



1

2 **Supporting Information for** 3 **Economic Preferences Predict Education Choices**

4 **Thomas Epper, Ernst Fehr, Kristoffer B. Hvidberg, Claus T. Kreiner, Søren Leth-Petersen, Basit Zafar**

5 **Corresponding Author name.**
6 **E-mail: ernst.fehr@econ.uzh.ch**

7 **This PDF file includes:**

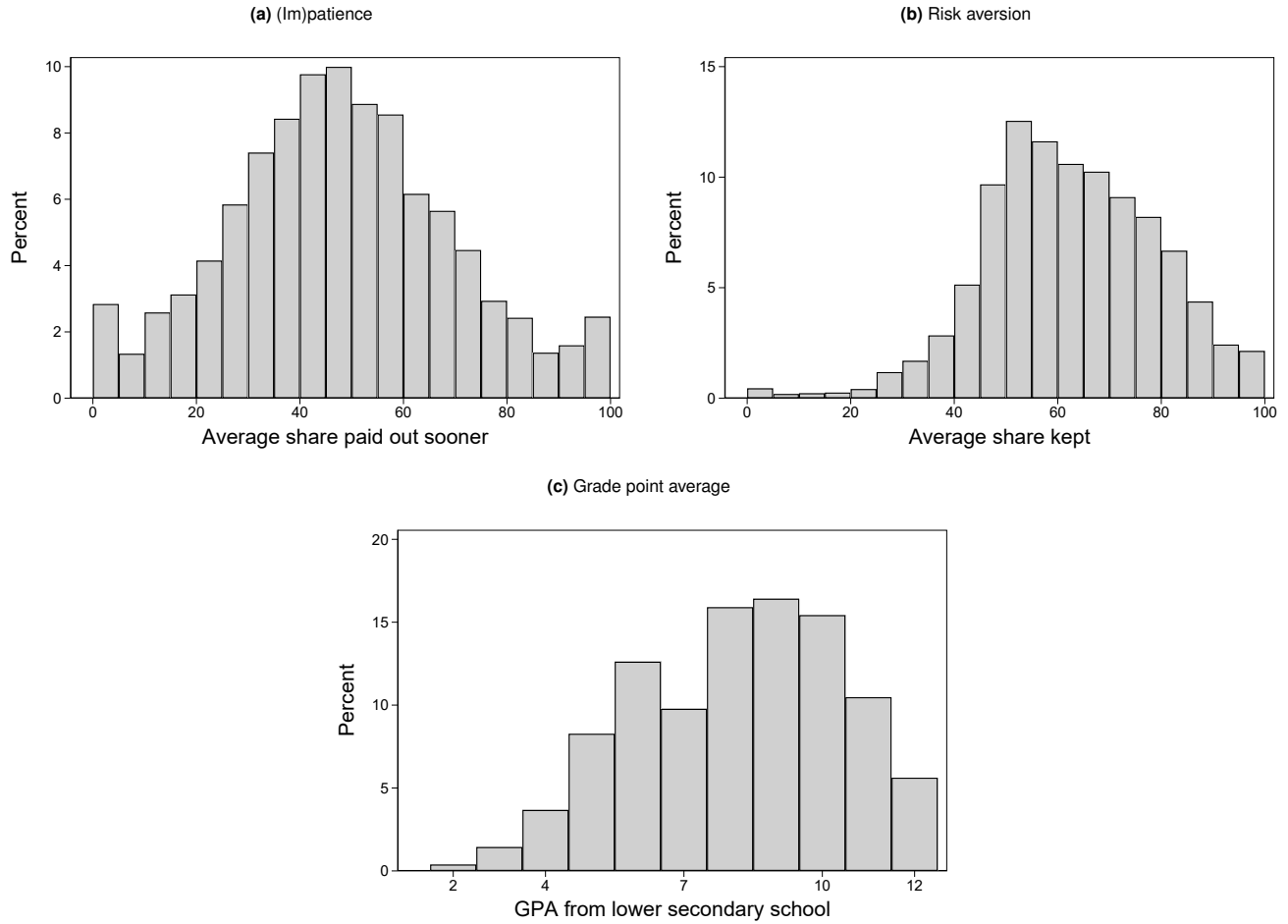
- 8 Figs. S1 to S3
- 9 Tables S1 to S8
- 10 SI References



