Test scores and month of birth

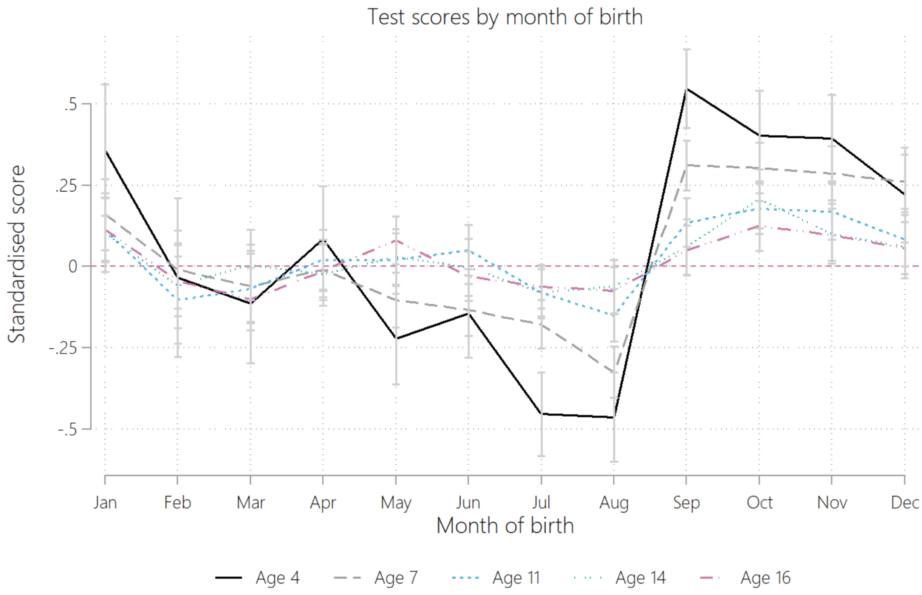

Figure 1: Standardised test scores at different ages by month of birth.

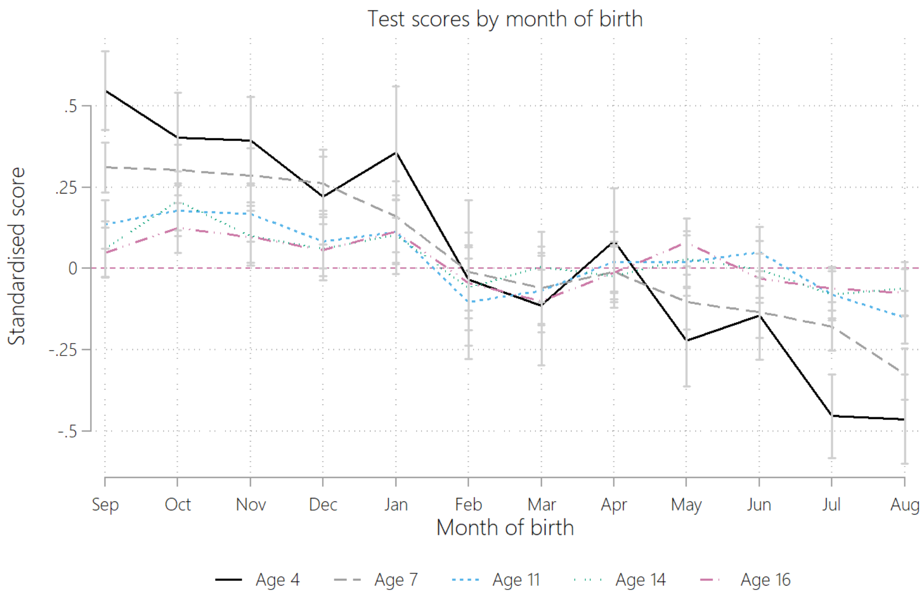

Figure 2: Standardised test scores at different ages by month of birth - showing approx linear trend.

$$Y_i = \sum_{m=1}^{11} \beta_m MoB_m + \lambda PGS_i + \alpha female_i + \sum_{k=1}^{10} \phi_k PC_k + u_i$$

where Y_i is standardised test score of individual i, and β_m are the 11 month-specific coefficients ($1 = Sep, 2 = Oct, ..., 11 = Jul$, as in Figure ??). We control for the child's EA polygenic score, standardized to have mean 0 and std dev 1. Also control for gender and the first 10 principal components.

