Online data supplement to: Class *et al*. Fetal growth and psychiatric and socioeconomic problems: population-based sibling comparison. *Br J Psychiatry* doi: 10.1192/bjp.bp.113.143693

Description of Swedish population-based registers with references and participant flow

The data were obtained by linking information available in the following government-maintained, Swedish population-based registries: (1) the Medical Birth Registry includes data on more than 99% of pregnancies in Sweden since 1973 [1, 2]; (2) the Multi-Generation Register [3] contains information about biological and adoptive relationships living in Sweden since 1933; (3) the Migration Register contains dates for migration into or out of Sweden; (4) the Cause of Death Register contains dates and causes of all deaths since 1958; (5) the National Patient Registry [4] provides data on all psychiatric inpatient admissions in Sweden since 1973 and outpatient care since 2001. Every record includes the discharge date, primary discharge diagnosis, and up to seven secondary diagnoses assigned by the treating medical doctor using WHO's ICD-10 codes since 1997 [5]; (6) the National Crime Register includes detailed information about all criminal convictions in lower court since 1973 on those aged 15 (the age of criminal responsibility) and older [6]; (7) the National School Register [7] includes grades across subjects for students at the end of grade nine (approximately age 16 years) since 1983; (8) the Education Register contains information on highest level of completed formal education between 1988 and 2008; (9) the longitudinal integrated database for health insurance and social studies (LISA) [8] contains yearly assessments of income, marital status, unemployment status, social welfare status, and education for all individuals 16 years of age and older since 1990.

Table DS1 International Classification of Disease (ICD) code used to classify outcomes with outcome description.

Outcome	Data Source	ICD Version	ICD Codes	Description
Psychiatric Problems				
ADHD	PR	9,10	314, F90	Hyperkinetic syndrome and attention-deficit hyperactivity disorders
ASD	PR	9, 10	299, F84	Includes disintegrative psychosis, Heller's syndrome, and schizophrenic syndrome of childhood
Psychotic or Bipolar Disorder	PR	8, 9, 10	295, F20 296.1, 296.3, 296A-296E, 296W, F30-F31 291, 292, 296.0, 296.2, 296.9, 297-299, 296B, 296X, F32.3 x.5 in F10-F19	Schizophrenia Bipolar disorder Other non-organic psychoses
Suicide Attempt	PR	8, 9, 10	E950-E959, E980-E989, X60-X84, Y870, Y10-Y34, Y872	Certain and uncertain attempts including violent, non-violent, other
Substance Use Problem	PR	8, 9, 10	303, 304, 305A, 305X, F10 (except x.5), F11-F19 (except x.5)	Alcohol and drug abuse (excludes nicotine)
Criminality	NCR	NA	NA	Earliest conviction date for any criminal act
Socioeconomic Outcomes				
Failing Grades	NSR	NA	NA	Poor school performance across all 16 academic subjects in grade 9 (about age 16)
Education under 10 years	ER	NA	NA	Low educational attainment
Higher Education	ER	NA	NA	Three or more years of postsecondary education
Social Welfare Benefits	LISA	NA	NA	Age at first receipt of government social welfare subsidies

Note: ASD = Autism spectrum disorder; PR = Patient Register; NCR = National Crime Register; NSR = National School Register; ER = Education register; MBR = Medical Birth Register

Table DS2 Unstandardised linear and quadratic regression coefficients for the baseline, adjusted, and fixed effects models.

	Baseline Model				Adjusted Model				Fixed Effects Model			
	Linear term		Quadratic term		Linear term		Quadratic term		Linear term		Quadratic term	
Outcomes	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE
Psychiatric Problems												
ASD	-0.003	0.005	**0.011**	**0.001**	0.004	0.005	**0.010**	**0.001**	**-0.072**	**0.011**	**0.012**	**0.002**
ADHD	**-0.020**	**0.003**	**0.007**	**0.001**	-0.004	0.003	**0.007**	**0.001**	**-0.038**	**0.009**	**0.007**	**0.002**
Psychotic or Bipolar Disorder	**-0.011**	**0.004**	**0.003**	**0.001**	-0.001	0.004	**0.003**	**0.001**	-0.017	0.012	0.004	0.002
Suicide Attempt	**-0.031**	**0.003**	-	-	**-0.012**	**0.003**	-	-	-0.003	0.008	-	-
Substance Use Problem	**-0.041**	**0.003**	-	-	**-0.022**	**0.003**	-	-	0.009	0.006	-	-
Criminality	**-0.033**	**0.001**	**-0.001**	**0.000**	**-0.014**	**0.001**	**-0.001**	**0.000**	**0.016**	**0.003**	**-0.002**	**0.001**
Socioeconomic Outcomes												
Failing Grades	**-0.177**	**0.002**	**-0.011**	**0.000**	**-0.153**	**0.002**	**-0.011**	**0.000**	**-0.137**	**0.004**	**-0.013**	**0.001**
Education Under 10 yrs	**-0.047**	**0.001**	**0.003**	**0.000**	**-0.031**	**0.001**	**0.003**	**0.000**	-0.004	0.003	**0.003**	**0.001**
Social Welfare Receipt	**-0.065**	**0.001**	**0.001**	**0.000**	**-0.038**	**0.001**	**0.001**	**0.000**	-0.003	0.003	0.000	0.000

Note: ADHD=Attention Deficit Hyperactivity Disorder; ASD=Autism Spectrum Disorder; b=maximum likelihood estimate of the unstandardised regression coefficient; SE = estimated standard error; A dash(-) indicates that the p-value of the Wald chi-square test statistic for the quadratic parameter is greater than 0.05, and therefore not included in the model; bold coefficients have p-value < 0.05.

Table DS3 Comparison of Akaike information criterion fit indices across baseline linear and quadratic candidate models.

The model selection Table DS3 compares the Akaike information criterion (AIC) for the baseline model with linear (L) birth weight only and the baseline model with both linear and quadratic (L+Q) birth weight. The column labeled "AIC-min" indicates which of the two candidate models (L or L+Q) yielded the lowest AIC. The observed difference, $\Delta AIC = AIC_L - AIC_{L+Q}$, provides a measure of relative merit that is free of scaling constants and can be interpreted as strength of evidence for model selection purposes [9].

Table DS3. Comparison of AIC values for linear and quadratic candidate models.

