Navarra Center for International Development

Working Paper nº 01/2021

Parental Human Capital Investment Responses to Children's Disability

Anastasia Terskaya

University of Navarra

Navarra Center for International Development

WP-01/2021

Parental Human Capital Investment Responses to Children's Disability*

Anastasia Terskaya †
University of Navarra

March 29, 2021

Abstract

This paper proposes a simple theoretical framework and an estimation strategy aimed at investigating whether parental decisions to invest in the education of disabled children are driven by equality or efficiency. Even if parents are inequality averse, they may still choose to invest more in non-disabled children than in disabled children if there are additional costs of education associated with disability. This implies that comparisons of parental investments across siblings with different health conditions (such as the ones underlying sibling fixed effects models) do not necessary yield an unambiguous conclusion about parental inequality aversion. By means of a general preference model, I show that variation in family size and children's disabilities can be used to infer whether parents are averse to inequality or if, instead, they care more about efficiency. In particular, I exploit the fact that parents of only children cannot possibly exhibit inequality aversion. I apply this identification strategy to Mexican cross-sectional data and find evidence that equality is important for parents.

JEL classification: D13, I12, I14, J13, J14

Keywords: Inequality aversion, disability, parental responses, inta-household allocation, human capital formation, entropy balancing.

*I am grateful to Antonio Cabrales, Lola Collado, Anna Sanz-de-Galdeano, Nuria Rodriguez-Planas for their feedback. I would like to thank participants at the SEHO 2019, ESPE 2018, EALE 2017, 2017 American-European Health Economics Study Group - II Edition Meeting, the Applied Economics Seminar at CRES University Pompeu Fabra (UPF), and the Applied Economics Seminar at the University of Alicante for comments that helped to improve the paper. All possible mistakes are my own.

†Department of Economics, University of Navarra, Campus Universitario, 31009 Pamplona, Spain. email: aterskaya@unav.es

1 Introduction

A growing body of empirical evidence indicates that people with disabilities are more likely to experience social exclusion and socio-economic disadvantage than those without disabilities, especially in developing countries (see Mitra *et al*. 2013; OECD 2009, among others).[1] One of the important channels through which a disability may lead to diminishing well-being and the likelihood of poverty is the lack of education for people with disabilities. According to Filmer (2008), in some developing countries, the school participation deficit associated with a disability is more than 50 percentage points. Although one of the Sustainable Development Goals is to ensure equal access to education for all, including people with disabilities, the disability gap in educational attainment has increased in developing countries during the last few decades and many children with disabilities are never enrolled in school (Male and Wodon, 2017). Since the decision to enroll in school and to continue education are mainly made by parents, understanding how parents respond to children's disabilities might shed light on the primary causes of the disability schooling gap. This is precisely the goal of this paper.

The aim of this work is to infer whether the disability schooling gap can be partially explained by parental responses to children's disabilities. In particular, parents might invest differently in the education of disabled and non-disabled children depending on whether parental behavior is driven by efficiency or equality concerns. If parental decision-making is driven by efficiency concerns, then parents will allocate resources in order to maximize total expected earnings of their children (Behrman *et al*., 1982). In this case, parents may provide more resources to children with higher expected returns from education and, therefore, they may invest less in disabled than in non-disabled children, reinforcing the disability schooling gap. On the other hand, if parental decisions are driven by equality concerns, then parents will allocate resources in order to reduce differences in endowments between their children.

To analyze parental responses to children's disabilities, I provide a simple parental preference model built upon seminal contributions of Becker and Tomes (1976) and Behrman *et al*. (1982). The model allows the cost of education to differ with children's endowment levels, while it also incorporates parental inequality aversion. It predicts that if the cost of education is higher for disabled than for non-disabled individuals, even inequality averse parents might provide more resources to non-disabled than to disabled children. Additionally, the model predicts that parental inequality aversion affects only multi-child families, since parents of single child families do not have other children to whom they can reallocate the resources. Therefore, the disability schooling gap of only children cannot be explained by parental preferences, but, for example, by differences in the costs of education between disabled and non-disabled individuals. In contrast, the disability schooling gap of children from multi-child families can be affected by both parental preferences and the costs of education.

I use these theoretical predictions to build an empirical strategy relying on the variation

[1] The World Health Organization describes disability as an umbrella term for impairments, activity limitations, and participation restrictions as part of a broader classification scheme covering three main domains: body functioning and structure, activities and participation, and environmental factors (https://www.who.int/topics/disabilities/en/).

in the number of children and in the disability status. In particular, under the assumption that the number of children and children's disability status are independent, the presence of parental inequality aversion would imply that the disability schooling gap is lower in multi-child families than in single-child families. It also implies that having a disabled sibling is associated with worse educational outcomes since inequality averse parents reallocate resources from non-disabled toward disabled children. In contrast, if parents care more about efficiency than equality, the disability schooling gap will be lower in single-child families than in multi-child families, and having a disabled child would be associated with better educational outcomes.

Most empirical studies about parental response to differences in children's endowments use sibling or twins comparisons (Aizer and Cunha, 2012; Behrman *et al.*, 1982; Bharadwaj *et al.*, 2018; Cabrera-Hernández and Orraca-Romano, 2016; Datar *et al.*, 2010; Garcia Hombrados, 2017; Grätz and Torche, 2016; Hsin, 2012; Rosales-Rueda, 2014; Yi *et al.*, 2015).[2] That is, these studies compare parental investments between low and high-endowed siblings, mostly using birth weight, health related outcomes, or cognitive ability test scores as endowment measures. Behrman *et al.* (1994) and Savelyev *et al.* (2019) instead rely on comparisons between the within-twin correlations of human capital outcomes of monozygotic and dizygotic twins. Berry *et al.* (2020) conduct a lab-in-the-field experiment to identify parental preferences for equality *verus* efficiency. Their results suggest that parents have strong preferences for equality in inputs as well as for maximizing expected earnings of children.

While sibling or twin comparisons can indicate whether parents follow a "reinforcing" or a "compensating strategy", these models generally cannot distinguish whether parental behavior is driven by parental preferences (*i.e.*, equality versus efficiency concerns), or by differences in the costs of investing in high *versus* low-endowed children.[3] In turn, the difference in costs of education between disabled and non-disabled individuals could be substantial, rendering sibling comparison uninformative about parental preferences. Therefore, this paper contributes to the previous literature by providing an alternative empirical strategy, which, under certain assumptions, allows one to infer the presence or absence of parental inequality aversion while allowing the cost of parental inputs to vary with the levels of children's endowments.

I apply this empirical strategy to a large sample from Mexican Census data. I match individual characteristics using the Entropy Balancing reweighting method (Hainmueller, 2012) in order to achieve a balance of observable characteristics between single-child and multi-child families with disabled and non-disabled children. The results suggest that the disability schooling gap is lower in multi-child families than in single-child families, and that siblings of disabled individuals have lower educational attainment than siblings of non-disabled individuals with similar characteristics. The totality of the evidence is consistent with parental inequality aversion. In particular, parental inequality aversion reduces the disability schooling

[2]See also Almond and Mazumder (2013) for a review of empirical studies on parental responses to endowments.

[3]I say that parents follow a "reinforcing strategy" if they devote more resources to increasing the quality of the better endowed child; parents follow a compensating strategy if they provide more resources to a child with a lower endowment; and parents follow a neutral strategy if they devote equal resources to their children.

gap by about 13 percent and induces a decrease in years of education of non-disabled individuals who have a disabled sibling by about 2 percent. I also explore alternative explanations of these results, such as a spillover effect across siblings and omitted variable bias related to endogenous fertility decisions. Reassuringly, my results are not supportive of these alternative explanations. The heterogeneity analysis suggest that the effect is statistically distinguishable from zero only in males.

The remainder of the paper is organized as follows. The next section presents the parental preference model that guides my empirical analysis. Section 3 lays out the empirical strategy. Section 4 describes the data and Section 5 discusses the results. Section 6 concludes.

2 Theoretical Model

In this section I provide a static parental preference model which motivates and guides my identification strategy. The aim is to show that variation in family size and endowments can be used to test whether parents are inequality averse. This follows from the model's implication that parental aversion to inequality does not affect families with just one child. Therefore, if parents exhibit inequality aversion, then the education gap between low and high-endowed children is lower in multi-child families than in single-child families. The opposite is true if parents care more about efficiency than equality.

2.1 Preference model

Preference models are models of constrained utility maximization where parental preferences —in particular, parental aversion to inequality in the distribution of wealth among their children—play a central role in determining the distribution of parental investments among siblings. The theoretical framework is built upon the classical intra-household allocation models of Becker and Tomes (1976) and Behrman *et al.* (1982). I assume that parental preferences can be represented by the utility function $U_p = U_p(c, V_1, .., V_n)$, where c denotes parental consumption and V_i is the quality of child i. Following Behrman *et al.* (1982), I assume that parental preferences are separable from consumption and, therefore, the problem of parental investment in children can be rewritten as the following utility maximization:

$$U = U(V_1, .., V_n)$$

I specify parental preferences using a CES utility function:

$$U = \left\{ \sum_{i=1}^{n} V_i^{\rho} \right\}^{\frac{1}{\rho}} \tag{1}$$

The main advantage of this utility form is that ρ measures the degree of parental inequality aversion across children. When $0 < \rho < 1$, parents do not dislike inequality and, instead, care about efficiency. In this case, parents follow a "reinforcement strategy." When $\rho < 0$, parents dislike inequality and, hence, are more concerned about equality than efficiency. In this case, parents may compensate the less-endowed child if marginal returns to education are positively correlated with endowments. When $\rho = 0$, parents trade off equality and efficiency.

