I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16447

Childhood Health Shocks and the Intergenerational Transmission of Inequality

Tine M. Eriksen
Amanda P. Gaulke
Jannet Svensson
Niels Skipper
Peter R. Thingholm

SEPTEMBER 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16447

Childhood Health Shocks and the Intergenerational Transmission of Inequality

Tine M. Eriksen
VIVE and IZA

Amanda P. Gaulke
Kansas State University and IZA

Jannet Svensson
Steno Diabetes Center Copenhagen

Niels Skipper
Aarhus University

Peter R. Thingholm
Aarhus University

SEPTEMBER 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Childhood Health Shocks and the Intergenerational Transmission of Inequality*

We examine the role of health shocks in childhood and parental background in transmitting intergenerational inequality. We use Danish administrative registry data (a setting with universal access to health care) and the quasi-random onset of Type 1 Diabetes in childhood to document substantial penalties in adult employment and labor market income at age 30. We document wide disparities in treatment effects and show that high-socioeconomic parents mitigate the adverse impacts of the health shock. This gradient is partly driven by differential impacts on health and human capital across the socioeconomic distribution. Maternal educational attainment matters for adoption of new and more advanced treatment regimens.

JEL Classification: I12, I14, I24

Keywords: intergenerational transmission of inequality, childhood health shocks, labor market outcomes

Corresponding author:
Amanda Gaulke
Kansas State University
327 Waters Hall
1603 Old Claflin Place
Manhattan
KS 66506
USA

E-mail: gaulke@ksu.edu

* This research was funded by Independent Research Fund Denmark grant 8019-00055B. The study was approved by the Danish Data Protection Agency. The Danish Registry of Childhood and Adolescent Diabetes (DanDiabKids) was approved by the Danish Health Data Authority (file No. 14/915976). These approvals constitute the necessary legal requirements, and informed consent is not required.

1. Introduction

Health is considered an important driver of economic inequality and of the intergenerational transmission of poverty (see e.g., Aizer and Currie, 2014; Bowles and Gintis, 2002). Both Case, Lubotsky, and Paxson (2002) and Currie and Stabile (2003) establish that the income-health gradient, previously documented among adults, also exists in early childhood. These results suggest that childhood health shocks may have an important role in the transmission if inequality. Moreover, Currie and Stabile (2003) suggest that the social gradient in child health disparities is driven by low-socioeconomic-status children experiencing more negative health shocks. However, questions remain about the role parents play in mitigating or exacerbating the effect of childhood health shocks. For example, Almond, Currie, and Duque (2018) state that "a greater understanding of the way that shocks and disadvantage interact, and of the role of parents in responding to them, is highly desirable." In this paper, we provide empirical evidence showing a clear socioeconomic gradient in effects from the same childhood health shock on later labor market outcomes. This suggests that differential impact by family background is also an important mechanism in driving the intergenerational transmission of inequality.

We examine the intergenerational transmission of inequality by studying how households across the income and education distribution respond to a uniform, randomly distributed health shock. The health shock we examine is Type 1 Diabetes (T1D), one of the most common chronic physical health conditions in childhood (Imperatore et al., 2018). We utilize rich administrative registry data from Denmark and the causal forest methodology. This approach allows us to study whether the consequences of a health shock differ between children of different socioeconomic status while addressing endogeneity in health shocks and holding access to health care constant across the socioeconomic distribution.

We provide extensive evidence that verifies findings from the medical literature that the onset of T1D is orthogonal to socioeconomic position and to underlying health status. Since T1D represents a near-random health shock and because the entire population is equally at risk, we can study the effects over the entire socioeconomic distribution. Consistent with previous literature that finds negative average treatment effects of T1D (see, e.g., Persson et al., 2016), we find negative average treatment effects on labor market income at age 30 (a penalty of DKK 32,422 or 14 percent) and negative impacts on employment at age 30 (an 8.7 percentage point (pp.) decline or 11% decline). These negative average treatment effects are likely explained by increased school absenteeism (Thingholm et al., 2020) and worse educational outcomes at age 16 and age 20 (Lindkvist et al., 2021).

We document considerable heterogeneity in effects that are missed by solely focusing on the average treatment effect by estimating the full distribution of conditional average treatment effects (CATEs) using the causal forest methodology (Wager and Athey, 2018). The estimates from the causal forest reveal significant heterogeneity. Ranked by their predicted treatment effects, individuals in the bottom quartile (largest income penalties) experience a penalty of DKK 42,694, while for the top quartile (smallest income penalties), the penalty is only DKK 9,027. The difference between the two groups is noteworthy as it is comparable to the average treatment effect of DKK 32,422. When studying the extensive margin, we find largely the same pattern: Individuals in the bottom quartile of predicted effects are 15.12 pp. less likely to be employed, while among those in the top quartile the corresponding penalty is only 8.68 pp. Again, this difference is comparable in terms of magnitude to the overall average treatment effect.

Not only do we document substantial heterogeneity in treatment effects, but we also find evidence that high-socioeconomic (SES) parents can mitigate the negative impacts. Having a parent who is in a higher income quartile leads to smaller penalties in labor market outcomes.

This pattern of effects is similar for both maternal and paternal income quartiles. Having a more educated mother also leads to smaller penalties in labor market outcomes. The Danish institutional setting implies that we can rule out that our results are driven by variation in whether the child or family is covered by insurance, variation in financial costs of treatment, or variation in access to diabetes-specific medical care. Thus, some parents can mitigate the impact of a childhood health shock, even in a setting with free and universal access to health care.

We provide suggestive evidence that disease management partly explains the SES gradient in impacts on labor market outcomes. We show that individuals with fathers or mothers in lower income quartiles experience worse T1D-related health outcomes in adulthood. This indicates that disparities in disease management by SES in childhood documented in Nielsen et al. (2019) translate to disparities in adulthood as well. We also provide evidence that maternal education matters for whether children are using new diabetes treatments. Specifically, we find that children of mothers with higher education levels have a higher rate of adopting a newer and more advanced treatment regimen known as multiple daily injections. Despite these families having universal access to health care, we find evidence that parental socioeconomic characteristics mitigate impacts of childhood health shocks. We also find that reduced accumulation of human capital partly explains the differences in labor market penalties.

Ultimately, we contribute to several strands of literature. First, we contribute to an extensive literature on health and the intergenerational transmission of inequality by showing low and high SES children are impacted differentially by the same health shock. Second, we contribute to a large literature on adverse health events in childhood and parental responses. While the related literature finds variation in direct measures of parental investments in response to health shocks (see, e.g., Hsin, 2012; Restrepo, 2016; and Guo and Zhang, 2021),

we show variation in adult outcomes by parental characteristics. Third, we add to a new and developing literature on using machine learning techniques in causal inference (Wager and Athey, 2018; Athey and Wager, 2019).

The rest of the paper is structured as follows. Section 2 provides background on T1D and the institutional setting in Denmark. Section 3 describes the Danish administrative registry data and provides descriptive statistics. In section 4 we present our empirical strategy. Section 5 discusses the results, and in section 6 we provide evidence for a range of potential mechanisms. Section 7 summarizes and concludes.

2. Background

2.1 T1D

T1D[1] is a chronic health condition in which the immune system kills insulin-producing beta cells in the pancreas. Insulin is a metabolic hormone needed to allow sugar (glucose) to enter the cells in the body to produce energy. In healthy individuals, the glucose concentration in the bloodstream is constantly kept within a narrow interval. This is done by secreting insulin when glucose levels rise, i.e., when eating and drinking (primarily carbohydrates), to lower glucose levels, and by releasing sugar (primarily from the liver) outside of meals (or when fasting). In individuals with T1D, this endogenous insulin supply is absent.

Type 1 diabetes is—in contrast to the more common type 2 diabetes—not related to lifestyle. It most often presents in childhood, and it is, second only to asthma, the most common chronic physical health condition in children and adolescents in most of the Western

[1] T1D was formerly known as juvenile diabetes or insulin-dependent diabetes.

world.[2] The exact cause of the disease is unknown. While there is ongoing research to determine the origin, Regnell and Lernmark (2017) note, "The aetiology of beta cell autoimmunity is still unclear". Genetics play a smaller role than for other chronic health conditions as Pociot and Lernmark (2016) note there is only a three percent risk of T1D in children of mothers with the condition. Whether maternal and paternal socioeconomic status (income and education) are effective predictors of T1D onset in Denmark was tested using full population data by Prætorius, Urhoj, and Andersen (2022). They conclude there is no socioeconomic gradient in T1D onset in children and adolescents. The condition is characterized by a rapid onset in children and adolescents and affected individuals will not go undiagnosed.[3]

In this paper, we verify that the onset of T1D is sudden and that children diagnosed with T1D do not differ from children without T1D in terms of underlying health prior to diagnosis or in terms of parental characteristics. Later in the paper we will show there are strikingly similar pre-diagnosis trends regarding hospital admissions, visits to the general practitioner (primary care physician), and the probability of pharmacy claims between diagnosed and comparison children. We also find no differences in parental characteristics at birth between the treatment and comparison groups (discussed in more detail in the data section).

These findings corroborate previous findings by Thingholm et al. (2020), who use an event-study design to compare school absenteeism among children who were eventually

[2] In the United States, around 1.25 million children and adults live with T1D, with an estimated annual cost of US$ 14.4 billion (Tao et al. (2010)). In Denmark, approximately 32,000 people are diagnosed with T1D, of which 3,500 are children.

[3] Before insulin treatment was discovered in the early 20th century, a patient newly diagnosed with T1D had an average life expectancy of around 2 years (Hakim et al., 2013). Thus, our control group is unlikely to contain individuals with undiagnosed T1D.

diagnosed with T1D to matched comparison children (who are then assigned a pseudo diagnosis date) based on sex and date of birth (Figure 1). The graph is informative in several ways that are key to our identifying assumption of T1D constituting a near-random health shock. It illustrates that the child's symptoms (severe enough to affect school absenteeism) are only present in the months very close to diagnosis (four-months prior is the first month with a significant difference). Related, there seem to be no differences in the underlying health of children who develop T1D based on their pre-diagnosis levels of school absenteeism (i.e., more than four months before onset). Further, Figure 1 shows that after the onset of diabetes, these children have systematically and significantly more school absenteeism (roughly 50% more than the comparison children). This points to the fact that something is now different for these children.

Consistent with previous literature finding parental income is not related to T1D onset, Eriksen et al. (2021) shows no differences in the wage level or trend prior to onset when comparing mothers and fathers who have a child with T1D with mothers and fathers who do not have a child with T1D. Further, Eriksen et al. (2021) show that children who are diagnosed with T1D had no significant differences in 5-minute APGAR scores compared with children who were not diagnosed and that both groups are equally likely to be of low birth weight.

Managing T1D is time consuming and complex. As there is no cure, the aim is to keep glucose levels as close to normal as possible by injecting insulin to counter the rises in glucose caused by the individual's intake of food and the endogenous release of sugar. The amount of insulin needed depends on several factors, for example, food intake[4] and exercise.

[4] The amount of insulin needed is typically calculated as an insulin-to-carbohydrate ratio, i.e., a specific amount of insulin is needed to match a certain amount of carbohydrates. However, this ratio is different from individual

In addition, numerous small blood samples from finger pricks or a continuous glucose monitor must be used to check glucose levels and adjust insulin dosages, as necessary.

Even with insulin treatment, individuals with T1D have chronically elevated glucose levels (known as hyperglycemia) compared with individuals without T1D. Hyperglycemia can lead to so-called late complications such as kidney disease/failure, blindness, amputation, and cardiovascular disease (DCCT, 1993; Nathan, 1993). At the other end of the spectrum, administering too much insulin causes low blood sugar (known as hypoglycemia), which can lead to seizures and coma. Thus, glucose levels need to be checked frequently, including during the night. Hypoglycemia is treated by ingesting a sugary snack or beverage. The constant alertness required to keep the blood sugar in check has been shown to negatively affect parental sleep (Cobry and Jaser, 2019; Pillar et al., 2003); thus, keeping the condition well-managed is taxing.

The impacts of T1D are not limited to health outcomes. Previous research has found that childhood onset results in an increase in school absenteeism of 50% compared with matched controls (Thingholm et al., 2020). Also, a T1D diagnosis leads to lower 9th grade exit exam GPAs, a higher relative risk of not completing 9th grade, and a higher risk of not having completed or being enrolled in upper secondary education by age 20 (Lindkvist et al., 2021). Thus, labor market outcomes are expected to be impacted through reduced human capital accumulation and poor health capital.

2.2 Institutional Setting

In Denmark the financial cost of managing T1D is low, especially compared with the costs faced in the United States. There is no cost for in-patient or out-patient care, and all medical

to individual and even varies over the time of day. The type of carbohydrate and the protein and fat content of the food has implications for the correct insulin dosage.

devices are provided free of charge. However, medicine is not free and thus insulin has an associated out-of-pocket cost. The median yearly out-of-pocket expenditure for insulin in 2016 was approximately $229. When a child is diagnosed with T1D, a caseworker visits the family at the hospital. The caseworker then determines if the family qualifies for a means-tested program where a low enough income results in any remaining out-of-pocket costs being reimbursed by the government. The family would also receive a subsidy to cover the expenses of glucose tablets and juice, which are essential for treating low blood sugar.

In Denmark extensive legislation exists to ensure equal rights both in- and outside of the labor market concerning aspects related to gender, health, age, etc. Workers are not obliged to disclose the diagnosis to their employer if it does not intervene with their ability to attend to their job, and in general, discrimination based on health conditions is illegal. This applies to both formal and informal discrimination; however, the latter can be hard to verify.

3 Data and descriptive statistics

The data are from several Danish registries, and we can match information across registries because each child is identified through a person-specific ID (equivalent to a social security number). All children diagnosed with T1D in Denmark since 1976 were identified through the National Patient Registry (*Landspatientregisteret*) through ICD-10 codes and before that ICD-8 codes. Since 1996, clinical information has been included in a national registry called DanDiabKids. This registry contains information on all Danish children and adolescents diagnosed with T1D who are seen at pediatric endocrinology clinics. Transfer to an adult endocrinology clinic usually happens at age 18, although some transfer at age 16. These data provide us with the exact date of diagnosis as well as the clinical characteristics of the children and adolescents collected at annual follow-up visits in pediatric endocrinology

clinics (out-patient clinics). To ensure that the children are old enough for us to examine their long-run labor market outcomes, we focus on the cohorts of children born from 1977 to 1987.