Fig. S1. Screenshots of online experiment eliciting time and risk preferences

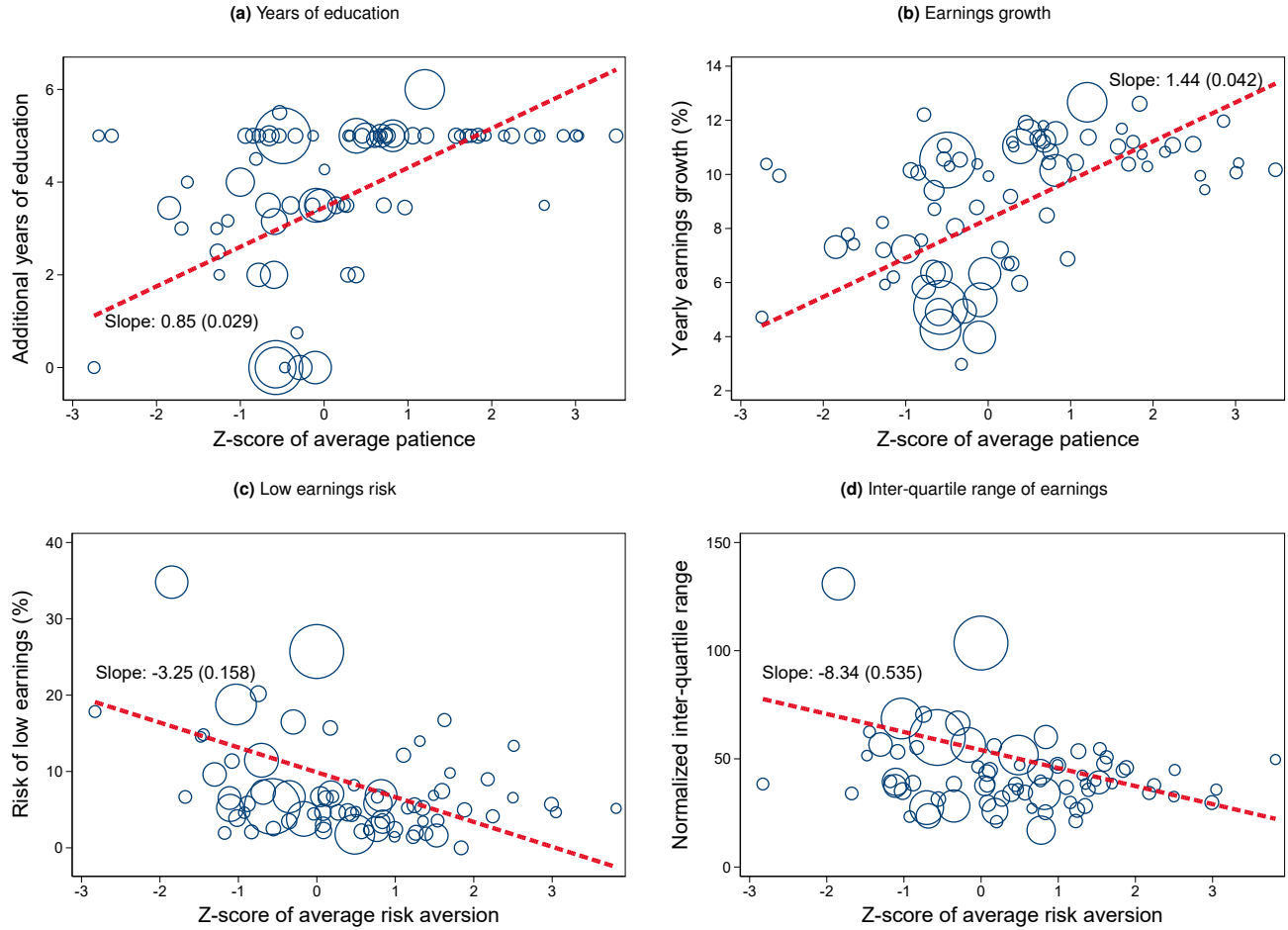
Notes: Panel (a) shows an example of a choice situation used to elicit time preference. The participant can choose how much of 250 DKK to receive in 8 weeks, and how much to save with interest to be paid out in 16 weeks. In the example, the participant keeps 150 and saves 100. The participant thus receives 150 DKK in 8 weeks and 102.4 DKK ($100 \times (1 + 0.024)$) in 16 weeks. Panel (b) shows an example of a choice situation used to elicit risk preference. The participant can choose how much of 250 DKK to keep in a safe account, and how much to invest in a risky lottery with a 40 percent probability of loss and a 60 percent probability of gain. In the example, the participant keeps 50 and invests 200. With 40% probability, the participant receives 130 DKK ($50 + 80$), and with 60% probability, the participant receives 350 DKK ($50 + 300$). See all choice situations in SI Tables [S1](#) and [S2](#).

