The causal effect of family difficulties during childhood on adult labour market outcomes

Emanuele Millemaci*
Università di Messina

Dario Sciulli[§]
Università di Chieti-Pescara and CEEAplA

July 8, 2010

Abstract

A propensity score matching approach and the National Child Development Study cohort database are used to evaluate the total causal effect of family difficulties during childhood on adult labour market outcomes. We find statistically significant evidence of a negative and long-lasting impact on employment probabilities and wages. Our estimates suggest that the occurrence of family problems in childhood reduces the chances of being employed by about 6 % and employees' hourly wages by about 8.4 percent. Moreover, this effect appears not to decline over the cohort working life. Looking at specific family difficulties, we find that economic difficulties determine the greatest disadvantages in terms of future labour market outcomes. These results are consistent with respect to estimations with standard parametric methods. Economic and social policies aimed to prevent poor labour market performances, and possibly consequent social exclusion and immobility, in adulthood, should also take into account the role of the various factors affecting family environment during childhood.

Keywords: family difficulties, childhood, propensity score matching, labour market outcomes, causal effects.

JEL codes: J12, J13, C21

---

*Corresponding author. Dipartimento di Economia "V. Pareto", Università degli studi di Messina, via Cannizzaro 278, 98100 Messina. Tel. +39 090 6411070, E-mail: millemaci@unime.it.

[§]DMQTE, Università di Chieti-Pescara, E-mail: d.sciulli@unich.it .

I. Introduction

Individuals experiencing disadvantages in the childhood are possibly exposed to direct and indirect consequences in their life development leading to negative performances in terms of different life outcomes, as cognitive/non-cognitive developments, educational attainments, labour market outcomes and social behaviors.

In this paper, using data from the National Child Development Study (NCDS), a unique cohort database of British individuals born in 1958, we focus on the long-term effects of family difficulties during childhood on labour market outcomes, in different instants of the working life. The long-term implication of negative events during childhood, including family difficulties, may be relevant not only for its direct effects on life outcomes, but they also may represent some of the underlying motives leading to social immobility and social exclusion.

Possibly due to the reaffirmed relevance in economic studies of children development to determine life outcomes (Cunha and Heckman, 2006), as well as the increasing availability of long panel datasets or cohort studies, the interest on the long standing impact of events experienced during childhood and/or adolescence on economic outcomes during adolescence and/or adulthood has significantly increased in the last years. This has resulted in a number of empirical studies focusing on the long-term effect of many aspects of children development, using different econometric techniques. Possibly the more studied issues have concerned the long-term effects of health problems (or disability) and of educational patterns on different economic outcomes, as educational achievements, cognitive and non-cognitive developments, health and labour market outcomes in adulthood. In this context, Case, Fertig and Paxson (2005) quantify the lasting effects of childhood health and economic circumstances on adult health, employment and socioeconomic status. Lindeboom, Llena-Nozal and van der Klaauw (2006) find that early childhood conditions are important in explaining adult health and socioeconomic outcomes. Smith (2009) investigates the impact of childhood health on adult socio-economic status using information on siblings. Chevalier and Viitanen (2003), investigate the long-run labour market effects of teenage motherhood. Fletcher and Wolfe (2008) examine the effect of mental health during childhood on human capital accumulation, while Glewwe et al. (2001) focused on the relationship between early childhood nutrition and academic achievements. About educational patterns and cognitive/non-cognitive development, Goodman and Sianesi (2005) have evaluated the effects of undergoing any early education and of pre-school on cognitive and non-cognitive abilities, educational attainment and labour market performances, while Heckman, Stixrud and Urzua (2006) have examined the effects of cognitive and non-cognitive abilities on labour market performances and social behavior. Other

economists have paid attention to the effect of antisocial behavior or conduct disorder on human capital formation (Koning et al., 2010), labour market outcomes (Healey, Knapp and Farrington, 2004) or both (Le et al., 2005). Slade and Wissow (2007) have investigated the role of victimization during childhood on academic performance during adolescence. Other authors have focused on the long-term effect of family environment on adult life outcomes. In this context economists have mainly focused on the effects of family structure changes, due to parents' death or parental divorce, on different life consequences as marital/fertility status, earning and income (Corak, 2001), education and income (Gruber, 2004) and students' performance (Sanz de Galdeano and Vuri, 2007).

Most of these empirical studies have shown that events during childhood have a relevant impact on economic outcomes in adulthood, even though the econometric model adopted play a role in determining the magnitude and the significance of the estimation results.

We are not the first using the family difficulties NCDS data to on the long-term effects of family difficulties during childhood using NCDS data. For example, Goodman and Sianesi (2005), estimating the impact of early education, uses family difficulties information in a comparative perspective, to highlight the relative importance of pre-school treatment on cognitive and non-cognitive development with respect to family difficulties, father's social class and mother's years of education. Their OLS estimates show that family difficulties are responsible of the greatest negative impact on cognitive and non-cognitive development. Gregg and Machin (2000) examine the relationship between childhood disadvantages, including financial difficulties and father's unemployment, and life outcomes (educational attainments, juvenile delinquency and labour market performances) in early adulthood. Using standard OLS and Probit models, they find evidence that both financial troubles and father's unemployment tend to reduce educational and labour performances and increase the contact with police.

