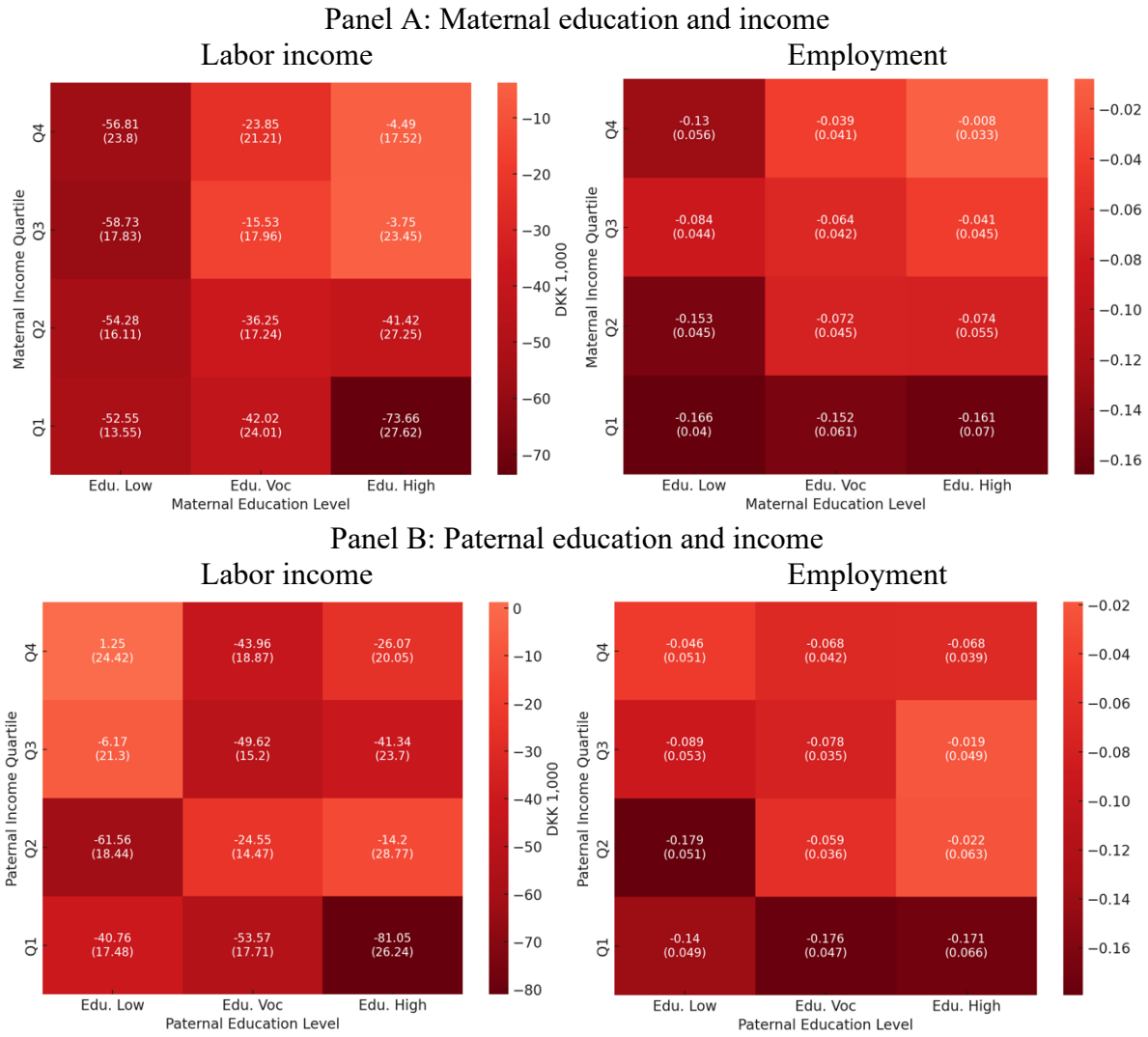
Online Appendix

Childhood Health Shocks and the Intergenerational Transmission of Inequality

BY TINE M. ERIKSEN, AMANDA P. GAULKE, JANNET SVENSSON, NIELS SKIPPER AND PETER R. THINGHOLM

Appendix A: Health gradients by parental characteristics

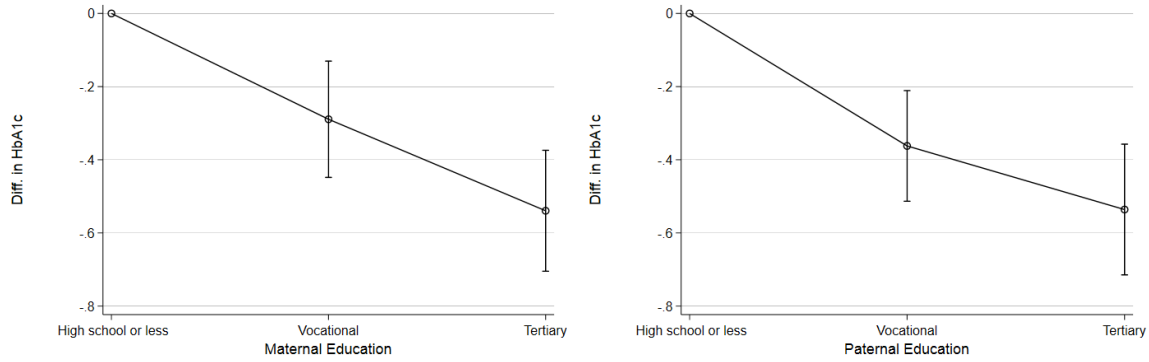
Figure A1: Treatment effects on employment outcomes by parental education and income



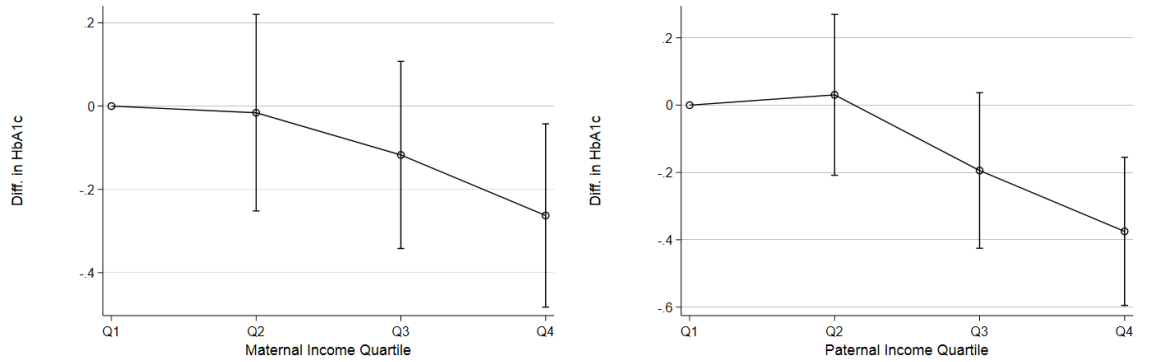