Through population registries within Statistics Denmark, the diagnosis data is augmented with information on demographic and socioeconomic characteristics. Further, children can be linked to their parents. We observe age, sex, educational attainment, income, and immigrant or descendant status for all individuals. For the children, we further observe birth order and the number of siblings. All data are recorded at the yearly level. We restrict the sample to include only native Danes to ensure that we have 1) accurate information about the child's medical history as we can only observe health records for care provided in Denmark and 2) information on parental background characteristics prior to the child's birth so we can compare the means of our treatment group and comparison group without worrying that characteristics are being impacted by the T1D diagnosis. This is important because Eriksen et al. (2021) find that a child's T1D diagnosis impacts parental labor supply and mental health. Using income measured in the year prior to the child's birth does result in some missing data (one third have missing data), since we can only go back to 1980 for income data and some cohorts were born before that. We have, however, tested for whether there are differences in missing income data across our treatment and comparison parents and find no statistical difference.

As the information on other background characteristics is only available from 1986 and onwards, we limit our data to children and adolescents who are diagnosed from 1986 to 2004. We identify 1,810 children who are diagnosed with T1D before they turn 18. The mean age at onset is 11.8 years (SD 3.8). For each child with T1D, we identify and match five comparison children (9,050 individuals) who had not been diagnosed with T1D before their 30^{th} birthday. The matching is performed on the exact date of birth and sex.

To lend further credibility to the assumption that T1D is a quasi-random health shock, we use an event study analysis to compare healthcare utilization among individuals with and without a T1D diagnosis each month from two years prior to diagnosis (or pseudo-diagnosis for the control group) through two years after diagnosis, see Figure 2. Panel A shows there are no differences in the probability of admission to a hospital prior to diagnosis, Panel B shows there is no difference in visits to the general practitioner prior to diagnosis, and Panel C shows there are no differences in pharmacy claims prior to diagnosis. In the month of diagnosis and the months after diagnosis, children with a T1D diagnosis are more likely to be admitted to the hospital, more likely to visit the general practitioner, and are more likely to have pharmacy claims than children not diagnosed with T1D. Note that in Denmark patients can receive more than one month of medicine and medical supplies at a time and thus do not have to renew their prescription every month. These results support the claim that there were no differences in underlying health for children who are diagnosed with T1D.

We next compare observable characteristics of the children and parents between the treatment and comparison groups; see Table 1.[5] We regress an indicator variable for T1D onset on the characteristics listed in Table 1 to see if we can predict which children will have childhood onset of T1D. Results are shown in Table 2. While there are some significant coefficients, this is to be expected given the number of coefficients we are estimating, and the joint F-test suggests no overall differences between the two groups. The lack of differences by family socioeconomic status is consistent with Prætorius, Urhoj, and Andersen (2022). This analysis supports the assumption that T1D onset is not related to socioeconomic status.

4 Empirical Strategy

[5] Statistics Denmark does not track race and thus we cannot test for racial differences.

We next present our empirical strategy to estimate the impact of a childhood health shock on adult labor market outcomes. To estimate the treatment effects, we rely on the quasi-random variation in the T1D diagnosis. Our analysis consists of two main parts. We first want to estimate the average impact of being diagnosed with T1D in childhood. Then we investigate the differential impact of being diagnosed with T1D in childhood across parental socioeconomic status by employing a causal forest methodology. The causal forest is well suited for our setting, as we have a relatively large number of covariates and no a priori theoretical ranking of their importance to the outcome.

Our observed data consists of (X_i, Y_i, D_i), where X_i are observed characteristics, Y_i is observed outcome, and D_i is an indicator equal to 1 if individual *i* is diagnosed with T1D in childhood. Letting $Y_i^D, D \in (0,1)$ denote potential outcomes if the individual is diagnosed with T1D ($D_i = 1$) and outcomes if the individual is not diagnosed ($D_i = 0$). The treatment effect we seek to estimate is defined as

$$\tau = E[Y_i^1 - Y_i^0].$$

In this paper, we leverage the independence between T1D and potential outcomes. $(Y_i^1, Y_i^0) \perp D_i$ to estimate the homogenous impact of diabetes on labor market outcomes at age 30 in a linear model using ordinary least squares (OLS)

$$Y_i = \tau D_i + \boldsymbol{X}_i\beta + \varepsilon_i \qquad (1)$$

where Y_i is either labor income or employment (defined as having a positive labor income) measured at age 30. D_i is an indicator for being diagnosed with T1D before age 18, and τ is our parameter of interest. $\boldsymbol{X}_i$ is a set of parental demographic covariates and ε_i is a normally distributed error term.

In the second part of our analysis, we are interested in quantifying the degree of heterogeneity in treatment effects and how they vary with parental background. Hence, we

estimate models where the treatment effect is allowed to be a function of a set of parental demographic covariates, X_i:

$$\tau(x) = E[Y_i^1 - Y_i^0 \,|X_i = \text{x }]$$

Heterogeneous treatment effects can be estimated by separately estimating (1) for each value of $\boldsymbol{X}_i$:

$$Y_i = \tau(x)D_i + \varepsilon_i \; if \; \boldsymbol{X}_i = x \qquad (2)$$

While this is a very flexible approach, as it does not impose any additional structure on the relationship between Y_i and D_i, there are several issues with this approach. First, increasing the dimension of $\boldsymbol{X}_i$ can lead to regions with very sparse density and, for continuous X_i's, perfectly subsetting is impossible. Second, no ex-ante guidance exists on how to subset the data or on how to discretize the continuous variables, and a researcher ex-ante must motivate the split and ex-post correctly for multiple hypothesis testing.

In this paper, we choose a different approach and estimate the $\tau(x)$ by causal forests. As described by Wager and Athey (2018), this methodology resembles a k-nearest neighbor method, where observations are divided into subsets or *leaves* using data-driven sample splits. In contrast to traditional methods, where a prespecified metric such as Euclidean distance is used, the closest observations to a given point are the observations that end up in the same leaf.

In practice, we implement the causal forest with honest trees (Athey and Imbens, 2016), grow 2,000 trees, and impose that at a minimum 5% of individuals in any leaf must be either treated or non-treated (α =0.05). We use out-of-bag prediction to predict the CATEs, which means *i*'s covariates were not used in estimating the model that was used to predict $\hat{\tau}$ for individual *i*. In Appendix B, we conduct a series of checks to verify that our estimated heterogeneity does not reflect sampling variation.

To gain insight into the characteristics that are associated with the treatment heterogeneity, we proceed by grouping the sample by quartiles of the predicted treatment effects. Although $\hat{\tau}$ for individual *i* is estimated using a leave-out approach (out-of-bag prediction), when ranking individuals *i* and *j* we further want to avoid using either individual in estimating the causal forest. Hence, we use *cross-fitting* where the data is split into *K*=10 folds of equal size. Then, in an iterative procedure, we estimate the causal forest using the *K*-1 folds and subsequently predict the CATEs for the left-out fold and rank the observations into quartiles based on the prediction. We then compare characteristics across the predicted quartiles.

5 Results

5.1 Homogenous treatment effects of T1D diagnosis in childhood

We begin by presenting evidence on the homogeneous treatment effects of being diagnosed with T1D during childhood. That is, we estimate equation (1) by OLS. Table 3 shows the results for our two main outcomes: Labor income (Panel A) and employment[6] (Panel B), both measured at age 30. In columns (1) to (4) the conditioning set is gradually expanded. Individuals who were diagnosed with T1D in childhood on average earn DKK 32,422 ($4,381) less (shown in column 4). This corresponds to a T1D income penalty of 14%. The likelihood of employment (ie, having positive labor income is 8.7 pp. lower in the T1D group, which corresponds to a relative difference of 11%. It is worth noting that adding more control variables does not affect the estimated treatment effects much but increases precision slightly.

[6] Defined as having any positive labor market income.

To rule out the possibility that these differences in outcomes are driven by specific events at age 30, we depict in Figure 3 how labor market income and the probability of employment evolve from age 18 to 30. Observing Panels A and B, we note an income difference between the two groups already at age 18. Although the difference in labor market income at age 18 is statistically insignificant, it corresponds roughly to an 8% difference in earnings. From age 19 and onwards, the difference remains statistically significant. The gap increases over time, both in absolute and relative terms: At age 30 the income difference is around DKK 33,000 which is a relative difference of 14%. On the extensive margin, there is a relatively stable difference in the probability of employment of 6-8 pp.; see Figure 3, Panel C. In summary, the findings from Figure 3 suggest that the diabetes penalty already exists at the time of labor market entry, resulting in lower labor market participation. The earnings penalty appears to grow over the life cycle.

We next put the results above into context of the related literature that estimates the impact of a childhood shock. It must be noted that whereas the related literature has largely focused on in-utero exposure and very early life exposure, this paper focuses on a health shock that can occur throughout childhood. The most comparable setting would be that in Gensowski, Nielsen, Nielsen, Rossin-Slater, and Wüst (2019), which uses the 1952 polio epidemic in Denmark as a childhood health shock to find positive impacts on affected individuals. Schwandt (2018) finds that in-utero maternal exposure to influenza in Denmark leads to reductions in adult earnings of 9%, which is 64% of our estimated effect of 14%. Nilsson (2017) uses variation in Swedish alcohol policy to find that individuals exposed to excessive alcohol in utero and born to mothers younger than 21 years old had 20% lower earnings, which is 1.4 times the magnitude of our effect size.

Related literature has also found effects of childhood shocks outside of Scandinavia, although comparisons with the United States may be less than ideal. Despite the differences

in settings, our results are similar in magnitude to Adhvaryu, Bednar, Molina, Nguyen, and Nyshadham (2020), who find that the rollout of iodized salt across the United States led to an 11% increase in income. Our results are also similar in magnitude to Smith (2009), who finds that controlling for education, which the author notes is negatively impacted by health, leads to income being 13% higher for individuals who report their health to be excellent or very good. Our results are larger in magnitude to those in Beach, Ferrie, Saavedra, and Troesken (2016), which finds that eliminating early life exposure to typhoid fever led to a nine percent increase in earnings using an IV strategy. Our results are also larger than Isen, Rossin-Slater, and Walker (2017), who find the Clean Air Act of 1970 led to a 1% increase in annual earnings for the affected cohorts. Our estimate for labor force participation is about 13 times as large as the estimates in Adhvaryu et al. (2020) and 3 times as large as in Saez (2021), which studies the reduction in exposure to pneumonia in infancy in Chile.

5.2 Differential impact of T1D diagnosis in childhood

To investigate the heterogeneity in treatment effects, we now turn to our causal forest analysis. In Figure 4, we graphically show the distribution of the estimated CATEs for both main outcomes. Panel A shows the distribution for labor income. The solid red line represents the mean of the CATEs, whereas the dotted blue line is the OLS estimate (or homogeneous treatment effect). The OLS estimate and the mean of the CATEs are not significantly different. The variation in the CATEs is considerable: The 10th percentile[7] is DKK -48,683.78 ($-6,572) and the 90th percentile is DKK -18,313.65 ($-2,472).[8] This is also

[7] Percentiles are here calculated as means of the five nearest observations of the actual percentile due to data aggregation requirements from Statistics Denmark.

[8] To address the concern there is something special about labor income at age 30, we also run a variation where labor income is defined as the average of labor income at age 29, 30, and 31 and find the results are robust.

true when looking at the likelihood of having positive labor income (-0.13 and -0.05 for the 10th and 90th percentile, respectively; see Figure 4, Panel B). These results suggest that some subgroups of children are much less impacted by childhood-onset T1D; thus, some parents can mitigate the effect.

An immediate concern is whether the heterogeneity is in fact essential heterogeneity or merely reflects sampling variation. As a heuristic test for this, we proceed by partitioning the data into quartiles of predicted treatment effects using the cross-fitting approach. With the data split into groups (quartiles) of predicted treatment effects, we simply 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.[9] Among the individuals predicted to be in Q1 (i.e., numerically largest penalties), the average treatment effect (standard error) is DKK -42,694.7 (9,160.7) vs. DKK -9,027.5 (9,365.2) in Q4 for labor market income (the difference is significant at the 5% level). For the outcome employment, we have -15.12 (2.05) pp. and -8.68 (1.77) pp. for Q1 and Q4, respectively (significant at the 5% level).

These differences are also significant in an economic sense. For 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. The full set of GATEs is presented in Appendix Figure

[9] 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).

B1. We also perform the *best linear predictor* test formalized in Chernozhukov et al. (2018) and follow the implementation outlined in Athey and Wager (2019). We find strong evidence supporting that the estimated heterogeneity in treatment effects represents true heterogeneity (results are available in Appendix B).

To describe how covariates are associated with treatment heterogeneity, we now compare the observable characteristics of those with larger predicted effects (Q1) vs. those with smaller predicted effects (Q4). These results are reported in Table 4. Panel A shows the mean characteristics from this analysis of labor market income. Being a firstborn is associated with larger predicted penalties (the mean share of firstborns in Q1 is 0.49 vs. 0.41 in Q4). This larger impact on first-born children is consistent with the larger educational spillovers stemming from childhood-onset T1D for younger siblings than older siblings found in Eriksen et al. (2023). Among those with the highest predicted effects, the ratio of males to females is roughly the same; however, males are much more likely to have smaller predicted income penalties.

Turning to the maternal characteristics, we observe that children of less-educated mothers are more likely to have larger predicted effects. The share of mothers with only primary education as highest attained education is 0.62 in Q1 vs. 0.24 in Q4—the share of mothers with medium tertiary education in Q1 is 0.04 vs. 0.42 in Q4. An age gradient is also at play as we find that larger effects are associated with younger mothers. Larger predicted effects are also correlated with lower maternal income. The paternal characteristics generally follow the same patterns as the maternal characteristics; however, the educational gradient is less clear for the lowest levels of educational attainment.

The same pattern broadly emerges when looking at the CATEs regarding the employment outcome. For example, males are underrepresented among those with numerically smaller penalties (share of males 0.42 vs. 0.64 in Q1 vs. Q4). We do note that here, the share of

firstborns is large for those with smaller treatment effects, which seems inconsistent with the findings from the results on labor income.