Our paper contributes to the literature focusing on the long-run labour market consequences of family difficulties in the childhood in various aspects. First, we examine the labour market outcomes in different points in time of the individuals' adult life to study the evolution of the family difficulties impact. Second, besides of considering the family difficulties as a unique problem, we also consider them separately, identifying four homogenous groups according to the nature of the family difficulty. Specifically, we single out economic difficulties, health (and disability) difficulties of family members, family structure difficulties (including parental divorce), and social isolation difficulties and compare them to identify heterogeneous effects according to the type of problem. Third, besides of using standard econometric models in a comparative perspective, we

adopt matching estimators for nonexperimental data to identify the causal effect of family difficulties on labour market outcomes. Standard econometric techniques are usually based on strong assumptions that possibly undermine the credibility of estimation results in case of their violation. Specifically, the violation of the common support condition as well as the misspecification of the functional form possibly leads to biased estimates. Recently, to relax such assumptions and reduce estimation bias problems, micro-econometricians (see Becker and Ichino, 2002, Black and Smith, 2004) have adopted techniques usually used in epidemiological studies (Rosenbaum and Rubin, 1983) to analyze the effect of an intervention (treatment) or problem to evaluate its effect on outcomes of interest. These non-parametric (or semi parametric) techniques include matching procedures based on the "selection on observables" assumption for which there exists a set of observed variables such that conditional on these, the impact of treatment is independent of the outcome that would occur without treatment (Conditional Independence Assumption, CIA) as Caliendo and Kopeinig (2008) underlined.

The empirical analysis uses information from five sweeps of NCDS database. The 1958 sweep (originally titled Perinatal Mortality Survey) and the 1965 sweep provide information about family difficulties at age 7, that is our treatment, and numerous pre-determined variables about family background and individual' s characteristics making the CIA credible. Finally, from 1991, 2000 and 2009 NCDS sweeps, we draw information about labour market outcomes, namely wages and employment probabilities. Besides to use standard OLS and probit estimates, we adopt propensity score matching estimators. Specifically, we use both Kernel Matching (KM) and Nearest Neighbor Matching (NNM) that differ in the way they deal with the trade-off between bias and efficiency.

Our findings suggest that propensity score matching results diverge somewhat from those of the standard econometric models, hence their use potentially preserves from estimation bias problems. Overall, we find evidence that family difficulties during childhood decrease both employment probabilities and wages with parameters' magnitudes and p values that differ somewhat according to the matching method used. Moreover, effects appear not to decline over the cohort working life. On average, the occurrence of family problems in childhood reduces the chances of being employed by about 6 % and employees' hourly wages by about 8.4%. Importantly, we also find that specific difficulties act differently. Economic and family structure difficulties determine the greatest disadvantage in terms of future labour market outcomes. While economic difficulties act both on employment probabilities and, overall wages, the effect of family structure problems on wages is smaller and not statistically significant. Health problems of family members and social isolation difficulties show smaller parameters and lower t-statistics, especially for estimations on wages.

This paper is organized as follows. Section II describes the data. Section III presents the econometric section. Section IV discusses the results and, finally, Section V concludes.

II. Data

The impact of family difficulties on adult labour market outcomes is investigated using information from the National Child Development Studies (NCDS). The NCDS is a cohort study that follows the lives of all those living in England, Scotland and Wales who were born in the first week in 1958. The main aim of the study is to improve the understanding of the factors affecting human development over the whole lifespan. The NCDS has its origin in the Perinatal Mortality Survey (PMS) that collected information on a cohort of about 17000 children. Successively, the PMS became the NCDS that has gathered information on about 11000 individuals of them in different points in time (1965, 1969, 1974, 1981, 1991, 1999-2000, 2004-2005 and 2008-2009).

We use five sweeps of the NCDS database. From the original 1958 and 1965 sweeps we draw information to identify treated and untreated individuals and suitable covariates to control for non-random selection into treatment, namely family difficulties during childhood. 1991, 2000 and 2009 NCDS sweeps are used to recover information useful to determine labour market performances in adulthood, namely employment and wages.

Family difficulties are identifiable using social environmental information gathered when cohort member was seven years old. Differently from many variables contained in the NCDS database, family difficulties variables are derived for completion of the health visitor report (from statutory or voluntary organizations), without questioning to the family, with the aim to determine the social environment in which children were growing up. Family difficulties include housing, financial, unemployment, physical illness or disability, mental illness or neurosis, mental sub-normality, death of child's parent(s), divorce, separation or desertion, domestic tension, in-law-conflict, alcoholism and other difficulties. This information is considered, in turn, as a whole to identify a unique and general family difficulties indicator or, in alternative, as four homogenous family difficulties sub-groups. Specifically, we determine the following family difficulties sub-groups: economic difficulties, health/disability difficulties of family members, family structure difficulties, and social isolation difficulties.

NCDS provides a large set of detailed pre-treatment information including those on cohort members and their parents. This richness allows us to identify a number of observable variables affecting both treatments and outcomes, making the CIA credible. With this in mind, we select the following

controls: sex of the cohort member, his/her birth weight, no walking alone by 1.5 years, speech talking by 2 years, wet by night after 5 years, disabling condition at age 7, number of cigarettes smoked prior pregnancy by the mother of the cohort member, English spoken at home, father's and mother's education, mother's age at the born of the cohort member, father's social class when the cohort member was 7 years old, parents' marital status, and regional dummies. Given the peculiarity of health and disability problems - with respect to the others -, a different set of covariates is selected. It is constituted by information about experiences of past stillbirths and neonatal death, a dummy indicating if the mother stopped smoking in pregnancy, and dummies indicating if the cohort member parents read newspapers.

