PRS OF SCHIZOPHRENIA AND WHERE YOU LIVE

Replication in the UKB sample

19/03/2018

baptiste.couvyduchesne@gmail.com

DATA

POSTCODE IN THE UK			
Outward Code		Inward Code	
Postcode Area	Postcode District	Postcode Sector	Postcode Unit
SW	1W	0	NY

Considering the precision of the participants localisation (1km), we can only infer safely their postcode district: e.g. **SW1W**

Population density per **postcodes district** from the 2011 census

For England and Wales

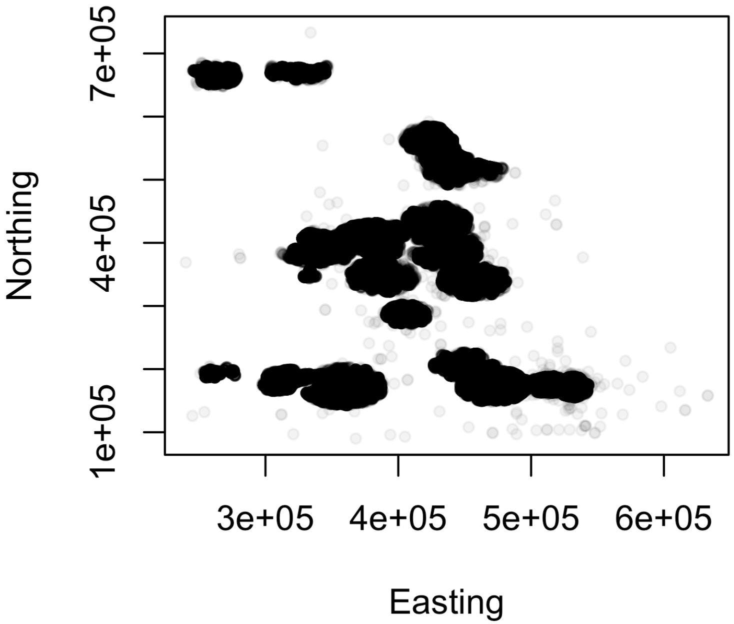
https://www.nomisweb.co.uk/census/2011/key_statistics

For Scotland

http://www.scotlandscensus.gov.uk/ods-analyser/

REVERSE GEO CODING

The biobank provides Easting and Northing coordinates (rounded per km) for each participants' current address. This set of coordinates is known as OSNG (Ordnance Survey) and may be easily transformed into longitude and latitude if needed.



=> Transformation into latitude-longitude

We recognise the UK a bit more

REVERSE GEO CODING – IN PRACTICE

Finding an address from coordinates is called reverse geo-coding and the best at doing that is google (maps).

There are R packages that allow to do this (*ggmap*), though they are limited to 2,500 searches per day. Above that google is making you pay.

https://www.doogal.co.uk/BatchReverseGeocoding.php
I found this brilliant website that gets around this problem (still uses google). You can export the results into a csv file.

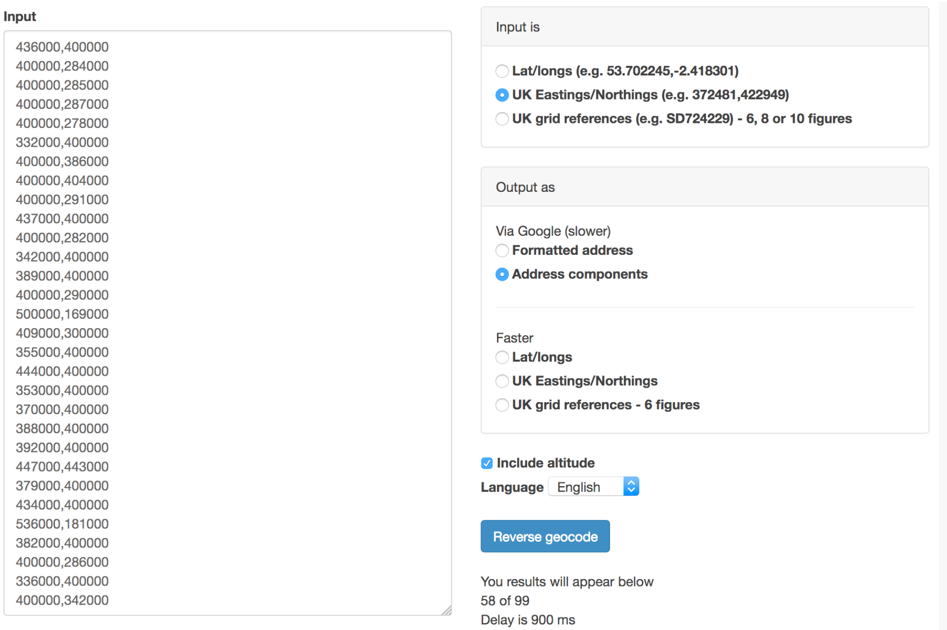
I exported the 15,000 east-north combinations present in the Biobank in one afternoon

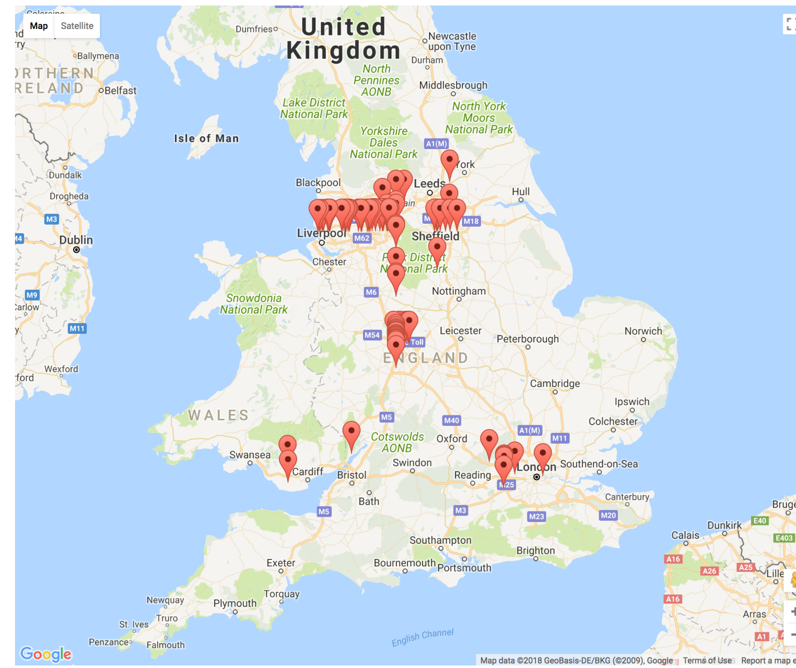
I checked that the results align with those from *ggmap* for a random subset of coordinates (they do)