	Candidate Model			
Outcome	Linear BW	with Quadratic BW	AIC-min	ΔAIC
Psychiatric Problems				
ADHD	844471.23	844370.62	L + Q	100.61
ASD	452237.10	452102.20	L + Q	134.90
Psychotic or Bipolar Disorder	539246.73	539235.29	L + Q	11.44
Suicide Attempt	863605.97	863607.04	L	-1.07
Substance Use Problem	1303851.20	1303852.90	L	-1.70
Criminality	7995423.10	7995413.10	L + Q	10.00
Socioeconomic Outcomes				
Failing Grades	1347757.50	1346626.00	L + Q	1131.50
Education under 10 yrs	1839981.20	1839871.00	L + Q	110.20
Social Welfare Receipt	10651879.00	10651870.00	L + Q	9.00

Note: BW = birth weight; ADHD = Attention Deficit Hyperactivity Disorder; ASD = Autism Spectrum Disorder; L = model with linear birth weight only; L+Q = baseline model with both linear and quadratic birth weight

Table DS4. Odds or Cox hazard regression parameter estimates for baseline and fixed effects models using ordinal birth weight.

Table DS4 presents the unstandardized regression coefficients with standard errors and the point estimates, either Odds Ratios or Hazard Ratios, with 95% confidence intervals associated with the ordinal bins of birth weight across baseline and fixed effects models. The baseline estimates presented here correspond with the point estimates presented in Fig. 1 within the main paper.

Estimates for fixed effects models using ordinal representation of birth weight provide a sensitivity analysis to examine the sibling comparison results absent of assumptions about the underlying pattern (i.e., linear or quadratic) of the associations between birth weight and the indices of mortality and morbidity. Figures 1 and 2 in the main paper provide a graphical comparison of the baseline and fixed effects models using ordinally represented birth weight. The fixed effects results using ordinal representation of birth weight give commensurate results with analyses based on linear and quadratic modeling presented in the main analyses. It can be noted, however, that the confidence intervals around fixed effects estimates using ordinal bins are larger than those presented in the main analyses due to the reduced statistical power in moving from a continuous representation of birth weight to ordinal bins. These results suggest that assumptions about the shape of model fitting using families with multiple offspring (which are the only informative families for the sibling comparison estimates) do not account for the fixed effects results using the continuous index of birth weight.

Table DS4 Odds or Cox hazard regression parameter estimates for baseline and fixed effects models using ordinal birth weight.

				Maximum Likelihood Parameters						
	Outcome	Model	BW bin (gm)	b	SE	ChiSq	Pr > ChiSq	HR/OR	95%LCL	95%UCL
Psychiatric Problems										
	ADHD	**Baseline**	**≤ 2500**	0.433	0.03	159.864	<.0001	1.542	1.442	1.649
			2501-3000	0.249	0.02	149.024	<.0001	1.283	1.232	1.335
			3001-3500	0.099	0.01	44.834	<.0001	1.104	1.072	1.136
			≥ 4001	0.059	0.02	11.599	0.001	1.061	1.025	1.097
		Fixed effects	**≤ 2500**	0.498	0.08	37.288	<.0001	1.646	1.403	1.932
			2501-3000	0.217	0.05	20.517	<.0001	1.243	1.131	1.365
			3001-3500	0.120	0.03	14.237	0.000	1.127	1.059	1.199
			≥ 4001	-0.001	0.04	0.001	0.971	0.999	0.927	1.076
	ASD	**Baseline**	**≤ 2500**	0.582	0.04	168.073	<.0001	1.790	1.639	1.955
			2501-3000	0.242	0.03	74.486	<.0001	1.274	1.206	1.346
			3001-3500	0.055	0.02	7.198	0.007	1.056	1.015	1.099
			≥ 4001	0.131	0.02	31.805	<.0001	1.140	1.089	1.193

	Fixed effects	**≤ 2500**	0.890	0.10	76.280	<.0001	2.435	1.994	2.973
		2501-3000	0.469	0.06	59.083	<.0001	1.598	1.418	1.801
		3001-3500	0.162	0.04	16.107	<.0001	1.176	1.087	1.273
		≥ 4001	-0.022	0.05	0.224	0.636	0.978	0.893	1.072
Psychotic or Bipolar Disorder	**Baseline**	**≤ 2500**	0.172	0.04	16.019	<.0001	1.188	1.092	1.293
		2501-3000	0.102	0.02	17.859	<.0001	1.108	1.056	1.162
		3001-3500	0.020	0.02	1.210	0.271	1.020	0.985	1.056
		≥ 4001	-0.004	0.02	0.037	0.849	0.996	0.952	1.041
	Fixed effects	**≤ 2500**	0.214	0.10	4.465	0.035	1.239	1.016	1.511
		2501-3000	0.076	0.06	1.822	0.177	1.079	0.966	1.206
		3001-3500	0.062	0.04	2.752	0.097	1.064	0.989	1.146
		≥ 4001	0.002	0.05	0.001	0.970	1.002	0.913	1.099
Suicide Attempt	**Baseline**	**≤ 2500**	0.176	0.04	24.390	<.0001	1.192	1.112	1.279
		2501-3000	0.166	0.02	75.531	<.0001	1.180	1.137	1.225
		3001-3500	0.075	0.01	28.430	<.0001	1.078	1.049	1.108
		≥ 4001	-0.025	0.02	1.852	0.174	0.975	0.940	1.011
	Fixed effects	**≤ 2500**	-0.058	0.08	0.578	0.447	0.944	0.813	1.095
		2501-3000	0.074	0.04	3.335	0.068	1.077	0.995	1.166
		3001-3500	-0.021	0.03	0.578	0.447	0.980	0.929	1.033
		≥ 4001	-0.030	0.03	0.751	0.386	0.970	0.906	1.039
Substance Use Problem	**Baseline**	**≤ 2500**	0.237	0.03	64.672	<.0001	1.267	1.196	1.342
		2501-3000	0.185	0.02	135.179	<.0001	1.203	1.166	1.241
		3001-3500	0.092	0.01	64.077	<.0001	1.097	1.072	1.122
		≥ 4001	-0.068	0.01	21.424	<.0001	0.934	0.907	0.961
	Fixed effects	**≤ 2500**	-0.072	0.06	1.575	0.210	0.930	0.831	1.041
		2501-3000	-0.039	0.03	1.537	0.215	0.962	0.904	1.023
		3001-3500	-0.012	0.02	0.318	0.573	0.988	0.948	1.030
		≥ 4001	-0.013	0.03	0.246	0.620	0.987	0.936	1.040
Criminality	**Baseline**	**≤ 2500**	0.139	0.01	113.851	<.0001	1.149	1.120	1.179
		2501-3000	0.141	0.01	447.394	<.0001	1.151	1.136	1.167
		3001-3500	0.072	0.00	240.863	<.0001	1.075	1.065	1.084
		≥ 4001	-0.063	0.01	123.894	<.0001	0.939	0.929	0.950
	Fixed effects	**≤ 2500**	-0.140	0.03	28.312	<.0001	0.869	0.825	0.915
		2501-3000	-0.077	0.01	30.061	<.0001	0.926	0.901	0.952
		3001-3500	-0.026	0.01	8.378	0.004	0.974	0.957	0.992
		≥ 4001	0.019	0.01	2.764	0.096	1.019	0.997	1.041
Socioeconomic Outcomes									
Failing Grades	**Baseline**	**≤ 2500**	0.509	0.01	1402.943	<.0001	1.664	1.620	1.708