The labour market performances are monitored by observing employment status and wages at different points in time of adulthood, i.e. when the cohort member is, respectively, 33, 42 and 51 years old. Employment status includes all employees or self-employed, either full-timers or part-timers. The individual wage is referred to the logarithm of the net hourly pay (at constant prices of 2009) received by an employee. It is calculated using information about the net pay, the period covered and the usual hours (including overtime) worked per week. The resultant hourly wage variable was subjected to top and bottom coding at 1% to reduce bias from outliers, and for the same reason we excluded from our sample individuals declaring to work less than 7 hours per week or more than 84 hours per week. Since we are interested in examining the evolution of the impact of family difficulties, we focus on individuals for which we have no missing information about the outcomes across the years investigated. This leaves us repeated cross-sectional information on 8008 individuals for the employment equations and 3872 individuals for the wage equations.

Table 1 contains descriptive information about the treatment identifiers, while table 2 contains descriptive statistics regarding employment and wage sub-samples distinguishing according to the total family difficulties variable.

[Table 1 about here]

[Table 2 about here]

Finally, table 3 displays observed average employment probabilities and wages comparing between treated and untreated individuals and referring to the total family difficulties treatment. T-test about the significance of the differentials between the two groups is also reported. Both observed employment and wage differentials remain quite constant across the period under investigation.

Specifically, observed employment differential is about 5% in 1991 and 2000, and about 6% in 2009, while observed log-real wage differential is, respectively, 0.09, 0.11 and 0.10. In all cases, differences are statistically significant at 1% level.

[Table 3 about here]

III. The model

We are interested in estimating the causal effect of family difficulties during childhood on adult labour market outcomes. Where family difficulties are taken as the treatment, the causal effect we wish to estimate corresponds to the average treatment effect on the treated (ATT). Ideally, we would need to compare the adult labour market outcomes of children experiencing family difficulties (the treated) to the same children had they lived in family without any difficulties. However, as we can observe each child only in one state, the outcomes for treated had they not been treated is an unobserved counterfactual. The average treatment effect (ATE) corresponds to the ATT only if the occurrence of any family difficulties is unrelated to outcomes. As we cannot exclude that there exist some factors or characteristics that affect both the occurrence of family difficulties and children outcomes, probably ATT and ATE will differ. An unbiased estimate of ATT can be obtained if treatment satisfies the Conditional Independence Assumption (CIA). In other words, treatments are at random once controlling for a suitable set of covariates. It remains possible that we are not provided with relevant information that affects both treatments and outcomes (selection on unobservables) but we are confident that the remaining source of selection is substantially reduced as the information provided to us from NCDS is detailed and we are controlling for many channels of indirect correlation.

More formally, it is possible to observe $Y_1|D=1$, the outcome of treated Y_1 if having family difficulties as a child (D=1), and $Y_0|D=0$, the outcome of untreated Y_0 if not having family difficulties as a child (D=0). The difference between the two outcomes corresponds to the ATE:

$$(1)\quad ATE = E(Y_1|D=1 - Y_0|D=0)$$

while we would like to estimate the ATT

(2) $ATT = E(Y_1|D=1 - Y_0|D=1) = E(Y_1 - Y_0|D=1)$

That is, the mean effect of experiencing family difficulties rather than not on the children who occurred family difficulties - the impact of treatment on the treated. However, $Y_0|D=1$ is not observable and if impacts are heterogeneous ATE and ATT diverge. Conditioning on an adequate set of covariates, it is possible to remove all systematic differences in outcomes in the untreated state (Conditional independence assumption, CIA)

(3) $(Y_0 \perp D) \mid X$

The outcome of untreated is independent of the treatment conditional on some set of observed covariates X.

Most of the existing applied studies estimated effects by standard parametric methods like OLS, and logit. However, these methods require strong assumptions on the functional form like linearity and additivity of regressors. The assumption of a linear or logistic function permits data from all observations to be combined into one estimate, but the validity of that estimate is suspect when one deals with people having very different characteristics. We rely on the propensity score matching (PSM) technique firstly proposed by Rosenbaum and Rubin (1983). PSM is a method in which no functional form restrictions on the relation between outcome, treatment, and control variables need be made. This technique matches each treated individual with an untreated individual having identical observable characteristics. The match does not need to be on each characteristic as it is equivalent to a match on the estimated probability of being in the treated group. Individuals whose predicted probabilities are similar are matched. Therefore, treated observations whose characteristics are not similar to anybody belonging to the control group - falling outside the common support region - get dropped from the analysis[1]. To examine the support condition, we plotted propensity scores of the two groups in Fig. 1. In the first quadrant, the top histogram reports observations who experienced at least a family difficulty (the D=1 group), while the bottom histogram represents those without any family problem in childhood (the D=0 group). The horizontal axis defines intervals of the propensity score and the height (or depth) of each bar on the vertical axis indicates the fraction of the relevant sample with scores in the corresponding interval. Similarly, we reported propensity scores for the case in which the treatment under consideration is

[1] For a complete discussion on matching methods, see also Dehejia and Wahba (2002).

represented by more homogeneous subgroups of family difficulties (economic, social isolation, family structure and health problems). Fortunately, the Figure 1 shows that in all cases the overlapped region is wide and it is not needed to eliminate a relevant number of observations[2].

[Figure 1 about here]

PSM method requires the balancing property to be satisfied. It is achieved when observations with the same propensity score have the same distribution of observable characteristics independently of treatment status. If this property is not satisfied this means that the two groups are too different in terms of observables and additional information would be needed[3].