REVERSE GEO CODING – IN PRACTICE

Example for 100 coordinates from the UKB

The website shows you the locations on a map and you can dowload the csv file, with postcode, altitude, region etc…





Easting,Northing,Number,Street,Locality,Postal town,Admin level 3,Admin level 2,Admin level 1,Postcode,Country,Altitude
436000,400000,29,Parkside Road,Hoyland,Barnsley,,South Yorkshire,England,S74 0AL,United Kingdom,139
400000,284000,25,Rutters Meadow,Quinton,Birmingham,,West Midlands,England,B32 1SH,United Kingdom,199
400000,285000,647,Hagley Road West,Quinton,Birmingham,,West Midlands,England,B32 1BY,United Kingdom,205

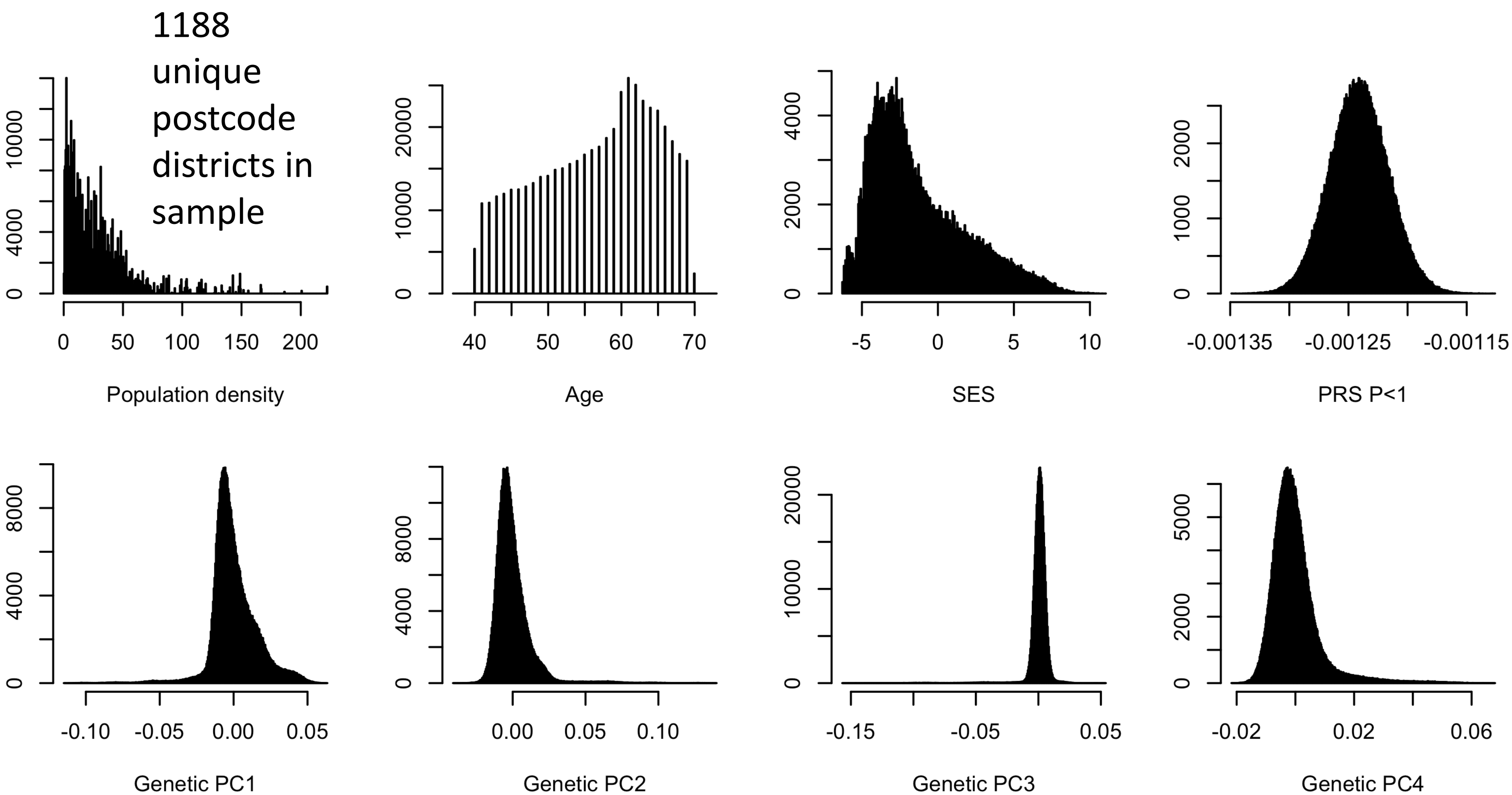
OTHER VARIABLES

PRS: calculated by **Maciej Trzaskowski** from the Ripke, 2014 paper
We use the "P>1" threshold as per our analysis
The calculation was restricted to participants of European ancestry (N=345,258)

Genetic PCs: at the moment I am using in sample PCs calculated using FastPCA (**Kathryn Kemper, Zhili Zheng**)
$\Rightarrow$ The first PC already represent within UK population structure

SES: Townsend deprivation index (provided by the UKB)
"Based on the preceding national census output areas. Each participant is assigned a score corresponding to the output area in which their postcode is located". A higher score indicate greater levels of deprivation.

VARIABLES DESCRIPTION



RESULTS FOR THE FULL SAMPLE (N= 342608) – WITHOUT SES

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00123	2.76E-06	-446	0	NA	NA
factor(sex)1	-1.22E-05	4.04E-06	-3.01	0.002594688	NA	NA
popDensityPostCode	**1.64E-08**	**1.59E-09**	**10.3**	**7.23E-25**	**0.0177**	**0.000313**
age	-3.97E-07	1.01E-07	-3.92	8.78E-05	-0.13	0.0169
age2	3.62E-09	9.13E-10	3.97	7.28E-05	0.131	0.0172
ageSex	4.36E-07	1.48E-07	2.95	0.00320311	0.508	0.258
age2Sex	-3.85E-09	1.33E-09	-2.89	0.003866339	-0.273	0.0745
PC1	0.000147	2.62E-06	56.3	0	0.0916	0.00839
PC2	0.000398	3.19E-06	125	0	0.204	0.0416
PC3	-0.000101	4.06E-06	-24.9	2.01E-136	-0.0378	0.00143
PC4	9.05E-06	4.75E-06	1.9	0.057002127	0.00308	9.49E-06

