

Supplement to:

Lahtinen, Hannu, Pekka Martikainen, Kaarina Korhonen, Tim Morris, and Mikko Myrskylä. 2024. "Educational tracking and the polygenic prediction of education." *Sociological Science* 11: 186-213.

Online supplement for the article: Lahtinen Hannu, Pekka Martikainen, Kaarina Korhonen, Tim Morris and Mikko Myrskylä. 2024. "Educational tracking and the polygenic prediction of education." *Sociological Science*.

Table A1. The association between one standard deviation change in education PGS and achieved years of education before and after the comprehensive school reform by reform region

Region	Pre-reform		Post-reform		Difference-in-difference		N (pre; post-reform)
	b	95% CI	b	95% CI	b	95% CI	
1972	0.66	(0.39; 0.92)	0.65	(0.41; 0.89)	-0.01	(-0.36; 0.35)	654 (335; 319)
1973	0.40	(0.22; 0.58)	0.64	(0.45; 0.82)	0.23	(-0.02; 0.49)	1,025 (556; 469)
1974	0.55	(0.41; 0.69)	0.58	(0.43; 0.72)	0.02	(-0.18; 0.22)	1,587 (801; 786)
1975	0.53	(0.34; 0.72)	0.67	(0.46; 0.87)	0.13	(-0.14; 0.41)	1,006 (508; 498)
1976	0.81	(0.61; 1.01)	0.66	(0.47; 0.84)	-0.15	(-0.42; 0.11)	1,065 (536; 529)
1977	0.82	(0.55; 1.10)	1.20	(0.95; 1.45)	0.38	(0.00; 0.75)	609 (296; 313)
Meta-analytic	0.59	(0.51; 0.67)	0.68	(0.61; 0.76)	0.08	(-0.03; 0.19)	5946 (3032; 2914)

Notes: Estimates are based on reform region (1972–1977) -specific linear regression models predicting years of education. Independent variables include reform indicator, education PGS, and their interaction; adjusted for gender, year of birth, the first ten principal components of the genome, study collection round and genotyping batch. Meta-analytic estimate is an inverse-variance weighted fixed-effect estimate on these region-specific estimates.

Table A2. The association between one standard deviation change in education PGS and achieved years of education before and after the comprehensive school reform by reform region, heterogeneous effects by gender and family education