Obtaining a specification that satisfies the balancing property does not assure us that we are credibly addressing the possible "selection on unobservables". This would be achieved by taking into account all relevant factors that affect both family difficulties and labour market outcomes. We are confident that the unobserved selection bias is reduced due to the uniquely rich source of information provided by the NCDS dataset. The dataset contains detailed information on the child condition and characteristics and on parents' characteristics and behaviors, taken from parents and health visitors reports at the time when the child born or was 7 years old. For what concerns the child, we include covariates containing information on the child gender, birth weight and on physical, cognitive and non-cognitive abilities (see in Table 2 the variables speech, walking, handicap and wet). Moreover, we account for information on parents' demographics and socio-economic characteristics like mother's age, age-squared, parents' years of schooling, mother's marital status, father's social class, number of cigarettes smoked by the mother prior pregnancy and language commonly spoken at home[4]. All these factors we are confident can be assumed as unaffected by treatment.

The causal effect we estimate (ATT) corresponds to the *total* effect - the summation of direct and the indirect effects - because we believe it to be more interesting from a public policy perspective.

[2] The number of observations dropped because not satisfying the common support condition ranges between 3.6 and 8.9% depending on the treatment.

[3] We check that this condition holds by means of the Stata command "pscore" employed by Becker and Ichino (2002) set at default parameters.

[4] This set of variables is common to all specifications but the one where the effect of health problem is estimated. In that case, the variables we consider are past stillbirths and neonatal death, a dummy indicating if the mother stopped smoking in pregnancy, and dummies indicating if the cohort member parents read newspapers.

We expect the direct effect of treatment to have a negative impact on the labour market outcomes. With regard to the indirect effects, on the one hand we expect that family problems may impact negatively on the human capital accumulation. On the other hand, these problems may trigger children efforts and determination at school and make them achieve higher and/or quicker outcomes at labour market. However, our view is that the net average indirect effect is negative like the direct one. Thus, the total effect we expect to be negative but we are not able to make prediction on the mid and long term intensity of such effects.

A variety of different methods can be used for implementing matching. All methods construct an estimate of the expected unobserved counterfactual for each treated observation by taking a weighted average of the outcomes of the untreated observations. What differs is the specific form of the weights. In order to check that our results are not driven by the kind of PSM technique chosen, we use two widely used methods that deal very differently with the trade-off between bias and variance: Gaussian Kernel and Nearest Neighbor Matching (NNM). Kernel matching can be seen as a weighted regression of the counterfactual outcome on an intercept with weights given by the kernel weights. Weights depend on the distance between each individual from the control group and the treated observation for which the counterfactual is estimated (see Smith and Todd, 2005). One major advantage of these approaches is the smaller variance, which is achieved because more information is used. A drawback of these methods is that also observations that are bad matches may be used. The second method is the most straightforward matching estimator. An individual from the comparison group is chosen as a matching partner for a treated individual that is closest in terms of propensity score[5].

Also standard parametric methods are employed in order to compare results. Depending on the outcome under observation, we use two alternative estimators. We performed logit for employment status and linear regression for the log of hourly wage. The set of covariates corresponds to the one used for estimating propensity score.

We restrict analysis to children participating at all five sweeps in order to compare the same individuals at each point in time and estimate changes in effects between mid and long term. The possible selection in the data due to non-response can be considered not to be a major problem as more than 70% of the children belonging to the cohort of 1958 participated in all five sweeps under consideration and the exits between 1991 and 2009 are few (the 85 % of those participating in 1991 is present in 2009).

[5] For a detailed discussion, see Caliendo and Kopeining (2008).

IV. Estimation results

Our estimation results using Gaussian kernel matching and Nearest Neighbor Matching techniques[6] are reported in tables 4 and 5. As mentioned above, we consider two different labour market types of outcome: employment status and employees log of hourly wages. For each treatment (namely, total, economic, health, social isolation and structure family difficulties) and child's age considered (33, 42 and 51 years old), we report estimated average treatment effects of treated (ATT), t-statistics (resulting from bootstrapped standard errors), and the number of treated and controls used by each matching technique. For what concerns employment status (Table 4), we find that the estimated ATT, for whatever type of family difficulty, is almost always statistically significant at 1% level with Kernel matching and significant parameters range between -0.062 and -0.111. Economic and family structure problems are the major predictors of adult poor labor employment rates, as on average the chances of being employed are less than their corresponding untreated by about the 7 percent. The corresponding NNM estimates show slight smaller point estimates. Total, economic and family structure difficulties are still statistically significant, while the coefficients on the remaining sub-groups are less precisely estimated. This is not surprising as the two considered methods balance very differently between the bias and variance trade-off, with the NNM minimizing bias at the cost of larger variance. This is due to the fact that the number of untreated observations matched with treated is by far larger with the Kernel than with the NNM.

[Table 4 about here]

Results from Table 4 show that family difficulties have long term lasting effects that do not tend to disappear even after more than forty years. Moreover, the estimated ATTs do not show any declining path through the two decades under observation (1991-2009). It is also interesting to note that the magnitude is relevant as the occurrence of whatever family difficulty in childhood steadily reduces the probability to find a job as an adult by on average 6%.

[6] Estimations are performed by STATA's commands "attk" and "attnd" employed by Becker and Ichino (2002), set at default parameters, even including bandwidth, and with options "logit" and "comsup". The latter is enabled to so that ATT estimations only use observations inside the common support. Standard errors are computed using the bootstrap technique with replications set at 500.

