondary education enrollment after lower secondary school.

	Males				Females			
	Type of upper secondary enrollment				Type of upper secondary enrollment			
Upper secondary educational	General		Vocational		General		Vocational	
attainment by age 25	N	Pct	N	Pct	N	Pct	N	Pct
Completed general program	43,657	87.13	3,122	6.21	60,403	90.91	3,525	11.28
Completed vocational program	2,430	4.85	29,861	59.44	2,248	3.38	16,186	51.78
Did not complete any program	4,018	8.02	17,256	34.35	3,795	5.71	11,547	36.94
Total	50,105	100.00	50,239	100.00	66,446	100.00	31,258	100.00

Note. If a student had completed (or enrolled in) both a general and a vocational upper secondary education program, they are categorized by their first completed (started) program.

When enrollment (instead of completion) is the treatment, mean outcomes at age 28 and 40 are less favorable, especially for those enrolled in vocational programs, because the sample includes those who enrolled in, but did not complete, an upper secondary education (compare Table C8 to Table 1). The first-stage results for enrollment into vocational instead of general education are fairly similar to the main results for completion (compare Table C9 to Table 5). The distance instruments are clearly significant.

Table C8. Earnings by gender and type of upper secondary *enrollment* ($1,000)

	Males		Females		All
	General	Vocational	General	Vocational	
Earnings age 28	46.519	42.564	39.498	26.170	39.949
Predicted earnings age 40	82.338	56.697	63.474	38.320	62.557
N	50,105	50,239	66,446	31,258	198,048

Note. If a student enrolled in both a general and a vocational upper secondary education program before age 25, they are categorized by their first enrollment.

Table C9. First stage logit model. Decision on *enrolling* in vocational instead of general upper secondary education. Average marginal derivatives

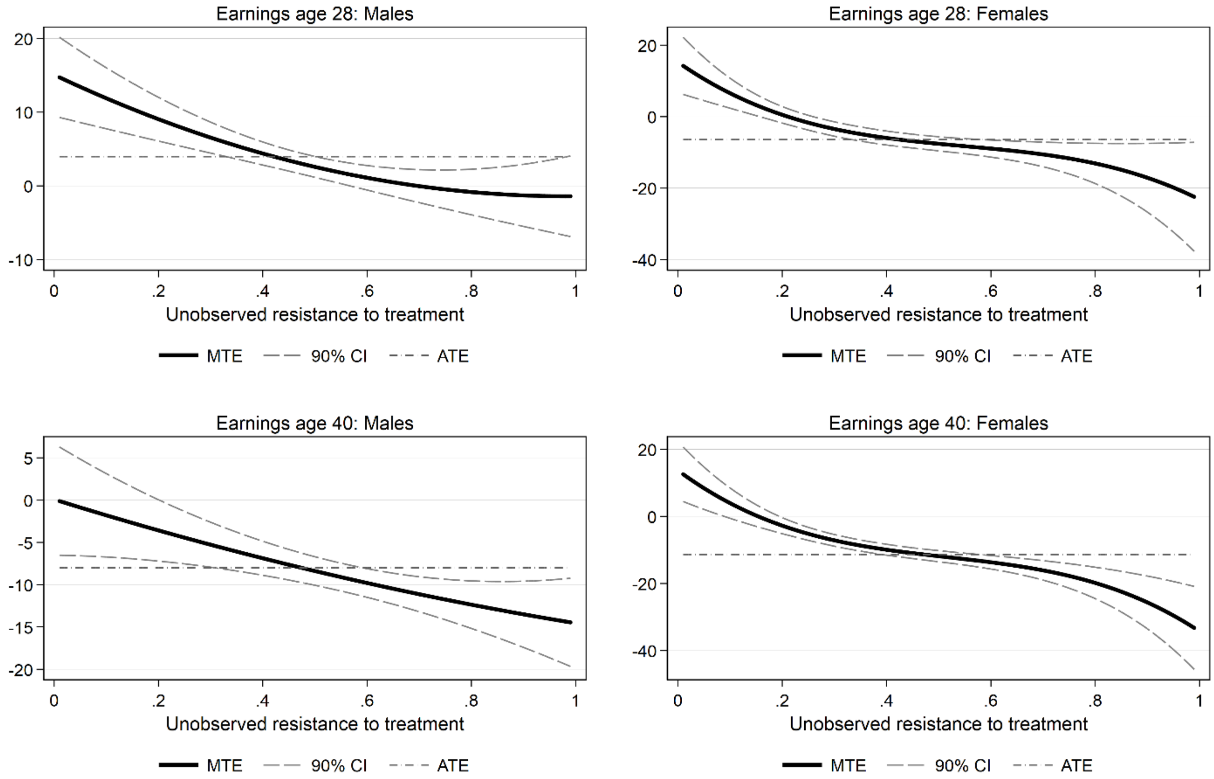
	(1)	(2)	(3)	(4)
	Male	Male	Female	Female
Distance to vocational education (km)	-0.0005	-0.0005	-0.0005	-0.0005
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Distance to general education (km)	0.0030***	0.0031***	0.0016***	0.0017***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Cross terms: distance×math score	X		X	
Control variables	X	X	X	X
N	100,344	100,344	97,704	97,704
Chi squared, exclusion restrictions	260.1	178.6	112.4	46.0
p-value, exclusion restrictions	2.836e-53	1.657e-39	6.448e-22	1.042e-10
Degrees of freedom, exclusion restrictions	6	2	6	2
Chi squared, cross terms	79.9		68.0	
p-value, cross terms	1.808e-16		6.092e-14	
Degrees of freedom, cross terms	4		4	