<b>Men</b>							
Region	Pre-reform		Post-reform		Difference-in-difference		N (pre; post-reform)
	b	95% CI	b	95% CI	b	95% CI	
1972	0.82	(0.39; 1.26)	0.23	(-0.13; 0.59)	-0.60	(-1.14; -0.05)	269 (148; 121)
1973	0.26	(0.02; 0.50)	0.77	(0.47; 1.06)	0.50	(0.13; 0.88)	469 (247; 222)
1974	0.41	(0.21; 0.62)	0.52	(0.32; 0.73)	0.11	(-0.17; 0.40)	720 (363; 357)
1975	0.47	(0.17; 0.76)	0.81	(0.49; 1.14)	0.35	(-0.08; 0.77)	456 (237; 219)
1976	0.76	(0.44; 1.08)	0.70	(0.43; 0.97)	-0.06	(-0.48; 0.35)	477 (238; 239)
1977	0.82	(0.39; 1.26)	1.44	(1.06; 1.82)	0.62	(0.03; 1.20)	289 (150; 139)
Meta-analytic	0.49	(0.37; 0.61)	0.69	(0.57; 0.80)	0.17	(0.01; 0.34)	2680 (1383; 1297)
<b>Women</b>							
Region	Pre-reform		Post-reform		Difference-in-difference		N (pre; post-reform)
	b	95% CI	b	95% CI	b	95% CI	
1972	0.52	(0.17; 0.87)	0.91	(0.63; 1.19)	0.39	(-0.07; 0.84)	385 (187; 198)
1973	0.52	(0.27; 0.78)	0.54	(0.30; 0.78)	0.01	(-0.33; 0.36)	556 (309; 247)
1974	0.68	(0.49; 0.87)	0.62	(0.41; 0.82)	-0.06	(-0.34; 0.22)	867 (438; 429)
1975	0.59	(0.34; 0.84)	0.54	(0.27; 0.81)	-0.05	(-0.40; 0.31)	550 (271; 279)
1976	0.82	(0.57; 1.07)	0.65	(0.40; 0.90)	-0.17	(-0.52; 0.18)	588 (298; 290)
1977	0.78	(0.41; 1.14)	1.01	(0.68; 1.35)	0.24	(-0.27; 0.75)	320 (146; 174)
Meta-analytic	0.66	(0.55; 0.76)	0.67	(0.57; 0.78)	0.01	(-0.14; 0.15)	3266 (1649; 1617)
<b>Basic family education</b>							
Region	Pre-reform		Post-reform		Difference-in-difference		N (pre; post-reform)
	b	95% CI	b	95% CI	b	95% CI	
1972	0.49	(0.15; 0.83)	0.42	(0.11; 0.74)	-0.07	(-0.54; 0.40)	401 (208; 193)
1973	0.37	(0.15; 0.58)	0.68	(0.44; 0.91)	0.31	(-0.01; 0.63)	670 (370; 300)
1974	0.41	(0.24; 0.58)	0.52	(0.33; 0.70)	0.11	(-0.14; 0.35)	928 (492; 436)
1975	0.43	(0.17; 0.69)	0.63	(0.32; 0.95)	0.20	(-0.20; 0.60)	480 (255; 225)
1976	0.61	(0.36; 0.86)	0.77	(0.47; 1.06)	0.16	(-0.23; 0.55)	494 (275; 219)
1977	0.45	(0.05; 0.86)	0.73	(0.30; 1.16)	0.28	(-0.29; 0.84)	263 (144; 119)
Meta-analytic	0.44	(0.35; 0.54)	0.60	(0.49; 0.71)	0.16	(0.02; 0.31)	3236 (1744; 1492)
<b>More than basic family education</b>							
Region	Pre-reform		Post-reform		Difference-in-difference		N (pre; post-reform)
	b	95% CI	b	95% CI	b	95% CI	
1972	0.97	(0.50; 1.43)	0.81	(0.42; 1.20)	-0.16	(-0.76; 0.45)	253 (127; 126)
1973	0.48	(0.19; 0.78)	0.48	(0.15; 0.81)	-0.01	(-0.44; 0.43)	355 (186; 169)
1974	0.62	(0.38; 0.87)	0.57	(0.35; 0.79)	-0.06	(-0.38; 0.27)	659 (309; 350)
1975	0.43	(0.13; 0.72)	0.58	(0.32; 0.83)	0.15	(-0.23; 0.53)	526 (253; 273)
1976	0.85	(0.55; 1.16)	0.50	(0.27; 0.73)	-0.35	(-0.73; 0.02)	571 (261; 310)
1977	0.97	(0.60; 1.34)	1.21	(0.90; 1.52)	0.24	(-0.26; 0.74)	346 (152; 194)
Meta-analytic	0.67	(0.54; 0.79)	0.64	(0.53; 0.76)	-0.04	(-0.21; 0.13)	2710 (1288; 1422)

Notes: Estimates are based on reform region (1972–1977) -specific linear regression models predicting years of education. Independent variables include reform indicator, education PGS, and their interaction; adjusted for gender (unless stratified), year of birth, first ten principal components of the genome, study collection round and genotyping batch. Meta-analytic estimate is an inverse-variance weighted fixed-effect estimate on these region-specific estimates.

Table A3 The association between one standard deviation change in education PGS and achieved years of education before, during and after the comprehensive school reform by reform region

Region	Pre-reform cohorts		1 <sup>st</sup> Reform cohort		Subsequent cohorts	
	b	95% CI	b	95% CI	b	95% CI
1972	0.66	(0.39; 0.92)	0.64	(0.16; 1.13)	0.66	(0.38; 0.94)
1973	0.40	(0.22; 0.58)	0.81	(0.44; 1.18)	0.59	(0.37; 0.81)
1974	0.55	(0.41; 0.69)	0.74	(0.46; 1.01)	0.53	(0.36; 0.69)
1975	0.53	(0.34; 0.72)	0.85	(0.46; 1.24)	0.60	(0.37; 0.84)
1976	0.81	(0.61; 1.01)	0.47	(0.11; 0.82)	0.71	(0.50; 0.93)
1977	0.83	(0.55; 1.10)	1.27	(0.89; 1.64)	1.17	(0.85; 1.49)
Meta-analytic	0.59	(0.51; 0.67)	0.79	(0.64; 0.94)	0.65	(0.56; 0.74)
Difference-in-difference						
	1 <sup>st</sup> Reform cohort vs. pre-reform		Subsequent cohorts vs. pre-reform			
Region	b	95% CI	b	95% CI		
1972	-0.01	(-0.56; 0.53)	0.00	(-0.39; 0.39)		
1973	0.41	(0.00; 0.82)	0.19	(-0.10; 0.47)		
1974	0.19	(-0.12; 0.49)	-0.03	(-0.25; 0.19)		
1975	0.32	(-0.12; 0.75)	0.07	(-0.22; 0.37)		
1976	-0.34	(-0.74; 0.06)	-0.09	(-0.38; 0.19)		
1977	0.44	(-0.03; 0.91)	0.35	(-0.08; 0.77)		
Meta-analytic	0.17	(0.00; 0.33)	0.05	(-0.07; 0.17)		