Following Behrman *et al.* (1982), I assume that a child's quality function has the following form:

$$V(e_i, S_i) = e_i^{\alpha_e} S_i^{\alpha_s} \tag{2}$$

where e_i denotes the endowment of child i and S_i denotes i's years of schooling. Diminishing returns to schooling requires $0 < \alpha_s < 1$, and positive returns to endowments imply that $\alpha_e > 0$. Note that with this function the marginal returns to education are positively related to endowments.

Finally, parental budget constraint has the following form:

$$\sum_{i=1}^{n} p_i S_i = I \tag{3}$$

, where p_i denotes the cost of education of child i, and$I$ denotes total investments in children.[4] Furthermore, I allow the cost of education to differ with children's initial endowments e assuming that $p(e)$ is not increasing in e.[5] In other words, I assume that education is not more costly for children with higher initial endowments. For example, due to reasons of poor health it might be more difficult to choose an appropriate school for a child with a low level of endowments than for a healthy child.

The household's optimization problem yields the following optimality condition:

$$\frac{\partial U}{\partial S_i} / \frac{\partial U}{\partial S_j} = p_i / p_j \tag{4}$$

, which implies that in the families where all children have the same initial endowment ($e_i = e_j \Rightarrow p_i = p_j \; \forall i, j \in 1, ..n$), all children will get the same amount of education.

2.2 Children's endowments and resource reallocation

Consider a population, which consists of two types of children: low-endowed ($T = L$) and high-endowed ($T = H$), such that initial endowments satisfy $e^H > e^L$, and costs of education are such that $p^H \leq p^L$. For simplicity I assume that there can only be one low-endowed child in a family. Therefore, I consider single-child families with a low-endowed child ($F = L1$), single-child families with a high-endowed child ($F = H1$); k-child families with all high-endowed children ($F = Hk$); and k-child families with one low-endowed child and $k - 1$ high-endowed children ($F = Lk$). By S_F^T I denote schooling of children who have a type T from a family F.

[4]Note that both the cost of education and total investment include monetary and non-pecuniary expenditures such as time.

[5]This assumption is weaker than the assumption in Behrman *et al.* (1982) that prices are independent from initial endowment.

In this model intra-household resource allocation depends on sibsize, on the distribution of siblings' initial endowments, on parental preferences, on the costs of education, and on returns to parental investments.

The optimal amount of schooling for low- and high-endowed children from single-child families satisfies the following condition

$$\frac{S_{H1}^{H}}{S_{L1}^{L}} = \frac{p_L}{p_H} \geq 1 \tag{5}$$

Therefore, if the price of education is the same for low- and high-endowed children, then only children should get the same amount of schooling. However, if the costs of schooling depend on initial endowments, the level of education will vary between low- and high-endowed children.

Solving the utility maximization problem for families with k equally high-endowed children yields the following optimal amount of schooling for each child

$$S_{Hk}^{H} = \frac{I}{kp_H} \tag{6}$$

Equations (5) and (6) show that the schooling levels of only children and schooling levels of equally endowed siblings (S_{Hk}^{H}) are not affected by parental inequality aversion (ρ).

The schooling of children from multi-child families where one child has low level of endowments satisfy

$$\frac{S_{Lk}^{L}}{S_{Lk}^{H}} = \gamma \tag{7}$$

, where $\gamma = \left\{ \frac{p_L}{p_H} \left(\frac{e_H}{e_L} \right)^{\rho\alpha_e} \right\}^{\frac{1}{\alpha_s\rho - 1}}$.

If $\gamma < 1$, we say that parents follow a "reinforcing strategy" and provide more schooling to the child with higher endowments. In contrast, when $\gamma > 1$, parents follow a "compensating strategy" since they provide more schooling to the child with lower endowments. If parents care more about efficiency $(0 < \rho < 1)$, they will always follow a "reinforcing strategy" ($\gamma < 1$) since $\alpha_e > 0$, $0 < \alpha_s < 1$, $p_L/p_H \geq 1$, and $e_H/e_L > 1$. When parents are inequality averse ($\rho < 0$), we have that $\left(\frac{e_H}{e_L} \right)^{\rho\alpha_e} < 1$, and γ can be greater or smaller than 1, depending on the relative price of investment and on the degree of inequality aversion. Then, even inequality averse parents may follow a "reinforcing strategy" if the cost of investing in low-endowed children is significantly higher than the cost of investing in high-endowed children. This implies that the comparison of parental investments between siblings with different endowments does not lead to unambiguous conclusions about parental inequality aversion if the cost of one unit of investment depends on children's initial endowment.

The comparison of the educational gap between low- and high-endowed children in single and multi-child families yields

$$\frac{S_{Hk}^{H}/S_{Lk}^{L}}{S_{H1}^{H}/S_{L1}^{L}} = \frac{1}{k} + \frac{(k-1)p_H}{k\gamma p_L} \tag{8}$$

Proposition 1 *In this framework the following conditions hold:*

(i) $\frac{S^H_{Hk}}{S^L_{Lk}} > \frac{S^H_{H1}}{S^L_{L1}}$ *if and only if* $\rho > 0$.

(ii) $\frac{S^H_{Hk}}{S^L_{Lk}} < \frac{S^H_{H1}}{S^L_{L1}}$ *if and only if* $\rho < 0$.

(iii) $\frac{S^H_{Hk}}{S^L_{Lk}} = \frac{S^H_{H1}}{S^L_{L1}}$ *if and only if* $\rho = 0$.

Proof. Let me start by proving item (i). Note that from equation (8) it follows that:

$$\frac{S^H_{Hk}}{S^L_{Lk}} > \frac{S^H_{H1}}{S^L_{L1}} \Leftrightarrow \frac{p_H}{\gamma p_L} = \left\{ \left(\frac{p_L}{p_H}\right)^{\rho\alpha_s} \left(\frac{e_H}{e_L}\right)^{\rho\alpha_e} \right\}^{\frac{1}{1-\alpha_s\rho}} > 1$$

Since $e_H > e_L$, the assumption that prices are not increasing in initial endowments implies that $\frac{p_L}{p_H} \geq 1$. Moreover, returns to initial endowments are positive ($\alpha_e > 0$) and returns to schooling are diminishing ($0 < \alpha_s < 1$). Therefore, if I assume that $\frac{p_H}{\gamma p_L} > 1$, then it must be the case that $\rho \in (0; \frac{1}{\alpha_s})$. However, since $\rho < 1 < 1/\alpha_S$, the last condition is satisfied whenever $\rho > 0$. On the other hand, if $\rho > 0$ then $\frac{S^H_{Hk}/S^L_{Lk}}{S^H_{H1}/S^L_{L1}} > 1$ is trivially satisfied.

The same argument is applied to (ii) and (iii). ■

Proposition 1 demonstrates that the schooling gap is greater in single-child families than in multi-child families when parents are inequality averse. This is due to the fact that inequality averse parents provide some extra inputs to low-endowed children who have high-endowed siblings, while low-endowed only children cannot be possibly affected by inequality aversion. On the other hand, when parents care more about efficiency ($\rho > 0$), the schooling gap is larger in multi-child families than in single-child families since, in this case, parents reallocate resources from low-endowed to highly-endowed children in order to increase the total expected earnings of their children.

For subsequent analyses, let me define the effect of parental inequality aversion on the schooling of low-endowed children who have high-endowed siblings as

$$\theta^{eq} = \left[log(S^L_{Lk}) - log(S^H_{Hk})\right] - \left[log(S^L_{L1}) - log(S^H_{H1})\right] \tag{9}$$

, where $\theta^{eq} = -log(\frac{1}{k} + \frac{(k-1)p_H}{k\gamma p_L})$. From Proposition 1 it follows that this effect is positive if parents are inequality averse ($\theta^{eq} > 0$ iff $\rho < 0$), and it is negative when parents care more about efficiency than equality ($\theta^{eq} < 0$ iff $\rho > 0$).

Another intuitive implication of the model is that when parents are inequality averse, high-endowed children in multi-child families who have a low-endowed sibling get less schooling than high-endowed children in multi-child families who have only high-endowed siblings ($\frac{S^H_{Hk}}{S^H_{Lk}} > 1$ iff $\rho < 0$). Let me define the effect of parental inequality aversion on the schooling of high-endowed children who have low-endowed siblings versus those who have only high-endowed siblings as

$$\psi^{eq} = log(S^H_{Lk}) - log(S^H_{Hk}) \tag{10}$$

It can be shown that $\psi^{eq} > 0$ if and only if $\rho > 0$ (parents value more efficiency than equality) and $\psi^{eq} < 0$ if and only if $\rho < 0$ (parents are inequality averse).

In sum, this theoretical model predicts that when parents are inequality averse (i) the negative effect of low genetic endowments on educational attainment is stronger for children from single-child families than for children from multi-child families; and (ii) high-endowed children get less education if they have low-endowed siblings than if they have high-endowed siblings. The opposite is true when parental decisions are driven by efficiency concerns. I use these predictions of the theoretical model to guide the empirical strategy aiming to infer whether parents are inequality averse or care more about efficiency. In particular, in Section 3 I focus on the empirical identification of θ^{eq} and ψ^{eq}.