Turning our view to wages (Table 5), we find that statistically significant ATT estimates on the log of employees' hourly wage imply an average reduction in salaries of about 8.4%, ranging between - 6.2 and - 11 percent. Again economic difficulties in childhood seem to be the more important to persistently determine adult wages through the working life with average impact of - 9.2%. The effect of health and disability problems is large and significant only with the Gaussian kernel method. In general, the effect on wages appears to be somewhat smaller and less precisely estimated than the effect on employment status. This can be probably attributed to the fact that the number of usable observations shrink by about the half and to the presence of greater measurement error in wages data.

[Table 5 about here]

Now we examine results from regression-based methods to assess whether using standard parametric methods lead to biased estimates and - if it is the case - assess its quantitative relevance. Table 6 shows estimated marginal effects of the parameter of interest for both the labour market outcomes considered. With regard to employment status, we find that almost always the coefficient on treatment status is statistically significant at conventional levels. The probability of being employed is lowered by the occurrence of any family problem on average by 4.5%. Economic and family structure problems have the major negative impact (-5.3 and -6.4%, respectively). Overall, our evidence on the employment outcome suggests that regression-based methods underestimate the true negative impact as we find slightly smaller parameter values in absolute terms than using matching methods (on average, -4.5% rather than -6%). Similarly, the estimated parameters on wages appear to be smaller than those obtained from matching estimates with average impact of about -6.6% (in place of 8.4% of previous estimates). Consistently with previous findings, such effects seem to last persistently over the entire working lives.

[Table 6 about here]

V. Conclusions

This paper focuses on the causal effect of family difficulties during childhood on adult labour market outcomes - employment and wage - applying the propensity score matching approach to

NCDS data. As family difficulties are taken as the treatment, the causal effect we estimate corresponds to the average treatment effect on the treated. The use of propensity score matching estimators preserve us from the usual estimation bias arising from the use of standard econometric techniques. A second point we examine is the evolution of the causal effect on labour market outcomes, evaluating it at different ages in adulthood. Finally, besides of considering the family difficulties as a unique problem, we also consider them separately, identifying four homogenous groups according to the nature of the family difficulties (economic, health/disability, family structure and social isolation).

Estimation results clarify that experiencing family difficulties during childhood negatively affect both employment probabilities and wages. Importantly, we also find that specific family difficulties act differently. Economic and family structure difficulties determine the greatest disadvantage in terms of future labour market outcomes. However, while economic difficulties act on both employment probabilities and overall wages, the effect of family structure problems is not statistically significant in wage equations. Results on health problems of family members and social isolation difficulties show more heterogeneous evidence. Health problems are significant only with one of the two matching estimators adopted (Gaussian kernel). Social isolation is statistically relevant only with Gaussian kernel method and in employment equation.

The negative impact of family difficulties during childhood on adult labour market outcomes has further implications on socio-economic conditions of individuals. For example, it may be seen as an underlying motive of social exclusion and social immobility. Because of their consequences both at economic and psychological level, poor labour market performances are likely to be associated with social exclusion. Moreover, social immobility may be seen as a direct consequence of the causal relationship between economic difficulties during childhood and poorer labour market performances in the adulthood.

Our findings suggest that disadvantaged positions on the labour market could be prevented, at least partially, implementing suitable policies aimed to reduce the direct and indirect impact of family difficulties during childhood. Moreover, social policies aimed to fight social exclusion and immobility should not ignore that these problems are in part determined in the individual's far past.

References

Becker S. and A. Ichino (2002) "Estimation of average treatment effects based on propensity scores", *Stata Journal*, vol. 2(4), pp. 358-377.

Black D.A. and J.A. Smith (2004) "How robust is the evidence on the effects of college quality? Evidence from matching", *Journal of Econometrics*, vol. 121, pp. 99-124.

Caliendo M. and S. Kopeinig (2008) "Some practical guidance for the implementation of propensity score matching", *Journal of Economic Surveys*, vol. 22(1), pp.31-72.

Case A., A. Fertig and C. Paxson (2005) "The lasting impact of childhood health and circumstance", *Journal of Health Economics*, vol. 24, pp. 365-389.

Chevalier A. and T.K. Viitanen (2003) "The long-run labour market consequences of teenage motherhood in Britain", *Journal of Population Economics*, vol. 16, pp. 323-343.

Corak, M. (2001) "Death and divorce: the long-term consequences of parental loss on adolescents", *Journal of Labor Economics*, vol. 19(3), pp. 682-715.

Cunha, F. and J.J. Heckman (2006) "Investing in our young people", Unpublished manuscript, University of Chicago, Department of Economics.

Dehejia, R. and S. Wahba (2002) "Propensity score matching methods for nonexperimental causal studies", *Review of Economics and Statistics*, vol. 84 (1), 151–161.

Fletcher, J. and B.L. Wolfe (2008) "Child mental health and human capital accumulation: the case of ADHD revisited", *Journal of Health Economics*, vol. 27, pp. 794–800.

Glewwe P., H.G. Jacoby and E.M. King (2001) "Early childhood nutrition and academic achievement: a longitudinal analysis", *Journal of Public Economics*, vol. 81(3), pp.345-368.

Goodman A. and B. Sianesi (2005) "Early education and children's outcomes: how long do the impacts last", *Fiscal Studies*, vol. 26(4), pp. 513-548.

Gregg, P. and S. Machin (2000) "Child development and success or failure in the youth labour market", NBER Comparative Labour Market Series (pp. 247-288), Chicago: University of Chicago Press.

Gruber, J. (2004) "Is making divorce easier bad for children? The long run implications of unilateral divorce", *Journal of Labor Economics*, vol.22 (4), pp. 799-834.

