



King's Research Portal

DOI:

[10.1038/s41562-018-0332-5](https://doi.org/10.1038/s41562-018-0332-5)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Rimfeld, K., Krapohl, E., Trzaskowski, M., Coleman, J. R. I., Selzam, S., Dale, P. S., Esko, T., Metspalu, A., & Plomin, R. (2018). Genetic influence on social outcomes during and after the Soviet era in Estonia. *Nature Human Behaviour*, 269–275. <https://doi.org/10.1038/s41562-018-0332-5>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

1 **Genetic influence on social outcomes during and after the Soviet era in Estonia**

2
3 Kaili Rimfeld^{1*}, Eva Krapohl¹, Maciej Trzaskowski², Jonathan R.I. Coleman^{1,3}, Saskia Selzam¹, Philip S.
4 Dale⁴, Tonu Esko⁵, Andres Metspalu⁵ & Robert Plomin¹

5
6 ¹ MRC Social, Genetic and Developmental Psychiatry Centre
7 Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK

8
9 ² Queensland Brain Institute, The University of Queensland, Brisbane, Australia

10
11 ³ National Institute for Health Research Biomedical Research Centre, South London and Maudsley National
12 Health Service Trust, UK

13
14 ⁴ University of New Mexico, Department of Speech and Hearing Sciences, Albuquerque, NM, 87131, USA

15
16 ⁵ Estonian Genome Centre, University of Tartu, Estonia

17
18
19
20
21 * Corresponding author: Kaili Rimfeld, MRC Social, Genetic and Developmental Psychiatry Centre
22 Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK; email:
23 Kaili.rimfeld@kcl.ac.uk

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 The etiology of individual differences in educational attainment and occupational status includes genetic as
55 well as environmental factors¹⁻⁵ and can change as societies change^{3,6,7}. The extent of genetic influence on
56 these social outcomes can be viewed as an index of success in achieving meritocratic values of equality of
57 opportunity by rewarding talent and hard work, which are to a large extent influenced by genetic factors,
58 rather than rewarding environmentally driven privilege. To the extent that the end of the Soviet Union and
59 the independence of Estonia led to an increase in meritocratic selection of individuals in education and
60 occupation, genetic influence should be higher in the post-Soviet era than in the Soviet era. Here we
61 confirmed this hypothesis: DNA differences (single-nucleotide polymorphisms, SNPs) explained twice as
62 much variance in educational attainment and occupational status in the post-Soviet era compared to the
63 Soviet era in both polygenic score analyses and SNP heritability analyses of 12 500 Estonians. This is the
64 first demonstration of a change in the extent of genetic influence in the same population following a massive
65 and abrupt social change – in this case, the shift from a communist to a capitalist society.

66

67

68

69

70

71

72

73

74

75

76

77

78

79 Socioeconomic status (SES), a composite index of educational attainment and occupational status, has been
80 shown to be associated with a range of life outcomes from life satisfaction and happiness, to physical and
81 mental health, and even life expectancy⁸⁻¹². Individual variation in SES in a population has often been
82 assumed to be explained entirely by environmental factors. Twin and adoption studies, however, suggest
83 that individual differences in SES are substantially genetic in origin¹⁻⁵, with heritability estimates from twin
84 studies of about 50%, meaning that around half of the individual differences in SES can be explained by
85 inherited differences in individual's DNA sequence. It is now possible to estimate heritability directly from
86 DNA using hundreds of thousands of DNA differences (single nucleotide polymorphisms, SNPs) genotyped
87 on microarrays (SNP chips) in samples of thousands of unrelated individuals¹³. Data of this sort are
88 available for many traits, including SES, as a by-product of genome-wide association (GWA) studies.
89 Unlike GWA analysis, which aims to identify specific SNPs associated with a trait, SNP heritability relates
90 overall similarity between individuals across all SNPs on a SNP chip to the individuals' phenotypic
91 similarity on a trait, without knowing which SNPs are associated with the trait.

92
93 SNP heritabilities have been estimated as about 20% for educational attainment, occupational status, and
94 combined SES^{4,14-18}. SNP heritability (20%) is less than heritability estimates from twin studies (50%)
95 because SNP heritability, like GWA analysis, is limited to the additive effects of common SNPs included on
96 SNP chips. For this reason, SNP heritability is the ceiling for GWA studies.

97
98 GWA data can also be used to create genome-wide polygenic scores (GPS) that aggregate thousands of SNP
99 associations across the genome to predict the trait of interest. Individual SNP associations typically account
100 for less than 0.1% of the variance, so are not individually useful for prediction. GPS can be created for each
101 individual and correlated with a trait in an independent sample, which yields an index of what could be
102 called *GPS heritability*, the extent to which GPS can explain variance in a trait. A GPS from a GWA study
103 of educational attainment (*EduYears*)¹⁹ predicts 4% of the variance of educational attainment in independent
104 samples¹⁹⁻²². No GWA studies of occupational status have been reported, but educational attainment and

105 occupational status correlate about 0.50 phenotypically²³⁻²⁵, and the *EduYears* GPS for educational
106 attainment predicts 2% of the variance of occupational status²¹, 2% of the variance of SES^{21,26}, and 7% of
107 the variance of family SES using children's DNA²⁷. GPS heritability (2-7%) is lower than SNP heritability
108 (20%) in part because GPS heritability is limited to specific SNPs shown to be associated with a trait and it
109 includes the trait's measurement error.

110

111 Heritability -- including GPS, SNP and twin heritability -- refers to the proportion of individual differences
112 that can be explained by inherited differences in individuals' DNA in a particular population at a particular
113 time. It describes what is, not what could be²⁸. The reported heritability of educational attainment and
114 occupational status from twin studies differs across birth cohorts and across countries^{2,3,5,19,6,7,29}. Specifically
115 it has been hypothesized that heritability of educational attainment can change following reform in
116 educational policy^{2,6}. Higher heritability estimates in twin studies have been noted in countries where
117 educational curriculum is highly standardized, such as the UK, because the standardization reduces
118 environmental differences between schools³⁰. However, research so far has yielded mixed results, with
119 some studies showing change in heritability estimates following a change in curriculum, or changes in the
120 heritability of achievement across birth cohorts, and other studies not showing such an effect^{3,6,29}. The major
121 limitation to date is that most research has been greatly underpowered; the twin method requires several
122 thousand twin pairs to achieve sufficient power to detect such gene-environment interactions³¹.

