Article:
Armstrong-Carter, E, Trejo, S, Hill, LJB orcid.org/0000-0002-4069-5121 et al. (3 more authors) (2020) The Earliest Origins of Genetic Nurture: The Prenatal Environment Mediates the Association Between Maternal Genetics and Child Development. Psychological Science, 31 (7). pp. 781-791. ISSN 0956-7976

https://doi.org/10.1177/0956797620917209

The Earliest Origins of Genetic Nurture:

Prenatal Environment Mediates the Association Between

Maternal Genetics and Child Development

Supplemental Information

1. Additional Methods

1A. Measures

Measures used in this study are described in Table A1. Additional details are available via the BiB website, https://borninbradford.nhs.uk/research/documents-data/.

Our health and SES composites were computed after separately standardizing all variables used for construction of each composite. Loadings for the components of the composites are:

- SES composite. Single -0.39, neighborhood deprivation -0.39, financial difficulties -0.36, government benefits -0.30, maternal leave -0.01, education 0.46, employment status 0.51.
- Health composite. cigarette use -0.51, smoke exposure -0.46, mental health composite -0.41, sleep problems -0.35, caffeine -0.29, illicit drug use -0.26, alcohol consumption -0.18, BMI 0.01, vitamin usage 0.23.

1B. Genetic Diversity amongst BiB participants

Figure A1 shows the first two PCs with color indicating self-reported ancestral background. Our main analytic sample consisted of those respondents who both (a) self-reported British ancestry

and (b) had their first two PCs fall within the red box (i.e., these are the red dots in the red box). A similar procedure was used to identify Pakistani respondents (i.e., blue dots in the blue box).

1C. Sample Comparisons

Table A2A compares the full BiB sample to the self-reported white British sample and our analytic sample (which is comprised of respondents with genetic data that we included based on the rules described above in 1B). Our analytic sample clearly differs from the full sample, which is expected given that it is an ancestrally homogenous subsample pulled from a diverse cohort. Differences between the two samples are potentially reflective of cultural differences that may fall along ancestral lines (i.e., the analytic sample is not uniformly "healthier" for example). For example, mothers in the analytic sample live in neighborhoods of lower disadvantage but are more likely to smoke than mothers in the full BiB sample. Children in the analytic sample have higher levels of both development and academic performance compared to those in the full BiB sample.

Focusing just on a comparison of the analytic sample (n=2077) to the full sample of respondents who self-report British ancestry (n=2210), our analytic sample is similar in terms of the child outcomes and maternal characteristics.

1D. Further description of Analytic Sample

Descriptive statistics are available in Table A2. Histograms for key variables in analytic sample are shown in Figure A2. Correlations amongst all variables are given in Figure A3.

As can be seen in Table A2, there are different levels of missingness in our variables. We further discuss missingness on three crucial variables in our analytic sample: child PGS, EYFSP, and KS1.

- In the analytic sample, 91 children did not have genetic data from which to compute a polygenic score. We did not observe a significant difference in the means of the maternal polygenic scores comparing those mothers whose children do have genetic data to those who do not.
- We did not observe EYFSP scores for 387 students. Missingness on the EYFSP is driven largely by students moving outside of the Bradford area. We did not observe a significant difference in the means of the maternal polygenic scores comparing those mothers whose children do EYFSP scores as compared to those who do not.
- We did not observe KS1 scores for 742 students. Missingness on the KS1 is due to both moving (as with the EYFSP) and also to the fact many BiB children were not yet old enough to have yet taken the KS1. Of the 742 students missing the KS1, nearly half (n=306, 41.2%) started school in 2015/2016 and thus would not have yet taken the KS1 by the time relevant data collection was complete.

Additional comparisons of the analytic sample to our minimum complete data sample are shown in Table A2 Panel B. There are a few differences (e.g., the full analytic sample had a slightly higher mean EYFSP than those in the minimum data sample), but the two groups are comparable across many dimensions (e.g., similar profiles of maternal education).

1E. Power Analysis

We conducted a sensitivity power analysis meant to determine our statistical power, given the BiB sample size, to detect associations of different size. In particular, we analyzed our power to detect associations in the presence of a covariate with known correlation structure (i.e., can we detect associations between the maternal PGS and some outcome given the fact that we also control for the child PGS which is correlated with the mother PGS at roughly 0.5). Figure A4 shows results of this analysis. Even for our analyses of KS1 in Table 2 (with N=1267), we have reasonable power (>0.8) to detect associations starting if effect sizes are larger than roughly 0.06.