Healey A., M. Knapp and D.P. Farrington (2004) "Adult labour market implications of antisocial behavior in childhood and adolescence: findings from a UK longitudinal study", *Applied Economics*, vol. 36(2), pp. 93-105.

Heckman, J.J., J. Stixrud, and S. Urzua (2006) "The effects of cognitive and non-cognitive abilities on labour market outcomes and social behavior", *Journal of Labor Economics*, vol. 24(3), pp. 411-481.

Koning P., D. Webbink, S. Vujic and N.G. Martin (2010) "The effect of childhood conduct disorder on human capital", IZA discussion paper n. 4940.

Le, A.T., P.W. Miller, A.C. Heath and N. Martin, (2005) "Early childhood behaviors, schooling, and labour market outcomes: estimates from a sample of twins", *Economics of Education Review*, vol. 24, pp. 1-17.

Lindeboom M., A. Llena-Nozal and B. van der Klaauw (2006) "Disability and work: the role of health shocks and childhood circumstances", IZA discussion paper n. 2096.

Rosenbaum P. and D. Rubin (1983) "The central role of the propensity score in observational studies for causal effect", *Biometrika*, vol.70 pp. 41-50.

Sanz de Galdeano A. and D. Vuri (2007) "Parental divorce and students' performance: evidence from longitudinal data", *Oxford Bulletin of Economics and Statistics*, vol. 69(3), pp. 321-338.

Slade, E.P. and L.S. Wissow, (2007) "The influence of childhood maltreatment on adolescents academic performance", *Economics of Education Review*, vol. 26(5), pp. 604-614.

Smith, J A. and Todd, P.E. (2005). “Does matching overcome Lalonde’s critique of nonexperimental estimators?”, *Journal of Econometrics*, vol. 125, 1-2, 305-353.

Smith J. (2009) “The impact of childhood health on adult labor market outcomes”, *The Review of Economics and Statistics*, vol. 91(3), pp.478-489.

Table 1. Treatments: descriptive statistics

	Employment equations		Wage equations	
Treatment	Mean	Std.Dev.	Mean	Std.Dev.
Family difficulties (FD)	0.166	0.372	0.159	0.366
Economic FD	0.091	0.287	0.085	0.280
Health FD	0.058	0.234	0.054	0.225
Social isolation FD	0.069	0.254	0.066	0.249
Family structure FD	0.039	0.193	0.035	0.185

Note. Our elaboration based on NCDS data.

Table 2. Descriptive statistics by labour market outcome

	Employment equations				Wage equations			
	Control group		Treatment group		Control group		Treatment group	
	(obs. 6676)		(obs. 1332)		(obs. 3181)		(obs. 601)	
Variable's name	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Birth weight	0.949	0.220	0.923	0.267	0.948	0.221	0.932	0.252
English spoken at home	0.752	0.432	0.748	0.434	0.767	0.423	0.745	0.436
Disabling condition at age 7	0.023	0.149	0.056	0.231	0.022	0.146	0.052	0.221
No walking alone by 1.5 years	0.030	0.172	0.043	0.202	0.028	0.165	0.037	0.188
Speech-talking by 2 years	0.812	0.390	0.923	0.267	0.824	0.381	0.923	0.266
Wet by night after 5 years	0.078	0.269	0.133	0.340	0.083	0.276	0.158	0.365
Number of cigarettes prior pregnancy	3.487	5.853	4.857	6.670	3.464	5.762	4.979	6.771
Male	0.481	0.500	0.456	0.498	0.494	0.500	0.479	0.500
Father education	2.988	2.331	2.492	1.961	3.021	2.283	2.511	1.982
Mother education	3.024	2.155	2.733	1.888	3.040	2.073	2.687	1.834
Missing father education	0.265	0.442	0.303	0.460	0.251	0.434	0.304	0.461
Missing mother education	0.252	0.434	0.260	0.439	0.235	0.424	0.265	0.441
Mother's age	26.175	8.117	26.696	7.829	26.250	8.136	26.857	7.836
Mother's age square	751.00	359.35	773.93	379.24	755.24	365.34	782.59	378.37
Missing mother's age	0.051	0.220	0.035	0.183	0.049	0.217	0.035	0.184
Father is manual social class at 7	0.528	0.499	0.681	0.466	0.545	0.498	0.689	0.463
Married	0.926	0.262	0.917	0.276	0.928	0.259	0.915	0.279
North	0.067	0.249	0.079	0.270	0.071	0.256	0.098	0.298
North-West	0.107	0.310	0.128	0.335	0.112	0.315	0.133	0.340
East & West Riding	0.077	0.266	0.074	0.262	0.077	0.267	0.090	0.286
North Midlands	0.073	0.260	0.068	0.252	0.077	0.266	0.077	0.266
Midlands	0.082	0.275	0.103	0.304	0.082	0.274	0.093	0.291
East	0.079	0.270	0.071	0.257	0.073	0.260	0.052	0.221
South-East	0.163	0.369	0.180	0.384	0.150	0.357	0.173	0.379
South	0.059	0.235	0.073	0.260	0.056	0.229	0.055	0.228
South-West	0.061	0.240	0.044	0.204	0.059	0.236	0.040	0.196
Wales	0.048	0.213	0.070	0.255	0.047	0.212	0.068	0.252
Scotland	0.089	0.285	0.110	0.313	0.105	0.306	0.121	0.327
Mother stopped smoking in pregnancy	0.069	0.253	0.073	0.260	0.071	0.257	0.079	0.270
Past stillbirths and neonatal death	0.113	0.316	0.148	0.355	0.114	0.318	0.163	0.370
Father's read-newspaper	0.666	0.472	0.623	0.485	0.673	0.469	0.606	0.490
Mother's read-newspaper	0.741	0.438	0.717	0.451	0.751	0.432	0.739	0.440
Husband' s age	28.254	10.219	30.281	10.839	28.210	10.353	31.010	9.436
Missing husband's age	0.078	0.268	0.071	0.257	0.079	0.269	0.044	0.206

Note. Our elaboration based on NCDS data.