123

124 Few studies have investigated changes in SNP heritability as a function of environmental change^{4,19}; this
125 method requires several thousand unrelated individuals to detect gene-environment interactions. Only one
126 study has explored secular changes in GPS heritability. Using *EduYears* GPS, GPS heritability of
127 educational attainment was reported to be greater in older as compared to younger cohorts in Sweden¹⁹. This
128 decline in heritability is opposite to the results found in a twin study in Norway² and also in recent meta-
129 analyses of twin data³. However, no evidence has yet been reported for significant changes in GPS or SNP
130 heritability estimates following a major and abrupt social change.

131

132 Here we use GPS heritability and SNP heritability to estimate genetic influence on individual differences in
133 educational attainment and occupational status for 12 500 adults participating in the Estonian Genome
134 Centre, University of Tartu (EGCUT). EGCUT affords the unique opportunity to compare heritabilities in a
135 single population before and after the collapse of the Soviet Union. Estonia was occupied by the Soviet
136 Union after World War II and regained independence in 1991³².

137

138 The post-Soviet era is generally assumed to be more meritocratic in the sense that access to education and
139 occupation is to a greater extent based on ability^{32,33}. Given that education- and occupation-related abilities
140 are substantially due to inherited DNA differences between individuals, the greater equality of opportunity
141 implied by meritocracy should diminish the impact of environmental inequalities such as privilege or
142 privation. Inherited DNA differences will remain and will account for a relatively larger portion of
143 differences among individuals. In this sense, heritability can be viewed as an index of equality of
144 opportunity and meritocracy. In an entirely genetically driven meritocracy, genetic differences in ability
145 would account for all individual differences in educational attainment and occupational status.
146 Environmental differences that convey privilege or privation would account for none.

147

148 We used the EGCUT sample to test the hypothesis that heritability of educational attainment and
149 occupational status differs after a major environmental change. We compared SNP heritability and GPS
150 heritability for educational attainment and occupational status before and after the collapse of the Soviet
151 Union in Estonia. If independence led to greater meritocracy in terms of increased environmental
152 opportunity, the heritability of educational attainment and occupational status should be higher for
153 individuals who lived the majority of their studying and working lives in independent Estonia as compared
154 to those who lived during the Soviet Union.

155

156 Supplementary Table 1 shows means and standard deviations for height, educational attainment,
157 occupational status and SES for the whole sample, males and females separately and for historical eras

158 separately. ANOVA results indicate that historical group and sex explained up to 4% variance for the SES
 159 variables. For subsequent analyses, we controlled for sex effects by using sex-regressed standardized
 160 residuals.

161

162 Figure 1 compares GPS heritability in the Soviet and post-Soviet eras for the *EduYears* GPS (see Methods).
 163 For the whole sample, GPS heritability was 1.9% for occupational status and 2.3% for educational
 164 attainment (Figure 1). Using the less stringent cut-off of 15 years (Figure 1a), GPS heritability was
 165 significantly greater in the post-Soviet era compared to Soviet era for occupational status and educational
 166 attainment (see Supplementary Table 2 for all comparisons). These results are based on a GPS calculated at
 167 a 0.1 GWA study p-value threshold, which provided on average the best prediction across phenotypes and
 168 across historical eras. (Supplementary Figure 1 shows variance explained across multiple thresholds.)

169

170 The more stringent cut-off of 10 years yielded even larger GPS heritability differences (Figure 1b). For
 171 occupational status, GPS heritability was significantly greater in the post-Soviet era (5.6%) compared to the
 172 Soviet era (1.7%). Similarly for educational attainment, GPS heritability was significantly greater in the
 173 post-Soviet era (6.1%) than the Soviet era (2.1%). (See Supplementary Table 2 for all comparisons,
 174 including the composite SES score.)

175

176

177

 << Insert Figure 1 here >>

178 The GPS heritability estimates for composite SES (see Supplementary Figure 1) in the post-Soviet era
 179 (~7%) are in line with the GPS heritability estimates obtained in the UK²⁷, a meritocratic society, for family
 180 SES using offspring GPS. The difference arises from a significantly lower GPS heritability in the Soviet era.
 181 The results were very similar when additional analyses were run using variables that were not sex corrected
 182 (Supplementary Figure 2) and taking the transition period between Soviet and post-Soviet era into account
 183 (Supplementary Figure 3).

184

185 GPS heritability was also calculated for males and females separately (Supplementary Figure 4). The
 186 difference between GPS heritability in the Soviet and post-Soviet era was substantially greater for females
 187 compared to males, especially when a stricter cut-off of 10 years was used. This finding suggests that
 188 increased meritocracy after the Soviet era especially favored women, although the sample size and therefore
 189 the power of analyses were reduced when the sample was divided by gender.

190

191 We explored the extent to which the difference in GPS heritability between the Soviet and post-Soviet era
 192 differs by birth cohort. We divided the sample into birth cohorts using 10-year and 5-year intervals
 193 (Supplementary Figure 5). The difference in GPS heritability was greatest between the oldest and youngest
 194 birth cohort, the two birth cohorts that most clearly represent the Soviet versus post-Soviet. During the
 195 Soviet era, GPS heritability estimates fluctuate across birth cohorts but do not show a general trend of
 196 increasing GPS heritability, which could suggest that birth order itself underlies the Soviet versus post-
 197 Soviet GPS heritability difference. (See Supplementary Figure 6 for the distribution of sample size and SES
 198 for the Soviet and post-Soviet birth cohort groups and Supplementary Figure 7 for the distribution of
 199 *EduYears* GPS for the Soviet and post-Soviet birth cohort groups.)

200

201 We also calculated GPS scores using summary statistics from a GWA analysis of household income and
 202 social deprivation¹⁴, although this study was conducted using only the UK Biobank sample (N~112,000).
 203 However, these GPS scores are much less powerful predictors, explaining less than 1% of variance in
 204 independent samples. For this reason, these GPS scores explained less than 1% of the variance in our SES
 205 variables regardless of the historical era (Supplementary Figure 8-9).

206

207 We also used height as a control variable. *EduYears* GPS heritability was less than 1% regardless of the
 208 historical era (Supplementary Figure 10). This slight association is to be expected because height correlates

209 significantly but slightly with SES variables. For example, the genetic correlation between household
210 income (a good proxy for SES) and height has been shown to be around 0.2¹⁴.

211

212 Turning to SNP heritability, it should be noted our sample had much less power to detect SNP heritability
213 differences between the Soviet and post-Soviet groups. For the whole sample, SNP heritabilities were 15%
214 (SE 0.03) for occupational status and 18% (SE 0.03) for educational attainment (Figure 2). Despite having
215 less power to detect SNP heritability, SNP heritabilities were almost twice as high in the post-Soviet than
216 the Soviet era for educational attainment using age 15 as a cut-off (Figure 2). In the Soviet era, SNP
217 heritabilities were 17% (SE 0.04) for occupational status and 18% (SE 0.04) for educational attainment. In
218 contrast, in the post-Soviet era, SNP heritabilities were 23% (SE 0.16) and 37% (SE 0.14), respectively.
219 Although SNP heritabilities were larger in the post-Soviet era, these differences were not significantly
220 different as is evident from the standard errors.