Notes: Panel A shows group average treatment effects by maternal income and education for labor income and employment. Panel B shows the corresponding effects by paternal income and education. Standard errors are reported in parentheses.

Figure A2: Difference in HbA_{1C} for People with T1D by parental characteristics

Panel A: Difference by parental education



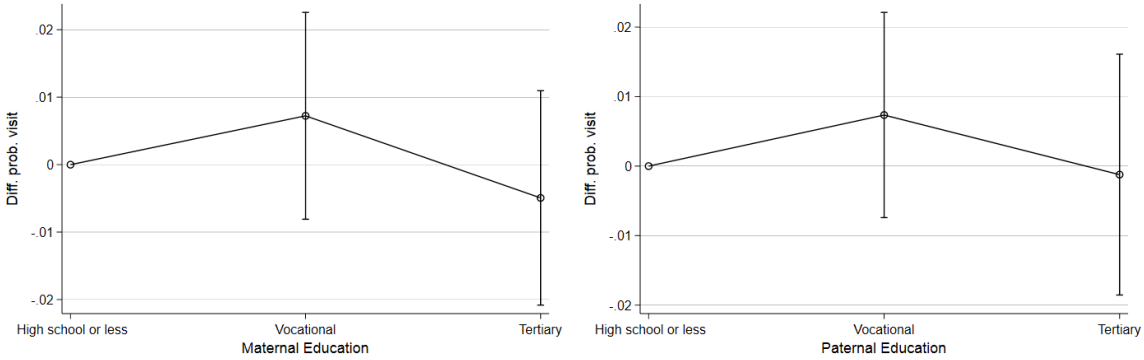
Panel B: Difference by parental income quartile