		2501-3000	0.412	0.01	2985.285	<.0001	1.509	1.487	1.532
		3001-3500	0.299	0.01	2931.030	<.0001	1.348	1.333	1.363
		≥ 4001	-0.495	0.01	3746.348	<.0001	0.610	0.600	0.620
	Fixed effects	**≤ 2500**	0.065	0.03	5.199	0.023	1.067	1.009	1.128
		2501-3000	0.104	0.02	42.666	<.0001	1.110	1.076	1.145
		3001-3500	0.167	0.01	242.267	<.0001	1.182	1.157	1.207
		≥ 4001	-0.446	0.01	917.245	<.0001	0.640	0.622	0.659
Education under 10 yrs	**Baseline**	**≤ 2500**	0.376	0.01	927.635	<.0001	1.456	1.421	1.492
		2501-3000	0.253	0.01	1576.803	<.0001	1.288	1.272	1.305
		3001-3500	0.112	0.00	627.380	<.0001	1.118	1.108	1.128
		≥ 4001	-0.057	0.01	108.709	<.0001	0.945	0.934	0.955
	Fixed effects	**≤ 2500**	0.161	0.03	37.388	<.0001	1.175	1.116	1.237
		2501-3000	0.054	0.01	14.857	0.000	1.055	1.027	1.085
		3001-3500	0.011	0.01	1.497	0.221	1.011	0.993	1.029
		≥ 4001	0.017	0.01	2.371	0.124	1.017	0.995	1.039
Social Welfare Receipt	**Baseline**	**≤ 2500**	0.417	0.01	1664.692	<.0001	1.518	1.488	1.548
		2501-3000	0.321	0.01	3453.146	<.0001	1.379	1.364	1.394
		3001-3500	0.143	0.00	1272.565	<.0001	1.154	1.145	1.163
		≥ 4001	-0.090	0.01	299.761	<.0001	0.914	0.905	0.924
	Fixed effects	**≤ 2500**	-0.001	0.02	0.003	0.956	0.999	0.954	1.045
		2501-3000	0.005	0.01	0.125	0.724	1.005	0.980	1.030
		3001-3500	-0.007	0.01	0.651	0.420	0.993	0.977	1.010
		≥ 4001	0.004	0.01	0.138	0.710	1.004	0.983	1.026

Note: b = unstandardized regression coefficient; SE=standard error

BW = birthweight in grams

ADHD=Attention Deficit Hyperactivity Disorder; ASD=Autism Spectrum Disorder

Figure DS1 Fixed effects parameter estimates when limiting sample to full term births only.

Compared with parameter estimates from the main analyses, which included all gestational ages, results from analyses limited to full term births did not substantially alter the results (Figure DS1). This suggests that associations presented in main analyses were not biased by extremely premature or late births. Figure DS1 presents main analyses figures as well as those limited to full term births only (right column). As can be seen in Figure DS1, when restricted to full term births only (right figure), parameters corresponding to the smallest ordinal category of birth weight were attenuated as compared with main analyses (left figure). Small sample size may also contribute to this attenuation.

Figure DS1 Comparison of birth weight predicting psychiatric and socioeconomic outcomes across all gestational ages (left column) and full term only (right column) births.