2. Ancillary Analyses on British respondents

2A. Associations with outcomes measured at birth

In Table A3, we examined associations between polygenic scores and outcomes measured at birth in parallel to Table 2 of main text. We considered gestational age, APGAR scores, birthweight in grams and a low birth indicator (additional information on measures in Table A1); results were null.

2B. Associations between prenatal exposures and child development and academic performance

Our composite measures of prenatal exposures are highly associated with child development and academic performance (Table A4). The SES composite has estimated associations of around 0.28 with both the EYFSP and the KS1. The health composite has estimated associations with both measures of around 0.15.

2C. Mediation via individual environmental pathways

Table A5 replicates results along the lines of Table 3 from the main text including each environmental variable separately. Maternal education is an especially important mediator. Note that maternal education has, as expected, strong associations with the maternal PGS (Figure A5). However, we do note that the direct effect of maternal PGS on offspring development (b=0.082, 95% CI=[0.025, 0.140], p<0.005) remained highly significant in our mediation model. For academic performance, associations were weaker (b=0.055, 95% CI=[-0.003, 0.120], p=0.074).

3. Analysis of a Pakistani ancestry subsample of the BiB cohort

We considered analysis of the genetically identified respondents of Pakistani ancestry (e.g., the blue dots in the blue box in Figure A1). We first computed health and SES prenatal composites in the same manner as before. We then looked at associations between the maternal education PGS and the prenatal composites net of the first 10 PCs computed in the entire genetic sample. Results were null. The PGS for the mothers in the Pakistani sample was not robustly predictive of either the health or SES prenatal composite (see Table A6). We then looked at associations between the maternal education PGS and the child outcomes net of both the PCs and the child PGS. Results were again null. The maternal PGS was not associated with either child development or academic performance.

4. Syntax

In the interest of reproducibility, we provide the syntax for all of our statistical analyses. Note that we prepared data using Stata Version 14 (StataCorp, 2015) and analyzed data using R (Version 3.5.2).

[link removed for submission to maintain anonymity; included in letter to editor]

Table A1. Description of study variables.

Measures	Description
Prenatal health	
Body Mass Index (BMI)	BMI was directly assessed at the hospital by nursing staff during upon study enrollment
Mental Health	General Health Questionnaire (GHQ; Goldberg & Hillier, 1979) is 28-item scaled questionnaire that assessed mother's somatic symptoms, anxiety and insomnia, social dysfunction and depression. A General Health factor score was the mean of all items and standardized.
Cigarette use	A single item "do you smoke cigarettes?" indexed mother's cigarette use. Mothers responded yes or no.
Indirect Smoke Exposure	A single item index: "During pregnancy have you been exposed to other peoples' cigarette smoke at work or at home?" Mothers responded yes or no.
Alcohol use	A single item: "Have you drunk alcohol in the past 3 months? Mothers responded yes or no
Caffeine consumption.	Mothers responded to 28 items of decaffeinated and caffeinated
Drug use	A single item index: "Have you used any drugs like marijuana or ecstasy during pregnancy or in the three months before pregnancy?" Mothers indicated yes or no.
Vitamin use	A single item index: "Have you taken any dietary supplements including vitamins or iron tablets in the last 4 weeks of pregnancy?" Mothers indicated yes or no.
Sleep problems	Sleep problems were assessed with two items: "have you Lost much sleep over worry?" and "have you had difficulty staying asleep once you are off?". Mothers responded on a Likert-type scale from 1 (not at all) to 4 (much more than usual).
Prenatal socio-economic conditions	
Single	A single item index: Mothers responded whether they were currently living with the baby's father, living with another partner, or not living with another partner. If mothers lived with a partner of any kind, they were coded as 0 = not single. If they did not live with a partner, they were coded as 1 = single (N_{single} = 505, $N_{partnered}$ = 1486).
Employment	A single item index: "are you currently employed?". Mothers responded yes or no
Maternal leave	A single item index: "Are you currently on maternity/sick leave?". Mothers indicated yes or no
Neighborhood deprivation	The Index of Multiple Deprivation (IMD) is an official measure of neighborhood affluence in England based on income, employment, health and disability, education, barriers to housing and services, crime, and living environment. IMD rankings within Bradford only were included in order to illuminate the full variation among Bradford neighborhoods. There were ten neighborhoods ranked from one to ten in Bradford, with 10 indicating relatively more deprived neighborhoods.
Governmental benefits	Checklist of ten governmental benefits, indicating which ones they received and their partner received (e.g., child tax credit, income support, disability living allowance). Mothers responded yes or no to each item. Principle components analysis was used to create a composite score with higher scores indicating more governmental benefits received
Perceived financial difficulty	A single item: "How well are you and your partner managing financially?" Mothers responded on a 5-item response set ranging from "living comfortably" to "finding it very difficult".
Characteristics at birth of child	
APGAR score	APGAR scores at birth were determined by a hospital nurse. Two scores were provided for each child: one within the first minute of life, and the other within the first five minutes of life. We calculated an average score.
Gestational age	Child's gestational age was obtained from medical records.
Gestational weight	Birth weight was directly assessed by hospital staff and was recorded in grams.