PCs 5-20 not shown

RESULTS FOR THE FULL SAMPLE (N= 342608) – WITH SES

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00123	2.76E-06	-446	0	NA	NA
factor(sex)1	-1.19E-05	4.04E-06	-2.95	0.003169852	NA	NA
popDensityPostCode	**6.31E-09**	**1.78E-09**	**3.55**	**0.000384554**	**0.00682**	**4.65E-05**
age	-3.78E-07	1.01E-07	-3.73	0.000187826	-0.124	0.0154
age2	3.48E-09	9.14E-10	3.81	0.00014146	0.126	0.0159
ageSex	4.25E-07	1.48E-07	2.87	0.004066736	0.496	0.246
age2Sex	-3.74E-09	1.33E-09	-2.81	0.00501404	-0.265	0.0702
ses	1.96E-07	1.54E-08	12.7	4.69E-37	0.0245	6.00E-04
PC1	0.000145	2.62E-06	55.4	0	0.0903	0.00815
PC2	0.000396	3.20E-06	124	0	0.203	0.0412
PC3	-9.97E-05	4.06E-06	-24.5	6.18E-133	-0.0374	0.0014
PC4	1.07E-05	4.76E-06	2.26	0.024111484	0.00365	1.33E-05

PC5-20 not shown

PRS ANALYSES IN ADMIXED POPULATION

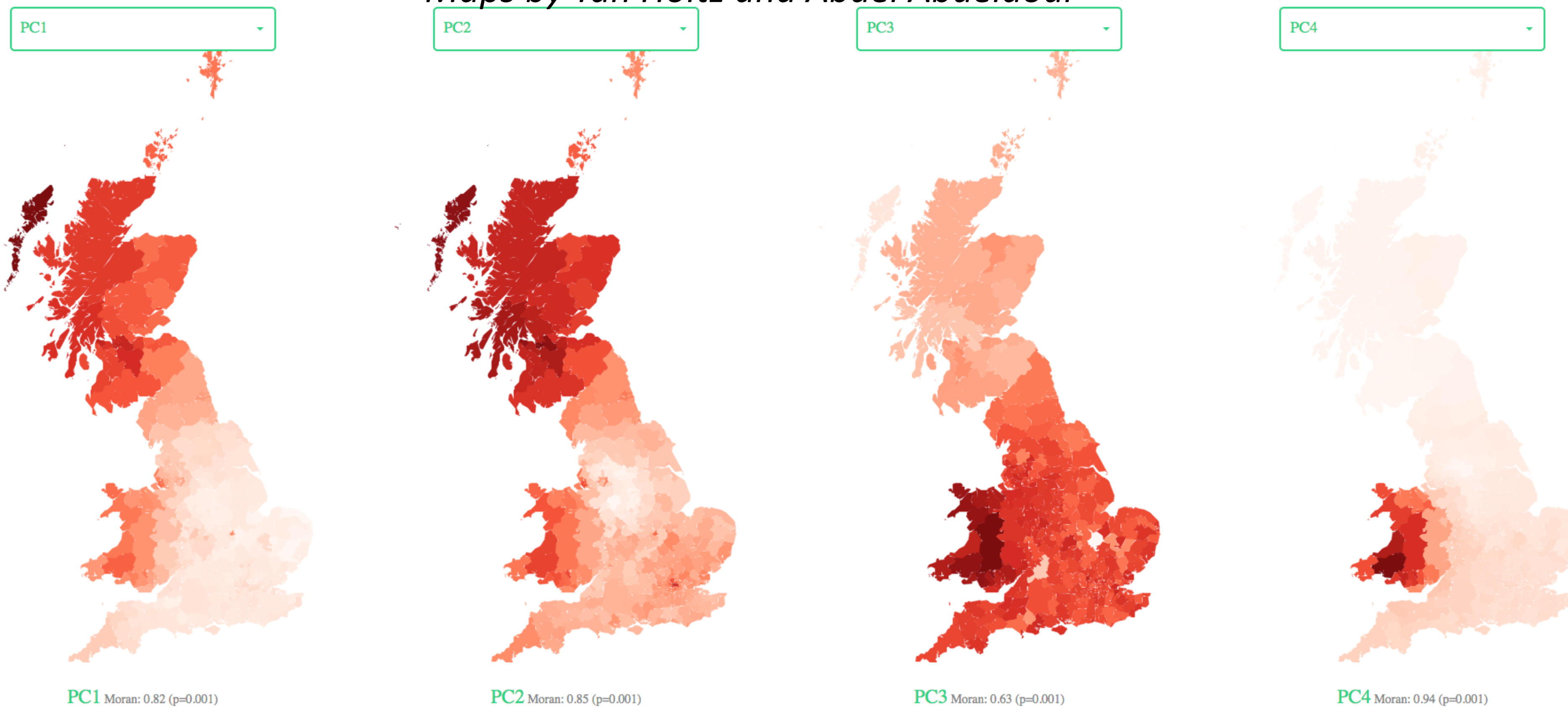
Rationale: PRS accuracy may differ by ancestry groups (e.g. differences in variance or mean), which may create spurious correlations

Q: is the association between PRS and population density confounded by differences in predictiveness of the scores across the different countries?