3 Empirical Model

In this section, I conduct an empirical analysis of parental responses to children's disabilities, guided and motivated by the theoretical model described above. I use a child's disability as a measure of endowments, so that the presence of a disability is associated with low initial endowments. In particular, I want to test whether the effect of disability on parental inputs differs between only children and children who have siblings.

The framework I propose is particularly useful to study parental responses to children's disabilities since it takes into account that the cost of parental investments in human capital (i.e., education) might be higher for disabled than for non-disabled children. Since access to education for people with disabilities is limited, this will result in higher costs of education for persons with disabilities than for non-disabled persons. Therefore, frequently used empirical models based on sibling comparison would not be able to identify whether parental responses to children's disabilities are driven by efficiency or equality concerns since differences in parental investments (i.e., education) between disabled and non-disabled siblings depend on both parental preferences for equality and the "price effect".

To measure parental inputs, I use educational attainment as Ejrnæs and Pörtner (2004). As in the theoretical model, education depends on disability and family size. I limit my analysis to families with no more than one disabled child.

I specify the model as

$$\begin{aligned} log(S_{i,T,f}) = \alpha_0 + \alpha_1 D_T + \alpha_2 Multi_f + \tilde{\theta}^{eq}(Multi_f \times D_T) + \\ \tilde{\psi}^{eq}(Multi_f \times DS_T) + v_{i,T,f} \end{aligned} \tag{11}$$

, where i indexes an individual, $T = \{D, N, DS\}$ indexes the type of individual with D denoting disabled, N denoting non-disabled, and DS denoting non-disabled with a disabled sibling; and $f = \{1, k\}$ indexes the family size. D_T is a dummy variable that takes the value 1 if the child is disabled, DS_T is a dummy variable that takes the value 1 if the child has a disabled sibling. $Multi_f$ is a multi-child family indicator.

In this model α_1 captures the effect of disability on schooling of only children and α_2 captures the effect of the family size on non-disabled individuals. The main coefficient of interest is $\tilde{\theta}^{eq}$, which captures the difference in the disability schooling gap between single-child families and multi-child families. $\tilde{\psi}^{eq}$ captures the effect of having a disabled sibling on the schooling of non-disabled children relative to non-disabled children who have non-disabled siblings.

Positive $\tilde{\theta}^{eq}$ together with negative $\tilde{\psi}^{eq}$ indicate that parents are inequality averse and provide additional education to disabled children when they have non-disabled children. Since inequality averse parents redistribute resources from non-disabled children to children with disabilities, parental inequality aversion has a negative impact on the education of non-disabled children who have siblings with disabilities ($\tilde{\psi}^{eq} < 0$).

If $v_{i,T,f}$ is uncorrelated with disability status and family size, the OLS estimator of $\tilde{\theta^{eq}}$ is unbiased.

$$\mathbb{E}\left[log(S_{i,D,k}) - log(S_{i,N,k})\right] - \mathbb{E}\left[log(S_{i,D,1}) - log(S_{i,N,1})\right] = \tilde{\theta^{eq}} \tag{12}$$

, which is the empirical equivalent of θ^{eq} as defined in equation (9).

Let me allow $v_{i,T,f}$ to be correlated with disability status and with family size and specify this relationship as follows

$$v_{i,T,f} = \gamma_T + \delta_f + \omega_{T,f} + \epsilon_i \tag{13}$$

, where γ_T denotes the omitted variables that are correlated only with the disability status, δ_f denotes omitted variables that are correlated only with the family size, and $\omega_{T,f}$ denotes omitted variables that are correlated with both family size and disability status, and ϵ_i denotes a disturbance term. Then, the OLS estimator of $\tilde{\theta^{eq}}$ may be biased since

$$\begin{gathered}\mathbb{E}\left[log(S_{i,D,k}) - log(S_{i,N,k})\right] - \mathbb{E}\left[log(S_{i,D,1}) - log(S_{i,N,1})\right] = \\ \tilde{\theta^{eq}} + \mathbb{E}(\omega_{D,k} - \omega_{N,k}) - \mathbb{E}(\omega_{D,1} - \omega_{N,1})\end{gathered} \tag{14}$$

Note that omitted variables that are correlated only with the disability status (γ_T) and only with the family size (δ_f) cannot possibly bias the estimate of $\tilde{\theta^{eq}}$ since equation (11) controls for the disability dummy and for the multi-child family indicator. To estimate consistently $\tilde{\theta^{eq}}$, I need to assume that the unobservable factors affecting schooling may differ for disabled and non-disabled children but this difference cannot depend on family size. For example, disabled individuals may exert less effective effort than non-disabled, and effort levels may also depend on family size. However, the difference between average effective effort exerted by disabled and non-disabled has to be the same in single-child families and in multi-child families. The main identification assumption can be specified as

$$\mathbb{E}(\omega_{D,k} - \omega_{N,k}) - \mathbb{E}(\omega_{D,1} - \omega_{N,1}) = 0 \tag{15}$$

Another coefficient in which I am interested is $\tilde{\psi^{eq}}$ (the effect of inequality aversion on non-disabled individuals with disabled siblings). Note that if $v_{i,T,f}$ is uncorrelated with disability status and family size, then $\tilde{\psi^{eq}} = \mathbb{E}(log(S_{i,DS,k})) - \mathbb{E}(log(S_{i,N,k}))$. However, if $v_{i,T,f}$ is defined as in (13), then

$$\mathbb{E}\left[log(S_{i,DS,k}) - log(S_{i,N,k})\right] = \tilde{\psi^{eq}} + \mathbb{E}(\gamma_{DS} - \gamma_N) + \mathbb{E}(\omega_{DS,k} - \omega_{N,k}) \quad (16)$$

Hence, to consistently identify $\tilde{\psi^{eq}}$, I need to assume that non-disabled individuals who have disabled siblings do not differ in individual and family characteristics from non-disabled individuals who have non-disabled siblings. This assumption may not be plausible and hence one should be cautious about giving $\tilde{\psi^{eq}}$ a causal interpretation of the effect of parental inequality aversion on the education level of non-disabled individuals. Therefore, I will be mainly discussing the effect of parental inequality aversion on the schooling level of disabled individuals who have non-disabled siblings ($\tilde{\theta^{eq}}$) by estimating (11).

4 Data Description and Sample Construction

I use individual and household-level data for Mexico from Integrated Public Use Microdata Series International (IPUMS-I) for 2010. The data was originally produced by the Mexican National Institute of Statistics, Geography, and Informatics. The data set contains information on a wide range of characteristics, including family interrelationships, education, and disability.

For this analysis, I select households with children and the sample of children (those who report to be sons or daughters of the head of household) includes 5,174,463 individuals. About 2.4 percent of individuals report to have some form of disability. I define the number of children in the family as the number of children reported by the mother. Alternatively, I could have defined the number of children as the number of children who live in the household. However, older children are likely to live separately from their parents, and therefore the number of children would be undermeasured. In section 5.2.2 I test whether the results are robust to this alternative definition of the number of children. Since the number of siblings is one of the key variables in this analysis, I eliminate 152,823 observations with missing information on the number of children. I also eliminate 1,447 observations with missing age and 537,646 observations with missing educational attainment. Since the main outcome is years of schooling, I restrict the sample to individuals older than 8 years and younger than 30 years, which leaves me with 3,238,833 observations.[6] Next, I restrict the sample to households with both parents inhabiting and with no more than 4 children, which leaves me with 1,289,545 observations.[7] I select only disabled individuals whose disabilities are of congenial origin in order to address the potential endogeneity of disability and to control for the timing of disability occurrence.

[6]Primary school in Mexico starts when the student's age is 6 or 7. Therefore, I do not consider individuals younger than 7. Note that, usually, when analyzing years of schooling, researchers consider only individuals who have supposedly finished secondary school (older than 15 years old) (Acemoglu and Angrist, 1999; Angrist and Krueger, 1991; Maccini and Yang, 2009). Instead, I also consider younger individuals, since family structure might also have an effect on lower levels of educational attainment. In fact, there is a considerable proportion of disabled individuals in the sample who have never attended school, although secondary school (grade 9) education is compulsory by law in Mexico.

[7]In the main analysis I compare single-child families with two-child families and I also test whether the results hold when I compare single-child families with three-child families and four-child families.

Table 1: Years of Schooling

	(1)	(2)	(3)	(4)	(5)	(6)
		Age 8-30			Age 16-30	
	N	Mean	St. Dev	N	Mean	St. Dev
Disabled	4,472	3.987	3.683	1479	5.732	4.938
Disabled single	789	3.503	3.597	285	4.754	4.742
Disabled two-child	3,683	4.091	3.693	1194	5.965	4.957
Non-disabled	376,176	6.495	4.063	117261	11.160	3.028
Non-disabled single	48,341	6.275	4.041	15276	10.703	3.331
Non-disabled two-child	327,835	6.527	4.065	101985	11.229	2.974

Note: All statistics are for individuals from single-child and two-child families.

Particularly, while disabilities caused by accidents or diseases might be affected by parental investments in education, congenial disabilities cannot possibly be affected by parental post-natal investments. This restriction leaves me with 1,278,845 observations. Finally, I consider only households with no more than one disabled child, leaving 1,275,502 observations from 697,977 households.

Table B.1 reports summary statistics for the final sample. Individuals are, on average, 14.2 years old with 6.7 years of schooling. Most individuals are literate (97.1 percent), 43.2 percent have completed primary education, 10.9 percent have completed secondary education, and 2.8 percent have tertiary education. 47.2 percent of the sample are females. Disabled individuals constitute 1.1 percent of the sample. Most disabled individuals have a mental disability. Most families in my sample are multi-child families (93.3 percent) with 2.8 children on average.