Table 3. Observed employment probabilities and wages

Employment						
	1991		2000		2009	
Group	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Control	0.818	0.386	0.879	0.326	0.871	0.335
Treatment	0.768	0.422	0.830	0.376	0.812	0.391
t-test H0: $\mu 1=\mu 0$	4.262		4.840		5.725	
Log-Real Wages						
	1991		2000		2009	
Group	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Control	1.976	0.389	2.057	0.563	2.268	0.438
Treatment	1.885	0.368	1.944	0.523	2.169	0.414
t-test H0: $\mu 1=\mu 0$	5.293		4.555		5.133	

Note. Our elaboration based on NCDS data. Real wages are at constant prices of 2009.

Table 4. Estimation Results. Labour Market outcomes: Employment Status. Propensity Score Matching Estimates.

Labour Market Outcome: Employment Status								
	Gaussian Kernel Matching				*NNM*			
	treat.	*contr.*	*ATT*	*t*	*treat.*	*contr.*	*ATT*	*t*
				Any Family Difficulties				
1991	1332	6015	**-0.045**	-3.516	1332	1315	***-0.032***	-1.665
2000	1332	6015	**-0.045**	-4.034	1332	1315	**-0.034**	-2.013
2009	1332	6015	**-0.054**	-4.738	1332	1315	-0.024	-1.423
				Economic Family Difficulties				
1991	725	6577	**-0.074**	-4,358	725	851	**-0.063**	-2,360
2000	725	6577	**-0.053**	-3,560	725	851	**-0.048**	-2,070
2009	725	6577	**-0.083**	-5,247	725	851	**-0.062**	-2,639
				Health Family Difficulties				
1991	467	7481	***-0.033***	-1.663	467	700	-0.014	-0.405
2000	467	7481	**-0.058**	-2.909	467	700	***-0.057***	-1.928
2009	467	7481	**-0.041**	-2.322	467	700	-0.026	-0.878
				Social Isolation Difficulties				
1991	556	6750	-0.027	-1.537	556	617	-0.017	-0.579
2000	556	6750	**-0.041**	-2.505	556	617	-0.013	-0.524
2009	556	6750	**-0.063**	-3.548	556	617	***-0.045***	-1.695
				Family Structure Difficulties				
1991	311	7479	**-0.049**	-2.015	311	348	**-0.089**	-2.265
2000	311	7479	**-0.067**	-2.805	311	348	***-0.065***	-1.943
2009	311	7479	**-0.096**	-4.145	311	348	**-0.117**	-3.156

Note: Our elaboration based on NCDS data. Propensity score matching estimations are performed by means of the STATA commands attk and attnd using default parameters and options "logit" and "comsup". Statistically significant at 1 and 5 % are reported in bold; significant at 10% in bold and italic. T-stats are obtained by using standard errors bootstrapped with 500 replications.

Table 5. Estimation Results. Labour Market outcomes: Log of Employees Hourly Wages. Propensity Score Matching Estimates.

Labour Market Outcome: Log Hourly Wage								
	Gaussian Kernel Matching				*NNM*			
	treat.	*contr.*	*ATT*	*t*	*treat.*	*contr.*	*ATT*	*t*
				Any Family Difficulties				
1991	601	2874	**-0.066**	-4.45	601	556	***-0.046***	-1.712
2000	601	2874	**-0.081**	-3.698	601	556	-0.038	-1.042
2009	601	2874	**-0.070**	-3.689	601	556	-0.022	-0.71
				Economic Family Difficulties				
1991	323	3146	**-0.081**	-4,353	323	336	***-0.064***	-1,879
2000	323	3146	**-0.106**	-4,006	323	336	**-0.092**	-2,005
2009	323	3146	**-0.097**	-4,477	323	336	**-0.111**	-2,866
				Health Family Difficulties				
1991	203	3501	**-0.087**	-3.118	203	282	-0.063	-1.396
2000	203	3501	-0.054	-1.417	203	282	-0.040	-0.597
2009	203	3501	**-0.067**	-2.133	203	282	-0.071	-1.362
				Social Isolation Difficulties				
1991	251	3185	***-0.037***	-1.643	251	252	-0.043	-1.017
2000	251	3185	-0.035	-1.039	251	252	-0.006	-0.1
2009	251	3185	-0.044	-1.554	251	252	***-0.076***	-1.625
				Family Structure Difficulties				
1991	134	3639	***-0.062***	-1.927	134	131	-0.065	-1.259
2000	134	3639	-0.071	-1.5	134	131	-0.024	-0.306
2009	134	3639	-0.020	-0.516	134	131	-0.012	-0.174