Fig. S2. Distributions of explanatory variables



Notes: Panels (a) and (b) show the distributions of the average choices in the experiments for each participant across all choice situations. Panel (c) shows the distribution of the grade point average in Danish and Math from the final written exams in lower secondary school. $N = 3,132$.

Fig. S3. Z-scores of economic preferences and education and labor market outcomes



Notes: The figure is similar to Figure 1, but instead of using ranks within the distribution of preferences we use z-scores. We compute the program z-scores as follows. For the 3,132 participants, we use the average shares paid out and kept in the experiments and the GPA from lower secondary school. We then compute the applied-for program means of these three variables. Finally, we standardize the applied-for program means using the 3,132 observations to compute the mean and standard deviation. Each scatter represents an educational program. The size of the scatters reflects the number of participants applying for the program. On the *x*-axis we plot the program z-scores. On the *y*-axis we plot the program outcomes. For years of education, it is the average number of years studying full time required to graduate from the program. The labor market outcomes are based on the actual outcomes of graduates from the programs. The red line illustrates the ordinary least squares prediction from a regression of the outcome on the z-score at the program level, using frequency weights. In Panel (d), we do not show two outlier observations with normalized inter-quartile ranges of more than 180 and 250, respectively. We include the observations in all estimations. $N = 3,132$. Robust standard errors in parentheses.

Table S1. Choice situations in the time experiment

Situation	Low stakes		High stakes		t_1	t_2	Rate
	x_1	x_2	x_1	x_2			
1	250	251	7500	7530	0	8	0.024
2	250	256	7500	7680	0	8	0.153
3	250	261	7500	7830	0	8	0.295
4	250	266	7500	7980	0	8	0.451
5	250	271	7500	8130	0	8	0.622
6	250	276	7500	8280	0	8	0.811
7	250	281	7500	8430	0	8	1.016
8	250	286	7500	8580	0	8	1.242
9	250	251	7500	7530	8	16	0.024
10	250	256	7500	7680	8	16	0.153
11	250	261	7500	7830	8	16	0.295
12	250	266	7500	7980	8	16	0.451
13	250	271	7500	8130	8	16	0.622
14	250	276	7500	8280	8	16	0.811
15	250	281	7500	8430	8	16	1.016
16	250	286	7500	8580	8	16	1.242

Notes: x_1 is the amount the participant can get paid out sooner and x_2 is the amount the participant can get paid out later. They differ by whether the participant was assigned to the low or high stakes treatment. t_1 indicates the sooner payout time (either within 24 hours (0) or in 2 months) while t_2 indicates the later payout time (either in 8 weeks or in 16 weeks). The user interface displayed the delays in weeks to avoid confounds by payments at different weekdays. Rate is the annualized interest rate the participant gets on the amount saved for two months. For instance, in Situation 1, Rate=0.024 refers to a annual interest rate of 2.4%. Panel (a) of SI Figure S1 illustrates situation 10 with low stakes.

Table S2. Choice situations in the risk experiment

Situation	DKK	p	Good	Bad
1	250	0.5	1.21	0.81
2	250	0.2	1.41	0.91
3	250	0.8	1.11	0.61
4	250	0.5	1.31	0.71
5	250	0.2	1.61	0.86
6	250	0.8	1.16	0.41
7	250	0.5	1.35	0.75
8	250	0.2	1.65	0.90
9	250	0.8	1.20	0.45
10	250	0.6	1.50	0.40
11	250	0.4	1.72	0.62
12	250	0.6	1.45	0.35
13	250	0.4	1.67	0.57
14	250	0.5	1.51	0.50
15	250	0.5	1.61	0.60

Notes: *DKK* is the amount the participant can keep or invest in the lottery. p is the probability that the lottery will give the good state. *Good* is the rate of return of the investment in the good state while *Bad* is the rate of return of the investment in the bad state. Panel (b) of SI Figure S1 illustrates situation 10.