Notes: Estimates are based on reform region (1972–1977) -specific linear regression models including reform status in three categories, education PGS, and their interaction; adjusted for gender, year of birth, first ten principal components of the genome, study collection round and genotyping batch. Meta-analytic estimate is an inverse-variance weighted fixed-effect estimate on these region-specific estimates.

Table A4. The association between one standard deviation change in education PGS and years of education before, during and after the comprehensive school reform by reform region, heterogeneous effect by gender

<b>Men</b>						
Region	Pre-reform cohorts		1 <sup>st</sup> Reform cohort		Subsequent cohorts	
	b	95% CI	b	95% CI	b	95% CI
1972	0.82	(0.38; 1.26)	0.23	(-0.46; 0.92)	0.23	(-0.24; 0.69)
1973	0.26	(0.02; 0.50)	0.97	(0.46; 1.48)	0.70	(0.33; 1.07)
1974	0.41	(0.20; 0.62)	0.83	(0.34; 1.31)	0.44	(0.22; 0.66)
1975	0.46	(0.17; 0.76)	1.07	(0.36; 1.77)	0.73	(0.39; 1.07)
1976	0.76	(0.44; 1.08)	0.92	(0.22; 1.62)	0.64	(0.35; 0.93)
1977	0.83	(0.39; 1.27)	0.95	(0.30; 1.60)	1.67	(1.19; 2.16)
Meta-analytic	0.49	(0.37; 0.61)	0.84	(0.60; 1.09)	0.63	(0.50; 0.76)
Difference-in-difference						
Region	Reform cohort vs. pre-reform		Subsequent cohorts vs. pre-reform			
	b	95% CI	b	95% CI		
1972	-0.59	(-1.38; 0.20)	-0.60	(-1.22; 0.03)		
1973	0.71	(0.16; 1.27)	0.44	(-0.01; 0.89)		
1974	0.42	(-0.12; 0.95)	0.03	(-0.27; 0.32)		
1975	0.60	(-0.14; 1.35)	0.26	(-0.18; 0.70)		
1976	0.16	(-0.61; 0.93)	-0.12	(-0.54; 0.31)		
1977	0.12	(-0.66; 0.90)	0.84	(0.18; 1.51)		
Meta-analytic	0.33	(0.05; 0.60)	0.11	(-0.06; 0.29)		
<b>Women</b>						
Region	Pre-reform		1 <sup>st</sup> Reform cohort		Subsequent cohort	
	b	95% CI	b	95% CI	b	95% CI
1972	0.52	(0.17; 0.87)	0.88	(0.26; 1.49)	0.90	(0.59; 1.22)
1973	0.52	(0.27; 0.78)	0.68	(0.16; 1.20)	0.50	(0.23; 0.78)
1974	0.68	(0.49; 0.87)	0.70	(0.38; 1.01)	0.58	(0.33; 0.83)
1975	0.59	(0.34; 0.84)	0.70	(0.20; 1.20)	0.49	(0.18; 0.80)
1976	0.81	(0.56; 1.06)	0.26	(-0.13; 0.64)	0.79	(0.49; 1.09)
1977	0.78	(0.41; 1.14)	1.54	(1.02; 2.05)	0.76	(0.31; 1.21)
Meta-analytic	0.66	(0.55; 0.76)	0.72	(0.54; 0.90)	0.65	(0.52; 0.77)
Difference-in-difference						
Region	Reform cohort vs. pre-reform		Subsequent cohorts vs. pre-reform			
	b	95% CI	b	95% CI		
1972	0.36	(-0.35; 1.07)	0.39	(-0.09; 0.87)		
1973	0.15	(-0.42; 0.72)	-0.02	(-0.39; 0.35)		
1974	0.02	(-0.35; 0.38)	-0.10	(-0.42; 0.22)		
1975	0.11	(-0.45; 0.67)	-0.10	(-0.48; 0.29)		
1976	-0.55	(-1.01; -0.10)	-0.02	(-0.41; 0.37)		
1977	0.76	(0.11; 1.40)	-0.02	(-0.60; 0.56)		
Meta-analytic	0.04	(-0.17; 0.25)	-0.01	(-0.17; 0.15)		