Small for gestational age	Small for gestational age was coded yes/no: Yes if birthweight is below 10th percentile on UK WHO fetal growth charts for sex and gestational week at birth. This measure was only calculated for singletons.
Large for gestational age	Large for gestational age was coded yes/no: Yes if birthweight was above 90th percentile on UK WHO fetal growth charts for sex and gestational week at birth. This measure was only calculated for singletons.
Child outcomes	
Child development	We used children's scores on the Early Years Foundation Stage Profile (EYFSP), a teacher-led observational assessment conducted towards the end of the child's first year at school. The version of this assessment analysed was used from the start of the 2012/2013 academic year onwards in English schools, and is completed at the end of the child's first year in school, when they are usually 4-5 years old. The profile measures children's attainment in seven main areas of learning: communication and language; expressive arts and design; literacy, mathematics; physical development; personal, socio and emotional development and understanding the world. The assessment is designed not as an academic test but to assess children's development in these areas, compared to the average child at the end of one year's schooling. Teachers completed the assessment for each child based on their knowledge and observations of that child. The measure is intended to provide a complete picture of children's development, not just a snapshot of what happens at school. We standardized each subscale and calculated a mean composite score for child development, with higher scores indicating relatively greater development.
Academic performance	We used children's scores on the Key Stage 1 Assessment, a standardized test conducted under exam conditions and set by the Standards and Testing Agency in England. This version was used from the start of the 2015/2016 academic year onwards, and is completed towards the end child's third year in school when the child was 6-7 years old. The Key Stage 1 Assessment includes math, reading and science subscales. For math and reading, children were graded on a five-point scale: level 1, just into level 2, securely at level 2, top end of level 2, and level 3. For science, children were graded on a three-point scale: levels 1, 2 and 3. We standardized each subscale and calculated a mean composite score for academic performance, with higher scores indicating relatively better performance.

Table A2. Sample Comparisons

A. Comparison between full BiB cohort, the white British sample, and our analytic sample (dyads for whom genetic and test data were available and European ancestry only).

	Full Sample (N=6124)			Self-reported White British (N=2210)			Analytic sample (genotyped respondents of British ancestry, N=2077)			p-value of test of difference in means between full British sample (N=2210) and analytic sample (N=2077)
	Mean	SD	N	Mean	SD	N	Mean	SD	N	
Child Characteristics										
Child Development (EYSFP)	0	1	5056	0.184	0.955	1800	0.18	0.955	1690	9.10E-01
Academic Performance (Key Stage 1)	0	1	4023	0.029	0.967	1424	0.021	0.971	1335	8.34E-01
Maternal Characteristics										
Mom's Education			4087			2014			1892	
GCSE 1	0.285	0.451		0.216	0.412		0.22	0.414		8.50E-01
GCSE 2	0.316	0.465		0.382	0.486		0.379	0.485		8.80E-01
GCSE 3	0.185	0.389		0.2	0.4		0.198	0.398		9.24E-01
University Degree	0.215	0.411		0.202	0.402		0.203	0.403		9.45E-01
Maternal Age	27.499	5.608	5405	26.844	6.121	2210	26.856	6.102	2077	8.77E-01
BMI	26.344	5.732	5163	27.131	6.031	2115	27.139	6.029	1987	9.11E-01
Mental Health	0.082	0.815	5042	0.033	0.798	2102	0.042	0.801	1989	7.64E-01
Vitamin use	0.411	0.492	5392	0.295	0.456	2209	0.295	0.456	2077	9.99E-01
Indirect Smoke Exposure	0.316	0.465	5382	0.423	0.494	2207	0.425	0.494	2075	9.49E-01
Cigarette Use	0.153	0.36	6124	0.338	0.473	2210	0.339	0.474	2077	9.29E-01
Alcohol Consumption	0.175	0.38	6124	0.423	0.494	2210	0.428	0.495	2077	8.02E-01
Caffeine Consumption (mg)	61.403	99.138	4598	90.323	130.255	1929	89.892	130.742	1813	2.50E-01
Drug Use	0.011	0.104	5211	0.02	0.141	2118	0.021	0.145	2002	9.21E-01
Single	0.162	0.368	5397	0.268	0.443	2208	0.268	0.443	2076	9.89E-01
Employed	0.396	0.489	6124	0.644	0.479	2210	0.644	0.479	2077	9.88E-01
Maternal Leave	0.074	0.262	3213	0.055	0.229	1736	0.058	0.233	1632	8.90E-01