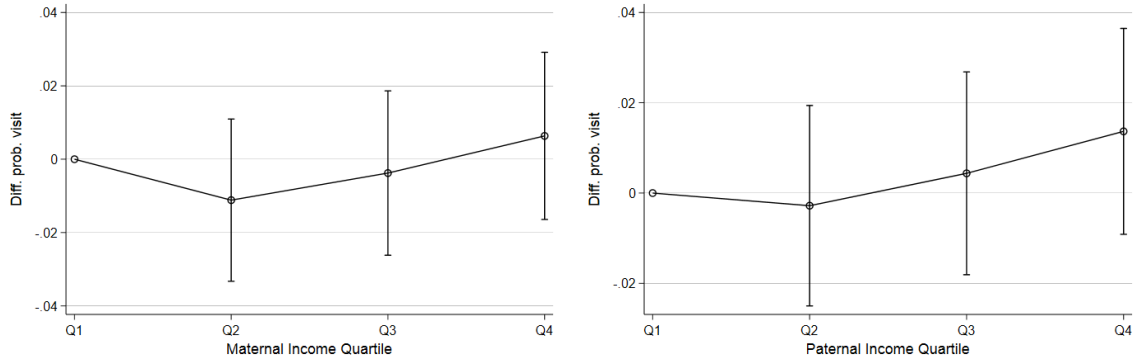
Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. Glycated hemoglobin (HbA_{1c}) is a measure of how well the glucose levels are managed with lower values indicating better disease management. Mean differences relative to the reference group (Panel A: High school or less; Panel B; Q1) are reported with 95% CI. The outcome (HbA1c) mean is 8.2.

Figure A3: Difference in probability of receiving ambulatory care related to T1D by parental characteristics

Panel A: Difference by parental education



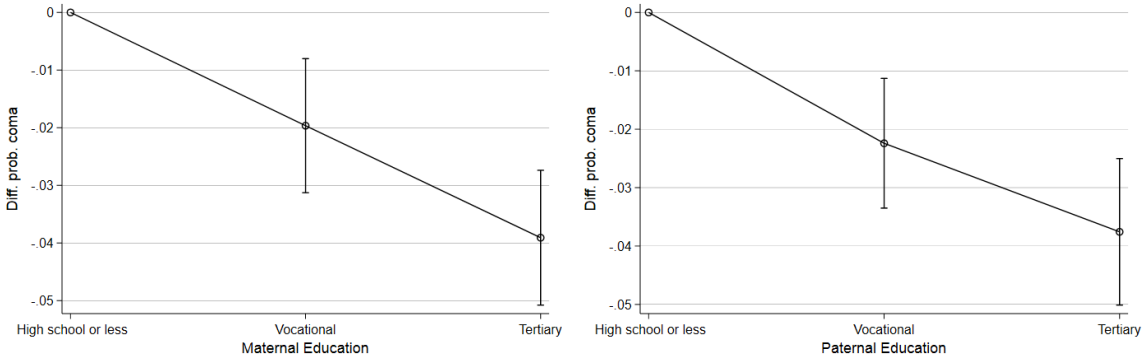
Panel B: Difference by parental income quartile



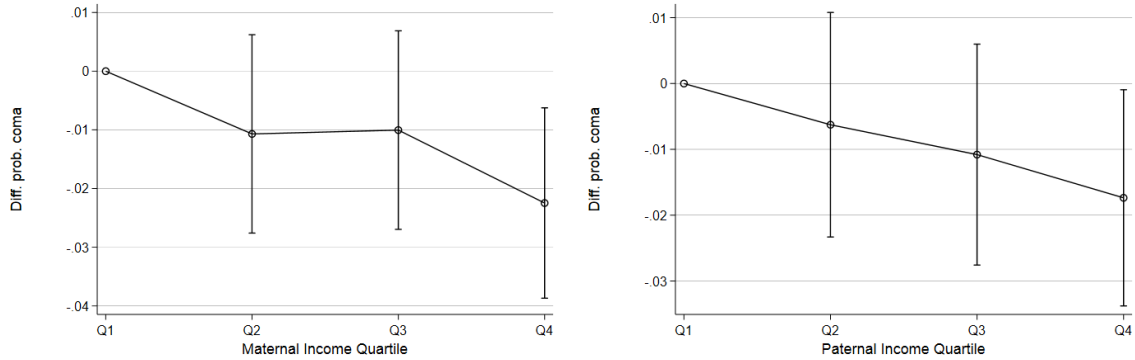
Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. The outcome is the probability of receiving specialized ambulatory care. Mean differences relative to the reference group (Panel A: High school or less; Panel B; Q1) are reported with 95% CI. The outcome mean (Probability of ambulatory visit) is 0.77

Figure A4: Difference in probability of hospital admission with diabetes related acute conditions by parental characteristics