Note. The dependent variable is an indicator for enrolling in a vocational instead of a general upper secondary education. Average marginal derivatives based on the logit model estimates are shown for the two distance measures. For each individual, we calculate the effect of increasing the distance by one unit (1 km), while holding all control variables fixed, on the probability of enrolling in vocational instead of general education. In Columns 2 and 4 the exclusion restrictions are the two distance measures. In Columns 1 and 3 they also include cross terms between each distance variable and two control variables, namely the math test score and an indicator variable for missing math test score. The first chi-square test is for all exclusion restrictions, and the second is for only the cross terms in Columns 1 and 3. All control variables are included in the models: Socio-demographics, test scores and cohort and municipality dummies. Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

For both genders and both outcomes, the MTE curves are decreasing; see Figure C1.

Figure C1. Marginal treatment effects on earnings of *enrolling* in vocational instead of general education ($1,000)



Note. The two upper panels show MTEs on earnings at age 28, and the two lower panels show MTEs on predicted earnings at age 40. The MTEs are evaluated at the sample means of covariates. The order of the polynomials in the propensity score is reported in Table C10.

Table C10 shows the treatment effect estimates when enrollment is the treatment. The ATT effects of enrolling in vocational instead of general education are smaller than the corresponding main estimates for completion in Table 6, which can be explained by the high dropout rates for those enrolling in vocational education, and the lower ability and more disadvantaged backgrounds of dropouts. The ATT for predicted earnings at age 40 is small and not significant for males in Table C10. It is \$4,300 for females, whereas in Table 6 it is \$8,100. In contrast, the ATUT and ATE effects are approximately the same in the two tables. The only exception is for males' earnings at age 28, for which the ATUT is -\$1,800 when enrollment is treatment, but \$3,300 when completion is treatment.

As with the main results in Table 6, the estimates for enrollment as treatment in Table C10 are consistent with selection on gains, because the ATT estimates are substantially and significantly larger than the ATUT estimates. The PRTE estimates for enrollment in vocational instead of general education tend to be smaller (or more negative) than the corresponding estimates for completion effects in Table 6.

The ATT estimates are smaller in Table C10 than in Table 6. This difference is caused not by the change in the definition of the treatment (from completion to enrollment in vocational education), but by the larger sample in the enrollment analysis with many low-ability dropouts from vocational education. Thus, if we restrict the sample to be the same as in the main analysis by excluding individuals who do not complete an upper secondary education by age 25, the ATT, ATUT and ATE estimates using enrollment in vocational instead of general education as treatment become very similar to our main estimates in Table 6 with completion as treatment. These further results are not shown, but are available on request.