Table 1: OLS of test scores on MoB (dummies)

	(1) Age 4	(2) Age 7	(3) Age 11	(4) Age 14	(5) Age 16
Oct	-0.141***	-0.036***	0.025***	0.111***	0.063***
	(0.005)	(0.001)	(0.002)	(0.002)	(0.002)
Nov	-0.168***	-0.045***	0.022***	0.030***	0.038***
	(0.008)	(0.002)	(0.002)	(0.002)	(0.002)
Dec	-0.310***	-0.070***	-0.042***	-0.001	0.003*
	(0.004)	(0.002)	(0.003)	(0.002)	(0.002)
Jan	-0.207***	-0.186***	-0.038***	0.005**	0.049***
	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)
Feb	-0.575***	-0.317***	-0.235***	-0.120***	-0.092***
	(0.006)	(0.003)	(0.001)	(0.003)	(0.002)
Mar	-0.611***	-0.382***	-0.208***	-0.064***	-0.150***
	(0.009)	(0.002)	(0.002)	(0.002)	(0.001)
Apr	-0.445***	-0.349***	-0.133***	-0.115***	-0.086***
	(0.009)	(0.001)	(0.002)	(0.002)	(0.002)
May	-0.743***	-0.460***	-0.157***	-0.075***	0.012***
	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)
Jun	-0.681***	-0.462***	-0.103***	-0.085***	-0.087***
	(0.006)	(0.002)	(0.002)	(0.002)	(0.002)
Jul	-1.003***	-0.522***	-0.242***	-0.169***	-0.117***
	(0.007)	(0.002)	(0.003)	(0.002)	(0.002)
Aug	-0.972***	-0.654***	-0.302***	-0.146***	-0.130***
	(0.008)	(0.002)	(0.003)	(0.001)	(0.002)
PGS	0.135***	0.244***	0.309***	0.311***	0.121***
	(0.026)	(0.009)	(0.012)	(0.010)	(0.012)
R2	0.187	0.138	0.113	0.110	0.035
Observations	1826	5734	6075	5240	7268

Robust standard errors clustered by MoB in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To allow for linear trend in MoB, i.e.:

$$Y_i = \beta MoB + \lambda PGS_i + \alpha female_i + \sum_{k=1}^{10} \phi_k PC_k + u_i$$

where MoB is coded as $1 = Sep, 2 = Oct, \ldots, 11 = Jul$. This gives the following estimates:

Table 2: OLS of test scores on MoB (linear)

	(1) Age 4	(2) Age 7	(3) Age 11	(4) Age 14	(5) Age 16
Treated	0.684***	0.451***	0.186***	0.142***	0.112***
	(0.084)	(0.035)	(0.041)	(0.028)	(0.027)
Treated*MoB	0.117***	0.076***	0.038**	0.026***	0.011
	(0.029)	(0.013)	(0.016)	(0.008)	(0.013)
MoB	-0.101***	-0.057***	-0.026*	-0.017***	-0.007
	(0.024)	(0.011)	(0.013)	(0.005)	(0.011)
PGS	0.136***	0.244***	0.309***	0.312***	0.121***
	(0.025)	(0.009)	(0.012)	(0.011)	(0.012)
R2	0.171	0.136	0.110	0.108	0.033
Observations	1826	5734	6075	5240	7268

Robust standard errors clustered by MoB in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note that the R^2 is much lower for the KS4 specification (column 5). Not sure why this is. The KS4 variable is measured slightly differently, but should still pick up EA. Would have to look into this more if we want to go ahead with this, or let's just focus on Age 4–14 for now.

I prefer this second specification with linear MoB. I think this model would be preferred statistically too, e.g. very similar R2 using fewer df.

Next, we run the following regression, allowing for an interaction between G and E:

$$Y_i = \sum_{m=1}^{11} \eta_m MoB_m \times PGS_i + \sum_{m=1}^{11} \beta_m MoB_m + \lambda PGS_i + \alpha female_i + \sum_{k=1}^{10} \phi_k PC_k + u_i$$