221

222

<< Insert Figure 2 here >>

223

224 Height was also used as control variable for analyses of SNP heritabilities. SNP heritability was 32% for
225 height in the whole sample. For the Soviet era, SNP heritabilities was 33% for height, however, the post-
226 Soviet estimates were not significantly different (40%) (Supplementary Figure 11).

227

228 Our main finding is that heritabilities are higher for SES variables in the post-Soviet era as compared to the
229 Soviet era in the same Estonian population. GPS heritability for the composite SES measure (mean of
230 educational attainment and occupational status) was 7.5% in the post-Soviet era and 2.3% in the Soviet era
231 using the more stringent cut-off of 10 years. The variance in SES explained by the *EduYears* GPS seems
232 small compared to the twin study estimates of about 50% and SNP heritability estimates of about 25%.

233 However, we are only in the early stages of GPS research and the predictions are becoming stronger. SNP

234 heritabilities showed a similar trend as GPS heritabilities: SNP heritabilities for educational attainment were
235 twice as high in the post-Soviet era (37%) as compared to the Soviet era (17%).

236

237 A possible explanation for the increased heritability is increased meritocracy in Estonia following the
238 restoration of independence in 1991. By meritocracy, we refer to equal opportunity for access to education
239 and occupation and, when selection occurs, to meritocratic selection based on talent and effort, which are
240 substantially influenced by genetic factors, rather than on environmentally driven privilege or
241 discrimination. A meritocratic mechanism for the increased heritability of educational attainment and
242 occupational status in the post-Soviet era would be genotype-environment correlation in the sense that
243 individuals with equal opportunities are better able to select or to be selected for educational and
244 occupational environments correlated with their genetic propensities. When environmental differences in
245 access to education and occupation diminish, genetic differences increasingly account for educational
246 attainment and occupational status.

247

248 There are of course other possible explanations for increased GPS heritability in the post-Soviet era. The
249 largest increase in GPS heritability was observed for the participants who were in the youngest cohort when
250 Estonia regained the independence. Much has changed in the society after the collapse of the Soviet Union,
251 including wealth, culture, values -- all of which might contribute to the change in GPS heritability for the
252 cohort who lived, studied and worked the majority of their lives in independent Estonia. Migration and
253 changing population dynamics could also have affected the study results, although it should be noted that
254 there was substantial migration during the Soviet era (within the Soviet Union) as well as after the Soviet
255 era. However, we see no substantive hypothesis about the increased heritability following the collapse of the
256 Soviet Union as obvious as increased meritocracy, although this cannot be definitely tested. One point in
257 favor of the meritocracy hypothesis is that GPS heritability for SES in modern post-Soviet Estonia is similar
258 to GPS heritability in the UK, presumably a meritocratic society. The difference is that GPS heritability for
259 SES is lower in the Soviet era.

260 Another possible explanation is methodological. GPS scores were calculated for *EduYears* on the basis of a
261 meta-analytic GWA of heterogeneous cohorts. If the GWA discovery sample weights were closer to the
262 post-Soviet sample in the present study, then more variance would be explained in the post-Soviet compared
263 to Soviet sample.

264

265

266 *Equal educational opportunities*

267 The meritocracy hypothesis assumes that educational and occupational success was less meritocratic in the
268 Soviet era. In the Soviet era, access to primary education was universal and universal secondary education
269 was introduced in the 1960s. However, the quality of teaching and even the curricula varied widely across
270 schools^{34,35}. Within schools, students were divided into one of the three different tracks, with limited
271 movement between tracks: vocational training, secondary education and (special) secondary education³⁶.
272 This tracking was partly done based on merit (school achievement), but social-political ranking played a
273 significant part as well. The number of students admitted to each track depended on the economic and
274 social goals of central planning at the time; individual aspirations and ability were not considered to be as
275 important³⁵. Access to tertiary education from lower ‘ranks’ in the social-political system was limited; for
276 example, students who were religious were not admitted^{34,36}. In this way, the Soviet education system
277 created environmental inequalities both directly and indirectly³⁵. Importantly, university education was not
278 as highly valued in society as it is now and this was accompanied by limited competition for university
279 places, with an average of only two applicants per position. Admissions to university remained low
280 throughout the Soviet era, which restricted any selection, meritocratic or not.

281

282 Since regaining independence, education in Estonia has become more meritocratic in terms of educational
283 opportunity. Many educational reforms were introduced after the collapse of the Soviet Union with the aim
284 of building a more egalitarian and effective educational system. Currently, almost all students complete
285 elementary education and the rate of completing secondary education is among the highest in the OECD
286 countries. Estonian equality in education is now above the OECD average, with limited variation in teaching

287 standards between schools. The quality of teaching is considered to be excellent according to international
288 standards and Estonia is ranked among the highest performing educational systems according to PISA
289 surveys in 2012 and 2015^{37,38}. This overall educational excellence, and the limited number of selective or
290 private schools, suggests that there is equal opportunity and access to good education for all at primary and
291 secondary level of education. We hypothesized that equality of opportunity should increase the heritability
292 of educational achievement by making it possible for children to select, modify and choose educational
293 experiences correlated with their education-related genetically influenced propensities, which include
294 appetites as well as abilities. Educational achievement in turn contributes importantly to eventual
295 educational attainment and occupational status.

296

297 For tertiary education, in addition to self-selection, students are now selected for university largely on the
298 basis of ability and prior achievement, rather than environmentally driven privilege. Selection is not based
299 on socio-political or religious considerations as in the Soviet era. Nor is selection based on the ability to pay
300 tuition, because almost all university education is free. There is also greater opportunity for selection for
301 university admission in the post-Soviet era because university applications and admissions increased
302 exponentially in the 1990s; for example, applications to University of Tartu have increased threefold
303 compared to the Soviet era³⁴.

304

305 *Equal access to occupation*

306 During the Soviet era, the economy and labor market was mainly characterized by centralized control, with
307 the majority of workforce assigned to jobs in manufacturing and agriculture. Occupational status was
308 determined more by loyalty to the communist party than by ability, achievement or qualifications.

309 Recommendations for job positions and promotion always came from party leaders, although educational
310 qualifications were also needed for certain positions³⁹. The economy and labor market had very limited
311 workforce mobility³⁶.