Neighborhood Deprivation	7.119	2.369	5332	6.152	2.519	2154	6.153	2.515	2023	9.92E-01
Sleep Problems	0	1.217	5037	0.075	1.167	2099	0.085	1.171	1986	7.67E-01
Financial Difficulties	2.124	0.934	5385	2.134	0.932	2207	2.13	0.934	2074	9.06E-01
Receipt of Governmental Benefits	0	1.368	5394	-0.085	1.412	2207	-0.07	1.415	2075	6.83E-01

Table 2B. Comparison of analytic sample to minimal complete data sample.

	Analytic sample (genotyped respondents of British ancestry, N=2077)			Analytic sample with both child PGS and KS1 (n=1267)			p-value of test of difference in means between analytic sample (N=2077) and complete data sample (n=1267)
	Mean	SD	N	Mean	SD	N	
Child Characteristics							
Child Development (EYSFP)	0.18	0.955	1690	0.088	0.927	1257	1.09E-02
Academic Performance (Key Stage 1)	0.021	0.971	1335	0.012	0.974	1267	8.09E-01
Maternal Characteristics							
Maternal Education			1892			1145	
GCSE 1	0.22	0.414		0.236	0.425		5.12E-01
GCSE 2	0.379	0.485		0.396	0.489		5.24E-01
GCSE 3	0.198	0.398		0.188	0.391		6.73E-01
University Degree	0.203	0.403		0.181	0.385		3.33E-01
Maternal Age	26.856	6.102	2077	26.721	6.11	1267	1.24E-01
BMI	27.139	6.029	1987	27.358	6.12	1218	1.49E-02
Mental Health	0.042	0.801	1989	0.039	0.794	1211	9.44E-01
Vitamin use	0.295	0.456	2077	0.272	0.445	1267	3.39E-01
Indirect Smoke Exposure	0.425	0.494	2075	0.439	0.496	1267	5.70E-01
Cigarette Use	0.339	0.474	2077	0.364	0.481	1267	3.22E-01
Alcohol Consumption	0.428	0.495	2077	0.424	0.494	1267	8.67E-01
Caffeine Consumption (mg)	89.892	130.742	1813	96.112	135.13	1075	1.95E-44
Drug Use	0.021	0.145	2002	0.023	0.15	1218	9.14E-01
Single	0.268	0.443	2076	0.292	0.455	1266	3.07E-01
Employed	0.644	0.479	2077	0.641	0.48	1267	8.93E-01
Maternal Leave	0.058	0.233	1632	0.055	0.228	1033	8.99E-01

Neighborhood Deprivation	6.153	2.515	2023	6.083	2.523	1262	2.22E-01
Sleep Problems	0.085	1.171	1986	0.079	1.158	1208	8.82E-01
Financial Difficulties	2.13	0.934	2074	2.148	0.92	1266	6.09E-01
Receipt of Governmental Benefits	-0.07	1.415	2075	0.029	1.445	1265	1.98E-02

Table A3. Estimated associations between maternal PGS and child PGS with birth characteristics (controlling for 10 PCs and maternal age)

	Maternal PGS		Child PGS		
	Estimate	95% CI	Estimate	95% CI	N
Gestational Age (days)	0.024	-0.026, 0.075	0.002	-0.047, 0.051	1985
APGAR Score	0.016	-0.035, 0.067	-0.036	-0.087, 0.015	1975
Birthweight (g)	0.044	-0.008, 0.096	0.042	-0.007, 0.091	1984
Small for gestational age	-0.066	-0.116, -0.015	0.000	-0.046, 0.045	1951

Table A4. Associations between prenatal composites & child academic and developmental outcomes (net of 10 PCs and maternal age).

Outcome	Predictor	Estimate	95% CI
EYFSP	SES Composite	0.290	0.241, 0.339
	Health Composite	0.159	0.111, 0.207
KS1	SES Composite	0.279	0.223, 0.335
	Health Composite	0.143	0.088, 0.199

Table A5. Mediation Analysis: Proportion of maternal PGS association on outcome (EYFSP or KS1) mediated by individual environmental measures.