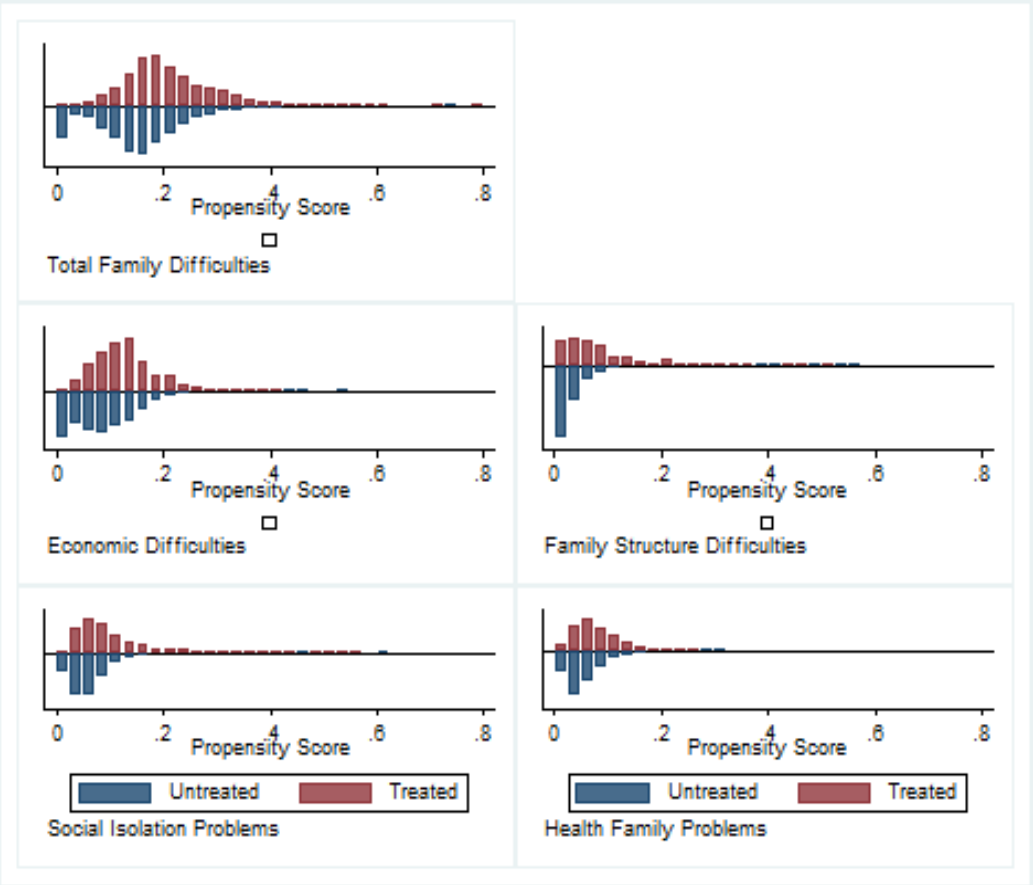
Note. Our elaboration based on NCDS data. Propensity score matching estimations are performed by means of the STATA commands attk and attnd using default parameters and options logit and comsup. Reported coefficients of the log of real wage equations were modified to be interpreted as the percent change in wages due to treatment by using this simple formula: exp(treatment coefficient) - 1. T-stats are obtained by using standard errors bootstrapped with 500 replications. Statistically significant at 1 and 5 % are reported in bold; significant at 10% in bold and italic. Real wages are at constant prices of 2009.

Table 6. Estimation Results. Labour Market outcomes: Employment Status and Log of Employees Real Hourly Wages. Standard Parametric Methods (Logit and OLS).

	Standard Parametric Methods					
	Employment			ln hourly wage		
	1991	*2000*	*2009*	*1991*	*2000*	*2009*
Any Family Difficulties	**-0.036**	**-0.036**	**-0.043**	**-0.057**	**-0.071**	**-0.062**
	(-3.22)	(-3.78)	(-4.45)	(-3.84)	(-3.25)	(-3.53)
R-squared	0.105	0.054	0.032	0.193	0.123	0.144
N	8008	8008	8008	3782	3782	3782
Economic FD	**-0.057**	**-0.041**	**-0.062**	**-0.062**	**-0.082**	**-0.073**
	(-4.07)	(-3.44)	(-5.23)	(-3.36)	(-3.02)	(-3.27)
R-squared	0.105	0.053	0.032	0.192	0.122	0.143
N	8008	8008	8008	3782	3782	3782
Health FD	-0.026	**-0.044**	***-0.028***	**-0.070**	-0.025	-0.045
	(-1.46)	(-3.14)	(-1.86)	(-2.58)	(-0.61)	(-1.42)
R-squared	0.006	0.008	0.01	0.052	0.035	0.045
N	8008	8008	8008	3782	3782	3782
Social Isolation FD	-0.024	**-0.032**	**-0.052**	-0.029	-0.034	-0.039
	(-1.47)	(-2.39)	(-3.82)	(-1.25)	(-1.08)	(-1.47)
R-squared	0.103	0.051	0.03	0.191	0.121	0.141
N	8008	8008	8008	3782	3782	3782
Structure FD	***-0.04***	**-0.05**	**-0.078**	**-0.059**	***-0.074***	-0.026
	(-1.95)	(-2.93)	(-4.65)	(-2.00)	(-1.71)	(-0.72)
R-squared	0.103	0.051	0.031	0.19	0.121	0.14
N	8008	8008	8008	3782	3782	3782

Note. Our elaboration based on NCDS data. Reported coefficients of the log of wage equations were modified to be interpreted as the percent change in wages due to treatment by using this simple formula: exp(treatment coefficient) - 1. Statistically significant at 1 and 5 % are reported in bold; significant at 10% in bold and italic. Real wages are at constant prices of 2009.

Figure 1. Propensity scores by grouped family difficulties



Note. Our elaboration based on NCDS data.