Table C10. Effects of *enrolling* in vocational instead of general education on earnings ($1,000)

	(1)	(2)	(3)	(4)
	Males		Females	
	Earnings age 28	Earnings age 40	Earnings age 28	Earnings age 40
Conventional treatment effects				
ATE	3.969***	-8.004***	-6.403**	-11.343***
	(1.038)	(1.178)	(1.989)	(1.651)
ATT	9.715***	0.292	2.511+	4.325**
	(1.413)	(1.802)	(1.377)	(1.631)
ATUT	-1.791	-16.320***	-10.612***	-18.748***
	(1.528)	(1.394)	(2.833)	(2.340)
Policy-relevant treatment effects				
PRTE: p + 0.01	4.704***	-7.968***	0.731	-3.222
	(1.352)	(1.515)	(2.228)	(2.177)
PRTE: Distance general + 1 km	4.279***	-7.537***	-2.202	-3.975*
	(1.012)	(1.255)	(1.514)	(1.551)
PRTEs by 9th grade math scores				
PRTE: p + 0.01, math low	6.459***	1.136	-4.766+	-4.153*
	(1.724)	(1.685)	(2.562)	(2.057)
PRTE: p + 0.01, math medium	5.462***	-6.129***	3.365	-0.006
	(1.121)	(1.323)	(2.096)	(2.209)
PRTE: p + 0.01, math high	5.173*	-14.584***	6.691+	-2.559
	(2.113)	(2.529)	(3.960)	(3.954)
N	100,344	100,344	97,704	97,704
Order of polynomials in p	2	2	3	3
p-value, observed heterogeneity	<0.0001	<0.0001	<0.0001	<0.0001
p-value, unobserved heterogeneity	0.0009	0.0175	<0.0001	<0.0001

Note. The MTE models include polynomials in the propensity score. We report results for a polynomial of order *n* if the *n* order terms are significant (for either k_0 or k), and if higher order terms are not significant. The models include all control variables: Socio-demographics, test scores and cohort and municipality dummies. The test for observed heterogeneity is a test that $\beta_1 - \beta_0 = 0$. The test for unobserved heterogeneity is a test that coefficients of the polynomial for k are equal to zero. The mean propensity score is 0.501 for males and 0.320 for females. The five PRTEs are effects per net individual shifted for five different policy-related shifts in the propensity score. The first policy augments the propensity score by 1 percentage point for all observations. The second policy increases the distance to general schools by 1 km for everybody. This results in an average increase in the propensity score by 0.3 percentage points for males and 0.2 percentage points for females. The last three sets of PRTEs are for policies which increase the propensity score by 0.01, but only for those with low, medium, and high 9th grade math scores, respectively. Bootstrapped standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C.5. Validity of the distance instruments

We now discuss some specification checks related to the validity of the distance instruments. One issue is that the instruments might predict selection into the estimation sample. Thus, in the main analysis we condition on completing an upper secondary education (vocational or general) by age 25, and in the alternative analysis discussed in Appendix C.4 we condition on having enrolled in an upper secondary education by age 25. The full sample (including those who do not enroll in any upper secondary program) consists of 210,339 individuals, of whom 5.8% do not enroll in any upper secondary program and 23.3% do not complete any upper secondary program (see Appendix Table A1). As discussed in Appendix C.4, most dropouts were enrolled in vocational programs. If enrollment in vocational programs causally increases the risk of not completing any upper secondary program, our main analysis that conditions on completion would miss an important negative effect of vocational programs.

Consequently, it is important to test whether the instruments predict selection into the estimation sample. We check whether the instruments are significant in a regression on the full sample where the dependent variable is a dummy for having completed (or enrolled in) an upper secondary education by age 25. For simplicity, we check this using OLS models which do not include cross terms between distance and other variables, and which are estimated for males and females together. An advantage of using OLS models is that we can report the F-test statistics for the exclusion restrictions which are conventionally used to assess whether instruments are weak. The two distance instruments are not significant at the 5% level. In the model for completion, the F-test statistic is 2.5 with a p-value of 0.084, and in the model for enrollment the F statistic is 1.9 with a p-value of 0.150; see Table C11. This is in sharp contrast to the clearly significant estimates in our first stage for completion of (enrollment in) vocational rather than general education. Here, the corresponding F statistics are 129.6 and 91.8. Thus, the distance instruments do not predict the selection into the estimation sample.