A. Child development (EYFSP)

Mediator	Total Effect (maternal PGS on outcome)	95% CI	Proportion mediated	95% CI	N
BMI	0.118	0.060, 0.171	0.013	-0.018, 0.066	1539
Mental Health	0.110	0.054, 0.165	0.001	-0.014, 0.024	1549
Indirect Smoke Exposure	0.115	0.061, 0.168	0.078	0.008, 0.185	1610
Cigarette use	0.115	0.061, 0.177	0.099	0.033, 0.219	1611
Alcohol Consumption	0.115	0.060, 0.173	0.000	-0.028, 0.026	1611
Caffeine Consumption	0.115	0.053, 0.172	0.060	0.011, 0.179	1408
Drug use	0.112	0.054, 0.169	0.000	-0.022, 0.027	1562
Vitamin use	0.115	0.061, 0.171	0.031	-0.020, 0.097	1611
Sleep Problems	0.110	0.054, 0.166	0.006	-0.025, 0.051	1546
Maternal Education	0.121	0.063, 0.176	0.319	0.188, 0.627	1470
Single	0.114	0.057, 0.168	0.029	-0.003, 0.095	1610
Employed	0.115	0.058, 0.173	0.092	0.014, 0.204	1611
Maternal Leave	0.125	0.056, 0.184	0.000	-0.017, 0.025	1258
Subjective Financial Difficulty	0.114	0.056, 0.171	0.004	-0.050, 0.053	1610
Neighborhood Deprivation	0.114	0.062, 0.167	0.078	-0.012, 0.212	1602
Receipt of Governmental Benefits	0.114	0.059, 0.171	0.139	0.050, 0.307	1609

B. Academic Performance (KS1)

Mediator	Total Effect (maternal PGS on outcome)	95% CI	Proportion mediated	95% CI	N
BMI	0.096	0.032, 0.162	0.010	-0.044, 0.087	1218
Mental Health	0.083	0.020, 0.147	0.008	-0.051, 0.084	1211
Indirect Smoke Exposure	0.087	0.021, 0.148	0.066	-0.003, 0.283	1267
Cigarette use	0.087	0.026, 0.151	0.113	0.035, 0.405	1267
Alcohol Consumption	0.088	0.025, 0.152	0.008	-0.020, 0.083	1267
Caffeine Consumption	0.115	0.051, 0.183	0.049	-0.002, 0.169	1075
Drug use	0.084	0.021, 0.146	0.001	-0.058, 0.045	1218
Vitamin use	0.088	0.026, 0.147	0.015	-0.018, 0.095	1267
Sleep Problems	0.080	0.020, 0.143	0.015	-0.039, 0.112	1208
Maternal Education	0.097	0.037, 0.162	0.432	0.224, 1.053	1145
Single	0.087	0.023, 0.148	0.012	-0.042, 0.112	1266
Employed	0.087	0.027, 0.146	0.111	-0.023, 0.398	1267
Maternal Leave	0.081	0.012, 0.150	0.000	-0.040, 0.065	1033
Subjective Financial Difficulty	0.087	0.028, 0.148	-0.009	-0.117, 0.069	1266
Neighborhood Deprivation	0.089	0.030, 0.152	0.089	-0.003, 0.323	1262
Receipt of Governmental Benefits	0.087	0.025, 0.144	0.174	0.031, 0.532	1265

Table A6. Association estimates in East Asian subsample.

Outcome	Controls	Association	95% CI	N
SES Composite	10 PCs & Age	0.050	0.008, 0.091	2196
Health Composite	10 PCs & Age	-0.035	-0.078, 0.007	2196
EYFSP	10 PCs, Age, and Child PGS	0.007	-0.049, 0.063	1852
KS1	10 PCs, Age, and Child PGS	0.023	-0.039, 0.085	1473

Figure A1. Principle components plots for polygenic scores in genetic sample.