312

313 Inequality in occupations during the Soviet era was even more dramatic for females than males. During the
314 Soviet era there was an increase in participation of women in workforce, meaning that both men and women
315 were largely employed. However, this did not lead to occupational equality; women often did jobs requiring
316 lower level of skills⁴⁰. Although Soviet ideology argued for gender equality, this was not carried out in
317 practice⁴¹.

318
319 The transition from the Soviet Union to a prosperous independent Estonia was more difficult than
320 anticipated. After the restoration of independence in Estonia the living standards were low, the economy
321 was struggling, and the situation worsened with a major recession until 1994 when Estonia joined the
322 European Union^{32,33}. Equality of opportunity increased as the Estonia became more integrated with the
323 west⁴².

324
325 These historical events may explain why *EduYears* GPS did not explain more variance in SES in the
326 transition time compared to the Soviet era. Our results suggested that *EduYears* GPS heritability is greatest
327 for the youngest participants who had lived, studied and worked in independent Estonia the longest.
328 Gender equality in Estonia started to improve, albeit gradually, after the collapse of the Soviet Union⁴³. This
329 was mirrored by an interesting facet of the results in the present study showing that GPS heritability
330 increased more dramatically for females compared to males following the collapse of the Soviet Union.
331 These results further support the meritocratic hypothesis specifically in relation to gender.

332
333 *Future research directions*

334 The present analyses excluded participants who were younger than 25 at the time of data collection because
335 they may not yet have achieved their highest educational qualifications or reached their highest occupational
336 status. Linking the EGCUT database with data from the Estonian Department of Education will make it
337 possible in the future to include those individuals who were excluded as they complete their education and
338 reach their ultimate occupational status. This will increase the size of our post-Soviet sample and thus the

339 power of our SNP and GPS heritability comparisons. Because these individuals grew up completely in the
340 post-Soviet era, we predict that they will show even greater heritability of SES. Increased sample size would
341 also provide greater power to investigate further gender differences in GPS heritability.

342

343 Another interesting direction for research concerns the relationship between education and fecundity.

344 Decreased fecundity in Iceland among highly educated citizens has been reported to result in lower GPS

345 scores for *EduYears*, although the effect is very small²⁰. According to Statistics Estonia, the population in

346 Estonia has been decreasing for decades (<http://www.stat.ee/news-release-2017-008>), although it increased

347 for the first time in 2016. We plan to investigate the extent to which decreasing fecundity comes

348 disproportionately from highly educated individuals, in which case we might expect lower average GPS in

349 the most recent birth cohorts. Our preliminary analyses did not support this hypothesis in that the average

350 *EduYears* GPS did not differ across birth cohorts (Supplementary Figure 12), although we did not study

351 fecundity here.

352

353 Studying parent-offspring resemblance to understand intergenerational social mobility is also part of our

354 future research plans in EGCUT. Intergenerational social mobility is often assumed to be solely due to

355 environmental factors. For example, the OECD uses parent-offspring resemblance in SES outcomes to

356 assess intergenerational social mobility, assuming that this resemblance is environmentally mediated. Our

357 current results and results from other studies show that educational and occupational outcomes are partly

358 explained by genetic factors. Because parents and offspring are on average 50% similar genetically, parent-

359 offspring resemblance is also likely to show genetic influence for SES. From this perspective, parent-

360 offspring resemblance could be viewed as an index of equality rather than inequality. In other words, if

361 environmental inequalities were eliminated, genetic resemblance between parents and offspring would

362 completely account for parent-offspring resemblance.

363

364 While our analyses provided evidence for changes in GPS and SNP heritabilities following the major social
365 change from a communist to a capitalist society, no definite conclusions can be drawn. It will be necessary
366 to replicate the results of the present analyses using data from a different country that has gone through
367 similar abrupt social change. A country that used to be part of the Soviet Union and has regained
368 independence would be ideal; however, we are not aware of such a replication sample available at this time.
369 We hope that our results lead to future molecular genetic studies researching gene-environment interactions
370 of this sort that are now possible using GPS scores.

371

372 Another direction for future research is to consider intermediate phenotypes such as cognitive abilities that
373 might mediate these changes in the distal outcomes of educational attainment and occupational status. In
374 addition, the precision and power of all of these SNP and GPS analyses will increase as the power of GWA
375 studies increases.

376

377 *Meritocracy or social justice?*

378

379 In closing, we wish to emphasize that we are not advocating meritocracy, although these issues are more an
380 issue of values than science. At first glance meritocracy seems unquestionably good, but it could have
381 unintended consequences such as creating social inequalities if societal rewards such as wealth are doled out
382 on the basis of genetically driven abilities. The word meritocracy was coined by Michael Young whose
383 book, *The Rise and Fall of the Meritocracy*⁴⁴, was meant as a cautionary tale about the dangers of
384 meritocracy. The value system underlying meritocracy is that the point of education is to get better test
385 scores in order to get better jobs, and that the point of occupations is to achieve high status and make lots of
386 money. A different way to look at education is as a time to learn basic skills but also to learn how to learn
387 and to enjoy learning. It is a decade when children can find out what they like to do and what they are good
388 at doing, finding their genetic selves. If education were universally good, there would be no need for
389 selection, especially at the level of primary and second education, and thus there would be no need to apply
390 meritocratic criteria.

391

392 Similarly with occupations, where selection cannot be avoided, we will end up with a lot of frustrated people
393 if we only value high-status occupations that earn lots of money. Society needs people who are good care
394 workers, nurses, plumbers, public servants, and people in the service industry. To the extent that selection is
395 necessary it should be meritocratic, but it is possible to imagine an occupational system that is not driven so
396 much by monetary reward. For example, society could choose to reduce income inequality with a tax system
397 that redistributes wealth.

398

399 In his book, *The Myth of Meritocracy*, James Bloodworth (2016)⁴⁵ argues that meritocracy leads to an
400 inherent inequality of opportunity and reward based on genetic differences. He suggests that we need to
401 replace meritocracy with what he calls a just society in which everyone could live well.

402

403 **Methods**

404 **Sample**

405 The sample for the present study was drawn from the Estonian Genome Centre, University of Tartu
406 (EGCUT) sample. Ethical approval was granted by the Research Ethics Committee of the University of
407 Tartu (approval 245/T-16).

408 EGCUT is a population-based study with a sample size of over 52 000 individuals (all participants ≥ 18 years
409 of age), which comprises 5% of the adult population in Estonia. Genome-wide genetic data are available for
410 approximately 20 000 of these individuals. EGCUT has been shown to be representative of the Estonian
411 population in terms of age and geographical location while females are overrepresented, 66% female as
412 compared to 55% in the adult population in Estonia⁴⁷. EGCUT is also reasonably representative in terms of
413 educational attainment when compared to national figures from the Department of Statistics Estonia
414 (<http://www.stat.ee/phc2011>) (Supplementary Table 4). The initial sample for the present study included all
415 participants with available genotypic and phenotypic data. All individuals who were 25 or younger were
416 excluded from the analyses, as it is possible that these young individuals had not yet reached their highest

417 educational level and highest occupation. The sample size before exclusions included 17 990 participants (7
418 409 males and 10 581 females). After exclusions (removing participants who were under 25 at the time of
419 data collection and following quality control) the sample size was reduced to 12 490. Sample size for each
420 measure separately is presented in Supplementary Table 1.