Psychiatric Problems

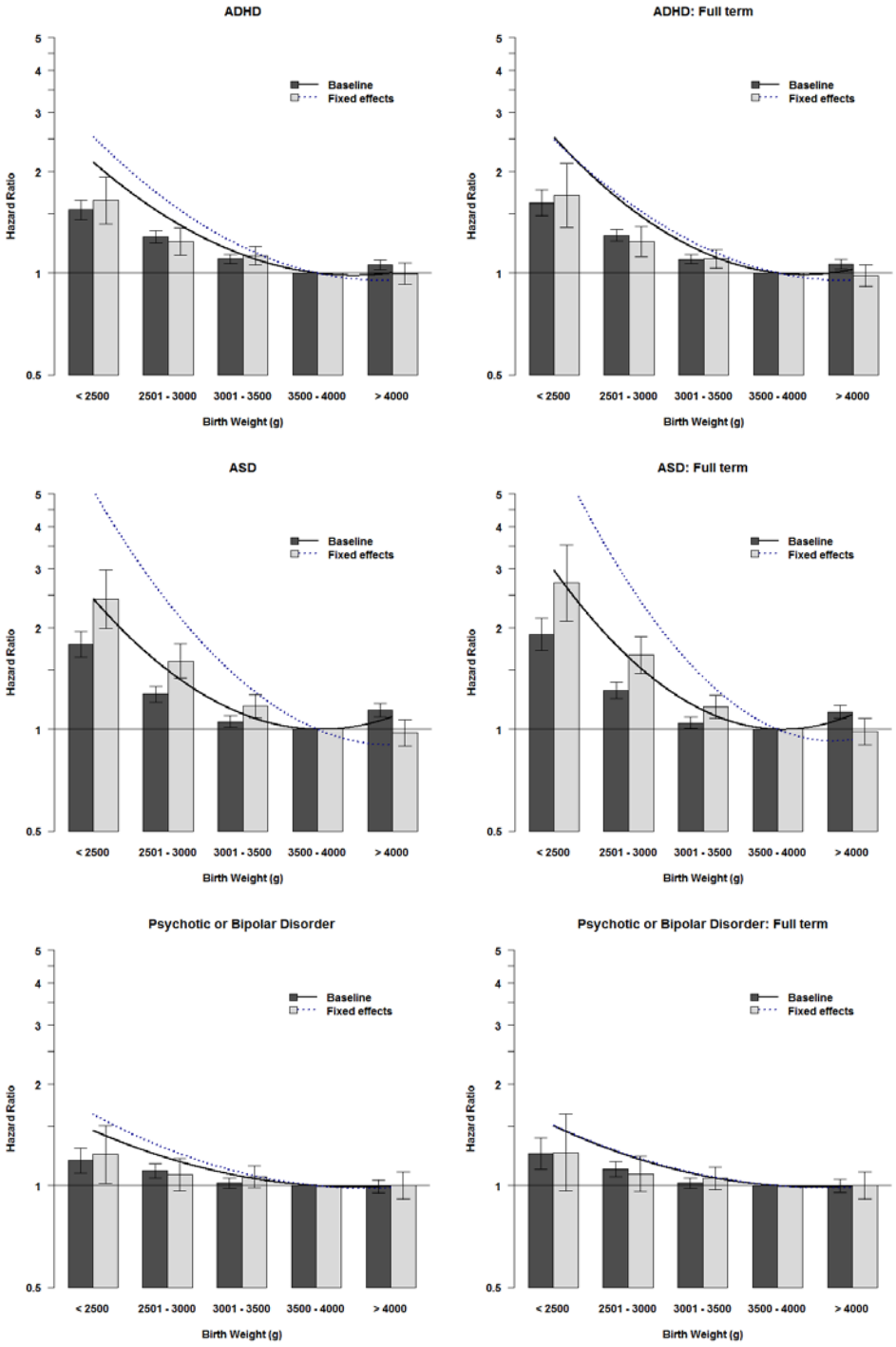


Figure DS1 Continued.