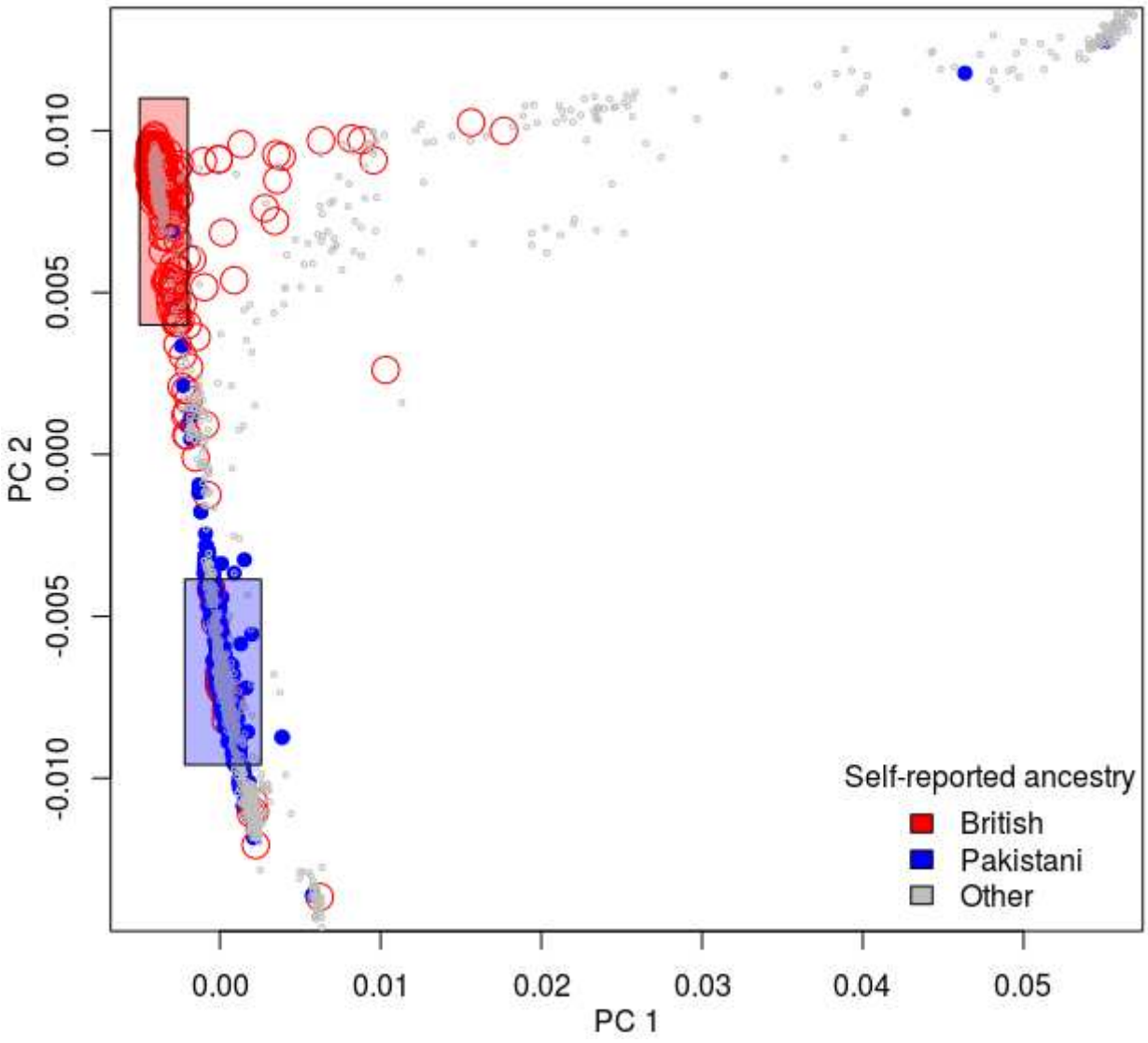

Figure A2. Histograms of (top row) children's development and academic performance and (bottom row) prenatal health and SES composites.