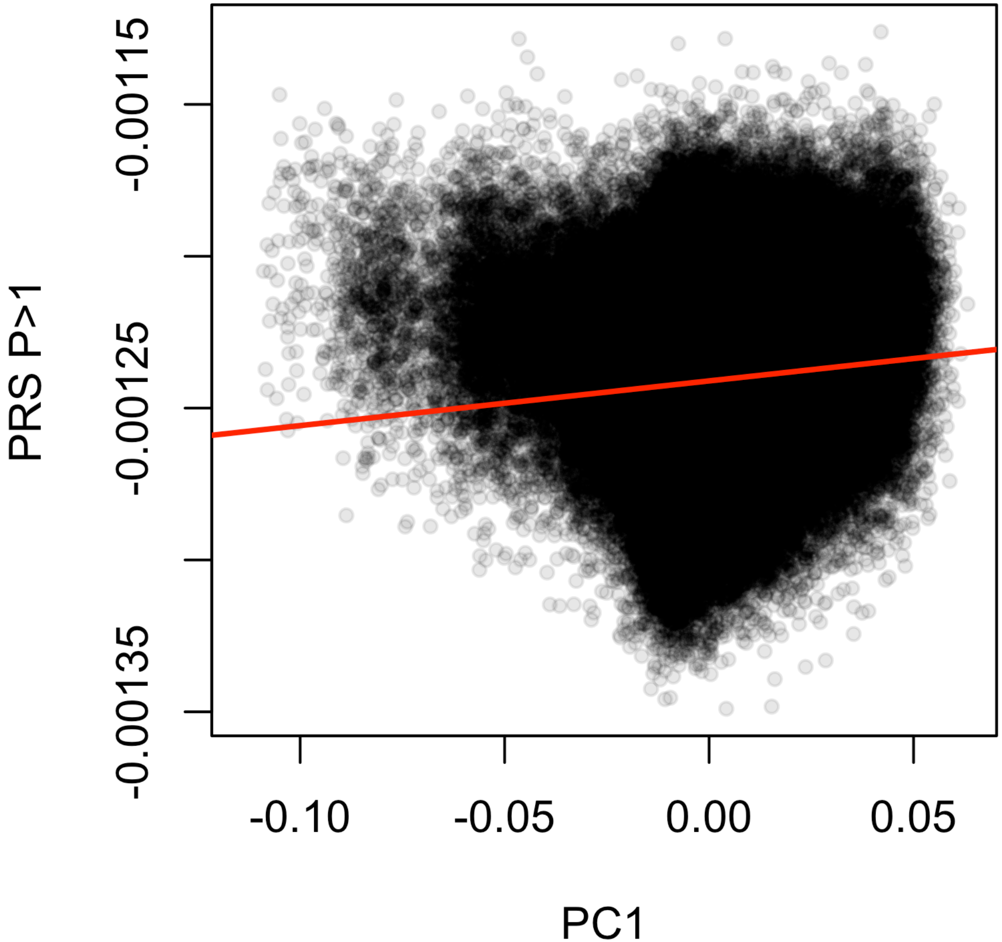
PRS ANALYSES IN ADMIXED POPULATION

Here, we see that some of the PC, clearly correlate with the UK countries
PC1: East – West ; PC3: North-South ; PC4: Wales – the rest of the UK

Maps by Yan Holtz and Abdel Abdelaoui



EXAMPLE OF PRS BIAS BY ANCESTRY

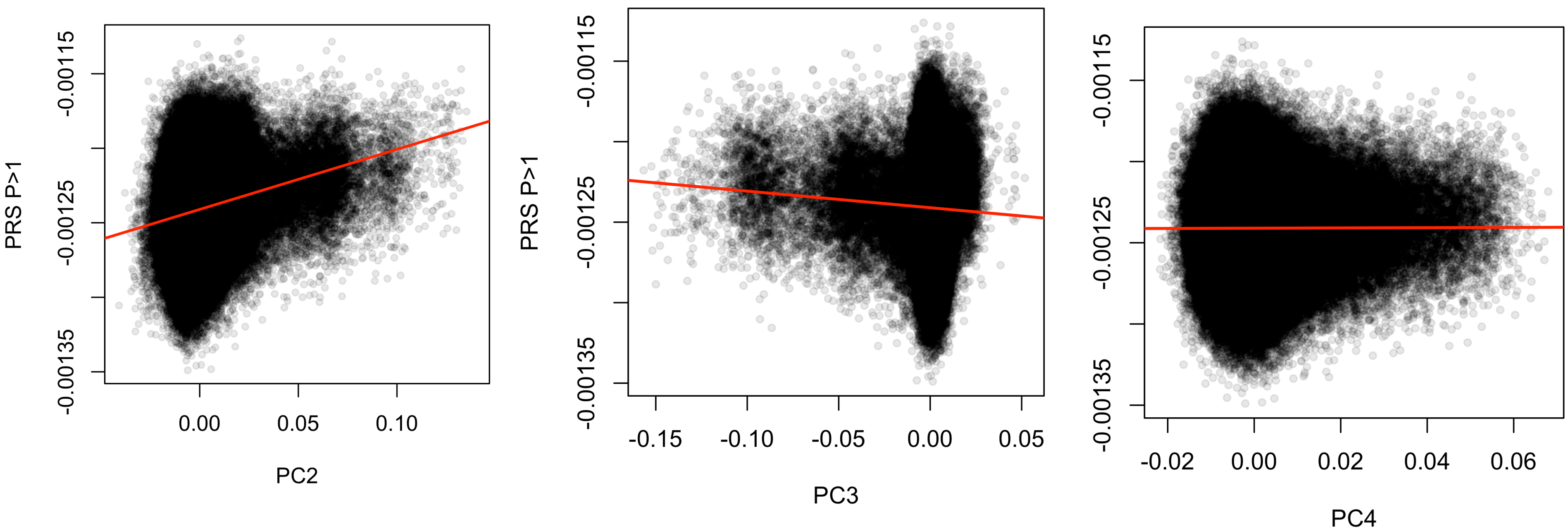


We see that the variance of the PRS (y axis) is not constant as a function of the PC

This suggests differences in prediction accuracy or PRS dispersion between ancestry

Potentially problematic => the mean PRS differences explained by PCs may be due to differences in accuracy (variance) and not to true mean differences

MORE EXAMPLE OF PRS BIAS BY ANCESTRY



For PC4 there is no mean differences so the difference in variance is less likely to confound the analysis

FROM OUR RESULTS

In our models we are fitting 20 PCs, hence we may account for some of the variance differences across ancestry

We can test if the residuals of our models are heteroscedastic (non constant variance) using studentized Breusch-Pagan test (one of the many heteroscedasticity test)

studentized Breusch-Pagan testdata: glm("prs49.s10 ~ factor(sex) + popDensityPostCode + age + age2 + ageSex + age2Sex + ***ses*** *+ PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC7 + PC8 + PC9 + PC10 + PC11 + PC12 + PC13 + PC14 + PC15 + PC16 + PC17 + PC18 + PC19 + PC20", family = gaussian(link = "identity"), data = dat)*

BP = 54.214, df = 27, p-value = 0.001432 Evidence of heteroscedasticity

studentized Breusch-Pagan testdata: glm("prs49.s10 ~ factor(sex) + popDensityPostCode + age + age2 + ageSex + age2Sex + PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC7 + PC8 + PC9 + PC10 + PC11 + PC12 + PC13 + PC14 + PC15 + PC16 + PC17 + PC18 + PC19 + PC20", family = gaussian(link = "identity"), data = dat)

BP = 40.214, df = 26, p-value = 0.03717 Evidence of heteroscedasticity

HOW TO GET AROUND THIS?