Table S3. Examples of education programs

Program	Obs.	Years	Earnings	Growth	Low	IQR
USS: General	293	0	12.8	5.1	25.7	103.5
USS: Preparatory	106	0	12.6	4	34.8	130.9
USS: Commercial	167	0	12.9	4.3	18.8	68.6
USS: Technical	59	0	13	4.9	16.5	66.5
AP in Financial management	74	2	13.1	4.9	5.2	39.4
AP in Marketing management	55	2	13	5.8	9.6	56.6
Physiotherapist	67	3	12.9	6.3	7.1	25.8
Professional BA in finance	52	3.5	13.2	7.3	6.5	37.5
Social educator	115	3.5	12.8	5.4	11.5	27.4
Social worker	60	3.5	12.9	6.4	7.3	23.6
Nurse	104	3.5	12.9	6.3	6.7	28.2
Teacher	80	4	13	7.2	5.8	17
Communication	40	5	13.1	9.4	6.7	37.6
Data science	53	5	13.4	11.2	1.6	39.1
Law	121	5	13.4	11	3.8	56.4
Political science	64	5	13.4	11.6	2.5	43.9
Psychology	101	5	13.1	10.1	6.9	33.9
Economics	54	5	13.5	11.5	3.4	60.1
Business economics	313	5	13.4	10.5	5.4	59.9
Medicin	157	6	13.6	12.7	1.8	51.8

Notes: The table shows the main outcomes for the 20 largest educational programs in terms of number of applicants among the participants in the experiments. USS is an abbreviation for Upper Secondary School. AP is an abbreviation for Academy Profession. *Obs.* is the number of participants we observe applying for the program among the sample of 3,132. *Years* is the number of years of full time study required to graduate from the program. *Earnings* is the log of median earnings in 2019 among graduates from the program. *Growth* is the annual growth rate of mean earnings from age 20 to age 42-47 for graduates from the program. *Low* is the probability that a graduate from the program has earnings in the bottom quintile of the earnings distribution in 2019. *IQR* is the inter-quartile range of earnings in 2019 among graduates from the program. The inter-quartile range is normalized by the median earnings and multiplied by 100.

Table S4. Pairwise correlations

	Pairwise correlations								
	1.	2.	3.	4.	5.	6.	7.	8.	9.
Explanatory variables									
1. Patience rank	1.00	-0.13	0.15	0.08	0.09	-0.05	-0.07	0.08	-0.07
2. Risk aversion rank	-0.13	1.00	0.13	0.10	0.11	-0.06	-0.08	0.07	-0.05
3. GPA rank	0.15	0.13	1.00	0.37	0.47	-0.14	-0.30	0.41	-0.21
Main outcomes									
4. Additional years	0.08	0.10	0.37	1.00	0.90	-0.64	-0.86	0.79	-0.57
5. Earnings growth	0.09	0.11	0.47	0.90	1.00	-0.37	-0.71	0.86	-0.52
6. Inter-quartile range	-0.05	-0.06	-0.14	-0.64	-0.37	1.00	0.84	-0.46	0.53
7. Low earnings risk	-0.07	-0.08	-0.30	-0.86	-0.71	0.84	1.00	-0.82	0.73
Additional outcomes									
8. Log median earnings	0.08	0.07	0.41	0.79	0.86	-0.46	-0.82	1.00	-0.66
9. Unemployment	-0.07	-0.05	-0.21	-0.57	-0.52	0.53	0.73	-0.66	1.00

Notes: The table shows the pairwise correlations at the individual level. *Patience rank* is the individual rank in the distribution of average share paid out later in the time experiment. *Risk aversion rank* is the individual rank in the distribution of the average share not invested in the risky lottery in the risk experiment. *GPA rank* is the individual rank in the distribution of grade point average in the final written exams in Danish and Math in lower secondary school. *Additional years* is the number of years of full time study required to graduate from the applied-for program. *Earnings growth* is the annual growth rate of mean earnings from age 20 to age 42-47 for graduates from the applied-for program. *Low* is the probability that a graduate from the program has earnings in the bottom quintile of the earnings distribution in 2019. *Inter-quartile range* is the inter-quartile range of earnings in 2019 among graduates from the applied-for program. The inter-quartile range is normalized by the median earnings. *Log median earnings* is the log of median earnings in 2019 among graduates from the applied-for program. *Unemployment* is the probability that a graduate from the applied-for program was unemployed for more than 3 months in a calendar year between 2015 and 2019. $N = 3,132$.