As discussed in Section III.A, the distance instruments affect the type of education completed conditional on enrollment. This is the case when we condition on enrollment in a vocational program, in a general program, or in either of the two types of programs. For each enrollment condition we have three completion outcomes: vocational, general or no program. Table C12 shows the results of estimating three sets of seemingly unrelated regression (SUR) models.[2] Thus, Columns 1-3 condition on enrollment in vocational education and show results for models in which the dependent variable is no completion, completion of vocational education and completion of general education. Similarly, Columns 4-6 condition on enrollment in general education, and Columns 7-9 condition on enrollment in either vocational or general education. The distance instruments do not significantly affect the probability of not completing any program (see the upper Chi-squared test statistics in Columns 1, 4 and 7 with p-values of 0.61, 0.12 and 0.25, respectively); this is consistent with the results discussed above concerning selection into the estimation sample. The distance instruments do, however, predict the probability of whether a vocational or general program is completed (see the test statistics in Columns 2, 3, 5, 6, 8 and 9, and the lower test statistics in Columns 1, 4 and 7). As discussed in Section III.A, this may question the validity of the instruments for the analysis of effects of enrollment in vocational instead of general programs.

In our main analysis of the effect of completing vocational instead of general education, the exclusion restriction might not hold if the distance instruments affect not only the type, but also the quality, of the upper secondary education completed. Consequently, it is

[2] Almost identical results are obtained by estimating multinomial logit models (results not shown).

reassuring that the distance instruments do not affect GPA in general upper secondary school; see Table C13.[3]

As placebo tests of instrument validity, we estimate choice models for completion of compulsory school (9th grade). We do not expect distance to upper secondary school to affect completion of lower secondary school except through possible confounders. We omit 9th grade test score controls from these regressions and consider four different definitions of completion: At least one non-missing 9th grade test score; non-missing test scores in math and Danish; passing test scores in math and Danish; and passing test scores in math, Danish, English and science. The two distance instruments are jointly non-significant for all four definitions of completing 9th grade when we use the full sample, and also when we condition on enrollment in (or completion of) an upper secondary education by age 25; see Table C14.

As another check of instrument validity, we estimate the first-stage models excluding controls for test scores (and teachers' marks for the year's work).[4] Large differences in the estimates of the coefficients on distance, compared to our main specification with test score controls, might indicate that the inclusion of additional controls for student ability (that are not in our data) would also have large effects on the coefficients and significance of the instruments. Estimated coefficients of distance variables tend to be a little smaller numerically when we exclude test score controls for math and science, but slightly larger numerically when we instead exclude test score controls for Danish and English, or when we exclude all test score controls; see Table C15. Although these differences are small and not statistically significant, they might suggest some endogeneity of the instruments, which should generate caution in the interpretation of results.

[3] We have no information on GPA or other quality indicators for vocational programs.

[4] Nybom (2017) apply a similar check for a first-stage selection model of choice of college, with distance as instrument.

Table C11. Tests of significance of the two distance instruments in four OLS selection models: (1) Enroll into an upper secondary education; (2) complete an upper secondary education; (3) enroll in vocational instead of general education; (4) complete vocational instead of general education

	(1) Enroll vs. not	(2) Complete vs. not	(3) Enroll in vocational vs. general	(4) Complete vocational vs. general
F-test, distance instruments	1.9	2.5	91.8	129.6
p-value	0.150	0.084	1.41e-40	5.75e-57
Degrees of freedom, constraints	2	2	2	2
N	210,339	210,339	198,048	161,432

Note. This table reports the F-tests for the significance of the two distance instruments in four OLS selection models. The two models in Columns 3 and 4 correspond to our first-stage equations reported in Table C9 and Table 5 (Columns 2 and 4) except for the functional form (OLS versus logit) and the fact that here we estimate the models for males and females together (with control for gender). These models are conditional on having enrolled in (completed) some upper secondary education by age 25. The models in Columns 1 and 2 are models for selection into these two estimation samples. All models include the full set of controls (see Table A5).