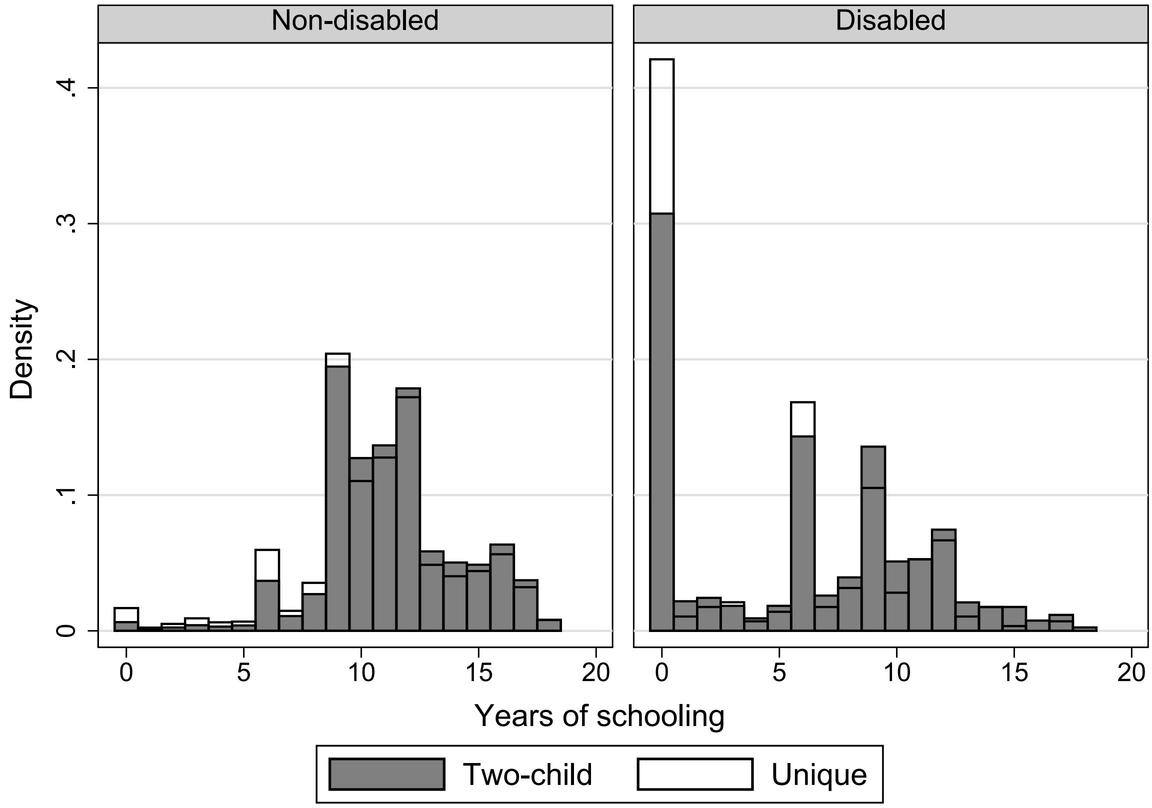
In the main analysis I compare single-child families with two-child families, so that the main estimation sample constitutes 380,648 observations of which 4,472 correspond to disabled individuals.

Table 1 reports the average years of education by sibsize and disability status. It shows that individuals between 8 and 30 years old with congenial disabilities have, on average, 4 years of education, while non-disabled individuals have an average of 6.5 years of education. The disability schooling gap constitutes 5.4 years of schooling for 16-30 year olds.

Figure 1 shows the distribution of years of schooling for disabled and non-disabled individuals from single-child and two-child families aged 16-30. A considerable share of disabled individuals have never attended school. This may suggest that disabled children in Mexico face many barriers to accessing education. The distribution of years of schooling is more right-skewed for disabled only children than for disabled children with siblings. In contrast, for non-disabled individuals, the distribution of years of schooling is similar for single-child and multi-child families.

In sum, this descriptive evidence suggests that the education gap between non-disabled and disabled individuals is lower in multi-child families, which is consistent with parental inequality aversion. However, these differences may be driven by differences in family and individual characteristics, which I take into account in the subsequent analysis.

Figure 1: Years of Education by Disability Status and Family Size



Notes. This table reports the average years of education in single-child and two-child families by disability status.

4.1 Balancing of observable characteristics

To check whether the main identification assumption holds for observable characteristics, I conduct a set of balancing tests. Specifically, I regress each control variable on the disability dummy of the two-child family indicator and on its interaction. Then I test whether the interaction coefficient is statistically distinguishable from zero.

In order to achieve a better balance between observations from different groups I reweight observations, such that the covariate distribution of the control group that maximizes its similarity with the covariate distribution of the treatment group.[8]

To generate weights, I use the entropy balancing (EB) method developed by Hainmueller (2012), which produces a set of observation-level weights that balance covariate distributions across groups. One of the advantages of EB over the popular propensity score weighting (PS) is that EB guarantees that all the covariate moments included in reweighting are equally balanced. In contrast, PS can lead to a worse balance on some covariate moments, while improving balance on others (Iacus *et al.*, 2012). Besides, EB allows one to directly incorporate

[8]In the context of my identification strategy, the control groups are those unaffected by parental inequality aversion: single-child families and multi-child families in which all children have similar initial endowments. Disabled individuals who have non-disabled siblings are potentially affected by parental inequality aversion and hence they constitute the treatment group.

covariate balance, so there is no need to check covariate balance iteratively to avoid model misspecification as with PS (Diamond and Sekhon, 2013; Zhao and Percival, 2016). The details on weights construction are provided in Appendix A.

Table B.2 in the Appendix reports mean and standard deviations of observable characteristics by the disability status in single- and two-child families and the difference-in difference coefficient before and after EB reweighting. Column (5) reports the differences before EB reweighting, suggesting that there are statistically significant differences in age, age of mother and father, and parental disability status. In particular, disabled individuals from single-child families are older than those from two-child families. Column (6) reports the differences after EB reweighting, showing that EB balances the distributions of observable characteristics appropriately since all the differences are not statistically distinguishable from zero.

My identifying assumption (15) also requires that disabled individuals from two-child families are similar to disabled individuals from single-child families. Therefore, I make the comparisons before and after EB reweighting.

Table B.3 in the Appendix compares the distributions of the type of disability in single- and two-child families. It is crucial to control for the type of disability because the cost of education is likely to depend on it. For instance, disabled individuals from single-child families may have fewer years of schooling than those from two-child families simply because their disability is more severe. I reweight observations for disabled individuals from single-child families so that the distribution of their type of disability is similar to the distribution of the type of disability of disabled individuals from two-child families. Differences after EB reweighting are reported in the last column of Table B.3 and, reassuringly, they are no longer significant. In section 5.1.4 I conduct the analysis according to the disability type separately.

5 Results

In this section I estimate equation (11) for single-child and two-child families.

Table 2 provides the estimates before and after EB reweighting. Column (1) of Table 2 provides the estimated results without including additional controls and before EB. Column (2) provides the results after including controls, and Columns (3-4) report the estimates after EB reweighting with and without controls respectively. The main coefficient of interest is the coefficient of the interaction term $Two-child \times Disabled$, which indicates that the gap in schooling between disabled and non-disabled children is, on average, 13.5 percent smaller in two-child families than in single-child families. The estimates of the coefficient for $Two-child \times Disabled$ reduce after observations are reweighted with EB weights. On the other hand, the inclusion of controls does not affect the magnitude of the estimates.

In order to assess the degree to which differences in unobservable characteristics may drive the results, I follow the methodology proposed by Oster (2019) and compute the relative degree of selection on unobservables (with respect to observables). The Oster ratios are reported in Table 2. I find that the influence of unobserved factors would have to be at least 26.8 times stronger than the influence of observed factors (listed in Tables B.2 and B.3) in order to explain

Table 2: The Effect of Disability and Family Size on Schooling

	(1)	(2)	(3)	(4)
	OLS		OLS with EB weights	
Two-child × Disabled	0.165***	0.163***	0.131***	0.135***
	(0.044)	(0.041)	(0.049)	(0.042)
Disabled Sibling	-0.000	-0.012**	-0.015	-0.019***
	(0.010)	(0.005)	(0.011)	(0.006)
Disabled	-0.836***	-0.302***	-0.797***	-0.338***
	(0.040)	(0.054)	(0.045)	(0.063)
Two-child	0.053***	0.022***	0.012***	0.011***
	(0.004)	(0.002)	(0.004)	(0.002)
Blind		0.079*		0.118**
		(0.040)		(0.055)
Deaf		-0.002		0.044
		(0.049)		(0.071)
Mute		-0.561***		-0.485***
		(0.039)		(0.051)
Lower extremities		-0.229***		-0.217***
		(0.040)		(0.054)
Mental		-0.909***		-0.840***
		(0.040)		(0.053)
Covariates included	No	Yes	No	Yes
State fixed effects	No	Yes	No	Yes
N	380,648	380,648	380,648	380,648
R^2	0.011	0.795	0.140	0.539
Oster ratios				
Two-child × Disabled		232.262		-26.812
Disabled Sibling		-3.925		-4.350

Notes: The reported estimates correspond to regressions of the inverse hyperbolic sine one transformation of years of schooling (used to approximate log of schooling) on the disability dummy, the indicator of two-child family, the indicator of disabled sibling, and the interaction between the disability dummy and the two-child family dummy. Columns (2) and (4) also include the set of controls listed in Table B.2, age fixed effects, and state fixed effects. In Columns (2)-(3) observations are weighted using EB weights. Standard errors clustered at household level in parentheses. Oster ratios are relative degrees of selection under proportional selection of observable and unobservable factors computed as proposed by Oster (2019). $^{*}\ p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

away the interaction coefficient $Two - child \times Disabled$.