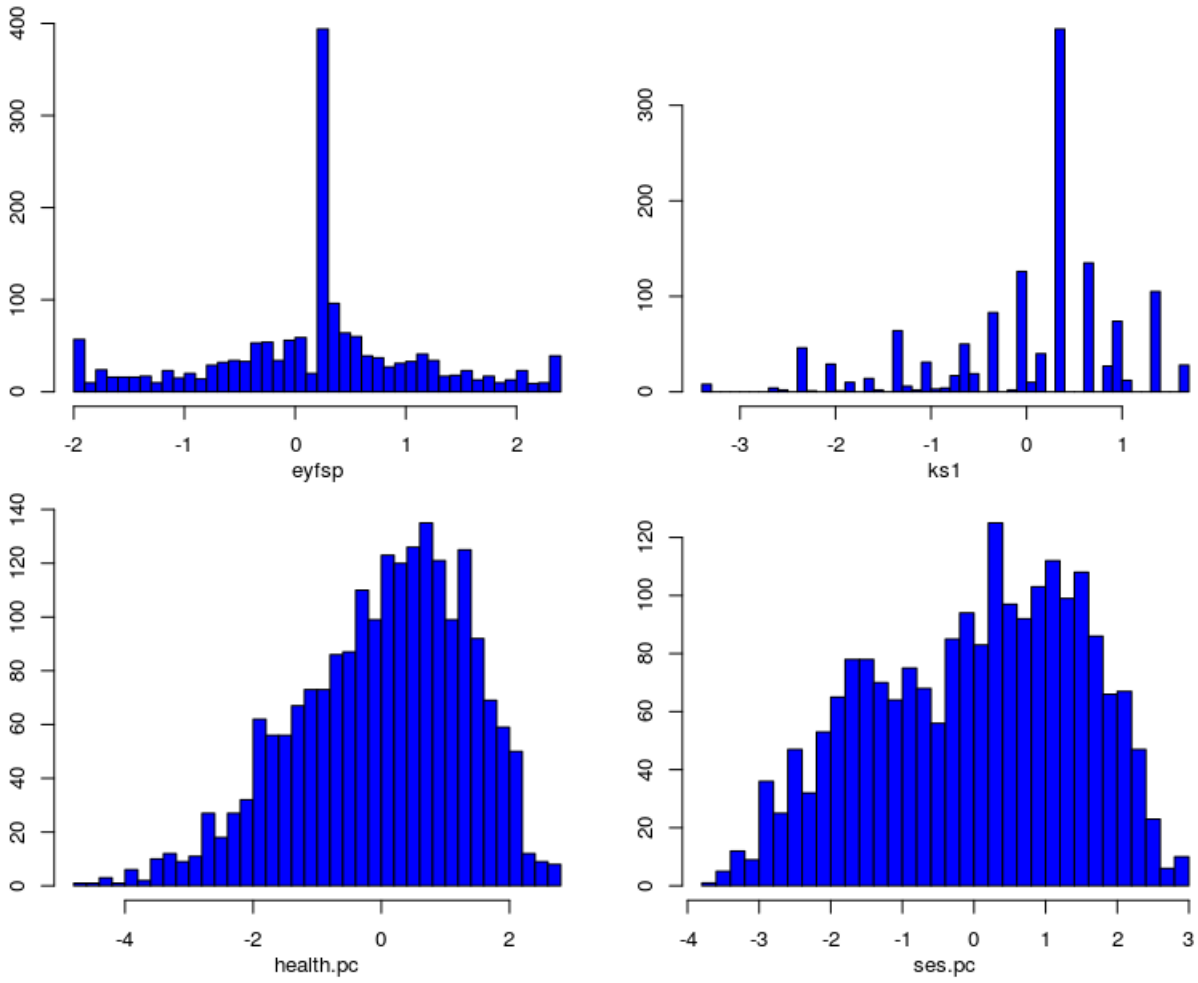

Figure A3. Correlations between key study variables.