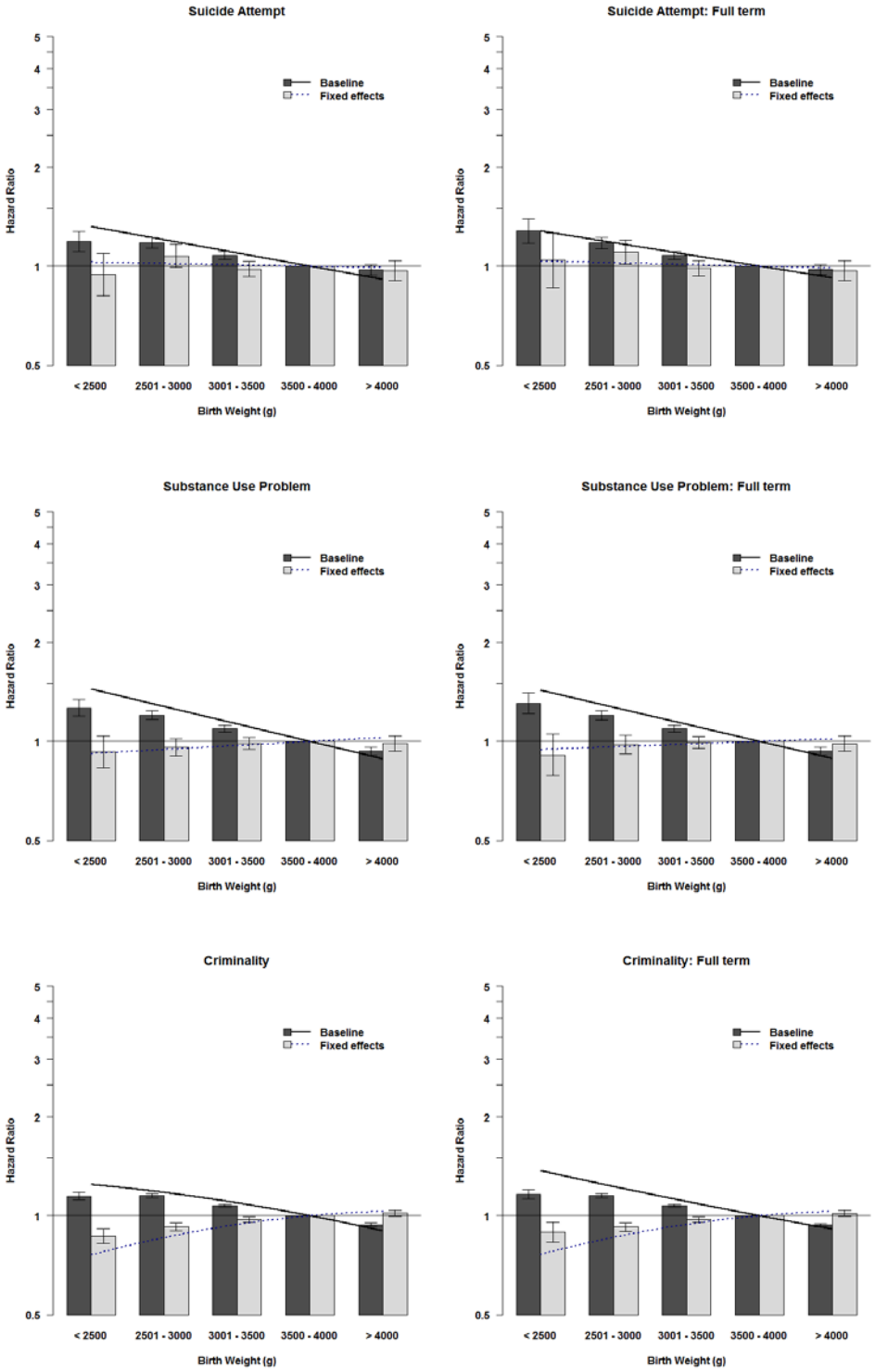


Figure DS1. Continued

Socioeconomic Outcomes

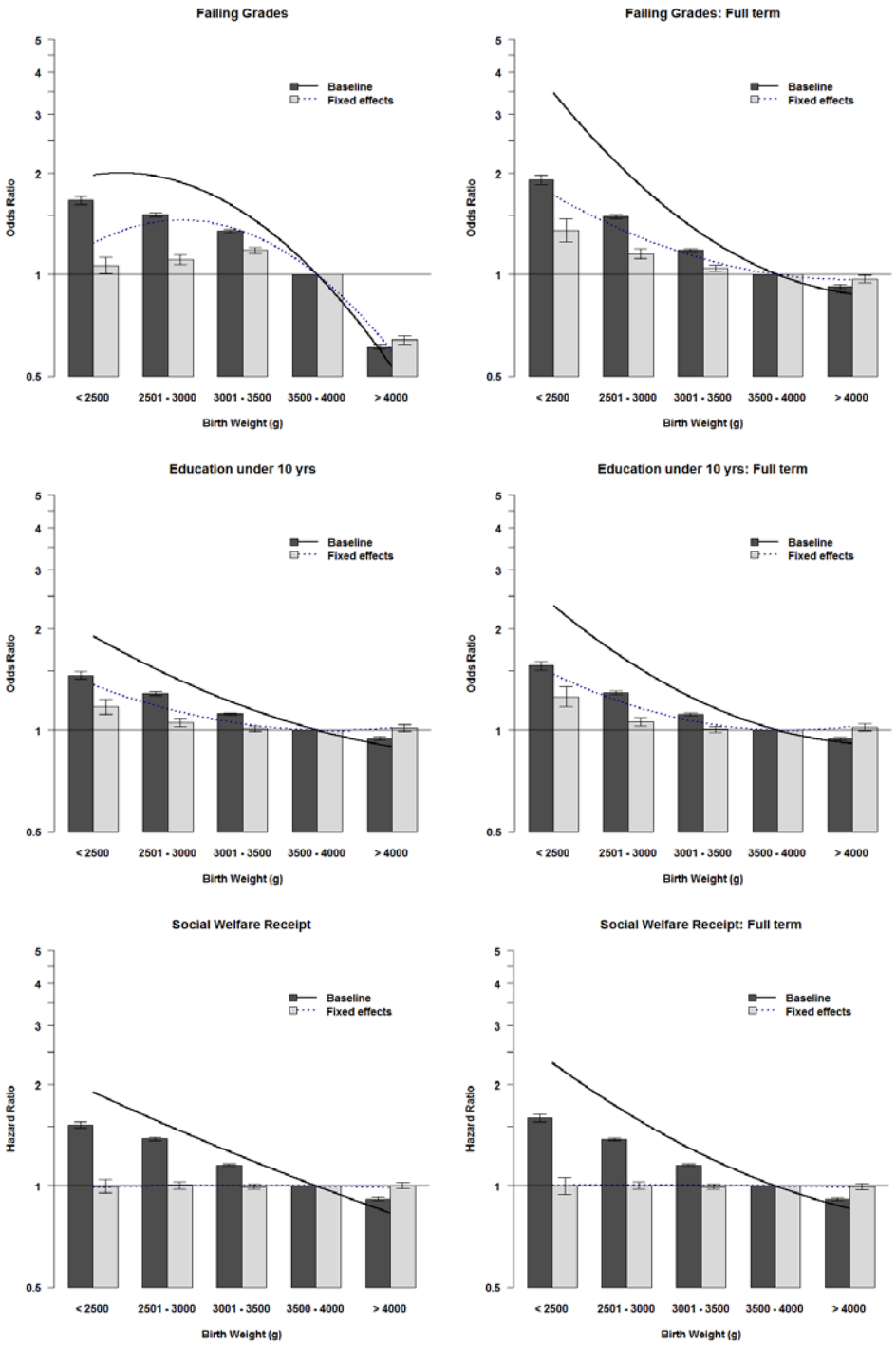


Figure DS2. Continuous and ordinal baseline and fixed effects figures predicting low income and higher education.

We predicted two additional socioeconomic outcomes to examine if we could provide converging evidence with the outcomes predicted in the main analyses. We predicted *low income*, from the LISA database, defined as being in the lowest 20th percentile income bracket for 2 consecutive years. *Higher education* was defined as three or more years of postsecondary education and was gathered from the Education Register. Only offspring born 1973-1983, whose age made it possible to achieve this level of education, were included in this sample.

As can be seen in Figure DS2, below, the baseline and fixed effects findings for these outcomes support those presented in the main analyses; lower birth weight is associated with increased odds of *Low Income* only in baseline analyses. Similar to *Social Welfare Receipt*, the relation was fully attenuated following fixed effects modeling. Low birth weight was also found to decrease the odds of achieving a *Higher Education*, and similar to *Failing Grades* and *Education Under 10 years*, this association was consistent in the fixed effects model.

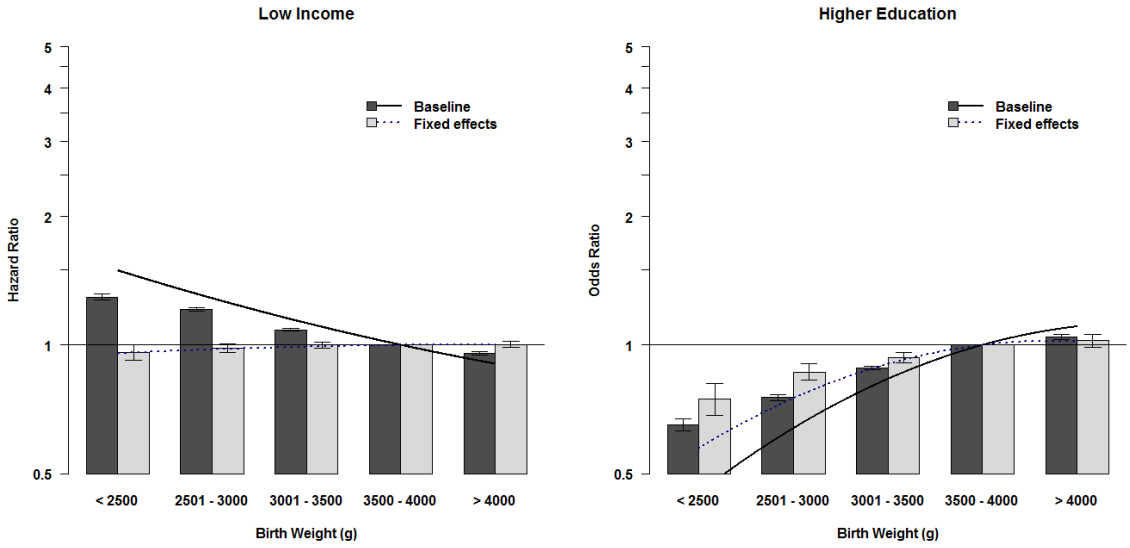


Figure DS3. Comparison of baseline parameter estimates when estimating separately for (a) families with more than one child and (b) only one child.