The estimated effect of having a disabled sibling on schooling of non-disabled individuals (Row 2 of Table 2) is negative and implies that non-disabled children who have disabled siblings receive on average 1.9 percent less education than non-disabled children who have no disabled siblings. The Oster ratio for the disabled sibling indicator is approximately 4, suggesting that the influence of unobserved factors would have to be at least 4 times stronger than the influence of observed factors in order to explain away the effect of having a disabled sibling.

Overall, the results are consistent with parents being inequality averse.

5.1 Heterogeneous parental response

5.1.1 By gender

In this section I analyze whether the effect of parental inequality aversion differs by children's gender. The effect may vary by gender if parents are not gender neutral. For instance, Dahl and Moretti (2008) have found evidence supporting the notion that parents in the U.S. favor boys over girls. In contrast, Baccara *et al.* (2014) have identified significant preferences favoring girls. Behrman *et al.* (1986) have shown that parental preferences either slightly favor girls or are neutral.

On the other hand, the effect of parental inequality aversion on schooling may depend on children's gender when there are gender differences in the returns to parental inputs even if parental preferences are gender neutral. In fact, there is evidence that cognitive and noncognitive development of boys is more responsive to parental inputs than that of girls (Bertrand and Pan, 2013; Brooks-Gunn *et al.*, 2002; Hill and Duncan, 1987; Leibowitz, 1974; Moore *et al.*, 2004). For instance, Bertrand and Pan (2013) show a substantial part of the gender gap in disruptive behaviors can be explained by gender differences in returns to parental inputs.

In addition, the evidence on gender differences in emotionality and sociability in children with disabilities suggests that females with autism display better social skills than males with autism (Head *et al.*, 2014; Lai *et al.*, 2011). Therefore, for males with disability, any kind of social interaction, such as those involved with school attendance, can be costlier than for females. This can be incorporated into the model by allowing the non-pecuniary costs of education for disabled females to be lower than those for disabled males. Then, the model predicts a greater effect of parental inequality aversion on educational attainment for males than for females, since the effect of inequality aversion ($\tilde{\theta}^{eq}$) on the cost of education for disabled children increases when parents exhibit inequality aversion (see equation (9)).

I report the estimates of equation (11) in males and females separately in Columns (1-2) of Table 3. The results suggest that the interaction coefficient ($Two-child \times Disabled$) is large and statistically significant in males but not statistically distinguishable from zero for females.

5.1.2 By birth order

A number of empirical studies predict a negative relationship between birth order and parental investments (Behrman and Taubman, 1986; Black *et al.*, 2005; Iacovou, 2001; Lehmann *et al.*, 2018; Price, 2008). I explore whether the effect of parental inequality aversion differs by birth order by reestimating equation (11) for first and second-born children separately. The results are reported in Columns (3-4) of Table 3. The interaction coefficient appears to be statistically significant and positive for both firstborn and later-born children. However, the

Table 3: Heterogeneous Effect of Disability and Family Size on Schooling

	(1)	(2)	(3)	(4)	(5)	(6)
	By Gender		By Birth Order		By Family Size	
	Females	Males	Firstborn	Second-born	Three-child	Four-child
Multi-child×Disabled	0.094	0.176**	0.187***	0.092*	0.146***	0.106**
	(0.066)	(0.070)	(0.052)	(0.050)	(0.046)	(0.051)
Disabled Sibling	-0.018*	-0.018**	-0.016**	-0.028***	-0.015***	-0.019***
	(0.010)	(0.008)	(0.007)	(0.010)	(0.005)	(0.006)
Disabled	-0.420***	-0.296***	-0.363***	-0.239***	-0.390***	-0.494***
	(0.099)	(0.095)	(0.075)	(0.077)	(0.068)	(0.076)
Multi-child	0.015***	0.008***	0.011***	0.001	0.010***	0.003
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Blind	0.240***	-0.058	0.150**	0.034	0.135**	0.154**
	(0.078)	(0.089)	(0.062)	(0.073)	(0.061)	(0.069)
Deaf	0.197*	-0.022	-0.048	0.060	-0.036	0.093
	(0.117)	(0.085)	(0.099)	(0.086)	(0.084)	(0.087)
Mute	-0.562***	-0.459***	-0.407***	-0.462***	-0.487***	-0.523***
	(0.084)	(0.069)	(0.062)	(0.062)	(0.053)	(0.057)
Lower extremities	-0.152*	-0.183**	-0.247***	-0.302***	-0.245***	-0.196***
	(0.082)	(0.080)	(0.065)	(0.070)	(0.056)	(0.062)
Mental	-0.826***	-0.851***	-0.831***	-0.733***	-0.865***	-0.918***
	(0.082)	(0.078)	(0.062)	(0.067)	(0.056)	(0.060)
N	178,668	201,977	238,886	151,035	573,005	420,109
R^2	0.549	0.573	0.551	0.542	0.526	0.514

Notes: The reported estimates correspond to regressions of the inverse hyperbolic sine one transformation of years of schooling (used to approximate log of schooling) on the disability dummy, on the indicator of multi-child families, the indicator of disabled siblings, and the interaction between the disability dummy and the multi-child family dummy. All specifications include the set of controls listed in Table B.2, age fixed effects, and state fixed effects. Observations are weighted using EB weights. Standard errors clustered at household level in parentheses. * $p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

effect of parental inequality aversion on the educational attainment of firstborn children is greater than the effect on their second-born counterparts. Specifically, the estimates imply that the disability schooling gap is 18.7 percent smaller for firstborns who have non-disabled siblings than for only children. For later-born children, this difference constitutes 9.2 percent. These results are in line with findings in the economic literature showing a negative correlation between birth order and children's human capital outcomes (Black *et al.*, 2005; Lehmann *et al.*, 2018; Rosales-Rueda, 2014).

5.1.3 By family size

Up to now, this paper has compared the educational gap between non-disabled and disabled in single-child families and two-child families. To analyze whether the effect of parental inequality aversion varies with the family size, I estimate equation (11) for three and four-child families. The resulting estimates are reported in Columns (1-2) of Table 3, suggesting that the disability schooling gap is 14.6 percent smaller for children from three-child families than for only children and 10.6 percent smaller for children from four-child families than for only children. The results are consistent with parental inequality aversion.

5.1.4 By type of disability

In order to analyze whether parents respond differently to different types of disabilities I reestimate equation (11) separately for mental and non-mental types of disability, and for blind, mute, or deaf individuals.

The results reported in Table 4 suggest that the interaction term is not statistically distinguishable from zero when only mental disabilities are considered. This might be due to the fact that we do not observe the severity of mental disability, but only the presence or absence of such disability. While, for instance, blindness does not vary in severity, the severity of mental disability might vary substantially. In fact, if children with a mental disability from multi-child families have, on average, more severe conditions than those from single-child families, this would introduce a negative bias on the estimate of the interaction term. On the other hand, if disabled only children have more severe conditions, this would introduce a positive bias on the estimate of the interaction term.

Column 2 of Table 4 reports the results for children with non-mental disabilities, suggesting that in this subsample the interaction term is positive and statistically different from zero.

Finally, Column (3) of Table 4 reports the results for blind, mute, or deaf individuals. I analyze this group separately because these disability types do not vary in severity and, therefore, the cost of education for individuals with these types of disabilities cannot vary with the family size. The estimated interaction term is positive and significant at the 10 percent level of confidence.[9]

5.2 Robustness checks

5.2.1 Parental Disability

Column 5 of Table B.2 shows than maternal and paternal disability indicators failed the balancing test before applying EB weights. In this section I test whether the results hold if I drop families with disabled parents from the analysis. Column 1 of Table B.4 reports the estimated coefficients for the sample with non-disabled parents. The estimated effect of inequality aversion on education for the disabled reported in the first row remains positive and statistically significant. The effect of having a disabled sibling (reported in the second row) is negative and statistically significant. This suggests that the previously documented differences

[9] In all regressions I control for the type of disability using the EB weights so the share of mute, blind and deaf individuals are constrained to be equal in single-child and multi-child families.

Table 4: The Effect of Disability and Family Size on Schooling by Type of Disability

	(1) Mental	(2) Non-Mental	(3) Blind or Mute or Deaf
Two-child $\times$ Disabled	0.056	0.138***	0.106*
	(0.087)	(0.051)	(0.060)
Disabled Sibling	-0.023***	-0.017***	-0.019***
	(0.008)	(0.006)	(0.006)
Disabled	-1.226***	-0.505***	-0.459***
	(0.083)	(0.048)	(0.058)
Two-child	0.014***	0.008***	0.007**
	(0.003)	(0.003)	(0.003)
N	377,881	378,934	377,989
R^2	0.522	0.520	0.538

Notes: The reported estimates correspond to regressions of the inverse hyperbolic sine one transformation of years of schooling (used to approximate log of schooling) on the disability dummy, the indicator of two-child family, the indicator of disabled sibling, and the interaction between the disability dummy and the two-child family dummy. All specifications include the set of controls listed in Table B.2, age fixed effects, and state fixed effects. Observations are weighted using EB weights. Standard errors clustered at household level in parentheses. * $p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

in education between disabled only children and disabled children who have siblings are unlikely to be driven by differences in parental disability status.