Of course, all these comparisons of characteristics are partial in that we do not keep all the other characteristics fixed. Having documented a SES gradient in the estimated heterogeneous treatment effects, it is worth noting that these treatment effects are in levels, and that the baseline mean labor market income or probability of being employed can vary considerably across subgroups. Thus, we proceed by quantifying the estimated CATEs relative to these baselines based on selected maternal and paternal characteristics.

To further investigate whether high-SES parents can mitigate the impact, we estimate the CATE within subgroups across maternal and paternal background characteristics, see Figures 5 and 6 (maternal and paternal). Looking at maternal education, we see a monotonic decrease in the (numeric) treatment effect such that more education leads to smaller income penalties. The CATE for labor income among individuals where the mother's educational attainment is high school or less is DKK -44,527 ($-6,009), while it is DKK -29,239 ($-3,946) when the mother has attained a tertiary education. Taken relative to the income mean in each group, this corresponds to relative effects of -19.8% and -7.3%, respectively (Figure 5, Panel A, left side). We also find a similar relationship between maternal education and whether the child is employed. The CATE for having any labor income is -6.9% relative to the mean when the mother has a tertiary education and -12.9% when the mother's educational attainment is high school or less (Figure 6, Panel A, left side).

When considering differences across the maternal income distribution, we find the predicted earnings' penalty is greatest for children of mothers who were in the bottom income quartile the year before the child was born. At age 30, the CATE for these individuals is DKK -53,637 ($-7,238) versus DKK -21,934 ($-2,960) for those in the upper quartile. Relative to the mean, this corresponds to -24.7% and -7.7% (Table 4, Panel A and Figure 5,

Panel B, left side). This income gradient also holds for the probability of employment; see Figure 6, Panel B, left side.

We further investigate the relationship between maternal age at the time of the child's birth and the outcomes. The older the mother was at the time of birth, the smaller the labor market penalty. The differences in CATEs across the age distribution are on par with those found for maternal income and education. When investigating these characteristics, we do not keep the other characteristics fixed; therefore, parental age at child's birth is likely correlated with parental educational attainment and parental income.

We similarly compare relative effects across paternal education levels, paternal income quartile, and paternal age at childbirth (Figures 5 and 6, right side). When the outcome is labor market income, we do not observe the same clear gradient in paternal education levels as we did for maternal education levels (Figure 5, Panel A, right side). However, for 'employment' individuals whose fathers had the lowest educational attainment seem to fare worse. There is, however, evidence of a strong gradient in paternal income quartile for our main outcomes of interest. The pattern for paternal age is comparable to the results by maternal age.

So far, we have clearly demonstrated a social gradient in the impact of childhood-onset T1D. At age 30, individuals who were diagnosed with T1D in childhood have larger penalties if they were more disadvantaged to begin with in terms of having a mother or father in a lower income quartile. While previous literature has documented a clear income-health gradient in childhood (see, e.g., Case, Lubotsky, and Paxson (2002), and Currie and Stabile (2003)), we add to the literature by showing a clear socioeconomic gradient in adult labor market outcomes stemming from the same childhood health shock.

6 Mechanisms

In this section, we investigate potential drivers of the variation in treatment effects by comparing individuals who have the highest and the lowest predicted treatment effects on characteristics that are measured in adulthood. That is, characteristics that are not used to estimate the causal forest, but rather are used to test for potential mechanisms driving the treatment heterogeneity. In this section we show that those with the largest (absolute) treatment effects tend to have worse physical and mental health in adulthood, have lower educational attainment and have less attachment to the labor market. There is also evidence of differential take-up of a new diabetes treatment in childhood by parental SES.

6.1 Diabetes management outcomes and adoption of new treatment regimen

A potential mechanism behind the heterogeneity in adult labor market outcomes could be health capital. Previous literature has documented clinically meaningful differences in glucose control among children by SES in Denmark. For example, Nielsen et al. (2019) document large gaps in glycated hemoglobin (HbA_{1c}) levels by maternal education. Based on slightly younger cohorts born from 1987 and onwards, they find that although children have similar HbA1c levels at the time of diagnosis, the gap becomes statistically significant starting two years after diagnosis (as shown in Figure 7, Panel A). This matters because a large literature, starting with DCCT (1993), has shown a very strong relationship between higher HbA_{1c} (i.e., higher average glucose concentrations) and the risk of diabetes related complications such as retinopathy (leading to blindness), nephropathy (leading to kidney failure), and cardiovascular disease.

Nielsen et al. (2019) document a robust correlation between maternal education and the frequency of daily glucose tests performed, as illustrated in Figure 7, Panel B. Measuring and monitoring glucose levels are crucial for optimal diabetes treatment, as individuals affected by the condition may respond negatively to both excessively high and excessively low

glucose concentrations. Five years after diagnosis, the average number of daily glucose tests is five when the mother has completed high school or less, compared to six when the mother holds at least a master's degree. This suggests that even when financial barriers to accessing test equipment are absent, there are variations in the time investments made in disease management across the SES distribution.

Another reason for differences in health capital could be variation in who utilizes advancements in treatment regimens. Until the late 1980s, the standard treatment of T1D consisted of two daily doses of so-called mixed insulins that involved using a mixture of short-acting and intermediate acting insulins. This regimen provided a fixed insulin dose that did not allow for flexibility in adjusting insulin levels based on individual needs, often with poor glucose control as a result. However, with advances in insulin formulations and the recognition of the importance of achieving tight glycemic control, a new and more advanced treatment regimen became available: Multiple daily injections (MDI). This new insulin delivery method involves administering basal (long-acting) insulin to provide a steady background insulin level throughout the day and bolus (rapid-acting) insulin doses to cover meals and correct high blood glucose levels. The method allows for greater flexibility and customization of insulin doses and takes into account factors such as carbohydrate intake, physical activity, and blood glucose levels. In the hugely influential study Diabetes Control and Complications Trial from 1993 (DCCT, 1993), the treatment arm was MDI, and the control arm was the traditional therapy of only two daily injections. The study concluded that MDI led to better glucose management and fewer diabetes related complications. The widespread adoption of MDI as an alternative to the simpler two-dose regimen with mixed insulins occurred over the course of the 1990s and early 2000s. Since then, MDI has become the standard treatment for most individuals with T1D. Relevant for our study, this is the period where many of the individuals in our sample were followed at pediatric treatment

centers. In Figure 8, we show the share of individuals who were on MDI treatment in childhood by maternal education (high school or less vs. tertiary education). From the onset of T1D and extending at least five years after the diagnosis, children of mothers with tertiary education are approximately 10 pp. more likely to be on MDI treatment. While the confidence bands are overlapping, the difference is 10.6 pp. from a baseline of 60.5%. in the high school or less group (significant at the 1% level) taken over the entire data period. This suggests that differential usage of modern diabetes treatments may play a role in creating disparities. This is consistent with Glied and Lleras-Muney (2008), who find that technological innovation leads to inequality in health because individuals with higher levels of education are better able to take advantage of cancer-related medical advancements.

We test for differences in diabetes health outcomes in adulthood by comparing the means in Q1 and Q4 for the subset of our sample who were diagnosed (i.e., the treatment group). Results are shown in the top portion of Table 5.[10] The largest difference in characteristics for labor market income is for being hospitalized with diabetic coma or diabetic ketoacidosis (DKA). Forty-one percent of individuals in Q1 have been hospitalized with diabetic coma or DKA as opposed to 33% in Q4, indicating an 8 pp. gap between the highest and lowest quartile. For 'employment', the biggest difference between Q1 and Q4 is again being hospitalized due to diabetic coma or DKA. In Q1 40% of individuals have been hospitalized with diabetic coma or DKA and 32% in Q4 have been hospitalized, resulting in an 8 pp. gap. We do not observe any difference across Q1 vs. Q4 on other health and utilization metrics, including LDL cholesterol levels, the probability of having ever smoked, and the probability of being treated for high cholesterol or high blood pressure.

[10] As part of our analysis on the diabetes only sample, we also test for differences in penalties by age at onset where we test for early onset (age six or younger) versus later onset (seven or above). The results suggest that early onset is associated with larger penalties.

In appendix figures A1-A4 we emulate Figures 5 and 6 and graphically show the social gradient in HbA_{1c} levels, hospitalizations related to diabetic coma or DKA, specialty ambulatory care, and late diabetes complications in adulthood. While the confidence intervals overlap, our results provide suggestive evidence that children from less advantaged backgrounds have larger long-run penalties because of worse disease management in childhood and poorer health status in adulthood.

Taken together, these results suggest that both maternal and paternal socioeconomic status matters for T1D health-related outcomes in adulthood and that better adult health is tied to better adult labor market outcomes. Children of more educated mothers and children of mothers in higher income quartiles may have better disease management as proxied by HbA_{1c} levels and hospitalizations due to diabetic coma or DKA. We find a similar suggestive pattern for paternal education and paternal income quartile. Finding that parental characteristics matter for child health is consistent with medical literature documenting the role of parental behavior in children's glucose control (see, e.g., Davis, Delamater, Shawn, La Greca, Eidson, Perez-Rodiguez, and Nembery (2001) and Thompson, Auslander, and White (2001)). High-SES parents may be better at producing diabetes-management-related behaviors due to differences in knowledge and cognitive abilities, and because their social networks include people who are more able to navigate the healthcare system.[11]

6.2 Educational Outcomes

[11]Cutler and Lleras-Muney (2010) determine that 30% of the education-health behavior gradient can be explained by differences in knowledge and cognitive ability, while another 10% can be explained by social networks.

Since health capital impacts human capital (see, e.g., Bhalotra, and Venkataramani (2015), Karbownik, and Wray (2019) and Saez (2021)), we next discuss how T1D impacts educational outcomes. Previous research has documented a clear relationship between glucose control and T1D educational penalties. For example, Skipper et al. (2019) show that, on average, children diagnosed with T1D have similar test scores compared with their peers, but that children with poorer glucose control have poorer test scores and children with the best glucose control have better test scores. Eriksen et al. (2020) study the association of T1D and school well-being among middle school children and find similar outcomes (except children with T1D reported higher levels of headaches). However, those with poorer glucose control were again found to have worse outcomes. Lindkvist et al. (2021) find that being diagnosed with T1D leads to lower 9th grade exit exam scores, a higher relative risk of not completing 9th grade by age 16 (end of compulsory school in Denmark), and a higher risk of not being enrolled in or graduated from upper secondary school by age 20. The results also indicate larger penalties for those with poorer glucose control. Taken together, this research suggests that poorly controlled T1D in childhood can negatively impact a wide variety of educational outcomes.

Even individuals with well-controlled T1D could be negatively impacted in the longer term through increased absenteeism. For example, Thingholm et al. (2020) find that after diagnosis there is an increase in school absenteeism of around 50% more than the matched comparison children. Even children with well-controlled glucose must miss school to attend medical appointments at a pediatric endocrinology clinic multiple times a year. There may be cumulative effects of consistently missing more school than peers without T1D.

Using our data, we now test for the role of educational attainment in explaining the labor market results. For each labor market outcome, we compare the mean of different levels of educational attainment for Q1 versus Q4. Results are shown in the bottom portion of Table 5.

For labor market income, we find evidence of higher levels of educational attainment in Q4 than Q1. For example, 22% of individuals in Q1 have attained only a primary level of education, whereas the corresponding figure for Q4 is only 14%. Eighteen percent of individuals in Q4 have attained a long tertiary education while only 9% of individuals in Q1 have attained a long tertiary education. Similarly, for employment we find evidence of higher levels of educational attainment among those in Q4 than in Q1. For example, 23% of individuals in Q1 have only attained a primary level of education while this number is only 14% in Q4. In terms of long tertiary education, only 9% of individuals in Q1 have attained this level of education, while 17% of individuals in Q4 have this educational attainment. Altogether, these results suggest impacts on educational attainment is likely an important mechanism in driving the heterogeneity in labor market outcomes.

6.3 Differences in employment-related outcomes

In Table 5, we also report differences between more and less severely affected individuals for several employment related outcomes. Motivated by Deshpande (2016), who finds that removing low-income youths with a disability from Supplemental Security Income leads to increased income volatility, we test whether individuals with the largest estimated penalties have more employment volatility. Focusing on labor market income, we find that individuals in Q1 diagnosed with T1D spend a larger share of time between age 25 and age 30 as unemployed (2.9% vs. 1.7%) and have more unemployment spells from age 25 to age 30 than those in Q4 (1.48 vs. 1.19). The group with the largest predicted penalties is also more likely to have taken anti-depressants or anti-anxiety medication post diagnosis (27% in Q1 versus 21% in Q4) and is more likely to be on disability insurance post diagnosis (4% in Q1 versus 2% in Q4). We also find evidence of differences in the intensity of work at age 30. A total of 16% of individuals in Q1 work part-time versus 14% in Q4. Additionally, Q1 is also more

likely to take long-term sickness absence.[12] The patterns are very similar when focusing on employment as the outcome.

7 Conclusion

In this paper we use the onset of T1D in childhood, Danish administrative registry data, and causal forests to show that differences in how families across the socioeconomic distribution respond to childhood health shocks plays an important role in creating intergenerational inequality. The higher the mother's educational attainment or the higher the mother's or father's income quartile leads to smaller T1D penalties in adult employment and labor market income. We provide some evidence that children of high-SES parents have better diabetes-related health outcomes. This is perhaps due to their higher take-up rates of modern treatment regimens, even when there are no out-of-pocket costs associated with this. This indicates that access to healthcare is not sufficient to eliminate disparities stemming from childhood health shocks; that is, social determinants of health matter.

It should be noted that conducting a similar analysis using data from another country may result in a different distribution of treatment heterogeneity. Variation in access to health insurance, availability, and access to endocrinology care, costs of medicine and medical devices, worker and student protection, and anti-discrimination laws related to health and disability status could all impact the results. However, our results do suggest that it is important for research using data from other countries to also test for variation in long-run impacts from childhood health shocks.

[12] Note that this does not include short-term sickness absences, but only extended sickness-related absences that are reported to the government due to firms having the right to receive wage-compensation from the government in the case of long-term sickness absence spells. Data regarding short-term sickness absence belongs to firms, and thus we do not have access to such information.

Future research could focus on pinpointing what high-SES parents do differently than low-SES parents in terms of investing in the child's health capital. If there are differences in disease management knowledge or an inability to follow prescribed treatment plans, then maybe children with less educated mothers or fathers and children of lower income mothers and fathers would benefit from increased clinic visits, telehealth meetings or at-home support. Differences in parental preferences or discount rates may also be driving the effects, in which case rules surrounding when additional clinical interventions are implemented may need to be reevaluated.