Table 3: OLS of test scores on MoB (dummies) and interactions with PGS

	(1) Age 4	(2) Age 7	(3) Age 11	(4) Age 14	(5) Age 16
Oct*PGS	0.111***	0.065***	-0.009**	0.074***	0.030***
	(0.015)	(0.005)	(0.003)	(0.006)	(0.003)
Nov*PGS	0.142***	0.068***	0.055***	0.099***	0.036***
	(0.013)	(0.006)	(0.004)	(0.007)	(0.004)
Dec*PGS	0.092***	0.053***	0.051***	0.045***	-0.027***
	(0.011)	(0.005)	(0.005)	(0.007)	(0.004)
Jan*PGS	0.036*	-0.022***	-0.057***	0.032***	-0.030***
	(0.017)	(0.004)	(0.005)	(0.008)	(0.003)
Feb*PGS	0.221***	0.064***	0.109***	0.080***	0.013**
	(0.017)	(0.006)	(0.004)	(0.007)	(0.005)
Mar*PGS	-0.102***	0.101***	0.126***	-0.009**	0.043***
	(0.012)	(0.005)	(0.003)	(0.004)	(0.005)
Apr*PGS	-0.014	0.060***	0.022***	0.002	0.024***
	(0.015)	(0.007)	(0.004)	(0.005)	(0.004)
May*PGS	0.065***	0.096***	0.082***	0.041***	0.094***
	(0.011)	(0.004)	(0.004)	(0.005)	(0.003)
Jun*PGS	0.217***	0.021***	0.009***	-0.025***	-0.067***
	(0.012)	(0.004)	(0.002)	(0.004)	(0.003)
Jul*PGS	0.162***	0.062***	-0.015***	0.025***	0.015***
	(0.014)	(0.004)	(0.003)	(0.005)	(0.003)
Aug*PGS	0.048***	0.069***	0.026***	0.045***	0.008
	(0.010)	(0.006)	(0.003)	(0.005)	(0.004)
Oct	-0.140***	-0.025***	0.030***	0.118***	0.068***
	(0.009)	(0.003)	(0.002)	(0.005)	(0.003)
Nov	-0.158***	-0.041***	0.020***	0.039***	0.040***
	(0.010)	(0.003)	(0.004)	(0.006)	(0.004)
Dec	-0.290***	-0.061***	-0.040***	0.007*	0.004
	(0.008)	(0.002)	(0.004)	(0.003)	(0.003)
Jan	-0.191***	-0.180***	-0.039***	0.012**	0.052***
	(0.010)	(0.004)	(0.005)	(0.005)	(0.003)
Feb	-0.534***	-0.307***	-0.230***	-0.112***	-0.086***
	(0.007)	(0.004)	(0.003)	(0.005)	(0.005)
Mar	-0.638***	-0.370***	-0.205***	-0.061***	-0.145***
	(0.010)	(0.003)	(0.002)	(0.003)	(0.002)
Apr	-0.434***	-0.342***	-0.133***	-0.108***	-0.081***
	(0.010)	(0.003)	(0.004)	(0.005)	(0.004)
May	-0.730***	-0.454***	-0.160***	-0.067***	0.012***
	(0.010)	(0.003)	(0.004)	(0.004)	(0.003)
Jun	-0.671***	-0.457***	-0.101***	-0.081***	-0.079***
	(0.007)	(0.002)	(0.002)	(0.003)	(0.003)
Jul	-0.993***	-0.513***	-0.237***	-0.160***	-0.113***
	(0.011)	(0.003)	(0.003)	(0.005)	(0.002)
Aug	-0.963***	-0.646***	-0.301***	-0.142***	-0.128***
	(0.008)	(0.002)	(0.002)	(0.004)	(0.003)
PGS	0.060**	0.160***	0.268***	0.272***	0.101***
	(0.020)	(0.015)	(0.014)	(0.016)	(0.013)
R2	0.200	0.142	0.116	0.114	0.038
Observations	1826	5734	6075	5240	7268

Robust standard errors clustered by MoB in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

I don't think this is very informative. Some are positive, others negative, which would suggest that in some months having a high PGS increases inequalities in EA between Aug and Sept births, but it reduces inequalities in other months. It also doesn't seem particularly robust across specifications. Therefore, allowing for a linear trend in MoB:

Table 4: OLS of test scores on MoB (linear) and interactions with PGS

	(1) Age 4	(2) Age 7	(3) Age 11	(4) Age 14	(5) Age 16
Treated	0.682***	0.450***	0.185***	0.142***	0.112***
	(0.083)	(0.034)	(0.041)	(0.028)	(0.027)
Treated*PGS	-0.019	-0.025	0.009	0.028	0.004
	(0.066)	(0.021)	(0.028)	(0.021)	(0.023)
Treated*MoB	0.112***	0.077***	0.038**	0.026***	0.011
	(0.030)	(0.012)	(0.016)	(0.007)	(0.013)
MoB*PGS	0.029	-0.005	-0.019*	0.008**	-0.012**
	(0.020)	(0.004)	(0.009)	(0.003)	(0.005)
MoB*PGS*Treated	-0.032	0.009	0.021*	-0.006	0.013**
	(0.021)	(0.006)	(0.011)	(0.004)	(0.006)
MoB	-0.097***	-0.057***	-0.026*	-0.017***	-0.007
	(0.025)	(0.011)	(0.013)	(0.005)	(0.011)
PGS	0.178**	0.219***	0.278***	0.297***	0.103***
	(0.069)	(0.017)	(0.018)	(0.016)	(0.024)
R2	0.179	0.139	0.111	0.111	0.035
Observations	1826	5734	6075	5240	7268

Robust standard errors clustered by MoB in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

hmmm... bit of a shame really! But a null-result is also a result... We haven't controlled for the school and cohort fixed effects yet.