5.2.2 Sibsize Definition

In the main analysis I have defined the number of children as the number of children reported by their mother. However, if some children do not live with their parents, parents might not react to the differences between these children the same way they react to the differences between children who live with them. To address this point, I reestimate equation (11) for families with all the children residing in the household. The results are reported in Column (2) of Table B.4 and these estimates are similar to the estimates in the full sample.

5.2.3 Sibling spillover effects

The model described above has one important limitation: it is unable to distinguish between parental inequality aversion and the direct sibling spillover effects. Specifically, there might be a direct effect of children on their siblings in addition to the indirect effect that is mediated by intra-household allocation of parental investments across siblings. In fact, changes in parental investments can be seen as a mechanism through which sibling spillovers are working.

It is difficult to separate the effect of parental inequality aversion and the direct sibling spillover effect empirically without detailed information about parental investments and sibling interaction. Specifically, the identification problem arises if sibling spillovers affect disabled and non-disabled children unequally. In particular, if the sibling spillover effect is positive and greater in magnitude for disabled children than for non-disabled children (or negative and smaller in magnitude for disabled children), then the estimated effect of parental inequality aversion will be inflated. In contrast, if the sibling spillover effect is positive and smaller in magnitude on disabled children than on non-disabled children (or negative and larger in magnitude on disabled children than on non-disabled children), then the estimated effect of parental inequality aversion on schooling will be attenuated.

There are several studies that analyze sibling spillover effects on school achievement. Black *et al.* (2021) is particularly relevant since they analyze the effect of having a younger sibling with a disability. Specifically, the authors use the fact that birth order influences the amount of time which a child spends with their siblings and compare second-born and firstborn children who have a third-born sibling with a disability to children who have a non-disabled third-born sibling. Their results indicate that the second-born child is negatively affected by a younger sibling's disability. However, their approach does not allow the effect of parental investments to be separated from the direct sibling spillover effect. In fact, the authors argue that the effect is at least partially mediated by changes in parental investment.

Qureshi (2011), Nicoletti and Rabe (2014),Joensen and Nielsen (2018), and Karbownik and Özek (2019) find positive spillover effects from older to younger siblings on school achievement, but not *vice versa*. Blake *et al.* (1991) did not find any significant effect of the number of siblings on sociability.

Interestingly, Hindes (2006) has found that children with mental disability experience lower rates of positive peer experiences and higher rates of negative peer experiences than their non-disabled counterparts. For my results this would imply underestimation of the effect of parental inequality aversion. However, if sibling spillovers have a stronger effect on younger siblings that on older siblings, as was found in Joensen and Nielsen (2018); Karbownik and Özek (2019); Nicoletti and Rabe (2014); Qureshi (2011), then comparing estimates of $\tilde{\theta^{eq}}$ by birth order may be useful to assess the direction of this bias. Results from Columns (3-4) of Table 3 suggest that $\tilde{\theta^{eq}}$ is actually larger for older children, who are likely to be less affected by sibling spillovers, than for younger children. This finding suggests that the effect of parental inequality aversion might be underestimated for later-born children. Hence, my estimates would be conservative.

5.2.4 Exogenous fertility

The main identification assumption requires that all unobservable family characteristics that affect schooling are not correlated with family size and children's disability status. However, this assumption can be violated if fertility decisions are partially determined by a firstborn's disability status.

To take the problem of endogenous fertility into account, I present results for the sample of second-born children only in Table 3 and the estimated effects are similar in magnitude to the ones obtained using my main specification. The similarity of the point estimates in the sample of second-born and in the full sample suggests that the results for the full-sample are unlikely to be biased.

Another way to address this issue is to fix the fertility decision by considering as multi-child families only families with twins. This method is frequently used in the economic literature in order to achieve exogenous variation in family size (see Black *et al*. 2005; Rosenzweig and Zhang 2009; Yi *et al*. 2015, among others). Therefore, I consider the following groups: non-disabled individuals from single-child families; disabled individuals from single-child families; non-disabled individuals who have non-disabled twins; non-disabled individuals who have disabled twins; and disabled individuals who have no disabled twins. The problem with this approach is that the size of the treatment group (disabled individuals who do not have a disabled twin) consists of only 30 observations. I provide the estimation results for this sample in Column 3 of Table B.4. The point estimates for the sample of twins are almost identical to the estimates from Table 1. However, the estimated effect of parental inequality aversion is not statistically significant given the small number of disabled twins. The point estimates in the sample of twins have similar sign and magnitude as the estimates in the full sample, which suggests that the main results are unlikely to be biased.

5.2.5 Falsification Tests

Finally, I run a set of placebo tests to verify that the results I obtain are not driven by pure chance. To do so, I randomly assign to each child a placebo disability status using the observed probability of disability. Specifically, I generate a binomial random variable that takes value one with a probability equal to the share of disabled individuals in the sample. Using this placebo variable, I estimate equation (11) including controls. I repeat this procedure 500 times,

The distribution of placebo t-values is illustrated in Figure B.1 in the Appendix. In fewer than 5 percent of cases the estimated interaction coefficient is significant at the 5 percent level, which suggests that the results are unlikely to be driven by chance.

6 Conclusions

In this paper I study how parents respond to their children's disability and, in particular, whether parents are averse to inequality in the distribution of quality among their children. I assess the impact of parental responses to their children's disability on the children's education. By means of a general preference model, I show that the variation in disability status and in family size can be used to infer the presence of parental inequality aversion. This is due to the fact that only parents from multi-child families whose children have different endowments can possibly display inequality aversion.

My theoretical framework and identification strategy take into account that the cost of adding to quality may depend on children's initial endowments, an issue that cannot be addressed when considering siblings comparisons. When I apply this identification strategy to Mexican data, the results indicate that parents in Mexico are subject to inequality aversion and, therefore, they attenuate the negative effects of disability on their children's educational attainment. However, parental inequality aversion significantly affects only the educational attainment of males. I also show that, in Mexico, the education gap between disabled and non-disabled individuals constitutes about 5.4 years of schooling, which may have dramatic implications for the labor market outcomes of disabled individuals.

One of the limitations of this analysis is that the results for Mexico cannot be generalized to other countries, since parental preferences for equality may differ across countries depending on pension systems, culture, informal institutions, and other factors. However, the method of testing for parental inequality aversion proposed in this paper can be easily applied to other contexts, since it requires data that are generally easily available.

Lastly, policymakers may want to take into account that, in the presence of inequality aversion, compensatory education policies may be less effective for disadvantaged children from multi-child families than for those from single-child families, whose parents are not subject to inequality aversion. This is due to the fact that by increasing the endowment of disabled individuals from multi-child families, such policies induce inequality-averse parents to redistribute resources away from the child being compensated and toward themselves or other children.

References

Acemoglu, D. and Angrist, J. (1999): "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws," Working Paper 7444, National Bureau of Economic Research.

Aizer, A. and Cunha, F. (2012): "The Production of Human Capital: Endowments, Investments and Fertility," NBER Working Papers 18429, National Bureau of Economic Research, Inc.

Almond, D. and Mazumder, B. (2013): "Fetal Origins and Parental Responses," *Annual Review of Economics*, 5, 37–56.

Angrist, J. D. and Krueger, A. B. (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106, 979–1014.

Baccara, M., Collard-Wexler, A., Felli, L. *et al.* (2014): "Child-Adoption Matching: Preferences for Gender and Race," *American Economic Journal: Applied Economics*, 6, 133–58.

Becker, G. S. and Tomes, N. (1976): "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84, S143–62.

Behrman, J. R., Pollak, R. A. and Taubman, P. (1982): "Parental Preferences and Provision for Progeny," *Journal of Political Economy*, 90, 52–73.

——— (1986): "Do Parents Favor Boys?" *International Economic Review*, 27, 33–54.

Behrman, J. R., Rosenzweig, M. R. and Taubman, P. (1994): "Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment," *Journal of Political Economy*, 102, 1131–1174.

Behrman, J. R. and Taubman, P. (1986): "Birth order, schooling, and earnings," *Journal of Labor Economics*, 4, S121–S145.

Berry, J., Dizon-Ross, R. and Jagnani, M. (2020): "Not Playing Favorites: An Experiment on Parental Fairness Preferences," Tech. rep., National Bureau of Economic Research.

Bertrand, M. and Pan, J. (2013): "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior," *American Economic Journal: Applied Economics*, 5, 32–64.

Bharadwaj, P., Eberhard, J. P. and Neilson, C. A. (2018): "Health at birth, parental investments, and academic outcomes," *Journal of Labor Economics*, 36, 349–394.

Black, S. E., Breining, S., Figlio, D. N. *et al.* (2021): "Sibling spillovers," *The Economic Journal*, 131, 101–128.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005): "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *Quarterly Journal of Economics*, 120, 669–700.

Blake, J., Richardson, B. and Bhattacharya, J. (1991): "Number of siblings and sociability," *Journal of Marriage and the Family*, 271–283.

Brooks-Gunn, J., Han, W.-J. and Waldfogel, J. (2002): "Maternal employment and child cognitive outcomes in the first three years of life: The NICHD study of early child care," *Child Development*, 73, 1052–1072.