Table C12. Effect of distance on completion conditional on enrollment; SUR models.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Enrollment:	Vocational			General			Vocational or general		
Completion:	No	Vocational	General	No	Vocational	General	No	Vocational	General
Distance to vocational	-0.0003	0.0001	0.0002	-0.0003	-0.0000	0.0003	-0.0003	-0.0003	0.0006*
	(0.0005)	(0.0005)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
Distance to general	-0.0000	0.0019***	-0.0019***	0.0004*	0.0003	-0.0007*	0.0004^{+}	0.0026***	-0.0030***
	(0.0005)	(0.0005)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
N	81,497			116,551			198,048		
Chi squared	1.0	44.7	110.0	4.2	5.0	8.8	2.7	186.4	250.0
p-value	0.6126	<0.0001	<0.0001	0.1244	0.0839	0.0126	0.2535	<0.0001	<0.0001
Degrees of freedom	2	2	2	2	2	2	2	2	2
Chi squared	118.0			9.9			278.4		
p-value	<0.0001			0.0425			<0.0001		
Degrees of freedom	4			4			4		

Note. This table shows estimates for the distance instruments from three sets of SUR (seemingly unrelated regression) models including all covariates (and data for both genders). For each equation we show the coefficient estimates and a Chi-squared test for both coefficients being zero, and for each of the three sets of SUR models we show also the overall test of zero coefficients. The models in Columns 1-3 are conditioned on the sample whose first enrollment was in a vocational program; Columns 4-6 are conditioned on the sample whose first enrollment was in a general program, and Columns 7-9 are for the full enrollment sample. In each case, the completion outcome can be No (neither a vocational nor a general program was completed), Vocational (the first completed program was vocational) or General (the first completed program was general). Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C13. Effect of distance instruments on GPA in general upper secondary school. OLS models

	(1) First enrollment and completion of general school	(2) First completion of general school	(3) All with non-missing GPA from general school
Distance to vocational education (km)	0.0040	0.0092	-0.0008
	(0.0176)	(0.0172)	(0.0175)
Distance to general education (km)	0.0061	-0.0037	0.0138
	(0.0176)	(0.0172)	(0.0175)
N	102,765	109,297	114,386
F-test of distance instruments	0.4	0.2	0.8
p-value	0.6613	0.8159	0.4561
No. of restrictions of the test	2	2	2

Note. The three OLS models include all control variables, and are estimated for both genders together. They differ only in terms of the sample used: Column 1 includes all individuals who enrolled in a general upper secondary education after 9th grade and completed a general upper secondary program as their first upper secondary degree; Column 2 includes in addition those who first enrolled in vocational school but then switched to general and completed; Column 3 includes all individuals with a non-missing GPA from upper secondary school (including those who first completed a vocational education and then a general). The F-test is a test that the two distance variables both have zero coefficients. Our variable for GPA in general upper school has a mean of 66 and a standard deviation of 24. It is constructed by multiplying the recorded GPA by a factor of 10. The GPA is based on grades in all subjects on the student's study program according to a 7-point grading scale. A different 10-point scale was used in the Danish education system before 2007. For degrees obtained before 2007, we have converted the GPA to the 7-point scale equivalent using the official conversion table (https://www.optagelse.dk/vejledninger/pdf/Fra-13-skalaen-til-7.pdf). Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C14. Placebo test of significance of the two distance instruments in OLS choice models for completing compulsory school: p-value of F-test

Sample	(1) Full sample	(2) Enrollment	(3) Completion
Dependent variable (completing compulsory school):			
At least one non-missing 9th grade test score	0.84346	0.22523	0.13284
Non-missing test scores in math and Danish	0.91426	0.36963	0.19729
Passing test scores in math and Danish	0.39449	0.29984	0.25187
Passing test scores in math, Danish, English and science	0.17906	0.40590	0.26180
N	210,339	198,048	161,432

Note. This table shows p-values of F-tests for the two distance instruments in linear probability models for completion of compulsory school (9th grade). The p-values are shown for four different definitions of completion, and for three samples: The full sample; the sample restricted to those who had enrolled in upper secondary education by age 25; and the main sample restricted to those who had completed an upper secondary education by age 25.