421

422 The sample was divided into two historical eras: the Soviet era and the post-Soviet era. Estonia regained
423 independence in 1991; consequently, all participants who were born on or after 1976 went into secondary or
424 further education in the post-Soviet era (i.e., they were aged 15 or younger when Estonia regained
425 independence) and the rest of the sample was aged 16 or older when Estonia regained independence. This is
426 an arbitrary cut-off that does not take into account the transition time between communist to capitalist
427 society since societal changes take time to have an effect on people's lives. We assumed that young
428 individuals were in the middle of their educational career, still making decisions about their universities and
429 post-graduate degrees. We therefore repeated the analyses allowing for a transition period before and after
430 the collapse of the Soviet Union assigning participants who were 16-25 year olds in 1991 to a 'transition'
431 group. In addition, we used another cut-off to define the Soviet and post-Soviet groups, assigning all
432 participants who were aged 10 or younger at the time of the restoration of independence in Estonia to the
433 post-Soviet group and participants who were older than 10 years to the Soviet group.

434

435 **Measures**

436 *Educational attainment*

437 Educational attainment was assessed using a 10-point self-reported scale from no elementary education to
438 postgraduate degree. The measure and scoring followed closely the International Standard Classification of
439 Education (ISCED: [http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-
440 education.aspx](http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx)). However, some participants were studying towards an undergraduate or postgraduate
441 degree at the time of the data collection, so additional points were added to the scale. Our measure included
442 the following 10 categories (rather than the 8 categories that were in the original scale) for educational

443 attainment: (1) no educational qualifications, (2) elementary school education, (3) basic education/ junior
444 grade of high school, (4) secondary school/high school education, (5) vocational qualification/community
445 college, (6) professional higher education, (7) studying towards university degree, (8) university degree, (9)
446 studying towards postgraduate degree, (10) postgraduate degree.

447 *Occupational Status*

448 Occupational status was assessed with two questions: “What is your professional status right now?” and
449 “What has been your main professional status (the occupation you kept the longest)?” These occupational
450 status responses were scored according to the International Standard Classifications of Occupations (ISCO:
451 <http://www.ilo.org/public/english/bureau/stat/isco/>). ISCO is a widely used and reliable measure⁴⁸⁻⁵¹. ISCO
452 classification assigns occupational status to broad groups (as well as more specific subgroups), taking into
453 account the skills and education level required for occupation as well as the potential earnings. The present
454 study used nine occupational status groups, classified in ISCO as the following categories, scored from 1 to
455 9 respectively: (1) elementary occupations (cleaners, helpers, laborers), (2) plant and machine operators,
456 assemblers, (3) craft and related trades workers, (4) skilled agricultural, forestry and fishery workers, (5)
457 service and sales workers, (6) clerical support workers, (7) technicians and associate professionals, (8)
458 professionals, (9) legislators, senior officials and managers. The current occupational status and the main
459 occupational status correlated 0.46. Both the current and the main occupational status had missing data;
460 therefore, to increase power and sample size, a composite measure of occupational status was created by
461 taking the mean of current and the longest held occupations; if only one measure were available then that
462 measure was used. The same measure was used for both the Soviet and post-Soviet eras. Although, the
463 classification of occupational status and the potential pay could have been different during the Soviet era, we
464 assume that occupational positions (and the prestige of them) still fit into the broad ISCO categories.

465 *SES*

466 Because educational attainment and occupational status correlated 0.62, we calculated a mean as an index of
467 general socioeconomic status (SES). SES is usually operationalized as a composite measure that includes
468 income as well as occupational status and educational attainment. Although the measure of SES used in the

469 present study does not include family income, occupational classification takes into account the potential
470 earnings and prestige of the occupation. Therefore, we consider our composite measure of occupational
471 status and educational attainment to be a reasonable index of SES.

472 *Height*

473 Height was used as control variable in the analyses; we had no hypothesis about changes in the SNP or GPS
474 heritabilities following the shift from a communist to a capitalist society. Height was assessed in person by
475 the researchers and was measured in cm.

476 *Genotyping*

477 Venous blood was collected from all 52 000 participants of EGCUT. DNA and plasma were immediately
478 extracted from the blood and stored in EGCUT Core Laboratory of EGCUT in Tartu, Estonia. Genome-wide
479 genotyping was assayed for 20 000 participants using three Illumina arrays: Illumina HumanCoreExome,
480 Illumina Human370 CNV and Illumina OmniExpress in the Core Laboratory of EGCUT in Tartu, Estonia.
481 Data were harmonized across the three arrays and harmonized data were used for all analyses (see Quality
482 Control).

483 *Quality Control*

484 Genotype quality control were performed using Illumina GenomeStudio 3.1 and PLINK 1.07⁵². Standard
485 quality control analyses were conducted at both the individual level and the SNP level excluding individuals
486 with genotype call rate < 95%, sex discrepancies (using the heterozygosity rate of X-chromosome) and
487 excess heterozygosity (mean±3SD). Additionally, duplicates and multidimensional-scaling (MDS) outliers
488 were excluded. At the SNP level, we excluded SNPs with minor allele frequency (MAF) < 1%, call rate <
489 95%, failure of the Hardy-Weinberg Equilibrium (HWE) exact test (threshold 1×10^{-6}), A/T or C/G and sex
490 chromosome SNPs were removed. Phasing and imputation of the cleaned data were performed using
491 ShapeIT v2⁵³ and IMPUTE v2.3.1⁵⁴ with 1000 Genomes Phase 3 Oct 2014 imputation reference panel based
492 on 5 008 haplotypes⁴ (www.1000genomes.org). IMPUTE2 builds custom-reference panels for each
493 individual to be imputed and so is the best-suited software for imputing genotype data from Estonians, for
494 whom no population-specific reference panel exists.

495

496 After imputation, further quality control was carried out. SNPs with MAF < 1%, and SNPs with poor
 497 imputation quality (info score < 0.30) or failure of the HWE exact test (threshold 1×10^{-6}) were removed. We
 498 harmonized the genotyped datasets across the 3 arrays removing duplicate individuals and duplicate
 499 markers. Other standard quality control methods were applied removing SNPs and samples with call rate
 500 <0.97. The quality control was performed on each array separately, and was repeated after harmonization.
 501 After harmonization and quality control the final sample included 4 052 281 variants and 16 397 individuals
 502 (see Supplementary Table S5 for number of SNPs dropped after each step of quality control).