Cabrera-Hernández, F. and Orraca-Romano, P. (2016): "The accident of birth: effect of birthweight on educational attainment and parent's compensations among siblings," Working paper, Centro de Investigación y Docencia Económicas. Available at https://www.academia.edu/34744698/The_accident_of_birth_effects_of_birthweight_on_educational_attainment_and_parents_compensations_among_siblings.

Dahl, G. B. and Moretti, E. (2008): "The Demand for Sons," *Review of Economic Studies*, 75, 1085–1120.

Datar, A., Kilburn, M. and Loughran, D. (2010): "Endowments and parental investments in infancy and early childhood," *Demography*, 47, 145–162.

Diamond, A. and Sekhon, J. S. (2013): "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies," *Review of Economics and Statistics*, 95, 932–945.

Ejrnæs, M. and Pörtner, C. C. (2004): "Birth order and the intrahousehold allocation of time and education," *Review of Economics and Statistics*, 86, 1008–1019.

Filmer, D. (2008): "Disability, Poverty, and Schooling in Developing Countries: Results from 14 Household Surveys," *The World Bank Economic Review*, 22, 141–163.

Garcia Hombrados, J. (2017): "Cognitive Skills and Intra-Household Allocation of Schooling," Working paper, Department of Economics, University of Sussex Business School. Available at https://www.sussex.ac.uk/webteam/gateway/file.php?name=wps-18-2017.pdf&site=24.

Grätz, M. and Torche, F. (2016): "Compensation or reinforcement? The stratification of parental responses to children's early ability," *Demography*, 53, 1883–1904.

Hainmueller, J. (2012): "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, 20, 25.

Head, A. M., McGillivray, J. A. and Stokes, M. A. (2014): "Gender differences in emotionality and sociability in children with autism spectrum disorders," *Molecular Autism*, 5, 19.

Hill, M. S. and Duncan, G. J. (1987): "Parental family income and the socioeconomic attainment of children," *Social Science Research*, 16, 39–73.

Hindes, A. R. (2006): "The Buffering Effect of Sibling Relationships on Problems with Peer Experiences and Psychological Functioning in Children with Cognitive Disabilities," Working paper, Georgia State University. Available at https://scholarworks.gsu.edu/psych_diss/20/.

Hsin, A. (2012): "Is Biology Destiny? Birth Weight and Differential Parental Treatment," *Demography*, 49, 1385–1405.

Iacovou, M. (2001): "Family composition and children's educational outcomes," Tech. rep., ISER Working Paper Series.

Iacus, S. M., King, G. and Porro, G. (2012): "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 20, 1.

Joensen, J. S. and Nielsen, H. S. (2018): "Spillovers in education choice," *Journal of Public Economics*, 157, 158–183.

Karbownik, K. and Özek, U. (2019): "Setting a good example? Examining sibling spillovers in educational achievement using a regression discontinuity design," Tech. rep., National Bureau of Economic Research.

Lai, M.-C., Lombardo, M. V., Pasco, G. *et al.* (2011): "A behavioral comparison of male and female adults with high functioning autism spectrum conditions," *PloS One*, 6, e20835.

Lehmann, J.-Y. K., Nuevo-Chiquero, A. and Vidal-Fernandez, M. (2018): "The early origins of birth order differences in children's outcomes and parental behavior," *Journal of Human Resources*, 53, 123–156.

Leibowitz, A. (1974): "Home investments in children," *Journal of Political Economy*, 82, S111–S131.

Maccini, S. and Yang, D. (2009): "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, 99, 1006–26.

Male, C. and Wodon, Q. T. (2017): "Disability gaps in educational attainment and literacy (English). The price of exclusion : disability and education," Tech. rep., Washington, D.C. : World Bank Group. Available at http://documents.worldbank.org/curated/en/396291511988894028/Disability-gaps-in-educational-attainment-and-literacy.

Mitra, S., Posarac, A. and Vick, B. (2013): "Disability and poverty in developing countries: a multidimensional study," *World Development*, 41, 1–18.

Moore, Q., Schmidt, L. *et al.* (2004): "Do Maternal Investments in Human Capital Affect Children's' Academic Achievement," Tech. rep., Citeseer.

Nicoletti, C. and Rabe, B. (2014): "Sibling spillover effects in school achievement," *Journal of Applied Econometrics*.

OECD (2009): "Sickness, disability and work: keeping on track in the economic downturn," Tech. rep.

Oster, E. (2019): "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 37, 187–204.

Price, J. (2008): "Parent-child quality time does birth order matter?" *Journal of Human Resources*, 43, 240–265.

Qureshi, J. (2011): "Additional Returns to Investing in Girls Education: Impact on Younger Sibling Human Capital," *PhD thesis, Harris School of Public Policy Studies. University of Chicago*.

Rosales-Rueda, M. F. (2014): "Family investment responses to childhood health conditions: Intrafamily allocation of resources," *Journal of Health Economics*, 37, 41–57.

Rosenzweig, M. R. and Zhang, J. (2009): "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy," *Review of Economic Studies*, 76, 1149–1174.

Savelyev, P. A., Ward, B., Krueger, R. F. *et al.* (2019): "Health Endowments, Schooling Allocation in the Family, and Longevity: Evidence from US Twins," Working paper, The College of William and Mary, Department of Economics. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3193396.

Yi, J., Heckman, J. J., Zhang, J. *et al.* (2015): "Early Health Shocks, Intra-household Resource Allocation and Child Outcomes," *The Economic Journal*, 125, F347–F371.

Zhao, Q. and Percival, D. (2016): "Entropy balancing is doubly robust," *Journal of Causal Inference*.

Appendix A Entropy Balancing Weights

In the EB, every control unit $(i|C)$ gets a weight that satisfies a set of balance constraints and each treated unit $(i|T)$ gets either a weight $\omega_i = 1$ or $\omega_i = s_i$, where s_i is a sampling weight associated with i. Specifically, the weights for each control unit are chosen in order to minimize the loss function:

$$\min_{\omega_i} H(\omega) = \sum_{i|C} h(\omega_i) = \sum_{i|C} \omega_i log(\omega_i / s_i) \tag{A.1}$$

subject to balance and normalizing constraints

$$\sum_{i|C} \omega_i c_{ri}(X_i) = m_r \text{ with } r \in 1, ..., R \text{ and} \tag{A.2}$$

$$\sum_{i|C} \omega_i = 1 \tag{A.3}$$

$$\omega_i \geq 0 \text{ for all } i|C \tag{A.4}$$

, where $c_{ri}(X_i) = m_r$ describes a set of R balance constraints imposed on the covariate moments of the reweighted control group. For this analysis, a balance constraint is formulated with m_{rj} containing the rth order moment of a given variable x_j for the treatment group (group of disabled individuals from multi-child families), whereas the moment functions $c_{ri}(X_i)$ are specified for the control group. Therefore, weights (ω_i) are chosen in a way that the weighted $1_{st}, .., R_{th}$ moments of the covariates in the control group are equal to the correspondent moments of the covariates in the treatment group. The loss function $H(\omega)$ measures the distance between the distribution of estimated control weights $(\omega = [\omega_1, .., \omega_n])$ and the distribution of the base weights ($S = [s_1, ..s_n]$); it is non-negative and it decreases the closer ω is to S (the unconstrained minimum would be achieved at zero if $\omega = S$). These properties of the loss function imply that while weights are adjusted as far as needed to fulfill the balance constraints (A.2), they are maintained as close as possible to the base weights to sustain information about the control group. Therefore, another advantage of EB over PS is that with EB extreme weights are less likely.

I apply the EB algorithm to generate four sets of weights for each of the following control groups: (C_1) disabled individuals from single-child families; (C_2) non-disabled individuals from multi-child families who have no disabled siblings; (C_3) non-disabled individuals from single-child families; and (C_4) non-disabled individuals from multi-child families who have disabled siblings. As the treatment group (T) I consider disabled individuals from multi-child families. Then I use the obtained EB weights to estimate coefficients from equation (11) by OLS. I impose the EB constraints on the first and second moments of observable family and individual characteristics listed in Table B.2 and age fixed effects.

Appendix B Tables and Figures

Table B.1: Summary Statistics

	Mean	St. Dev.
Panel A: Individual Level Statistics		
Years of Schooling	6.675	3.977
Literate	0.971	0.166
Less than primary	0.431	0.495
Primary	0.432	0.495
Secondary	0.109	0.311
Tertiary	0.028	0.164
Age	14.260	4.930
Age 8-15	0.647	0.478
Age 16-30	0.353	0.478
Female	0.472	0.499
Disabled	0.011	0.103
Mental	0.004	0.064
Blind	0.002	0.044
Deaf	0.001	0.030
Mute	0.003	0.055
Lower extremities	0.002	0.050
N	1,275,502	
Panel B: Household Level Statistics		
Number of children	2.773	0.874
One-child family	0.067	0.250
Two-child family	0.320	0.467
Three-child family	0.386	0.487
Four-child family	0.227	0.419
Rural	0.410	0.492
Maternal Age	38.389	8.349
Paternal Age	41.773	9.382
Years of Schooling of Mother	7.607	4.212
Years of schooling of Father	7.932	4.481
Mother is Employed	0.300	0.458
Father is Employed	0.896	0.305
Mother's earnings (1000 peso)	1.249	4.259
Father's earnings (1000 peso)	4.779	8.964
Mother is Disabled	0.025	0.156
Father is Disabled	0.042	0.200
Dwelling ownership	0.845	0.361
Car ownership	0.458	0.498
N	697,977	

Note: sample mean and standard deviation are reported.