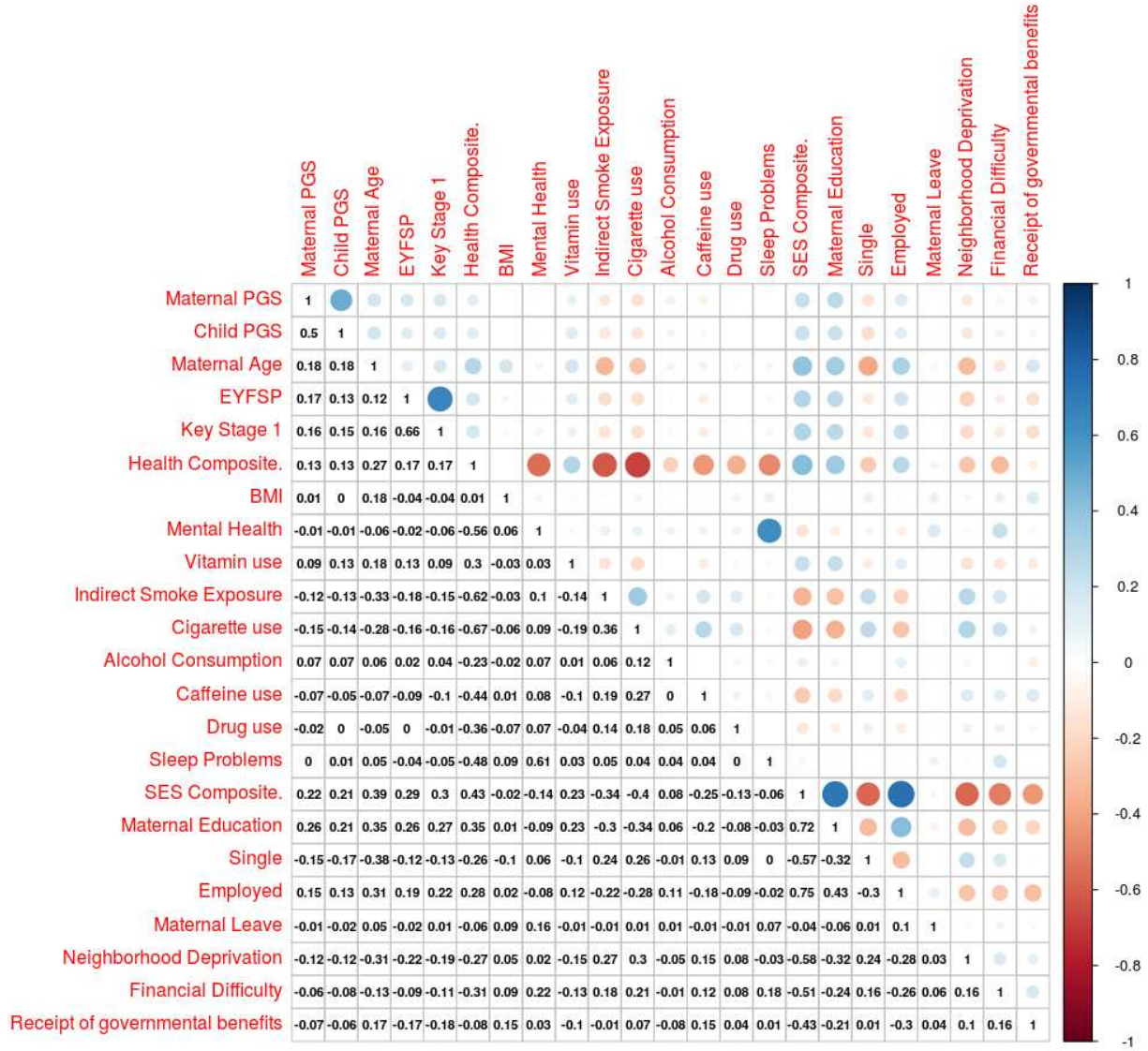

Figure A4. Power Analysis

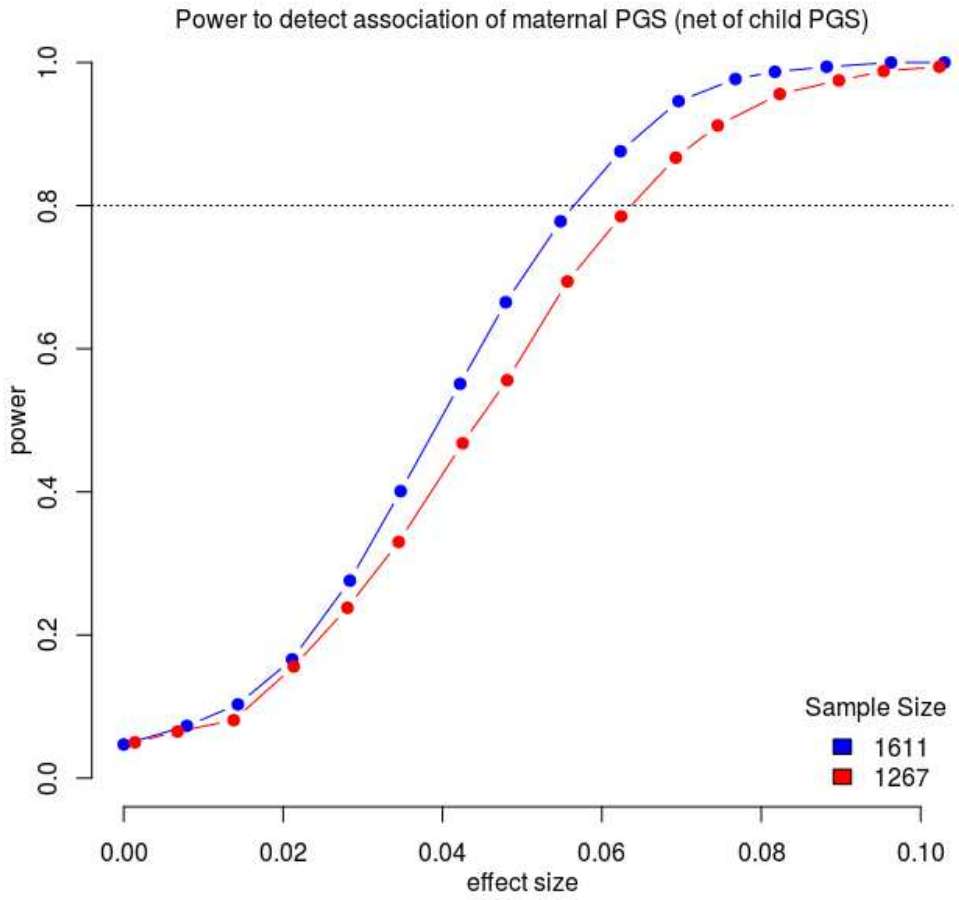

Figure A5. Mean polygenic scores as a function of level of maternal education.