Panel A: Difference by parental education



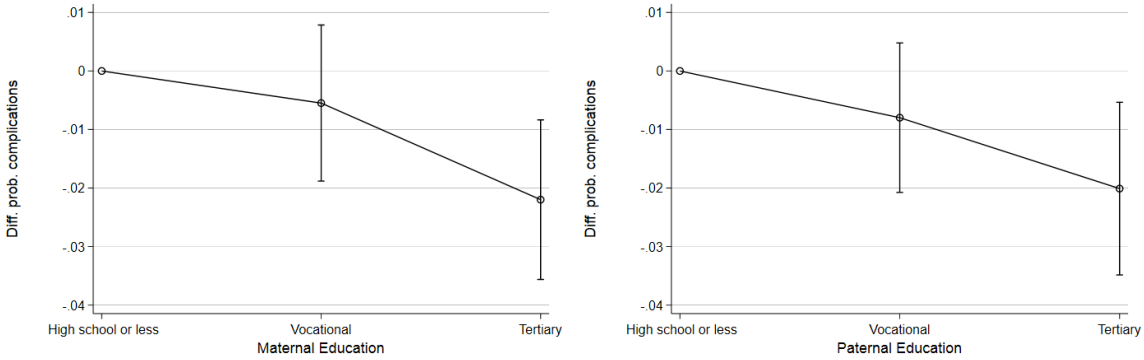
Panel B: Difference by parental income quartile



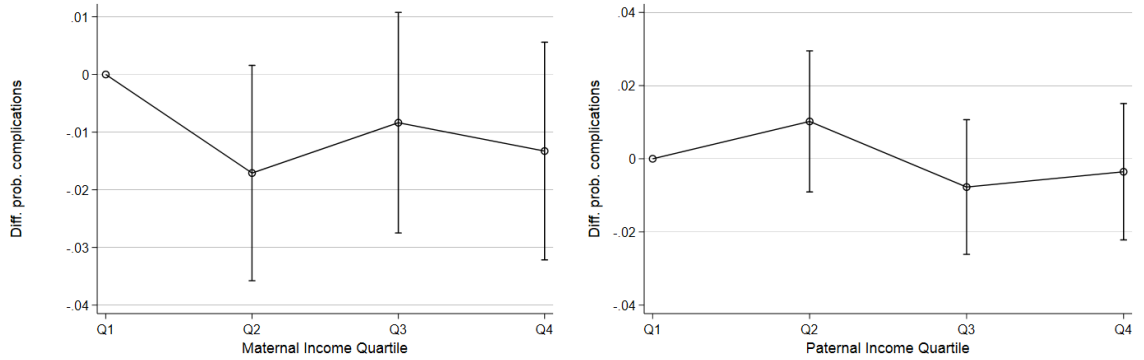
Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. The outcome is the probability of having been admitted to the hospital for diabetes related acute conditions (diabetic ketoacidosis or hypoglycemic coma). Mean differences relative to the comparison group (Panel A: High school or less; Panel B; Q1) are reported with 95% CI. The outcome mean is 0.36.

Figure A5: Difference in probability of diabetes related complications by parental characteristics

Panel A: Difference by parental education



Panel B: Difference by parental income quartile



Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. The outcome is the probability of having been diagnosed with any late complication by age 30. Mean differences relative to the reference group (Panel A: High school or less; Panel B; Q1) are reported with 95% CI. The outcome mean (probability of diabetes related complications) is 0.49.

Appendix B: Tests for heterogeneity

In this appendix, we conduct a series of tests to assess whether the heterogeneity we observe is in fact essential heterogeneity or merely sampling variation. As a first test, we return to the quartiles of predicted treatment effects outlined in section 5.1. With the data split into groups (quartiles) of predicted treatment effects, we calculate the group average treatment effect (GATE). Intuitively, if our estimated model is successful in detecting treatment heterogeneity, we should observe meaningful differences in the average treatment effect across the groups.[1] The results are reported in Figure B1. Among the individuals predicted to be in Q1, the average treatment effect (S.E) is DKK -42,694.7 (9,160.7) vs. DKK -9,027.5 (9,365.2) in Q4 for labor market income. For the outcome 'employment', we have -15.12 (2.05) pp. and -8.68 (1.77) pp. for Q1 and Q4, respectively. Even though the confidence intervals for Q1 and Q4 are slightly overlapping for both outcomes in Figure B1, the difference in treatment effects between Q1 and Q4 is statistically significant at the 5% level in a formal test.

The differences are also significant in an economic sense. For the labor market income, the treatment effect is four times as large in Q1 vs. Q4. The difference corresponds to the overall average treatment effect reported in Table 3. Regarding 'employment', the difference is a factor 2, or 75% of the average treatment effect reported in Table 3. We note that we do not see a monotonic increase in the GATEs for the outcome 'employment', as the estimated GATE is numerically smallest in Q3.