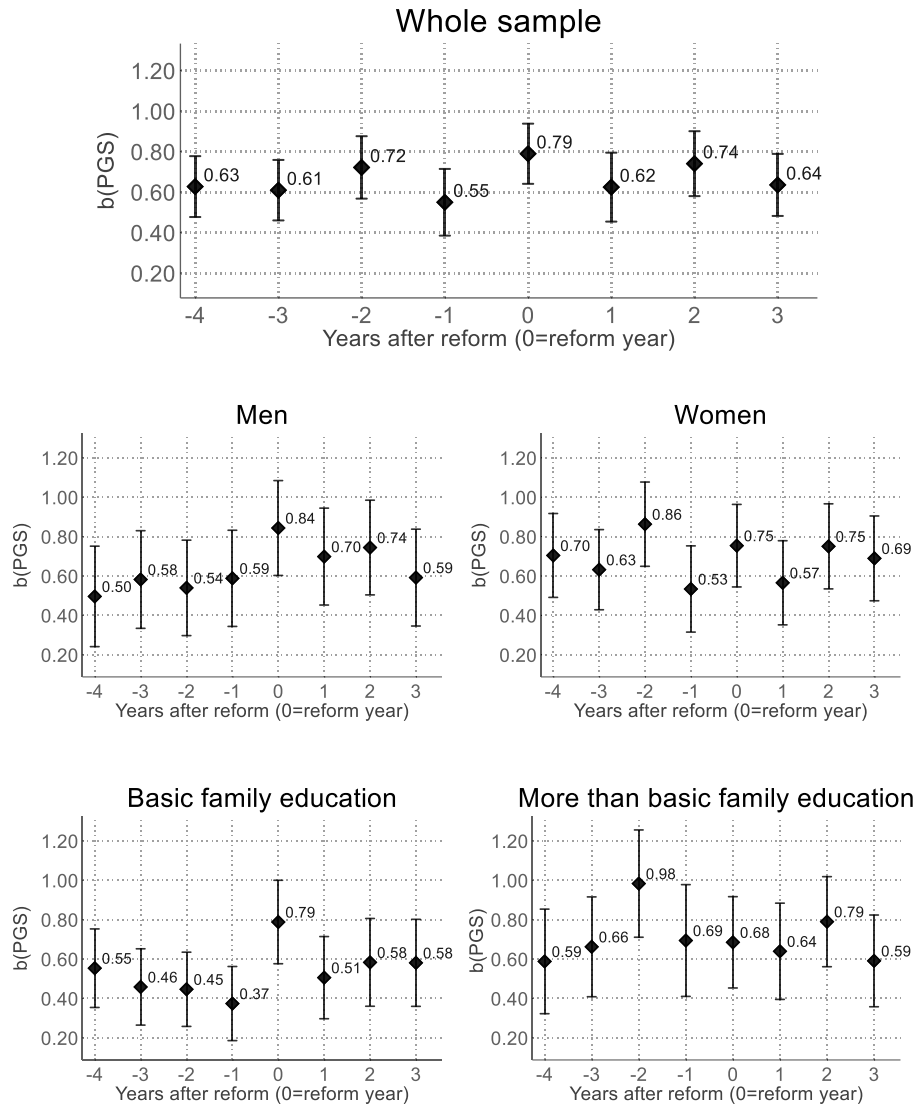
Notes: Estimates are based on reform region (1972–1977) -specific linear regression models including reform status in three categories, education PGS, and their interaction; adjusted for year of birth, first ten principal components of the genome, study collection round and genotyping batch. Meta-analytic estimate is an inverse-variance weighted fixed-effect estimate on these region-specific estimates.