We need to break down the sample into more homogeneous ancestry group or account for the sources of heteroscedasticity present in the sample

a) Include more PCs
b) Fit a GRM
c) Exclude ancestry-PC outliers (admixed participants)
d) Break down by country

A) RESULTS FOR THE FULL SAMPLE – WITH SES & 40 PCS

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00123	2.76E-06	-446	0 NA	NA	
factor(sex)1	-1.19E-05	4.03E-06	-2.94	0.003241621 NA	NA	
popDensityPostCode	**6.51E-09**	**1.78E-09**	**3.66**	**0.000249153**	**0.00703**	**4.94E-05**
age	-3.77E-07	1.01E-07	-3.73	0.000195115	-0.123	0.0151
age2	3.47E-09	9.13E-10	3.8	0.000142014	0.126	0.0159
ageSex	4.24E-07	1.48E-07	2.87	0.004080042	0.495	0.245
age2Sex	-3.75E-09	1.33E-09	-2.81	0.004920618	-0.265	0.0702
ses	1.93E-07	1.54E-08	12.6	3.28E-36	0.0242	0.000586
PC1	0.000145	2.62E-06	55.4	0	0.0903	0.00815
PC2	0.000396	3.19E-06	124	0	0.203	0.0412
PC3	-9.97E-05	4.06E-06	-24.6	4.40E-133	-0.0374	0.0014
PC4	1.07E-05	4.75E-06	2.25	0.024319686	0.00365	1.33E-05

PC5-40 not shown

studentized Breusch-Pagan test: BP = 63.422, df = 47, p-value = 0.05524

A) RESULTS FOR THE FULL SAMPLE – NO SES & 40 PCS

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00123	2.76E-06	-446	0 NA	NA	
factor(sex)1	-1.21E-05	4.03E-06	-3.01	0.00264113 NA	NA	
popDensityPostCode	**1.65E-08**	**1.59E-09**	**10.3**	**4.24E-25**	**0.0178**	**0.000317**
age	-3.95E-07	1.01E-07	-3.91	9.19E-05	-0.129	0.0166
age2	3.62E-09	9.12E-10	3.97	7.34E-05	0.131	0.0172
ageSex	4.35E-07	1.48E-07	2.95	0.003200413	0.508	0.258
age2Sex	-3.86E-09	1.33E-09	-2.9	0.003778694	-0.273	0.0745
PC1	0.000147	2.62E-06	56.3	0	0.0916	0.00839
PC2	0.000398	3.19E-06	125	0	0.204	0.0416
PC3	-0.000101	4.05E-06	-24.9	1.60E-136	-0.0378	0.00143
PC4	9.05E-06	4.75E-06	1.91	0.056705493	0.00308	9.49E-06

PC5-40 not shown

studentized Breusch-Pagan test: BP = 50.045, df = 46, p-value = 0.316

B) RESULTS FOR THE FULL SAMPLE – GRM

With the current memory and time available we cannot run a REML estimation (and fixed effect estimation) in GCTA on the full sample

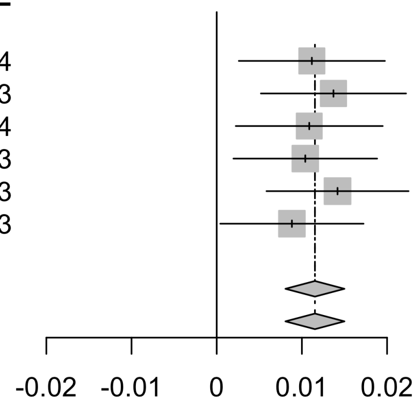
Solution: break down the sample (of unrelated individuals) into smaller chunks of 50,000 participants and meta-analyse the results.

$\Rightarrow$ Breaking down the GRM is memory efficient using the new GCTA
$\Rightarrow$ Model estimation is Fast ~3h
$\Rightarrow$ Requires <100GB memory

B) RESULTS FOR THE FULL SAMPLE – GRM

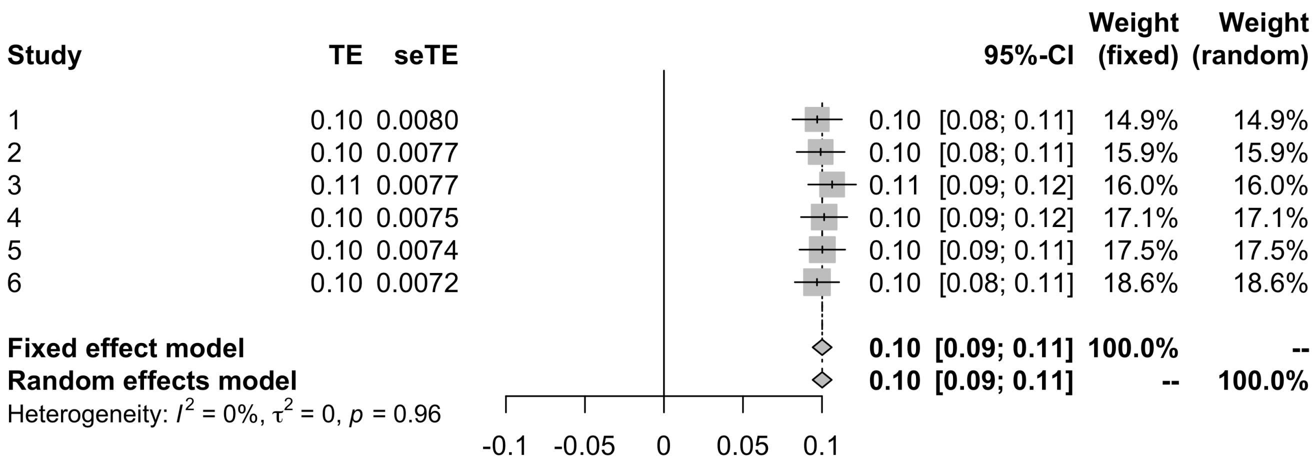
Study	TE	seTE		95%-CI	Weight (fixed)	Weight (random)
1	0.01	0.0044	0.01	[0.00; 0.02]	16.3%	16.3%
2	0.01	0.0043	0.01	[0.01; 0.02]	16.5%	16.5%
3	0.01	0.0044	0.01	[0.00; 0.02]	16.1%	16.1%
4	0.01	0.0043	0.01	[0.00; 0.02]	16.9%	16.9%
5	0.01	0.0043	0.01	[0.01; 0.02]	17.2%	17.2%
6	0.01	0.0043	0.01	[0.00; 0.02]	17.0%	17.0%
Fixed effect model			**0.01**	**[0.01; 0.01]**	**100.0%**	**--**
Random effects model			**0.01**	**[0.01; 0.01]**	**--**	**100.0%**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.95$



Correlation between PRS and Population density (random effect meta, variance weighted) **R=0.0115 (SE=0.00176), pvalue=6.6e-11**