Table C15. First stage logit models with inclusion of all, some or no controls for test scores in 9th grade. Decision on completing vocational instead of general upper secondary education. Average marginal derivatives.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Males				Females			
	All	Not Math or Science	Not Danish or English	No	All	Not Math or Science	Not Danish or English	No
Distance to vocational	-0.0004	-0.0004	-0.0004	-0.0010*	-0.0007*	-0.0006+	-0.0010**	-0.0011**
	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Distance to general	0.0042***	0.0041***	0.0048***	0.0054***	0.0019***	0.0017***	0.0022***	0.0020***
	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
N	79,070	79,070	79,070	79,070	82,362	82,362	82,362	82,362
Chi squared	256.8	236.7	303.6	241.8	49.7	40.0	55.0	29.2
p-value	1.745e-56	4.049e-52	1.210e-66	3.169e-53	1.644e-11	2.065e-09	1.161e-12	4.500e-07
Degrees of freedom	2	2	2	2	2	2	2	2

Note. The models in Columns 1 and 5 are the baseline first stage except that the distance variables are not interacted with math test scores. They are identical to the models in Columns 2 and 4 of Table 5. The models in Columns 2 and 6 exclude controls for test scores in math and science in 9th grade (and similar variables for teacher assessed marks for the year's work and related dummies for missing information). The models in Columns 3 and 7 exclude instead controls for test scores in Danish and English. The models in Columns 4 and 8 do not include any controls for 9th grade test scores or teacher assessed marks for the year's work. The dependent variable is an indicator for completing a vocational instead of a general upper secondary education. Average marginal derivatives based on the logit model estimates are shown for the two distance measures. For each individual, we calculate the effect of increasing the distance by one unit (1 km), while holding all control variables fixed, on the probability of completing vocational instead of general education. The Chi squared test statistic and its p-value are for a test for joint significance of the two distance instruments. Heteroskedasticity robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix D. Data availability

This study uses anonymized micro data from Statistics Denmark based on administrative registers. These data are available only at protected research servers at Statistics Denmark. It is not permitted to extract any micro data from Statistics Denmark. To gain access to micro data through Statistics Denmark, researchers must be affiliated with a Danish authorized research environment; authorization is undertaken by Statistics Denmark. For any research project, Statistics Denmark must approve a project description of the purpose of the project, the study population, the data needed, and the researchers involved. Statistics Denmark then makes the data available to the named researchers at a protected research server. Access is provided through the researcher's PC over the Internet. To replicate the analyses presented in this paper, then, a researcher must affiliate with an authorized Danish research environment and apply to Statistics Denmark for access to the relevant data. For more details, see https://www.dst.dk/en/TilSalg/Forskningsservice.

The authors are committed to providing guidance on obtaining access to the data, and to providing the Stata code which can be used for replication.

References for the Online Appendix

Athey, Susan, Raj Chetty, Guido W. Imbens, and Hyunseung Kang. 2019. "The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely." NBER Working Paper No. 26463.

Björklund, Anders, and Marianne Sundström. 2006. "Parental separation and children's educational attainment: A siblings analysis on Swedish register data." *Economica* 73: 605-24.

Brinch, Christian N., Magne Mogstad and Matthew Wiswall. 2017. "Beyond LATE with a discrete instrument." *Journal of Political Economy* 125: 985-1039.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. "Estimating marginal returns to education." *American Economic Review* 101(6): 2754-81.

Carneiro, Pedro, Michael Lokshin, and Nithin Umapathi. 2017. "Average and marginal returns to upper secondary schooling in Indonesia." *Journal of Applied Econometrics* 32: 16-36.

Heckman, James J. 2010. "Building bridges between structural and program evaluation approaches to evaluating policy." *Journal of Economic Literature* 48(2): 356-98.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and estimation of local average treatment effects." *Econometrica* 62: 467-75.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters. 2021. "The causal interpretation of two-stage least squares with multiple instrumental variables." *American Economic Review* 111(11): 3663-98.

Nybom, Martin. 2017. "The distribution of lifetime earnings returns to college." *Journal of Labor Economics* 35(4): 903-52.