503

504 To control for ancestral stratification, principal component analyses were performed after pruning to remove
 505 markers in linkage disequilibrium (200kb window using $R^2 > 0.05$). The first 10 principal components were
 506 used as covariates in the genetic analyses.

507

508 **Statistical Analyses**

509 Means and variances for measures were calculated, comparing the Soviet era and post-Soviet era, as well as
 510 sex differences. Mean differences were tested using ANOVA (Supplementary Table 1). Because significant,
 511 though small, sex differences emerged for both occupational status and educational attainment, explaining 2-
 512 4% of the variance in SES measures, we corrected the measures for mean sex differences using the
 513 regression method. In addition, we repeated the analyses without sex correction and calculated the variance
 514 explained by GPSs created separately for males and females. No correction for multiple testing was done, as
 515 all analyses tested just one hypothesis and we were interested in the effect size rather than the significance
 516 level.

517 *Genome-wide polygenic scores*

518 Genome-wide polygenic scores (GPSs) aggregate the effects of individual SNPs shown to be associated
 519 with the trait in a GWA study⁵⁵. GPSs were calculated for 16 398 participants using p-values and β - weights
 520 obtained from summary statistics from the Okbay et al (2016) GWA analysis¹⁹ of years of education

521 (*EduYears*) with the PRSice program⁵⁶ using multiple p-value thresholds (0.001; 0.05; 0.1; 0.2; 0.3; 0.4;
 522 0.5). Of the 293 723 participants in the *EduYears* GWAS, the present study excluded 23andMe participants,
 523 for legal reasons, and excluded all participants from EGCUT, resulting in a sample of 208 596 individuals
 524 (see Supplementary Table 6 for cohort description). SNPs were clumped in PRSice for linkage
 525 disequilibrium, using a cut-off of $R^2=0.1$ within a 250-kb window. GWA summary statistics were obtained
 526 from the sample of 208 596 individuals, and p-values and β - weights were used to calculate the *EduYears*
 527 GPS. Delta R^2 are reported as the estimates of variance explained by adding the GPS to the regression model
 528 that included 10 principal components to control for population stratification.
 529 We also calculated GPS scores using p using p-values and β - weights obtained from summary statistics from
 530 the Hill et al (2016) GWA analysis¹⁴ of household income and social deprivation with the PRSice
 531 program⁵⁶ using the same procedure.

532 The difference in GPS heritabilities was evaluated using Fisher's exact test with Z to r transformation that
 533 assesses the significance in the difference in correlation coefficients in independent samples using both the
 534 effect sizes and sample sizes in the two samples⁵⁷.

535 *SNP heritability*

536 SNP heritability estimates genetic and residual (environmental) components of variance directly from DNA
 537 using unrelated individuals and hundreds of thousands of SNPs (single nucleotide polymorphisms) from
 538 thousands of individuals⁵⁸. Using GCTA software, a genetic relatedness matrix was calculated weighting the
 539 pairwise genetic similarities with allele frequencies across all genotyped SNPs^{58,59}. Individuals found to be
 540 even remotely related (relatedness >0.05) were removed from the analyses. We repeated the analyses when
 541 using the more stringent cut-off of 0.025, but this did not make any difference in SNP heritability estimates.
 542 This matrix of pair-by-pair genetic similarities were then compared to the matrix of pair-by-pair phenotypic
 543 similarity using residual maximum likelihood estimation^{58,59}. This method only assesses additive effects
 544 captured by the common SNPs genotyped on the DNA array, and does not take into account gene-gene or
 545 gene-environment interactions or rare DNA variants, but these are unlikely to have a strong influence on the
 546 phenotype^{58,60}. Prior to SNP heritability analyses we adjusted educational attainment and occupational status

547 for sex using regression; standardized residuals were used in all analyses. To correct for the slight skew in
 548 the data, all measures were transformed to a normal distribution using the van der Waerden rank-based
 549 transformation^{61,62}.

550 *Statistical power*

551 Power for estimating SNP and GPS heritability was estimated using the online tool GCTA-GREML power
 552 calculator⁶³ and AVENGEME R code^{55,64}. Our sample provided more than 80% power to detect GPS
 553 associations that explained 4% variance under the following circumstances: GWAS discovery sample size of
 554 208 596, our target sample of 12 500 participants (the power did not change when we calculated power with
 555 a target sample of 2100 or a target sample of 680 for post-Soviet subgroups); number of independent SNPs
 556 in the GPS=20,000; proportion of variance explained in discovery sample =4%, covariance between genetic
 557 effect sizes in the discovery and target sample =4%; and proportion of SNPs with no effects on the
 558 discovery trait = 99%; range of p-values from GWA summary statistics= 0.00- 0.5). These assumptions are
 559 somewhat arbitrary, but the power calculations did not change when parameters for the power calculations
 560 were changed (for example, changing the proportion of SNPs with no effects on the trait in the discovery
 561 sample to 50%). In addition, the power of our sample sizes to detect the expected GPS effect is supported by
 562 a much simpler approach: *EduYears* GPS predicts around 4% of variance in independent samples, a
 563 correlation of 0.20, which requires a sample size of only 150 for 80% power ($p = .05$, one-tailed
 564 (<http://www.sample-size.net/correlation-sample-size/>).

565

566 Power for estimating SNP heritability is 99% to detect a SNP heritability of 20% for the whole sample. For
 567 the Soviet-era subsample, we had 99% power to detect a SNP heritability of 20%, but power was only 24%
 568 in the post-Soviet era (the power to detect heritability of 35% was 64% in the post-Soviet era). Therefore,
 569 little confidence is warranted for assessing differences in SNP heritability in the Soviet and the post-Soviet
 570 groups.

571

572 **Supplementary information** accompanies this article.
 573

574 **Acknowledgments**

575 We gratefully acknowledge the ongoing contribution of the participants in the Estonian Genome Centre
576 University of Tartu. RP is supported by the UK Medical Research Council [MR/M021475/1 and previously
577 G0901245], with additional support from the US National Institutes of Health [HD044454; HD059215].
578 KR, EK and SS are supported by a Medical Research Council studentship. RP is supported by a Medical
579 Research Council Research Professorship award [G19/2] and a European Research Council Advanced
580 Investigator award [295366]. JRIC is funded by the NIHR Biomedical Research Centre at South London and
581 Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the
582 author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. T.E. & A.M. were
583 supported by Grant from the Est.RC IUT 20-60 (A.M.), PUT-1660 (T.E) and by CoEx for Genomics and
584 Translational Medicine (GENTRANSMED). The funders had no role in study design, data collection and
585 analysis, decision to publish, or preparation of the manuscript.