As a second test of heterogeneity, we turn to the best linear predictor (BLP) test formalized in Chernozhukov et al. (2025), and follow the implementation outlined in Athey and Wager (2019). We estimate the best linear predictor of the true CATE based on the estimated $\hat{\tau}$ using a transformed outcome approach. Specifically, we estimate the following model:

$$Y_i - \hat{\mu}(X_i) = \alpha\bar{\tau}\big(D_i - \hat{e}(X_i)\big) + \beta(\hat{\tau}(X_i) - \bar{\tau})\big(D_i - \hat{e}(X_i)\big) + \epsilon \qquad \text{(B1)}$$

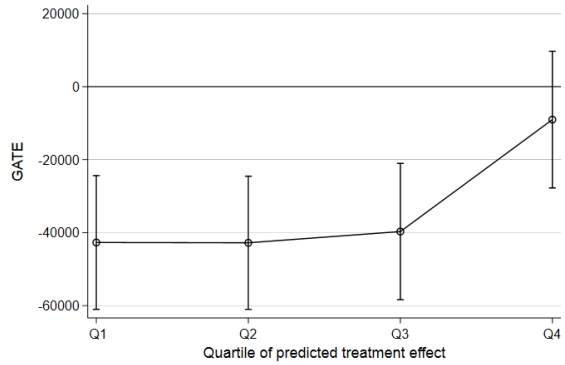
Where $\hat{\mu}(X_i)$ is from (3), $\hat{\tau}(X_i)$ is the estimated treatment effect from (6), $e(x) = \mathbb{P}[D_i|X_i=\text{x}]$ (i.e., the propensity score), and $\bar{\tau}$ is the mean of the out-of-bag treatment effects. $\hat{\mu}$, $\hat{\tau}$ and $\hat{e}$ are estimated using out-of-bag prediction. The estimated α and β coefficients are informative about the performance of the estimated CATEs. α measures if the average

[1] This exercise can also be performed using a doubly robust approach with augmented inverse probability weighting (AIPW) scores. AIPW scores are preferable in observational studies with unconfoundedness when the covariates do not balance across the treatment and control group. We get quantitively and qualitatively similar results using this approach (results available upon request).

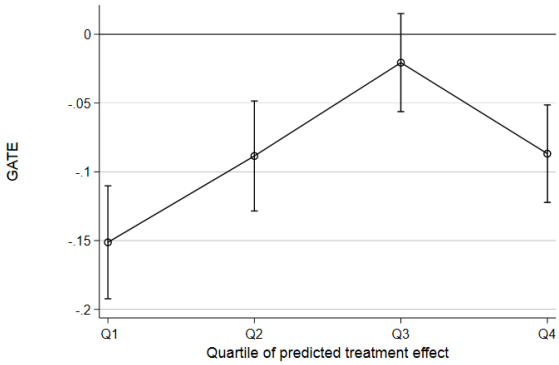
prediction produced by the causal forest is correct, with a value of 1 indicating that it is. β measures if the model adequately captures the underlying heterogeneity. If $\beta = 1$, the predictions from the causal forest adequately capture the underlying heterogeneity. Put in another way, β measures how well the CATE predictions covary with the true CATE (Chernozhukov et al., 2025). When β is positive, and significantly greater than 0, we can reject the null of no heterogeneity. The results from estimating (B1) on the outcomes are reported in Table B1. The top panel (A) shows the estimates for labor market income. The estimate of β is 0.65 and significantly greater than 0 at the 5% level and α is estimated to be 1.01. While the estimate of β is different from 1, which indicates that the estimated CATEs do not perfectly correlate with the true (unobserved) CATEs, the fact that it is different from 0 is clear evidence of the presence of essential heterogeneity. For the employment outcome, β is 1.08 and statistically significantly greater than 0 at all conventional levels of significance and α is estimated to be 1.00. This provides strong support of the presence of treatment effect heterogeneity.

Figure B1: Group Average Treatment Effects (GATEs) by predicted quartile of CATE

Panel A: Labor market income



Panel B: Employment



Notes: This figure presents the group average treatment effect by the predicted CATE for the two outcomes. Mean and 95% CI.

Table B1: Best linear predictor test of heterogeneity

		Coef.	S.E.	t statistic	Pr(>t)
Panel A: Labor market income					
	α	1.01	0.13	7.75	<0.0001
	β	0.65	0.39	1.66	0.048
Panel B: Employment					
	α	1.00	0.12	8.06	<0.0001
	β	1.08	0.35	3.09	0.001

Notes: Model parameters are estimated using the 'test_calibration' function part of the GRF package in R. The test is a one-sided test as we are testing against non-zero values of beta.