Table S5. Preferences, GPA, and background characteristics

	Patience rank		Risk aversion rank		GPA rank	
	(1) Bivar.	(2) Multi.	(3) Bivar.	(4) Multi.	(5) Bivar.	(6) Multi.
Male (=1)	-8.62 (1.01)	-8.90 (1.03)	-0.67 (1.02)	-0.84 (1.04)	-3.32 (1.01)	-4.44 (0.97)
Immigrant/descendant (=1)	-6.82 (1.34)	-4.76 (1.89)	-3.03 (1.36)	-1.92 (1.96)	-16.64 (1.31)	-9.77 (1.77)
Father edu.: High (=1)	3.29 (1.19)	1.57 (1.58)	2.26 (1.19)	-0.08 (1.61)	19.45 (1.13)	11.18 (1.45)
Mother edu.: High (=1)	3.49 (1.21)	1.43 (1.66)	5.70 (1.21)	4.93 (1.67)	19.39 (1.18)	10.22 (1.54)
Income rank of father	0.04 (0.14)	0.02 (0.02)	0.02 (0.14)	-0.00 (0.02)	0.22 (0.13)	0.11 (0.02)
Income rank of mother	0.06 (0.14)	0.04 (0.02)	0.05 (0.14)	0.03 (0.02)	0.21 (0.13)	0.09 (0.02)
Observations		3132		3132		3132
R^2		0.031		0.006		0.124
Mean outcome		50.5		50.5		50.5

Notes: The table shows coefficients from ordinary least squares regressions of preferences and grade point average on background characteristics. Regressions are at the individual level. Columns 1, 3, and 5 are bivariate regressions, i.e., each outcome is regressed on one explanatory variable at a time. Remaining columns show results from multivariable regressions. *Patience rank* is the individual rank in the distribution of average share paid out later in the time experiment. *Risk aversion rank* is the individual rank in the distribution of the average share not invested in the risky lottery in the risk experiment. *GPA rank* is the individual rank in the distribution of grade point average in the final written exams in Danish and Math in lower secondary school. *Male* and *Immigrant/descendant* are indicators for gender and immigrant status. *Father/mother edu.: High* are indicators for parents having obtained Master's or PhD degrees. The parental income ranks are the percentile ranks in the distribution of total income within all fathers/mothers to individuals born in 1999. Robust standard errors in parentheses.

Table S6. Individual level regressions

	Additional years of education			Earnings growth (%)		
	(1) Bivar.	(2) Multi.	(3) Multi.	(4) Bivar.	(5) Multi.	(6) Multi.
Patience rank	0.0055 (0.0012)	0.0023 (0.0012)	0.0017 (0.0012)	0.0093 (0.0017)	0.0034 (0.0016)	0.0039 (0.0016)
Risk aversion rank	0.0067 (0.0013)	0.0037 (0.0012)	0.0036 (0.0012)	0.011 (0.0017)	0.0057 (0.0016)	0.0058 (0.0015)
GPA rank	0.026 (0.0011)	0.025 (0.0011)	0.024 (0.0012)	0.046 (0.0014)	0.044 (0.0014)	0.042 (0.0015)
R^2		0.144	0.179		0.222	0.259
Mean outcome	3.46	3.46	3.46	8.35	8.35	8.35
	Low earnings (%)			Inter-quartile range (%)		
	(7) Bivar.	(8) Multi.	(9) Multi.	(10) Bivar.	(11) Multi.	(12) Multi.
Patience rank	-0.021 (0.0057)	-0.0090 (0.0056)	-0.0096 (0.0056)	-0.049 (0.020)	-0.034 (0.020)	-0.022 (0.020)
Risk aversion rank	-0.024 (0.0057)	-0.014 (0.0055)	-0.014 (0.0054)	-0.063 (0.019)	-0.049 (0.019)	-0.048 (0.018)
GPA rank	-0.094 (0.0057)	-0.091 (0.0058)	-0.090 (0.0061)	-0.15 (0.021)	-0.14 (0.021)	-0.14 (0.022)
R^2		0.091	0.115		0.023	0.037
Mean outcome	9.90	9.90	9.90	54.1	54.1	54.1
Observations	3132	3132	3132	3132	3132	3132
Controls			✓			✓

Notes: The table is similar to Table 3, but instead of using frequency weighted groups as the level of observation, we use individual observations. The table shows coefficients from ordinary least squares regressions of outcomes on preferences and grade point average. Columns 1, 4, 7, and 10 are bivariate regressions, i.e., each outcome is regressed on one explanatory variable at a time. Remaining columns show results from multivariable regressions. *Controls* include program-level shares of males, immigrants and descendants of immigrants, and students with no highly educated parents, as well as the program-average income rank of fathers and mothers. Robust standard errors in parentheses.