Sibling-comparison studies assume that findings from families with multiple offspring generalize to families with only one offspring. Therefore, the interpretation of the sibling-comparison results could be confounded if the population-based associations were different in offspring who had siblings than in those that are only children. If systematic magnitude differences are found between offspring with siblings and only children, then the reduction or increase in association magnitude found in the fixed effects models may be due to alternate explanations.

To help assess whether a bias was introduced by analyzing families with multiple offspring, we estimated the population-based estimates between birth weight and offspring outcomes in (a) offspring without siblings and (b) offspring with siblings. Each figure below presents these two baseline models. One model (grey bars with 95% confidence intervals) estimated on the sub-sample of offspring from families with only one offspring within the dataset. The second model (white bars with 95% confidence intervals) was estimated on the sub-sample of offspring from families with more than one child.

Figures DS3 show that the baseline associations are comparable for the two sub-samples of offspring. The figures also suggest that differences between the sub-samples do not account for differences in the sibling-comparison estimates as compared with the population estimates presented in the main paper. Across outcomes, associations in the two sub-samples are in the same direction and the magnitudes of association greatly overlap. Additionally, we found no pattern where magnitudes were always larger in one sub-sample. Overall, this sensitivity analysis suggests that the sibling-comparison results that showed changes in magnitude from the population analyses are not due to different population-based estimates in offspring with siblings than in offspring who are only children.