References:

Aizer, A. and Currie, J. (2014), "The intergenerational transmission of inequality: Maternal disadvantage and health at birth", *Science*, vol. 344, 856-861

Athey, S., and Imbens, G. (2016), "Recursive Partitioning for Heterogeneous Causal Effects," *Proceedings of the National Academy of Sciences*, 113, 7353–7360.

Athey, S. and Wager, S. (2019). "Estimating treatment effects with causal forests: an application". *arXiv*:1902.07409.

Adhvaryu, Bednar, Molina, Nguyen, Nyshadham, (2020). " When It Rains It Pours: The Long-Run Economic Impacts of Salt Iodization in the United States" *The Review of Economics and Statistics*, vol 102(2), pages 395-407.

Almond, Currie, and Duque, (2018). "Childhood Circumstances and Adult Outcomes: Act II," Journal of Economic Literature, 56(4), 1360–1446

Beach, Brian, Joseph Ferrie, Martin Saavedra, and Werner Troesken, (2016). "Typhoid Fever, Water Quality, and Human Capital Formation." Journal of Economic History 76 (1): 41–75.

Bhalotra, Sonia, and Atheendar Venkataramani. (2015). "Shadows of the Captain of the Men of Death: Health Innovation, Human Capital Investment, and Institutions." Working paper.

Bowles and Gintis (2002). The Inheritance of Inequality. *Journal of Economic Perspectives* Vol 16, Number 3, p3-30

Case, Lubotsky and Paxon (2002). Economic Status and Health in Childhood: The Origins of the Gradient. *The American Economic Review.* Vol. 92, No. 5, p1334-1308

Chernozhukov, V., Demirer, M., Duflo, E. & Fernández-Val, I. (2018) Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. *NBER Working Paper* no. 24678.

Cobry, E and Jaser, S. (2019). Brief Literature Review: The Potential of Diabetes Technology to Improve Sleep in Youth With Type 1 Diabetes and Their Parents: An Unanticipated Benefit of Hybrid Closed-Loop Insulin Delivery Systems. *Diabetes Spectrum* DOI**:** 10.2337/ds18-0098

Cutler and Lleras-Muney. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics.* Volume 29, Issue 1, p 1-28.

Currie and Stabile (2003). Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children? *The American Economics Review.* Vol 93, No. 5, p 1813-1823

Davis, Delamater, Shaw, La Greca, Eidson, Perez-Rogdriguez, Nemery. (2001). Parenting styles, regimen adherence, and glycemic control in 4- to 10-year-old children with diabetes. *J Pediatric Psychology* Mar;26(2):123-9. doi: 10.1093/jpepsy/26.2.123

DCCT 1993 (The Diabetes Control and Complications Trial Research Group), (1993) The Effect of Intensive Treatment of Diabetes of the Development and Progression of Long-Term Complications in Insulin-Dependent Diabetes Mellitus, *New England Journal of Medicine*, vol. 329: 977-986.

Deshpande, M. (2016). Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls, *American Economic Review*, 106(11): 3300–3330.

Eriksen, Gaulke, Skipper and Svensson (2021). The Impact of Childhood Health Shocks on Parental Labor Supply. *Journal of Health Economics*. Vol 78

Eriksen, Gaulke, Thingholm, Svensson and Skipper (2020). Association of type 1 diabetes and school wellbeing: a population-based cohort study of 436,439 Danish schoolchildren. *Diabetologia*. https://doi.org/10.1007/s00125-020-05251-z

Eriksen, Gaulke, Thingholm, Svensson and Skipper (2023). Educational Consequences of Sibling's Disability. *Economics of Education Review*. Volume 94.

Gensowski, Nielsen, Nete M. Nielsen, Rossin-Slater and Wüst, (2019). "Childhood health shocks, comparative advantage, and long-term outcomes: Evidence from the last Danish polio epidemic". *Journal of Health Economics*, volume 66, pages 27-36 https://doi.org/10.1016/j.jhealeco.2019.03.010

Glied, Sherry and Adriana Lleras-Muney; Technological innovation and inequality in health. *Demography* 1 August 2008; 45 (3): 741–761. doi:

Guo and Zhang, (2021). Altruistic or Exchange Motive? Evidence on the effect of children's health shocks on intra-household resource allocation. Working paper.

Imperatore, G., Mayer-Davis, E. J., Orchard, T. J. and Zhong, V. W. (2018). "Prevalence and Incidence of Type 1 Diabetes Among Children and Adults in the United States and Comparison With Non-U.S. Countries". Diabetes in America. 3rd edition. Cowie CC, Casagrande SS, Menke A, et al., editors.
Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases (US); 2018 Aug.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker. (2017). "Every Breath You Take—Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970." *Journal of Political Economy* 125(3): 848–902.

Karbownik, Krzysztof and Anthony Wray, (2019). "Educational, Labor-market and Intergenerational Consequences of Poor Childhood Health" NBER Working Paper No. 26368

Lindkvist, Thorsen, Paulsrud, Thingholm, Eriksen, Gaulke, Skipper and Svensson (2021). Association of type 1 diabetes and educational achievement in 16–20-year-olds: A Danish nationwide register study. *Diabetic Medicine.* DOI: 10.1111/dme.14673

Nathan, D. M. (1993), “Long-term complications of diabetes mellitus”, *New England Journal of Medicine*, vol. 328, 1676-1685.

Nielsen, Gaulke, Eriksen, Svensson, and Skipper (2019). Socioeconomic inequality in metabolic control among children with type 1 diabetes: a nationwide longitudinal study of 4,079 Danish children", *Diabetes Care*, 2019 Aug; 42(8): 1398-1405.

Nilsson, J. Peter, (2017). “Alcohol Availability, Prenatal Conditions, and Long-Term Economic Outcomes.” *Journal of Political Economy*, 125 (4): 1149–207.

Persson, Gerdtham and Carlsson (2016). “Labor market consequences of childhood onset type 1 diabetes” Economics and Human Biology. 23, p. 180-192.

Pillar, Schuscheim, Weiss, Malhotra, McCowen, Shlitner, Peled, and Shehadeh. (2003). Interactions between hypoglycemia and sleep architecture in children with type 1 diabetes mellitus, *The Journal of Pediatrics*, Volume 142, Issue 2, Pages 163-168, https://doi.org/10.1067/mpd.2003.66.

Pociot and Lernmark. (2016). Genetic risk factors for type 1 diabetes, *The Lancet*, Volume 387, Issue 10035, Pages 2331-2339, https://doi.org/10.1016/S0140-6736(16)30582-7.

Prætorius, K., Urhoj, S. K., & Andersen, A.-M. N. (2022). Parental socio-economic position and the risk of type 1 diabetes in children and young adults in Denmark: A nation-wide register-based study. *Scandinavian Journal of Public Health*, Epub ahead of print.

Regnell, S., and Lernmark, A. (2017). Early Prediction of Autoimmune (type 1) diabetes. *Diabetologia*, 1370-1381.

Restrepo, Brandon J. (2016). Parental investment responses to a low birth weight outcome: Who compensates and who reinforces? Journal of Population Economics, 29(4):969–989.

Saez, Camila. (2021). Long-term effects of early life exposure to pneumonia. Working paper.
Schwandt, Hannes, (2018). "The Lasting Legacy of Seasonal Influenza: In-Utero Exposure and Labor Market Outcomes." Centre for Economic Policy Research Discussion Paper DP12563

Smith, J.P. (2009). The Impact of Childhood Health on Adult Labor Market Outcomes. *The Review of Economics and Statistics*, August, 91(3): 478–489

Skipper, N., Gaulke, A., Sildorf, S.M., Eriksen, T.M., Nielsen, N.F., Svensson, J. (2019). Association of Type 1 Diabetes with Standardized Test Scores of Danish Schoolchildren. *JAMA.* 321(5), 484–492. doi:10.1001/jama.2018.21819

Tao, B., Pietropaolo, M., Atkinson, M., Schatz, D., and Taylor, D. (2010). Estimating the cost of Type 1 Diabetes in the US: A Propensity Score Matching Method. *Plos ONE*, July 2010

Thingholm, P. R., Gaulke, A., Eriksen, T. M., Svensson, J., Skipper, N. (2020). Association of Prodromal Type 1 Diabetes With School Absenteeism of Danish Schoolchildren: A Population-Based Case-Control Study of 1,338 Newly Diagnosed Children, *Diabetes Care*, Nov. 43(11), 2886-2888, 2020,

Thompson, Auslander and White (2001). Comparison of Single-Mother and Two-Parent Families on Metabolic Control of Children with Diabetes. *Diabetes Care*;24(2):234–238 https://doi.org/10.2337/diacare.24.2.234

Wager and Athey, (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, doi:10.1080/01621459.2017.1319839

Table 1: Descriptive statistics

		Diabetes	No diabetes
Number of observations		1810	9050
Child characteristics			
Birth order	First born (0/1)	0.45	0.47
	Second born (0/1)	0.40	0.38
	Third born or later (0/1)	0.16	0.15
Sex	Male (0/1)	0.53	0.51
Paternal Characteristics			
Education	Primary education (0/1)	0.33	0.31
	Secondary education (0/1)	0.03	0.03
	Vocational (0/1)	0.44	0.43
	Short Ter. (0/1)	0.02	0.03
	Medium Ter. (0/1)	0.11	0.13
	Long Ter. (0/1)	0.07	0.07
Income	First quartile (0/1)	0.17	0.17
	Second quartile (0/1)	0.16	0.17
	Third quartile (0/1)	0.16	0.17
	Fourth quartile (0/1)	0.17	0.16
Immigration status	Native (0/1)	0.97	0.97
	Immigrant or descendent (0/1)	0.03	0.03
Age	<25 (0/1)	0.07	0.09
	25-29 (0/1)	0.30	0.31
	30-34 (0/1)	0.35	0.34
	35+ (0/1)	0.27	0.26
Maternal Characteristics			
Education	Primary education (0/1)	0.41	0.39
	Secondary education (0/1)	0.04	0.03
	Vocational (0/1)	0.31	0.31
	Short Ter. (0/1)	0.03	0.03
	Medium Ter. (0/1)	0.18	0.20
	Long Ter. (0/1)	0.03	0.03
Income	First quartile (0/1)	0.18	0.17
	Second quartile (0/1)	0.16	0.17
	Third quartile (0/1)	0.17	0.17
	Fourth quartile (0/1)	0.16	0.17
Immigration status	Native (0/1)	0.98	0.98
	Immigrant or descendent (0/1)	0.02	0.02
Age	<25 (0/1)	0.21	0.22
	25-29 (0/1)	0.39	0.40
	30-34 (0/1)	0.28	0.27
	35+ (0/1)	0.12	0.11

Notes: Descriptive statistics of individuals diagnosed with diabetes before the age of 18 and a group matched on birthday and onset year.

Table 2: Linear regression of diabetes status on observable characteristics

		Coeff.	S.E.
Child characteristics			
Birth order	First born (0/1)	Ref.	
	Second born (0/1)	0.002	(0.009)
	Third born or later (0/1)	-0.009	(0.012)
Sex	Male (0/1)	0.014**	(0.007)
Paternal characteristics			
Education	Primary education (0/1)	Ref.	
	Secondary education (0/1)	0.002	(0.022)
	Vocational (0/1)	-0.004	(0.009)
	Short Ter. (0/1)	-0.035	(0.022)
	Medium Ter. (0/1)	-0.025*	(0.013)
	Long Ter. (0/1)	-0.016	(0.018)
Income	First quartile (0/1	Ref.	
	Second quartile (0/1)	-0.005	(0.013)
	Third quartile (0/1)	-0.007	(0.013)
	Fourth quartile (0/1)	0.008	(0.014)
Immigration status	Immigrant or descendent (0/1)	-0.018	(0.022)
Age			
	<25 (0/1)	Ref.	
	25-29 (0/1)	0.028**	(0.014)
	30-34 (0/1)	0.035**	(0.015)
	35+ (0/1)	0.035**	(0.017)
Maternal characteristics			
Education	Primary education (0/1)	Ref.	
	Secondary education (0/1)	0.005	(0.022)
	Vocational (0/1)	-0.012	(0.009)
	Short Ter. (0/1)	-0.016	(0.021)
	Medium Ter. (0/1)	-0.026**	(0.011)
	Long Ter. (0/1)	-0.013	(0.023)
Income	First quartile (0/1)	Ref.	
	Second quartile (0/1)	-0.015	(0.013)
	Third quartile (0/1)	-0.015	(0.013)
	Fourth quartile (0/1)	-0.014	(0.014)
Immigration status	Immigrant or descendent (0/1)	-0.022	(0.024)
Age	<25 (0/1)	Ref.	
	25-29 (0/1)	0.006	(0.011)
	30-34 (0/1)	0.017	(0.014)
	35+ (0/1)	0.02	(0.017)
	Observations	10,860	
	R-squared	0.003	
	Joint F-test	0.589	
	Prob > F	0.995	

Notes: OLS regression of diabetes status on observable characteristics. The OLS additionally controls for cohort, onset-year and municipality fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the match group level.

Table 3: Average Treatment Effects of Diabetes onset in childhood.

Panel A	(1) Income (DKK)	(2) Income (DKK)	(3) Income (DKK)	(4) Income (DKK)
Diabetes	-33,102***	-32,978***	-32,453***	-32,422***
	(4,540)	(4,482)	(4,487)	(4,460)
Observations	10,418	10,418	10,418	10,418
Mean Outcome	239,468.45	239,468.45	239,468.45	239,468.45
Paternal Char.	No	Yes	No	Yes
Maternal Char.	No	No	Yes	Yes
R-squared	0.047	0.067	0.069	0.078

Panel B	(1) Employment (0/1)	(2) Employment (0/1)	(3) Employment (0/1)	(4) Employment (0/1)
Diabetes	-0.0875***	-0.0875***	-0.0866***	-0.0867***
	(0.0110)	(0.0109)	(0.0109)	(0.0109)
Observations	10,418	10,418	10,418	10,418
Mean Outcome	0.81	0.81	0.81	0.81
Paternal Char.	No	Yes	No	Yes
Maternal Char.	No	No	Yes	Yes
R-squared	0.043	0.054	0.058	0.064

Notes: Standard errors in parentheses. *** p<0.01. ** p<0.05. * p<0.1. Estimates of treatment impact of diabetes onset in childhood. The estimates are coefficients from separate OLS regressions of income, or an indicator for employment defined as having any income. Column 1 controls for child characteristics, cohort, onset-year and municipality fixed effects. Column 2 and 3 additionally controls for paternal or maternal characteristics. Column 4 controls for both paternal and maternal characteristics. Mean outcomes are reported for the control group.