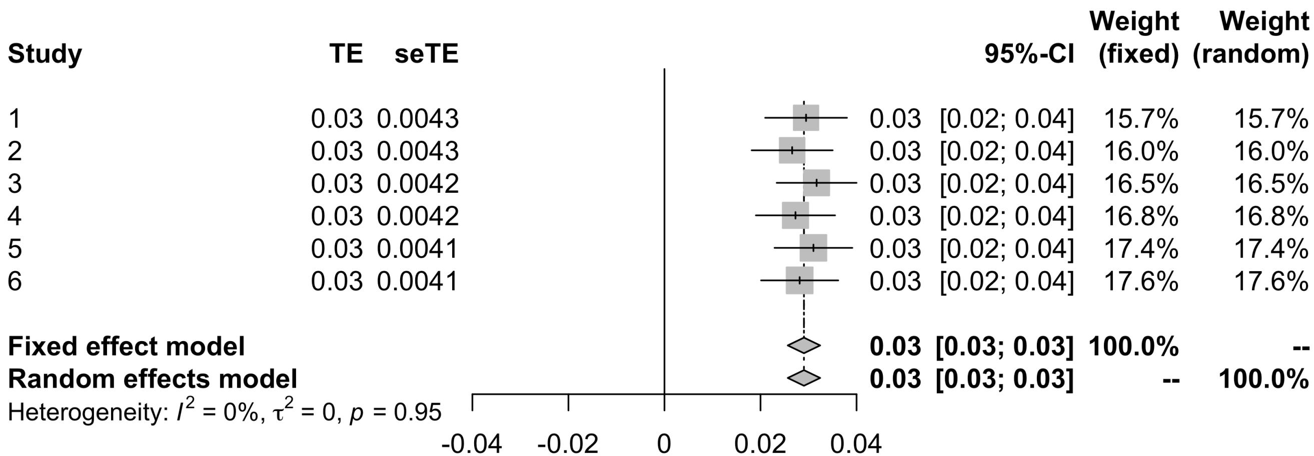
B) SNP HERITABILITY FOR POPULATION DENSITY



Heritability: h2=0.10 (SE=0.0031), pvalue=1.6e-230

Note this is regressing age sex and ses (Townsend index)

B) CORRELATION BETWEEN SES AND PRS – GRM

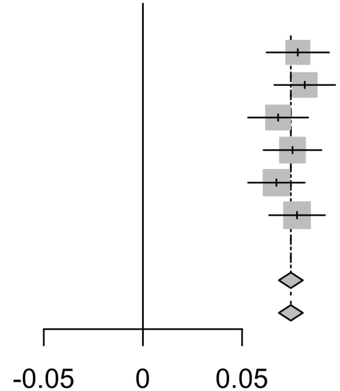


Correlation between PRS and SES
R=0.029 (SE=0.00172), pvalue=8.6e-64

B) SNP HERITABILITY FOR SES

Study	TE	seTE		95%-CI	Weight (fixed)	Weight (random)
1	0.08	0.0080	0.08	[0.06; 0.09]	14.8%	14.8%
2	0.08	0.0078	0.08	[0.07; 0.10]	15.6%	15.6%
3	0.07	0.0077	0.07	[0.05; 0.08]	15.9%	15.9%
4	0.08	0.0075	0.08	[0.06; 0.09]	17.1%	17.1%
5	0.07	0.0073	0.07	[0.05; 0.08]	18.0%	18.0%
6	0.08	0.0072	0.08	[0.06; 0.09]	18.5%	18.5%
Fixed effect model			**0.07**	**[0.07; 0.08]**	**100.0%**	**--**
Random effects model			**0.07**	**[0.07; 0.08]**	**--**	**100.0%**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.72$