Table A5 The association between one standard deviation change in education PGS and years of education before, during and after the comprehensive school reform by reform region, heterogeneous effect by family education

<b>Those with basic family education</b>							
Region	Pre-reform			1 <sup>st</sup> Reform cohort		Subsequent cohort	
	b	95% CI		b	95% CI	b	95% CI
1972	0.50	(0.16; 0.84)		0.61	(-0.03; 1.25)	0.35	(0.00; 0.7)
1973	0.37	(0.15; 0.59)		1.06	(0.56; 1.56)	0.57	(0.3; 0.84)
1974	0.41	(0.24; 0.58)		0.62	(0.23; 1.00)	0.48	(0.27; 0.69)
1975	0.43	(0.17; 0.69)		0.66	(0.00; 1.31)	0.61	(0.26; 0.96)
1976	0.60	(0.35; 0.85)		0.97	(0.41; 1.54)	0.74	(0.40; 1.07)
1977	0.47	(0.06; 0.88)		0.89	(0.16; 1.62)	0.57	(0.02; 1.11)
Meta-analytic	0.45	(0.35; 0.54)		0.79	(0.57; 1.01)	0.54	(0.42; 0.66)
Difference-in-difference							
Region	1 <sup>st</sup> Reform cohort vs. pre-reform			Subsequent vs. pre-reform			
	b	95% CI		b	95% CI		
1972	0.12	(-0.60; 0.83)		-0.15	(-0.65; 0.35)		
1973	0.69	(0.15; 1.23)		0.20	(-0.15; 0.55)		
1974	0.20	(-0.22; 0.62)		0.07	(-0.20; 0.34)		
1975	0.23	(-0.47; 0.92)		0.18	(-0.24; 0.6)		
1976	0.37	(-0.25; 0.99)		0.13	(-0.29; 0.55)		
1977	0.42	(-0.40; 1.24)		0.10	(-0.57; 0.76)		
Meta-analytic	0.34	(0.10; 0.58)		0.10	(-0.06; 0.26)		
<b>Those with more than basic family education</b>							
Region	Pre-reform			Reform cohort		Subsequent cohort	
	b	95% CI		b	95% CI	b	95% CI
1972	0.96	(0.50; 1.42)		0.57	(-0.2; 1.34)	0.95	(0.52; 1.39)
1973	0.48	(0.18; 0.78)		0.29	(-0.35; 0.93)	0.54	(0.14; 0.94)
1974	0.62	(0.38; 0.87)		0.80	(0.42; 1.19)	0.50	(0.24; 0.76)
1975	0.43	(0.13; 0.73)		0.65	(0.17; 1.12)	0.54	(0.26; 0.83)
1976	0.85	(0.55; 1.16)		0.22	(-0.18; 0.62)	0.60	(0.33; 0.87)
1977	0.96	(0.59; 1.34)		1.49	(1.02; 1.96)	1.07	(0.66; 1.49)
Meta-analytic	0.67	(0.54; 0.79)		0.70	(0.50; 0.89)	0.63	(0.50; 0.77)
Difference-in-difference							
Region	1 <sup>st</sup> Reform cohort vs. pre-reform			Subsequent vs. pre-reform			
	b	95% CI		b	95% CI		
1972	-0.39	(-1.29; 0.51)		-0.01	(-0.65; 0.64)		
1973	-0.20	(-0.89; 0.50)		0.06	(-0.43; 0.55)		
1974	0.18	(-0.27; 0.63)		-0.12	(-0.47; 0.23)		
1975	0.22	(-0.33; 0.77)		0.12	(-0.29; 0.52)		
1976	-0.63	(-1.14; -0.13)		-0.26	(-0.66; 0.15)		
1977	0.53	(-0.10; 1.16)		0.11	(-0.45; 0.67)		
Meta-analytic	-0.03	(-0.26; 0.21)		-0.04	(-0.22; 0.14)		

Notes: Estimates are based on reform region (1972–1977) -specific linear regression models including reform status in three categories, education PGS, and their interaction; adjusted for gender, year of birth, first ten principal components of the genome, study collection round and genotyping batch. Meta-analytic estimate is an inverse variance-weighted fixed-effect estimate on these region-specific estimates.

Figure A6. The association between one standard deviation change in education PGS and years of education by the years to reform



Notes: Sub-figures are based on linear regression including categorical years-after-reform variable, education PGS and the interaction between them. Capped bars are 95% confidence intervals. Models adjusted by gender (if not stratified), year of birth, first ten principal components of the genome, study collection round and genotyping batch.