Table 4: Mean child, maternal, and paternal characteristics by predicted CATE quartile, Q1 vs. Q4

		Panel A: Labor market income				Panel B: Employment			
Child characteristics		Q1	Q4	Difference	t stat	Q1	Q4	Difference	t stat
Birth order	First born (0/1)	0.49	0.41	0.08	6.00	0.36	0.59	-0.23	-17.00
	Second born (0/1)	0.37	0.45	-0.08	-6.24	0.53	0.27	0.26	19.56
	Third born or later (0/1)	0.14	0.14	0.00	0.23	0.12	0.15	-0.03	-2.66
Sex	Male (0/1)	0.51	0.55	-0.04	-2.61	0.42	0.64	-0.22	-16.50
Maternal characteristics									
Education	Primary education (0/1)	0.62	0.24	0.39	30.79	0.58	0.27	0.31	23.93
	Secondary education (0/1)	0.04	0.02	0.02	4.80	0.03	0.03	0.00	0.38
	Vocational (0/1)	0.24	0.27	-0.03	-2.37	0.24	0.34	-0.10	-8.05
	Short Ter. (0/1)	0.03	0.03	0.00	0.99	0.02	0.03	-0.01	-3.43
	Medium Ter. (0/1)	0.04	0.42	-0.39	-37.35	0.10	0.29	-0.18	-17.38
	Long Ter. (0/1)	0.02	0.03	0.00	-0.09	0.02	0.03	-0.01	-2.80
Age	<25 (0/1)	0.25	0.14	0.11	9.85	0.26	0.21	0.06	4.88
	25-29 (0/1)	0.52	0.33	0.19	14.20	0.51	0.29	0.23	17.06
	30-34 (0/1)	0.16	0.39	-0.23	-19.33	0.16	0.37	-0.20	-17.26
	35+ (0/1)	0.07	0.14	-0.07	-7.86	0.06	0.14	-0.08	-9.33
Income	First quartile (0/1)	0.34	0.05	0.29	28.29	0.43	0.02	0.41	40.07
	Second quartile (0/1)	0.24	0.09	0.15	14.94	0.20	0.10	0.09	9.74
	Third quartile (0/1)	0.19	0.13	0.06	5.76	0.13	0.15	-0.02	-1.85
	Fourth quartile (0/1)	0.10	0.23	-0.13	-12.79	0.07	0.24	-0.17	-17.19
Immigration status	Immigrant or descendent (0/1)	0.02	0.03	-0.01	-1.30	0.02	0.02	0.00	-0.66
Paternal characteristics									
Education	Primary education (0/1)	0.30	0.34	-0.04	-2.93	0.41	0.24	0.17	13.20
	Secondary education (0/1)	0.02	0.04	-0.02	-3.54	0.03	0.03	0.00	0.15
	Vocational (0/1)	0.54	0.32	0.22	16.43	0.41	0.48	-0.06	-4.63
	Short Ter. (0/1)	0.03	0.02	0.00	0.69	0.03	0.02	0.00	0.43
	Medium Ter. (0/1)	0.08	0.18	-0.10	-11.03	0.08	0.14	-0.06	-6.56
	Long Ter. (0/1)	0.03	0.10	-0.07	-9.67	0.04	0.09	-0.05	-7.53
Age	<25 (0/1)	0.12	0.05	0.07	9.29	0.12	0.08	0.04	5.04
	25-29 (0/1)	0.36	0.21	0.15	12.26	0.44	0.22	0.22	17.23
	30-34 (0/1)	0.32	0.42	-0.10	-7.67	0.29	0.41	-0.12	-9.21
	35+ (0/1)	0.20	0.32	-0.12	-9.98	0.16	0.30	-0.14	-12.09
Income	First quartile (0/1	0.30	0.05	0.25	25.02	0.36	0.03	0.33	33.13
	Second quartile (0/1)	0.21	0.11	0.09	9.17	0.20	0.13	0.07	7.04
	Third quartile (0/1)	0.21	0.13	0.08	7.54	0.17	0.16	0.01	0.81
	Fourth quartile (0/1)	0.14	0.17	-0.03	-2.60	0.10	0.17	-0.08	-8.14
Immigration status	Immigrant or descendent (0/1)	0.02	0.03	-0.01	-2.66	0.03	0.02	0.00	0.70

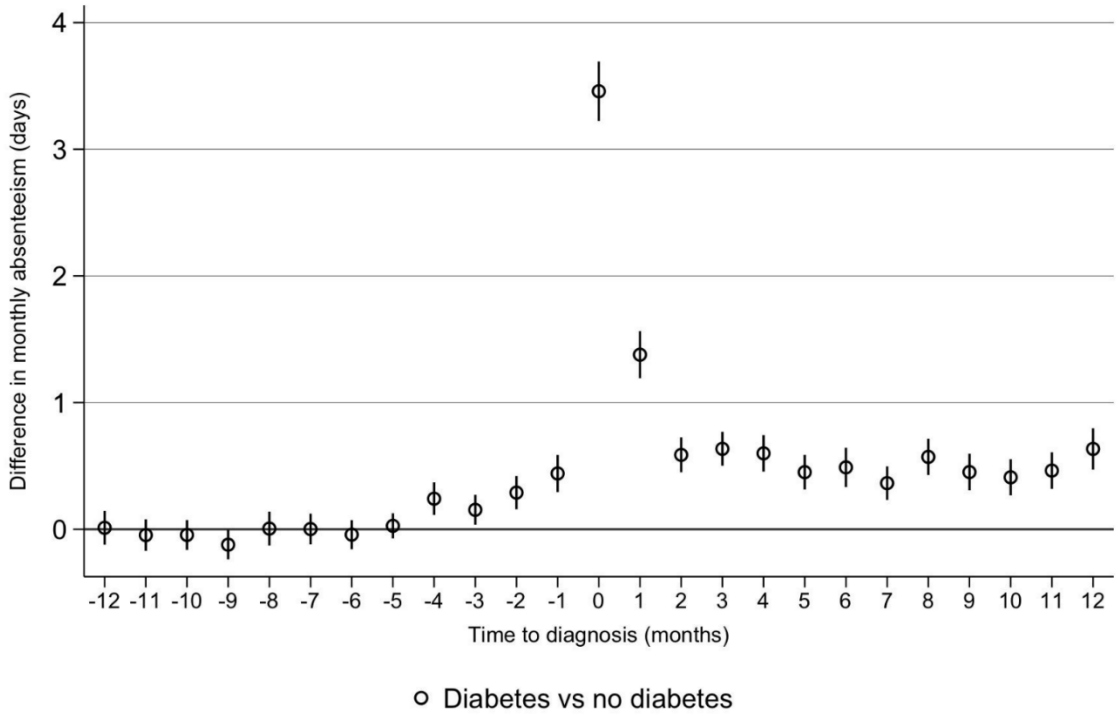
Notes: The table reports the means and differences in means for child, maternal, and paternal characteristics by predicted quartile of the conditional average treatment effects. Q1 represents the numerically largest effects and Q4 the smallest. Panel A shows the differences by labor market income and Panel B for the outcome employment defined as having any labor market income.

Table 5: Differences in selected health and labor market outcomes by predicted treatment quartile.

		Panel A: Labor market income				Panel B: Employment			
		Q1	Q4	Diff.	t stat	Q1	Q4	Diff.	t stat
Diabetes Sample									
Health	Diabetic coma/DKA (0/1)	0.41	0.33	0.08	2.31	0.40	0.32	0.08	2.34
	Late complications (0/1)	0.52	0.45	0.07	2.00	0.53	0.49	0.05	1.37
	HbA1c	8.20	8.07	0.13	1.27	8.28	8.16	0.12	1.28
	LDL cholesterol	2.56	2.53	0.03	0.51	2.61	2.60	0.02	0.28
	Ever smoker (0/1)	0.28	0.31	-0.03	-0.82	0.33	0.33	0.00	0.12
	Hypertension treatment (0/1)	0.17	0.16	0.00	0.17	0.20	0.19	0.01	0.48
	Lipid lowering treatment (0/1)	0.12	0.13	-0.01	-0.59	0.17	0.14	0.02	0.87
Timing of onset	Year of diabetes onset	1995.3	1993.9	1.37	4.50	1994.3	1994.0	0.26	0.87
	Early onset, <7 years (0/1)	0.14	0.05	0.09	4.55	0.19	0.05	0.13	6.13
	Late onset, >7 years (0/1)	0.86	0.95	-0.09	-4.55	0.81	0.95	-0.13	-6.13
Full sample									
Employment	Part time (0/1)	0.16	0.14	0.02	1.98	0.17	0.13	0.04	3.79
	Public sector (0/1)	0.37	0.42	-0.05	-3.09	0.41	0.39	0.02	1.39
	Share unemployed	0.029	0.017	0.01	5.62	0.030	0.019	0.01	4.89
	Unemployment spells	1.48	1.19	0.29	3.09	1.46	1.18	0.28	3.18
Health	Anti dep. or anxiety meds. (0/1)	0.27	0.21	0.06	5.19	0.29	0.22	0.08	6.52
	Sickness benefits (0/1)	0.16	0.14	0.02	1.91	0.17	0.14	0.03	2.54
	Disability programs (0/1)	0.04	0.02	0.02	4.12	0.05	0.02	0.03	4.99
Education	Primary education (0/1)	0.21	0.14	0.06	5.96	0.22	0.14	0.09	8.03
	Secondary education (0/1)	0.08	0.08	0.00	-0.56	0.07	0.08	-0.01	-1.10
	Vocational (0/1)	0.39	0.29	0.10	7.61	0.37	0.33	0.04	3.03
	Short Ter. (0/1)	0.04	0.06	-0.02	-2.66	0.04	0.06	-0.02	-2.72
	Medium Ter. (0/1)	0.18	0.24	-0.06	-5.40	0.18	0.23	-0.04	-3.62
	Long Ter. (0/1)	0.09	0.18	-0.08	-8.87	0.10	0.16	-0.06	-6.94

Notes: Differences in the probability of selected health outcomes for individuals with T1D by treatment effect quartile. Differences in educational attainment for the full sample by treatment effect quartile. DKA is diabetic ketoacidosis. The health outcomes are based on whether a person ever has that outcome after diagnosis (or pseudo-diagnosis). Diabetes sample refers to the treated individuals, and full sample refers to both treated and comparison individuals. Diabetic outcomes are only observed for diagnosed(treated) individuals.

Figure 1: Mean difference (95% CI) in days absent from school during a given month relative to diagnosis of type 1 diabetes (diabetes vs. no diabetes)

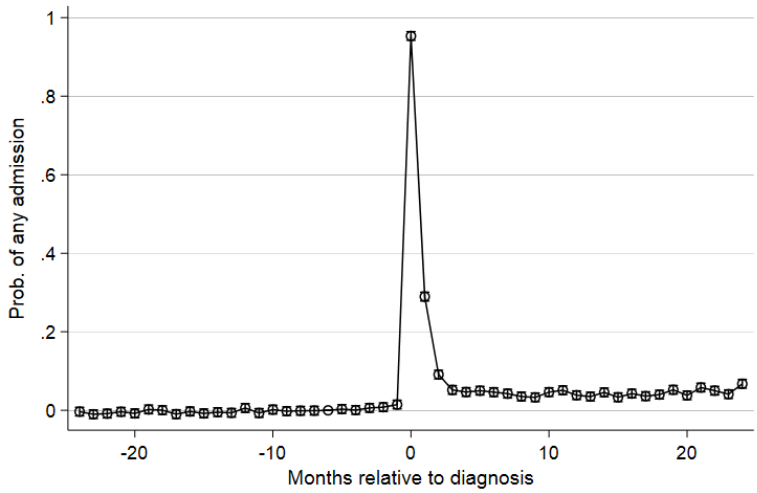


Notes: n= 1,338 children diagnosed with type 1 diabetes from August 1 2010 to June 30 2017 compared with n= 6,690 age and sex matched controls. Mean (95% CI) difference in number of days absent from school relative to diabetes diagnosis (month 0). The mean differences are adjusted for calendar-month and school grade specific effects. As the month of July is the only month of year with no school days in Denmark, it was left out of the analysis. Months -12 to -5 showed non-significant differences (with a level of significance at $p<0.05$).

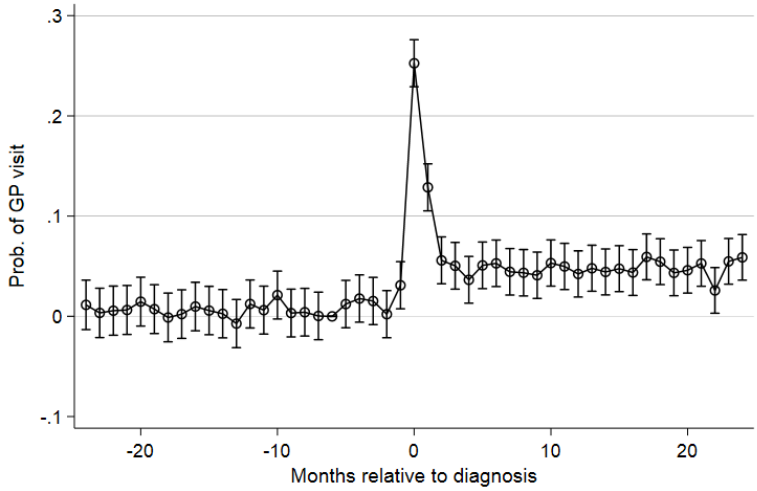
Published previously in Thingholm et al. Association of Prodromal Type 1 Diabetes with School Absenteeism of Danish Schoolchildren: A Population-Based Case-Control Study of 1,338 Newly Diagnosed Children. Diabetes Care 2020 Nov; 43(11): 2886-2888. Copyright 2020 by the American Diabetes Association.