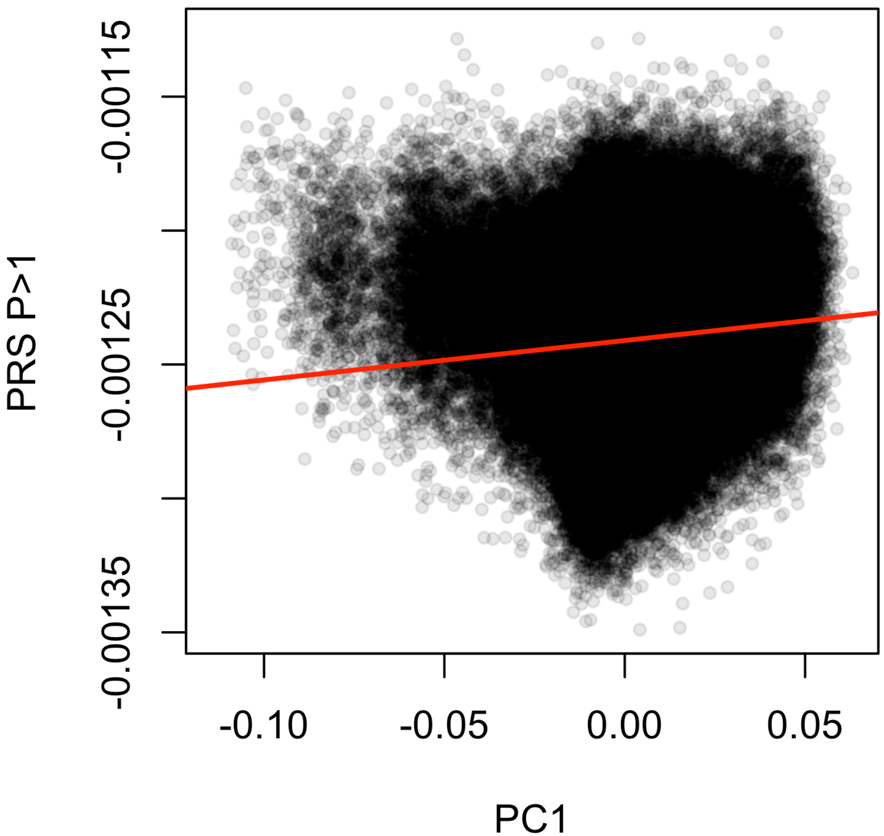


Heritability: h2=0.075 (SE=0.0031), pvalue=1.3e-128

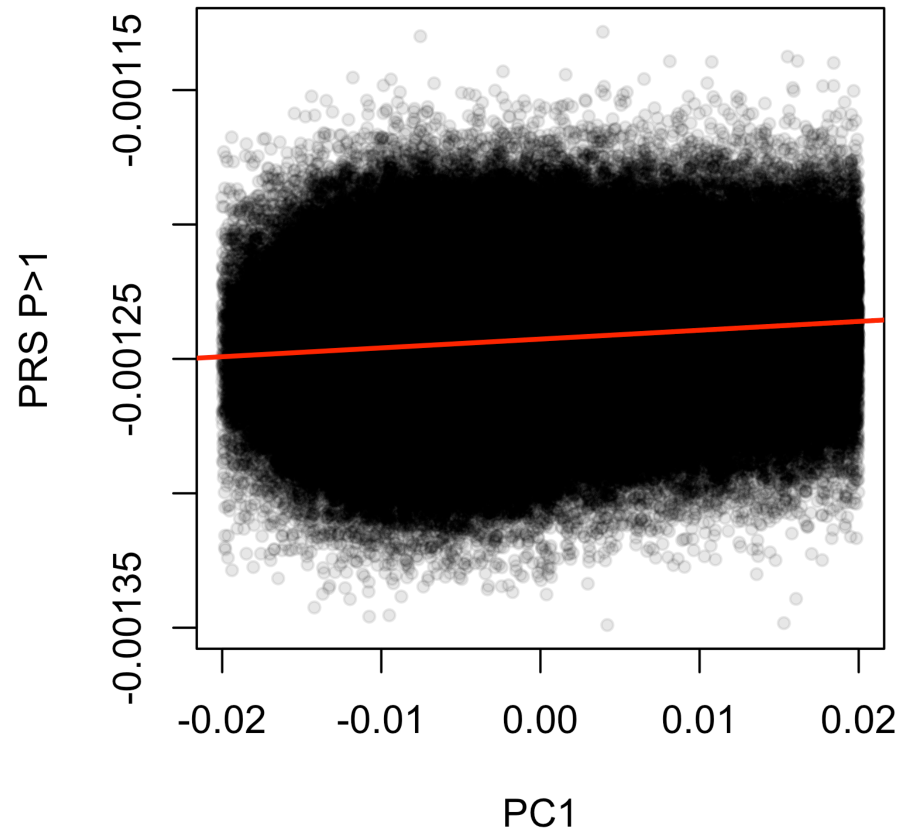
This is regressing out age, sex and population density

C) EXCLUDE PC OUTLIERS IN THE WHOLE SAMPLE

Exclusion of 86,141 participants, enforcing a exclusion criteria on 20 first PC (-0.02 – 0.02)



BP = 1480.3, df = 1, p-value < 2.2e-16



BP = 0.0095955, df = 1, p-value = 0.922

ANALYSIS WITH SES – 20 PC OUTLIERS EXCLUSION

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00123	3.11E-06	-396	0	NA	NA
factor(sex)1	-1.18E-05	4.53E-06	-2.6	0.009415864	NA	NA
popDensityPostCode	**7.70E-09**	**2.06E-09**	**3.74**	**0.000181503**	**0.00834**	**6.96E-05**
age	-3.44E-07	1.14E-07	-3.02	0.002537967	-0.115	0.0132
age2	3.18E-09	1.03E-09	3.1	0.001940581	0.118	0.0139
ageSex	4.11E-07	1.66E-07	2.48	0.013181119	0.492	0.242
age2Sex	-3.56E-09	1.49E-09	-2.39	0.016981533	-0.259	0.0671
ses	1.63E-07	1.75E-08	9.32	1.20E-20	0.0206	0.000424
PC1	0.00018	6.66E-06	27	1.53E-160	0.0652	0.00425
PC2	0.000341	8.90E-06	38.3	1.15018482351842e-320	0.0923	0.00852
PC3	-4.95E-05	1.23E-05	-4.04	5.37E-05	-0.00781	6.10E-05
PC4	2.35E-05	8.19E-06	2.87	0.004166661	0.00573	3.28E-05

PC5-20 not shown

BP = 51.057, df = 27, p-value = 0.003428

ANALYSIS BY COUNTRY

Rationale: the first 4 PCs delimitate pretty clearly the countries within the UK
Also there may be heterogeneous socio-economic, historical factors within each country

ANALYSIS WITH SES – 20 PC WALES (N=15039)

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00122	1.31E-05	-93.3	0	NA	NA
factor(sex)1	-4.10E-05	1.90E-05	-2.16	0.031158604	NA	NA
popDensityPostCode	-1.49E-08	1.52E-08	-0.984	0.325329347	-0.00801	6.42E-05
age	-8.14E-07	4.83E-07	-1.69	0.091441212	-0.265	0.0702
age2	7.89E-09	4.39E-09	1.8	0.072100356	0.283	0.0801
ageSex	1.53E-06	7.02E-07	2.18	0.029020085	1.8	3.24
age2Sex	-1.40E-08	6.37E-09	-2.19	0.028443009	-0.989	0.978
ses	8.10E-08	7.69E-08	1.05	0.292109382	0.00861	7.41E-05
PC1	0.000146	1.62E-05	9.02	2.08E-19	0.0795	0.00632
PC2	0.000431	1.87E-05	23.1	7.17E-116	0.191	0.0365
PC3	-0.000117	2.24E-05	-5.22	1.79E-07	-0.0391	0.00153
PC4	2.36E-05	1.76E-05	1.34	0.180153232	0.0151	0.000228
PC5-20 not shown						

BP = 29.803, df = 27, p-value = 0.323

ANALYSIS WITH SES – 20 PC ENGLAND (N=297538)

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00123	2.96E-06	-415	0 NA	NA	
factor(sex)1	-1.23E-05	4.33E-06	-2.85	0.004340354 NA	NA	
popDensityPostCode	**8.40E-09**	**1.87E-09**	**4.49**	**7.29E-06**	**0.00943**	**8.89E-05**
age	-3.93E-07	1.09E-07	-3.61	0.000301553	-0.129	0.0166
age2	3.55E-09	9.80E-10	3.62	0.000295106	0.129	0.0166
ageSex	4.33E-07	1.59E-07	2.73	0.006290112	0.507	0.257
age2Sex	-3.76E-09	1.43E-09	-2.63	0.008522345	-0.267	0.0713
ses	1.84E-07	1.70E-08	10.8	2.03E-27	0.0229	0.000524
PC1	0.000144	2.98E-06	48.3	0	0.0854	0.00729
PC2	0.000395	3.45E-06	114	0	0.202	0.0408
PC3	-1.00E-04	4.28E-06	-23.4	4.41E-121	-0.0382	0.00146
PC4	2.85E-06	5.97E-06	0.477	0.633353872	0.00084	7.06E-07
PC5-20 not shown						

BP = 38.061, df = 27, p-value = 0.07691

D) ANALYSIS WITH SES – 20 PC SCOTLAND (N=26052)