Figure DS3. Psychiatric Problems

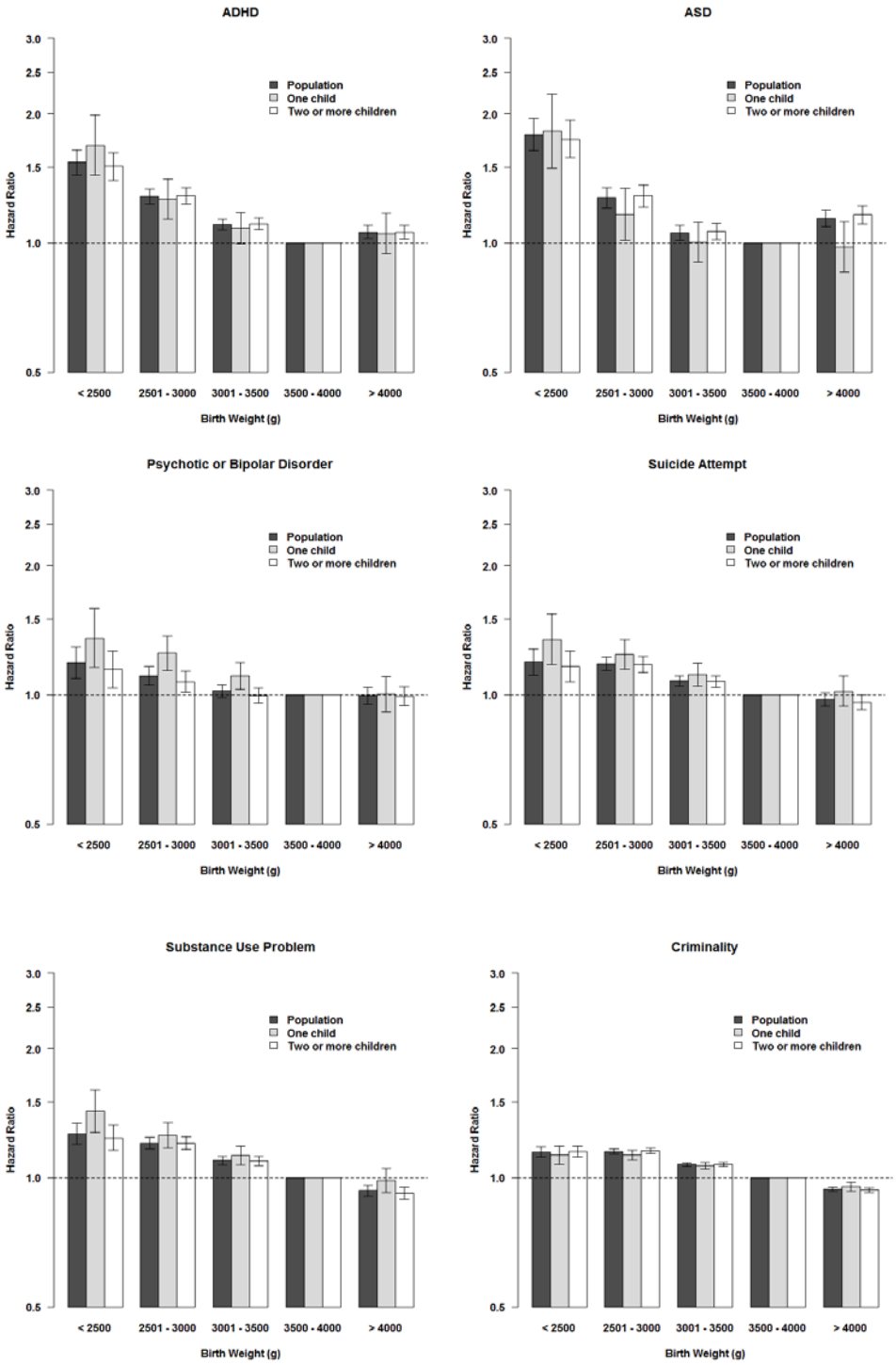


Figure DS3 Continued. Socioeconomic Outcomes

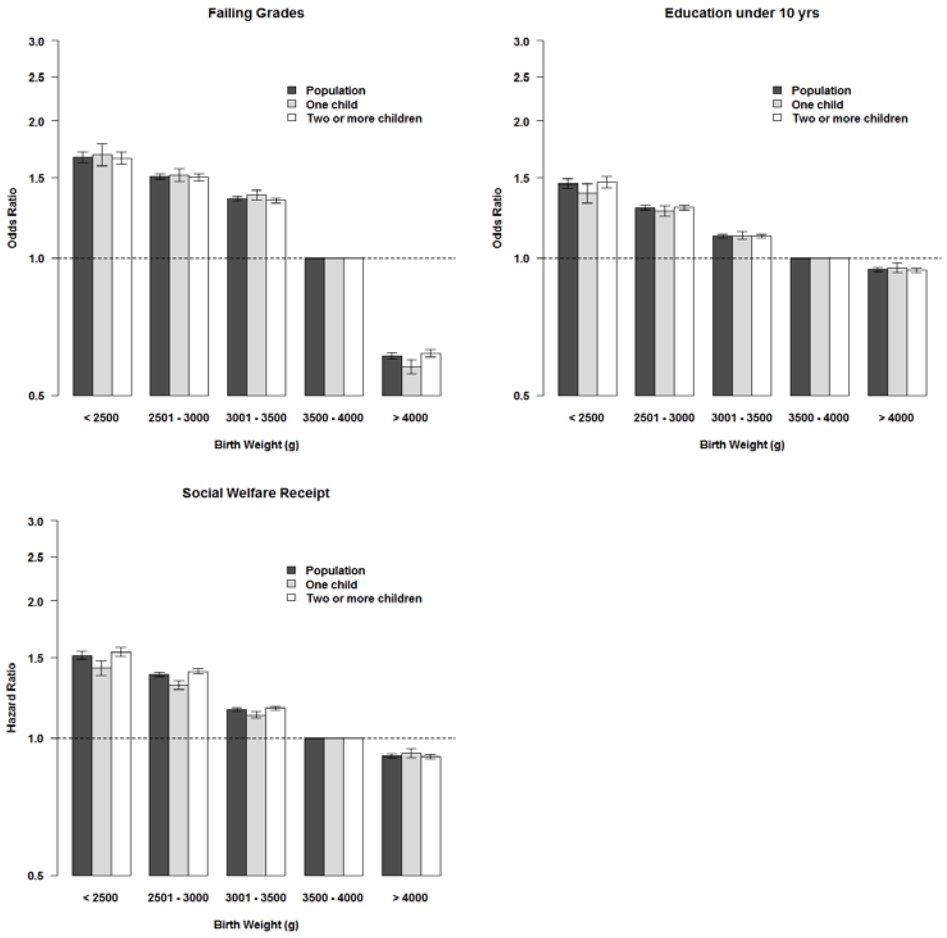


Figure DS4. Fixed effects parameter estimates for cousin comparisons.

To disentangle the source of possible confounding between birth weight and outcome, we conducted another test that utilizes a population that varies in their genetic relatedness. This was important because of inherent assumptions of the sibling-comparison approach, but also because individual genetic factors account for some variability in birth weight [10, 11]. While siblings share 50% of their genetic makeup on average, cousins, share 12.5 % of their genetic makeup on average. Therefore, we examined if the degree to which individuals share genetic risk moderates the association between birth weight and outcomes. If the associations are smaller when comparing relatives that share more genetic background (i.e., sibling associations are smaller than cousin associations), then genetic confounding is implicated. If the associations between birth weight and outcome are found to be the same magnitude across all relative groups, results may instead suggest the importance of environmental confounds. Although cousin versus sibling comparisons cannot by itself support or refute genetic confounding because these groups can also vary on their environmental risk "relatedness", when these results are combined with results from cousin-comparison designs that also vary in the degree of genetic relatedness, more evidence is gathered.

Figure DS4 show baseline, sibling- and cousin-comparison fixed effect (FE) results. Analyses were performed via stratification on the maternal grandmother of the target child. Overall, results for ASD, psychotic or bipolar disorder, and education under 10 years, support the findings that associations are consistent with a causal inference. For suicide attempt, substance use problem, and criminality, sibling and cousin comparisons similarly fully attenuate the association present in the population baseline analysis. For failing grades, social welfare receipt, it can be noted that the magnitudes of association are greater for cousin comparisons than for sibling comparisons, suggesting genetic confounding.

Figure DS4. Comparison of population baseline, fixed effects sibling-comparison, and fixed effect cousin-comparison for psychiatric morbidity and socioeconomic outcomes.