Figure 2: Differences in health care utilization around the time of diagnosis, diabetes vs. comparison

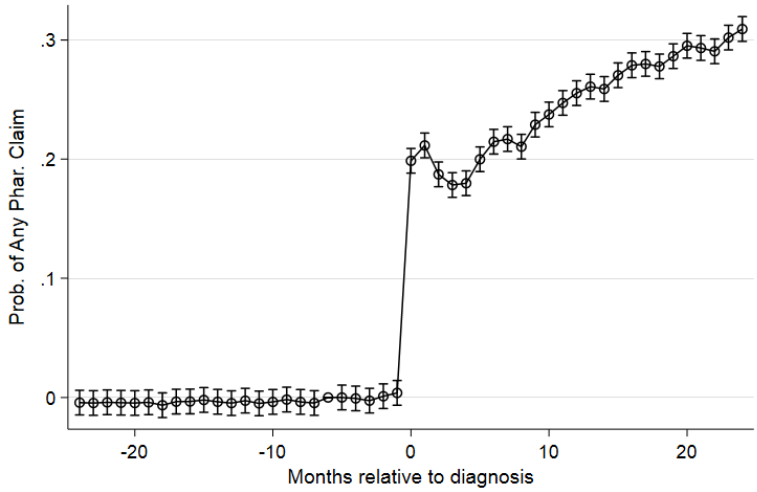
Panel A: Probability of any hospital admission



Panel B: Probability of any GP visit



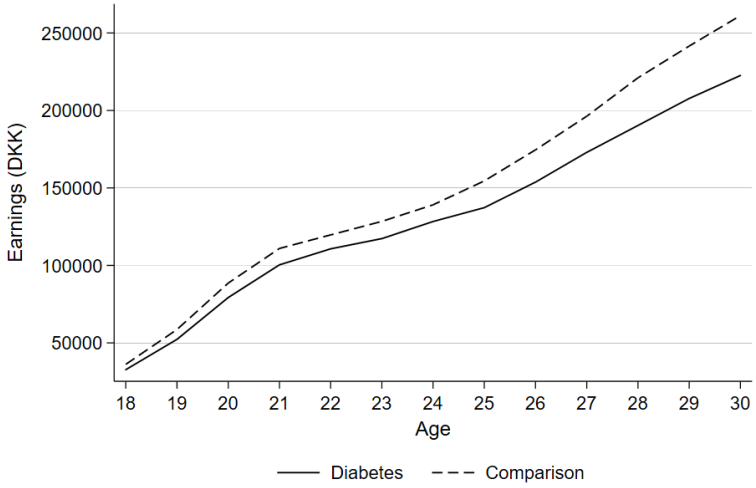
Panel C: Probability of any pharmacy claim



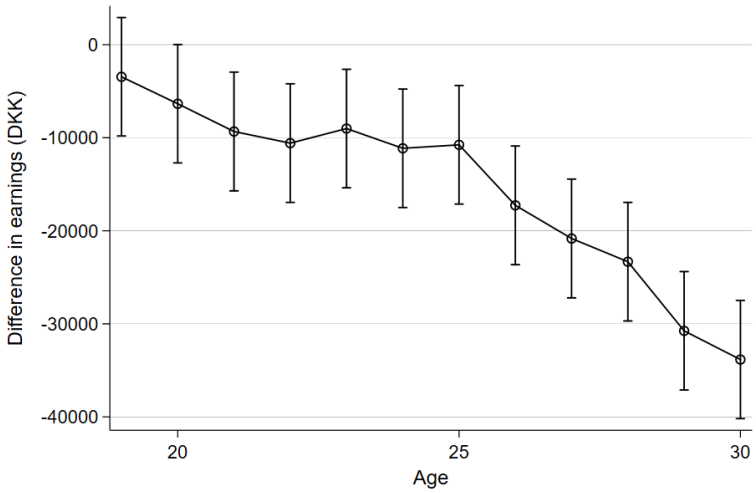
Notes: Panel A shows the mean difference (95% CI) in 'any hospital admission' by month relative to the diagnosis month for diabetes vs comparison individuals. In Panel B, the outcome is the probability of visiting the general practitioner, and Panel C shows the difference in the probability of having a pharmacy claim.

Figure 3: Labor market outcomes by age, diabetes vs. comparison

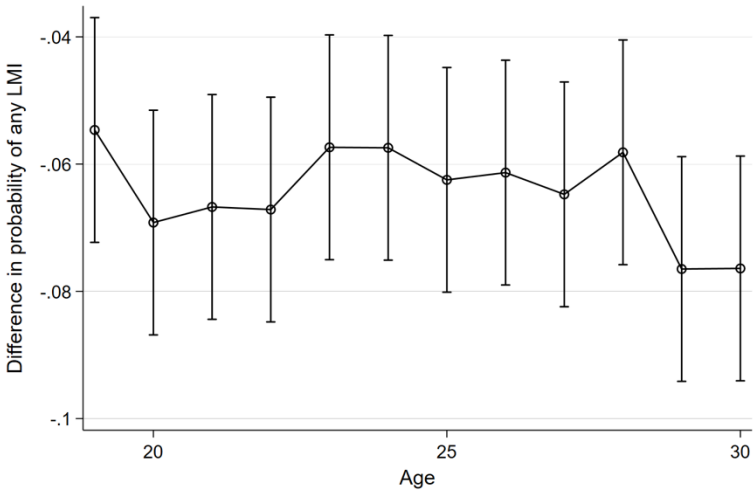
Panel A: Labor market income



Panel B: Difference in labor market income by age



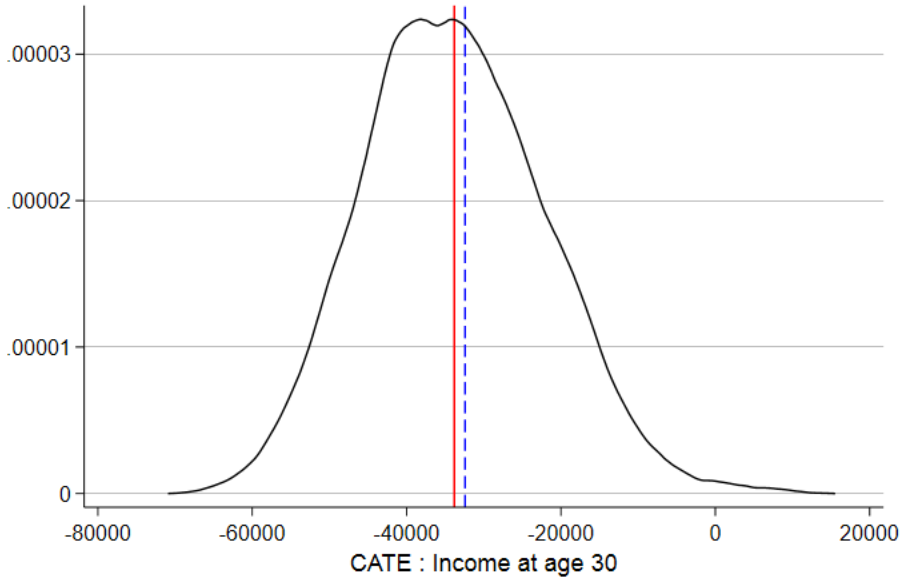
Panel C: Difference in the probability of employment



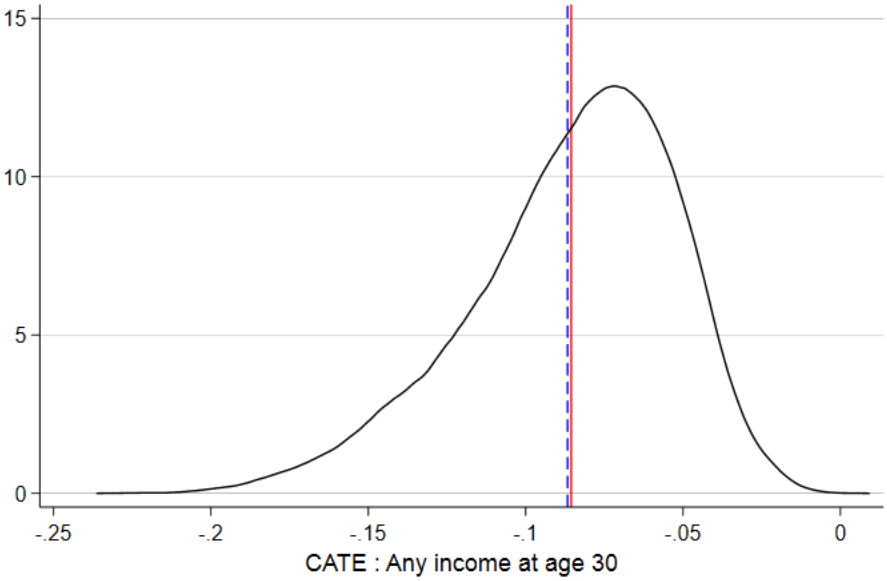
Notes: Panel A shows the development in mean labor market income (DKK) from age 18 to 30 for individuals with diabetes and the comparison group. Panel B shows the mean difference in labor market income (95% CI). Panel C shows the mean difference in employment defined as the probability of having positive labor market income (95% CI).

Figure 4: Estimated distributions of Conditional Average Treatment Effects

Panel A: Estimated distribution of CATEs for income at age 30



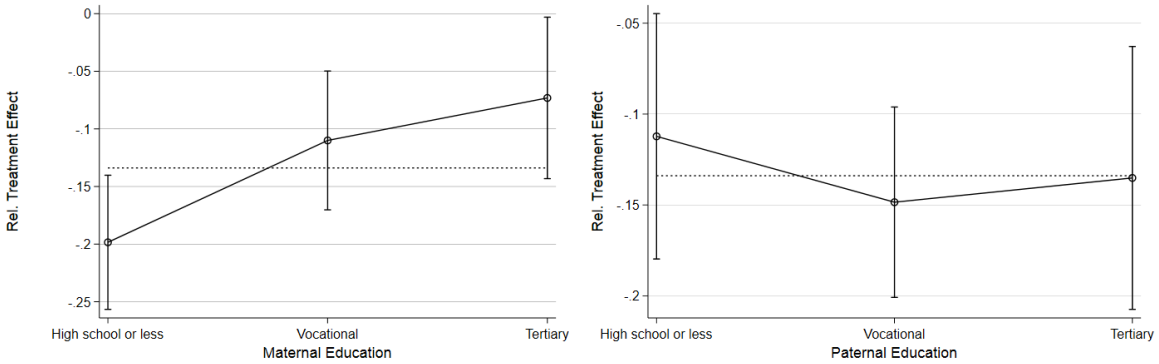
Panel B: Estimated distribution of CATEs for employment at age 30



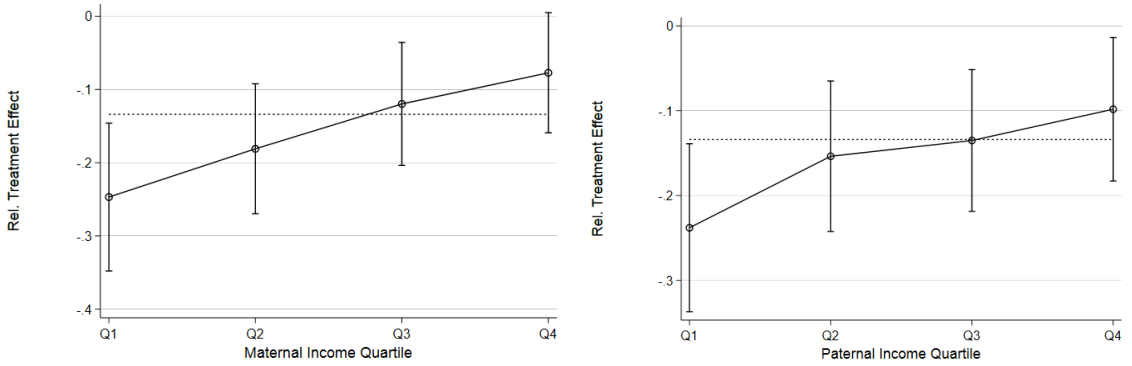
Notes: The solid (red) line indicates the average CATE in the sample. The dashed (blue) line is the treatment effect estimate from an OLS regression including the full set of controls (equal to the coefficient reported in column (4) in table 3). The average CATE is not statistically significantly different from the OLS treatment estimate for any of the outcomes. The estimated distributions are capped at the 1st and 99th percentiles.