586 **Author Contributions**

587 Conceived and designed the experiments: KR, RP. Analyzed the data: KR, MT, EK, JRIC. Wrote the paper:
588 KR, MT, EK, TE, AM, RP. All authors approved the final draft of the paper.

589

590 **Data availability**

591

592 For information on data availability, please see the Estonian Genome Centre, University of Tartu (EGCUT)
593 data access policy. This can be found at: <http://www.geenivaramu.ee/en/biobank.ee/data-access>

594

595

596 **Conflict of interest:** The authors declare no conflicting of interest.

- 598 1. Lykken, D. T., Bouchard Jr., T. J., McGue, M. & Tellegen, A. The Minnesota Twin Family Registry:
599 some initial findings. *Acta Genet Med Gemellol* **39**, 35–70 (1990).
- 600 2. Heath, A. C. *et al.* Education policy and the heritability of educational attainment. *Nature* **314**, 734–
601 736 (1985).
- 602 3. Branigan, A. R., Mccallum, K. J. & Freese, J. Variation in the heritability of educational attainment:
603 An international meta-analysis. *Soc. Forces* **92**, 109–140 (2013).
- 604 4. Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with
605 educational attainment. *Science*, **340** (6139), 1467–71 (2013).
- 606 5. Tambs, K., Sundet, J. M., Magnus, P. & Berg, K. Genetic and environmental contributions to the
607 covariance between occupational status, educational attainment, and IQ: A study of twins. *Behav.*
608 *Genet.* **19**, 209–222 (1989).
- 609 6. Colodro-Conde, L., Rijdsdijk, F., Tornero-Gómez, M. J., Sánchez-Romera, J. F. & Ordoñana, J. R.
610 Equality in educational policy and the heritability of educational attainment. *PLoS One* **10**, e0143796
611 (2015).
- 612 7. Lichtenstein, P., Pedersen, N. L. & McClearn, G. E. The origins of individual differences in
613 occupational status and educational level: a study of twins reared apart and together. *Acta Sociol.* **35**,
614 13–31 (1992).
- 615 8. Adler, N. E. *et al.* Socioeconomic status and health: the challenge of the gradient. *Am. Psychol.* **49**,
616 15–24 (1994).
- 617 9. Cutler, D. M. & Lleras-Muney, A. *Education and health: insights from international comparisons.*
618 *NBER Working Papers* (2012).
- 619 10. Cutler, D. M., Lleras-Muney, A. & Vogl, T. *Socioeconomic status and health: dimensions and*
620 *mechanisms. NBER Working Papers* (2008).
- 621 11. Batty, G. D., Deary, I. J. & Gottfredson, L. S. Premorbid (early life) IQ and Later Mortality Risk:
622 Systematic Review. *Ann. Epidemiol.* **17**, 278–288 (2007).
- 623 12. von Stumm, S., Deary, I. J. & Hagger-Johnson, G. Life-course pathways to psychological distress: a
624 cohort study. *BMJ Open* **3**, e002772 (2013).
- 625 13. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. Genome-wide complex trait analysis (GCTA):
626 Methods, data analyses, and interpretations. *Methods Mol. Biol.* **1019**, 215–236 (2013).
- 627 14. Hill, W. D. *et al.* Molecular genetic contributions to social deprivation and household income in UK
628 Biobank (n = 112,151). *Curr. Biol.* **26**, 3083–3089 (2016).
- 629 15. Marioni, R. E. *et al.* Molecular genetic contributions to socioeconomic status and intelligence.
630 *Intelligence* **44**, 26–32 (2014).
- 631 16. Benjamin, D. J. *et al.* The genetic architecture of economic and political preferences. *Proc. Natl.*
632 *Acad. Sci.* **109**, 8026–8031 (2012).
- 633 17. Davies, G. *et al.* Genome-wide association study of cognitive functions and educational attainment in
634 UK Biobank (N=112 151). *Mol. Psychiatry* **21**, 758–67 (2016).
- 635 18. Hyytinen, A., Ilmakunnas, P., Johansson, E. & Toivanen, O. *Heritability of lifetime income. Helsinki*
636 *Centre of Economic Research* (2013).
- 637 19. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational
638 attainment. *Nature* **533**, 539–542 (2016).
- 639 20. Kong, A. *et al.* Selection against variants in the genome associated with educational attainment. *Proc.*
640 *Natl. Acad. Sci.* **114**, E727-32 (2017).
- 641 21. Belsky, D. W. *et al.* The genetics of success: how single-nucleotide polymorphisms associated with