Table B.2: Balancing Tests

	(1)	(2)	(3)	(4)	Difference-in-Difference	
	Non-disabled		Disabled		$\{(4)-(3)\}-\{(2)-(1)\}$	
	Unique	Multi	Unique	Multi	Before EB	After EB
	mean/sd	mean/sd	mean/sd	mean/sd	beta/se	beta/se
Age	13.791	13.752	14.701	14.096	-0.566**	-0.005
	(5.063)	(4.855)	(5.853)	(5.302)	(0.227)	(0.228)
Female	0.467	0.470	0.417	0.414	-0.006	-0.000
	(0.499)	(0.499)	(0.493)	(0.493)	(0.019)	(0.022)
Rural	0.357	0.335	0.401	0.390	0.012	-0.000
	(0.479)	(0.472)	(0.490)	(0.488)	(0.019)	(0.022)
Dwelling ownership	0.818	0.835	0.803	0.806	-0.014	-0.000
	(0.385)	(0.371)	(0.398)	(0.395)	(0.016)	(0.017)
Car ownership	0.469	0.511	0.390	0.433	0.001	-0.000
	(0.498)	(0.499)	(0.487)	(0.495)	(0.019)	(0.022)
Mother's age	39.126	38.236	41.510	38.843	-1.776***	-0.013
	(9.492)	(7.710)	(11.132)	(8.406)	(0.422)	(0.333)
Mother's years of schooling	8.311	8.748	7.170	7.834	0.227	-0.003
	(4.700)	(4.277)	(4.789)	(4.094)	(0.185)	(0.166)
Mother's personal earnings	1.827	1.735	1.286	1.293	0.099	-0.000
	(4.977)	(5.155)	(4.006)	(4.396)	(0.162)	(0.213)
Mother is Employed	0.360	0.357	0.341	0.321	-0.016	-0.000
	(0.479)	(0.478)	(0.474)	(0.466)	(0.019)	(0.020)
Mother has no health insurance	0.314	0.283	0.291	0.269	0.008	-0.000
	(0.463)	(0.450)	(0.453)	(0.443)	(0.018)	(0.019)
Mother is disabled	0.032	0.021	0.124	0.072	-0.041***	-0.000
	(0.177)	(0.143)	(0.330)	(0.258)	(0.013)	(0.010)
Father's age	43.047	41.393	45.267	42.363	-1.250***	-0.014
	(10.557)	(8.702)	(11.930)	(9.478)	(0.455)	(0.386)
Father's years of schooling	8.420	9.026	7.163	8.004	0.235	-0.003
	(4.843)	(4.544)	(4.816)	(4.403)	(0.188)	(0.186)
Father's personal earnings	5.251	5.826	3.945	4.576	0.056	-0.001
	(9.506)	(10.160)	(5.678)	(6.543)	(0.234)	(0.319)
Father is employed	0.888	0.912	0.854	0.892	0.014	-0.000
	(0.314)	(0.283)	(0.352)	(0.309)	(0.014)	(0.013)
Fatherr has no health insurance	0.337	0.306	0.338	0.299	-0.008	-0.000
	(0.472)	(0.460)	(0.472)	(0.458)	(0.019)	(0.020)
Father is disabled	0.047	0.034	0.145	0.096	-0.036***	-0.000
	(0.212)	(0.180)	(0.352)	(0.294)	(0.013)	(0.012)
N	48341	322764	789	3683		

Note: Siblings of disabled individuals are excluded. Standard deviation and standard errors clustered at household level in parentheses for columns (1-4) and (5-6) respectively. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Balancing Tests. Type of Disability

	(1) Unique-child family	(2) Multi-child family	Difference (2)-(1) Before EB	After EB
Type of Disability	mean/sd	mean/sd	beta/se	beta/se
Blind	0.190	0.201	0.011	-0.000
	(0.393)	(0.401)	(0.015)	(0.018)
Deaf	0.077	0.087	0.010	0.000
	(0.267)	(0.283)	(0.011)	(0.013)
Mute	0.275	0.260	-0.015	-0.000
	(0.447)	(0.439)	(0.017)	(0.019)
Lower extremities	0.274	0.240	-0.033*	-0.000
	(0.446)	(0.427)	(0.017)	(0.018)
Mental	0.411	0.376	-0.035*	-0.000
	(0.492)	(0.484)	(0.019)	(0.021)
N	789	3683		

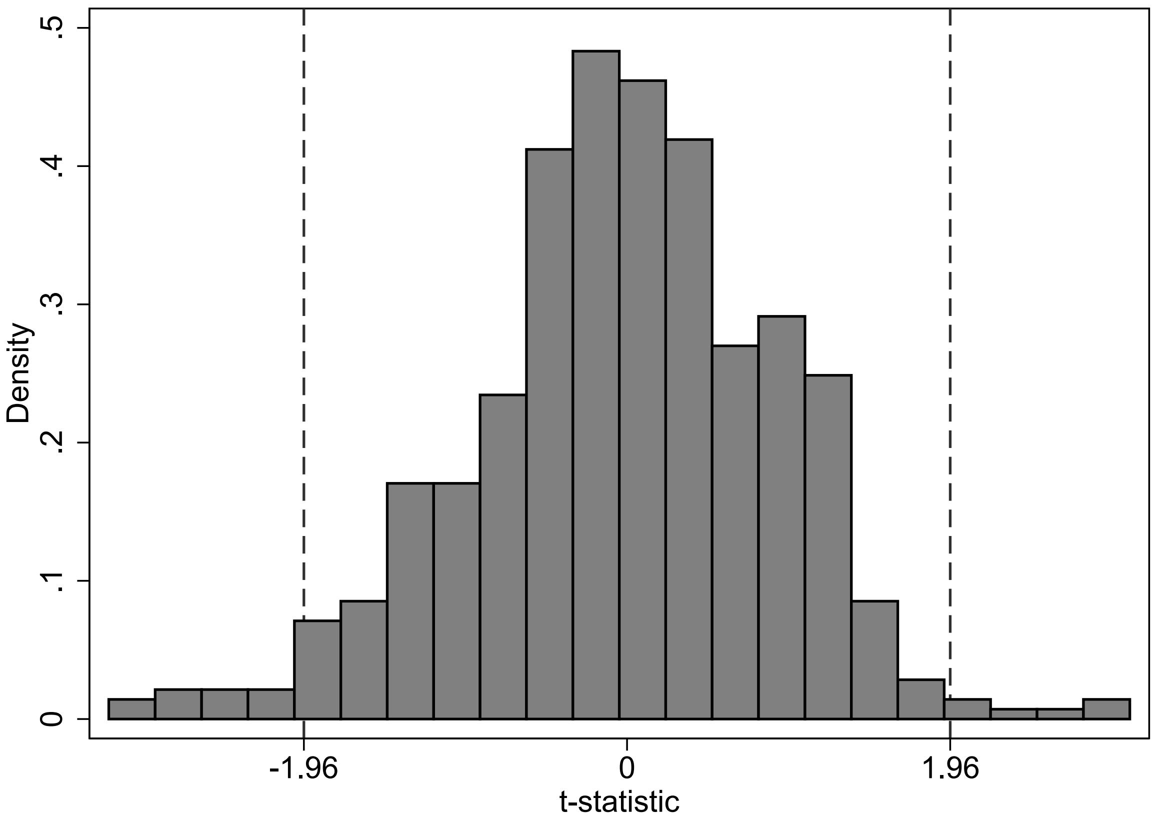
Note: Standard deviation and standard errors clustered at household level in parentheses for columns (1-2) and (3-4) respectively. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4: Robustness Checks

	(1) No disabled parents	(2) All children in the household	(3) Twins vs. Unique
Two-child×Disabled	0.151***	0.152***	0.083
	(0.047)	(0.045)	(0.141)
Disabled Sibling	-0.024***	-0.023***	-0.084
	(0.007)	(0.006)	(0.093)
Disabled	-0.383***	-0.315***	-0.289*
	(0.068)	(0.067)	(0.152)
Two-child	0.013***	0.007***	0.024
	(0.002)	(0.002)	(0.027)
Blind	0.154***	0.107*	-0.156
	(0.059)	(0.058)	(0.224)
Deaf	0.023	-0.020	0.359**
	(0.080)	(0.084)	(0.155)
Mute	-0.454***	-0.429***	-0.774***
	(0.054)	(0.055)	(0.159)
Lower extremities	-0.255***	-0.264***	-0.157
	(0.059)	(0.058)	(0.157)
Mental	-0.807***	-0.795***	-0.884***
	(0.057)	(0.056)	(0.172)
N	357,345	344,137	51,164
R^2	0.538	0.547	0.596

Notes: The reported estimates correspond to regressions of the inverse hyperbolic sine one transformation of years of schooling (used to approximate log of schooling) on the disability dummy, the indicator of two-child family, the indicator of disabled sibling, and the interaction between the disability dummy and the two-child family dummy. All specifications include the set of controls listed in Table B.2, age fixed effects, and state fixed effects. The estimates in Column (1) are for sample that excludes disabled parents. The estimates in Column (2) are for households where all children inhabit in the household. The estimates in Column (3) are for household with twins or with unique children. Observations are weighted using EB weights. Standard errors clustered at household level in parentheses. * $p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure B.1: Distribution of Placebo t-values



Notes. This graph reports the distribution of the t-values of the test $\tilde{\theta}^{eq} = 0$ obtained when estimating 500 placebo regression of equation (11). To obtain placebo values of the disability status, the actual value is replaced by randomly chosen.