Figure 5: Relative treatment effects on income at age 30 by parental characteristics

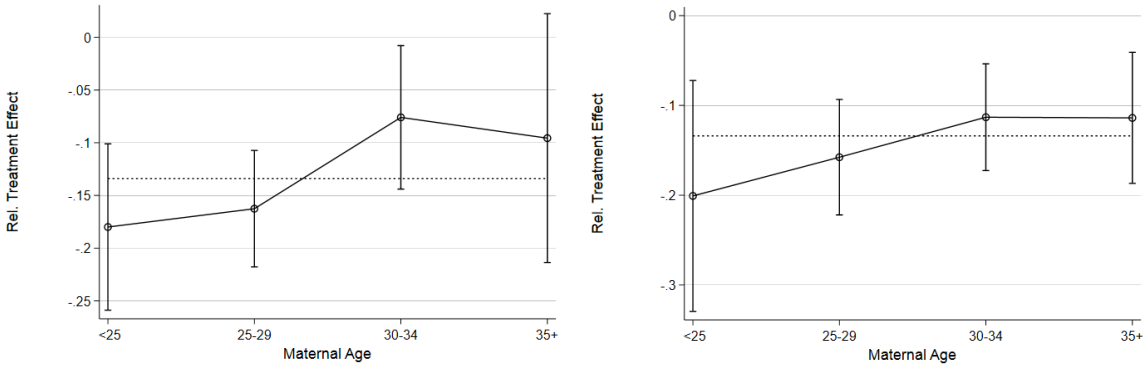
Panel A: Relative treatment effects by parental education



Panel B: Relative treatment effects by parental income quartile



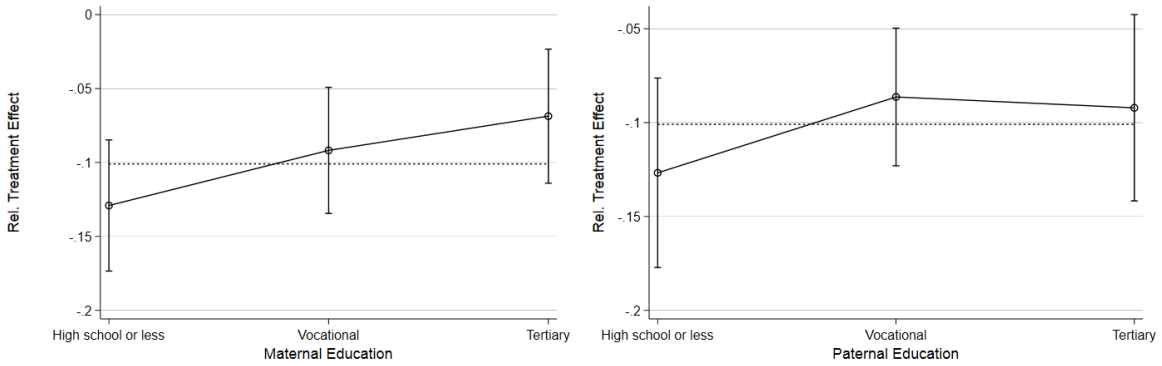
Panel C: Relative treatment effects by parental age



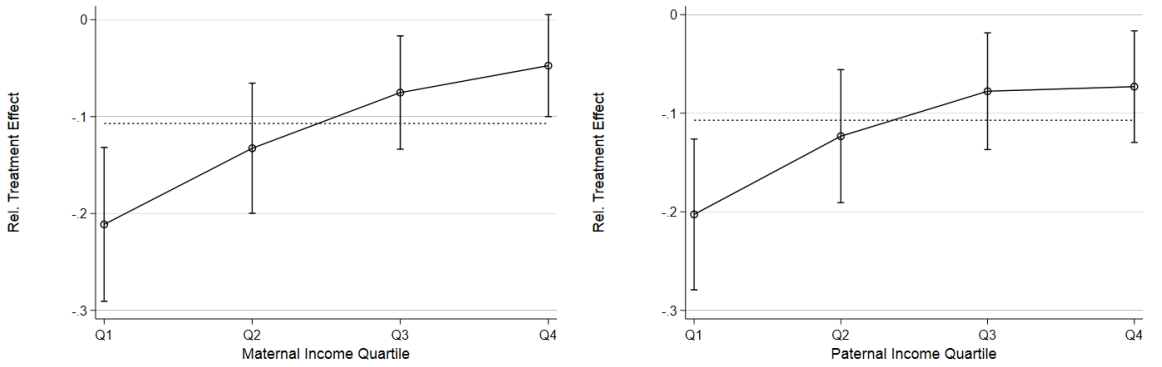
Notes: This figure presents subgroup effects and 95% CI for the impact of diabetes on income at age 30. The graph shows the average treatment effect from the causal forest analysis by maternal education, age, and income quartile. The mean is scaled by the mean outcome in the subgroup. The horizontal line indicates the relative treatment effect for the entire population.

Figure 6: Relative treatment effects on employment at age 30 by maternal characteristics

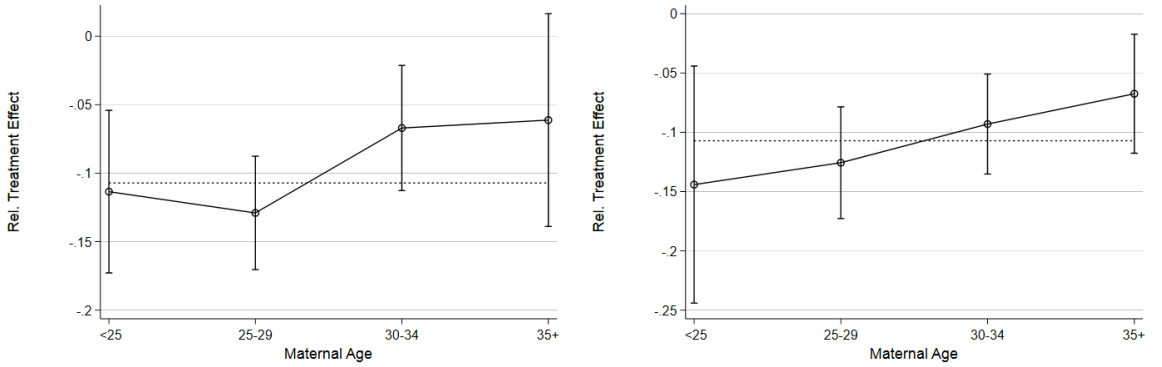
Panel A: Relative treatment effects by parental education



Panel B: Relative treatment effects by parental income quartile



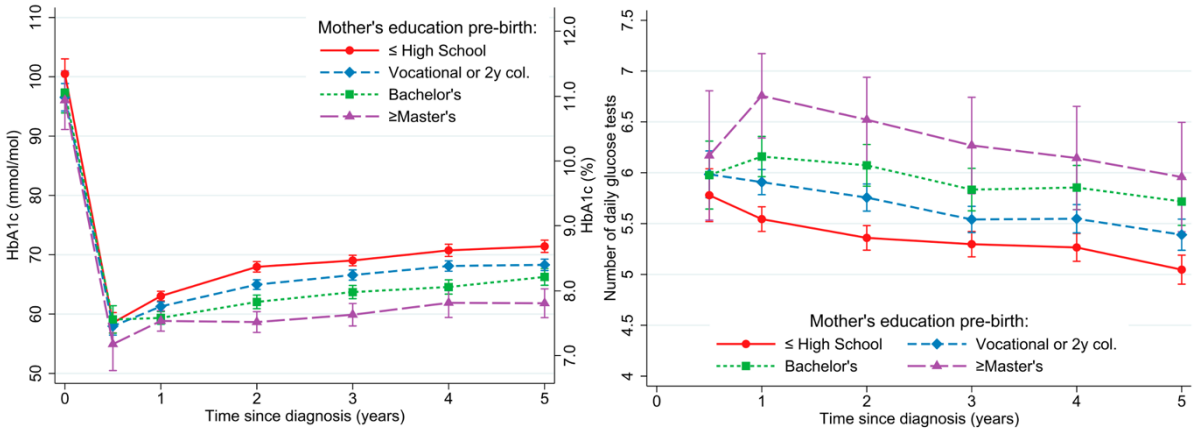
Panel C: Relative treatment effects by parental age



Notes: This figure presents subgroup effects and 95% CI for the impact of diabetes on employment defined as the probability of having any labor market income at age 30. The graph shows the average treatment effect from the causal forest analysis by maternal education, age, and income quartile. The mean is scaled by the mean outcome in the subgroup. The horizontal line indicates the relative treatment effect for the entire population.

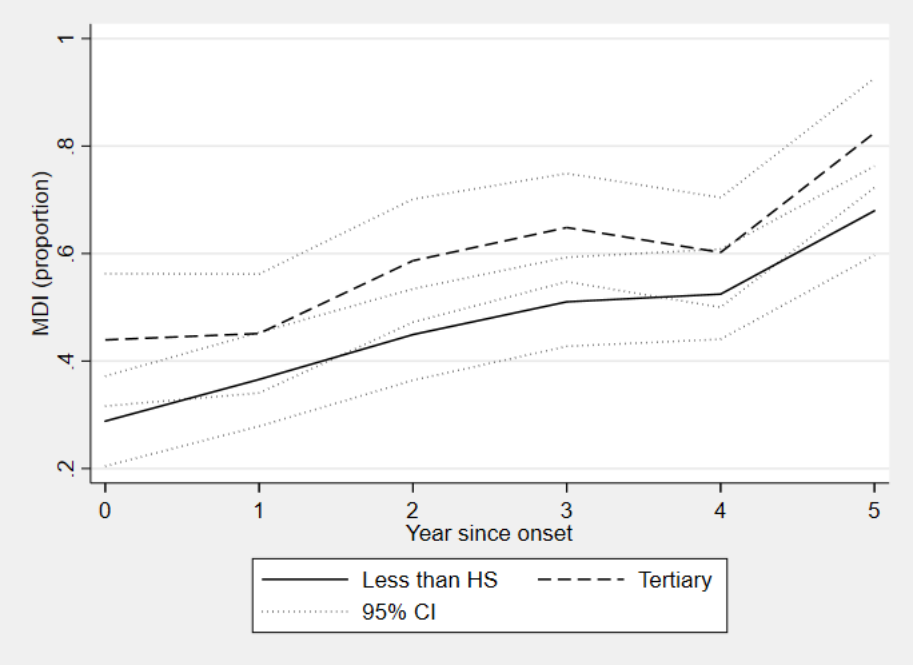
Figure 7: Mean difference (95% CI) levels of HbA_{1c} (Panel A) and number of daily glucose tests (Panel B) by time since T1D onset (years) across maternal education level.

Panel A **Panel B**



Panels A and B were published previously in Nielsen et al. Socioeconomic inequality in metabolic control among children with type 1 diabetes: a nationwide longitudinal study of 4,079 Danish children. Diabetes Care 2019 Aug; 42(8):1398-1405. Copyright 2019 by the American Diabetes Association.

Figure 8: Mean (95% CI) proportion of MDI users by maternal educational level by time since T1D onset (years)

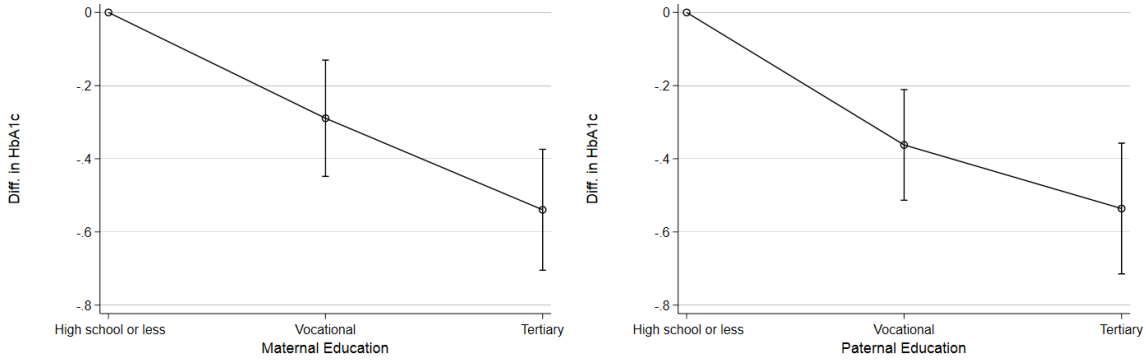


Note: This graph shows the proportion of individuals who were on a multiple daily injections (MDI) treatment regimen during childhood, categorized by maternal educational level and time since onset of diabetes.

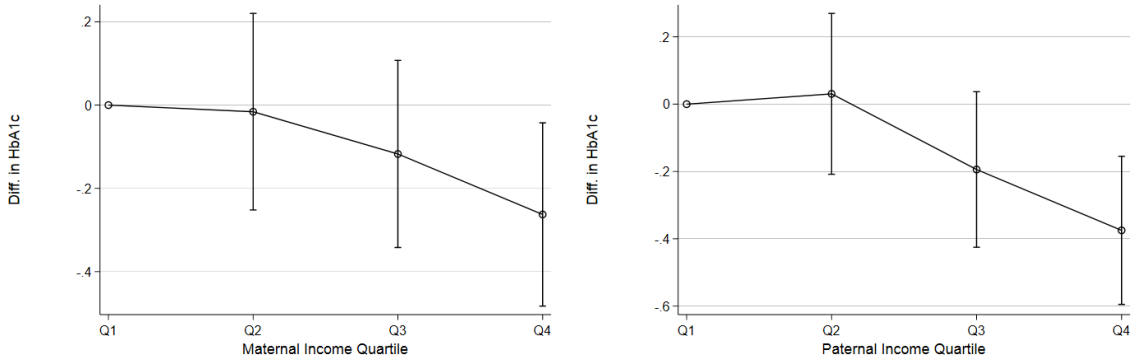
Appendix A : Figures and tables

Figure A1: Difference in HbA_{1C} for People with T1D by parental characteristics

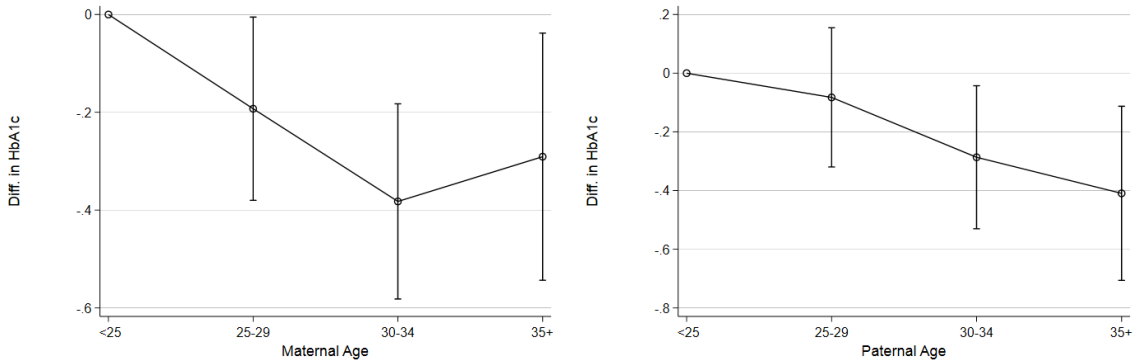
Panel A: Difference by parental education



Panel B: Difference by parental income quartile



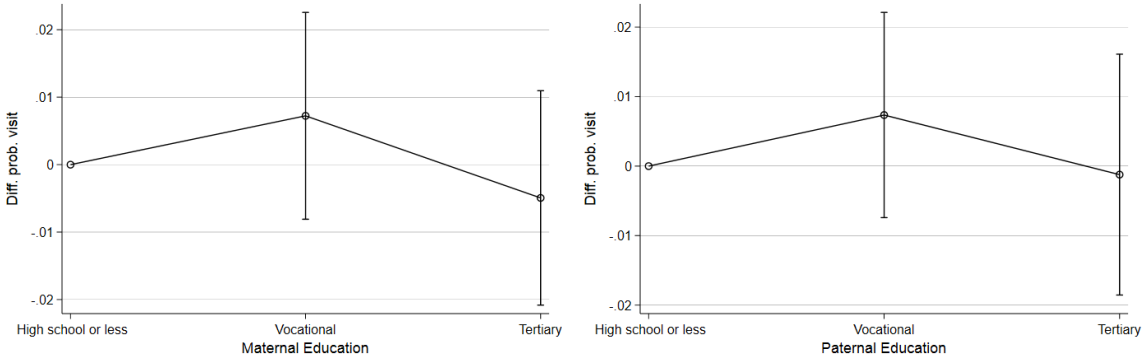
Panel C: Difference by parental age



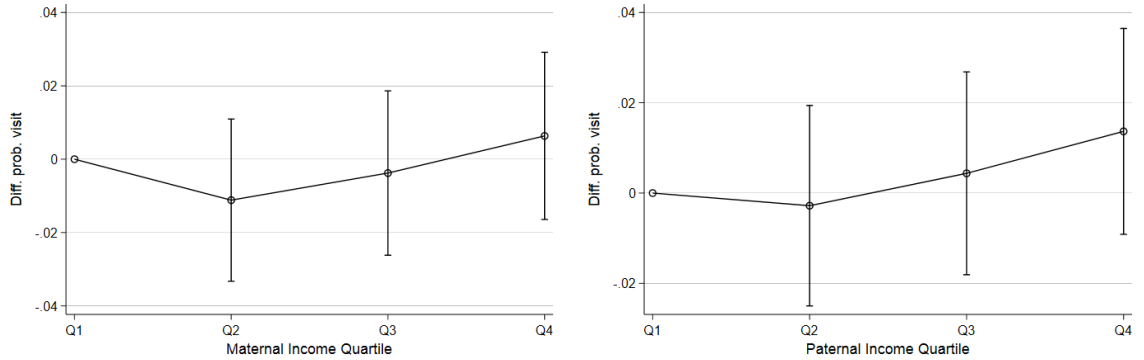
Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. Glycated hemoglobin (HbA_{1c}) is a measure of how well the glucose levels are managed with lower values indicating better disease management. Mean differences relative to the reference group (Panel A: High school or less; Panel B; Q1; Panel C: <25) are reported with 95% CI. The outcome (HbA1c) mean is 8.2.