	Estimate	Std. Error	t value	Pr(>|t|)	r	r2
(Intercept)	-0.00124	9.94E-06	-125	0	NA	NA
factor(sex)1	1.37E-06	1.48E-05	0.0922	0.926556887	NA	NA
PopulationDensity	-1.52E-08	8.35E-09	-1.83	0.06799606	-0.0112	0.000125
age	-1.62E-07	3.64E-07	-0.445	0.656506015	-0.0522	0.00272
age2	1.90E-09	3.28E-09	0.581	0.561378332	0.0682	0.00465
ageSex	-1.13E-08	5.42E-07	-0.0209	0.983323419	-0.013	0.000169
age2Sex	-2.21E-10	4.88E-09	-0.0453	0.963905163	-0.0153	0.000234
ses	2.39E-07	4.47E-08	5.33	9.87E-08	0.0337	0.00114
PC1	0.00017	1.02E-05	16.7	2.65E-62	0.108	0.0117
PC2	0.000414	1.50E-05	27.5	1.72E-164	0.18	0.0324
PC3	-0.000115	1.85E-05	-6.19	6.14E-10	-0.0356	0.00127
PC4	-8.02E-05	3.03E-05	-2.65	0.008166572	-0.0168	0.000282

PC5-20 not shown

BP = 27.822, df = 27, p-value = 0.4202

SAMPLE OVERLAP BETWEEN UKB AND SCZ GWAS

Sample overlap may inflate associations found using PRS

As per the LD score paper, the intercept of the genetic covariance (I_{rg}) is a function of the phenotypic correlation r, the overlapping sample size Ns and the GWAS sample sizes N_1 and N_2

$I_{rg}=r*N_s / sqrt(N_1N_2) => N_s = sqrt(N_1N_2)*I_{rg} / r$

GWAS of population density and SES are running, I will try with those

CONCLUSIONS

PRS for SCZ associated with population density in the UK, and especially within England

Without correcting for SES:

AUS: r=0.035, r2=0.12% UK: r=0.018, r2=0.032%

Correcting for SES:

AUS: r=0.028, r2=0.08% UK: r=0.012, r2=0.013%

In the UK stronger association between SES (Townsend deprivation index) and SCZ-PRS. This was not significant in the Australian sample when correcting for population density (pvalue=0.35, 0.00028% of PRS variance accounted for).

SES MEASUREMENT IN AUSTRALIA - USED

Australia: SEIFA (Socio Economic Indexes for Areas)

We used the Index of **Relative Socio-economic Advantage and Disadvantage (IRSAD)** that summarises information about the economic and social conditions of people and households within an area, including both relative advantage and disadvantage measures.

FINAL VARIABLE LIST

Variable	Description	Loading
INC_HIGH	% of people with stated household equivalised income greater than $52,000 per year	0.84
HIGHMORTGAGE	% of occupied private dwellings paying mortgage greater than $2,800 per month	0.70
DIPLOMA	% of people aged 15 years and over whose highest level of educational attainment is a diploma qualification	0.63
OCC_PROF	% of employed people classified as Professionals	0.62
HIGHBED	% of occupied private dwellings with four (4) or more bedrooms	0.52
OCC_MANAGER	% of employed people classified as managers	0.42
HIGHRENT	% of occupied private dwellings paying rent greater than $370 per week	0.40
SPAREBED	% of occupied private dwellings with one or more spare bedrooms	0.37
ATUNI	% of people aged 15 years and over at university or other tertiary institution	0.36
HIGHCAR	% of occupied private dwellings with three (3) or more cars	0.35
NOEDU	% of people aged 15 years and over who have no educational attainment	-0.37
OVERCROWD	% of occupied private dwellings requiring one or more extra bedrooms	-0.45
NOCAR	% of occupied private dwellings with no cars	-0.49
OCC_SERVICE_L	% of employed people classified as low skill Community and Personal Service workers	-0.51
OCC_DRIVER	% of employed people classified as Machinery Operators and Drivers	-0.57
SEP_DIVORCED	% of people aged 15 years and over who are separated or divorced	-0.57
LOWRENT	% of occupied private dwellings paying rent less than $166 per week (excluding $0 per week)	-0.67
DISABILITYU70	% of people under the age of 70 who have a long-term health condition or disability and need assistance with core activities	-0.67
UNEMPLOYED	% of people (in the labour force) who are unemployed	-0.69
ONEPARENT	% of one parent families with dependent offspring only	-0.69
OCC_LABOUR	% of employed people classified as Labourers	-0.78
CHILDJOBLESS	% of families with children under 15 years of age who live with jobless parents	-0.80
NOYEAR12ORHIGHER	% of people aged 15 years and over whose highest level of education is Year 11 or lower	-0.82
NONET	% of occupied private dwellings with no internet connection	-0.82
INC_LOW	% of people with stated household equivalised income between $1 and $20,799 per year	-0.89

SES MEASUREMENT IN AUSTRALIA – (NOT USED)

There is another SEIFA score

Index of Relative Socio-economic Disadvantage (IRSD), which is a general socio-economic index that summarises a range of information about the economic and social conditions of people and households within an area. Unlike the other indexes, this index includes only measures of relative disadvantage.

FINAL VARIABLE LIST

Variable	Description	Loading
ENGLISHPOOR	% of people who do not speak English well	-0.34
NOEDU	% of people aged 15 years and over who have no educational attainment	-0.44
OCC_SERVICE_L	% of employed people classified as low skill Community and Personal Service workers	-0.50
OCC_DRIVERS	% of employed people classified as Machinery Operators and Drivers	-0.52
OVERCROWD	% of occupied private dwellings requiring one or more extra bedrooms	-0.52
SEP_DIVORCED	% of people aged 15 years and over who are separated or divorced	-0.54
NOCAR	% of occupied private dwellings with no cars	-0.56
DISABILITYU70	% of people under the age of 70 who have a long-term health condition or disability and need assistance with core activities	-0.66
ONEPARENT	% of one parent families with dependent offspring only	-0.71
LOWRENT	% of occupied private dwellings paying rent less than $166 per week (excluding $0 per week)	-0.73
UNEMPLOYED	% of people (in the labour force) who are unemployed	-0.74
NOYEAR12ORHIGHER	% of people aged 15 years and over whose highest level of education is Year 11 or lower	-0.75
OCC_LABOUR	% of employed people classified as Labourers	-0.75
NONET	% of occupied private dwellings with no internet connection	-0.81
CHILDJOBLESS	% of families with children under 15 years of age who live with jobless parents	-0.85
INC_LOW	% of people with stated household equivalised income between $1 and $20,799 per year	-0.90

Conceptually more similar to the Townsend index ("index of deprivation") though we can expect the 2 SEIFA scores to correlate

SES MEASUREMENT IN THE UK

Townsend deprivation index

Wikipedia:
"The **Townsend index** is a measure of material deprivation within a population. It was first described by sociologist Peter Townsend in 1988.[1]
The measure incorporates four variables:

- Unemployment (as a percentage of those aged 16 and over who are economically active);
- Non-car ownership (as a percentage of all households);
- Non-home ownership (as a percentage of all households);
- Household overcrowding.