Table S7. Alternative measures of preferences

	Years of education		Earnings growth (%)		Low earnings (%)		Inter-quartile range (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Bivar.	Multi.	Bivar.	Multi.	Bivar.	Multi.	Bivar.	Multi.
Panel A: Z-scores								
Average patience (z-score)	0.85 (0.029)	0.093 (0.027)	1.44 (0.042)	0.16 (0.034)	-3.18 (0.15)	-1.10 (0.17)	-7.20 (0.70)	-5.73 (0.84)
Average risk aversion (z-score)	0.94 (0.026)	0.24 (0.025)	1.55 (0.036)	0.23 (0.028)	-3.25 (0.16)	-1.32 (0.19)	-8.34 (0.54)	-6.32 (0.91)
Average GPA (z-score)	1.43 (0.023)	1.14 (0.044)	2.49 (0.023)	2.08 (0.049)	-5.33 (0.17)	-5.40 (0.30)	-8.78 (0.80)	-13.3 (1.67)
R^2		0.670		0.845		0.513		0.272
Panel B: Structural parameters								
Average patience rank	0.12 (0.0052)	0.022 (0.0049)	0.17 (0.0071)	0.016 (0.0063)	-0.44 (0.036)	-0.22 (0.035)	-1.58 (0.16)	-1.27 (0.18)
Average risk aversion rank	0.16 (0.0047)	0.024 (0.0043)	0.27 (0.0069)	0.026 (0.0053)	-0.53 (0.024)	-0.15 (0.028)	-1.45 (0.074)	-1.16 (0.11)
Average GPA rank	0.095 (0.0015)	0.081 (0.0033)	0.17 (0.0014)	0.14 (0.0036)	-0.34 (0.011)	-0.36 (0.019)	-0.55 (0.050)	-0.81 (0.090)
R^2		0.666		0.845		0.495		0.275
Observations	3132	3132	3132	3132	3132	3132	3132	3132
# programs	76	76	76	76	76	76	76	76
Controls		✓		✓		✓		✓
Mean outcome	3.46	3.46	8.35	8.35	9.90	9.90	54.1	54.1

Notes: In Panel A, we use z-scores of the choices in the experiments and grade point average. We compute the program z-scores as follows. For the 3,132 participants, we use the average shares paid out and kept in the experiments and the GPA from lower secondary school. We then compute the applied-for program means of these three variables. Finally, we standardize the applied-for program means using the 3,132 observations to compute the mean and standard deviation. In Panel B, we use structurally estimated preference parameters and then compute the ranks based on these instead of the choices directly. For more details on the structural estimation, see the supplementary information for Epper et al. (1). The table shows coefficients from ordinary least squares regressions of outcomes on preferences and grade point average. Regressions are at the program level, weighted by the number of applicants. Columns 1, 3, 5, and 7 are bivariate regressions, i.e., each outcome is regressed on one explanatory variable at a time. Remaining columns show results from multivariable regressions. *Controls* include program-level shares of males, immigrants and descendants of immigrants, and students with no highly educated parents, as well as the program-average income rank of fathers and mothers. Robust standard errors in parentheses.

Table S8. Earnings and unemployment

	Log of median earnings		Unemployment > 3 m. (%)	
	(1) 2019	(2) 2015-19	(3) 2019	(4) 2015-19
Average patience rank	0.0055 (0.00071)	0.0033 (0.00053)	-0.064 (0.0077)	-0.22 (0.020)
Average risk aversion rank	-0.0014 (0.00073)	-0.0027 (0.00059)	-0.027 (0.0077)	-0.10 (0.023)
Average GPA rank	0.014 (0.00048)	0.012 (0.00037)	-0.035 (0.0040)	-0.11 (0.011)
R^2	0.806	0.840	0.303	0.347
Mean outcome	13.1	13.2	2.38	7.99
Observations	3132	3132	3132	3132
# programs	76	76	76	76
Controls	✓	✓	✓	✓

Notes: The table shows coefficients from ordinary least squares regressions of outcomes on preferences and grade point average. For *Log of median earnings*, we first compute the median earnings for graduates from each program—either using earnings in 2019 or average earnings from 2015 to 2019. We then compute the log of the median earnings. For *Unemployment > 3 m.*, we use an indicator for being unemployed for more than 3 months—either in 2019 or in at least one year from 2015 to 2019. *Controls* include program-level shares of males, immigrants and descendants of immigrants, and students with no highly educated parents, as well as the program-average income rank of fathers and mothers. Robust standard errors in parentheses.

11 **References**

- 12 1. T Epper, et al., Preferences predict who commits crime among young men. *Proc. Natl. Acad. Sci.* **119**, e2112645119 (2022).