Figure A2: Difference in probability of receiving ambulatory care related to T1D by parental characteristics

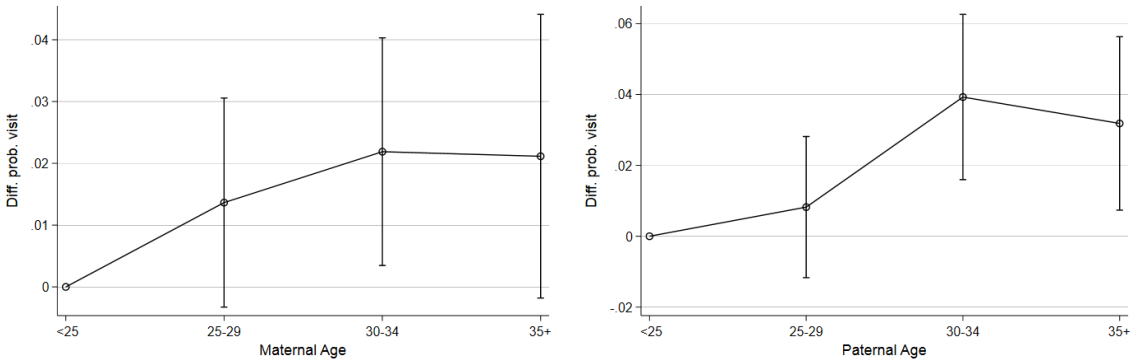
Panel A: Difference by parental education



Panel B: Difference by parental income quartile



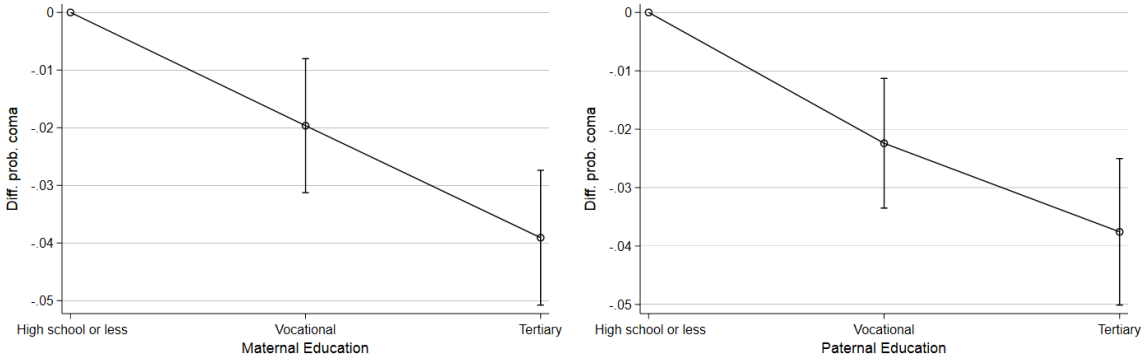
Panel C: Difference by parental age



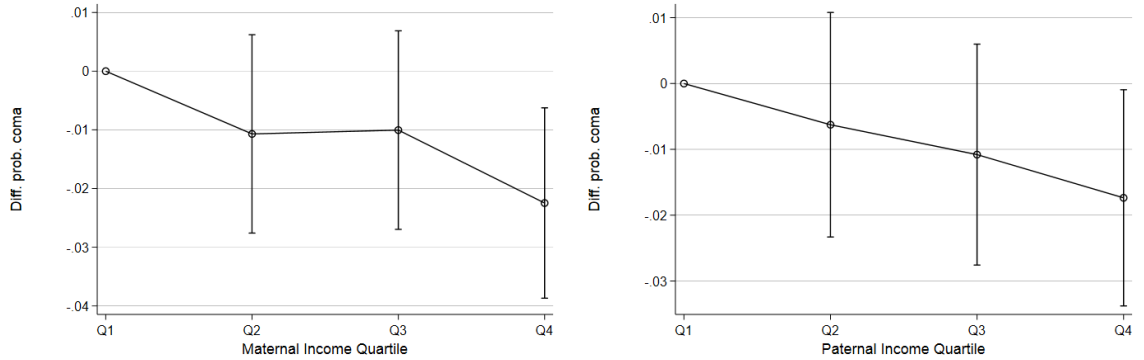
Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. The outcome is the probability of receiving specialized ambulatory care. Mean differences relative to the reference group (Panel A: High school or less; Panel B; Q1; Panel C: <25) are reported with 95% CI. The outcome mean (Probability of ambulatory visit) is 0.77

Figure A3: Difference in probability of hospital admission with diabetes related acute conditions by parental characteristics

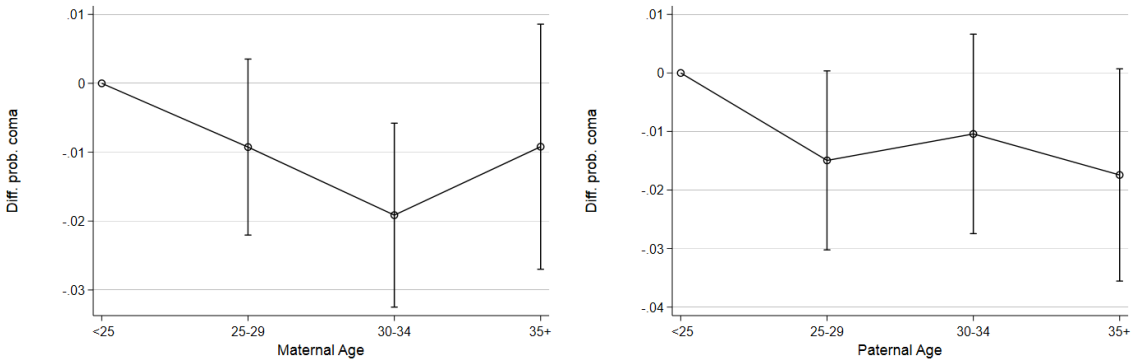
Panel A: Difference by parental education



Panel B: Difference by parental income quartile



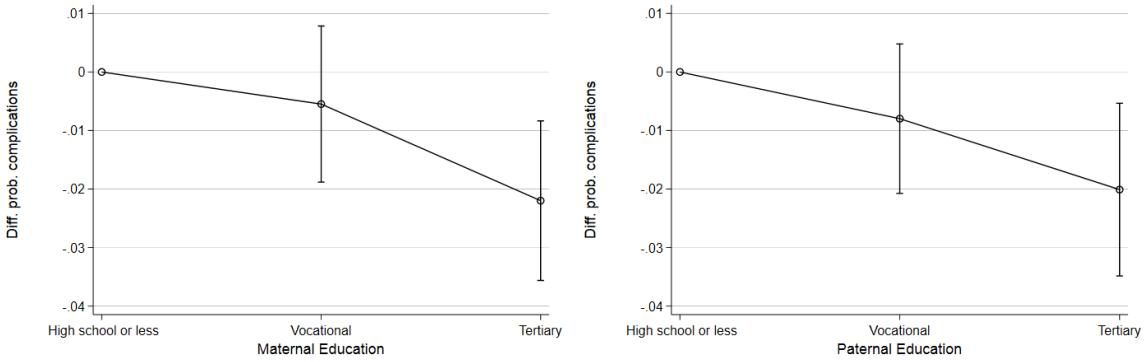
Panel C: Difference by parental age



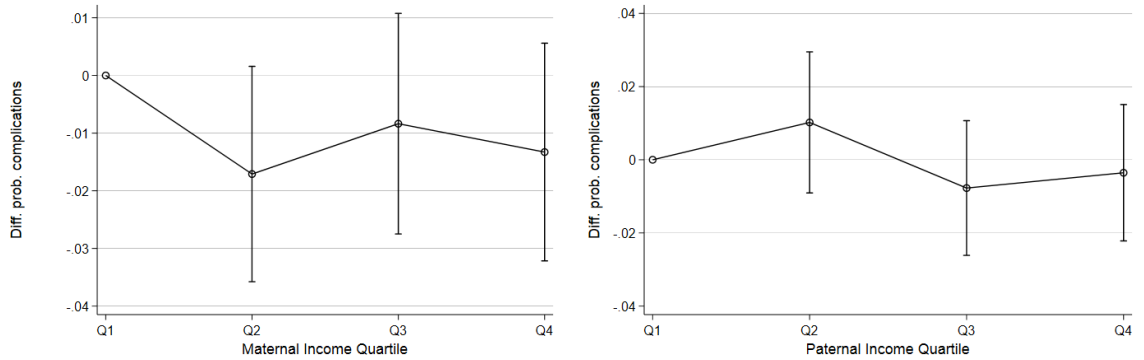
Notes: This figure presents subgroup differences in diabetes related outcomes among the individuals with diabetes. The outcome is the probability of having been admitted to the hospital for diabetes related acute conditions (diabetic ketoacidosis or hypoglycemic coma). Mean differences relative to the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

Figure A4: Difference in probability of diabetes related complications by parental characteristics

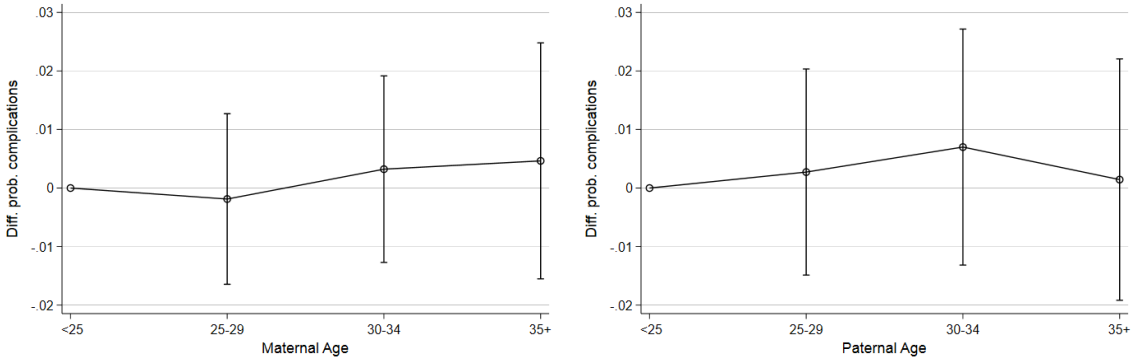
Panel A: Difference by paternal education



Panel B: Difference by paternal income quartile



Panel C: Difference by parental age



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; Panel C: <25) are reported with 95% CI. The outcome mean (probability of diabetes related complications) is 0.49.

Appendix B: Tests for heterogeneity

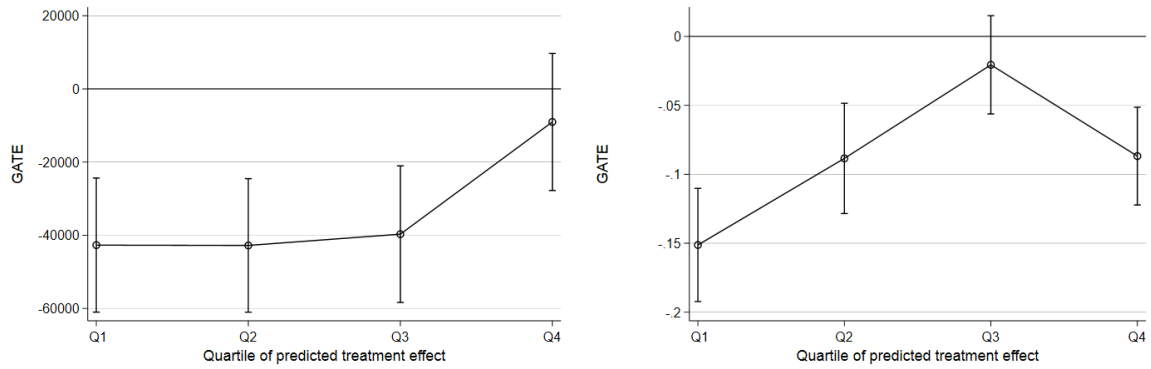
In this appendix, we conduct a series of tests to assess whether the heterogeneity we observe is in fact essential heterogeneity or merely sampling variation. As a first heuristic test, we return to the quartiles of predicted treatment effects outlined in section 5.1. With the data split into groups (quartiles) of predicted treatment effects, we simply 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.[13] 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 the labor market outcome. 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 CIs 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.

[13] 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).

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.

As a second test of heterogeneity, we turn to the best linear predictor (BLP) test formalized in Chernozhukov et al. (2018), 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 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. (2018)). When β is positive, and significantly greater than 0, we can reject the null of no heterogeneity. The results from estimating (7) 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 does 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.

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.