Psychiatric Problems

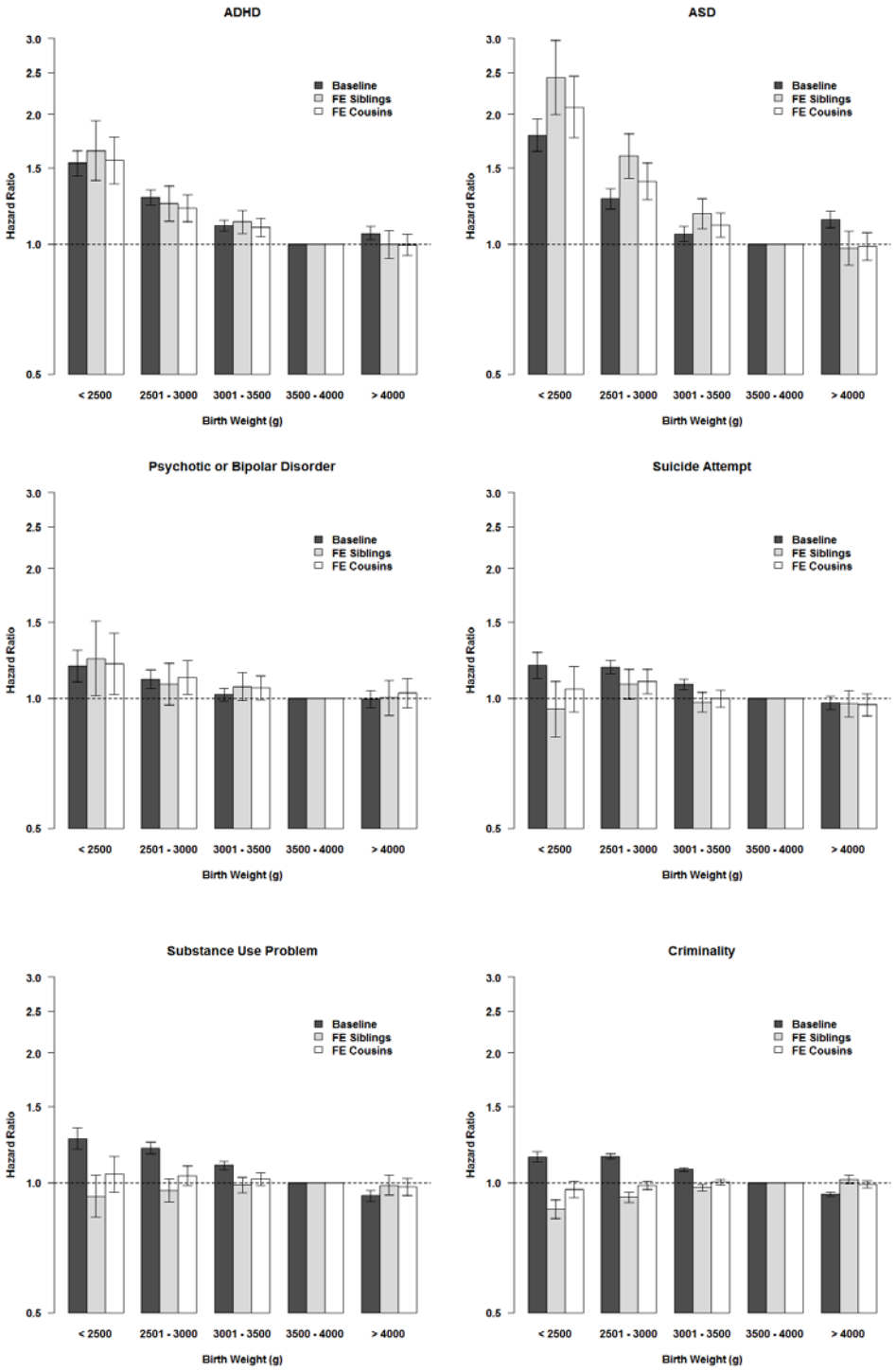
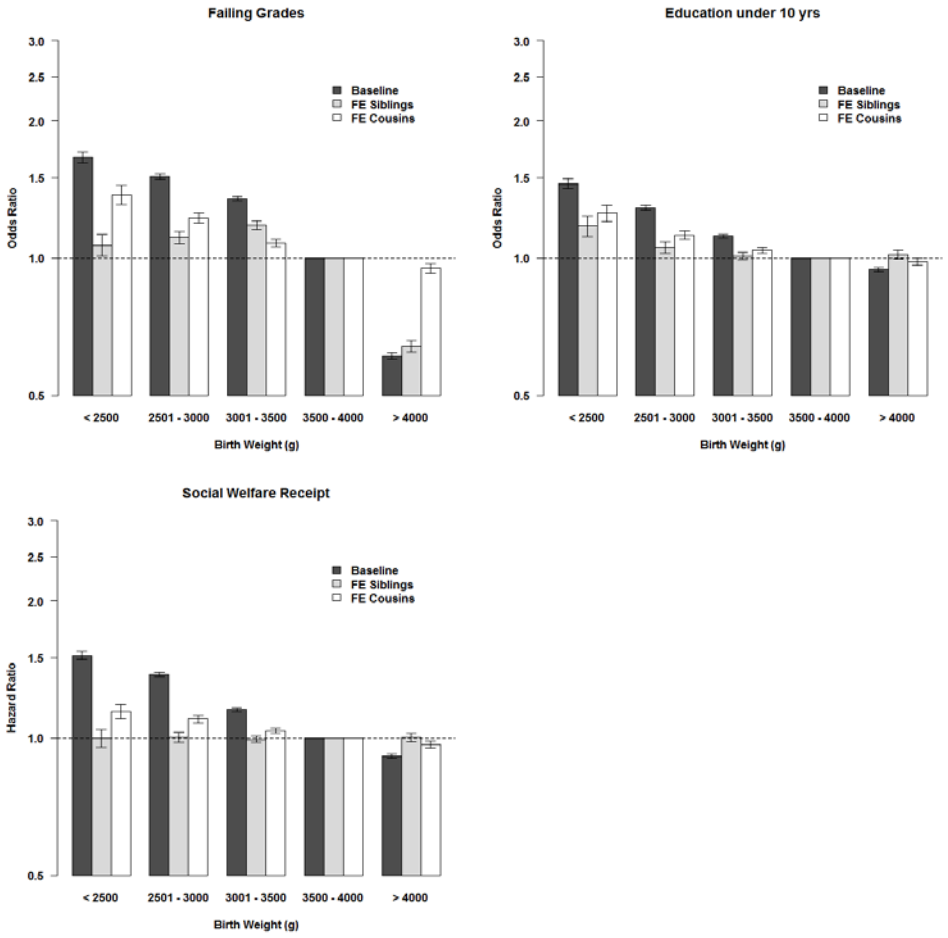


Figure DS4 continued. Socioeconomic Outcomes



Additional references

1. Centre for Epidemiology, *The Swedish Medical Birth Register - A summary of content and quality.* 2003.
2. Cnattingius, S., et al., *A quality study of a medical birth registry.* Scandinavian Journal of Social Medicine, 1990. **18**(2): p. 143-148.
3. Statistics Sweden, *Multi-generation register 2005 - A description of contents and quality*, 2006, Orebro: Statistics Sweden.
4. Centre for Epidemiology, *The Swedish hospital discharge register http:www.socialstyrelsen.se/en/statistics/statsbysubject/the+swedish+hospital+discharge+register.htm.* 2005.
5. World Health Organization, *International Classification of Diseases.* 1992.
6. Fazel, S. and M. Grann, *The population impact of severe mental illness on violent crime.* American Journal of Psychiatry, 2006. **163**(8): p. 1397-1403.
7. Swedish National Agency for Education, *http://www.skolverket.se/*.
8. LISA database, *http://www.scb.se/pages/list_257743.aspx*.
9. Burnham, K.P. and D.R. Anderson, *Model selection and multimodel inference: a practical information-theoretical approach*. 2nd ed2002, New York: Springer.
10. Lunde, A., et al., *Genetic and environmental infuences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data.* American Journal of Epidemiology, 2007. **165**: p. 734-741.
11. Svensson, A.C., et al., *Familial aggregation of small-for-gestational-age births: The importance of fetal genetic effects.* American Journal of Obstetrics and Gynecology, 2006. **194**(2): p. 475-479.