RUNNING TITLE: GENETIC MERITOCRACY IN ESTONIA

- 542 educational attainment relate to life-course development. *Psychol. Sci.* **27**, 957–972 (2016).
- 543 22. Hugh-Jones, D., Verweij, K. J. H., St. Pourcain, B. & Abdellaoui, A. Assortative mating on
544 educational attainment leads to genetic spousal resemblance for polygenic scores. *Intelligence* **59**,
545 103–108 (2016).
- 546 23. Hollingshead, A. Four factor index of social status. *Yale Journal of Sociology* **8**, 21–52 (1975).
- 547 24. Sirin, S. R. Socioeconomic status and academic achievement: a meta-analytic review of research.
548 *Rev. Educ. Res.* **75**, 417–453 (2005).
- 549 25. White, K. R. The relation between socioeconomic status and academic achievement. *Psychol. Bull.*
550 **91**, 461–481 (1982).
- 551 26. Domingue, B. W., Belsky, D. W., Conley, D., Harris, K. M. & Boardman, J. D. Polygenic influence
552 on educational attainment. *AERA Open* **1**, 2332858415599972 (2015).
- 553 27. Selzam, S. *et al.* Predicting educational achievement from DNA. *Mol. Psychiatry* **22**, 267–272
554 (2017).
- 555 28. Knopik, V. S., Neiderhiser, J. M., DeFries, J. C. & Plomin, R. *Behavioral Genetics. 7th ed.* (Worth
556 Publishers, New York, 2017).
- 557 29. Baker, L. A., Treloar, S. A., Reynolds, C. A., Heath, A. C. & Martin, N. G. Genetics of educational
558 attainment in Australian twins: Sex differences and secular changes. *Behav. Genet.* **26**, 89–102
559 (1996).
- 560 30. Samuelsson, S. *et al.* Environmental and genetic influences on prereading skills in Australia,
561 Scandinavia, and the United States. *Journal of Educational Psychology* **97**, 705–722 (2005).
- 562 31. Hanscombe, K. B. *et al.* Socioeconomic status (SES) and children’s intelligence (IQ): in a UK-
563 representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS One* **7**,
564 e30320 (2012).
- 565 32. Laar, M. *Estonia’s way.* (Pegasus, Tallinn, Estonia, 2007).
- 566 33. Laar, M. The Estonian economic miracle. *Backgrounder* **2060**, 1–12 (2007).
- 567 34. Saar, E. Changes in intergenerational mobility and educational inequality in Estonia: Comparative
568 analysis of cohorts born between 1930 and 1974. *Eur. Sociol. Rev.* **26**, 367–383 (2010).
- 569 35. Saar, E. Transitions to Tertiary Education in Belarus and the Baltic Countries. *Eur. Sociol. Rev.* **13**,
570 139–158 (1997).
- 571 36. Titma, M., Tuma, N. B. & Roosma, K. Education as a factor in intergenerational mobility in soviet
572 society. *European Sociological Review* **19**, 281–297+i (2003).
- 573 37. OECD. *Education policy outlook: Estonia.* (2016).
- 574 38. OECD. *Equity and quality in education - supporting disadvantaged students and schools.* (2011).
- 575 39. Titma, M. & Roots, A. Intragenerational mobility in successor states of the USSR. *Eur. Soc.* **8**, 493–
576 526 (2006).
- 577 40. Carnaghan, E. & Bahry, D. Political attitudes and the gender gap in the USSR. *Comp. Polit.* **22**, 379–
578 399 (1990).
- 579 41. Katz, K. *Gender, work and wages in the Soviet Union: a legacy of discrimination.* (Palgrave
580 Macmillan UK, 2001).
- 581 42. Boughton, J. *Tearing down walls: the International Monetary Fund, 1990-1999.* (2012).
- 582 43. Silova, I. & Magno, C. Gender equity unmasked: democracy, gender, and education in
583 Central/Southeastern Europe and the former Soviet Union. *Comp. Educ. Rev.* **48**, 417–442 (2004).
- 584 44. Young, M. *The rise and fall of the meritocracy.* (Penguin Books, 1965).
- 585 45. Bloodworth, J. *The myth of meritocracy.* (Biteback Publishing, 2016).

- 586 46. Piketty, T. *Capital in the twenty-first century*. (Harvard University Press., 2014).
- 587 47. Leitsalu, L. *et al.* Cohort profile: Estonian Biobank of the Estonian Genome Center, University of
588 Tartu. *Int. J. Epidemiol.* **44**, 1137–1147 (2015).
- 589 48. Ganzeboom, H. B. G. A New International Socio-Economic Index [ISEI] of occupational status for
590 the International Standard Classification of Occupation 2008 [ISCO-08] constructed with data from
591 the ISSP 2002–2007; with an analysis of quality of occupational measurement in ISS. *Annu. Conf. Int.*
592 *Soc. Surv. Program.*, Lisbon (2010).
- 593 49. Ganzeboom, H. B. & Treiman, D. J. in *Advances in Cross-National Comparison. A European*
594 *Working Book for Demographic and Socio-Economic Variables*. (eds. Hoffmeyer-Zlotnik, J. H. P. &
595 Christof Wolf) 159–193 (New York: Kluwer Academic Press., 2003).
- 596 50. Wolf, C. The ISCO-88 International Standard Classification Of Occupations in cross-national survey
597 research. *Bull. Methodol. Sociol.* **54**, 23–40 (1997).
- 598 51. Kromhout, H. The use of occupation and industry classifications in general population studies. *Int. J.*
599 *Epidemiol.* **32**, 419–428 (2003).
- 700 52. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage
701 analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- 702 53. Delaneau, O., Zagury, J.-F. & Marchini, J. Improved whole-chromosome phasing for disease and
703 population genetic studies. *Nat. Methods* **10**, 5–6 (2013).
- 704 54. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for
705 the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
- 706 55. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **9**, e1003348
707 (2013).
- 708 56. Euesden, J., Lewis, C. M. & O’Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics*
709 **31**, 1466–1468 (2014).
- 710 57. Fisher, R. . On the probable error of a coefficient of correlation deduced from a small sample. *Metron*
711 **1**, 3–32 (1921).
- 712 58. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat.*
713 *Genet.* **42**, 565–9 (2010).
- 714 59. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait
715 analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
- 716 60. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era — concepts and
717 misconceptions. *Nat. Rev. Genet.* **9**, 255–266 (2008).
- 718 61. Lehmann, E. *Nonparametric Statistical Methods Based on Ranks*. (Holden-Day, San Francisco, CA,
719 1975).
- 720 62. Van Der Waerden BL. On the sources of my book *Moderne Algebra*. *Hist. Math.* **2**, 31–40 (1975).
- 721 63. Visscher, P. M. *et al.* Statistical power to detect genetic (co)variance of complex traits using SNP data
722 in unrelated samples. *PLoS Genet.* **10**, e1004269 (2014).
- 723 64. Palla, L. & Dudbridge, F. A fast method that uses polygenic scores to estimate the variance explained
724 by genome-wide marker panels and the proportion of variants affecting a trait. *Am. J. Hum. Genet.*
725 **97**, 250–259 (2015).

726
727
728
729
730
731

732

733

734 **Figures**

735

736 **Figure 1.** Variance explained by *EduYears* GPS in the post-Soviet (PS) and Soviet (S) groups. We
 737 calculated GPS using a 0.1 GWA study p-value threshold for educational attainment (EA) and occupational
 738 status (OS) for the whole EGCUT sample (N(EA)=12 483; N(OS)= 11 419) and when divided into historical
 739 eras using two cut-offs: (a) The post-Soviet (PS) group included participants 15 or younger when
 740 independence was regained and the Soviet (S) group included the rest of the participants (N(EA_S)=10 381;
 741 N(EA_PS)=2 102; N(OS_S)= 9 417; N(OS_PS)= 2 002); (b) The post-Soviet (PS) group included
 742 participants 10 or younger when independence was regained and the Soviet (S) group included the rest of
 743 the participants (N(EA_S)=11 808; N(EA_PS)=675; N(OS_S)= 10 767; N(OS_PS)= 652).

744

745 **Figure 2.** SNP heritabilities showing the proportion of variance explained by additive effects of common
 746 SNPs (SE as error bars) for the whole EGCUT sample and for the Soviet and post-Soviet groups using a cut-
 747 off of 15 years. SNP heritabilities were adjusted for population stratification. (N(EA)=12 483; N(OS)= 11
 748 419; N(EA_S)=10 381; N(EA_PS)=2 102; N(OS_S)= 9 417; N(OS_PS)= 2 002).

749





